Retroactivity, Modularity, and Insulation in Synthetic Biology Circuits
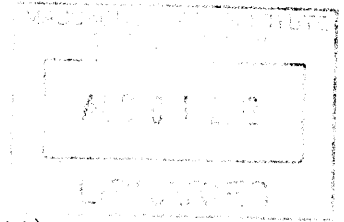
by

Allen Lin

B.S. Electrical Engineering and Computer Science (2011)
B.S. Chemical-Biological Engineering (2011)
Massachusetts Institute of Technology

Submitted to the Department of Electrical Engineering and Computer
Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2011

Author . . . .
Department of Electrical Engineering and Computer Science
August 22, 2011

Certified by. . . . .
Ron Weiss, Associate Professor
Thesis Supervisor

Certified by. . . . .
Domitilla Del Vecchio, Associate Professor
Thesis Supervisor

Accepted by . . . .
Dennis M. Freeman, Professor
Chairman, Masters of Engineering Thesis Committee

Retroactivity, Modularity, and Insulation in Synthetic Biology Circuits

by

Allen Lin

Abstract

A central concept in synthetic biology is the reuse of well-characterized modules. Modularity simplifies circuit design by allowing for the decomposition of systems into separate modules for individual construction. Complex regulatory networks can be assembled from a library of devices. However, current devices in synthetic biology may not actually be modular and may instead change behavior upon interconnections, a phenomenon called retroactivity. Addition of a new component to a system can change individual device dynamics within the system, potentially making time-consuming iterative redesign necessary. Another need for systems construction is the ability to rapidly assemble constructs from part libraries in a combinatorial, high-throughput fashion.

In this thesis, a multi-site assembly method that permits the rapid reshuffling of promoters and genes for yeast expression is established. Synthetic circuits in yeast to measure retroactivity and to act as an insulator that attenuates such effect are designed and modeled.

Thesis Supervisor: Ron Weiss
Title: Associate Professor

Thesis Supervisor: Domitilla Del Vecchio
Title: Associate Professor

Acknowledgments

First and foremost, I would like to thank my advisor, Ron Weiss, for providing me with the opportunity and the resources to work in the exciting field of synthetic biology. His vision and passion for the field has always been inspiring, and the balance he has given me between advice and a freedom to pursue my own interests has allowed me to grow scientifically. I also owe much gratitude to my co-advisor, Domitilla Del Vecchio, for her guidance and feedback.

I would like to thank the many labmates I have worked with this past year. Thanks to Deepak Mishra, who got me started in the Weiss Lab and with yeast work, and to Jingjing Sun and Liliana Wroblewska, who also provided me with help in working with yeast. Thanks to Jonathan Babb for being an enjoyable bay-mate in Tecan robot room. Thanks to Yinqing Li for his great companionship, Adrian Slusarczyk for his continual energy, and Xavier Dupont for his good humor. Thanks to Steve Firsing for keeping the lab neat and running smoothly. I have also benefited from many entertaining and academic interactions with all other past or current members of the Weiss Lab: Hattie Chung, Noah Davidsohn, Cristian Grecu, Saurabh Gupta, Patrick Guye, Adrian Randall, Adam Rubin, and Zhen Xie.

Outside of lab, many thanks to my dear Pearl Street housemates, Isaac Joseph, Patrick Yamane, Jabe Ziino, and, de facto, Aaron Thom. They created an always warm and fun environment to come home to. In addition, thanks to Bob Chen, Fei Chen, and Jonathan Gootenberg for their friendship and for being great biology conversation-mates. In addition, thanks to my academic advisors, Kris Prather and Bruce Tidor, for the invaluable advice they have provided during my time at MIT. And thanks to Ken Oye for many engaging discussions about the societal aspects of synthetic biology.

Last but not the least, I would like to thank my family, especially my parents Ling-ling Chiang and Chiachen Lin. Their love, support, and sacrifice have been instrumental to me being where I am today. I hope I can do something useful for them and for society in return for all of their efforts.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Thesis Statement

Synthetic biology is a rapidly growing discipline that seeks to bring engineering concepts into biological systems. This field aims to redesign cellular systems to perform useful functions, such as producing drugs, synthesizing materials, and even fighting tumors. A central concept in synthetic biology is the construction of reusable, well-characterized modules. Modularity simplifies circuit design by allowing engineers to decouple systems into separate modules and construct and test modules individually. Instead of designing genetic circuits from scratch with each project, complex regulatory networks can be quickly assembled from a library of functional devices. Characterization of external module behavior, such as with input/output signal response curves, provides a useful abstraction that does not require a biological engineer to know about internal device workings.

Modularity is an existing and essential concept in engineering fields such as electrical engineering and computer architecture, and this abstraction can be applied to synthetic biology. However, current devices constructed in synthetic biology may not actually exhibit modular behavior. Modularity fails when individual device behavior depends on the behavior of other connected components. In particular, device characteristics may change upon different interconnections, a phenomenon called retroactivity.

Retroactivity is undesirable in biological engineering, as it decreases the reliability and predictability of device behavior in different circuit settings. Every time a new component is added to a system, new connections are made which can affect input/output characteristics of all of the devices in the system. Therefore, iterative redesign of the network might be necessary to accommodate each new change, which is time-consuming and inefficient. Ideally, modules should function similarly whether tested individually or placed into larger systems.

Studying such natural phenomenon in synthetic biology is facilitated by the availability of quick and reliable DNA assembly technology. Currently, the field is limited in its physical ability to assemble new circuits and redesign existing ones. Being able to assemble large constructs from existing parts libraries in a combinatorial, high throughput fashion will advance the development of synthetic biology.

> In this thesis, I develop a multi-site assembly system that permits the rapid reshuffling of promoters and genes for yeast expression. Using this construction scheme, I design and model synthetic circuits in yeast to measure retroactivity and to act as an insulator to attenuate retroactivity.

Level	Description	Examples
Systems	multiple devices linked together	cancer killing bacteria, artificial tissue homeostasis
Devices	multiple parts linked together with higher level functionality	invertor, oscillator, switches
Parts	sequence of nucleotides with defined functionality	promoter, RBS, ORF, terminator
DNA	sequence of nucleotides	AAAGAGGAGAAA

Figure 1-1: Ideal abstraction hierarchy in synthetic biology. Engineering complexities should be contained within each abstraction level. For example, a systems engineer should be able to use a device just knowing its input/output characterization, without having to know how internal promoters were tuned by the device engineer who made it. Figure adapted from [30].

1.2 Modularity in Synthetic Biology

1.2.1 Synthetic Biology

Synthetic biology is a nascent bottom-up, forward engineering discipline that aims to rationally design and program biological systems. The field seeks to apply engineering principles to biology to improve human therapeutics, drug manufacturing, biofuel production, environmental remediation, and biosensing applications [63]. Individual parts are systematically and thoroughly characterized. These standardized parts can then be composed to be make devices and systems. Parts, once developed, can be re-used in multiple modules and systems, accelerating the process of assembling complex regulatory networks. Examples of parts and devices constructed to date include toggle switches [38, 66], oscillators [29, 35], protein switches [24, 26, 42, 45, 54, 97], RNA circuits [9, 56], multicellular systems [7, 8, 19, 123, 65, 17], and metabolic networks [32, 78]. By utilizing the principles of abstraction, standardization, and characterization, synthetic biology aims to bring engineering concepts present in other construction-oriented disciplines into biology.

1.2.2 Modularity and Retroactivity

The first step towards designing and constructing large-scale biological systems is engineering individual devices that act in a modular fashion. Modularity simplifies circuit design by allowing engineers to decouple complicated systems into separate modules and construct and test modules individually. Modularity also guarantees that the characteristics of individual devices do not change upon interconnections. As components are built and tested in isolation, it is desirable for their properties to stay the same when connected. Modularity is an existing and essential concept in engineering fields such as electrical engineering and computer architecture.

Use of modular abstraction will accelerate the development of synthetic biology. However, current devices constructed in synthetic biology may not actually exhibit modular behavior. In particular, device characteristics may change upon different interconnections, a phenomenon called retroactivity. Retroactivity, impedance-like effects, at module junctions can disrupt desired component functionality. This phenomenon occurs because of the physical binding of regulatory factors from the output of an upstream device to binding sites of downstream devices. The effect of retroactivity on an upstream device increases as the load increases, where load is the number of downstream binding sites.

Retroactivity is undesirable in engineering as it decreases the reliability and predictability of device behavior in different circuit settings. Every time a new component is added to a system, new connections are made which can affect input/output characteristics of all of the devices in the system. Therefore, iterative redesign of the network might be necessary to accommodate each new change, which is time-consuming and inefficient.

Retroactivity is observed in not just biological settings but also in electrical and hydraulic systems. In fact, retroactivity is analogous to the phenomenon termed "fanout" in electronic digital circuits, where a single logic gate can reliably drive only up to a certain number of downstream gates. Increasing the number of downstream gates will cause errors, as the particular upstream logic gate cannot deliver enough current

to the inputs of downstream gates.

1.2.3 Use of an Insulator

Ideally, modules should function similarly whether tested individually or placed into larger systems. To resolve this, an insulator can be placed between two components to insulate them from retroactivity. In this thesis, a phosphorylation-based insulation device is proposed that facilitates modular composition. Components connected through the insulator will maintain their isolated dynamical properties. This insulator is similar to non-inverting operational amplifiers in electronic systems. These operational amplifiers can provide infinite input impedance such that it does not draw in current (zero retroactivity to the input) and can greatly attenuate retroactivity to the output due to large amplification gains and negative feedbacks.

1.3 Summary of Contributions and Thesis Outline

In the following section, I summarize my contributions and outline the remainder of this thesis.

Novel Assembly Method for Yeast Expression

I established a rapid method of assembling multiple promoter-gene pairs into a single yeast-compatible plasmid. This method builds on two existing technologies, Gateway© cloning and Gibson cloning, to create a two-step assembly protocol. In the first step, pairs of promoters and genes are recombined together into positional vectors with Gateway© cloning. In the second step, these positional clones are assembled in a defined ordered into a single carrier vector. This assembly method does not require the use of gel extraction, allows for specification of parts order, does not require the use of PCR and thus preserves DNA sequence integrity, does not have forbidden sites in promoters or genes, utilizes parallel over serial/hierarchical assembly for speed, and can currently include eight promoter-genes pairs. This assembly method was established with Deepak Mishra in the Weiss Lab and is based of an assembly method for mammalian circuits developed by Patrick Guye and Yinqing Li, also in the Weiss

21

Lab.

Libraries of Promoters, Genes, and Yeast Vectors

To study retroactivity and its insulation, I constructed a library of promoters and genes compatible with the assembly method previously described. The promoter library includes high constitutive, low constitutive, inducible, and repressible promoters. The gene library includes inducible and repressible transcription factors, fluorescent proteins, signaling proteins, and an antibiotic marker. I also constructed a library of vectors with different auxotrophic markers and yeast origins of replication that have varied copy numbers. The assembly method allows for combinatorial linkage of these promoters and genes into single vectors. These libraries were constructed with Deepak Mishra.

Computational Evaluation of Previous Models of Retroactivity and Insulators

I computationally simulated models for retroactivity and a phosphorylation - dephosphorylation cycle based insulator, previously presented in [22], to evaluate the various assumptions made. In particular, I found that in a simple buffer system, whether a protein-DNA complex decays has a significant effect on model behavior. On the other hand, I found that in the insulator design, whether signaling proteins and protein-DNA complexes decay has a limited effect on model behavior at the degradation rates tested. Investigation of the assumptions made supports the model of the phosphorylation - dephosphorylation cycle based insulator as a buffer against retroactivity.

Design and Modeling of a Construct to Measure Retroactivity in Yeast

To quantify retroactivity in uninsulated interconnections between regulatory genetic networks, I designed and modeled a buffer system that has an adjustable load on the signaling transcription factor. I also created a deterministic model of the system with parameters obtained from literature such that the model could eventually be fitted to experimental data.

Design and Modeling of an Insulation Device in Yeast

I conducted background research on a proposed synthetic insulator that utilizes two existing signaling pathways, the high osmolarity glycerol pathway (HOG) in yeast and the JAK-STAT pathway in mammalian cells. I also created a deterministic model that demonstrates the that proposed insulator can robustly attenuate retroactivity.

Preliminary Data on a Repressible Constructs Created through Assembly Strategy

Lastly, I constructed and tested circuits to demonstrate the functionality of the assembly strategy developed in this thesis. Preliminary data collected suggests that tTA, a repressible activator to be used in the buffer and insulator constructs, is functional and that systems that are integrated are more stable than systems expressed from plasmids.

Chapter 2 discusses the novel assembly method for yeast expression and the libraries of promoters, genes, and yeast vectors. Chapter 3 discusses the computational evaluation of previous models of retroactivity and insulators. Chapter 4 discusses the design and modeling of the buffer construct and the insulation device in yeast and also shows preliminary data on constructs made.

Chapter 2

Assembly of Experimental Circuits

Synthetic biology seeks to engineer cells to perform particular tasks or change their metabolic processes through the introduction of biological DNA circuits into cells. Although short DNA oligonucleotides can be synthesized at a reasonable price, synthesis of multi-kilobases sequences, the typical length of such circuits, is still relatively expensive. Thus, there is a large effort within synthetic biology to develop methods that rapidly, accurately, and reliably assemble large strands of recombinant DNA. Since the field is currently limited by its capability to assemble new circuits and re-design existing circuits, the ability to assemble existing components into an ordered construct will accelerate the development of synthetic biology [28].

The Baker's yeast *Saccharomyces cerevisiae* is a model organism that has been used in multiple academic studies and industrial processes. It is a desirable model organism because it is the simplest unicellular eukaryotic organism that can be easily grown and genetically manipulated. Many of its pathways are similar to those found in mammalian cells, and libraries of single knockouts strains and single inducible expression strains have been developed [107, 121]. Yeast has also been extensively metabolically engineered for chemical and biofuel production. Advances in methods for assembling yeast gene circuits can greatly accelerate research in human diseases, drug discovery, and chemical production.

In Section 2.1, I review Gateway© cloning and Gibson cloning, which are used in the proposed assembly method that I discuss in Section 2.2.
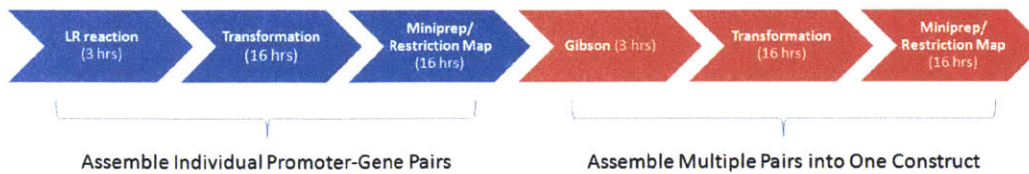
Figure 2-1: Assembly overview and flowchart. Two steps are involved in the assembly method presented. First, individual pairs of promoters and genes are assembled on separate plasmids with LR recombination. Second, these pairs are assembled into a single plasmid with Gibson cloning. The total process takes 4 days and currently can assemble up to 8 promoter-gene pairs into a single plasmid.

2.1   Review of Two Existing Assembly Technologies

The novel assembly technology developed in this thesis allows for the rapid assembly of multiple promoter-gene pairs into a single plasmid, which can then be expressed in yeast as a plasmid or stably integrated into a single locus site (Figure 2-1). This technology differs from existing yeast circuit assembly methods in that it permits the assembly of multiple promoter-gene pairs in a parallel, rather than sequential or hierarchical, reaction. The time to assemble circuits from an existing library is 4 days and does not require the use of gel extractions or polymerase chain reactions (PCR).

The benefits of assembling multiple promoter-gene pairs into a single plasmid include reduced cellular toxicity and component flexibility. The transformation of multiple centromeric plasmids into yeast is toxic to the cell and can significantly slow its growth [37]. Thus, it is desirable to assemble the circuit into as few plasmids as possible. Integration of single promoter-gene pairs at different sites can only be conducted sequentially and thus is not time efficient. In addition, sequential integration does not allow for the rapid swapping of integrated elements (e.g. changing the strength of a particular promoter). The assembly technology described below overcomes both of these shortcomings and relies on two existing cloning technologies, Gateway© cloning and Gibson cloning.
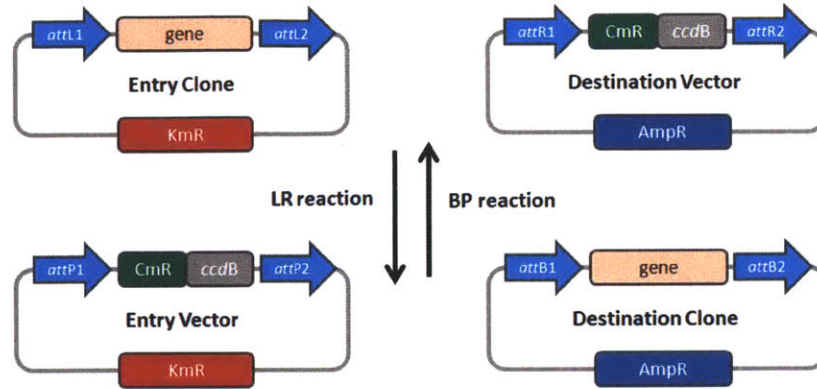
Figure 2-2: Gateway© reaction. The BP reaction recombines matching attP and attB sites to produce products with attL and attR sites, and the LR reaction recombines matching attL and attR sites to produce products with attP and attB sites. Entry vectors carry an KmR selection marker, and destination vectors carry an AmpR selection marker. The BP reaction creates an entry clone through the recombination of an entry vector and a gene flanked with attB1 and attB2 sites. The desired product is selected with Km. The LR reaction creates a destination clone through the recombination of an entry clone and a destination vector. The desired product is selected with Amp. Both reactions take 1 to 3 hours at room temperature. Figure is adapted from [2].

2.1.1 Gateway© Cloning

Gateway© cloning is an assembly method developed by Life Technologies that involves the use of bacteriophage λ recombinases, which are highly site-specific and involve no gain or loss of nucleotides [49, 69] (Figure 2-2). The system is based on two reactions, the BP reaction and the LR reaction [49]. In the BP reaction, an attB-flanked DNA fragment and an attP-containing donor vector are recombined to produce an attL-containing entry clone and an attR-containing by-product. The BP Clonase^{TM} enzyme mix, which contains the λ Integrase (Int) and *Escherichia coli* Integration Host Factor (IHF) proteins, catalyzes this reaction. In the LR reaction, an attL-containing entry clone and an attR-containing destination vector are recombined to produce an attB-expression entry clone and an attP-containing by-product. The LR Clonase^{TM} enzyme mix, which contains the λ Int and Excisionase (Xis) and *E. coli* Integration Host Factor (IHF) proteins, catalyzes this reaction.

In this system, directionality is mediated through the use of different mutant re-

combination sites. For example, *att*B1 will recombine only with *att*P1, and *att*B2 will recombine only with *att*P2. Selection of desired plasmids is achieved through the use of various antibiotics and the *ccd*B gene. Entry clones, which contain a DNA fragment of interest flanked by *att*L1 and *att*L2 sites, carry a kanamycin resistant (Km^R) marker, and destination vectors, which contain a chloramphenicol-*ccd*B cassette flanked by *att*R1 and *att*R2 sites, carry an ampicillin resistant (Ap^R) marker. *ccdB* encodes for a cytotoxic protein that targets *E. coli* DNA gyrase and results in cell death [11]. Only cells with particular mutations in their DNA gyrase can propagate *ccd*B-containing plasmids. Thus, entry clones created by BP reactions are selected for with kanamycin. Expression clones created by LR reactions are selected for with ampicillin. Cells that harbor destination vectors, although they carry an Ap^R marker, also carry *ccdB* and thus will die.

Destination vectors can carry features of interest, which can include different promoters, N- or C- terminal tags, origins of replications, or auxotrophic or antibiotic-resistance markers. Gateway$^©$ cloning allows for the rapid movement of genes of interest present in entry clones into various destination vectors.

2.1.2   Gibson Cloning

Gibson reaction is a single isothermal method to assemble multiple overlapping DNA molecules [41]. T5 5' exonuclease, Phusion DNA polymerase, and *Taq* DNA ligase are combined with linearized DNA fragments flanked by 40 bp homology regions. In the reaction, the exonuclease first chews back at the 5' ends of double-stranded DNA molecules, revealing single-stranded DNA ends. Subsequently, complementary single-stranded DNA ends anneal, the polymerase fills in the gap, and the ligase repairs the nicks. The product is a circularized DNA molecule that links multiple fragments through their homologous ends.

2.2   Rapid Assembly of Large Yeast Constructs

In this thesis, Gateway$^©$ cloning was extended and combined with Gibson cloning to allow for the rapid assembly of multiple promoter-gene pairs for yeast expression.
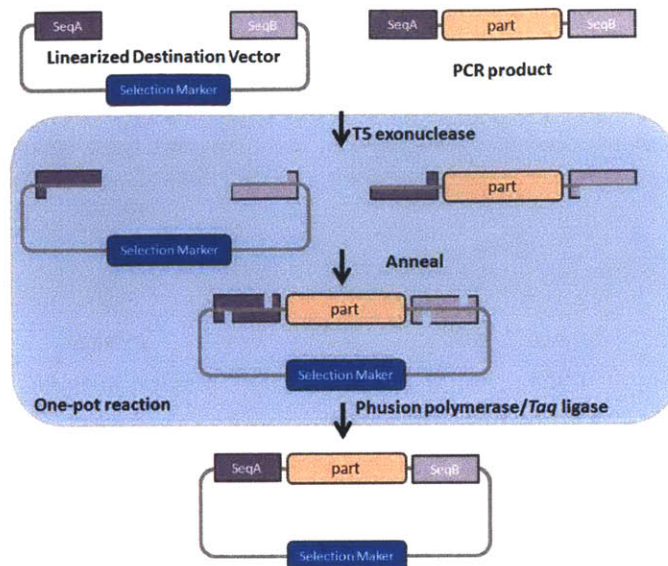
Figure 2-3: Gibson one-pot isothermal reaction. Multiple DNA molecules with overlapping ends are assembled through the action of a 5' exonuclease, a polymerase, and a ligase. The reaction takes 1 hour at 50 °C. Figure is adapted from [41].

This assembly method for yeast constructs was developed with Deepak Mishra in the Weiss Lab and is an extension of an existing assembly method for mammalian constructs, developed by Patrick Guye and Yinqing Li also in the Weiss Lab.

A suite of Gateway© cloning vectors for yeast has been previously developed [2]. However, promoters are fixed on destination vectors, and their destination vectors allow only for the expression of one gene on each plasmid.

Thus, standard Gateway© cloning was first expanded to allow for promoter swapping, through the addition of a pENTR-5' Multi-Site Gateway© Element. In addition to the using attL1 and attL2 containing entry clones, I also constructed attL4 and attR1 containing entry clones and attR4 and attR2 containing destination vectors. I constructed a library of promoters in attL4 and attR1 containing entry clones and a library of genes in attL1 and attL2 containing entry clones (Figures 2-4a-b and 2-5a-b).[1] As before, an attB1-gene-attB2 PCR fragment and a attP1-Cm^R-ccdB-attP2 donor vector (pDONR^{TM}221 from Life Technologies) were recombined using BP Clonase^{TM}

---

[1]In this thesis, a gene refers to a CDS (coding sequence). Thus, a gene starts with an ATG start codon and ends with a stop codon.

pEntr-L4-promoter-R1
TEF
CMV
ADH1
ADH1/lacO
$CYC1_{min}$/tetO7
SSRE
TR-SSRE

(a) Promoter entry clone library.

pEntr-L4-gene-R1
mLacI
tTA
rtTA
mJak2
mStat5b-HKRR
mKate
mCherry
eYFP
eCFP
eGFP
KanMX4

(b) Gene entry clone library.

pDest-R4-ccdb-R2
HIS3, CEN/ARSH
URA3, CEN/ARSH
HIS3, 2μ
LEU2, 2μ
TRP1, 2μ
URA3, 2μ

(c) Destination vector library.

pCarrier-R4-ccdb-R2
HIS3, CEN/ARSH
URA3, CEN/ARSH
HIS3, 2μ
LEU2, 2μ
TRP1, 2μ
URA3, 2μ

(d) Carrier vector library.

Figure 2-4: Entry, destination, and carrier vectors created in this thesis (with the exception of fluorescent protein entry vectors, which were created by Deepak Mishra). Promoters were cloned into an existing entry vector backbone at EcoRI and XhoI sites to create promoter entry clones. Genes with *att*B sites were recombined with pDONR^{TM}221 to create gene entry clones. The Gateway© *CmR-ccd*B cassette was cloned into pRS413, pRS416, and pRS423-426 at SacI and XhoI sites to create destination vectors. A *Seq1*-PacI recognition site-*SeqX* cassette was cloned into the same pRS plasmids at the same sites to create carrier vectors.

to produce an *att*L1-gene-*att*L2 entry clone. Kozak sequences of $AAAAAA$ were placed before the start codons of genes to increase their expression in yeast [47, 84]. Restriction digestion and ligation was used to clone a promoter of interest to create an *att*L4-promoter-*att*R1 entry clone. Subsequently, an *att*L4-promoter-*att*R1 entry clone, an *att*L1-gene-*att*L2 entry clone, and an *att*R4-Cm^R-*ccd*B-*att*R2 destination vector were recombined to produce a *att*B4-promoter-*att*B1-gene-*att*B2 expression clone.

I constructed a library of destination vectors by modifying the existing pRS suite of yeast shuttle vectors [106, 20]. Yeast shuttle vectors carry both an *E. coli* and yeast origin of replication so that the plasmid can be propagated in both organisms. The pRS suite contains Yeast Centromere Plasmids (YCp), which each have a yeast centromere sequence (CEN) and an autonomously replicating sequence (ARS), and Yeast Episomal Plasmids (Yep), which each have a high copy 2μ origin of replication. YCps maintain plasmid copy number at 1-2 copies per cell. In particular, these plasmids carry *ARSH4* and *CEN6*, originally chosen because of their small length and relative dearth of intervening restriction sites. Different YCps also carry different auxotrophic markers. YCps pRS413, pRS414, pR4315, and pR416 contain the *HIS3*, *TRP1*, *LEU2*, and *URA3* auxotrophic markers, respectively. In contrast, YEps exist in yeast at higher, more variable copy number (60 - 80 per cell) [36]. YEps pRS423, pRS424, pR4325, and pR426 contain the *HIS3*, *TRP1*, *LEU2*, and *URA3* autotrophic markers, respectively. In this thesis, six YCps and YEps were modified through the addition of the *att*R4-Cm^R-*ccd*B-*att*R2 cassette to form destination vectors (Figures 2-4c and 2-5c). Downstream of this cassette in centromere-based plasmids is a *CYC1* terminator, and downstream of this cassette in episomal-based plasmids is a *ADH1* terminator. Thus, a library of destination vectors was created. These destination vectors can only carry one promoter-gene pair.

Second, to extend the system to assemble multiple promoter-gene pairs through Gibson cloning, destination vectors were modified to become position vectors. A variable number of position clones, each containing a promoter-gene pair created through LR recombination, could then be assembled into a carrier vector through Gibson
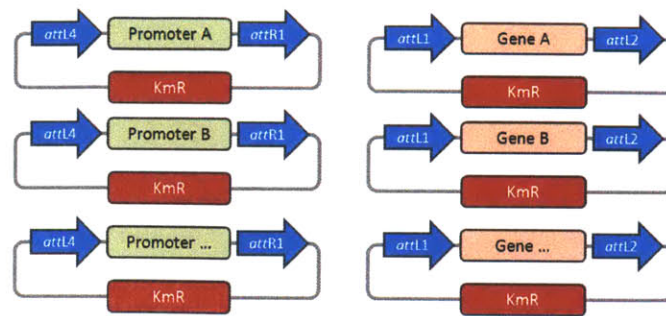
(a) Promoter clone library.

(b) Gene clone library.
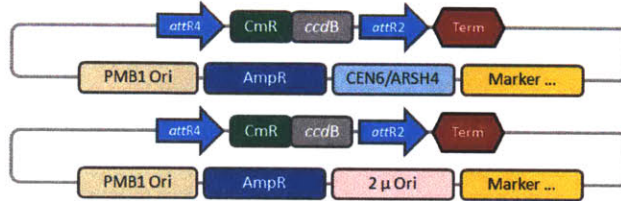
(c) Destination vectors.
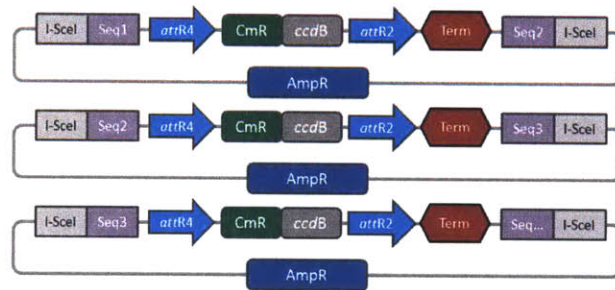
(d) Position vectors.
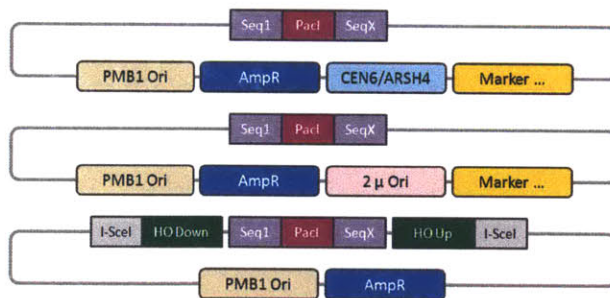
(e) Carrier vectors.

Figure 2-5: Cartoon representation of elements on (a)-(b) entry, (c) destination, (d) position, and (e) carrier vectors. Not depicted on entry clones and position vectors are *E. coli* origin of replications. Markers can be any one of HIS3, LEU2, TRP1, URA3 genes.
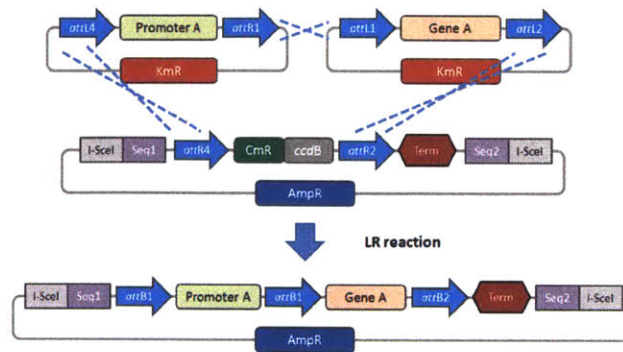
(a) Step 1: LR reaction.
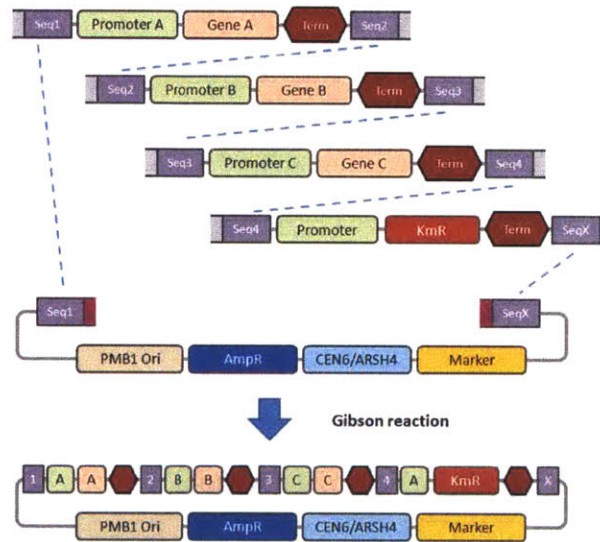
(b) Step 2: Gibson reaction.

Figure 2-6: Two step multi-gene pair assembly. (a) In the first step, LR reactions are carried out to create individual promoter-gene pairs on sequential position vectors and selected for with Amp. (In this example, Promoter A and Gene A are recombined into Position 1-2, Promoter B and Gene B into Position 2-3, and Promoter C and Gene C into Position 3-4.) (b) In the second step, position vectors carrying desired promoter-gene pairs are digested with I-SceI, along with a Km adapter. (In this example, the Km adapter is in Position 4-X). These linearized fragments are combined with a carrier vector cut with PacI and mixed with the Gibson reaction mix [41]. The final vector is selected for with Amp and Km.

cloning to create a single multi-promoter-gene construct (Figure 2-5d-(e)). Position vectors were created by Patrick Guye and Yinqing Li. Nine 45 bp position sequences, termed *Seq1* through *Seq8* and *SeqX*, were selected on the basis to contain as little homology with each other as possible. Position vectors were created from destination vectors through the addition of consecutive numerical position sequences in regions flanking the *att*R4-Cm^R-*ccd*B-*att*R2 cassette. To linearize the position vector such that the position sequences were at the exposed 5' and 3' ends, I-SceI restriction sites were placed immediately outside the region containing the cassette and the position sequences. I-SceI is a homing enzyme that has a 30 bp long recognition sequence [58]. This enzyme was used here because there is a very small probability that its recognition sequence would appear in any of the promoters or genes used.

Lastly, I created carrier vectors with a *Seq1*-PacI restriction site-*SeqX* cassette from the YCps pRS413, pRS414, pR4315, and pR416 (Figures 2-4d and 2-5e). A PacI site was included to linearize the carrier vector with the PacI restriction enzyme, such that *Seq1* and *SeqX* became exposed at either ends after digestion. To increase the reliability of Gibson cloning, vectors containing a kanamycin resistance marker, flanked by a numerical position sequence at one end and *SeqX* at the other end, were created. It was found that inclusion of this kanamycin adapter as one of the position vectors increased the efficiency of Gibson cloning, when the final construct was selected for with both ampicillin and kanamycin.

To summarize, to create a multi-promoter-gene assembly with this assembly method, individual promoter gene pairs are first recombined into sequential position vectors with LR Clonase^{TM} (Figure 2-6a). These position vectors are then digested with I-SceI, mixed with a linearized kanamycin adapter and a linearized carrier vector, and then added to the Gibson mix [41] (Figure 2-6b). Cells are transformed with the resulting product and selected on ampicillin-kanamycin media plates.

The entire process of assembling a yeast circuit from libraries of genes and promoters takes a total of 4 days and involves a total of two cloning steps. This process also involves no gel extractions, a time-consuming process that often involves significant loss of DNA. In addition, PCR is not used in any of the steps, so if sequence-verified

Correctly Restriction Mapped Clones
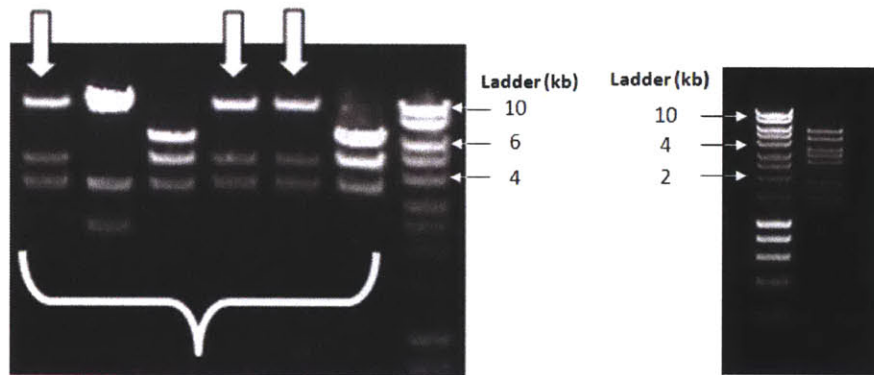
(a) Digestion with StuI.　　　　(b) Digestion with ScaI.

Figure 2-7: Restriction map of assembled construct in Figure 4-1b (System II) in a URA centromeric plasmid (derived from pRS416). (a) Six clones from the transformation were digested with StuI. Three clones restriction mapped correctly. Bands at 3.4, 3.4, 4.4, and 9.9 kb were expected. Hyperladder I was used, and the 4, 6, and 10 kb bands are highlighted in the figure. (b) Because two of the bands from the restriction mapping with StuI were close, one correctly restriction mapped clone was selected and digested with ScaI. Bands at 1.4, 1.8, 2.5, 2.8, 3.2, 4.3, and 5.3 were expected, and the 2, 4, and 10 kb bands are highlighted in the figure. The match between the bands in the gel and the expected bands suggests that the construct was assembled as expected.

entry clones are used, the final assembled circuit will be sequence-correct. The fact that this process does not use PCR is particularly beneficial in assembling large genetic circuits. PCR can introduce errors into the DNA sequence, and the probability that PCR will introduce errors into large circuits, with lengths in the tens of kilobases, is not negligible. A recent study that assembled the 16.3 kilobase mouse mitochondrial genome from 60 bp oligos with PCR found that the error rate of PCR was 3.78 10^{-4} [40]. As an example of the functionality of this assembly technology, a construct used to measure retroactivity to the output (Figure 4-1b) was constructed with this method (Figure 2-7).

An integration based carrier vector is currently being constructed. It was previously found that about 10% to 30% of cells in selective liquid culture do not contain plasmids and that cells in non-selective liquid culture lose their plasmids at a rate of 1% to 5% per doubling [36, 20]. Integration into the chromosome stabilizes the synthetic construct and increases the reliability of its expression, reducing the probability that cells will lose the circuit, even under selection. The integration based carrier vector will homologously recombine into the *HO* site in yeast strains. HO encodes for an endonuclease that facilitates conversion between **a** and α mating types and promotes diploidization [117]. Thus, almost all laboratory strains have a *HO* mutation. The *Seq1*-PacI-*SeqX* cassette will be flanked by regions homologous to *HO*, and I-SceI cut sites will be present outside of this region such that the plasmid can be easily linearized (Figure 2-5e). In preparation to use the integration based carrier vector, I created a series of position vectors with the *TEF* promoter and the *KanMX4* resistance gene. *KanMX4* confers resistance to the antiboitic G418, which will be used to select against cells that do not have the integrated construct [118].

2.3   Discussion

Establishment of a Gateway©-Gibson assembly method can expedite the construction and tuning of the synthetic circuits used not only in this project, but also in other yeast strain engineering projects. This assembly method does not require the use of gel extraction and allows for specification of parts order. In addition, the

method does not require the use of PCR and thus preserves DNA sequence integrity, as DNA strands are shuffled around only. Use of PCR in an assembly method can be problematic when the sequences are GC rich, and amplifying constructs that are greater than 10 kb can result in undesirable sequence mutations. Furthermore, the method does not have forbidden sites in promoters or genes since restriction enzymes with short restriction sites are not used. Therefore, site directed mutagenesis to remove internal restriction sites is not necessary. Also, the assembly method utilizes parallel over serial or hierarchical assembly for speed and can currently be scaled up to include eight promoter-gene pairs. Lastly, this process is amenable to automation and can be carried out with a liquid-handling robot, which allows for the high-throughput automated DNA assembly.

Currently, the assembly method does leave attB sites in between promoters, genes, and terminators. The presence of these sites is unknown, as the activity of parts, in particular promoters, can be sequence dependent. However, many other systems have created constructs with Gateway© cloning that leave attB sites in between the promoter and the gene [2, 76]. Therefore, if the sites do reduce promoter activity, they do not eliminate it. In addition, the current system does not allow for easy assembly of 3' or 5' protein fusions to genes. However, cloning strategies involving PCR of genes and restriction enzymes have been developed that allow for the creation of protein fusions in gene entry clones [4, 23].

Overall, this assembly method allows for rapid parallel and combinatorial assembly of constructs for yeast expression. Promoters and genes can be readily swapped for varied expression levels and functions, and large circuits can be assembled in a two step process that takes four days. This novel development can facilitate the advancement of not just this project on retroactivity and insulation, but of synthetic biology as a field.

Chapter 3

Modeling Retroactivity and Insulators

Retroactivity occurs because signal transmission in genetic circuits relies on the physical binding of proteins generated within a module to molecules outside of the module. When these signal proteins bind to outside molecules, they are effectively sequestered and can no longer transmit the signal. Thus, downstream components can affect the performance of the upstream module. In this sense, retroactivity can be thought of as the reverse signal from downstream to upstream components that occurs at a junction. An insulator is a device that, when placed at the interface between two devices, isolates the upstream and downstream devices from undesirable retroactive, impedance-like effects. The ideal insulator has zero retroactivity to the input and attenuates retroactivity to the output.

This chapter establishes the theoretical framework for investigating retroactivity in genetic regulatory networks and the performance of a generic phosphorylation - dephosphorylation insulator. Such theoretical understanding is necessary to understand what genetic devices can experimentally verify such phenomenon *in vivo*. In Section 3.1.1, I summarize the derivation of retroactivity presented in [22], which is rederived in Section A.1. In Section 3.1.2, I investigate how assumption made in [22] that transcription factor-DNA complexes do not degrade (where as unbounded transcription factors do) affects modeling of the system. I found that this assumption

greatly changes behavior of the system, and that retroactivity is higher in systems where the complex is assumed to degrade because of lower transcription factor steady state concentrations. In Section 3.2.1, I summarize the derivation of an insulator that utilizes a phosphorylation - dephosphorylation cycle presented in [22], which is rederived in A.2. In Section 3.2.2, I discuss parameter sensitivity studies of such a phosphorylation - dephosphorylation cycle based insulator. I also found that the assumption made in [22] that the phosphorylatable protein and the phosphatase do not decay does not greatly change insulator behavior, as long as the decay rate is small.

3.1 Retroactivity

3.1.1 Modeling Retroactivity

Retroactivity occurs when direct interconnection between two devices affects the performance of the upstream device. Because downstream components affect upstream device behavior, retroactivity can be thought of as the reverse signal from downstream to upstream components that occurs at a junction.

Let X be a transcription factor that is produced at rate $k(t)$ and decays at rate δ, as presented in Equation (3.1).[1]

$$
\emptyset \xrightarrow{k(t)} X
$$
$$
X \xrightarrow{\delta} \emptyset
$$

(3.1)

Then the differential equation that models the rate of change of the amount of X is presented in Equation (3.2).

$$
\frac{dX}{dt} = k(t) - \delta X
$$

(3.2)

Let p denote a binding site of X in a promoter. X can reversibly bind to p to form

---

[1]For convenience, we let X denote species itself when presented in a chemical equation, and let X denote the concentration of that species when presented in a rate equation. The subscript T, when used in a rate equation, denotes the total concentration of a particular species in the system, whether free or bound to other molecules.
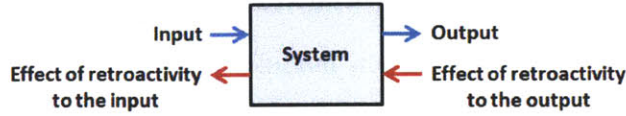
Figure 3-1: Retroactive flows in a system. Retroactivity is a "reverse" flow from downstream components to upstream components. This reverse flow is undesirable in engineering systems because it destroys component modularity, making the dynamics of a component dependent on its external interconnections. Figure is adapted from [22].

the complex C. The new system is presented in Equation (3.3).

$$\emptyset \xrightarrow{k(t)} X$$
$$X \overset{\delta}{\underset{}{\rightleftharpoons}} \emptyset \tag{3.3}$$
$$X + p \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} C$$

The associated differential equations in Equation (3.4) change because some of X is sequestered in its binding to p. Since the total number of promoters in the system is conserved, $p_T = p + C$, so $p = p_T - C$. The effect of retroactivity on the dynamics of X is represented by the boxed terms.

$$\frac{dX}{dt} = k(t) - \delta X + \boxed{k_{off}C - k_{on}(p_T - C)X}$$
$$\frac{dC}{dt} = -k_{off}C + k_{on}(p_T - C)X \tag{3.4}$$

Thus, if X was the output of module A, and p were binding sites in module B, then module B is referred to as being "downstream" of module A. The presence of module B creates "retroactivity to the output" in module A and module B itself experiences "retroactivity to the input." Thus, retroactivity can be thought of as adding two flows to the standard input/output flows of a device (Figure 3-1). Each device experiences a "retroactivity to the output" flow and creates a "retroactivity to the input" flow.

It was discovered in [22] (and rederived in Section A.1) that $\mathcal{R}(\bar{X}) = \frac{1}{1+\frac{(1+\bar{X}/K_d)^2}{p_T/K_d}}$ is a measure of the effect of retroactivity to the output on $\bar{X}$ dynamics.[2] In particular,

---

[2]$\bar{X}$ is the approximation of X after singular perturbation is applied, where ϵ is approximated as 0. See [22] or Section A.1 for details.

$\mathcal{R}(\bar{X})$ quantifies how much $\frac{d\bar{X}}{dt}$ changes when the system is isolated to when the system is connected to downstream modules. Note that $\mathcal{R}(\bar{X})$ is always between values 0 and 1. This value is large when $\frac{(1+\bar{X}/K_d)^2}{p_T/K_d}$ is small. This demonstrates that retroactivity has the greatest effect on the output in systems where (1) there is a heavy load (large p_T), (2) the signal carrier has a high affinity to the binding sites in downstream promoters (small K_d), or (3) the signal carrier is at low numbers relative to the number of binding sites (small $\bar{X}$). Figure 3-2 demonstrates that when the load is much greater than the steady state concentration of X, the effect of retroactivity on the dynamics of X becomes apparent.

3.1.2 Addition of Complex Decay in Modeling Retroactivity

In the previous analysis, it was assumed that the complex formed by the transcription factor and the promoter binding site did not decay. Here, I demonstrate that whether the complex degrades has a nontrivial effect on system behavior.

If the complex decays, the chemical equations presented in Equation (3.3) change to the following:

$$\emptyset \xrightarrow{k(t)} X$$
$$X \xrightarrow{\delta} \emptyset$$
$$C \xrightarrow{\delta} \emptyset \tag{3.5}$$
$$X + p \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} C$$

and the differential equations presented in Equation (3.4) change to the following:

$$\frac{dX}{dt} = k(t) - \delta X + \boxed{k_{off}C - k_{on}(p_T - C)X}$$
$$\frac{dC}{dt} = -k_{off}C + k_{on}(p_T - C) - \delta C \tag{3.6}$$

Although the boxed terms representing the effect of downstream devices on X dynamics remains the same, the dynamics of C change.

In the following plots, I found that retroactivity is much greater when it is assumed that the complex does decay (Figures 3-3, 3-5, and 3-6). This is because steady state levels of X become lower and dependent on the load. Under either assumption,

42
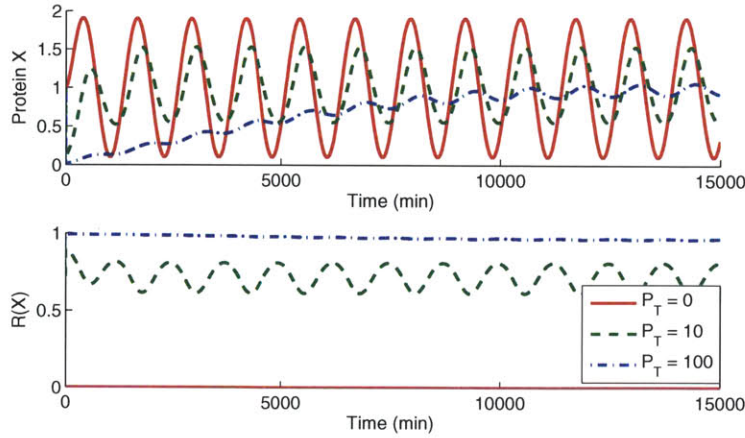
(a) Steady state X concentration of 1
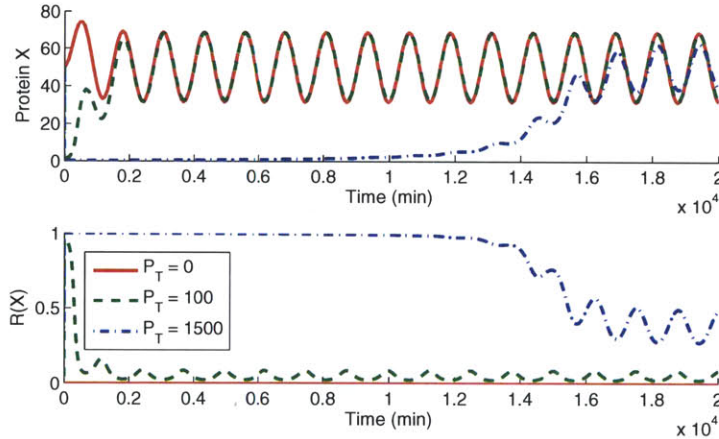
(b) Steady state X concentration of 50

Figure 3-2: Effect of retroactivity on buffer performance. An uninsulated buffer (Equation (3.4)) was simulated with two different parameter sets. Both simulations demonstrate that when the load is much greater than the steady state X concentration, the effect of retroactivity to the output on the dynamics of X becomes apparent. $\mathcal{R}(\bar{X})$ serves as a measure of the effect of retroactivity to the output on $\bar{X}$ dynamics. Unless otherwise noted, all simulations were conducted in Simulink by MATLAB with ODE23s. (a) Parameter set A: $k(t) = 0.01(1 + sin(\omega t))$, $\omega = 0.005$, $\delta = 0.01$, $X_{init} = 1$, $k_{on} = 10$, and $k_{off} = 10$, obtained from [22]. The choice of $\delta = 0.01$ results in a protein half-life of about 1 hr, and the choice of $\omega = 0.005$ results in oscillations with a period of about 20 hr, which is within the range of oscillatory clocks constructed in [5]. The steady state concentration of X in this system without load is 1. (b) Parameter set B: $k(t) = 0.1(1 + sin(\omega t))$, $\omega = 0.005$, $\delta = 0.002$, $X_{init} = 50$, $k_{on} = 10$, and $k_{off} = 10$. The steady state concentration of X in this system without load is 50.

retroactivity increases as the load increases (Figure 3-7). However, if complexes are assumed to not degrade, retroactivity introduces phase lag (Figures 3-4 and 3-7). If degradation rates are high, the effect of modeling complex degradation on the system is larger (Figure 3-8). In addition, modeling complex degradation is necessary to observe the expected relation between the protein-DNA disassociation constant and retroactivity (Figure 3-9). Lastly, if the complex degradation is modeled, then higher frequencies result in a sharper output amplitude versus load curve (Figure 3-10). Note that $\mathcal{R}(\bar{X})$ is not a very meaningful value when the complex is assumed to decay, as shown in Section A.2. These simulations demonstrate that the assumption that the bound transcription factor does or does not degrade is nontrivial.
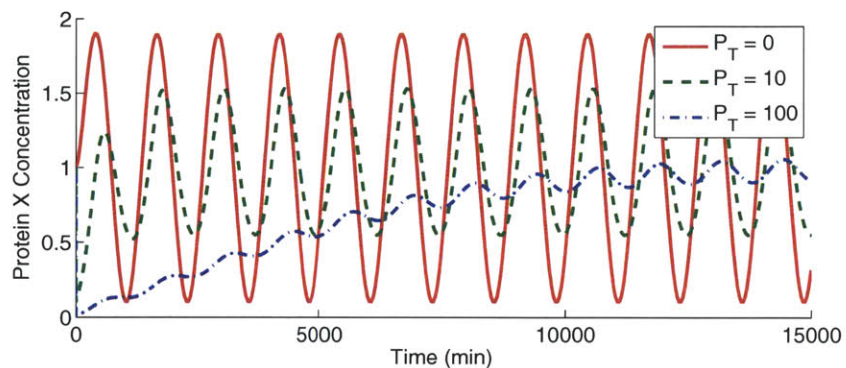
(a) Without C decay
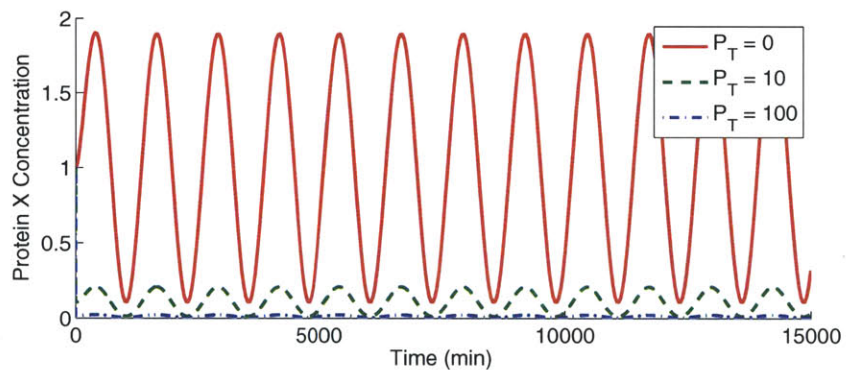
(b) With C decay

Figure 3-3: Effect of complex decay on buffer performance. Simulation of uninsulated buffer with parameter set A, with and without complex decay. The effect of retroactivity on the dynamics of X is much greater when complex C decays than when it does not. (a) When the complex does not degrade, a "reservoir" of X builds up that buffers against retroactivity to the output. (b) When the complex does degrade, complex formation no longer prevents the transcription factor from degradation. This leads to greater retroactivity.
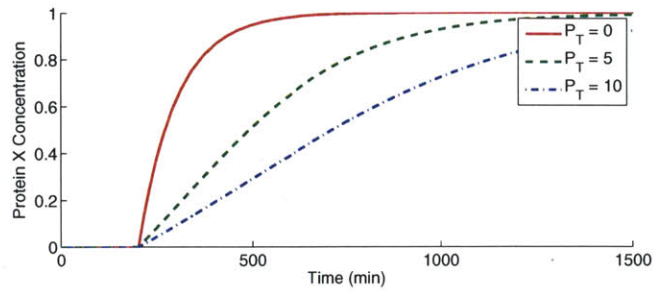
(a) Without C decay
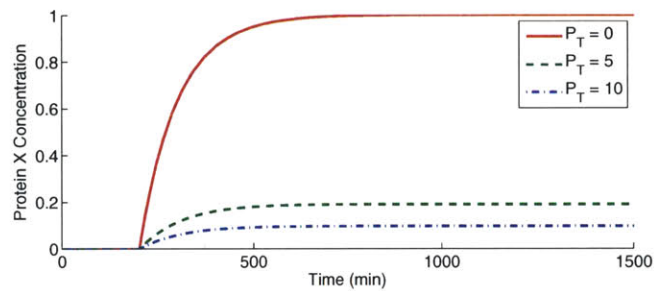
(b) With C decay
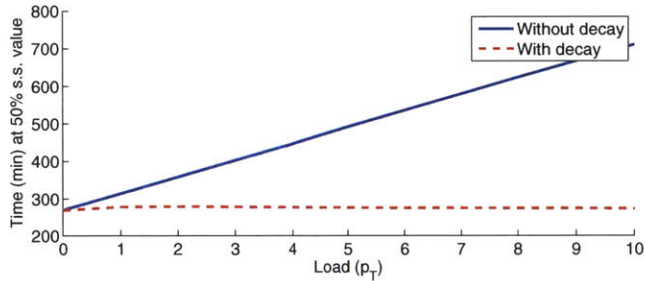
(c) At various loads

Figure 3-4: Change in dynamics of X with step function input. Simulation of uninsulated buffer with parameter set A, with and without complex decay. At 200 min, k(t) steps from 0 to 0.01. (a) When the complex does not degrade, the time it takes for the system to reach steady state increases with greater load. (b) When the complex does degrade, the time it takes for the system to reach steady state does not change with greater load. (c) The time it takes for the system to reach 50 percent of steady state is plotted. Note that this time includes the initial 200 min.

(a) $p_T = 10$, no C decay   (b) $p_T = 50$, no C decay   (c) $p_T = 100$, no C decay

(d) $p_T = 10$, with C decay   (e) $p_T = 50$, with C decay   (f) $p_T = 100$, with C decay
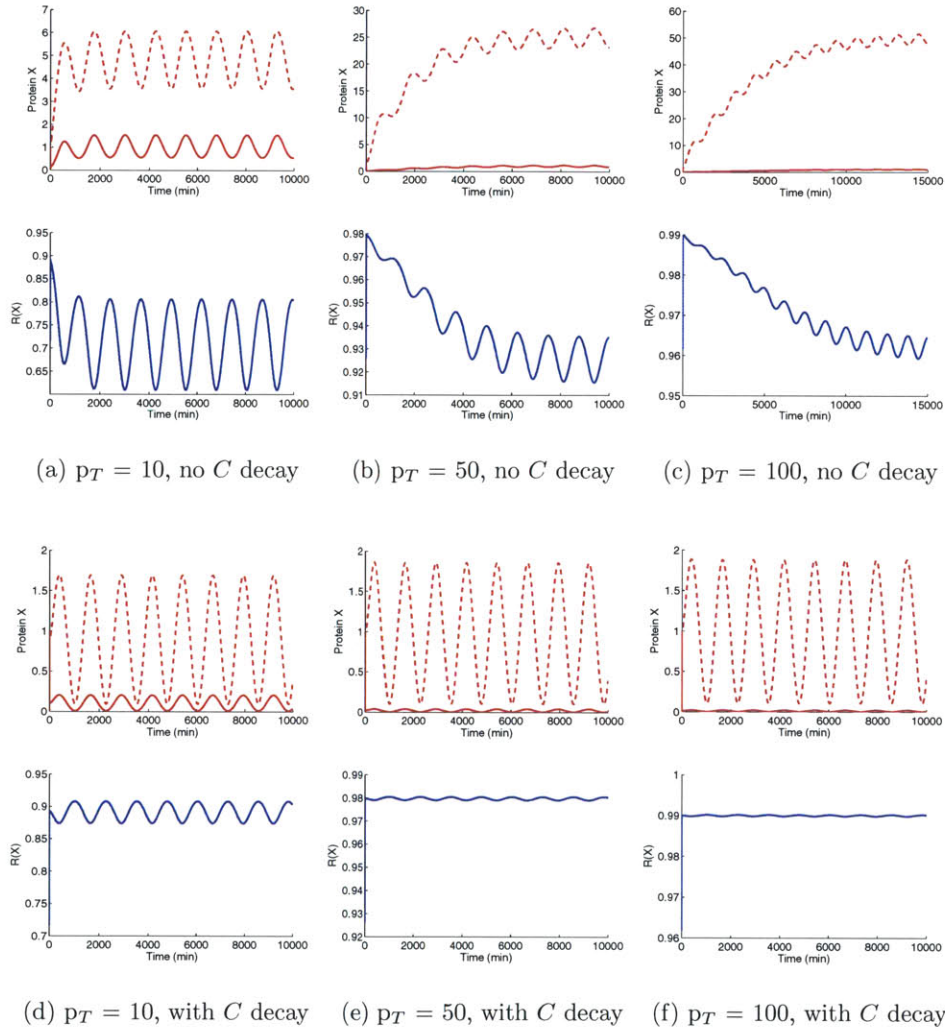
Figure 3-5: Effect of complex decay on buffer performance, parameter set A (no load s.s. $X = 1$). The solid red line represents the concentration of X, and the dashed red line represents the concentration of C. The load is increased from 10 to 50 to 100 binding sites. In (a)-(c), C does not degrade. The steady state value of X remains at 1 despite changes in the load. In (d)-(f), C does degrade. The steady state value of X decreases as the load increases.

(a) $p_T = 10$, no C decay   (b) $p_T = 50$, no C decay   (c) $p_T = 100$, no C decay

(d) $p_T = 10$, with C decay   (e) $p_T = 50$, with C decay   (f) $p_T = 100$, with C decay
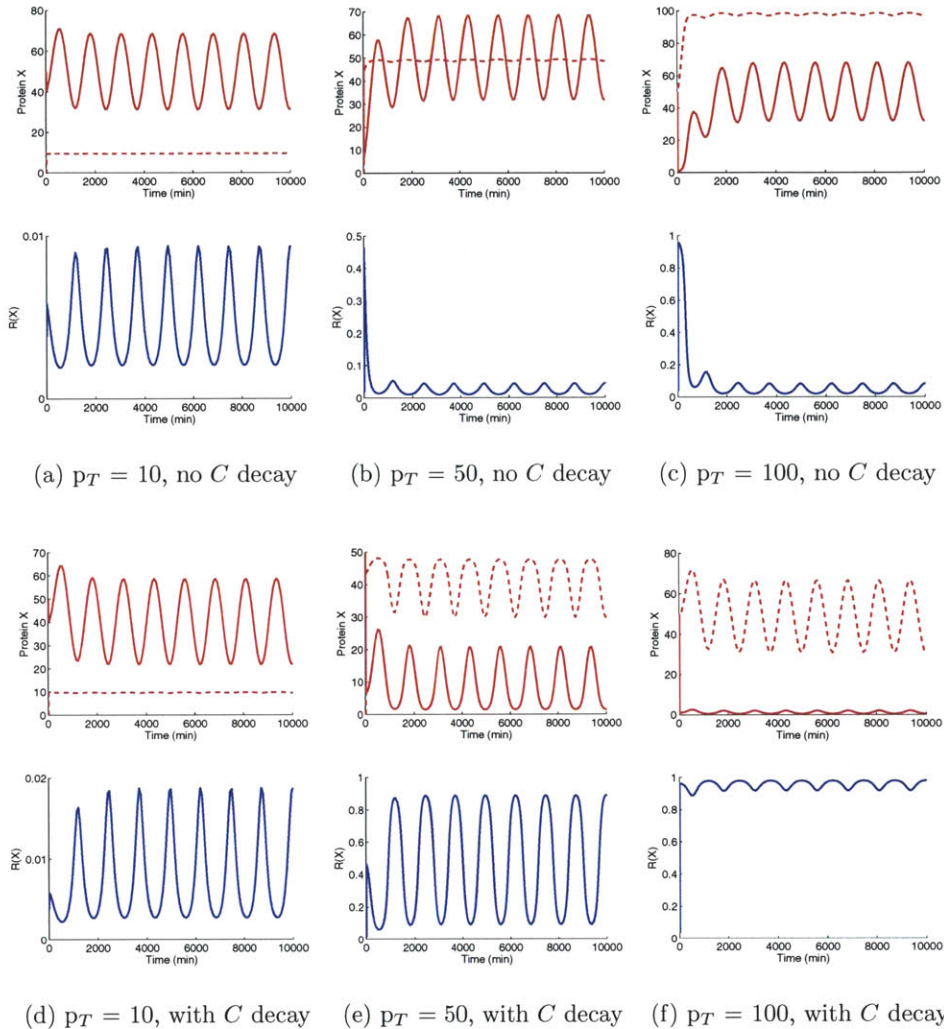
Figure 3-6: Effect of complex decay on buffer performance, parameter set B (no load s.s. $X = 50$). The solid red line represents the concentration of X, and the dashed red line represents the concentration of C. The load is increased from 10 to 50 to 100 binding sites. In (a)-(c), C does not degrade. The steady state value of X remains at 50 despite changes in the load. In (d)-(f), C does degrade. The steady state value of X decreases as the load increases. The change in retroactivity per additional binding site is greatest when numbers of binding sites and transcription factors are roughly equal.

48

(a) no load s.s. $[X] = 1$ (b) no load s.s. $[X] = 50$
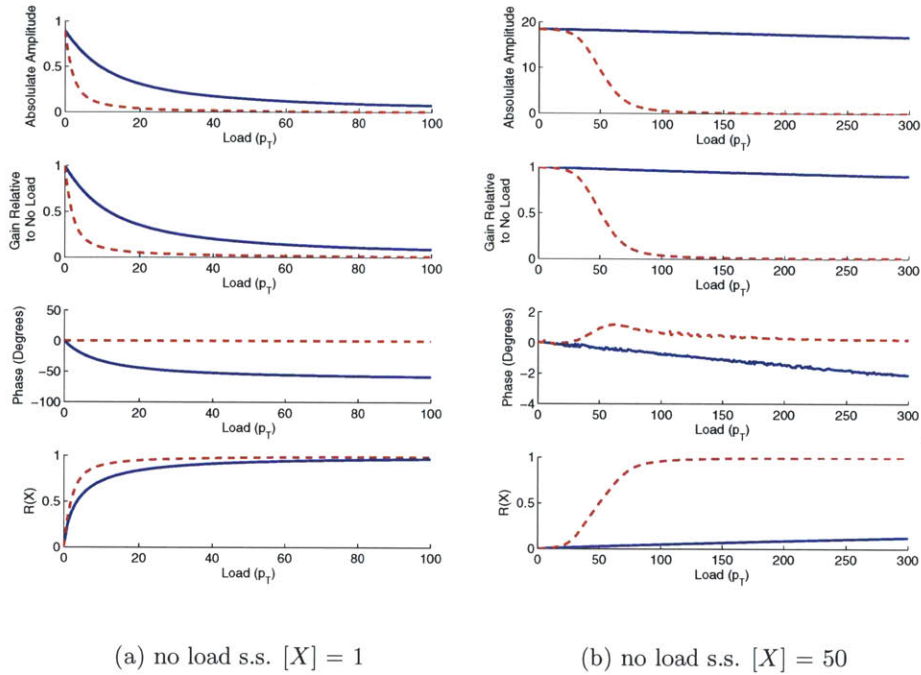
Figure 3-7: Gain, phase lag, and $\mathcal{R}(\bar{X})$ of buffer versus varied load. Solid blue line represents system behavior without complex decay. Dashed red line represents system behavior with complex decay. Complex decay increases the effect of retroactivity on the gain. A phase lag is present only when the complex is assumed to not decay. (a) Parameter set A. Steady state concentration of X without load is 1. (b) Parameter set B. Steady state concentration of X without load is 50. Assuming the complex decays, retroactivity to the output is minimal when there are fewer binding sites than transcription factors in the system. When the number of binding sites is roughly equal to or greater than the number of transcription factors in the system, retroactivity becomes significant.

(a) no C decay

(d) with C decay

(b) no C decay

(e) with C decay

(c) no C decay

(f) with C decay
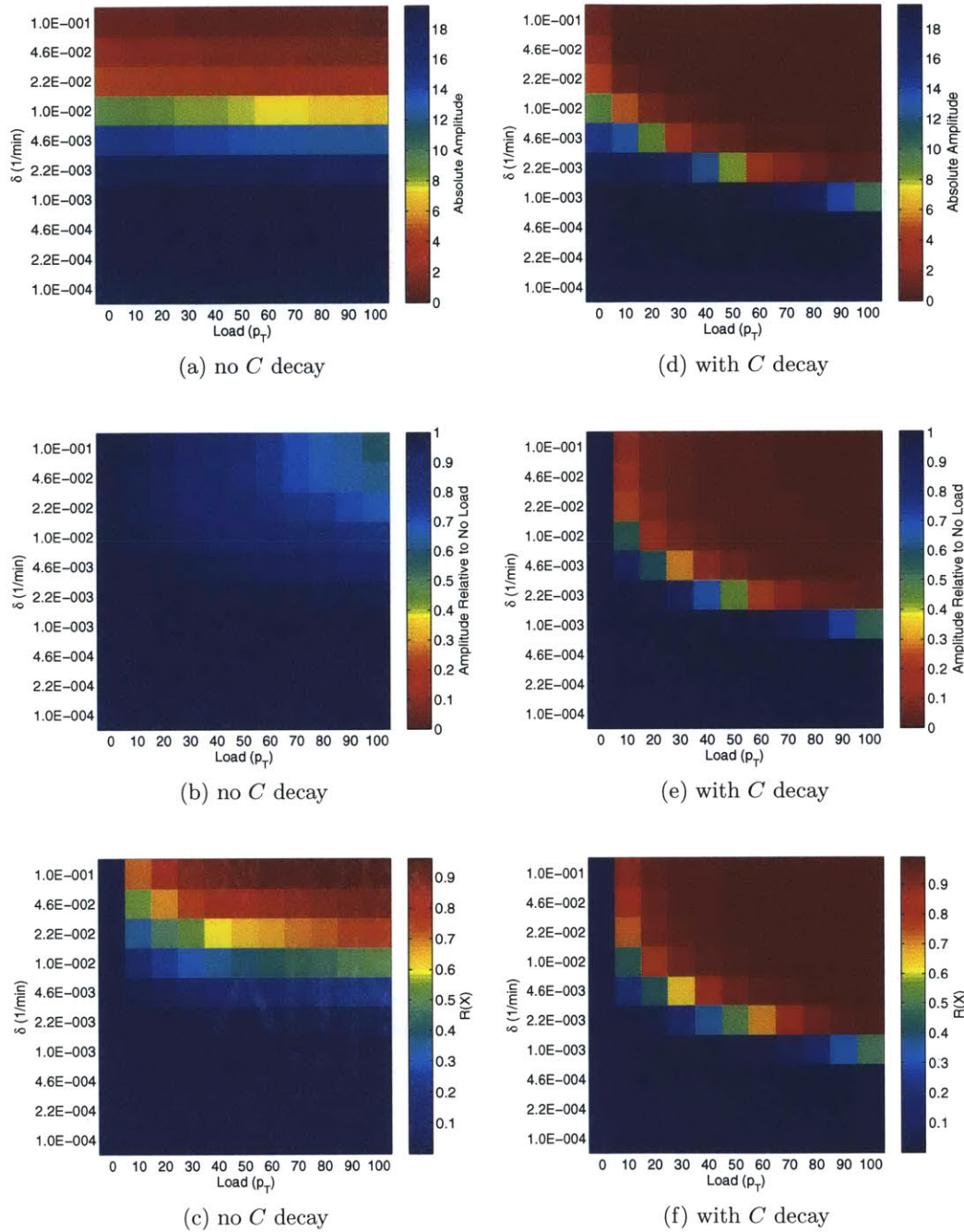
Figure 3-8: Effect of degradation rate on buffer performance. (a)-(c) The complex C is assumed to not degrade. (d)-(f) The complex C is assumed to degrade at rate δ. Increasing the degradation rate has the effect of lowering the steady state concentration of X. At high degradation rates (low steady state $[X]$), higher loads affect buffer performance more negatively when C is assumed to degrade, and at lower degradation rates, both systems behave similarly (because the effect of whether the complex degrades decreases).
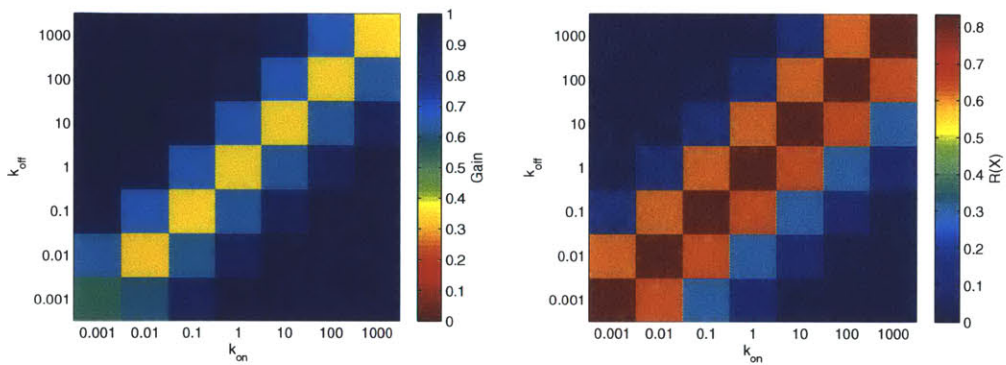
(a) Without C decay.
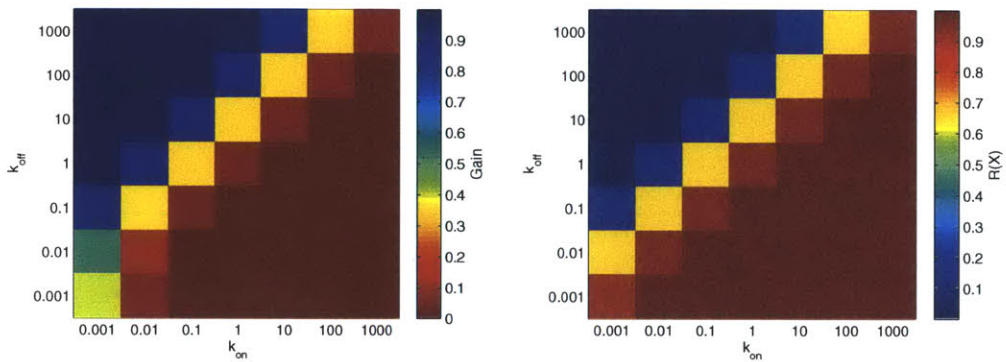
(b) With C decay.

Figure 3-9: Effect of dissociation constant on buffer performance. Only when the complex is assumed to decay does the expected effect of k_D on retroactivity appear. That is, as k_D decreases, retroactivity to the output is greater, since binding sites sequester a larger amount of transcription factors.

(a)   (b)

(c)
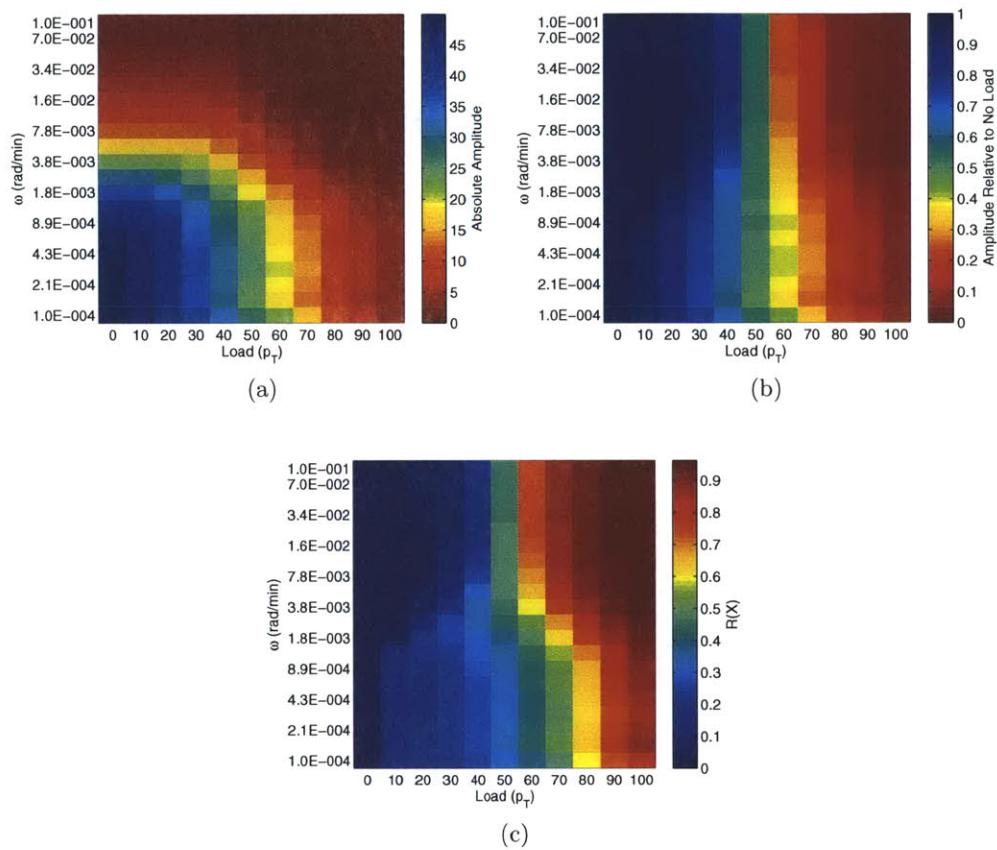
Figure 3-10: Frequency analysis of buffer with load. Simulations were carried out with parameter set B (no load s.s. $[X] = 50$) with the assumption that complexes degrade. The change in retroactivity per additional binding site when the load is equal to $[X]$ increases with higher frequencies.

3.2 Insulator

A retroactive insulation device decouples upstream and downstream transcriptional components such that device dynamics do not change upon interconnection. The insulator allows forwards propagation of the regulatory signal, while attenuating the reverse signal that occurs from retroactivity. In this section, the theoretical framework of an insulator is reviewed, and the effect of the assumption made in [22] is computationally investigated.

3.2.1 Modeling an Insulator

Block Diagram Representation
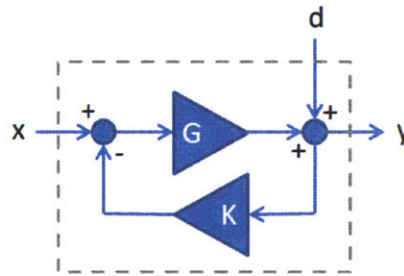
Figure 3-11: Insulator block diagram. x is the input signal into the insulator, y is the output signal. Retroactivity to the output d is modeled as an additive disturbance to the output. The system has a gain of magnitude G and a negative feedback of magnitude K. Figure is adapted from [22].

A block diagram is used to investigate the requirements needed to realize an insulator (Figure 3-11) [22]. Retroactivity to the output is represented as a disturbance added to the output. Equation (3.7) models the block diagram.

$$y = G(x - Ky) + d \tag{3.7}$$

Solving for the output y results in the following equation:

$$y = x\frac{G}{1 + GK} + \frac{d}{1 + GK} \tag{3.8}$$

53

The forward signal is represented by the first term, and retroactivity to the output is represented by the second term in Equation (3.8). Thus, the insulator should maximize the magnitude of the first term and minimize that of the second. To do so, the gain G should be as large of possible. As the magnitude of G increases, the first term approaches u/K, and the second term approaches 0. Thus, to attenuate retroactivity and preserve the unconnected dynamics of the upstream component, the insulator should have a large input amplification gain and a large negative feedback.

A phosphorylation-dephosphorylation cycle can work as an insulator because its dynamics operate on a much faster timescale than that of the genetic regulatory network it insulates. The transfer of a phosphate group occurs on the order of seconds, whereas protein production and decay occur on the order of minutes. The phosphorylation branch provides a large input amplification, and the dephosphorylation branch provides a large negative feedback. Another benefit of using a phosphorylation cycle based insulator is that it place a low metabolic strain on the cell [22]. Transfer of phosphate groups have a lower energy cost compared to the synthesis of regulator proteins.

Chemical Equations

The performance of an insulator with a two-step phosphorylation - dephosphorylation reactions is examined in this section. X is a protein that can be phosphorylated by the kinase X and dephosphorylated by the phosphatase Y. X and Z bind to form the complex C_1, and X and Y bind to form the complex C_2. It is assumed that only a free X_p molecule can bind to the promoter.

$$\emptyset \xrightarrow{k(t)} Z$$
$$Z \xrightarrow{\delta} \emptyset$$
$$X + Z \underset{\alpha_2}{\overset{\alpha_1}{\rightleftharpoons}} C_1 \xrightarrow{k_1} X_p + Z$$
$$X_p + Y \underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} C_2 \xrightarrow{k_2} X + Y$$
$$X_p + p \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} C$$

(3.9)
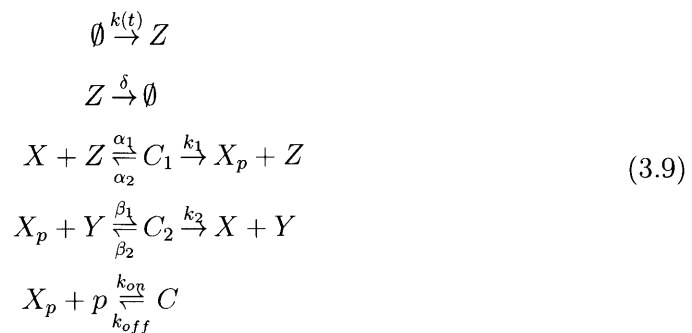
Again, total amounts of X and Y are assumed to be conserved. Z is the input signal, and X_p is the output signal. The following conservation equations result: $X_T = X + X_p + C_1 + C_2 + C$ and $Y_T = Y + C_2$. The differential equations that represent the system in (3.9) are as follows:

$$\frac{dZ}{dt} = k(t) - \delta Z \boxed{-\alpha_1 Z X_T (1 - \frac{X_p}{X_T} - \frac{C_1}{X_T} - \frac{C_2}{X_T} - \boxed{\frac{C}{X_T}}) + (\alpha_2 + k_1)C_1}$$

$$\frac{dC_1}{dt} = \alpha_1 Z X_T (1 - \frac{X_p}{X_T} - \frac{C_1}{X_T} - \frac{C_2}{X_T} - \boxed{\frac{C}{X_T}}) - (\alpha_2 + k_1)C_1$$

$$\frac{dC_2}{dt} = \beta_1 Y_T X_p (1 - \frac{C_2}{Y_T}) - (k_2 + \beta_2)C_2 \tag{3.10}$$

$$\frac{dX_p}{dt} = -\beta_1 Y_T X_p (1 - \frac{C_2}{Y_T}) + k_1 C_1 + \beta_2 C_2 + \boxed{k_{off}C - k_{on}X_p(P_T - C)}$$

$$\frac{dC}{dt} = -k_{off}C + k_{on}X_p(p_T - C)$$

Retroactivity to the input is represented by terms in the outer box in the equation for $\frac{dZ}{dt}$. The other boxes collectively represent retroactivity to the output.

3.2.2 Addition of Phosphorylatable Protein and Phosphatase Decay to Insulator Model

In Section 3.1.2, I found that assuming whether a protein bound to DNA degraded significantly affected the dynamics of an uninsulated buffer. In this subsection, I investigated whether the approximation made in [22], where total concentrations of X and Y were constant, affected the result of the modeling. If this simplification is

not made, Equation 3.9 becomes the following:

$$\emptyset \xrightarrow{k(t)} Z, \emptyset \xrightarrow{kX} X, \emptyset \xrightarrow{kY} Y$$

$$Z \xrightarrow{\delta} \emptyset, Y \xrightarrow{\delta} \emptyset, X \xrightarrow{\delta} \emptyset$$

$$X_p \xrightarrow{\delta} \emptyset, C1 \xrightarrow{\delta} \emptyset, C2 \xrightarrow{\delta} \emptyset$$

$$C \xrightarrow{\delta} p$$

$$X + Z \underset{\alpha_2}{\overset{\alpha_1}{\rightleftharpoons}} C_1 \xrightarrow{k_1} X_p + Z$$

$$X_p + Y \underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} C_2 \xrightarrow{k_2} X + Y$$

$$X_p + p \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} C$$

(3.11)
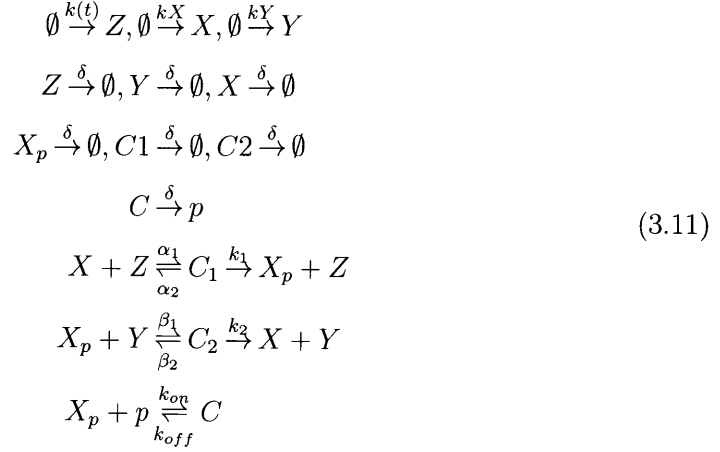
Then the differential equations that represent the system in (3.11) are as follows:

$$\frac{dZ}{dt} = k(t) - \delta Z - \alpha_1 Z X + (\alpha_2 + k_1)C_1$$

$$\frac{dC_1}{dt} = -\delta C_1 + \alpha_1 Z X - (\alpha_2 + k_1)C_1$$

$$\frac{dC_2}{dt} = -\delta C_2 + \beta_1 Y X_p - (k_2 + \beta_2)C_2$$

$$\frac{dX_p}{dt} = -\delta X_p - \beta_1 Y X_p + k_1 C_1 + \beta_2 C_2 + k_{off}C - k_{on}X_p(p_T - C)$$

$$\frac{dC}{dt} = -\delta C - k_{off}C + k_{on}X_p(p_T - C)$$

$$\frac{dX}{dt} = kX - \delta X - \alpha_1 Z X + \alpha_2 C_1 + k_2 C_2$$

$$\frac{dY}{dt} = kY - \delta Y + (k_2 + \beta_2)C_2 - \beta_1 X_p Y$$

(3.12)

In the following plots, I analyze the sensitivity of the insulator modeled with two-step phosphorylation - dephosphorylation reactions to changes in parameters, and I determine whether the sensitivity changes if X and Y are assumed to also decay. Plots for all of the simulations conducted with this model are presented in Section F.1. Here, I highlight only the more interesting findings.

I found that, in general, whether X and Y decay does not significantly affect the performance of the insulator (Figures 3-12 and 3-13) when the decay constant is $\delta = 0.01$, in contrast to what I found in Section 3.1.2 with an uninsulated buffer.

(a) Without X or Y decay.

(b) With X or Y decay.
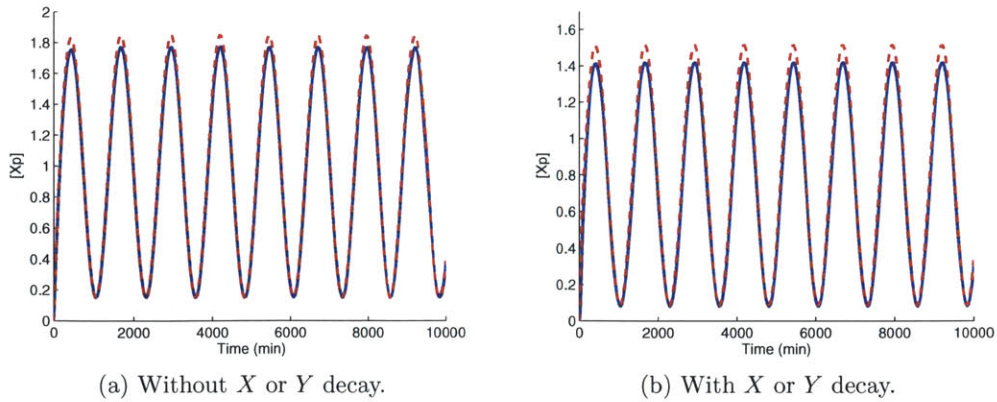
Figure 3-12: Insulator modeled with two-step phosphorylation - dephosphorylation reactions. Insulator attenuates retroactivity to the output well. Unless otherwise noted, parameters used in the two-step reaction models were: $k(t) = 0.01(1+sin(\omega t))$, $\omega = 0.005$, $\delta = 0.01$, $k_1 = 50$, $k_2 = 50$, $k_{on} = 10$, $k_{off} = 10$, $\alpha_1 = 0.01$, $\alpha_2 = 10$, $\beta_1 = 0.01$, $\beta_2 = 10$, $Y_T = 1500$, and $X_T = 1500$, obtained from [22].

However, when the decay constant is higher, the insulator does not perform as well. Thus, the assumption that the total amounts of X and Y remain constant does not hold when the degradation rate is high. In addition, I found that the insulator poorly attenuates retroactivity when the phosphorylation and dephosphorylation rates, k_1 and k_2, are small (e.g. equal to or less than 0.1/min) (Figure 3-14). In such a case, phosphorylation and dephosphorylation are no longer in a separate timescale from protein decay (0.01/min). Lastly, I found that if X and Y are assumed to decay, then as either the decay rates of X and Y increase or the production rates of X and Y decrease, insulator performance decreases because of lower levels of X and Y (Figures 3-15 and 3-16).

3.3 Discussion

I computationally simulated models for retroactivity and a phosphorylation - dephosphorylation cycle based insulator previously presented in [22] to evaluate the various assumptions made. In particular, I found that in a simple buffer system, whether a protein-DNA complex decays has a significant effect on model behavior. On the other hand, I found that in the insulator design, whether signaling proteins

(a) Without X or Y decay.

(b) With X or Y decay.
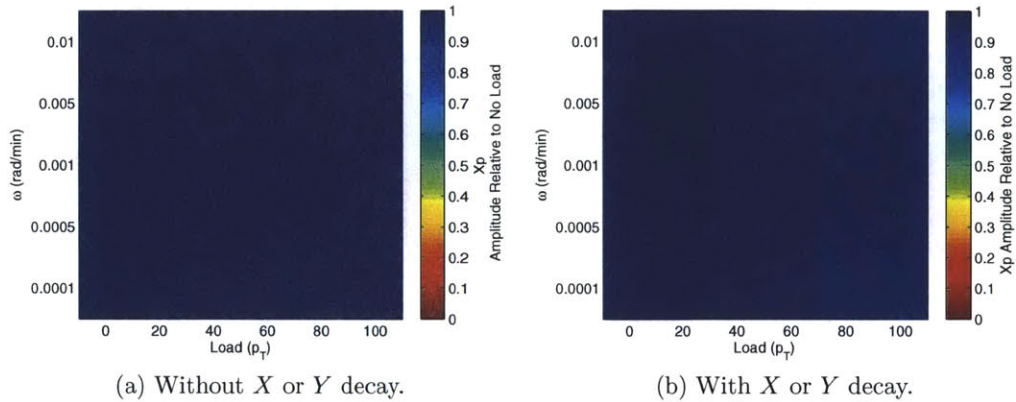
Figure 3-13: Insulator modeled with two-step phosphorylation - dephosphorylation steps with load and ω varied (a) without and (b) with Y and X decay. With either assumption, the insulator is able to attenuate retroactivity to the output. However, the performance of the insulator is slightly worse when X and Y are assumed to decay.

(a) Without X or Y decay.

(b) With X or Y decay.
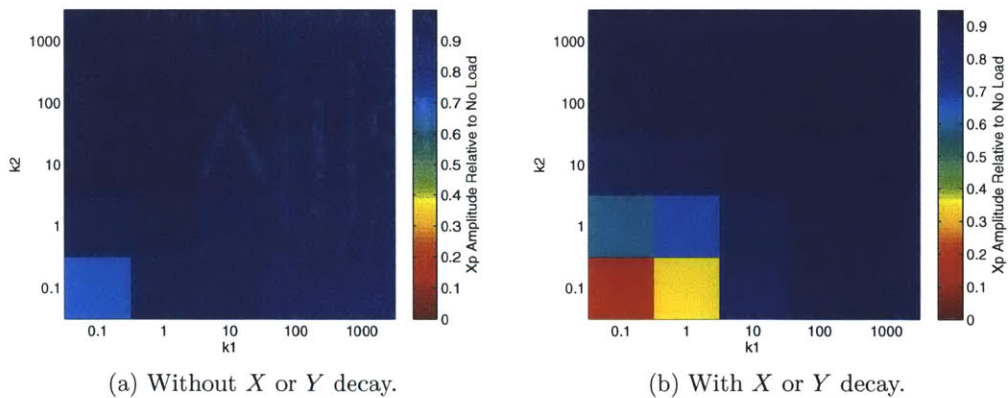
Figure 3-14: Insulator modeled with two-step phosphorylation - dephosphorylation steps with k_1 and k_2 varied (a) without and (b) with Y and X decay. With either assumption, performance of the insulator is poor at low k_1 and k_2 values. Low k_1 and k_2 values put phosphorylation and dephosphorylation in the same timescale as protein decay, breaking the insulator.

(a) Without X or Y decay.

(b) With X or Y decay.
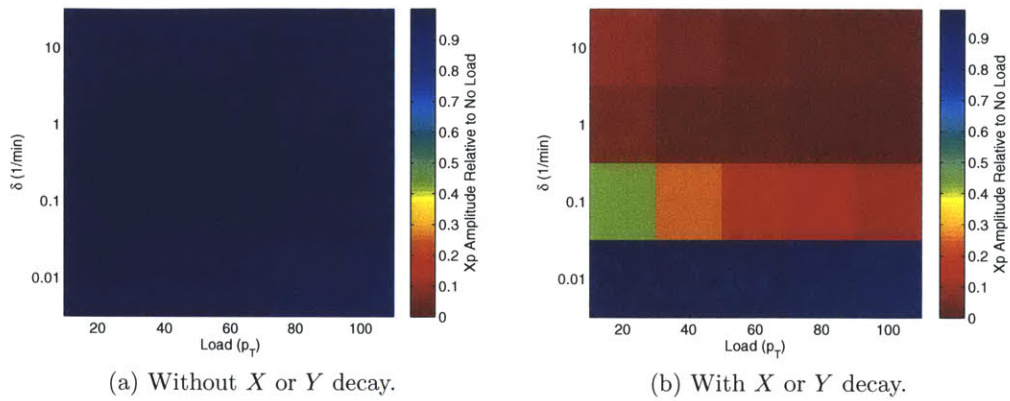
Figure 3-15: Insulator modeled with two-step phosphorylation - dephosphorylation steps with load and δ varied (a) without and (b) with Y and X decay. (a) The degradation rate (which affects only Z) does not affect the insulator performance noticeably. (b) The degradation rate has a significant impact on insulator performance, since increasing the degradation rate while keeping the kX and kY, the protein production rates, constant results in decreased levels of X and Y.

(a)

(b)
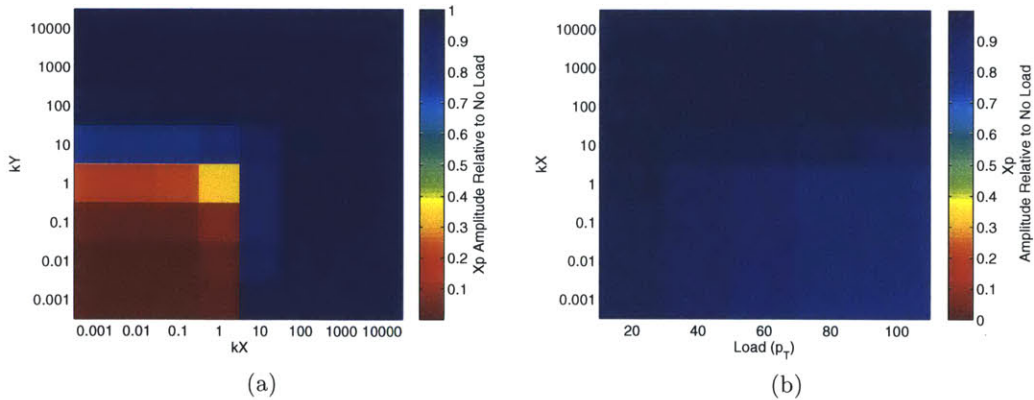
Figure 3-16: Insulator modeled with two-step phosphorylation - dephosphorylation steps with kX and kY varied. (a) Performance of the insulator suffers when kX and kY, production rates of X and Y, respectively, decrease. (b) At lower levels of X, the insulator does not attenuate retroactivity to the output as well.
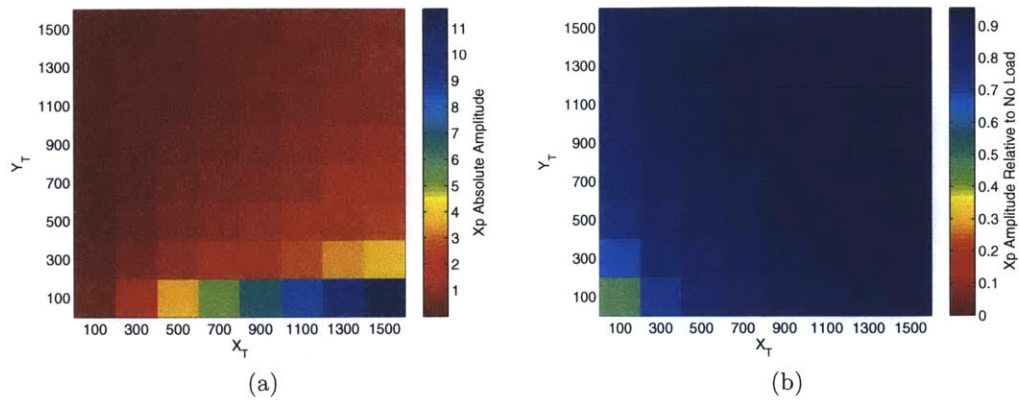
Figure 3-17: Insulator modeled with two-step phosphorylation - dephosphorylation steps with with X_T and Y_T varied. X_T and Y_T are a fixed constant (e.g. X and Y are assumed not to decay). Low values of X_T and Y_T hinder insulator performance, although the absolute amplitude of X_p is much more sensitive to X_T with this particular parameter set.

and protein-DNA complexes decay has a limited effect on model behavior at the degradation rate tested. Investigation of the assumptions made supports the model of the phosphorylation - dephosphorylation cycle based insulator as a buffer against retroactivity.

These simulations revealed that when the amount of binding sites is equal to or greater than the amount of transcription factor, retroactivity can be significant. Although it is unlikely for circuits that are currently being constructed to use hundreds of binding sites for a particular transcription factor, the amount of transcription factor in a cell can certainly be in the range of tens. The concentration of LacI in wild type E. coli is estimated to be about 10 [44]. (The size of E. coli is roughly $2 \ 10^{-12}$ cm^3, which means that the concentration of one molecule in an E. coli cell is about $1 \ 10^{-9}$ M [92]. The concentration of LacI was found to be in the nanomolar range [91].) In addition, some promoters have multiple binding sites. For example, the $pCYC1_{min}/tetO7$ promoter described in Chapter 4 has seven $tetO$ binding sites on each promoter. Thus, currently constructed circuits could have signal outputs and loads in the range where retroactivity noticeably affects device behavior.

Chapter 4

Genetic Circuit Designs

In this chapter, synthetic circuits to measure and attenuate retroactivity experimentally were designed and computationally modeled. In addition, preliminary results of circuits constructed with the assembly process described in Chapter 2 are presented.

In Section 4.1, I present a genetic construct to measure retroactivity in a buffer device, and in Section 4.2, I computationally model the buffer with parameters found in literature. In Section 4.3, I review background literature on the HOG pathway in yeast and the Jak-Stat pathway in mammalian cells. These two pathways are used in the proposed synthetic insulator design, which I present in Section 4.4. In Section 4.5, I computationally simulate the performance of the insulator. Lastly, in Section 4.6, I present flow cytometry data demonstrating the functionality of some components used in the genetic constructs.

4.1   Buffer Design to Measure Retroactivity

To experimentally demonstrate that high loads affect the module performance *in vivo*, the yeast circuit depicted in Figure 4-1 is able to measure retroactivity to the output. This system contains the following features: (1) an external inducer, (2) a transcription factor ("signal protein"), (3) a measurable fluorescent output, and (4) a adjustable load with factor binding sites.

The construction utilizes two transcriptional activators/repressors: tTA_2 and $LacI_4$. tTA_2 is a dimeric protein that is the fusion of the tetR DNA binding protein with
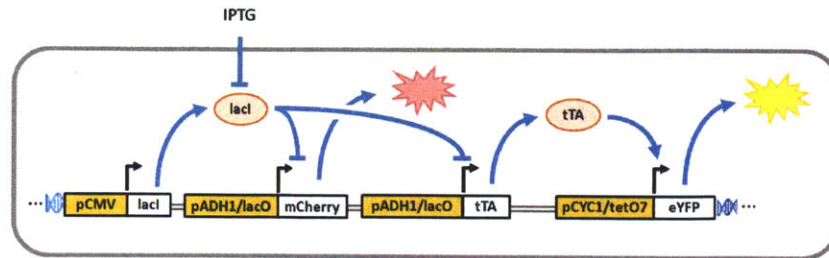
[IPTG]	Red?	Yellow?
0 mM	No	No
5 mM	Yes	Yes

(a) Buffer modulated with ITPG
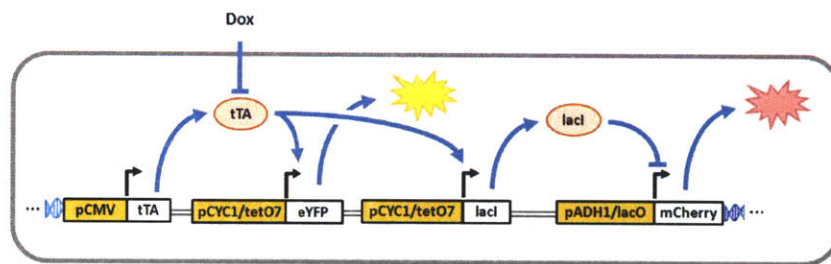
[Dox]	Red?	Yellow?
0 ug/mL	No	Yes
5 ug/mL	Yes	No

(b) Buffer modulated with Dox

Figure 4-1: Buffer constructs to investigate retroactivity. These constructs will be integrated into the chromosome, and a 2-μm plasmid, which exist at high copies in cells, with factor binding sites will be introduced to add load to the system (not depicted in figure). Expected phenotypes under different induction conditions are listed below each construct. (a) In one construction, IPTG is used to oscillate the system and tTA$_2$ is the signaling molecule that will be sequestered with additional binding sites. (b) In the other construction, Dox is used to oscillate the system and LacI$_4$ is the signaling molecule.

the transcriptional activator VP16 [39]. The tetR domain binds specifically to a *tetO* operator site, after which the VP16 domain recruits RNA polymerase to initiate transcription. tTA_2 cannot bind to the *tetO* operator in the presence of Dox, a tetracycline derivative [102]. Dox binds to allosteric sites in tTA_2 and disrupts its DNA binding functionality. $LacI_4$ is a tetrameric DNA binding protein that inhibits transcription [25]. $LacI_4$ binds to the *lacO* operator site and physically blocks the procession of RNA polymerase from the promoter towards the gene. *pCMV*, *pCYC1_{min}*, and *pAHD1* are yeast promoters. *pCMV* and *pCYC1_{min}* are very weak promoters, and *pADH1* is a moderately strong promoter [89, 39]. *pADH1/lacO* has a *lacO* binding site embedded in the promoter. In the absence of $LacI_4$, *pADH1/lacO* is a strong promoter. In its presence, *pADH1/lacO* should not be active. *pCYC1_{min}/tetO7* has seven *tetO* binding sites embedded in the promoter. In the absence of tTA_2, *pCYC1_{min}/tetO7* is a very weak promoter. In its presence, *pCYC1_{min}/tetO7* should be a strong promoter. mCherry and eYFP are yeast-enhanced fluorescent proteins that emit light in the wavelengths of red and yellow, respectively [105]. The buffer construct will be genomically integrated into the chromosome, such that only one copy of it exists in the cell at any time. In addition to this construct, a 2-μm plasmid with factor binding sites will be transformed (not depicted). 2-μm plasmids have copy numbers ranging from 60 to 80 [36].

In this system, the module consists of tTA_2 and $LacI_4$, which act as a buffer, converting a small molecule concentration into a transcription factor concentration. Examining the second buffer design in detail (Figure 4-1b), Dox is the input into the system that affects the binding of tTA_2 to the *tetO* sites, which in turns affects the production of $LacI_4$. $LacI_4$ is the output signal carrier of the module. In this system, the absence of $LacI_4$ allows for eYFP production, and the presence of $LacI_4$ blocks eYFP production. Thus, eYFP provides a visual output inversely correlated to the concentration of $LacI_4$. To summarize, in the absence of Dox, tTA_2 binds to the *tetO* sites, activating transcription of $LacI_4$, which inhibits eYFP production. In the presence of Dox, tTA_2 can no longer bind to the *tetO* sites, and thus $LacI_4$ is not produced, so eYFP is produced.
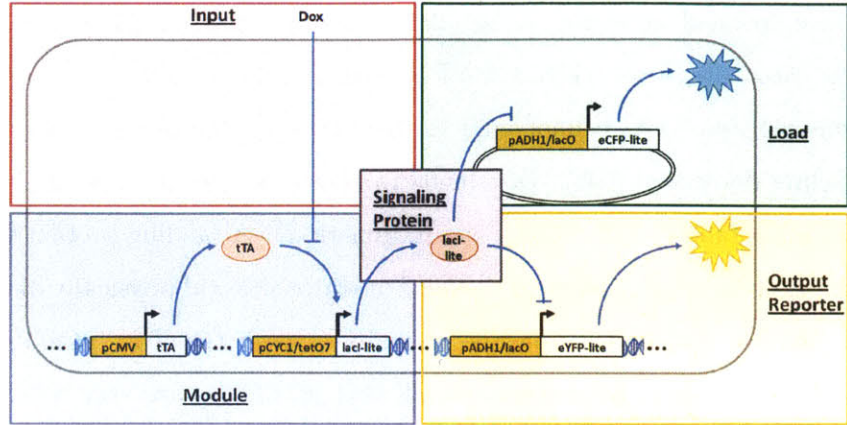
Figure 4-2: Modeled buffer design to investigate retroactivity.

The 2-μm plasmid acts as the variable load on the system. In the absence of the 2-μm plasmid, the $lacO$ site associated with mCherry is the only operator in the system that interacts with LacI$_4$. In the presence of the 2-μm plasmid, 60 to 80 additional $lacO$ sites are introduced into the system. The presence of these $lacO$ sites affects the availability of LacI$_4$ and the functionality of the module. The number of $lacO$ sites in the system can be varied by placing multiple copies of the $lacO$ site into the plasmid. For example, placing two copies of the $lacO$ site in 2-μm would introduce 120 to 160 additional $lacO$ sites into the system.

4.2   Analysis of Buffer Design

A preliminary ordinary differential equation model of the system was constructed with MATLAB's SimBiology toolbox. The particular system modeled is depicted in Figure 4-2, which is very similar to the system depicted in Figure 4-1b. Tables C.1 and C.2 lists the equations and parameters, respectively, used to simulate the system. Note that the equations assume that tTA$_2$ cannot bind to both Dox and DNA simultaneously. $ptTA$ is a constitutive promoter driving tTA. $LacOA$ is the total number of $lacO$ sites on all 2-μm plasmids in the model. $LacOB$ is the total number of $lacO$ sites integrated into the chromosome in the model. $LacOA$ drives CFP, and $LacOB$ drives eYFP. The load that the buffer drives is varied by changing the number of $LacOA$ sites in the model.

64
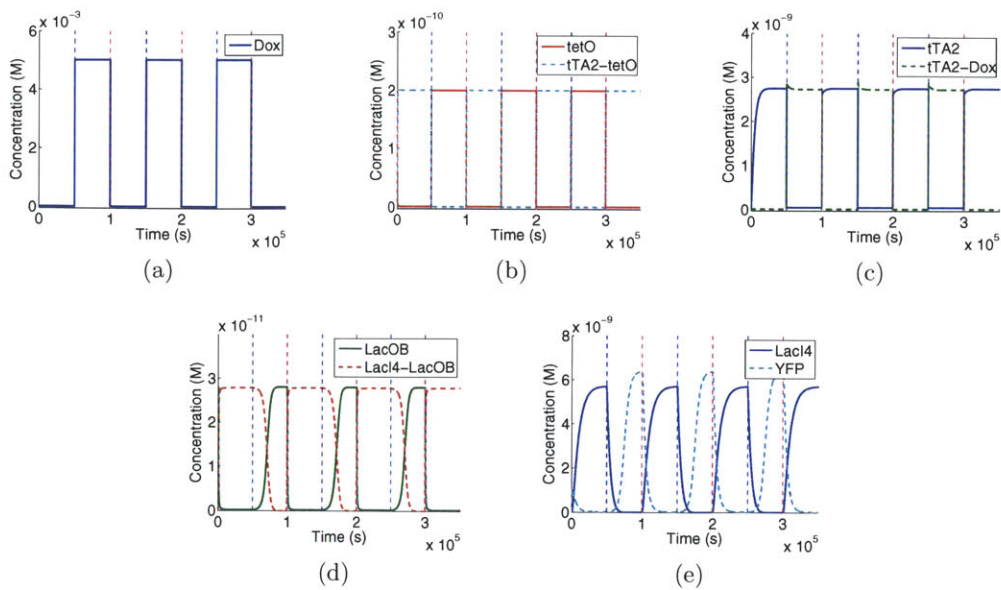
Figure 4-3: Simulation of buffer circuit. Dashed vertical blue lines indicate Dox induction, magenta lines indicate washing. Dox levels are oscillated as a square wave function. Concentrations of various species are shown as a function of time in the absence of a 2-μm plasmid (LacOA = 0).

The simulation results are presented in Figure 4-3. Dox inductions of 5×10^{-3} M and washes were incorporated into the SimBiology model as events, occurring every other 10^5 seconds. The model was run for a total of 3.5×10^5 seconds. Each cell was estimated to have an approximate volume of 6.0×10^{-14} L, which means that the concentration of one molecule in the cell is 2.8×10^{-11} M [113, 109, 60]. *ptTA* and *LacOB* concentrations were initialized at 2.8×10^{-11} M, and *tetO* concentration was initialized at 2.0×10^{-10} M because there are seven *tetO* sites in $pCYC1_{min}/tetO7$. All other species concentrations were initialized at 0 M. The deterministic ODE solver SUNDIALS was used [52].

The effect of increasing load (number of additional *lacO* sites, *LacOA*) on the system was investigated in the simulation. Figure 4-4 plots the resulting concentrations of LacI$_4$ and eYFP as the load was varied. These results suggest that retroactivity does occur in the buffer device modeled. Increasing the load sped up the increase in eYFP concentration after Dox induction (Figure 4-4). By increasing the load, more of the LacI$_4$ was sequestered from interacting with the *LacOB* operator site, which

65
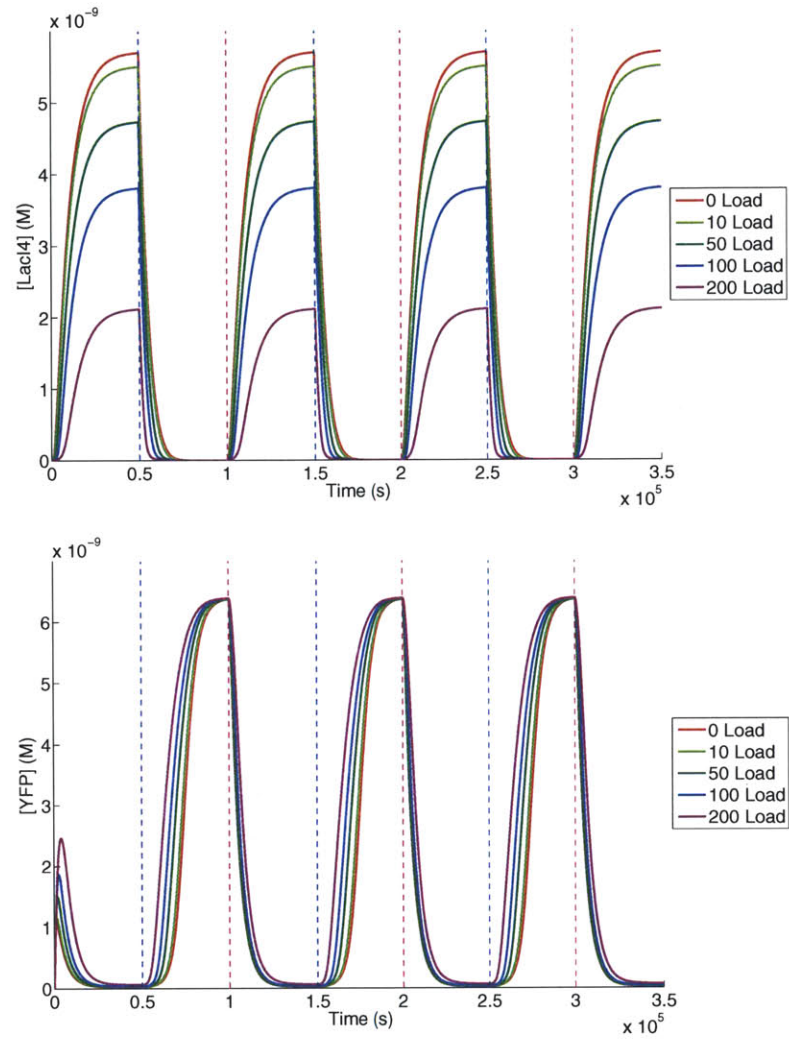
Figure 4-4: Simulation of buffer circuit with varied load. As seen, LacI$_4$ and eYFP concentrations vary over time at different loads, making it possible to see the effect of retroactivity on circuit dynamics.

drove the output reporter (eYFP). As the number of unbounded LacI$_4$ decreased, repression of eYFP production was reduced, causing an increase in transient eYFP concentration.

4.3 Signaling Pathways Used in Insulator

The proposed yeast insulator utilizes two pathways: the yeast high osmolarity glycerol (HOG) pathway and the murine JAK-STAT Pathway. The biology of these two pathways are reviewed in this section. Particular attention is paid to protein-protein interactions and methods of eliminating such interactions, to kinase rates and methods of increasing or decreasing kinase activity, and to localization of proteins and methods of changing their cytoplasm or nucleus localization. These details can be useful for engineering the pathways, as discussed in Section 4.4.

4.3.1 High Osmolarity Glycerol Pathway

SLN1-YPD1-SSK1/SKN7 is a phosphorelay signaling pathway that facilitates cell survival under osmotic stress (Figure 4-5). Hyperosmolarity reduces turgor in the cell and leads to hyperosmotic stress, and hypoosmolarity stresses the cell wall and leads to wall stress. Sln1p is a cell-surface osmosensor histidine kinase, and Ypd1p is a histidine phosphotransfer protein that shuttles across the nuclear membrane. The system then splits into two branches with Ssk1p and Skn7p, two nuclear aspartate response regulators. One branch (*SLN1-YPD1-SSK1*) regulates the hyperosmotic response, which triggers the HOG (high osmolarity glycerol) pathway. The other branch (*SLN1-YPD1-SKN7*) regulates the hypoosmotic response, which triggers genes involved in cell wall synthesis and metabolism.

In low osmolarity environments, Sln1p autophosphorylates and transfers its phosphate group to Ydp1p and in turn to Ssk1p, which inhibits Hog1p activity, and to Skn7p, which activates cell wall genes. In high osmolarity environments, Sln1p kinase activity is repressed, which leads to an accumulation of unphosphorylated Ssk1p and Skn7p. This activates the HOG pathway, which promotes synthesis of glycerol and transcription of other osmotic response genes [99, 73]. Thus, the Hog1 pathway
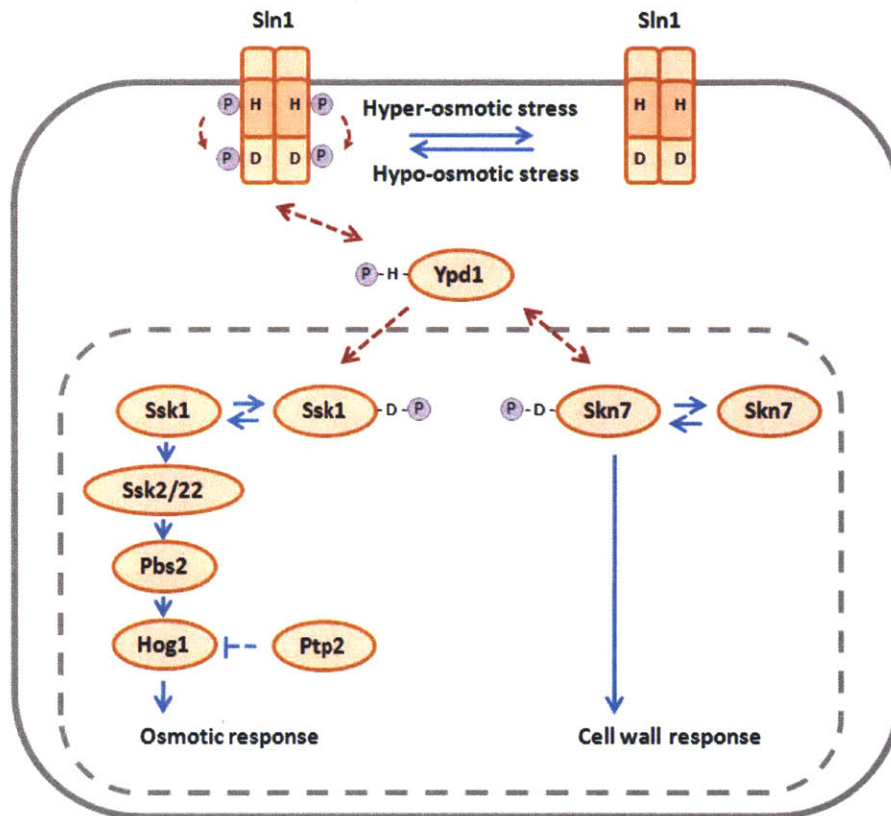
Figure 4-5: *SLN1-YPD1-SSK1/SKN7* phosphorylation pathway.

is active only in high osmolarity environments. Deletion of Sln1p or Ypd1p causes lethality because constitutive HOG pathway activation, under normal growth conditions, kills the cell. This lethality can be suppressed with a mutation in one of the downstream HOG components, which include Ssk1p, Ssk2p, Pbs2p, and Hog1p [75]. Mutants can also be rescued with tyrosine phosphatase Ptp2p overexpression [74]. Ptp2p suppresses Hog1 MAPK activity by dephosphorylating Hog1p.

SLN1

The osmoregulation pathway is a histidine-to-aspartate multi-component phosphorelay system. It is similar to two-component signal transduction systems found in bacteria, composed of a membrane-bound kinase sensor, which has an extracellular sensing domain and a cytoplasmic histidine autokinase, and a response regulator, which has a receiver domain with a conserved aspirate residue and a DNA binding domain [98]. In multi-component systems, a hybrid sensor exists along with histidine-containing phosphotransfer proteins and aspartate-containing response regulators.

Sln1p is composed of four regions [34, 110, 99]. At the N-terminus, Sln1p has an extracellular domain (ECD) between two transmembrane domains (TM1 and TM2). The rest of the protein contains a cytoplasmic linker region with two components (residues 537-569), a histidine kinase domain (HK, residues 570-947), and an aspartate response regulator domain (RR, residues 1059-1220). Sln1p is a hybrid kinase, since it contains both a kinase domain and a receiver domain. Sln1p exists as dimers in both activate and inactive states. Hypoosmotic stresses cause a conformational change in the dimer, which activates Sln1p [110]. The first event in response to hypoosmolarity is autophosphorylation of histidine 576 in the Sln1p kinase domain. This auto-phosphorylation has been observed to occur *in trans* [74]. The phosphate group is then irreversibly transferred to aspartate 1144 in the Sln1p receiver domain [110, 99, 53]. *sln1-H576Q* has non-functional kinase domain, and *sln1-D114N* has a nonfunctional receiver domain [99]. *sln1-ΔC*, which contains only the sensor domain of Sln1p, and *sln1-ΔN*, which contains only the receiver domain, by themselves cannot phosphorylate Ydp1p, but when expressed together can [99].

Sln1p requires dimerization and the coiled-coil linker for activation. The ECD facilitates dimerization of membrane bound Sln1p sensors [96]. Deletion of first transmembrane domain (TM1) results in a constitutively active kinase, whereas deletion of both TM1 and ECD eliminates kinase activity. Deletion of both TMs and ECD, such that the Sln1p mutant was cytosolic, also eliminates kinase activity. However, attachment of a leucine zipper in place of TM1 and ECD partially restores kinase activity, which suggests that dimerization of Sln1p is necessary for its activation. The coiled-coil linker is also required for kinase activity [110]. Two-hybrid assays were used to determine that the coiled-coil linker is necessary for dimerization [110]. A Sln1p fragment (residues 356-1077) dimerized, whereas the same fragment but lacking the coiled-coil linker (residues 356-1077; Δ 527-568) did not. Lack of the coiled-coil linker eliminated kinase activity, as expected.

Activated alleles of *SLN1*, *sln1**, have been constructed [33]. *sln1-T550I* has a mutation in the coiled-coil linker region, which changes its equilibrium between phosphorylated and unphosphorylated forms towards the phosphorylated form under non-stressed conditions [33]. Because this mutation is recessive and because Sln1p acts as a dimer, it was suggested that this mutation causes decrease in Sln1p-associated *in trans* phosphatase activity. A similar mutation, E543V, was also found to create an activated form of Sln1p [110]. Other recessive mutations in the receiver domain (P1148S and P1196L) have also been found to create activated forms of Sln1p. These mutations occur in exposed surfaces on the protein, suggesting that they may decrease phosphatase sensitivity.

YPD1

Ypd1p is a histidine-phosphorylated phosphotransfer protein that shuttles between the cytoplasmic side of the plasma membrane and the nucleus. The phosphorylation site in Ypd1p is histidine 64, and the mutant *ypd1-H64Q* is non-functional in phosphorelay [99]. Phosphotransfer rates between Ydp1p and response regulator domains of Sln1p and Skn7p showed that phosphotransfer reactions between Ydp1p and Sln1p or Skn7p are reversible [59]. A K67A mutation results in increased phospho-

transfer rate from Sln1p to Ydp1p but decreased binding affinity, and R90A, Q86A, or G68Q mutations result in decreased phosphotransfer rates.

SKN7

Ssk1p and Skn7p are two response regulator proteins, regulated by phosphorylation of an aspartate residue. They are the only proteins, other than Sln1p, known in yeast to have a receiver domain [53]. Skn7p is constitutively localized in the nucleus and has a role in coordinating cellular adaptation to global stresses, which include cell wall stress, oxidative stress, heat stress, and cell cycle control [100, 120]. In response to cell wall stress, Skn7p activates transcription of *OCH1*, a cell wall synthesis gene [72]. Overexpression of Skn7p also suppresses a *kre9*-Δ cell wall defect [14]. In response to oxidative and heat stress, Skn7p induces transcription of oxidative stress response and heat shock proteins [70, 100]. *skn7* mutants are sensitive to hydrogen peroxide and heat stress. Skn7p also has a role in regulating cell cycle, as overexpression of Skn7p suppresses a *swi6*-Δ cell cycle gene expression defect [86]. Lastly, Skn7p also regulates calcineurin signaling and interacts with calcineurin and the Crz1p transcription factor [120]. As part of its role in regulating stress adaptation, Skn7p has been discovered to have interactions with several transcription factors, which include Hsf1p (heat-shock), Crz1p (calcineurin-dependent), Mbp1p (cell-cycle), Rho1p (cell-wall), and Yap1p (oxidative stress) [53, 3]. Strong overexpression of Skn7p disturbs cell wall assembly and results in swollen unbudded cells, which is lethal to the cell [12].

Skn7p is composed of an N-terminal DNA binding domain (DBD) (residues 87-150), a coiled-coil (HR) domain (residues 222-303), a receiver domain (residues 378-497), and a C-terminal transcription activation domain [51, 87]. Its N-terminal DBD is highly homologous to that of the heat shock factor [15]. Mutations S137A and R140A, which are in the DBD, eliminate the ability of Skn7p to bind to the *OCH1* promoter [15, 72]. Ypd1p phosphorylates Skn7p at Asp427 in the response receiver domain [71]. Mutation of Asp427 to glutamate increases Skn7p activity, suggesting that it is constitutively activated in a phosphorylation-independent manner [71]. In

contrast, mutation of Asp427 to alanine and arginine decrease Skn7p activity [67]. Its HR domain is required for homodimerization and heterodimerization with other transcription factors, as determined through co-purification [13, 120, 100, 90]. Deletion of the coiled-coil domain disrupts Skn7 function [3].

Skn7p aspartyl phosphorylation is required for cellular responses to cell wall stress triggered by *SLN1-YPD1-SKN7* signal transduction, and for suppression of *kre9-Δ* cell wall defect and *swi6-Δ* cell-cycle defect. However, other Skn7p functions, in particular its role in oxidative stress response, are independent of its aspartyl phosphorylation. This was discovered through the use of *skn7-D427N*, which eliminates phospho-activation of Skn7p at residue 427. Expression of *skn7-D427N* can complement *skn7-Δ* strains to hydrogen peroxide sensitivity [87].

Skn7p appears to interact with Yap1p in its response to oxidative stress. Skn7p binds to the promoters of CCP1 (a peroxidase), NCA3 (nuclear control of ATPase), GPX2 (a peroxidase), OCH1 (a mannosyl-transferase), TRR1 (thioredoxin reductase), TRX2 (a thioredoxin), and TSA1 (a thio-specific antioxidant) [53, 94, 50, 70, 90]. TRX2, CCP1, TSA1, TRR1, and GPX2 are involved in the oxidative stress response. In some of these promoters, Yap1p binding sites are also present, and oxidative stress induces both Skn7p and Yap1p bind to these promoters. In contrast, OCH1 and NCA3 are involved in cell wall synthesis and contain Skn7p binding sites only.

It may be possible to separately affect the roles of Skn7p in cell wall stress and oxidative stress responses. It has been suggested that oxidative stress may induce Skn7p phosphorylation at serine/threonine sites other than D427, and this phosphorylation requires the presence of Yap1p [71, 51]. A T437A mutation in Skn7p was found to result in oxidative-stress sensitivity [51]. The T437A mutation in Skn7p disrupts Skn7p-Yap1p complex formation while still being responsive to osmotic-dependent Skn7p phosphorylation, suggesting that this mutation specifically affects oxidative stress response.

Skn7p DNA binding sites in theses promoters have been found to differ, with some resembling heat shock promoter elements, but most share a GGC repeat motif [87, 50].

Binding of Skn7p to the *OCH1* promoter site does not require phosphorylation at D426, but its activation of the *OCH1* promoter does [72]. Thus unphosphorylated Skn7p can still bind the *OCH1* promoter. Two binding sites in *OCH1*, named SSRE for *sln** response element, have been identified [72]. Called repeat A (ATTTGGC-CGGCCC) and repeat B (ATTTGGCTGGCC), they are separated by 11 bps and ~275 upstream of the start site of *OCH1*. Both have the same responsiveness to Skn7p, and effect of having both repeats in the promoter is additive. As expected, SSRE is responsive to osmotic but not oxidative stress.

4.3.2 JAK-STAT Pathway

Janus kinase (JAK) and signal transducer and activator of transcription (STAT) are a family of signaling protein found in higher order eukaryotes that activate particular genes upon binding of extracellular signals to transmembrane receptors [1]. The typical sequence of events in the JAK-STAT pathway is as follows: binding of extracellular signal to transmembrane cytokine receptors, receptor dimerization, JAK recruitment to dimerized receptors, JAK auto- and trans-phosphorylation of tyrosine residues on the receptors, and STAT recruitment and phosphorylation by JAKs. Once phosphorylated at tyrosine sites, STATs translocate into the nucleus to activate target gene promoters. There are four known mammalian JAKs (Jak1, Jak2, Jak3, and Tyk2), and seven known mammalian STATs (Stat1, Stat2, Stat3, Stat4, Stat5a, Stat5b, and Stat6).

JAKs

JAKs have seven homology domains (JH1-7) [103]. JH1 contains the kinase domain. JH2 has sequence similarity to kinase domains but lacks several residues necessary for catalytic activity. Thus JH2 is called a pseudokinase domain. JH3-7 contain cytokine receptor binding sites.

Dimerization of JAKs is sufficient for JAK activation [85]. Other tyrosine kinases are not necessary for JAK activation. Supporting that JAK dimerization causes its activation, co-expression of Jak2 kinase-inactive mutants (Jak2-K882E or Jak2-ΔJH1) with Jak2 suppresses Jak2 activation, probably through competition of the

Jak2 mutant in dimerization with Jak2 [103]. In mammalian cells, overexpression of JAKs results in ligand-independent interaction and autophosphorylation [27]. The same result has been observed in yeast [6]. Overexpression of Jak2 creates a pool of cytoplasmic Jak2 that can autophosphorylate in the absence of receptor interaction.

Jak2 is a 1129 residue protein that cannot distinguish been Stat5a or Stat5b [6]. In Jak2, JH2 (aa 548-804) has a negative regulatory function that suppresses Jak2 phosphorylation of Stat5 through *in trans* interactions with JH1 [103]. JH2 maintains Jak2 in an inactive state in the absence of cytokine receptor interactions. Deregulation of Jak2 through deletion of JH2 results in ligand-independent, constitutive Jak2 activation of Stat5 [6, 103].

The activities of truncated Jak2 kinases have been studied [103]. JH1-2, which contains just the JH1 and JH2 domains of Jak2, phosphorylates Stat5 at levels comparable to Jak2. Jak2-ΔJH2 and the JH1 domain alone activate Stat5 at higher levels than Jak2. Jak2-E665K, which has a mutation in the JH2 domain, activates Stat5 at higher levels than Jak2, but not as much as Jak2-ΔJH2 or JH1, suggesting that the mutation hinders, but does not eliminate, JH2 function. Fusions of Jak2 with other proteins have also been created. A Jak2-Stat5 fusion was found to possess kinase activity and was able to phosphorylate itself and as well as cytosolic Stat5 [10].

Negative regulators of the JAK-STAT pathway have been discovered. Tyrosine phosphatases (PTP), such as *Shp-1* and *Shp-2*, and SOCS proteins have been found to suppress signaling through interactions with phosphorylated receptors and Jaks [64, 122, 124].

STATs

STATs are phosphorylated by JAKs and, once activated, translocate into the nucleus to activate transcription of target genes. STATs contain an N-terminal domain (ND), a four-helix coiled-coil domain (CCD), an eight-stranded β-barrel DNA-binding domain (DBD), an α-helical linker domain, a SH2-domain, and a C-terminal transcriptional activation domain (TAD) [119]. Except for the TAD, STATs have a high homology in the other domains. The critical tyrosine in Stat5a is Tyr694 and in

Stat5b is Tyr699 [6]. Phosphotyrosine SH2 domains on both Stat5 monomers are required for Stat5 dimerization to occur. Mutation of Tyr699 in Stat5b to phenylalanine eliminates phosphorylation of Stat5b. Deletion of the TAD domain in Stat5 results in hyperphosphorylation due to decreased tyrosine dephosphorylation [88, 6]. SH2 domains bind to phosphotyrosines on phosphorylated cytokine receptors and to other phosphorylated STATs. The SH2 domain in Stat5 interacts with JH2 in Jak2 and is necessary for Stat5 phosphorylation by Jak2. R618K mutation in Stat5a SH2 domain eliminates phosphorylation of Stat5a by Jak2 [6]. Jak2-ΔJH2 and JH1 both activate Stat5a-R618L, but Jak2 and JH1-2 do not activate Stat5-R618L [10].

Inactive STATs were originally thought to be monomers. However, evidence has shown that inactive STATs exist in the cytoplasm as dimers. Once activated, the dimers translocate into the nucleus and bind to associated enhancer elements. Stat1 exists in mutually exclusive dimer configurations, depending on their phosphorylation state [83, 119]. The majority of phosphorylated Stat1 dimerize in a parallel configuration with reciprocal interactions between their phosphotyrosines and SH2 domains. The majority of unphosphorylated Stat1 dimerize in an antiparallel configuration with reciprocal interactions between their coiled-coil domains and DNA-binding domains, independent of SH2 [119, 77]. In particular, there appears to be two contact regions between unphosphorylated Stat1 dimers, one in the N-terminal domain mediated by residues F77 and L78, and one where residue F172 in the CCD inserts into a pocket created by residues Q340, L83, G382, T385, H406, and Q408 in the DBD. The single mutation F172W and the double mutation F77A/L78A greatly diminish unphosphorylated Stat1 dimer presence [77]. Inactive Stat5A dimers have been found to exist in an antiparallel configuration as well [95]. One of the contact regions in Stat5A is mediated by interactions among residues N361, V362, H363, M364, N365, S449, and S452, a collection of residues similar to the one found in Stat1 [95].

The N-terminal domain is responsible for dimerization of inactive STATs, cooperative DNA-binding, and nuclear translocation [83, 81, 116]. The N-terminal domain also has a role in spatial reorientation of Stat1 dimers, in order to exposure phosphotyrosine to dephosphorylation [83]. Mutations of this domain result in reduced level
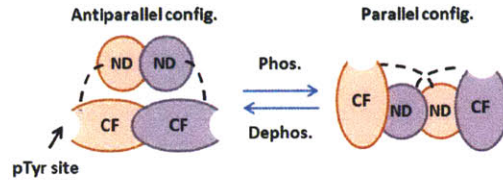
Figure 4-6: When Stat1 is unphosphorylated, it has greater tendency to exist in an antiparallel configuration. When Stat1 is phosphorylated, it has greater tendency to exist in a parallel configuration. Figure adapted from [77]

unphosphorylated Stat1 dimers and accumulation of phosphorylated Stat1 dimers, indicating resistance to dephosphorylation. Overexpression of the N-terminal domain with this mutant can facilitate dephosphorylation.

A functional nuclear localization signal (NLS) is present in only activated Stat1 dimers and not in inactive Stat1 dimers. Mutation of residue L407 to alanine in the DBD blocks recognition of the conformational NLS by the import shuttling receptor, which prevents Stat1-L407A nuclear translocation [80]. However, Stat3 and Stat5 do not need tyrosine phosphorylation or dimerization for nuclear translocation, and sequences in the coiled-coil domain (residues 152-163 in Stat3 and residues 145-150 in Stat5) have been found to be necessary for importin association and subsequent nuclear localization [57, 101].

4.4 Insulator Design to Attenuate Retroactivity

The *JAK-STAT* and *SLN1-YPD1-SKN7* pathways are used in the creation of an insulator in yeast. As a phosphorylation cycle is needed for the insulator, use of Skn7p activation was chosen over Ssk1p activation because phosphorylation of Skn7p by Ypd1p is reversible, whereas phosphorylation of Ssk1p by Ypd1p is irreversible. In order to read the output of Skn7p activation, synthetic Skn7p response element (SSRE) promoters were designed by Ming-Tang Chen in the Weiss Lab [18]. Double Skn7p binding sites were inserted into the *MEL1* promoter, replacing the promoter's native GAL4 UAS, to create SSRE [82]. In addition, the Swi4/Swi6 cell-cycle box that appears between the double Skn7p binding sites in the *OCH1* promoter was replaced with a 6-bp AgeI site to lower basal transcriptional activity [72]. A second synthetic
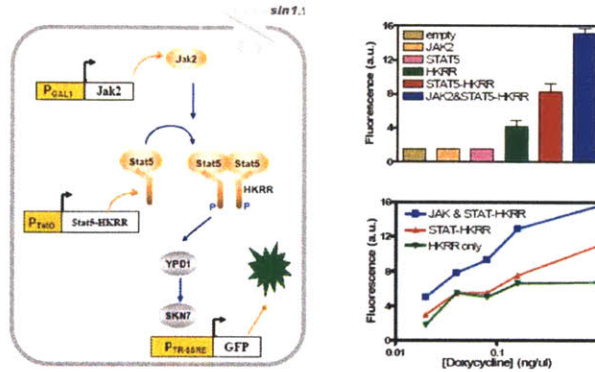
Figure 4-7: Synthetic *JAK2-STAT5B:HKRR-YPD1-SKN7* pathway results in GFP expression. Cells were induced with Dox to activate Stat5-HKRRp production, and galactose was added afterwards to induce Jak2p expression. The strain used has TR-SSRE:yEGFP3 integrated into the *SSK1* locus, *SLN1* knocked out, and harbored the plasmids p416G-mJak2 (containing pGAL1-Jak2) and p252-mStat5b-HKRR (containing pCMV-tTA and pCYC1$_{min}$/tetO7-mStat5b-HKRR). Data collected and figure produced by Ming-Tang Chen.

promoter, TR-SSRE, was created by inserting a tandem repeat of the double Skn7p binding sites to SSRE. These two promoters have similar sensitivities and gains, but TR-SSRE has a higher basal activity level (Figure 4-7).

A synthetic *JAK2-STAT5B:HKRR-YPD1-SKN7* signal transduction was previously created in yeast by Ming-Tang Chen. JAK and STAT are exogenous to yeast but have been previously expressed in yeast cells [6]. The use of an orthogonal system decreases the likelihood that endogenous yeast components will interfere with the synthetic signaling pathway. The cytoplasmic portion of Sln1p (HKRR, residues 530-1220) was fused to the C-terminal of Stat5b to create Stat5b-HKRRp with the linker LIK. Overexpression of Jak2p results in the phosphorylation and a dimer reconfiguration of Stat5-HKRRp, bringing the HKRR regions closer together. The HKRR regions autophosphorylate and activate the phosphorelay of Ypd1p and Skn7p, turning on the expression of genes driven by SSRE/TR-SSRE promoters (Figure 4-7). Note that although Stat5b-HKRRp can dimerize and autophosphorylate in the absence of Jak2p, addition of Jak2p increases the strength of the induced TR-SSRE promoter. The obtained dynamic range may suffice for insulator performance. If not, deletion of the JH2 domain in Jak2p and deletion of the C terminus of Stat5p have
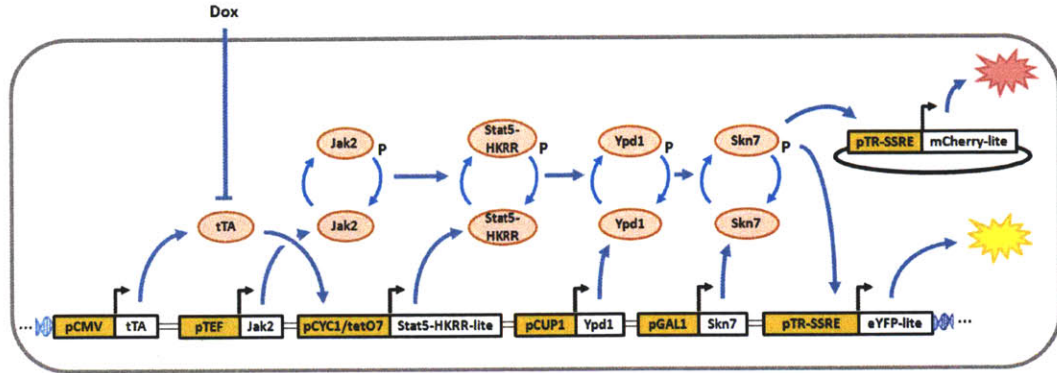
Figure 4-8: Design of an insulator device with the assembly method described in Chapter 2. Promoter-gene pairs at the bottom of the figure will be assembled into an integrating vector and subsequently integrated into the HO locus. In addition, variable load will be placed on the insulator by a plasmid carrying the TR-SSRE promoter. *SSK1*, *SLN1*, *YPD1*, and *SKN7* are knocked out in this strain.

both been observed to increase phosphorylation of Stat5p by Jak2p, which would increase the system dynamic range [6].

Although the *JAK2-STAT5B:HKRR-YPD1-SKN7* signaling pathway has already been constructed, it is not amenable to systematic studies because not all of the components are integrated at single copy and not all of the genes in the insulator are driven by tunable promoters. The proposed insulator design is depicted in Figure 4-8. This design notes that levels of Ypd1p and Skn7p may significantly affect the performance of the insulator. High levels of these proteins will provide good insulation between modules, where low levels will provide poor insulation. To affect the gains of the insulator, levels of Ypd1p and Skn7p will be modulated. Endogenous copies of *YPD1* and *SKN7* will be knocked out, and *YPD1* will be placed under the control of an inducible *CUP1* promoter, and *SKN7* under a inducible *GAL1* promoter. High levels of Cu^{2+} concentrations trigger the transcriptional activator Cup2p to bind the *CUP1* promoter [16]. Galactose induces the *GAL1* promoter, where as glucose represses the *GAL1* promoter [68]. Thus, levels of Cu^{2+} and galactose can be varied to investigate how the levels of Ypd1p and Skn7p affect the insulation capacity of the proposed device. In addition, load on the insulator can be modulated by a plasmid carrying binding sites for Skn7p. This design provides the construction flexibility

needed to study the effect of parameter variation on insulation capability, as promoters and genes can be easily switched to alternative mutant forms in the assembly process.

4.5 Analysis of Insulator Design

4.5.1 Chemical Equations

Here, the design presented in Figure 4-8 is analyzed. The input into the insulator is unphosphorylated Stat5b-HKRRp, and the output is phosphorylated Skn7p. In the following section, we abbreviate Stat5b-HKRRp as Z, Ypd1p as W, and Skn7p as X. Phosphorylated compounds are indicated with an asterisk. Only phosphorylated Skn7p can activate *SSRE/TR-SSRE*, but Skn7p does not need to be phosphorylated to bind to DNA. Skn7p is assumed to dimerize, and that both monomeric and dimeric Skn7p can bind to DNA. The production and degradation/dilution of Z is first modeled. Z is produced whenever the *pCYC1/tetO7* promoter is activated, modeled here as a variable zeroth-order rate $k(t)$. Z and Z^* undergo degradation/dilution with the first-order rate γ.

$$\emptyset \xrightarrow{k(t)} Z$$
$$Z \xrightarrow{\gamma} \emptyset$$
$$Z^* \xrightarrow{\gamma} \emptyset$$
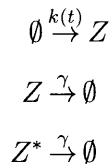
The phosphorylation of Z by Jak2 is modeled as a first-order rate k_2. Note that the amount of Jak2 present in the system has been lumped into k_2 for simplicity. In addition, Z^* can dephosphorylated by non-specific phosphatases at the first-order rate k_{dp}, where the amount of non-specific phosphatases in the system has been lumped into k_{dp}, again for simplicity.

$$Z \underset{k_{dp}}{\overset{k_2}{\rightleftharpoons}} Z^*$$
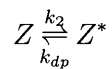
Non-specific dephosphorylation of activated W and X is modeled with rate k_{dp}. Complexes consisting of different species are assumed to not undergo dephosphorylation, because the phosphate group is presumably being transferred at an interior interface between the two species within the complex and, thus, out of reach of phosphatases. In addition, phosphorylated X bound to DNA is assumed not undergo dephosphorylation, because the phosphate group is presumably activating the promoter through interactions with other proteins. The subscript D represents a dimer.

$$W^* \xrightarrow{k_{dp}} W$$

$$X_D^{**} \xrightarrow{k_{dp}} X_D^*$$

$$X_D^* \xrightarrow{k_{dp}} X_D$$

$$X^* \xrightarrow{k_{dp}} X$$
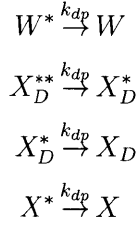
The reversible dimerization of inactive and activated X is modeled with the second-order rate k_f and first-order rate k_r.

$$X + X \underset{k_r}{\overset{k_f}{\rightleftharpoons}} X_D$$

$$X + X^* \underset{k_r}{\overset{k_f}{\rightleftharpoons}} X_D^*$$

$$X^* + X^* \underset{k_r}{\overset{k_f}{\rightleftharpoons}} X_D^{**}$$
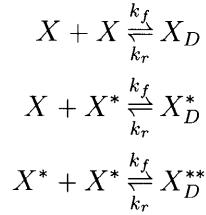
The reversible phosphotransfer between Z and W is modeled with second-order rates α_1 and α_4 and first-order rates α_2 and α_3. In addition, the reversible phosphotransfer between W and X is modeled with β_1 and β_4 and first-order rates β_2 and β_3.

$$Z^* + W \underset{\alpha_2}{\overset{\alpha_1}{\rightleftharpoons}} [W.Z^*] \underset{\alpha_4}{\overset{\alpha_3}{\rightleftharpoons}} Z + W^*$$

$$W^* + X \underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} [W^*.X] \underset{\beta_4}{\overset{\beta_3}{\rightleftharpoons}} W + X^*$$

$$W^* + X_D \underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} [W^*.X_D] \underset{\beta_4}{\overset{\beta_3}{\rightleftharpoons}} W + X_D^*$$

$$W^* + X_D^* \underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} [W^*.X_D^*] \underset{\beta_4}{\overset{\beta_3}{\rightleftharpoons}} W + X_D^{**}$$
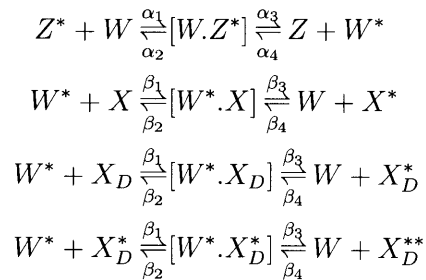
X bound to DNA is assumed to still undergo phosphorelay with W, with the same rates as X not bound to DNA.

$$W^* + [X.p] \underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} [W^*.X.p] \underset{\beta_4}{\overset{\beta_3}{\rightleftharpoons}} W + [X^*.p]$$

$$W^* + [X_D.p] \underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} [W^*.X_D.p] \underset{\beta_4}{\overset{\beta_3}{\rightleftharpoons}} W + [X_D^*.p]$$

$$W^* + [X_D^*.p] \underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} [W^*.X_D^*.p] \underset{\beta_4}{\overset{\beta_3}{\rightleftharpoons}} W + [X_D^{**}.p]$$
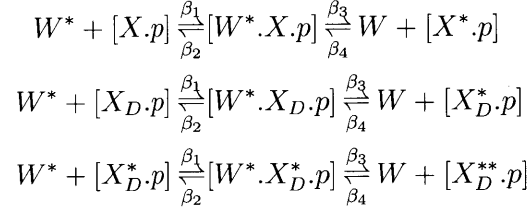
The binding of X to DNA is modeled with the second-order rate k_{on} and its disassociation with the first-order rate k_{off}.

$$X + p \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [X.p]$$

$$X^* + p \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [X^*.p]$$

$$X_D + p \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [X_D.p]$$

$$X_D^* + p \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [X_D^*.p]$$

$$X_D^{**} + p \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [X_D^{**}.p]$$
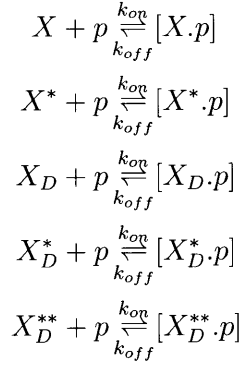
Lastly, a W X complex is assumed to interact with DNA with the same rates.

$$[W^*.X] + p \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [W^*.X.p]$$

$$[W^*.X_D] + p \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [W^*.X_D.p]$$

$$[W^*.X_D^*] + p \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [W^*.X_D^*.p]$$
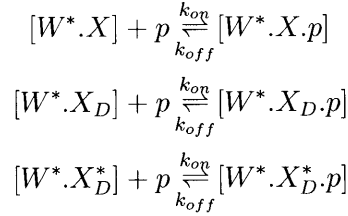
In addition, the total amount of inactive and activated W and X are assumed to stay constant and are termed W_T and X_T, respectively. Note that in this model, complex are assumed to not degrade. This assumption was made for simplicity of the model and on the basis that the assumption does not significantly affect model behavior at
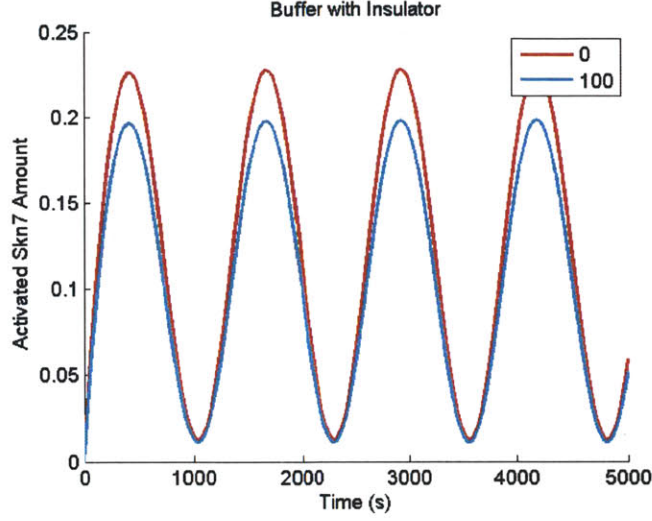
Figure 4-9: Simulation of synthetic insulator in yeast. The insulator is demonstrated to attenuate retroactivity to the output. Unless otherwise noted, the parameters used in the simulations were: $k(t) = 0.1(1 + sin(\omega t))$, $\omega = 0.005$, $\gamma = 0.01$, $k_1 = 1$, $k_2 = 0.1$, $k_{dp} = 0.01$, $a_1 = 0.01$, $a_2 = 20$, $a_3 = 48$, $a_4 = 4.8$, $b_1 = 1$, $b_2 = 5$, $b_3 = 48$, $b_4 = 4.8$, $k_f = 20$, $k_r = 20$, $k_{on} = 50$, $k_{off} = 50$, $X_T = 1500$, and $W_T = 1500$.

low decay rates, as discussed in Section 3.2.2.

$$W_T = W + W^* + [W.Z^*] + [W^*.X] + [W^*.X_D] + [W^*.X_D^*] +$$

$$[W^*.X.p] + [W^*.X_D.p] + [W^*.X_D^*.p] \tag{4.1}$$

$$X_T = X + [W^*.X] + 2[W^*.X_D] + 2[W^*.X_D^*] + [W^*.X.p] +$$

$$2[W^*.X_D.p] + 2[W^*.X_D^*.p] + X^* + 2X_D + 2X_D^* + 2X_D^{**} +$$

$$[X.p] + [X^*.p] + 2[X_D.p] + 2[X_D^*.p] + 2[X_D^{**}.p] \tag{4.2}$$

$$p_T = p + [W^*.X.p] + [W^*.X_D.p] + [W^*.X_D^*.p] + [X.p] + [X^*.p] +$$

$$[X_D.p] + [X_D^*.p] + [X_D^{**}.p] \tag{4.3}$$

The set of ordinary different equations that results from these chemical equations are presented in Section C.2.

These equations were modeled in the Simulink package of MATLAB with ode23s, and the insulator was demonstrated to buffer against retroactivity to the output at high loads *in silico* in Figure 4-9.

82
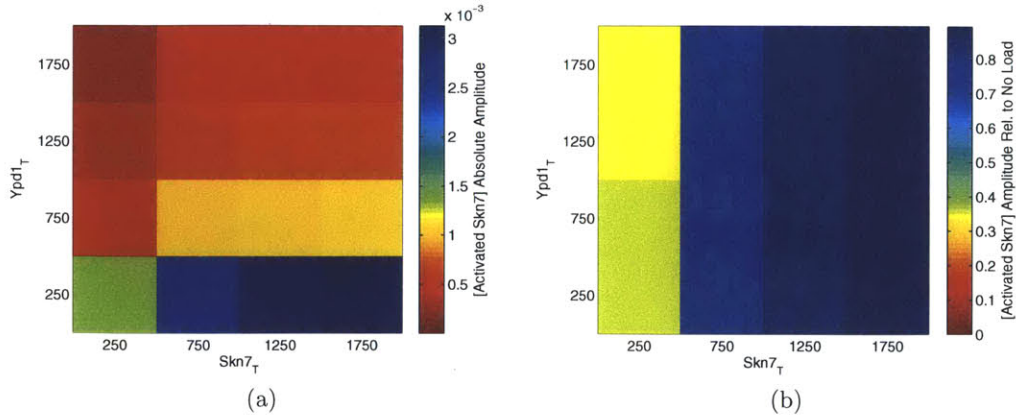
Figure 4-10: Synthetic insulator, X_T and W_T varied, absolute and relative amplitude to no load. Higher levels of Ypd1_T decrease the absolute amplitude of the output but have minimal effect on the relative amplitude of the output. When Ypd1p_T is low, greater Skn7p_T increases the absolute amplitude of the output. At all Ypd1p_T values, greater Skn7p_T increases the relative amplitude of the output. The absolute and relative amplitudes of each species (X^*, X_D^*, X_D^{**}, and $W^*X_D^*$) are presented in Figures F-16 and F-17.

4.5.2   Parameter Sensitivity

To determine the sensitivity of the insulator to various parameters, I chose pairs of parameters to vary simultaneously and measured the output amplitude. I assumed that phosphorylated Skn7 could activate the TR-SSRE promoter. Because it was unclear if Skn7 needed to dimerize to activate the promoter, I assumed that any complex that contained an phosphorylated Skn7 could activate the promoter. For simplicity, I assumed that each of these complexes activated TR-SSRE to the same level, and that a promoter could only be occupied by a maximum of one complex. Thus, the output of the model was the sum of the concentrations of X^*, X_D^*, X_D^{**}, and $W^*X_D^*$ ($Sln1p^*$, $Sln1p_D^*$, $Sln1p_D^{**}$, and $Ypd1^*Sln1p_D^*$), which I called "activated Skn7". The input to all of the systems was $k(t) = 0.1(1 + sin(\omega t))$, which controls the production rate of Z (unphosphorylated Stat5b-HKRRp). Unless otherwise specified, the parameters used in the simulation are listed in Figure 4-9, except for p_T which was set at 100.

I found that, in general, the insulator was robust to parameter perturbations over
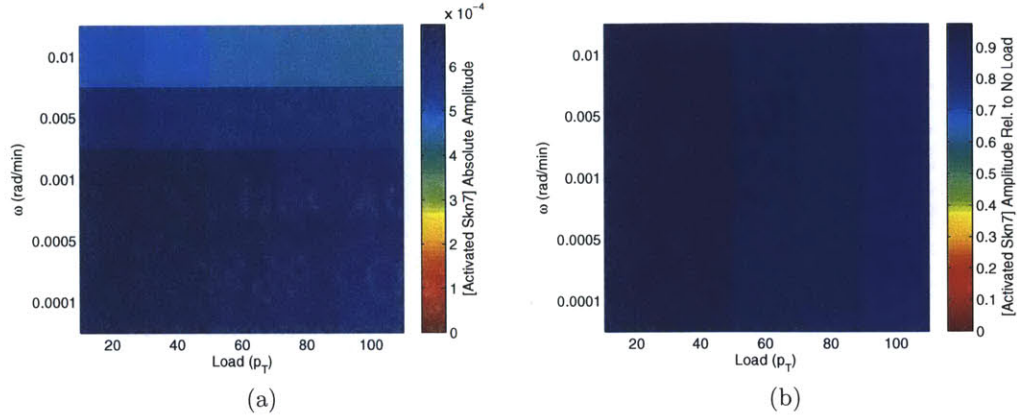
Figure 4-11: Synthetic insulator, ω and p_T varied, absolute amplitude and relative amplitude to no load.

a wide range. Note that in this study, only two select parameters were perturbed at once. Since other parameters remained at their default values, the entire space of parameter values was not explored. In the following simulations, both absolute amplitude of the output and the relative amplitude of the output to when no load was present were recorded. The relative amplitude is of greater interest, because it compares the behavior of the insulator when there is load to when there is no load. A good insulator is able to minimize retroactivity to the input, which is equivalent to keeping the relative amplitude of activated Skn7p to as close to 1 as possible.

The performance of the insulator appears to be fairly robust to changes in Ypd1p$_T$ and do poorly at low Skn7p$_T$ (Figure 4-10). Higher Skn7p$_T$ preserves a greater reservoir of activated Skn7p, which increases the capability of the insulator to attenuate the effect of higher loads. Higher loads results in lower relative output amplitude, indicating that although the insulator does attenuate retroactivity to the output, it does not completely eliminate it (Figure 4-11). Plots for variations in other parameters are presented in Section F.2. The insulator is robust to changes in all α parameters, which dictate the phosphorelay between Stat5b-HKRRp and Ypd1p. The insulator is affected to a greater extent by changes in β parameters, which dictate the phosphorelay between Ypd1p and Skn7p. Higher forward phosphotransfer rates from phosphorylated Ypd1p to unphosphorylated Skn7p (β_1 and β_3) and lower

84
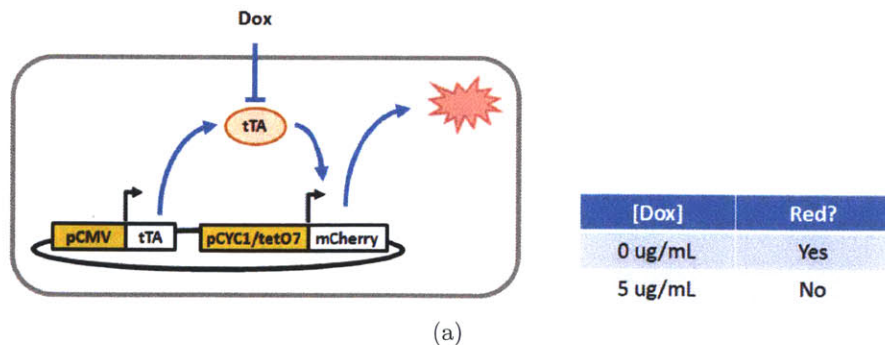
Figure 4-12: System 3 to test tTA functionality. (a) $pCMV$ constitutively drives tTA, which activates $pCYC1_{min}/tetO7$ driving mCherry, a red fluorescent protein. Dox inhibits tTA activation. (b) Qualitative expected fluorescent outputs for the systems.

reverse phosphotransfer rates (β_2 and β_4) result in better insulator performance. Interestingly, lower dimerization rates (k_f) and higher undimerization rates (k_r) result in better insulator performance. As expected, higher protein-DNA dissociation (k_D) results in better insulator performance. Lastly, higher degradation and phosphodecay rates (γ and k_{dp}), and lower phosphotransfer from Jak2p to Stat5b-HKRRp (k_2), decrease the absolute amplitude of the output, but do not affect the relative amplitude of the output. This is a result of the fact that only Stat5b-HKRRp, and not Ypd1p nor Skn7p, degrades. However, based on the findings in Section 3.2.2, if the degradation rate is low, this assumption should not greatly affect model behavior. Thus, overall, I found the insulator to be fairly robust to parameter perturbations. Some of parameter perturbations did decrease insulator performance. However, note that these parameters were varied over many magnitudes of order, such that assumptions about phosphotransfer reactions were no longer valid.

4.6 Preliminary Data

A construct to test the functionality of tTA was made with the two-step Gateway/Gibson assembly method (Figure 4-12). System 3 consists of $pCMV$, a low constitutive promoter, driving the expression of tTA. tTA activates a $pCYC1_{min}/tetO7$ promoter, which drives the production of mCherry, a red fluorescent protein. Addi-

(a) Negative control. (b) TEF-mCherry, integrated. (c) TEF-mCherry, CEN/ARSH plasmid.
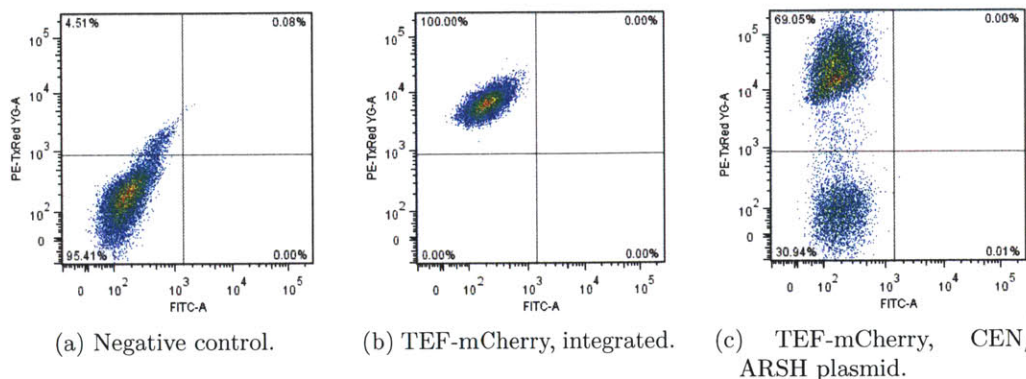
Figure 4-13: Flow cytometry data for TEF-mCherry construct (b) integrated and on a (c) URA selected CEN6/ARSH4 plasmid in yeast. Note that this mCherry part does not have a Kozak sequence. For comparison, data from a non-fluorescent strain is shown (a). Data was collected by Deepak Mishra.

tion of Dox inhibits tTA activation.

Flow cytometry (FCM) was used to collect data about the fluorescence levels of individual cells. (Method is in Section E.3.) Live cells (as opposed to dead cells or debris) were gated on the basis of forward and side scatter. Two fluorescent output channels were used to analyze cells. PE-TxRED detects fluorescence in the red light wavelength range, and FITC detects fluorescence in yellow light wavelength range. First, there was a noticeable difference between the fluorescent outputs of constructs that were stably integrated into the chromosome and of those on plasmids. In particular, the fluorescent levels of a strain with integrated TEF-mCherry and a strain harboring a centromeric plasmid with TEF-mCherry were compared (Figure 4-13). TEF is a strong constitutive promoter in yeast. Whereas cells with integrated mCherry form an unimodal fluorescent population, as shown along the PE-TxREF axis, cells with mCherry on plasmid form a bimodal population. Although the majority of cells fluoresce, a minority do not. The cells that are not fluorescing most likely do not harbor the plasmid. It has been previously shown that yeast cells can lose their plasmid, even under selection [36, 20].

Next, tTA activation activity was investigated with the system depicted in Figure 4-12a. As controls, fluorescence levels of cells harboring a centrometric plasmid without a fluorescent protein and those of cells harboring a TEF-mCherry construct

(a) Negative control, CEN/ARSH plasmid

(b) TEF-mCherry, CEN/ARSH plasmid

(c) Construct 3, CEN/ARSH plasmid, no Dox.

(d) Construct 3, CEN/ARSH plasmid, 5 ug/mL Dox.
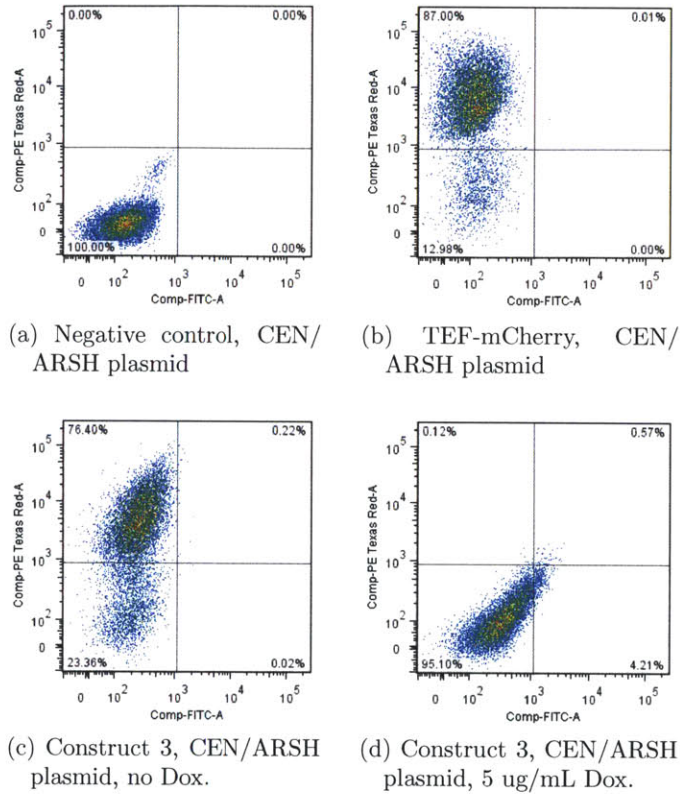
Figure 4-14: Flow cytometry data demonstrating tTA activation. As controls, data from (a) cells harboring a plasmid without a fluorescent protein and (b) cells harboring a plasmid with TEF-mCherry is presented. (c) Cells with construct 3 without Dox induction. tTA is able to activate the $pCYC1_{min}/tetO7$ promoter to levels almost comparable to the TEF promoter. (d) Cells with construct 3 with Dox induction. Dox blocks tTA activation, and fluorescence levels fall to levels comparable to the negative control. The data suggests that tTA is an effective activator and Dox is an effective inhibitor.

on a centromeric plasmid are shown in Figure 4-14a-b. Addition of Dox should decrease red fluorescence levels. Fluorescent levels do decrease with the addition of Dox, comparing Figures 4-14c-d. The data shows that tTA is a strong activator that can be inhibited by Dox.

The preliminary data presented in this section demonstrates the functionality of the assembly method described in Chapter 2. The assembly method, including the presence of *att*B sites in between the promoter, gene, and terminators still allowed the system to function as expected. In addition, the data suggests that experiments should be constructed with integrated systems when possible. When the system is placed on a centromeric plasmid, a portion of the cell population will not have the plasmid even if the plasmid is selected for. Many additional constructs, such as the ones presented in Figure 4-1, will be tested in a similar manner with Dox and IPTG.

4.7 Discussion and Future Work

In this chapter, constructs to measure and attenuate retroactivity were designed and modeled. These constructs can be made in a rapid manner from existing entry clones using the Gateway©-Gibson assembly method described earlier. As a demonstration of the functionality of the assembly strategy, preliminary data collected from a simple repressible system was shown.

Experimentation of the buffer construct will allow for quantitative measurement of retroactivity at an uninsulated interconnection between two regulatory genetic networks. Data collected at various loads can be fit to the model presented in Section 4.2 and compared to $\mathcal{R}(\bar{X})$, a measure of retroactivity to the output, presented in Section 3.1.1. The effect of varying the frequency and amplitude of inducer concentration (Dox or IPTG), the effect of mutating transcription binding sites to increase their k_D, and the effect of fusing different degradation tags with the signaling transcription factor (detailed in Appendix D) to the system output can be measured and compared to the theoretical model presented in Figures 3-10, 3-9, and 3-8, respectively. As another example, the strength of the promoter driving expression of the signaling transcription factor can also be changed. Retroactivity should be significantly greater

88

when the number of signaling transcription factor is equal to or less than the number of binding sites in the cell.

Likewise, experimentation of the insulation device will allow for measurement of retroactivity attenuation. If successful, this insulation device will be the first to experimentally verify the theoretical predictions in Section 3.2 on the use of a phosphorylation - dephosphorylation cycle as an insulator, which operates through the separation of the slow timescale of genetic regulatory networks and the fast timescale of protein phosphotransfer. The effectiveness of the insulation device can be varied through modifications of the *JAK2-STAT5B:HKRR-YPD1-SKN7* signaling cascade, as detailed in Section 4.3. For example, as a negative control, the residues that are phosphorylated in the signaling cascade can be knocked out. Instead of using the full length Jak2 protein, use of just its JH1 domain could result in higher activation of Stat5b, changing the dynamics of the insulator. Alternatively, the Sln1 HKRR can be linked differently, and various degradation tags can be placed on components in the signaling pathway to change their cellular expression levels, as described in Appendix D. Varying the properties of the insulator will result in better understanding of the features necessary to attenuate retroactivity.

The study of retroactivity and of insulators to attenuate such effect will further our understanding of device modularity. By increasing the reliability and predictability of device behavior in different circuit settings, synthetic biology will be better able to construct more predictable, large scale circuits. Iterative redesign of system will be less necessary when new components are added to a system, resulting in faster construction of systems in synthetic biology. Complex regulatory networks will be able to be quickly assembled from libraries of functional devices and behave with more predictable behavior. Such development will facilitate the ability of synthetic biology to tackle issues that include human therapeutics, drug manufacturing, materials and biofuel production, and environmental remediation.

Appendix A

Mathematical Analysis of Retroactivity and Insulator

In Section A.1, I rederive the derivation of retroactivity presented before in [22]. In Section A.2, I rederive the how a phosphorylation - dephosphorylation cycle buffers against retroactivity, also presented before in [22].

A.1   Quantification of Retroactivity

Development of a mathematical measurement of retroactivity across interconnections can elucidate engineering principles on how to limit retroactivity to inputs and outputs. Such a framework quantifies how sending a signal to downstream components affects the sender's dynamical behavior.

A technique called singular perturbation can be used to study the effect of retroactivity on X in the system presented in Equation (3.3). Singular perturbation can be used to simplify the analysis of systems that can be separated into different timescales. A system can be separated into two independent subsystems and analyzed separately, if one of the subsystems is assumed to act on a much faster timescale than the other.

In biological systems, protein production/decay and protein-DNA binding/unbinding have kinetic rates that differ by orders of magnitude. In particular, protein binding/unbinding, which takes on the order of seconds, is much faster than protein production/decay, which takes on the order of minutes. Thus, dynamics of proteins

can be approximated by its behavior on its slow manifold, dominated by the behavior of the slow acting subsystem (production/decay), as the fast acting subsystem (binding/unbinding) is approximated to be invariant.

A.1.1 Without Complex Decay

To carry out singular perturbation analysis, the parameter ϵ is used to represent that separation between the two timescales. In particular, epsilon is defined as follows:

$$\epsilon = \frac{\delta}{k_{off}} \tag{A.1}$$

Since the rate of protein production/decay is much slower than that of DNA-protein binding/unbinding, $k_{off}, k_{on} \gg \delta$, so $\epsilon \ll 1$. Fast kinetic parameters can then be determined in terms of ϵ: $k_{off} = \frac{\delta}{\epsilon}$ and $k_{on} = \frac{\delta}{\epsilon K_d}$, where $K_d = \frac{k_{off}}{k_{on}}$.

Using ϵ, Equation (3.4) can be rewritten as follows.

$$\begin{aligned}
\frac{dX}{dt} &= k(t) - \delta X + \frac{\delta}{\epsilon}C - \frac{\delta}{\epsilon K_d}(p_T - C)X \\
\frac{dC}{dt} &= -\frac{\delta}{\epsilon}C + \frac{\delta}{\epsilon K_d}(p_T - C)X
\end{aligned} \tag{A.2}$$

A change of variables $y = X + C$ is performed, where y is the slow variable that represents the total amount of X in the system, to obtain the following differential equations:

$$\frac{dy}{dt} = k(t) - \delta(y - C) \tag{A.3a}$$

$$\epsilon\frac{dC}{dt} = -\delta C + \frac{\delta}{K_d}(p_T - C)(y - C) \tag{A.3b}$$

The above equations are now in the correct form for singular perturbation analysis. Denote the expression on the right hand side of Equation (A.3b) as $g(C, y)$. Since protein-DNA binding/unbinding occurs much faster than protein production and decay, ϵ is approximated to be 0. Denote the approximations of X, y, and C as $\bar{X}$, $\bar{y}$, and $\bar{C}$. Let $\gamma(\bar{y})$ be the smallest zero of $g(C, y) = 0$, and set $\bar{C}$ equal to $\gamma(\bar{y})$. Note that $\gamma(\bar{y})$ is thus the quasi-steady state value of $\bar{C}$. The above equations now become

the following:

$$\frac{d\bar{y}}{dt} = k(t) - \delta(\bar{y} - \gamma(\bar{y})) \tag{A.4}$$

Because $\bar{X} = \bar{y} - \bar{C}$, $\frac{d\bar{X}}{dt} = \frac{d\bar{y}}{dt} - \frac{d\bar{C}}{dt}$. Note that $\frac{d\bar{C}}{dt} = \frac{d\gamma(\bar{y})}{d\bar{y}} \frac{d\bar{y}}{dt}$. Therefore, the following equation is obtained:

$$\begin{aligned}
\frac{d\bar{X}}{dt} &= (1 - \frac{d\gamma(\bar{y})}{d\bar{y}})\frac{d\bar{y}}{dt} \\
&= (1 - \frac{d\gamma(\bar{y})}{d\bar{y}})(k(t) - \delta\bar{X})
\end{aligned} \tag{A.5}$$

The implicit function theorem is used to evaluate $\frac{d\gamma(\bar{y})}{d\bar{y}}$ in Equation (A.6). To obtain the fourth equality below, $\bar{C} = \frac{\bar{X}p_T}{k_D + \bar{X}}$, which comes from $g(\gamma(\bar{y}), \bar{y}) = 0$ and $\bar{X} = \bar{y} - \bar{C}$.

$$\begin{aligned}
\frac{d\gamma(\bar{y})}{d\bar{y}} &= -\frac{\partial g/\partial \bar{y}}{\partial g/\partial \bar{C}} \\
&= \frac{\frac{1}{K_d}(p_T - \bar{C})}{1 + \frac{1}{K_d}(p_T - \bar{C}) + \frac{1}{K_d}(\bar{y} - \bar{C})} \\
&= \frac{1}{1 + \frac{K_d + \bar{y} - \bar{C}}{p_T - \bar{C}}} \\
&= \frac{1}{1 + \frac{K_d + \bar{X}}{p_T - \frac{\bar{X}p_T}{k_D + \bar{X}}}} \\
&= \frac{1}{1 + \frac{(K_d + \bar{X})^2}{p_T K_d}} \\
&= \frac{1}{1 + \frac{(1 + \bar{X}/K_d)^2}{p_T/K_d}}
\end{aligned} \tag{A.6}$$

$\frac{d\gamma(\bar{y})}{d\bar{y}}$ is always between values 0 and 1 and quantifies how much the dynamics of $\bar{X}$ change when the system is isolated to when the system is connected to downstream modules (compare Equations (3.2) and (A.5)). Thus, $\frac{d\gamma(\bar{y})}{d\bar{y}}$ is a measure of the effect of retroactivity to the output on $\bar{X}$ dynamics (see Section 3.1.1 for more).

Retroactivity to the input can be calculated in the similar fashion, replacing $\bar{X}$ with concentration of the particular transcription factor that activates the promoter driving X, and p_T with the total number of binding sites of that particular transcription factor in the system.

93

A.1.2 With Complex Decay

Equation (3.6) is analyzed in a similar manner as before with singular perturbation. Using ϵ, defined in Equation (A.1), Equation (3.6) can be rewritten as the following:

$$\frac{dX}{dt} = k(t) - \delta X + \frac{\delta}{\epsilon}C - \frac{\delta}{\epsilon K_d}(p_T - C)X$$
$$\frac{dC}{dt} = -\frac{\delta}{\epsilon}C + \frac{\delta}{\epsilon K_d}(p_T - C)X - \delta C \tag{A.7}$$

A change of variables $y = X + C$ results in the following:

$$\frac{dy}{dt} = k(t) - \delta X - \delta C$$
$$\epsilon\frac{dC}{dt} = -\delta C + \frac{\delta}{K_d}(p_T - C)X - \epsilon\delta C \tag{A.8}$$

Since $\bar{y} = \bar{X} + \bar{C}$, $\frac{d\bar{y}}{dt} = \frac{d\bar{X}}{dt} + \frac{d\bar{C}}{dt}$. Approximating ϵ as zero does not change the result for $\frac{d\bar{C}}{dt}$, which is equal to $\frac{d\gamma(\bar{X})}{d\bar{X}}\frac{d\bar{X}}{dt}$. Thus, the following equation results:

$$\frac{d\bar{y}}{dt} = \frac{d\bar{X}}{dt} + \frac{d\gamma(\bar{X})}{d\bar{X}}\frac{d\bar{X}}{dt}$$
$$\frac{d\bar{X}}{dt} = \frac{1}{1 + \frac{d\gamma(\bar{X})}{d\bar{X}}}\frac{d\bar{y}}{dt} \tag{A.9}$$
$$= \frac{1}{1 + \frac{d\gamma(\bar{X})}{d\bar{X}}}(k(t) - \delta\bar{X} - \delta\gamma(\bar{X}))$$

Note that $\frac{1}{1 + \frac{d\gamma(\bar{X})}{d\bar{X}}}$ evaluates to the same expression as $1 - \frac{d\gamma(\bar{y})}{d\bar{y}}$. Modeling complex decay adds an additional $-\delta\gamma(\bar{X})$ to the second multiplicand expression in the last equation in Equation (A.9). This addition alone increases the negative feedback of the system, which decreases the steady state and the time to reach steady state. However, the presence of the first multiplicand, $\frac{1}{1 + \frac{d\gamma(\bar{X})}{d\bar{X}}}$, is less than 1, which increases the time to reach steady state. Modeling complex decay will reduce the steady state value, but $\frac{1}{1 + \frac{d\gamma(\bar{X})}{d\bar{X}}}$ alone cannot determine if the system will reach steady state earlier or later in an unbuffered system.

Section 3.1.2 computationally analyzes how the assumption that the complex does or does not decay affects model behavior. Figure A-1 analyzes how the approximation

94

(a) s.s. $[X] = 1$, no C decay   (b) s.s. $[X] = 50$, no C decay

(c) s.s. $[X] = 1$, with C decay   (d) s.s. $[X] = 50$, with C decay
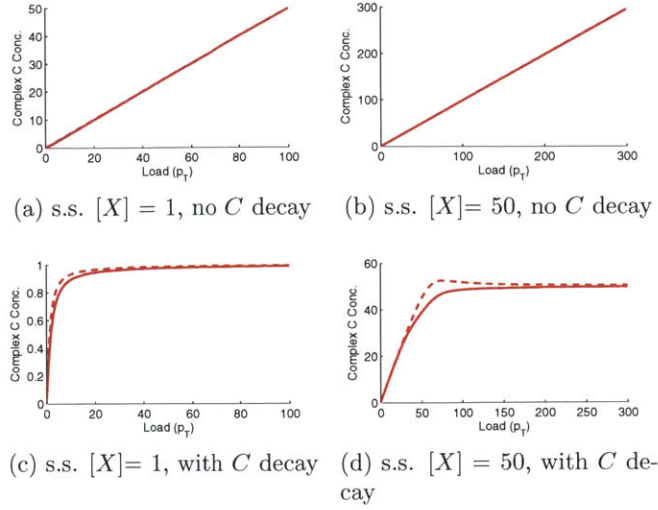
Figure A-1: Steady state concentration of protein-DNA complex. Solid line represents the approximated complex amount $\bar{C} = \frac{\bar{X}p_T}{k_D + \bar{X}}$ used in singular perturbation calculations in Section A.1. This equality results from setting $\epsilon = 0$ in Equation (A.3b). Dashed line represents the actual complex amount C. When the complex is assumed to not decay, the approximation is exact. When the complex is assumed to decay, the approximation is close only at very low and high loads. The degradation of C, although it is slower acting than X binding/unbinding, affects the steady state level of C.

$\bar{C} = \frac{\bar{X}p_T}{k_D + \bar{X}}$, used in Equation A.6, holds when the complex does decay.

A.2   Insulator to Attenuate Retroactivity

In this section, an insulator is first modeled with simplified one-step phosphorylation - dephosphorylation reactions, and simulation results with this model are shown. Theoretical analysis of the insulator modeled with two-step phosphorylation - dephosphorylation reactions, presented in Section 3.2.1, is then discussed. Simulation results of this more complicated model are shown in Section 3.2.2.

A.2.1   Modeling with One-step Phosphorylation - Dephosphorylation Reactions

The performance of an insulator is examined with simplified one-step phosphorylation - dephosphorylation reactions. X is a protein that can be phosphorylated by

the kinase X and dephosphorylated by the phosphatase Y. Only phosphorylated X (denoted X_p) can bind to DNA to form the complex C.

$$X + Z \xrightarrow{k_1} X_p + Z$$

$$X_p + Y \xrightarrow{k_2} X + Y \qquad (A.10)$$

$$X_p + p \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} C$$

In this model, total amounts of X and Y are conserved, whereas Z is the input signal. Thus, $X_T = X + X_p + C$. From Equation (A.10), and replacing X with $X_T - X_p - C$, the following differential equations are obtained:

$$\frac{dX_p}{dt} = k_1 Z(t) X_T \left(1 - \frac{X_p}{X_T} - \boxed{\frac{C}{X_T}}\right) - k_2 Y X_p + \boxed{k_{off}C - k_{on}X_p(p_T - C)}$$

$$\frac{dC}{dt} = -k_{off}C + k_{on}X_p(p_T - C) \qquad (A.11)$$

Retroactivity to the output is represented by terms in the box.

If Z is assumed to be a weak activator of X $(X_p \ll X_T)$ and the total amount X is much greater than the load $(p_T \ll X_T)$, Equation (A.11) simplifies to as follows:

$$\frac{dX_p}{dt} = k_1 Z(t) X_T - k_2 Y X_p + \boxed{k_{off}C - k_{on}X_p(p_T - C)}$$

$$\frac{dC}{dt} = -k_{off}C + k_{on}X_p(p_T - C) \qquad (A.12)$$

Using singular perturbation analysis, and setting $G = k_1 X_T$ and $G' = k_2 Y$, the system simplifies to the following:

$$\frac{dX_p}{dt} = (GZ(t) - G'Xp)(1 - d(t)). \qquad (A.13)$$

To attenuate retroactivity to the output, the gains G and G' should be large such that the dynamics of X_p approach the dynamics of X_p in the unconnected system:

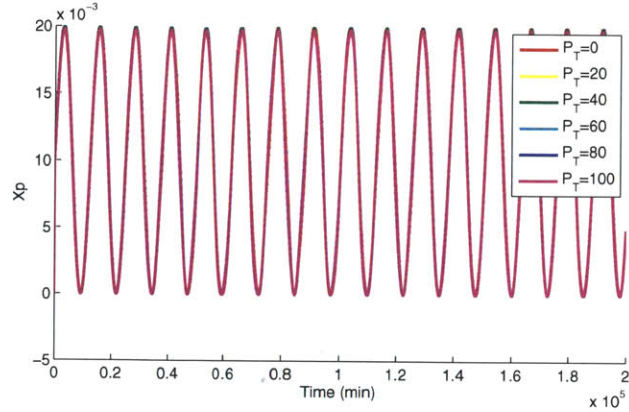$$\frac{dX_p}{dt} = (GZ(t) - G'Xp) \qquad (A.14)$$

Figure A-2: Insulator modeled with one-step phosphorylation - dephosphorylation reactions. Insulator attenuates retroactivity to the output well. Parameters used were: $k(t) = 0.01(1 + sin(\omega t))$, $\omega = 0.005$, $\delta = 0.01$, $k_1 = 50$, $k_2 = 50$, $k_{on} = 10$, $k_{off} = 10$, $X_T = 150$, and $Y_T = 150$.

Thus, in order to act as an insulator, the system should have large amounts of X_T and Y, or fast phosphorylation and dephosphorylation rate constants.

Simulations of the insulator modeled with one-step phosphorylation - dephosphorylation reactions indicate that the insulator is effective at buffering against retroactivity (Figure A-2). In this model, its performance is unaffected by changes in the frequency of the input and the number of downstream binding sites (Figure A-3). The system performs better with larger amounts of X_T and Y (Figure A-4) and fast phosphorylation and dephosphorylation rate constants (Figure A-5), confirming the analysis in Subsection A.2.1.
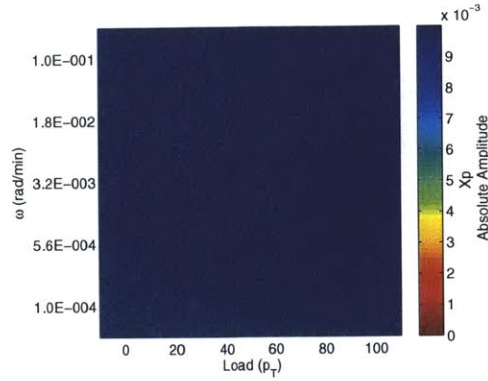
Figure A-3: Insulator modeled with one-step phosphorylation - dephosphorylation reactions with varied load and ω. In this model, the performance of the insulator is unaffected by changes in the frequency of the input and the number of downstream binding sites.
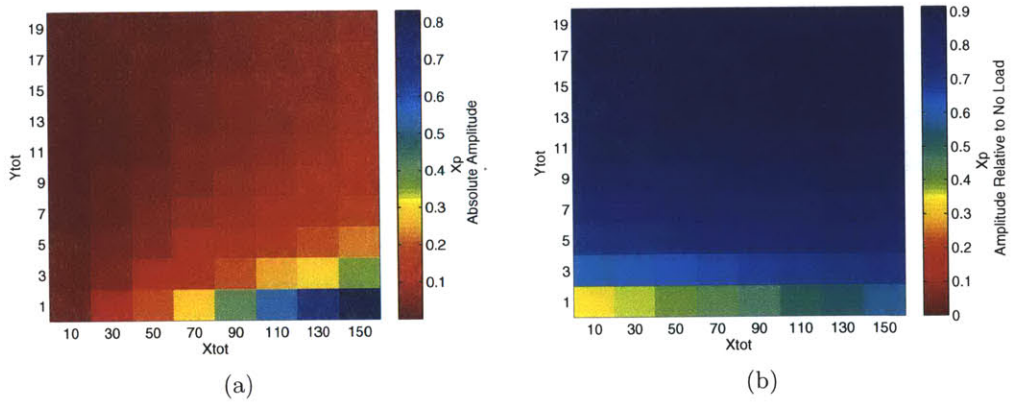

Figure A-4: Insulator modeled with one-step phosphorylation - dephosphorylation reactions with varied X_T and Y_T. In this model, larger concentrations of Y, the phosphatase, lower the absolute amount of phosphorylated X, and larger total concentrations of X better attenuate retroactivity to the output.
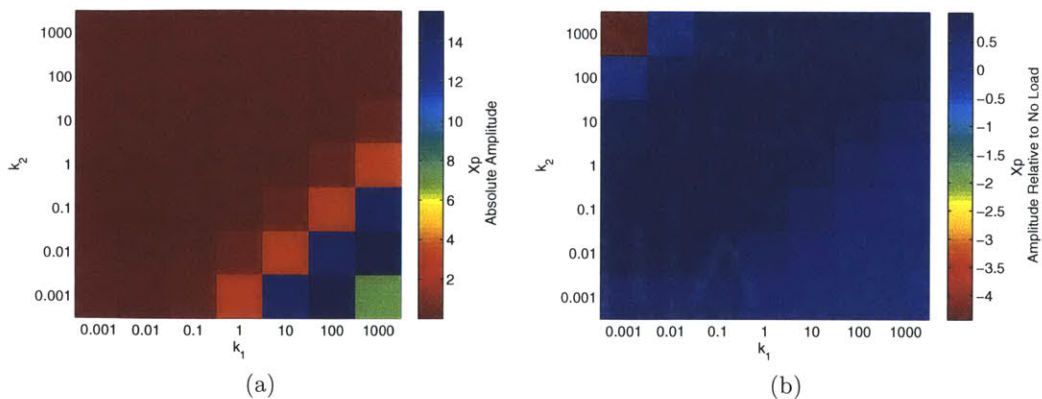
Figure A-5: Insulator modeled with one-step phosphorylation - dephosphorylation reactions with varied k_1 and k_2. When either k_1 or k_2 is small (slow phosphorylation or dephosphorylation), the insulator in this model does not perform well.

A.2.2 Modeling with Two-step Phosphorylation - Dephos-phorylation Reactions

The chemical equations for an insulator modeled with two-step phosphorylation - dephosphorylation reactions are presented in Section 3.2.1. Here, theoretical analysis of the system is reviewed.

Starting with Equation 3.9, the total amount X is assumed to be much greater than the load ($p_T \ll X_T$), such that C/X_T can be approximated as 0. To transform the equations into singular perturbation form, the variable $\epsilon = \frac{\delta}{k_{off}} \ll 1$ is introduced. $K_d = \frac{k_{off}}{k_{on}}$. Then $k_{off} = \frac{\delta}{\epsilon}$ and $k_{on} = \frac{\delta}{\epsilon K_d}$. The slow variable $z = Z + C_1$ is the total amount of kinase. With these variables, the system can be transformed into the

following form for singular perturbation analysis:

$$\frac{dz}{dt} = k(t) - \delta z \qquad (\text{A.15a})$$

$$\epsilon \frac{dC_1}{dt} = \epsilon \alpha_1 X_T (z - C_1)(1 - \frac{X_p}{X_T} - \frac{C_1}{X_T} - \frac{C_2}{X_T}) - \epsilon(\alpha_2 + k_1)C_1 \qquad (\text{A.15b})$$

$$\epsilon \frac{dC_2}{dt} = \epsilon \beta_1 Y_T X_p (1 - \frac{C_2}{Y_T}) - \epsilon(k_2 + \beta_2)C_2 \qquad (\text{A.15c})$$

$$\epsilon \frac{dX_p}{dt} = -\epsilon \beta_1 Y_T X_p (1 - \frac{C_2}{Y_T}) + \epsilon k_1 C_1 + \epsilon \beta_2 C_2 + \delta C - \frac{\delta}{K_d} X_p (P_T - C) \qquad (\text{A.15d})$$

$$\epsilon \frac{dC}{dt} = -\delta C + \frac{\delta}{K_d} X_p (p_T - C) \qquad (\text{A.15e})$$

Retroactivity to the Output

To approximate the system on its slow manifold, ϵ is set to 0, causing the terms in the box to drop out. Because the internal dynamics of the insulator occur on a timescale much faster than that of Z, the dynamics of output X_p is buffered from retroactivity to the output. Let $h_1(X_p)$ and $h_2(X_p)$ denote the expressions to the right hand side of the equation sign in Equations (A.15b) and (A.15c), respectively. To determine the quasi-steady state values of C_1 and C_2 with $\epsilon = 0$, the roots of $h_1(C_1, C_2, X_p) = 0$ and $h_2(C_2, X_p) = 0$ are solved. First, $h_2(C_2, X_p) = 0$ is solved to obtain an expression for C_2 as a function for X_p and called $g_2(X_p)$.

$$0 = \epsilon \beta_1 Y_T X_p (1 - \frac{C_2}{Y_T}) - \epsilon(k_2 + \beta_2)C_2$$

$$C_2 = \frac{\beta_1 Y_T X_p}{\beta_1 X_p + \beta_2 + k_2} = \frac{Y_T X_p / \gamma}{1 + X_p / \gamma} = g_2(X_p) \qquad (\text{A.16})$$

100

where $\gamma = \frac{\beta_2 + k_2}{\beta_1}$. Next, $h_1(C_1, C_2, X_p) = 0$ for C_1 is solved as a function for X_p and called $g_1(X_p)$. The previous result for C_2 is used in the solution.

$$0 = -\epsilon \beta_1 Y_T X_p (1 - \frac{C_2}{Y_T}) + \epsilon k_1 C_1 + \epsilon \beta_2 C_2$$

$$k_1 C_1 = (\beta_1 X_p - \beta_2) C_2 + \beta_1 Y_T X_p$$

$$k_1 C_1 = (\beta_1 X_p - \beta_2) \frac{Y_T X_p / \gamma}{1 + X_p / \gamma} + \beta_1 Y_T X_p \qquad (A.17)$$

$$k_1 C_1 = \frac{X_p Y_T (-\beta_2 / \gamma + \beta_1)}{1 + X_p / \gamma}$$

$$C_1 = \frac{X_p Y_T k_2 / (\gamma k_1)}{1 + X_p / \gamma} = g_1(X_p)$$

Equation (A.15b) is substituted with $\epsilon = 0$ and rearranged to obtain the following:

$$0 = \alpha_1 Z X_T (1 - \frac{X_p}{X_T} - \frac{C_1}{X_T} - \frac{C_2}{X_T}) - \epsilon(\alpha_2 + k_1) C_1$$

$$C_1(\alpha_1 Z + \alpha_2 + k_1) = \alpha_1 Z X_T - \alpha_1 Z X_p - \alpha_1 Z C_2 \qquad (A.18)$$

For simplicity, it is assume that $Xp \ll \gamma$. Then, $C_1 = X_p Y_T k_2 / (\gamma k_1)$ and $C_2 = Y_T X_p / \gamma$. These values are substituted into (A.18) to obtain the following:

$$(X_p Y_T k_2 / (\gamma k_1))(\alpha_1 Z + \alpha_2 + k_1) = \alpha_1 Z X_T - \alpha_1 Z X_p - \alpha_1 Z Y_T X_p / \gamma$$

$$X_p(Z) = \frac{\alpha_1 Z X_T}{\alpha_1 Z(1 + \frac{Y_T}{\gamma} + \frac{Y_T k_2}{\gamma k_1}) + \frac{Y_T k_2}{k_1 \gamma}(\alpha_2 + k_1)} \qquad (A.19)$$

To preserve linearity between the input and output signals, $X_p(Z)$ should be a linear function of Z. This occurs if the second term in the denominator of the expression for $X_p(Z)$ (A.19) is greater than the first, as follows:

$$\alpha_1 Z(1 + \frac{Y_T}{\gamma} + \frac{Y_T k_2}{\gamma k_1}) \ll \frac{Y_T k_2}{k_1 \gamma}(\alpha_2 + k_1)$$

$$Z \ll \frac{\frac{Y_T k_2}{k_1 \gamma}(\alpha_2 + k_1)}{\alpha_1(1 + \frac{Y_T}{\gamma} + \frac{Y_T k_2}{\gamma k_1})} \qquad (A.20)$$

$$Z \ll \frac{\frac{Y_T k_2 \gamma'}{k_1 \gamma}}{1 + \frac{Y_T}{\gamma} + \frac{Y_T k_2}{\gamma k_1}}$$

where $\gamma' = \frac{\alpha_2 + k_1}{\alpha_1}$. Then the following expression for $X_p(Z)$ is obtained:

$$X_p(Z) = \frac{Z X_t \gamma k_1}{Y_T \gamma' k_2} \tag{A.21}$$

The input/output gain is

$$\frac{X_p(Z)}{Z} = \frac{X_t \gamma k_1}{Y_T \gamma' k_2} \tag{A.22}$$

In order to prevent attenuation of the signal through the insulator, the input/output gain should be:

$$\frac{X_p(Z)}{Z} \geq 1 \tag{A.23}$$

To summarize, the three conditions for insulator preservation of signal in the face of increased load are $p_T \ll X_T$, Equation (A.20), and Equation (A.23).

Retroactivity to the Input

Retroactivity to the input can be calculated in a manner similar to Section 3.1.1. The dynamics of Z on the slow manifold are obtained by setting $\epsilon = 0$, which results in:

$$\frac{dZ}{dt} = (k(t) - \delta Z)(1 - \frac{dg_1}{dz}) \tag{A.24}$$

Here, $\frac{dg_1}{dz}$ quantifies effect of retroactivity to the input on Z dynamics.

$$\frac{dg_1}{dz} = \frac{dg_1}{dX_p}\frac{X_p}{dz} \tag{A.25}$$

Using the same three assumptions, the following results:

$$\frac{dF_1}{dX_p} = \frac{Y_T k_2}{\gamma k_1}$$
$$\frac{X_p}{dz} = \frac{X_t \gamma k_1}{Y_T \gamma' k_2} \tag{A.26}$$

Therefore,

$$\frac{dF_1}{dz} = \frac{X_t}{\gamma'} \tag{A.27}$$

Thus, retroactivity to the input is small if $\frac{X_t}{\gamma'} < 1$. This concludes the theoretical analysis of the insulator. Simulations of the insulator model are presented in Section 3.2.2.

Appendix B

Retroactivity in a Activator-Repressor Oscillator

In this chapter, I computationally find that high loads can break an oscillator, demonstrating retroactivity. I briefly demonstrate *in silico* that the insulator proposed in Section 4.4 can buffer an oscillator from retroactivity to the output.

As an example of how retroactivity can disturb expected module performance, the dynamics of a clock driving differing number of downstream components is explored. In particular, a activator-repressor clock, also known as a relaxation clock, is examined [5]. The activating protein A induces its own production and also that of the repressor protein R, which suppresses the production of A (Figure B-1). The system can be described by the differential equations in (B.1) [21]. μ is a positive parameter that represents the difference in rates between the activating loop and the repressing loop.
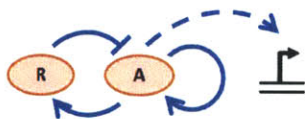

Figure B-1: Activator-repressor clock. In this scenario, the activator can also bind to promoters and activate genes.

$$\frac{dA}{dt} = \mu \left(\frac{\alpha A^2 + \alpha_0}{K_1 + R^2 + A^2} - \delta A \right)$$
$$\frac{dR}{dt} = \frac{\alpha A^2}{K_2 + A^2} - \delta R \tag{B.1}$$

This clock can be used to synchronize multiple downstream devices if A acts as the input to these devices. Thus, there is a "downstream" flow from the clock to the devices. However, A physically interacts with the promoters of downstream devices, and this interaction can affect the internal workings of the clock. Thus, there is an undesired "upstream" flow from the devices to the clock. Equation (B.2) characterizes the new system with downstream devices, and the boxed terms are the effect of retroactivity to the output on the system. Figure B-2 shows that high load can in fact destroy oscillator performance.

$$\frac{dA}{dt} = \mu \left(\frac{\alpha A^2 + \alpha_0}{K_1 + R^2 + A^2} - \delta A \right) - \boxed{k_{on2} \, A \, p_{in,T} - C + k_{off2} \, C}$$
$$\frac{dR}{dt} = \frac{\alpha A^2}{K_2 + A^2} - \delta R \tag{B.2}$$
$$\frac{dC}{dt} = k_{on2} \, A \, (p_{in,T} - C) - k_{off2} \, C$$

To determine if the proposed insulator could buffer an oscillator against retroactivity to the output, a system with the insulator attached to the output of the activator-repressor oscillator was simulated. Figure B-3 demonstrates that the proposed insulator *in silico* can attenuate the effect from being attached to many downstream components.
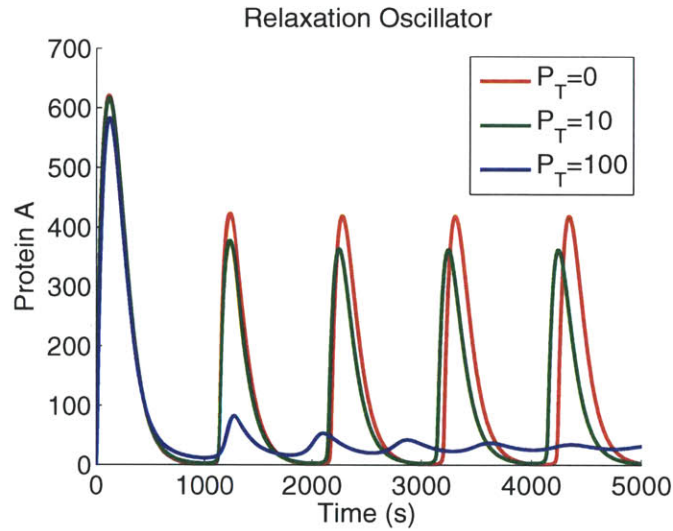
Figure B-2: Effect of load on oscillator performance. At low loads ($p_T = 10$), $[A]$ continues to oscillate. At higher loads ($p_T = 100$), enough A is sequestered away from the system such that the system no longer oscillates. Parameters used were $A_{init} = 0$, $R_{init} = 0$, $\mu = 1.5$, $\alpha_0 = 50$, $\alpha = 10$, $K_1 = 10$, $K_2 = 6000$, $k_{on2} = 1$, $k_o f f 2 = 20$, and $\delta = .006$.
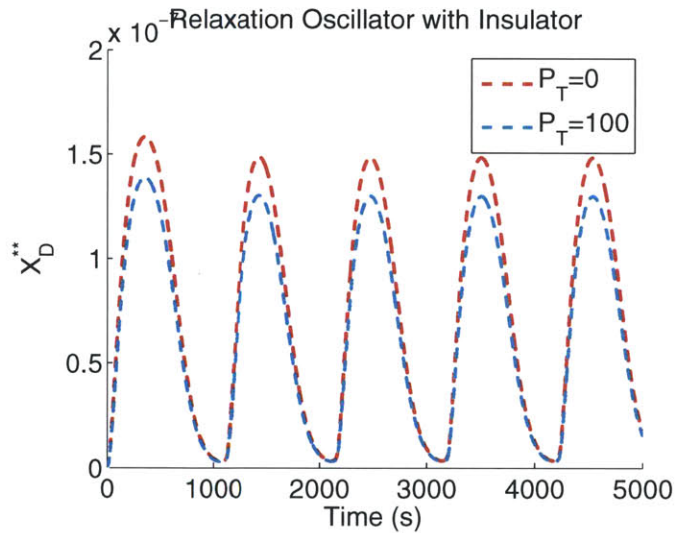

Figure B-3: Effect of load on insulated oscillator performance. At low and high loads, the output of the oscillator still oscillates. Parameters used for the oscillator are defined in Figure B-2.

107

Appendix C

Modeled Equations and Kinetics

C.1   Buffer Design

Equations and kinetic parameters used to generate the plots in Section 4.2 are presented in Tables C.1 and C.2. When possible, kinetic parameters were obtained from literature. Footnotes in Table C.2 are listed as follows: (1) Calculated from a half life of 70 minutes. The VP16 activation domain might be the reason for the increased decay rate compared to $LacI_4$ and GFP. (2) The PEST end of the G1 cyclin Cln2 is fused to decrease its half life from 7 hours to 30-90 minutes [79]. (3) Yeast cells are assumed to double every 120 minutes.

Table C.1: Modeled Biochemical Reactions

	Reaction Desc.	Reaction	Reaction Rate
1	tTA prod.	$p_{tTA} \rightarrow p_{tTA} + tTA$	$\alpha_1 {}^* p_{tTA}$
2	tTA dimer.	$2\ tTA \rightarrow tTA_2$	$k_1 {}^* tTA^2$
3	tTA$_2$ undimer.	$tTA_2 \rightarrow 2\ tTA$	$k_{-1} {}^* tTA_2$
4	tTA$_2$ + Dox bind.	$Dox + tTA_2 \rightarrow tTA_2\text{-}Dox$	$k_2 {}^* Dox {}^* tTA_2$
5	tTA$_2$-Dox unbind.	$tTA_2\text{-}Dox \rightarrow Dox + tTA_2$	$k_{-2a} {}^* tTA_2\text{-}Dox$
6	tTA$_2$-Dox undimer.	$tTA_2\text{-}Dox \rightarrow 2\ tTA + Dox$	$k_{-2b} {}^* tTA_2\text{-}Dox$
7	tTA$_2$ + tetO bind.	$tetO + tTA_2 \rightarrow tTA_2\text{-}tetO$	$k_3 {}^* tetO {}^* tTA_2$
8	tTA$_2$-tetO unbind.	$tTA_2\text{-}tetO \rightarrow tTA_2 + tetO$	$k_{-3a} {}^* tTA_2\text{-}tetO$
9	tTA$_2$-tetO undimer.	$tTA_2\text{-}tetO \rightarrow 2\ tTA + tetO$	$k_{-3b} {}^* tTA_2\text{-}tetO$
10	tTA$_2$-tetO + Dox bind.	$tTA_2\text{-}tetO + Dox \rightarrow tTA_2\text{-}Dox + tetO$	$k_4 {}^* tTA_2\text{-}tetO {}^* Dox$
11	tTA decay/dilution	$tTA \rightarrow \emptyset$	$\delta_1 {}^* tTA$
12	tTA$_2$ decay/dilution	$tTA_2 \rightarrow \emptyset$	$\delta_2 {}^* tTA_2$
13	tTA$_2$-Dox decay	$tTA_2\text{-}Dox \rightarrow Dox$	$\delta_3 {}^* tTA_2\text{-}Dox$
14	tTA$_2$-tetO decay	$tTA_2\text{-}tetO \rightarrow tetO$	$\delta_4 {}^* tTA_2\text{-}tetO$
15	LacI prod.	$tTA_2\text{-}tetO \rightarrow tTA_2\text{-}tetO + LacI$	$\alpha_2 {}^* tTA_2\text{-}tetO$
16	LacI dimer.	$2\ LacI \rightarrow LacI_2$	$k_5 {}^* LacI^2$
17	LacI$_2$ undimer.	$LacI_2 \rightarrow 2\ LacI$	$k_{-5} {}^* LacI_2$
18	LacI$_2$ dimer.	$2\ LacI_2 \rightarrow LacI_4$	$k_6 {}^* LacI_2^2$
19	LacI$_4$ undimer.	$LacI_4 \rightarrow 2\ LacI_2$	$k_{-6} {}^* LacI_4$
20	LacI$_4$ + LacOA bind.	$LacI_4 + LacOA \rightarrow LacI_4\text{-}LacOA$	$k_7 {}^* LacI_4 {}^* LacOA$
21	LacI$_4$-LacOA unbind.	$LacI_4\text{-}LacOA \rightarrow LacI_4 + LacOA$	$k_{-7} {}^* LacI_4\text{-}LacOA$
22	LacI$_4$ + LacOB bind.	$LacI_4 + LacOB \rightarrow LacI_4\text{-}LacOB$	$k_8 {}^* LacI_4 {}^* LacOB$
23	LacI$_4$-LacOB unbind.	$LacI_4\text{-}LacOB \rightarrow LacI_4 + LacOB$	$k_{-8} {}^* LacI_4\text{-}LacOB$
24	LacI decay/dilution	$LacI \rightarrow \emptyset$	$\delta_5 {}^* LacI$
25	LacI$_2$ decay/dilution	$LacI_2 \rightarrow \emptyset$	$\delta_6 {}^* LacI_2$
26	LacI$_4$ decay/dilution	$LacI_4 \rightarrow \emptyset$	$\delta_7 {}^* LacI_4$
27	LacI$_4$-LacOA decay	$LacI_4\text{-}LacOA \rightarrow LacOA$	$\delta_8 {}^* LacI_4\text{-}LacOA$
28	LacI$_4$-LacOB decay	$LacI_4\text{-}LacOB \rightarrow LacOB$	$\delta_9 {}^* LacI_4\text{-}LacOB$
29	CFP prod.	$LacOA \rightarrow LacOA + CFP$	$\alpha_3 {}^* LacOA$
30	YFP prod.	$LacOB \rightarrow LacOB + YFP$	$\alpha_4 {}^* LacOB$
31	CFP decay	$CFP \rightarrow \emptyset$	$\delta_{10} {}^* CFP$
32	YFP decay	$YFP \rightarrow \emptyset$	$\delta_{11} {}^* YFP$
33	Dox dilution	$Dox \rightarrow \emptyset$	$\delta_{12} {}^* Dox$

Table C.2: Kinetic Constants

	Reaction Desc.	Param	Param Value	Ref.
1	tTA prod.	α_1	1.0E-1M/sec	
2	tTA dimer.	k_1	1E9/Msec	[55, 108]
3	tTA_2 undimer.	k_{-1}	10/sec	[55, 108]
4	tTA_2 + Dox bind.	k_2	1.4E5/Msec	[62]
5	tTA_2-Dox unbind.	k_{-2a}	2.2E-2/sec	[62]
6	tTA_2-Dox undimer.	k_{-2b}	10/sec	[55, 108]
7	tTA_2 + tetO bind.	k_3	2.7E8/Msec	[48]
8	tTA_2-tetO unbind.	k_{-3a}	1.5E-3/sec	[48]
9	tTA_2-tetO undimer.	k_{-3b}	10/sec	[55, 108]
10	tTA_2-tetO + Dox bind.	k_4	1.4E5/Msec	[62]
11	tTA decay/dilution	δ_1	2.6E-4/sec	[112][1]
12	tTA_2 decay/dilution	δ_2	2.6E-4/sec	[112]
13	tTA_2-Dox decay	δ_3	2.6E-4/sec	[112]
14	tTA_2-tetO decay	δ_4	2.6E-4/sec	[112]
15	LacI prod.	α_2	3.0E-2/sec	
16	LacI dimer.	k_5	1E9/Msec	[55, 108]
17	$LacI_2$ undimer.	k_{-5}	10/sec	[55, 108]
18	$LacI_2$ dimer.	k_6	1E9/Msec	[55, 108]
19	$LacI_4$ undimer.	k_{-6}	10/sec	[55, 108]
20	$LacI_4$ + LacOA bind.	k_7	2E9/Msec	[25]
21	$LacI_4$-LacOA unbind.	k_{-7}	.04/sec	[25]
22	$LacI_4$ + LacOB bind.	k_8	2E9/Msec	[25]
23	$LacI_4$-LacOB unbind.	k_{-8}	.04/sec	[25]
24	LacI decay/dilution	δ_5	1.28E-4/sec	[93][2]
25	$LacI_2$ decay/dilution	δ_6	1.28E-4/sec	[93]
26	$LacI_4$ decay/dilution	δ_7	1.28E-4/sec	[93]
27	$LacI_4$-LacOA decay	δ_8	1.28E-4/sec	[93]
28	$LacI_4$-LacOB decay	δ_9	1.28E-4/sec	[93]
29	CFP prod.	α_3	5.0E-2/sec	
30	YFP prod.	α_4	5.0E-2/sec	
31	CFP decay	δ_{10}	2.19E-4/sec	[79][2]
32	YFP decay	δ_{11}	2.19E-4/sec	[79]
33	Dox dilution	δ_{12}	9.6E-5/sec	[3]

C.2 Insulator Design

Rate equations for the insulator design presented in Section 4.5.1 are displayed here.

$$\frac{dZ}{dt} = k(t) - \gamma Z - \alpha_4 Z W^* + \alpha_3 [W.Z^*] - k_2 Z + k_{dp} Z^* \tag{C.1}$$

$$\frac{dZ^*}{dt} = -\gamma Z^* - \alpha_1 Z^* W + \alpha_2 [W.Z^*] + k_2 Z - k_{dp} Z^* \tag{C.2}$$

$$\frac{dW}{dt} = -\alpha_1 Z^* W + \alpha_2 [W.Z^*] + k_{dp} W^* + \beta_3 ([W^*.X] + [W^*.X_D] +$$
$$[W^*.X_D^*] + [W^*.X.p] + [W^*.X_D.p] + [W^*.X_D^*.p]) -$$
$$\beta_4 W (X^* + X_D^* + X_D^{**} + [X^*.p] + [X_D^*.p] + [X_D^{**}.p]) \tag{C.3}$$

$$\frac{dW^*}{dt} = -\alpha_4 Z W^* + \alpha_3 [W.Z^*] - k_{dp} W^* + \beta_2 ([W^*.X] +$$
$$[W^*.X_D] + [W^*.X_D^*] + [W^*.X.p] + [W^*.X_D.p] +$$
$$[W^*.X_D^*.p]) - \beta_1 W^* (X + X_D + X_D^* + [X.p] +$$
$$[X_D.p] + [X_D^*.p]) \tag{C.4}$$

$$\frac{d[W.Z^*]}{dt} = \alpha_1 Z^* W - (\alpha_2 + \alpha_3)[W.Z^*] + \alpha_4 Z W^* \tag{C.5}$$

$$\frac{d[W^*.X]}{dt} = \beta_1 W^* X - (\beta_2 + \beta_3)[W^*.X] + \beta_4 W X^* - k_{on}[W^*.X]p +$$
$$k_{off}[W^*.X.p] \tag{C.6}$$

$$\frac{d[W^*.X_D]}{dt} = \beta_1 W^* X_D - (\beta_2 + \beta_3)[W^*.X_D] + \beta_4 W X_D^* - k_{on}[W^*.X_D]p +$$
$$k_{off}[W^*.X_D.p] \tag{C.7}$$

112

$$\frac{d[W^*.X_D^*]}{dt} = \beta_1 W^* X_D^* - (\beta_2 + \beta_3)[W^*.X_D^*] +$$

$$\beta_4 W X_D^{**} - k_{on}[W^*.X_D^*]p + k_{off}[W^*.X_D^*.p] \qquad \text{(C.8)}$$

$$\frac{d[W^*.X.p]}{dt} = \beta_1 W^*[X.p] - (\beta_2 + \beta_3)[W^*.X.p] +$$

$$\beta_4 W[X^*.p] + k_{on}[W^*.X]p - k_{off}[W^*.X.p] \qquad \text{(C.9)}$$

$$\frac{d[W^*.X_D.p]}{dt} = \beta_1 W^*[X_D.p] - (\beta_2 + \beta_3)[W^*.X_D.p] +$$

$$\beta_4 W[X_D^*.p] + k_{on}[W^*.X_D]p - k_{off}[W^*.X_D.p] \qquad \text{(C.10)}$$

$$\frac{d[W^*.X_D^*.p]}{dt} = \beta_1 W^*[X_D^*.p] - (\beta_2 + \beta_3)[W^*.X_D^*.p] +$$

$$\beta_4 W[X_D^{**}.p] + k_{on}[W^*.X_D^*]p - k_{off}[W^*.X_D^*.p] \qquad \text{(C.11)}$$

$$\frac{dX}{dt} = -k_f(2XX + XX^*) + 2k_r X_D + k_r X_D^* - \beta_1 W^* X +$$

$$\beta_2[W^*.X] - k_{on}Xp + k_{off}[X.p] \qquad \text{(C.12)}$$

$$\frac{dX^*}{dt} = -k_f(2X^*X^* + XX^*) + 2k_r X_D^{**} + k_r X_D^* -$$

$$\beta_4 W X^* + \beta_3[W^*.X] - k_{on}X^*p + k_{off}[X^*.p] \qquad \text{(C.13)}$$

$$\frac{dX_D}{dt} = k_f XX - k_r X_D - \beta_1 W^* X_D + \beta_2[W^*.X_D] -$$

$$k_{on}X_Dp + k_{off}[X_D.p] \qquad \text{(C.14)}$$

$$\frac{dX_D^*}{dt} = k_f XX^* - k_r X_D^* - (\beta_1 + \beta_4)W X_D^* +$$

$$\beta_3[W^*.X_D] + \beta_2[W^*.X_D^*] - k_{on}X_D^*p +$$

$$k_{off}[X_D^*.p] \qquad \text{(C.15)}$$

$$\frac{dX_D^{**}}{dt} = k_f X^*X^* - k_r X_D^{**} - \beta_4 W X_D^{**} +$$

$$\beta_3[W^*.X_D^*] - k_{on}X_D^{**}p + k_{off}[X_D^{**}.p] \qquad \text{(C.16)}$$

$$\frac{d[X.p]}{dt} = -\beta_1 W^*[X.p] + \beta_2[W^*.X.p] + k_{on}Xp - k_{off}[X.p] \qquad \text{(C.17)}$$

$$\frac{d[X^*.p]}{dt} = -\beta_4 W[X^*.p] + \beta_3[W^*.X.p] + k_{on}X^*p -$$

$$k_{off}[X^*.p] \qquad \text{(C.18)}$$

$$\frac{d[X_D.p]}{dt} = -\beta_1 W^*[X_D.p] + \beta_2[W^*.X_D.p] + k_{on}X_Dp - k_{off}[X_D.p] \qquad \text{(C.19)}$$

$$\frac{d[X_D^*.p]}{dt} = -(\beta_1 + \beta_4)W[X_D^*.p] + \beta_3[W^*.X_D.p] +$$

$$\beta_2[W^*.X_D^*.p] + k_{on}X_D^*p - k_{off}[X_D^*.p] \qquad \text{(C.20)}$$

113

$$\frac{d[X_D^{**}.p]}{dt} = -\beta_4 W[X_D^{**}.p] + \beta_3[W^*.X_D^*.p] + k_{on}X_D^{**}p - k_{off}[X_D^{**}.p] \qquad (C.21)$$

$$\frac{dp}{dt} = k_{on}p(X + X^* + X_D + X_D^* + X_D^{**} + [W^*.X] + [W^*.X_D] +$$

$$[W^*.X_D^*]) - k_{off}([X.p] + [X^*.p] + [X_D.p] + [X_D^*.p] +$$

$$[X_D^{**}.p] + [W^*.X.p] + [W^*.X_D.p] + [W^*.X_D^*.p]) \qquad (C.22)$$

$$W = W_T - W^* - [W.Z^*] - [W^*.X] - [W^*.X_D] - [W^*.X_D^*] - [W^*.X.p] -$$

$$[W^*.X_D.p] - [W^*.X_D^*.p] \qquad (C.23)$$

$$X = X_T - [W^*.X] - 2[W^*.X_D] - 2[W^*.X_D^*] - [W^*.X.p] - 2[W^*.X_D.p] -$$

$$2[W^*.X_D^*.p] - X^* - 2X_D - 2X_D^* - 2X_D^{**} -$$

$$[X.p] - [X^*.p] - 2[X_D.p] - 2[X_D^*.p] - 2[X_D^{**}.p] \qquad (C.24)$$

$$p = P_T - [W^*.X.p] - [W^*.X_D.p] - [W^*.X_D^*.p] - [X.p] - [X^*.p] -$$

$$[X_D.p] - [X_D^*.p] - [X_D^{**}.p] \qquad (C.25)$$

Appendix D

Potential Protein Engineering

This chapter briefly describes additional ways the proteins used in the *JAK2-STAT5B: HKRR-YPD1-SKN7* signaling cascade can be modified to tune parameters in the insulator.

Instead of using the JAK-STAT system, two Sln1p HKRR subunits could be connected with a flexible linker. This would simplify the system by eliminating the need of an activating protein (Jak). In addition, it is possible that two HKRR subunits constitutively linked together would result in higher levels of auto-phosphorylation than when attached to Stat5b. Lastly, a direct linkage without Stat5b would ensure that HKRRs stay in the cytoplasm. As reviewed in Section 4.3.2, Stat5b is able to shuttle between the cytoplasm and nucleus even when it is not phosphorylated. This might not be a problem, since Ypd1p also shuttles between the cytoplasm and nucleus. However, it is unclear if Ydp1p needs to be phosphorylated to be transported into the nucleus. As shown in Figure 4-7, Stat5-HKRRp does activate SSRE promoters, but activation might be higher if the HKRRs were contained within the cytoplasm. A possible linker could be the 21 amino acids from human methionine synthase which allow flexible domain movement [31].

Degradation tags could also be attached to Stat5b-HKRRp and fluorescent proteins to allow the system to better track the input. Three potential degradation tags are described here. One method is the attachment of a PEST-rich C-terminal residue. Fusion of 178 C-terminal residues from *Cln2*, a G_1 cyclin, to the C-terminus of *GFP*

reduces the half life of *GFP* from 7 hours to 30 minutes [79].

Another method is the attachment of a N-degron that utilizes the N-end rule degradation pathway [46]. Particular amino acids (D, E, F, H, I, K, L, N, Q, R, W, and Y) exposed at the N-terminus of proteins have been found to induce polyubiquitination followed by proteasomal proteolysis [115]. Cyan fluorescent proteins (CFPs) with shorter half-lives were constructed with this method [46]. Ubi-M-Δk-CFP, Ubi-E-Δk-CFP, and Ubi-Y-Δk-CFP were found to have half-lives of 75, 50, and 5 mins, respectively. In this degradation process, de-ubiquinating enzymes first cleave off the ubiquitin (Ubi) tag, which exposes the N-degron. Ubr1p, the N-end rule pathway ubiquitin ligase, then binds to the N-degron. The N-degron requires not only a destabilizing N-terminus amino acid but also lysine residues. A 40 amino acid sequence with a KRK motif from *E. coli* lac repressor (called *ek*) was found to function well as an N-degron signal. In the study, the authors found that CFP tagged with *ek* resulted in steady state levels too low to study. Thus, a 24 residue truncated version of *ek* called *δk*, a less efficient N-degron signal, was used. The attachment of Ubi-E-Δk or Ubi-Y-Δk to Stat-HKRR and fluorescent proteins could improve the measurements collected from the circuit.

Lastly, a third method is to use a bidirectional degron that is exposed only after cleavage by a tobacco etch virus (TEV) protease [61]. The bidirectional degron consists of a C-degron and a N-degron separated by a TEV protease cleavage site. The C-degron is from mouse ornithine decarboxylase which results in proteasomal polyubiquitination-independent proteolysis. The N-degron ends with a destabilizing phenylalanine (F) residue followed by an affinity domain that binds strongly to human protein p14 to enhance protease cleavage (the TEV protease used is fused to p14). TEV protease expression exposes both degrons, destabilizing both proteins at either side of the bidirectional degron. Without expression of TEV protease, degrons remain inactive. Such a bidirectional degron could be placed in between two Sln1p HKRR to result in its undimerization and rapid proteolysis after TEV protease expression. But since this degradation system involves the additional expression of TEV protease inside the cell, it is not as simple as the other proposed degradation tags.

Appendix E

Materials and Methods

E.1 Strains and Culture Conditions

E.1.1 Bacterial Strains and Culture Conditions

E. cloni © 10G Chemically Competent Cells (*F⁻ mcrA δ(mrr-hsdRMS-mcrBC) endA1 recA1 φ80dlacZδM15 δlacX74 araD139 δ(ara,leu)7697 galU galK rpsL nupGλ tonA*) (Lucigen Corporation, Middleton, WI) was used to clone and propagate all plasmids without *ccd*B. DB3.1 (*F⁻ gyrA462 endA1 δ(sr1-recA) mcrB mrr hsdS20(r⁻_B, m⁻_B) supE44 ara14 galK2 lacY1 proA2 rpsL20(Sm^r) xyl5 δleu mtl1*) was used to clone and propagate plasmids with *ccd*B, namely destination vectors (Invitrogen, Carlsbad, CA). *gyrA462* enables the strain to clone and propagate *ccd*B containing plasmids. All cells were grown in LB broth (Difco, Detroit, MI), with the addition of the following antibiotics (Simga-Aldrich, St. Louis, MO): 100 μg/mL ampicillin for cells harboring expression clones, position clones, or carrier vectors; 50 μg/mL kanamycin for cells harboring entry clones; 50 μg/mL chloramphenicol and 100 μg/mL ampicillin for cells harboring destination vectors; and 100 μg/mL ampicillin and 50 μg/mL kanamycin for cells harboring multi-promoter-CDS expression clones from Gibson reactions. Cultures were grown in a 37 °C shaking incubator at 300 rpm. Plasmids were introduced into cells through chemical transformation. Cells were incubated on ice for 30 mins, heatshocked at 42 °C for 30 sec, incubated on ice for 2

mins, resuspended in SOC, and incubated at 37 °C for 1 hr.

E.1.2   Yeast Strains and Culture Conditions

Yeast W303-1A ADE2+ (MATa *leu2-3,112 trp1-1 can1-100 ura3-1 his3-11,15*)
[111] was transformed with plasmids for expression studies. Plasmids were introduced
into yeast cells through LiAc/SS Carrier DNA/PEG transformation procedure [43].
Cells were grown at 30 °C in SD-His, SD-Leu, SD-Trp, or SD-Ura media, depending
on the auxotrophic marker on the plasmid harbored.

E.2   Plasmid Construction

Plasmids were constructed using standard molecular biology techniques [104].
PCR amplification with synthetic primers was used to create novel entry clones.
Primers and templates are listed in Figure E.1. Primers were designed with analy-
sis by Primer3Plus [114] and were ordered from Integrated DNA Technologies, Inc
(San Diego, CA). Phusion DNA polymerase (NEB, Ipswich, MA) was used for PCR
amplification from synthetic primers, and DpnI (NEB) was subsequently used for
template degradation. Restriction enzymes and buffers were acquired from NEB.
T4 ligase (NEB) with 2x Rapid Ligation Buffer (Promega, Madison, Wisconsin) was
used for ligations. QIAquick PCR Purification Kits and QIAEX II Gel Extraction
Kits (Qiagen, Hilden, Germany) were used for DNA purification. Plasmid DNA was
extracted with QIAprep Spin Miniprep Kits (Qiagen). Sequencing of DNA was done
by GeneWiz Inc. (New Brunswick, NJ)

For Gateway BP reactions, 1 uL of BP Clonase ^{TM} (Invitrogen) was mixed with
10 fmol of PCR amplified fragment and 10 fmol of pDONR221 and filled to a total
volume of 5 uL with TE buffer. For Gateway LR reactions, 1 uL of LR Clonase ^{TM}
(Invitrogen) was mixed with 10 fmol of each entry clone and 10 fmol of destination
or position vector and filled to a total volume of 5 uL with TE buffer. For Gibson
reactions, 10 fmol of each digested DNA fragment filled to 5 uL H_2O was mixed with
the 15 ul Gibson reaction mix [41] and incubated at 50 °C for 1 hour. Reagents
for Gibson reactions were obtains from the following companies: dNTP (Bioline,

Taunton, MA), T5 Exonuclease (Epicentre Biotechnologies, Madison, WI), Phusion DNA polymerase (NEB), NAD+ (NEB), Taq DNA ligase (NEB), PEG-800 (Sigma-Aldrich), and DTT (Sigma-Aldrich).

E.3 Flow Cytometry Measurements

BD FACS LSR II (Franklin Lakes, NJ) in the Flow Cytometry Core at the MIT Koch Institute was used to measure the fluorescence of individual cells. In the PE Texas Red channel, a 561 nm excitation laser and a 590-630 nm emission filter was used. In the FITC channel, a 488 nm excitation laser and a 500-560 nm emission filter was used. 10,000 events were collected for each sample in BD FACSDIVA software, and Flowjo (Tree Star, Ashland, OR) was used to analyze the data.

Table E.1: PCR Primers and Templates Used for Entry Clone Construction

Entry	Primer ID	Primer Sequence	Parent Template	Cloning Method
TEF	DMp026	AGTCTCGAGACCACACCTCTACCGGCAGATCC	pKT175	XhoI/EcoR1
	DMp027	CCCGAATTCCGCTGTGCTACTGGTGAGGCA		
CMV	ALp005	AGTCTCGAGGAGCTTGGCCCATTGCATAC	pCM185	XhoI/EcoR1
	ALp007	CCCGAATTCATGAATTAATTCGGGCCGCGGAGGCTG		
ADH1	ALp012	AGTCTCGAGAGCAGTTAAGCGTATTACTGAAAGTT	p413-ADH1-EGFP	XhoI/EcoR1
	ALp013	CCCGAATTCTTTAATACCCGGGGATCCTCTAGAGTCG		
ADH1/lacO	ALp012	AGTCTCGAGAGCAGTTAAGCGTATTACTGAAAGTT	p413-ADHi-EGFP	XhoI/EcoR1
	ALp013	CCCGAATTCTTTAATACCCGGGGATCCTCTAGAGTCG		
tetO7CYC1	ALp008	AGTCTCGAGCTACCGGCAGATCAATTCCT	pCM185	XhoI/EcoR1
	ALp009	CCCGAATTCCGAATTGATCCGGTAATTTAGT		
SSRE	ALp041	CCAACCAAGGTCTCCTCGAGGGCCAGTGCCAAGCTTTCTTTT	p413-SSRE	BsaI → XhoI/EcoR1
	ALp042	CCAACCAAGGTCTCGAATTCCCCGTCGTTGCTTTTATTACCG		
TR-SSRE	ALp041	CCAACCAAGGTCTCCTCGAGGGCCAGTGCCAAGCTTTCTTTT	p413-TR-SSRE-yEGFP3	BsaI → XhoI/EcoR1
	ALp042	CCAACCAAGGTCTCGAATTCCCCGTCGTTGCTTTTATTACCG		
mLacI	ALp010b	GGGGACAAGTTTGTACAAAAAAGCAGGCTGAAAAAAAAATGAAACCAGTAACGTTATACGA	p412-CYC1-mLacI	BP Clonase
	ALp011	GGGGACCACTTTGTACAAGAAAGCTGGGTATCACACCTTTCGCTTCTTCTT		
tTA	ALp003b	GGGGACAAGTTTGTACAAAAAAGCAGGCTGAAAAAAAAATGTCTAGATTAGATAAAAGTAA	pCM185	BP Clonase
	ALp004	GGGGACCACTTTGTACAAGAAAGCTGGGTACTTATTACGATCCCCCACCGTA		
rtTA	ALp001b	GGGGACAAGTTTGTACAAAAAAGCAGGCTGAAAAAAAAATGTCTAGATTAGATAAAAGTAA	p252-mSTAT5b	BP Clonase
	ALp002	GGGGACCACTTTGTACAAGAAAGCTGGGTACTACCCACCGTACTCGTCAAT		
mJak2	ALp035	GGGGACAAGTTTGTACAAAAAAGCAGGCTGAAAAAAAAATGGGAATGGCCTGCCTTA	p416-mJak2	BP Clonase
	ALp036	GGGGACCACTTTGTACAAGAAAGCTGGGTACTATACTGTCCCGCATTT		
mStat5b -HKRR	ALp037	GGGGACAAGTTTGTACAAAAAAGCAGGCTGAAAAAAAAATGGCTATGTGGATACAGG	p252-mSTAT5b-HKRR	BP Clonase
	ALp038	GGGGACCACTTTGTACAAGAAAGCTGGGTATCATTTGTTATTTTTCTTTCC		
kanMX4	ALp053	GGGGACAAGTTTGTACAAAAAAGCAGGCTGAAAAAAAAATGGGTAAGGAAAAGACTCACG	pYM-N1	BP Clonase
	ALp054	GGGGACCACTTTGTACAAGAAAGCTGGGTATTAGAAAAAACTCATCGAGCATCA		
Gateway Dest. Cassette	ALp033	CCAACCAAGAGCTCACCGTATCGATGTCGACGTTA	pDONR221	SacI/XhoI
	ALp034	CCAACCAACTCGAGTACCACTTTGTACAAGAAAGCTGAA		
Gibson Carrier Cassette	ALp028	CCAACCAACCAGAGCTCTGGAGGTTTACCGAGCTCTTATTGG	pZDonor_Seq1-PacI-SeqX	SacI/XhoI
	ALp029	CCAACCAACTCGAGGGGTTAGGCGACTGTTATAACTTACC		

Appendix F

Additional Plots

All plots generated from the parameter sensitivity study of the insulator modeled with two-step phosphorylation - dephosphorylation reaction in Section 3.2.2 and the study of proposed synthetic insulator in Section 4.5.2 are displayed in this chapter.
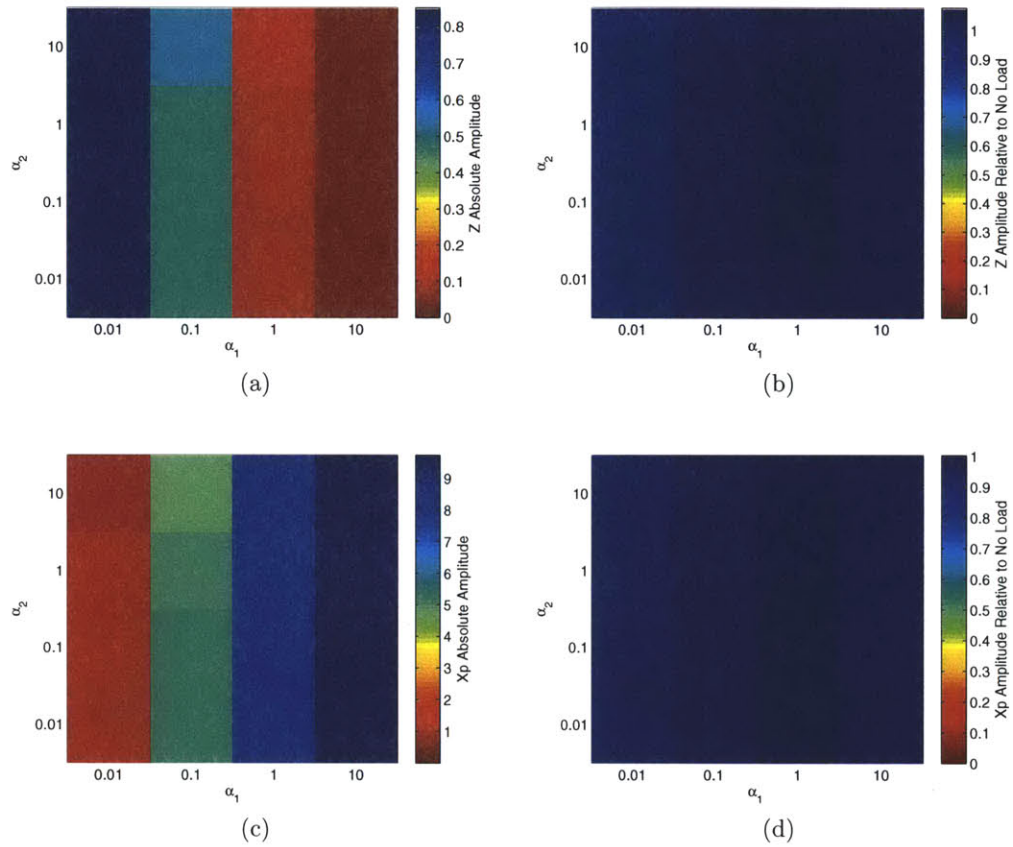
F.1   Insulator Simulations

Figure F-1: Two-step phosphorylation - dephosphorylation reactions, α_1 and α_2 varied, without Y and X decay.
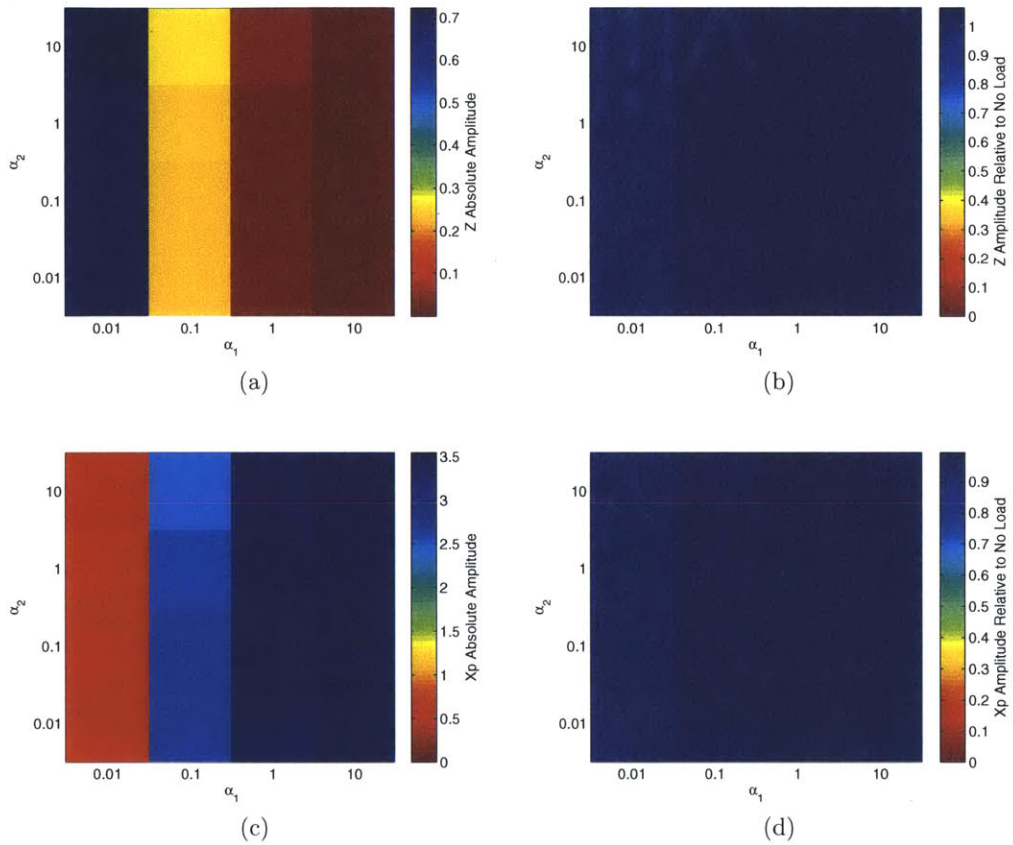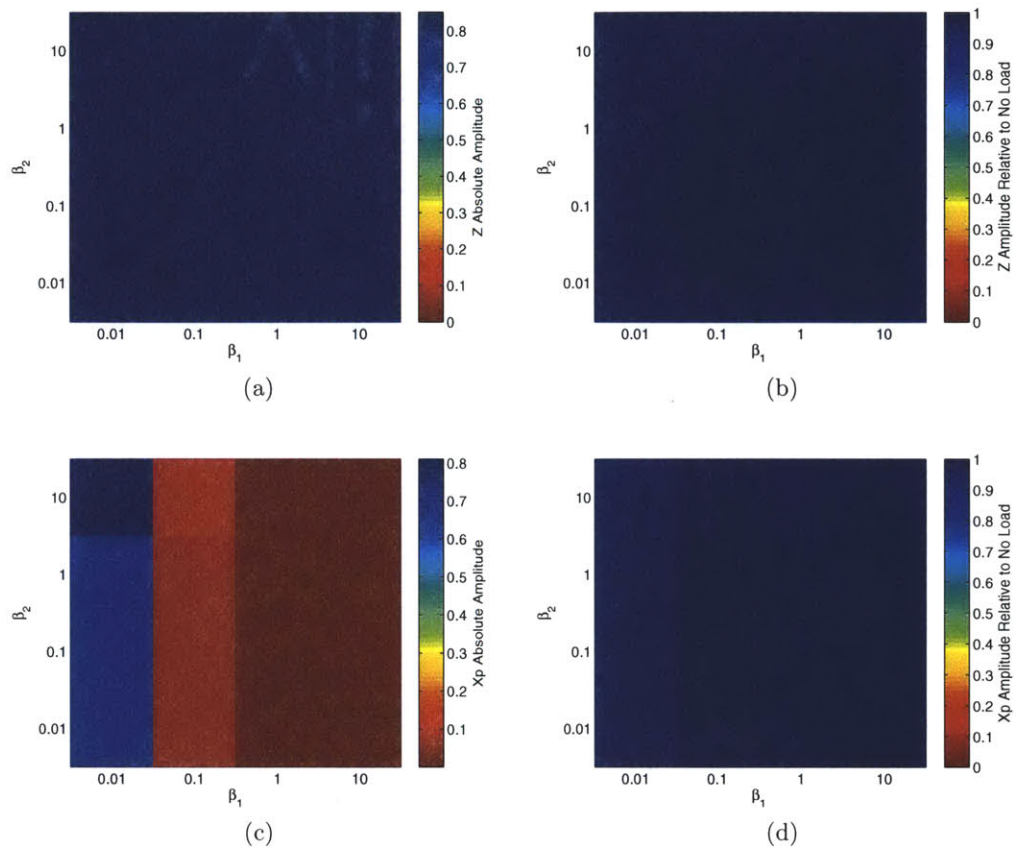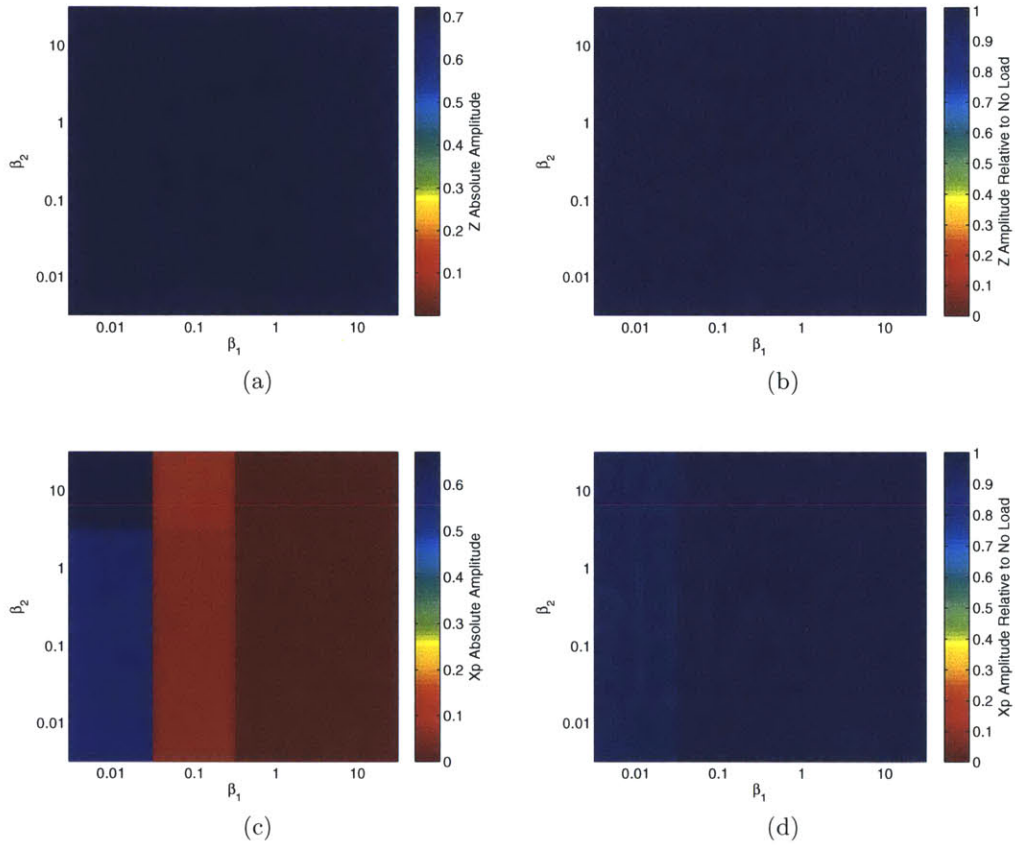
Figure F-2: Two-step phosphorylation - dephosphorylation reactions, α_1 and α_2 varied, with Y and X decay.

Figure F-3: Two-step phosphorylation - dephosphorylation reactions, β_1 and β_2 varied, without Y and X decay.
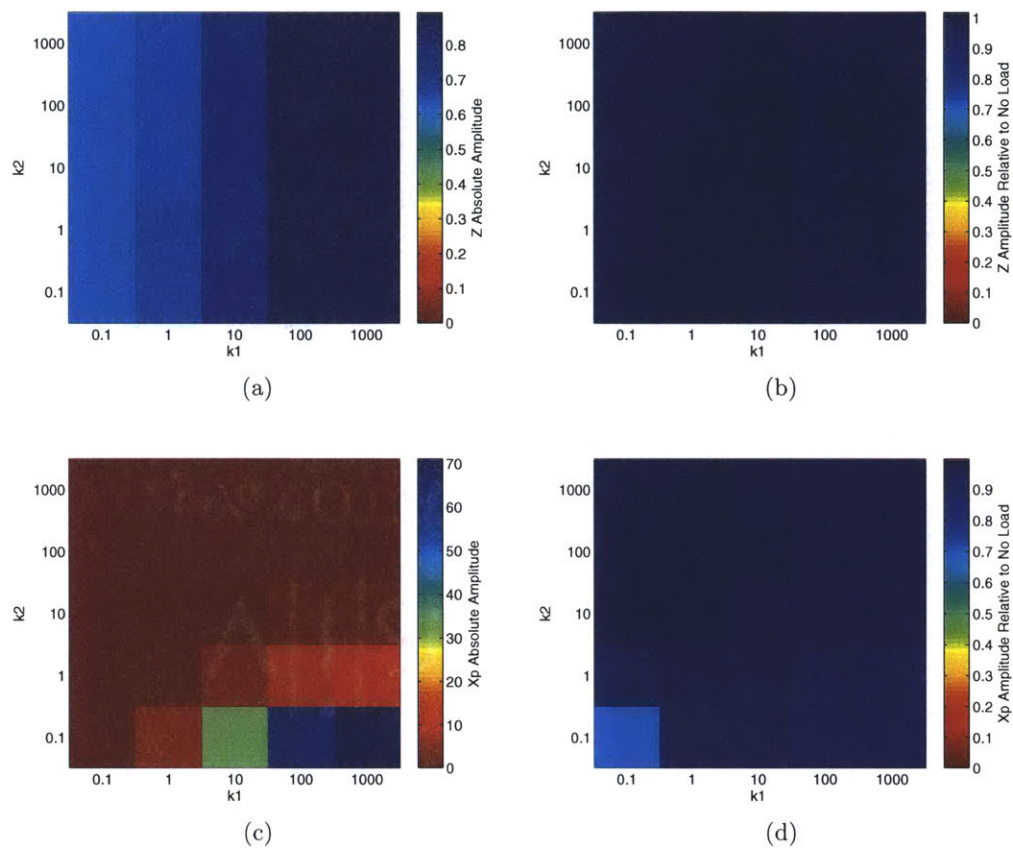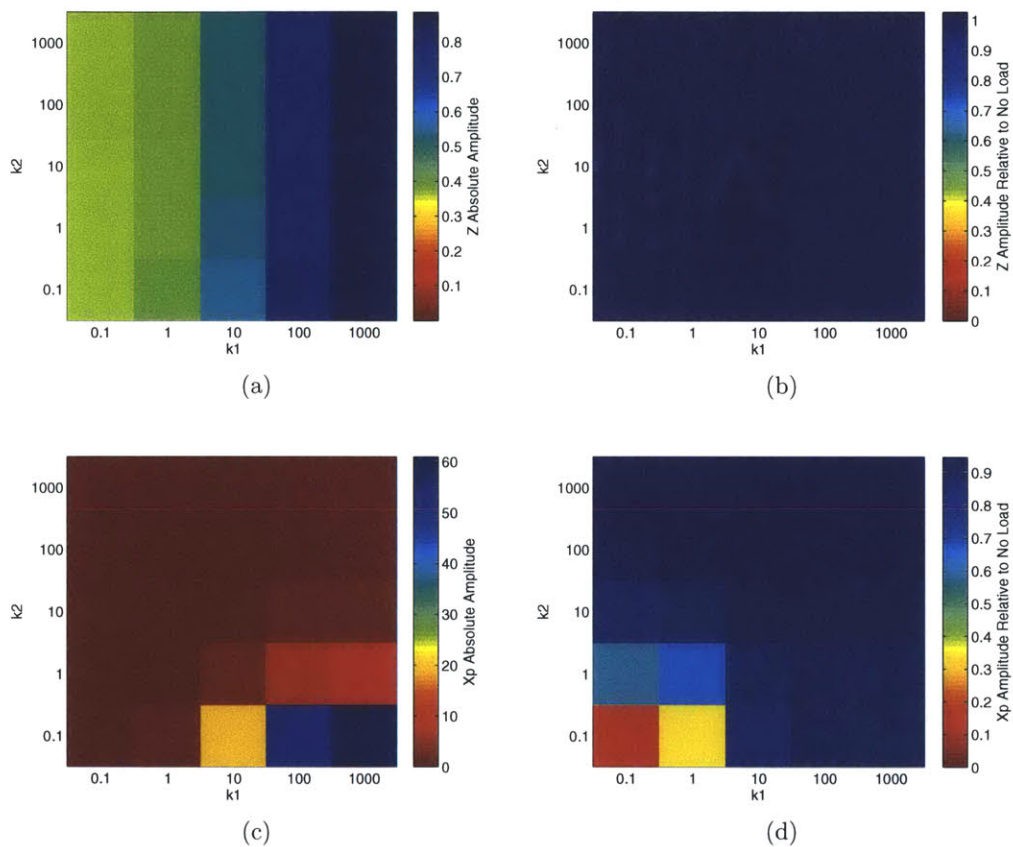
Figure F-4: Two-step phosphorylation - dephosphorylation reactions, β_1 and β_2 varied, with Y and X decay.

Figure F-5: Two-step phosphorylation - dephosphorylation reactions, k_1 and k_2 varied, without Y and X decay.

Figure F-6: Two-step phosphorylation - dephosphorylation reactions, k_1 and k_2 varied, with Y and X decay.
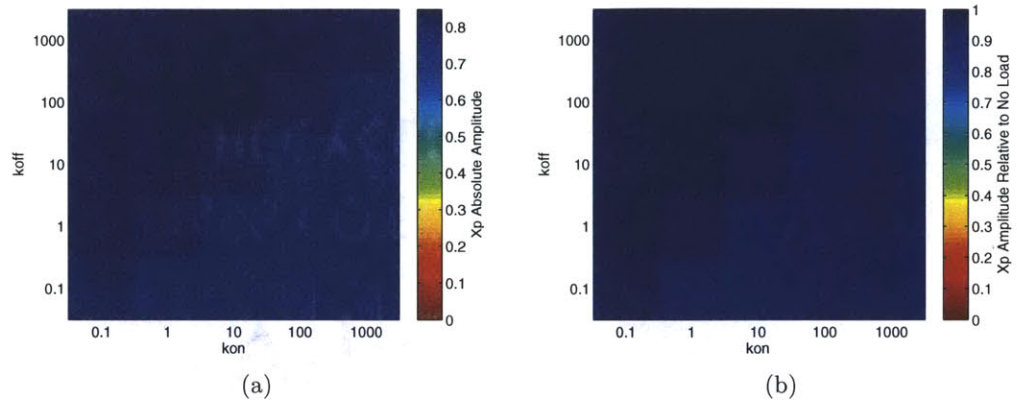
Figure F-7: Two-step phosphorylation - dephosphorylation reactions, k_{on} and k_{off} varied, without Y and X decay.
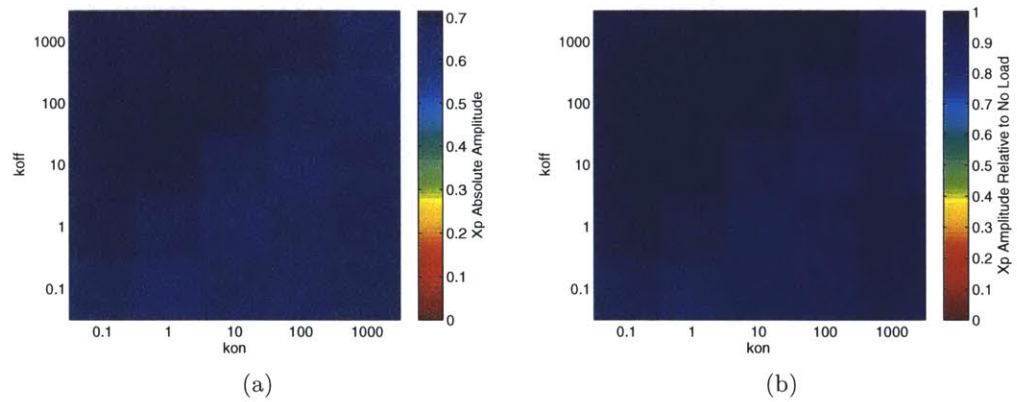

Figure F-8: Two-step phosphorylation - dephosphorylation reactions, k_{on} and k_{off} varied, with Y and X decay.
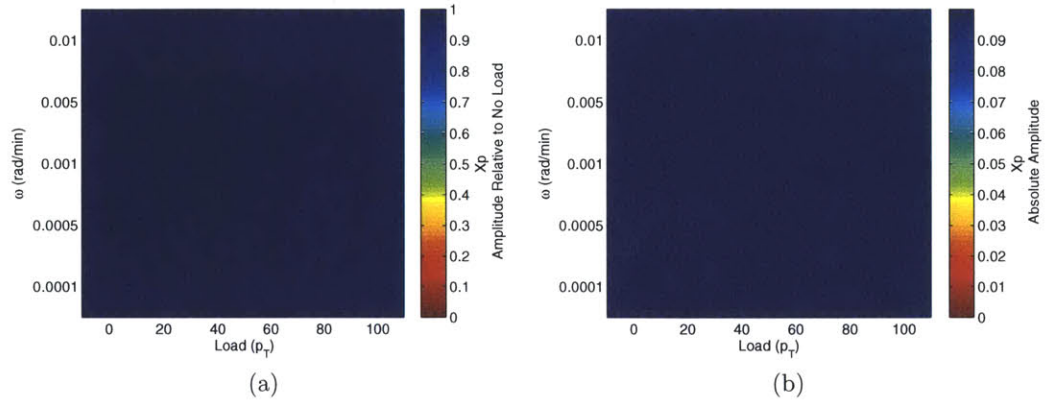
Figure F-9: Two-step phosphorylation - dephosphorylation reactions, load and ω varied, without Y and X decay.
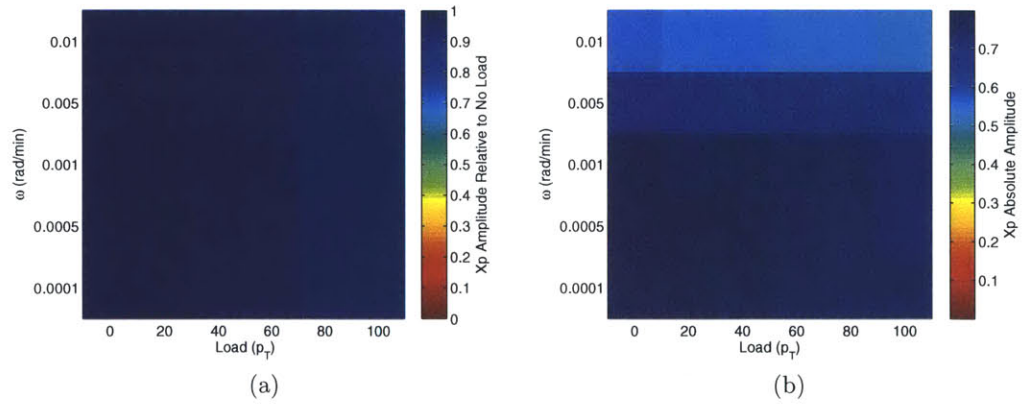

Figure F-10: Two-step phosphorylation - dephosphorylation reactions, load and ω varied, with Y and X decay.
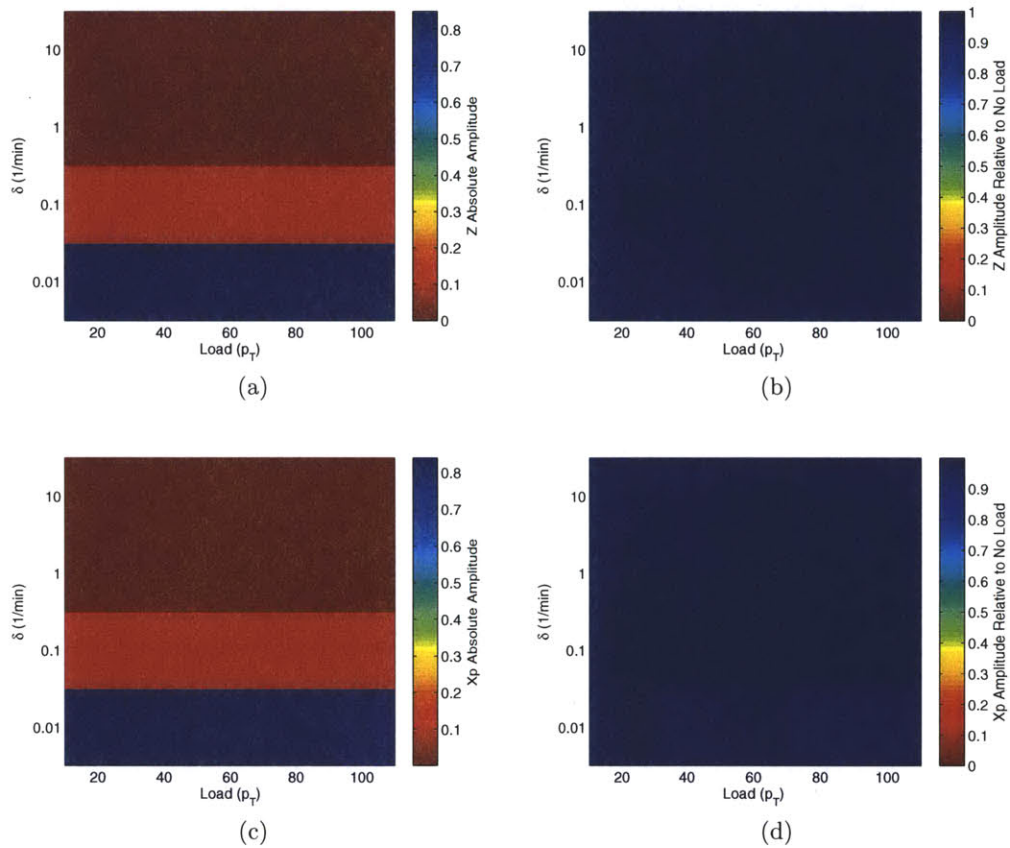
Figure F-11: Two-step phosphorylation - dephosphorylation reactions, δ and load varied, without Y and X decay.
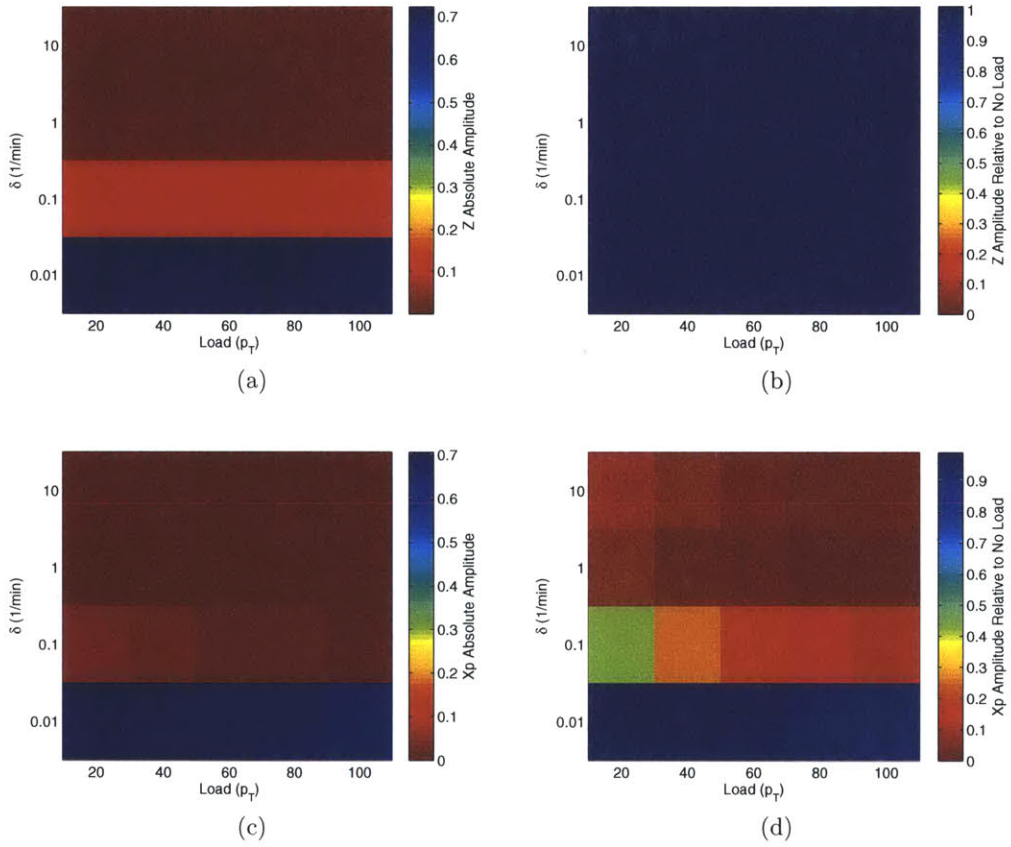
Figure F-12: Two-step phosphorylation - dephosphorylation reactions, δ and load varied, with Y and X decay.
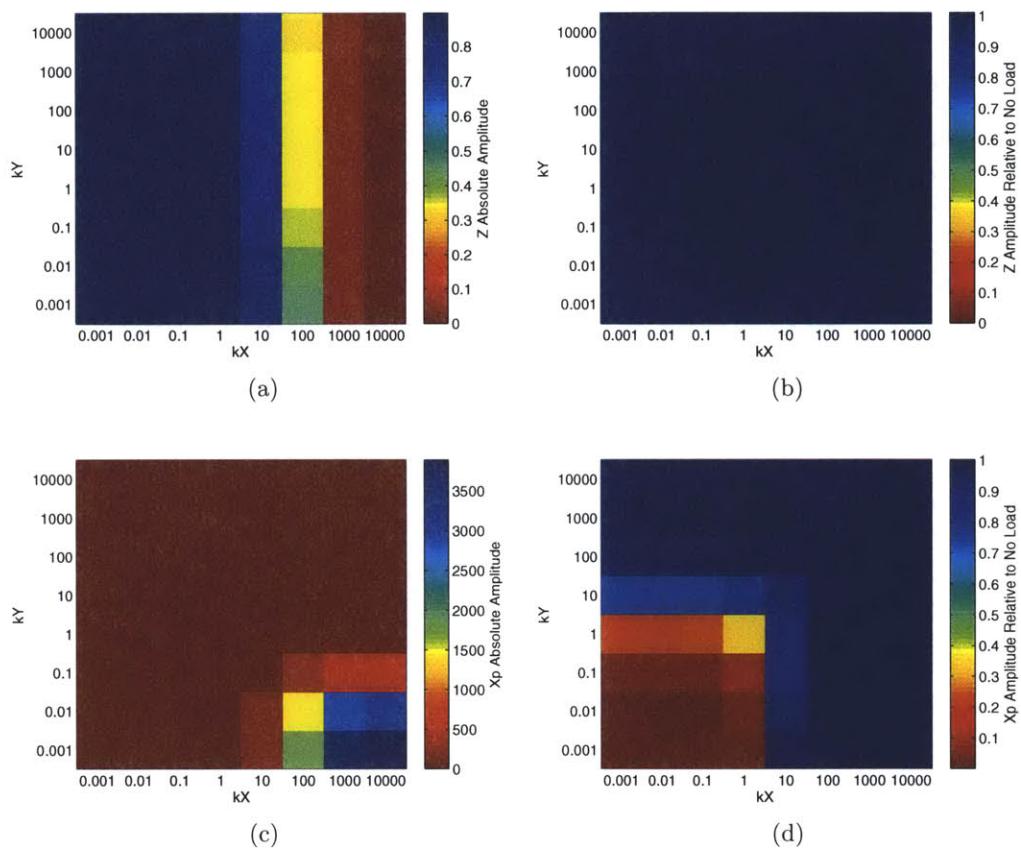
Figure F-13: Two-step phosphorylation - dephosphorylation reactions, kX and kY varied, with Y and X decay.
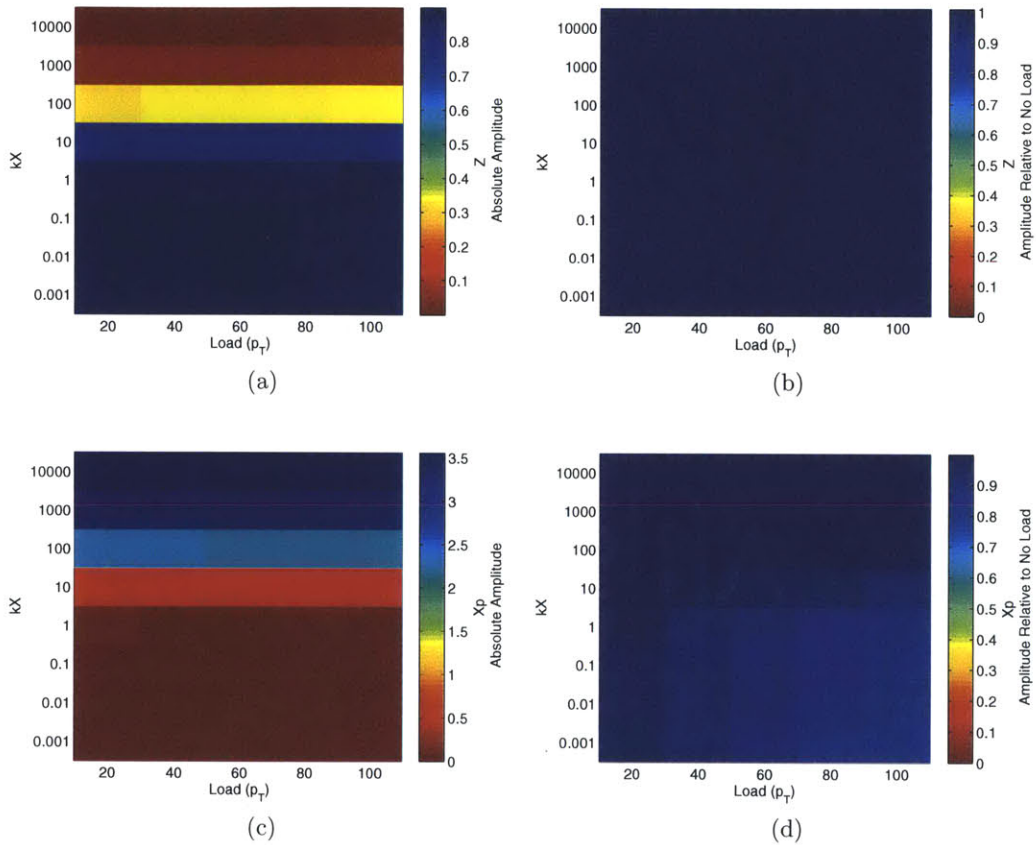
Figure F-14: Two-step phosphorylation - dephosphorylation reactions, kX and load varied, with Y and X decay.

F.2   Full Insulator Simulations

Figure F-15: Synthetic insulator, X_T and W_T varied, absolute and relative amplitude to no load.
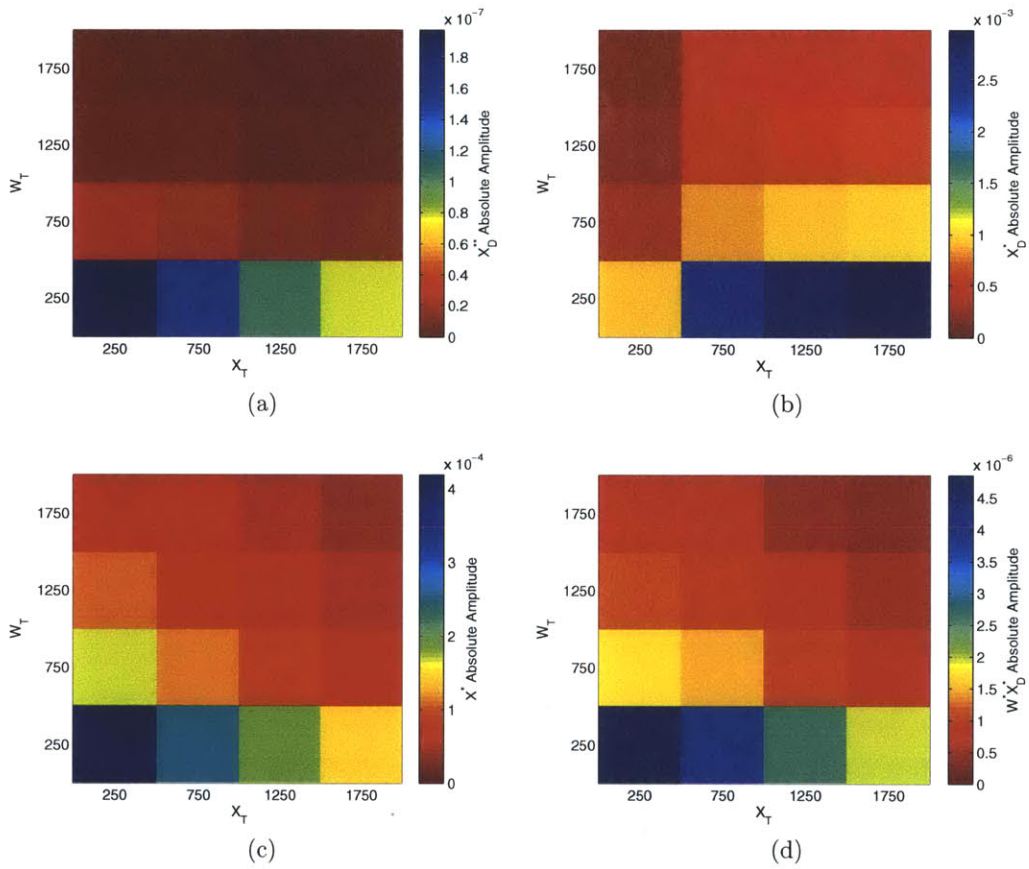
Figure F-16: Synthetic insulator, X_T and W_T varied, individual species, absolute amplitude.

135
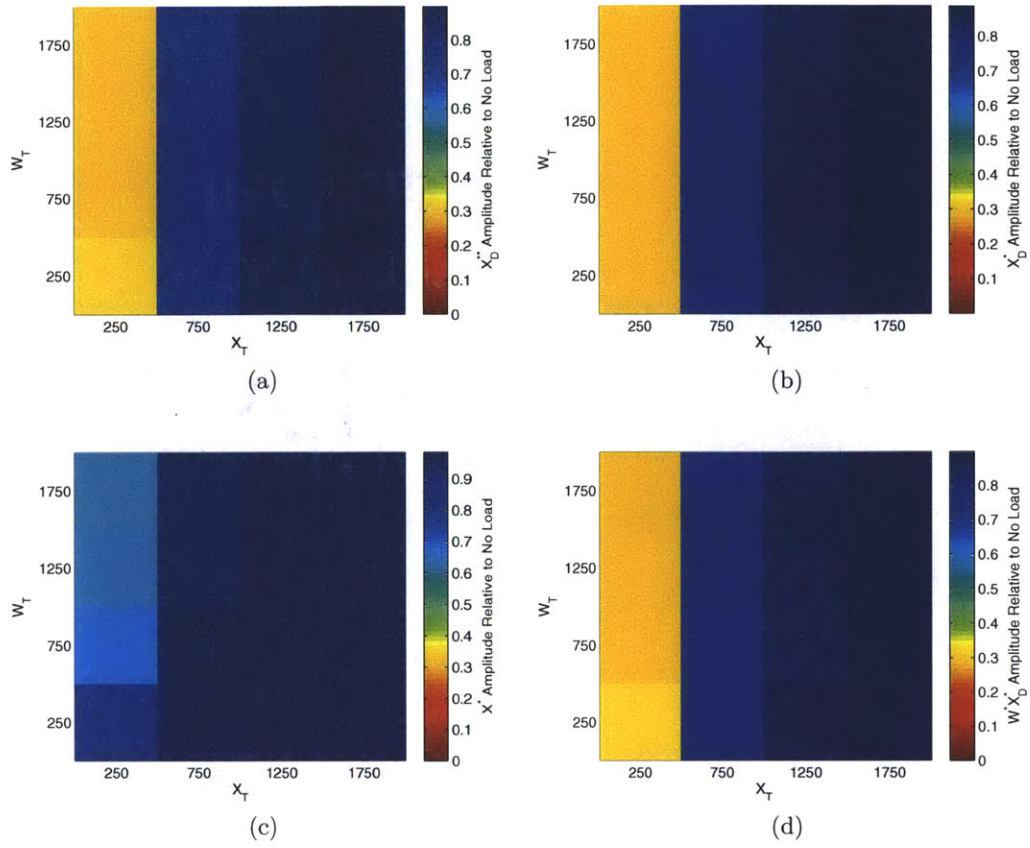
Figure F-17: Synthetic insulator, X_T and W_T varied, individual species, relative amplitude to no load.
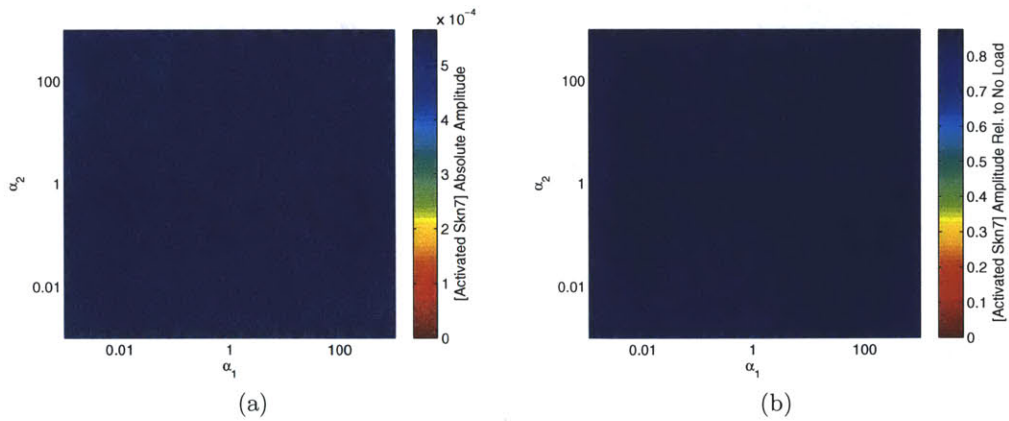
Figure F-18: Synthetic insulator, a_1 and a_2 varied, absolute amplitude and relative amplitude to no load.
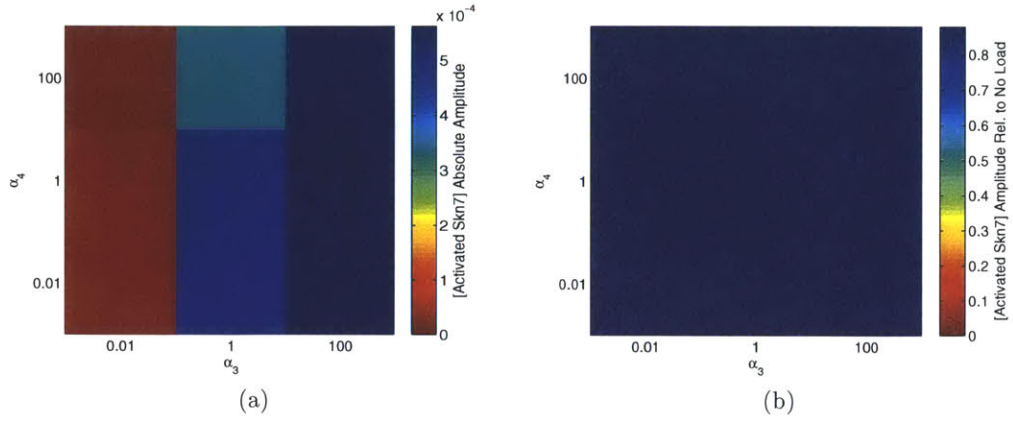
Figure F-19: Synthetic insulator, a_3 and a_4 varied, absolute amplitude and relative amplitude to no load.
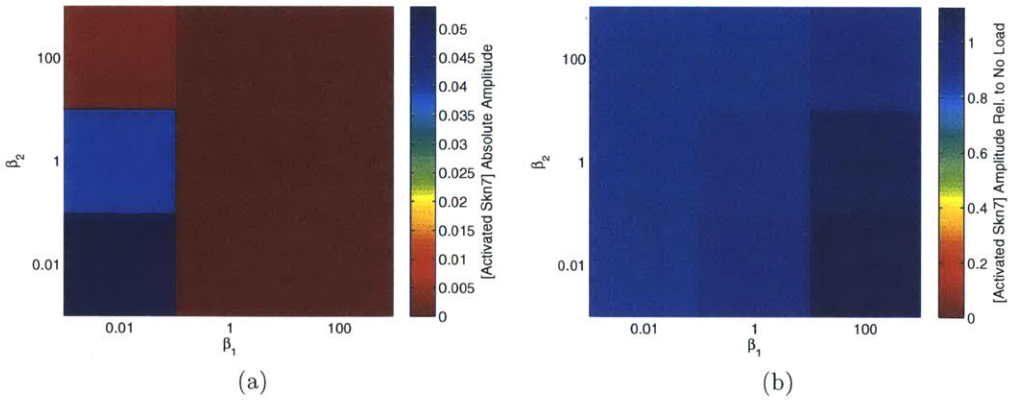
Figure F-20: Synthetic insulator, b_1 and b_2 varied, absolute amplitude and relative amplitude to no load.
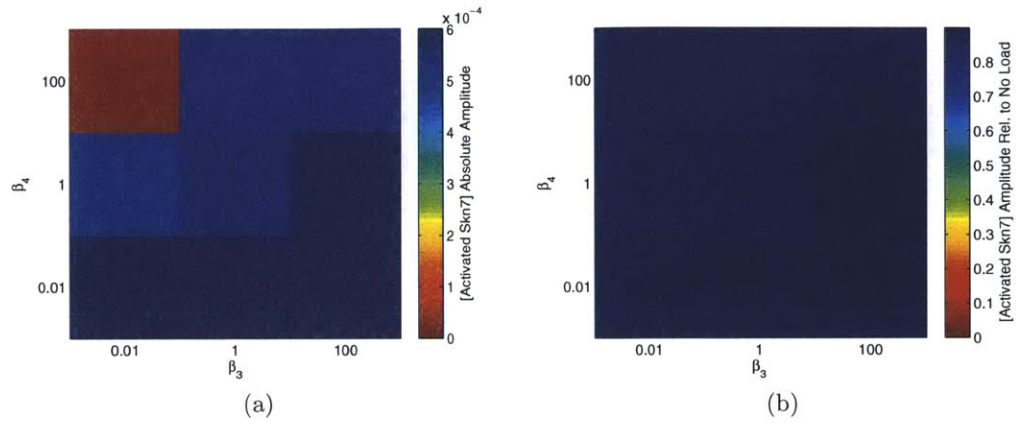
Figure F-21: Synthetic insulator, b_3 and b_4 varied, absolute amplitude and relative amplitude to no load.

Figure F-22: Synthetic insulator, k_{on} and k_{off} varied, absolute amplitude and relative amplitude to no load.
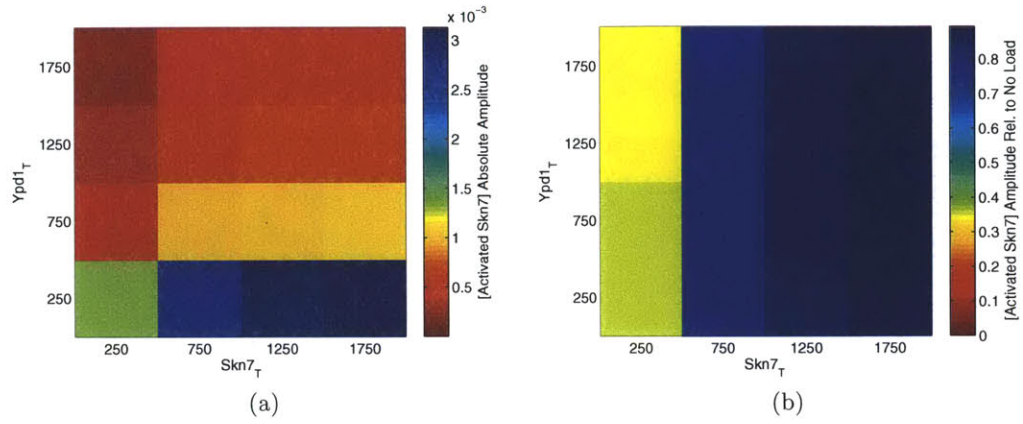
Figure F-23: Synthetic insulator, k_f and k_r varied, absolute amplitude and relative amplitude to no load.

Figure F-24: Synthetic insulator, k_2 and k_{dp} varied, absolute amplitude and relative amplitude to no load.

(a)   (b)
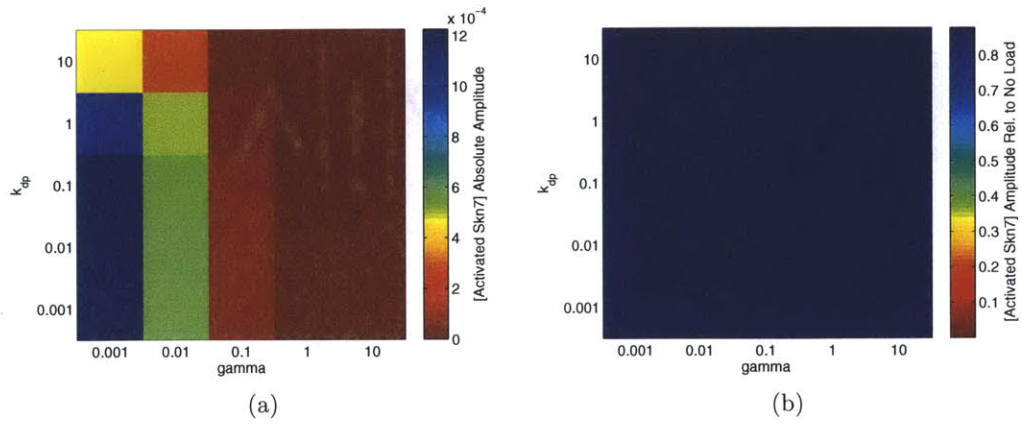
Figure F-25: Synthetic insulator, γ and k_{dp} varied, absolute amplitude and relative amplitude to no load.
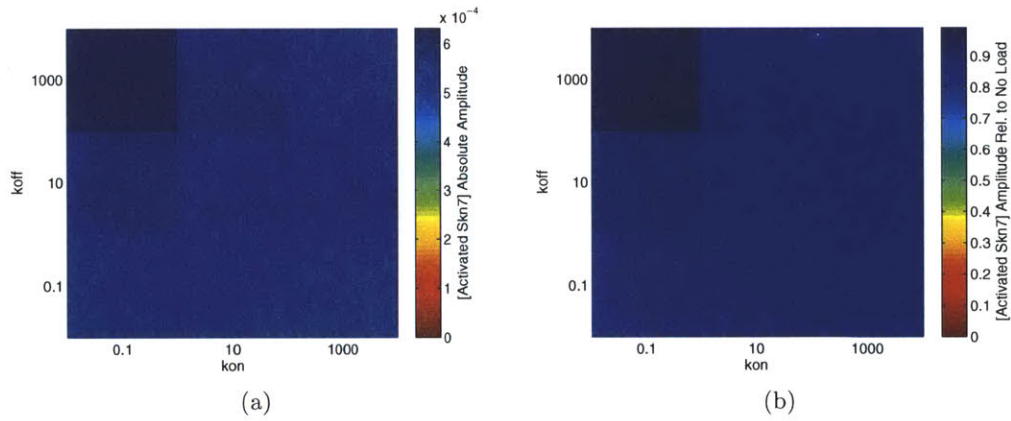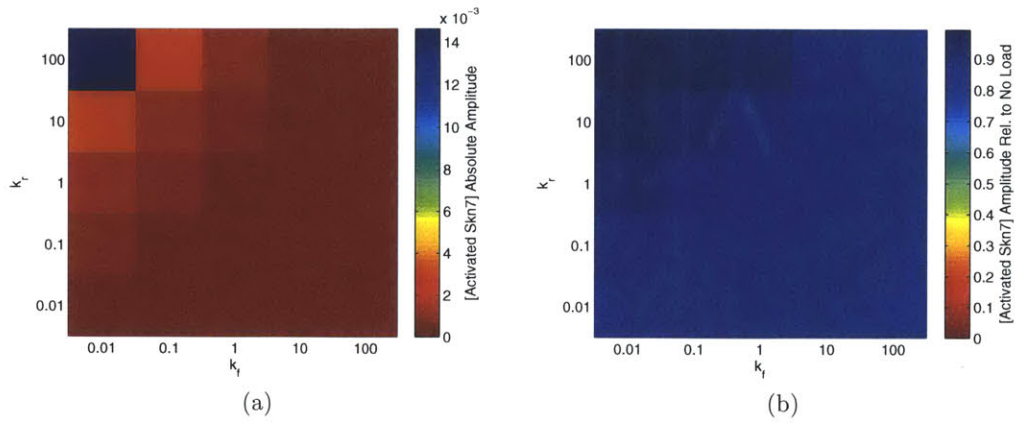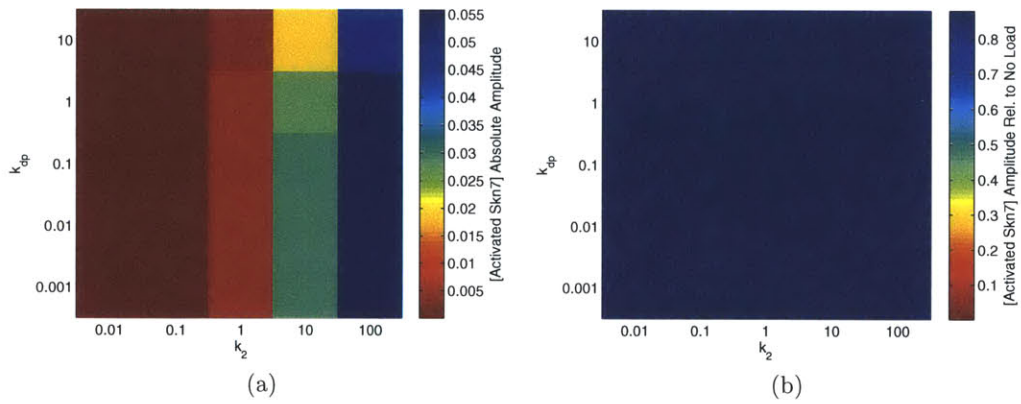

(a)   (b)
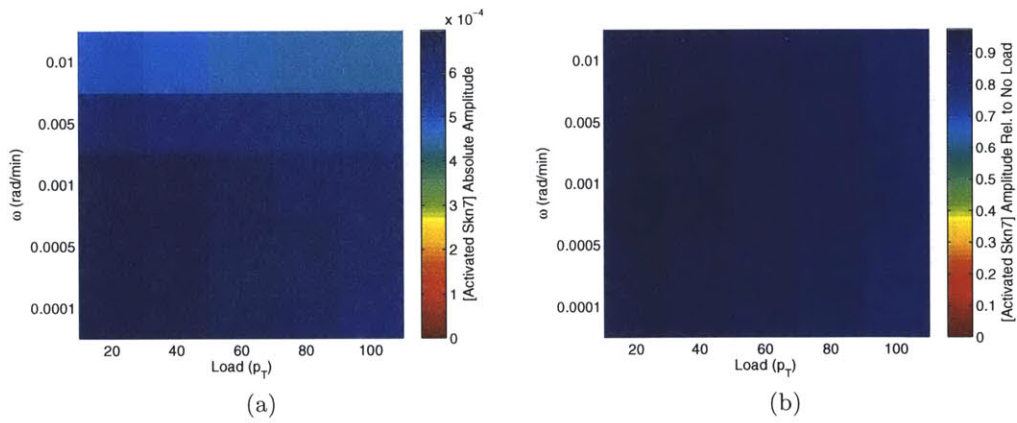
Figure F-26: Synthetic insulator, ω and p_T varied, absolute amplitude and relative amplitude to no load.

Bibliography

[1] David S. Aaronson and Curt M. Horvath. A road map for those who don't know jak-stat. *Science*, 296(5573):1653–1655, May 2002.

[2] Simon Alberti, Aaron D. Gitler, and Susan Lindquist. A suite of gateway cloning vectors for high-throughput genetic analysis in saccharomyces cerevisiae. *Yeast*, 24(10):913–919, October 2007.

[3] Arthur S. Alberts, Nicolas Bouquin, Leland H. Johnston, and Richard Treisman. Analysis of rhoa-binding proteins reveals an interaction domain conserved in heterotrimeric g protein beta subunits and the yeast response regulator protein skn7. *J. Biol. Chem.*, 273(15):8616–8622, April 1998.

[4] Ivan I Atanassov, Ilian I Atanassov, J Peter Etchells, and Simon R Turner. A simple, flexible and efficient pcr-fusion/gateway cloning procedure for gene fusion, site-directed mutagenesis, short sequence insertion and domain deletions and swaps. *Plant Methods*, 5(14), October 2009.

[5] Mariette R. Atkinson, Michael A. Savageau, Jesse T. Myers, and Alexander J. Ninfa. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in escherichia coli. *Cell*, 113(5):597–607, May 2003.

[6] Fariba Barahmand-Pour, Andreas Meinke, Bernd Groner, and Thomas Decker. Jak2-stat5 interactions analyzed in yeast. *J. Biol. Chem.*, 273(20):12567–12575, May 1998.

[7] Subhayu Basu, Yoram Gerchman, Cynthia H. Collins, Frances H. Arnold, and Ron Weiss. A synthetic multicellular system for programmed pattern formation. *Nature*, 434(7037):1130–1134, April 2005.

[8] Subhayu Basu, Rishabh Mehreja, Stephan Thiberge, Ming-Tang Chen, and Ron Weiss. Spatiotemporal control of gene expression with pulse-generating networks. *Proc. Natl. Acad. Sci. USA*, 101(17):6355–6360, April 2004.

[9] Travis S Bayer and Christina D Smolke. Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nat. Biotechnol.*, 23(3):337–343, March 2005.

[10] Susanne Berchtold, Richard Moriggl, Fabrice Gouilleux, Olli Silvennoinen, Christian Beisenherz, Edith Pfitzner, Manuela Wisslera, Elisabeth Stoklin, and Bernd Groner. Cytokine receptor-independent, constitutively active variants of stat5. *J. Biol. Chem.*, 272(48):30237–30243, November 1997.

[11] Philippe Bernard and Martine Couturier. Cell killing by the f plasmid ccdb protein involves poisoning of dna-topoisomerase ii complexes. *J. Mol. Biol.*, 226(3):735–745, August 1192.

[12] Nicolas Bouquin, Anthony L. Johnson, Brian A. Morgan, and Leland H. Johnston. Association of the cell cycle transcription factor mbp1 with the skn7 response regulator in budding yeast. *Mol. Biol. Cell*, 10(10):3389–3400, October 1999.

[13] Nicolas Bouquin, Anthony L. Johnson, Brian A. Morgan, and Leland H. Johnston. Association of the cell cycle transcription factor mbp1 with the skn7 response regulator in budding yeast. *Mol. Biol. Cell*, 10(10):3389–3400, October 1999.

[14] Jeffrey L. Brown and Howard Bussey. The yeast kre9 gene encodes an o glycoprotein involved in cell surface beta-glucan assembly. *Mol. Cell. Biol.*, 13(10):6346–6356, October 1993.

[15] Jeffrey L. Brown, Stan North, and Howard Bussey. Skn7, a yeast multicopy suppressor of a mutation affecting cell wall beta-glucan assembly, encodes a product with domains homologous to prokaryotic two-component regulators and to heat shock transcription factors. *J. Bacteriol.*, 175(21):6908–6915, November 1993.

[16] Carla Buchman, Petra Skroch, Juliet Welch, Seymour Fogel, and Michael Karin. The cup2 gene product, regulator of yeast metallothionein expression, is a copper-activated dna-binding protein. *Mol. Cell. Biol.*, 9(9):4091–4095, September 1989.

[17] Thomas Bulter, Sun-Gu Lee, Wilson WaiChun Wong, Eileen Fung, Michael R. Connor, and James C. Liao. Design of artificial cell-cell communication using gene and metabolic networks. *Proc. Natl. Acad. Sci. USA*, 101(8):2299–2304, February 2004.

[18] Ming-Tang Chen. *Artificial Cell-Cell Communication and Multi-cellular Pattern Formation in Yeast Saccharomyces cerevisiae.* PhD dissertation, Princeton University, Department of Electrical Engineering, November 2009.

[19] Ming-Tang Chen and Ron Weiss. Artificial cell-cell communication in yeast saccharomyces cerevisiae using signaling elements from arabidopsis thaliana. *Nat. Biotechnol.*, 23(12):1551–1555, November 2005.

[20] Thomas W. Christianson, Robert S. Sikorski, Michael Dante, James H. Shero, and Philip Hieter. Multifunctional yeast high-copy-number shuttle vectors. *Gene*, 110(1):119–122, January 1992.

[21] Domitilla Del Vecchio. 2.997 biomolecular feedback systems. Lecture conducted from MIT, Cambridge, MA., March 2011.

[22] Domitilla Del Vecchio, Alexander J. Ninfa, and Eduardo D. Sontag. Modular cell biology: retroactivity and insulation. *Mol. Syst. Biol.*, 4(161), February 2008.

[23] Manu J Dubin, Chris Bowler, and Giovanna Benvenuto. A modified gateway cloning strategy for overexpressing tagged proteins in plants. *Plant Methods*, 4(3), January 2008.

[24] John E. Dueber, Brian J. Yeh, Kayam Chak, and Wendell A. Lim. Reprogramming control of an allosteric signaling switch through modular recombination. *Science*, 301(5641):1904–1908, September 2003.

[25] Marietta Dunaways, John S. Olson, John M. Rosenberg, Olga B. Kallai, Richard E. Dickerson, and Kathleen Shive Matthews. Kinetic studies of inducer binding to lac repressor.operator complex. *J. Biol. Chem.*, 225(21):10115–10119, November 1980.

[26] Mary A Dwyera and Homme W Hellinga. Periplasmic binding proteins: a versatile superfamily for protein engineering. *Curr. Opin. Struct. Biol.*, 14(4):495–504, August 2004.

[27] Andreas Eilers, Kayoko Kanda, Birgit Klose, John Krolewski, and Thomas Decker. Constitutive stat1 tyrosine phosphorylation in u937 monocytes overexpressing the tyk2 protein tyrosine kinase does not induce gene transcription. *Cell Growth Diff.*, 7(6):833–840, June 1996.

[28] Tom Ellis, Tom Adie, and Geoff S. Baldwin. Dna assembly for synthetic biology: from parts to pathways and beyond. *Integr. Biol.*, 3(2):109–118, February 2011.

[29] Michael B. Elowitz and Stanislas Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, January 2000.

[30] Drew Endy. Foundations for engineering biology. *Nature*, 438(7067):449–453, November 2005.

[31] John C. Evans, Donald P. Huddler, Mark T. Hilgers, Gail Romanchuk, Rowena G. Matthews, , and Martha L. Ludwig. Structures of the n-terminal modules imply large domain motions during catalysis by methionine synthase. *Proc. Natl. Acad. Sci. USA*, 101(11):3729–3736, March 2004.

[32] William R. Farmer and James C. Liao. Improving lycopene production in escherichia coli by engineering metabolic control. *Nat. Biotechnol.*, 18(5):533–537, May 2000.

[33] Jan S. Fassler, William M. Gray, Cheryl L. Malone, Wei Tao, Hong Lin, and Robert J. Deschenes. Activated alleles of yeast sln1 increase mcm1-dependent reporter gene expression and diminish signaling through the hog1 osmosensing pathway. *J. Biol. Chem.*, 272(20):13365–13371, May 1997.

[34] Jan S. Fassler and Ann H. West. Genetic and biochemical analysis of the sln1 pathway in saccharomyces cerevisiae. *Methods Enzymol.*, 471:291–317, March 2010.

[35] Eileen Fung, Wilson W. Wong, Jason K. Suen, Thomas Bulter, Sun gu Lee, and James C. Liao. A synthetic gene-metabolic oscillator. *Nature*, 435(7038):118–122, May 2005.

[36] A. B. Futcher and B. S. Cox. Copy number and the stability of 2-micron circle-based artificial plasmids of saccharomyces cerevisiae. *J. Bacteriol.*, 157(1):283–290, January 1984.

[37] Bruce Futcher and John Carbon. Toxic effects of excess cloned centromeres. *Mol. Cell. Biol.*, 6(6):2213–2222, June 1986.

[38] Timothy S. Gardner, Charles R. Cantor, and James J. Collins. Construction of a genetic toggle switch in escherichia coli. *Nature*, 403(6767):339–342, January 2000.

[39] Eloi Gari, Lidia Piedrafita, Marti Aldea, and Enrique Herrero. A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in saccharomyces cerevisiae. *Yeast*, 13(9):837–848, July 1997.

[40] Daniel G Gibson, Hamilton O Smith, Clyde A Hutchison III, J Craig Venter, and Chuck Merryman. Chemical synthesis of the mouse mitochondrial genome. *Nat. Methods*, 7(11):901–903, October 2010.

[41] Daniel G Gibson, Lei Young, Ray-Yuan Chuang, J Craig Venter, Clyde A Hutchison III, and Hamilton O Smith. Enzymatic assembly of dna molecules up to several hundred kilobases. *Nat. Methods*, 6(5):343–345, May 2009.

[42] Astrid V Giesecke, Rui Fang, and J Keith Joung. Synthetic protein-protein interaction domains created by shuffling cys2his2 zinc-fingers. *Mol. Syst. Biol.*, 2(2006.0011), March 2006.

[43] R. Daniel Gietz and Robin A. Woods. Yeast transformation by the liac/ss carrier dna/peg method. *Methods Mol. Biol.*, 313:107–20, 2006.

[44] Walter Gilbert and Benno Muller-Hill. Isolation of the lac repressor. *Proc. Natl. Acad. Sci. USA*, 56(6):1891–1898, December 1966.

[45] Gurkan Guntas, Thomas J. Mansell, Jin Ryoun Kim, and Marc Ostermeier. Directed evolution of protein switches and their application to the creation of ligand-binding proteins. *Proc. Natl. Acad. Sci. USA*, 102(32):11224–11229, August 2005.

[46] Elizabeth A. Hackett, R. Keith Esch, Seth Maleri, and Beverly Errede. A family of destabilized cyan fluorescent proteins as transcriptional reporters in s. cerevisiae. *Yeast*, 23(5):333–349, April 2006.

[47] Robert Hamilton, Colin K. Watanabe, and Herman A. de Boer. Compilation and comparison of the sequence context around the aug start codons in saccharomyces cerevisiae mrnas. *Nucleic Acids Res.*, 15(8):3581–3593, April 1987.

[48] Dieter Hansen and Wolfgang Hillen. Tryptophan in alpha-helix 3 of tet repressor forms a sequence-specific contact with tet operator in solution. *J. Biol. Chem.*, 262(25):12269–12274, September 1987.

[49] James L. Hartley, Gary F. Temple, and Michael A. Brasch. Dna cloning using in vitro site-specific recombination. *Genome Res.*, 10(11):1788–1795, November 2000.

[50] Xin-Jian He and Jan S. Fassler. Identification of novel yap1p and skn7p binding sites involved in the oxidative stress response of saccharomyces cerevisiae. *Mol. Microbiol.*, 58(5):1454–1467, December 2005.

[51] Xin-Jian He, KariAn E. Mulford, , and Jan S. Fassler. Oxidative stress function of the saccharomyces cerevisiae skn7 receiver domain. *Eukaryot. Cell.*, 8(5):768–778, May 2009.

[52] Alan C. Hindmarsh, Peter N. Brown, Keith E. Grant, Steven L. Lee, Radu Serban, Dan E. Shumaker, and Carol S. Woodward. Sundials: Suite of nonlinear and differential/algebraic equation solvers. *ACM Transactions on Mathematical Software*, 31(3):363–396, September 2005.

[53] Stefan Hohmann. Osmotic stress signaling and osmoadaptation in yeasts. *Microbiol. Mol. Biol. Rev.*, 66(2):300–372, June 2002.

[54] Perry L. Howard, Marie C. Chia, Suzanne Del Rizzo, Fei-Fei Liu, and Tony Pawson. Redirecting tyrosine kinase signaling to an apoptotic caspase pathway through chimeric adaptor proteins. *Proc. Natl. Acad. Sci. USA*, 100(20):11267–11272, September 2003.

[55] Mark Hsieh and Michael Brenowitz. Comparison of the dna association kinetics of the lac repressor tetramer, its dimeric mutant laciadi, and the native dimeric gal repressor. *J. Biol. Chem.*, 272(35):22092–22096, August 1997.

145

[56] Farren J Isaacs, Daniel J Dwyer, Chunming Ding, Dmitri D Pervouchine, Charles R Cantor, and James J Collins. Engineered riboregulators enable post-transcriptional control of gene expression. *Nat. Biotechnol.*, 22(7):841–847, July 2004.

[57] Janaki Iyer and Nancy C. Reich. Constitutive nuclear import of latent and activated stat5a by its coiled coil domain. *FASEB J.*, 22(2):391–400, February 2008.

[58] Alain Jacquier and Bernard Dujon. An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell*, 41(2):383–394, June 1985.

[59] Fabiola Janiak-Spens, Paul F. Cook, and Ann H. West. Kinetic analysis of ypd1-dependent phosphotransfer reactions in the yeast osmoregulatory phosphorelay system. *Biochemistry*, 44(1):377–386, January 2005.

[60] Paul Jorgensen, Joy L. Nishikawa, Bobby-Joe Breitkreutz, and Mike Tyers. Systematic identification of pathways that couple cell growth and division in yeast. *Science*, 297(5580):395–400, July 2002.

[61] Marc Jungbluth, Christian Renicke, and Christof Taxis. Targeted protein depletion in saccharomyces cerevisiae by activation of a bidirectional degron. *BMC Syst. Biol.*, 4(176), December 2010.

[62] Sylwia Kedracka-Krok, Andrzej Gorecki, Piotr Bonarek, and Zygmunt Wasylewski. Kinetic and thermodynamic studies of tet repressor-tetracycline interaction. *Biochemistry*, 44(3):1037–1046, January 2005.

[63] Ahmad S. Khalil and James J. Collins. Synthetic biology: applications come of age. *Nat. Rev. Genet.*, 11(5):367–379, May 2010.

[64] Ursula Klingmuller, Ulrike Lorenz, Lewis C. Cantley, Benjamin G. Neel, and Harvey F. Lodish. Specific recruitment of sh-ptp1 to the erythropoietin receptor causes inactivation of jak2 and termination of proliferative signals. *Cell*, 80(5):729–738, March 1995.

[65] Hideki Kobayashi, Mads Kaern, Michihiro Araki, Kristy Chung, Timothy S. Gardner, Charles R. Cantor, and James J. Collins. Programmable cells: interfacing natural and engineered gene networks. *Proc. Natl. Acad. Sci. USA*, 101(22):8414–8419, June 2004.

[66] Beat P Kramer, Alessandro Usseglio Viretta, Marie Daoud-El Baba, Dominique Aube, Wilfried Weber, and Martin Fussenegger. An engineered epigenetic transgene switch in mammalian cells. *Nat. Biotechnol.*, 22(7):867–870, July 2004.

[67] Bernhard Krems, Christina Charizanis, and Karl-Dieter Entian. The response regulator-like protein pos9/skn7 of saccharomyces cerevisiae is involved in oxidative stress resistance. *Curr. Genet.*, 29(4):327–334, March 1996.

[68] Marc S. Lamphier and Mark Ptashne. Multiple mechanisms mediate glucose repression of the yeast gal1 gene. *Proc. Natl. Acad. Sci. USA*, 89(13):5922–5926, July 1992.

[69] Arthur Landy. Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Ann. Rev. Biochem.*, 58:913–949, 1989.

[70] Jaekwon Lee, Christian Godon, Gilles Lagniel, Daniel Spector, Jerome Garin, Jean Labarre, and Michel B. Toledano. Yap1 and skn7 control two specialized oxidative stress response regulons in yeast. *J. Biol. Chem.*, 274(23):16040–16046, June 1999.

[71] Sheng Li, Addison Ault, Cheryl L. Malone, Desmond Raitt, Susan Dean, Leland H. Johnston, Robert J. Deschenes, and Jan S. Fassler. The yeast histidine protein kinase, sln1p, mediates phosphotransfer to two response regulators, ssk1p and skn7p. *EMBO J.*, 17(23):6952–6962, December 1998.

[72] Sheng Li, Susan Dean, Zhijian Li, Joe Horecka, Robert J. Deschenes, and Jan S. Fassler. The eukaryotic two-component histidine kinase sln1p regulates och1 via the transcription factor, skn7p. *Mol. Biol. Cell*, 13(2):412–424, February 2002.

[73] Jade Mei-Yeh Lu, Robert J. Deschenes, and Jan S. Fassler. Saccharomyces cerevisiae histidine phosphotransferase ypd1p shuttles between the nucleus and cytoplasm for sln1-dependent phosphorylation of ssk1p and skn7p. *Eukaryot. Cell*, 2(6):1304–1314, December 2003.

[74] Atsuya Maeda, Susannah M. Wurgler-Murphy, and Haruo Saito. A two-component system that regulates an osmosensing map kinase cascade in yeast. *Nature*, 369(6477):242–245, May 1994.

[75] T. Maeda, M. Takekawa, and H. Saito. Activation of yeast pbs2 mapkk by mapkkks or by binding of an sh3-containing osmosensor. *Science*, 269(5223):554–558, July 1995.

[76] Enrico Magnani, Linnea Bartling, and Sarah Hake. From gateway to multisite gateway in one recombination event. *BMC Mol. Biol.*, 7(46), December 2006.

[77] Xiang Mao, Zhiyong Ren, Gregory N. Parker, Holger Sondermann, Michael A. Pastorello, Wei Wang, John S. McMurray, Borries Demeler, James E. Darnell Jr., and Xiaomin Chen. Structural bases of unphosphorylated stat1 association and receptor binding. *Mol. Cell*, 17(6):761–771, March 2005.

[78] Vincent J J Martin, Douglas J Pitera, Sydnor T Withers, Jack D Newman, and Jay D Keasling. Engineering a mevalonate pathway in escherichia coli for production of terpenoids. *Nat. Biotechnol.*, 21(7):796–802, July 2003.

[79] Carolina Mateus and Simon V. Avery. Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry. *Yeast*, 16(14):1313–1323, October 2000.

147

[80] Kevin M. McBride, Gregg Banninger, Christine McDonald, and Nancy C. Reich. Regulated nuclear import of the stat1 transcription factor by direct binding of importin-alpha. *EMBO J.*, 21(7):1754–1763, April 2002.

[81] Kevin M. McBride and Nancy C. Reich. The ins and outs of stat1 nuclear transport. *Sci. STKE*, 195(RE13), August 2003.

[82] Karsten Melcher, Balasubramanya Sharma, Wei Vivianne Ding, and Mark Noldena. Zero background yeast reporter plasmids. *Gene*, 247(1-2):53–61, April 2000.

[83] Claudia Mertens, Minghao Zhong, Ravi Krishnaraj, Wenxin Zou, Xiaomin Chen, and James E. Darnell Jr. Dephosphorylation of phosphotyrosine on stat1 dimers requires extensive spatial reorientation of the monomers facilitated by the n-terminal domain. *Genes Dev.*, 20(24):3372–3381, December 2006.

[84] Hitoshi Miyasaka. The positive relationship between codon usage bias and translation initiation aug context in saccharomyces cerevisiae. *Yeast*, 15(8):633–637, June 1999.

[85] Rumiko Mizuguchi and Masanori Hatakeyama. Conditional activation of janus kinase (jak) confers factor independence upon interleukin-3-dependent cells: Essential role of ras in jak-triggered mitogenesis. *J. Biol. Chem.*, 273(48):32297–32303, November 1998.

[86] B. A. Morgan, N. Bouquin, G. F. Merrill, and L. H. Johnston. A yeast transcription factor bypassing the requirement for sbf and dsc1/mbf in budding yeast has homology to bacterial signal transduction proteins. *EMBO J.*, 14(22):5679–5689, November 1995.

[87] Brian A. Morgan, Geoffrey R. Banks, W. Mark Toone, Desmond Raitt, Shusuke Kuge, and Leland H. Johnston. The skn7 response regulator controls gene expression in the oxidative stress response of the budding yeast saccharomyces cerevisiae. *EMBO J.*, 16(5):1035–1044, March 1997.

[88] Richard Moriggl, Valerie Gouilleux-Gruart, Ruth Jahne, Susanne Berchtold, Christoph Gartmann, Xiuwen Liu, Lothar Hennighausen, Athanassia Sotiropoulos, Bernd Groner, and Fabrice Gouilleux. Deletion of the carboxyl-terminal transactivation domain of mgf-stat5 results in sustained dna binding and a dominant negative phenotype. *Mol. Cell. Bio.*, 16(10):5691–5700, October 1996.

[89] Koki Moriyoshi. pbt, a novel vector for tetracycline-regulated yeast three-hybrid assay. *Nucleic Acids Res.*, 37(2):e11, February 2009.

[90] K. E. Mulford and J. S. Fassler. Association of the skn7 and yap1 transcription factors in the saccharomyces cerevisiae oxidative stress response. *Eukaryot. Cell*, 10(6):761–769, April 2011.

[91] Benno Muller-Hill. *The lac Operon: a short history of a genetic paradigm*, volume 2. Walter de Gruyter, Berlin, Germany, 1996.

[92] Benno Muller-Hill. The function of auxiliary operators. *Mol. Microbiol.*, 29(1):13–18, July 1998.

[93] Brian Munsky, Brooke Trinh, and Mustafa Khammash. Listening to the noise: random fluctuations reveal gene network parameters. *Mol. Syst. Biol.*, 5(318), October 2009.

[94] Sandhya S. Narang, Cheryl L. Malone, Robert J. Deschenes, and Jan S. Fassler. Modulation of yeast sln1 kinase activity by the ccw12 cell wall protein. *J. Biol. Chem.*, 283(4):1962–1973, November 2008.

[95] Dante Neculai, Ana Mirela Neculai, Sophie Verrier, Kenneth Straub, Klaus Klumpp, Edith Pfitzner, and Stefan Becker. Structure of the unphosphorylated stat5a dimer. *J. Biol. Chem.*, 280(49):40782–40787, September 2005.

[96] Darin B. Ostrander and Jessica A. Gorman. The extracellular domain of the saccharomyces cerevisiae sln1p membrane osmolarity sensor is necessary for kinase activity. *J. Bacteriol.*, 181(8):2527–2534, April 1999.

[97] Sang-Hyun Park, Ali Zarrinpar, and Wendell A. Lim. Rewiring map kinase pathways using alternative scaffold assembly mechanisms. *Science*, 229(5609):1061–1064, February 2003.

[98] Anne-Laure Perrauda, Verena Weissb, and Roy Gross. Signalling pathways in two-component phosphorelay systems. *Trends Microbiol.*, 7(3):115–120, March 1999.

[99] Francesc Posas, Susannah M. Wurgler-Murphy, Tatsuya Maeda1, Elizabeth A. Witten, Tran Cam Thai1, and Haruo Saito. Yeast hog1 map kinase cascade is regulated by a multistep phosphorelay mechanism in the sln1-ypd1-ssk1 two-component osmosensor. *Cell*, 86(6):865–875, September 1996.

[100] Desmond C. Raitt, Anthony L. Johnson, Alexander M. Erkine, Kozo Makino, Brian Morgan, David S. Gross, and Leland H. Johnston. The skn7 response regulator of saccharomyces cerevisiae interacts with hsf1 in vivo and is required for the induction of heat shock genes by oxidative stress. *Mol. Biol. Cell*, 11(7):2335–2347, July 2000.

[101] Nancy C. Reich and Ling Li. Tracking stat nuclear traffic. *Nat. Rev. Immunol.*, 6(8):602–612, August 2006.

[102] Sean E. Reichhelda, Zhou Yub, and Alan R. Davidson. The induction of folding cooperativity by ligand binding drives the allosteric response of tetracycline repressor. *Proc. Natl. Acad. Sci. USA*, 106(52):22263–22268, December 2009.

[103] Pipsa Saharinen, Kati Takaluoma, and Olli Silvennoinen. Regulation of the jak2 tyrosine kinase by its pseudokinase domain. *Mol. Cell Biol.*, 20(10):3387–3395, May 2000.

[104] Joseph Sambrook and David W. Russel. *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, third edition, January 2001.

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

[106] Robert S. Sikorski and Philip Hieter. A system of shuttle vectors and yeast host strains designed for efficient manipulation of dna in saccharomyces ceratisiae. *Genetics*, 112(1):19–27, May 1989.

[107] Richelle Sopko, Dongqing Huang, Nicolle Preston, Gordon Chua, and et al. Mapping pathways and phenotypes by systematic gene overexpression. *Mol. Cell*, 21(3):319–330, February 2006.

[108] Michail Stamatakis and Nikos V. Mantzaris. Comparison of deterministic and stochastic models of the lac operon genetic network. *Biophys. J.*, 96(3):887–906, February 2009.

[109] Hisanori Tamaki, Cheol-Won Yun, Tomohiro Mizutani, Takahiro Tsuzuki, Yukinobu Takagi, Makiko Shinozaki, Yukiko Kodama, Katsuhiko Shirahige, and Hidehiko Kumagai. Glucose-dependent cell size is regulated by a g protein-coupled receptor system in yeast saccharomyces cerevisiae. *Genes Cells*, 10(3):193–206, March 2005.

[110] Wei Tao, Cheryl L. Malone, Addison D. Ault, Robert J. Deschenes, and Jan S. Fassler. A cytoplasmic coiled-coil domain is required for histidine kinase activity of the yeast osmosensor, sln1. *Mol. Microbiol.*, 43(2):459–473, January 2002.

[111] Barbara J. Thomas and Rodney Rothstein. Elevated recombination rates in transcriptionally active dna. *Cell*, 56(4):619–630, February 1989.

[112] Tsz-Leung To and Narendra Maheshri. Noise can induce bimodality in positive transcriptional feedback loops without bistability. *Science*, 327(5969):1142–1145, February 2010.

[113] Christina B. Tyson, Peter G. Lord, and Alan E. Wheals. Dependency of size of saccharomyces cerevisiae cells on growth rate. *J. Bacteriol.*, 138(1):92–98, April 1979.

[114] Andreas Untergasser, Harm Nijveen, Xiangyu Rao, Ton Bisseling, Rene Geurts, and Jack A.M. Leunissen. Primer3plus, an enhanced web interface to primer3. *Nucleic Acids Res.*, 35(W71-4), July 2007.

[115] Alexander Varshavsky. The n-end rule: functions, mysteries, uses. *Proc. Natl. Acad. Sci. USA*, 93(22):12142–12149, October 1996.

[116] Uwe Vinkemeier. Getting the message across, stat! design principles of a molecular signaling circuit. *J. Cell Biol.*, 167(2):197–201, October 2004.

[117] Warren P. Voth, James D. Richards, Janet M. Shaw, and David J. Stillman. Yeast vectors for integration at the ho locus. *Nucleic Acids Res.*, 29(12):e59, June 2001.

[118] Achim Wach, Arndt Brachat, Rainer Pohlmann, and Peter Philippsen. New heterologous modules for classical or pcr-based gene disruptions in saccharomyces cerevisiae. *Yeast*, 10(13):1793–1808, December 1994.

[119] Nikola Wenta, Holger Strauss, Stefanie Meyer, and Uwe Vinkemeier. Tyrosine phosphorylation regulates the partitioning of stat1 between different dimer conformations. *Proc. Natl. Acad. Sci. USA*, 105(27):9238–9243, July 2008.

[120] Kim E. Williams and Martha S. Cyert. The eukaryotic response regulator skn7p regulates calcineurin signaling through stabilization of crz1p. *EMBO J.*, 20(13):3473–3483, July 2001.

[121] Elizabeth A. Winzeler, Daniel D. Shoemaker, Anna Astromoff, Hong Liang, Keith Anderson, and et al. Functional characterization of the s. cerevisiae genome by gene deletion and parallel analysis. *Science*, 285(5429):901–906, August 1999.

[122] Hideo Yasukawa, Hiroyuki Misawa, Hiroshi Sakamoto1, Masaaki Masuhara, Atsuo Sasaki, Toru Wakioka, Satoshi Ohtsuka, Tsutomu Imaizumi, Tadashi Matsuda, James N.Ihle, and Akihiko Yoshimura. The jak-binding protein jab inhibits janus tyrosine kinase activity through binding in the activation loop. *EMBO J.*, 18(5):1309–1320, March 1999.

[123] Lingchong You, III Robert Sidney Cox, Ron Weiss, and Frances H. Arnold. Programmed population control by cell-cell communication and regulated killing. *Nature*, 428(6985):868–871, April 2004.

[124] Min You, De-Hua Yu, and Gen-Sheng Feng. Shp-2 tyrosine phosphatase functions as a negative regulator of the interferon-stimulated jak/stat pathway. *Mol. Cell Biol.*, 19(3):2416–2424, March 1999.