In Vitro Synthetic Transcriptional Networks

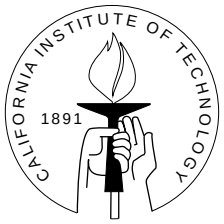
Thesis by

Jongmin Kim

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2007

(Defended Dec 6, 2006)

Acknowledgements

First of all, I would like to thank my excellent advisor, Erik Winfree, who is full of great ideas. His challenge of building a general biochemical circuit turned out to be quite difficult. However, his enthusiasm and insight along with incessant support saved the day. One of my fond memories is building a cheap fluorometer machine with Erik.

I thank my thesis committee members: Paul Sternberg, Barbara Wold, and Michael Elowitz. Paul's enthusiasm and humor lightened up committee meetings. Barbara had wonderful comments and practical tips when I started working on my projects. Michael's synthetic biology work has been a great inspiration for my research directions. I thank Judith Campbell for helpful comments during my candidacy and Stephen Quake for emphasis of detailed understanding during my committee meeting.

I have great affection for the DNA and Natural Algorithms group members: Paul Rothemund, Georg Seelig, Sungha Park, Peng Yin, Rebecca Schulman, David Soloveichik, Rizal Hariadi, Joseph Schaeffer, Dave Zhang, Nadine Dabby, Rob Barish, and all past members. I would like to thank all of them for discussion and experimental tips. We also had fun outside the lab. For me, the most memorable group activity was kayaking in Mono Lake where Dave and I began to swim with shrimps. I thank Rizal for suggesting a place to live and being a great neighbor, Dave Zhang and Sarina Mohanty for help on early experiments, and Dan Stick and Casimir Wierzynski for help on modeling. I also thank Elisa Franco, an extended lab member, for characterizing negative autoregulation. I thank Karolyn Young, our lab secretary, for organizing lab supplies.

I had wonderful undergraduates who worked with me: Pakpoom Subsoontorn, Ken Ho, and Kristin White. Pakpoom characterized a positive-feedback switch in detail, which helped my construction of oscillators. Ken modeled the mutually inhibitory bistable network in detail. Kristin pioneered fluorescence techniques for the transcriptional network.

I thank Bernard Yurke, John Hopfield, Friedrich Simmel, and Richard Murray for helpful discussion and thank Murray Deutscher and Arun Malhotra for sharing RNases. And I thank CDS 273 students for tackling modeling issues, especially Mary Dunlop and Ian Shapiro for their lasting interest in the project.

For technical help, I thank Seungyong Jung for measurement of small-volume reactions and

Tim Liedl for microfluidic chip construction.

For financial support I thank the following funding agencies: National Science Foundation and Office of Naval Research.

I would like to thank all my friends here at Caltech who helped me throughout the years.

Finally I thank all my family. My wife Elena filled my wildly varying schedule with fun activities and mostly love. My parents also supported me throughout my graduate studies and they nurtured my interest in science from the very beginning.

Abstract

Information processing using biochemical circuits is essential for survival and reproduction of natural organisms. Construction of synthetic biochemical circuits from simple components provides a useful approach to establish the minimal determinants required for complex logical functions. As stripped-down analogues of genetic regulatory networks in cells, we engineered artificial transcriptional networks consisting of synthetic DNA switches, regulated by RNA signals acting as transcription activators or repressors, and two enzymes, bacteriophage T7 RNA polymerase and *Escherichia coli* ribonuclease H. The synthetic switch design is modular with programmable connectivity and allows dynamic control of RNA signals through enzyme-mediated production and degradation. The switches support sharp and adjustable thresholds using a competitive hybridization mechanism, analogous to a biological threshold mechanism, "inhibitor ultrasensitivity," thus allowing arbitrary analog or digital circuits to be created in principle. Theoretical correspondence of our biochemical network to neural networks where synaptic weights and thresholds are encoded by concentrations of DNA strands greatly facilitates network design and analysis. Experimentally, we have constructed and analyzed several simple networks: positive and negative autoregulatory circuits, a mutual inhibitory circuit, and oscillators with positive and negative feedback. Reasonable agreement between experimental data and a simple mathematical model was obtained for switch input/output functions, phaseplane trajectories, the bifurcation diagram, and oscillation periods. A systematic quantitative characterization lead to identification of important network properties such as the saturation of degradation machinery and challenges to understand such as the interference by incomplete RNA signals. Construction of larger synthetic circuits provides a unique opportunity for evaluating model inference, prediction, and design of complex biochemical systems and could be used to control nanoscale devices and artificial cells.

Table of Contents

Chapter 1 Introduction

1.1 Cellular networks

What makes living systems unique? Albeit governed by the same physical laws, living systems take advantage of their environment unlike nonliving systems. For instance, cells have the ability to process information for survival and reproduction using functional circuits made up of many species of interacting molecules (Hartwell et al. 1999). Thus, information processing through regulatory networks lies at the heart of living systems. Taking a top-down view of protein-protein interactions, signaling pathways, and gene regulatory pathways, the basic architecture of biological networks has been analyzed (Bray 2003). Network description of cellular circuits allows application of tools and concepts developed in fields such as graph theory, physics, sociology, and engineering (Alon 2003). Remarkably, biological networks share the design principles of engineered networks: modularity, robustness, and recurring circuit elements. A module in a network is a set of nodes that have strong interactions and a common function (Hartwell et al. 1999). Modules in engineering, and presumably in biology, have special features that make them easily embedded in almost any system. Robustness to component tolerance dictates that the design must function under plausible fluctuations and interferences due to the components and the environment (Savageau 1971). Recurring network motifs for signal processing tasks such as filtering out input noise, accelerating throughput of the network, or temporal programming can be found in biological networks (Shen-Orr et al. 2002). The fact that a gene circuit has the above characteristics imposes severe constraints on its design and potentially helps delineate system architecture with limited data.

Interestingly, spontaneous evolution of modularity and network motifs has been demonstrated in computational evolution model of electronic circuits and neural networks (Kashtan and Alon 2005). Many computational models of biological evolution use networks in a population explored by means of mutations, crossover, and duplication to be selected for a defined goal. The evolved systems typically result in intricately wired nonmodular solutions because these nonmodular solutions are more optimized than human-engineered counterparts. Lack of modularity has been one of the reason why computational evolution can generate designs for simple tasks, but has difficulty in scaling up to more complex tasks. By imposing modularly varying goals for network evolution, the authors

achieved modular architecture and network motifs that work on subgoals but with more computation units (hence less optimal). The modularity decreased quickly when the network was trained on a single goal or nonmodularly varying goals. The authors suggest that biological networks have modularity because higher modularity has higher adaptability and organisms evolve in environments that require a certain set of basic functions in different combinations.

The overall architecture of cellular network is of great interest. For example, the number of connections per node for a regular or random network will have a roughly normal distribution with an average value that gives a characteristic scale to the network. By contrast, in a "scale-free" network, the number of molecules (N) with a given number of connections (k) falls off as a power law: $N(k)\tilde{}k^{-g}$, where no characteristic peak value can be found. Protein-protein interaction maps have the features of a scale-free network; their degree sequences (number of edges per node) often follow a long-tailed distribution (Strogatz 2001). In addition, natural networks often show the small-world property of short paths between nodes and highly clustered connections. However, the fact that a network has scale-free properties is of limited use, since power laws occur very widely in nature, possibly with different mechanistic origins (Bray 2003).

More recently, it has been suggested that biological networks have additional constraints that are beyond simple scale-free networks (Mattick and Gagen 2005). Networks that are simple connection networks, such as the Internet, are able to grow in an unconstrained way. In contrast, regulatory networks such as genetic regulatory networks in biology are networks that must be able to operate in a globally responsive way. To maintain global connectivity, the number of connections scales quadratically with network size. Such global responsiveness imposes an upper size limit on the complexity of integrated systems due to the costs incurred by the need for an increased number of connections and levels of regulation (Mattick and Gagen 2005). An alternative approach is to introduce dedicated hierarchies. However, each level of regulatory hierarchy introduces time delays, increases noise and stochastic errors (Hooshangi, Thiberge, and Weiss 2005). Regulatory proteins scale almost quadratically with genome size in prokaryotes (van Nimwegen 2003), and extrapolation of this relationship suggests that prokaryotes have reached their complexity limit by their reliance on protein-based regulatory architecture. Eukaryotes have a far more developed RNA processing and signaling system than prokaryotes, which appears to be linked to the more sophisticated pathway of gene regulation. Recently, it is suggested that in addition to being a digital storage medium, non-coding RNA themselves are actually transmitting digital signals (Mattick and Makunin 2006). In contrast, regulatory proteins act mainly as analog components because their signals are transmitted

as their concentrations. It is possible that the cellular network complexity limit was lifted by the use of both digital and analog signals. An analogous characteristic can be found in the neuronal networks of the brain. The cortex has many of the hallmarks of an energy efficient hybrid device. As in electronic hybrid devices, the synaptic inputs are integrated as analog modules and the action potential transmits the integrated signal digitally (Laughlin and Sejnowski 2003). Although the functional characterization of cellular network is a daunting task, investigating the network at different levels of complexity starting from basic network motifs would provide useful insights.

1.2 Synthetic approaches to understand cellular networks

How should we understand the behavior of cellular networks? In many cases, network description in an abstract sense would not be enough to understand cellular networks. The investigation of detailed kinetics and reaction mechanisms among the constituent macromolecules may be required. The reductionist approach attempts to explain the behavior of cellular networks in terms of the behavior of the components. Despite many molecular components of biological organisms being identified and characterized using genetic and biochemical techniques, it is still not possible to predict system behavior except in the simplest systems. This indicates that the great complexity of cellular network hinders prediction of system behavior from characterized components and that we need a better framework for understanding cellular network behavior and design principles.

Synthetic biology provides an alternative to the study of cellular networks. By constructing increasingly complex analogues of natural circuits, synthetic biology attempts to test sufficiency of mechanistic models and gain insights that observation and analysis alone do not provide (Benner and Sismour 2005). A common ground between the synthetic biology and engineering communities lies in the global strategy by which scientists come to understand their subject matter, synthesis. For engineering purposes, parts are most suitable when they contribute independently to the whole. This "independence property" allows one to predict the behavior of an assembly. In terms of independence, DNA molecules described by Watson–Crick model stand out because each nucleotide pair contributes independently to the stability of a duplex to a good approximation (SantaLucia 1998). No other molecular system can be described so simply. For example, the behavior of a protein is generally not a function of the behavior of its constituent amino acids, even as an approximation.

Although amino acids may be a poor unit to apply independence property, we can treat natural folded proteins as interchangeable parts. Several synthetic networks constructed by rearranging reg-

ulatory components in a cell have been characterized, including autoregulators (Becskei and Serrano 2000; Becskei et al. 2001), feedforward cascades (Basu et al. 2004; Hooshangi et al. 2005), bistable memory element (Gardner et al. 2000), and oscillators (Elowitz and Leibler 2000; Atkinson et al. 2003). For this type of network design to lead to an improved understanding of naturally occurring networks, detailed studies of the synthetic systems are needed (Benner and Sismour 2005), for example, through a systematic examination of the effects of parameter variations with quantitative modeling and analysis (Ozbudak et al. 2004). However, the quantitative predictions of mathematical models for many non-trivial synthetic network remain essentially untested because the enormous complexity of the cell hinders suitable perturbation and measurement approaches in these synthetic circuits.

For example, a bistable memory was constructed by Gardner et al. (2000) where two constitutively expressed repressors regulate the expression of each other. Several plasmid constructions with different promoters and ribosome binding sequences could show bistability. However, it was not determined how different experimental constructs correspond to mathematical model parameters and where stochastic noise that blurs bifurcation originates. The first synthetic oscillator was a ring oscillator constructed by Elowitz and Leibler (2000) where three repressors regulate the expression of other repressors. A tightly regulated promoter and a shorter protein half-life improves performance in the mathematical analysis, which they also implemented in their experimental design. Although the mathematical model provided insights into the system design, the phase diagram constructed by model was not experimentally verified. It remains unclear how the high variability of oscillation can be explained by stochastic noise or interaction with other cellular processes and why only 40% of the cells exhibited oscillation. The oscillator design of Atkinson et al. (2003) involves a positive regulator which activates its own expression and a repressor. This design did not involve a degradation sequence, and the experiments were performed in a continuous bioreactor under constant cell density. This oscillator displayed oscillation dynamics at population level, even though the oscillation was damped. The authors suggested various parameter variations such as messenger RNA stability and protein stability to achieve sustained oscillation. But none was pursued due to experimental difficulties. Of equal concern, the mechanism for synchronization remains unclear (Wong and Liao 2006). Combinatorial generation of small regulatory circuits yielded networks that implemented a variety of 2-input logic gates (Guet et al. 2002). But several of the networks' behaviors could not be explained in terms of network topology. A theoretical study concluded that a general quantitative model of gene expression is unable to explain the results of Guet

et al. (2002) without the added assumption of competition for degradation enzymes (Kim and Tidor 2003). Synthetic networks *in vivo* have been reusing previously used parts, not only because a single point mutation may alter the *in vivo* activity of the network, but also because we cannot predict how redesigned molecules such as synthetic promoters will behave (Ventura et al. 2006). Mutations can be a serious problem especially when a large population of cells are considered. A "population control" circuit (You et al. 2004) utilized bacterial quorum-sensing system linked to cell death signal to regulate cell density of *E. coli* population. A steady-state cell density in regulated cell culture was maintained about ten times lower than that of the control culture. The disadvantage in growth means that cells that acquired mutations to disrupt the synthetic circuit control will outgrow those of regulated cells. A microfluidic microreactor alleviated this problem by greatly reducing population size (Balagadde et al. 2005), and allowed monitoring of synthetic circuit behavior over hundreds of hours. Interestingly, limiting the population size also allowed observation of sustained oscillation in cell density unlike bulk samples. However, mutation and loss of control by synthetic network *in vivo* is inevitable unless the network confers selective advantage in cellular growth.

RNA molecules play important and diverse regulatory roles in the cell by virtue of their interaction with other nucleic acids, proteins and small molecules. For instance, diverse *cis* and *trans* gene regulation by noncoding RNA molecules such as microRNAs (Carrington and Ambros 2003) and antisense RNAs (Kramer et al. 2003) have been characterized in natural organisms. Researchers have engineered RNA molecules with new biological functions realized in bacteria and yeast (Isaacs et al. 2004; Bayer and Smolke 2005). Isaacs et al. (2004) achieved repression of a target gene by forming a hairpin structure in the 5' untranslated region of the mRNA (*cis*-regulator), sequestering the ribosome binding sequence. Expression of targeted *trans*-RNA activator allowed translation from modified mRNA by exposing the ribosome binding sequence. Bayer and Smolke (2005) developed RNA regulatory molecules that have an aptamer domain to recognize specific effector molecules and an antisense domain to control gene expression. Specific and dose-dependent switching responses of these regulatory RNA molecules have been demonstrated. The development of designed RNA switches can potentially be employed as cellular sensors and effectors to create programmable cells (Isaacs et al. 2006). However, these synthetic RNA regulation systems mainly demonstrated switching behavior rather than general network construction. The inputs and outputs of these systems are not homologous: inputs are small molecules and outputs are proteins or nucleic acids. Moreover, although thermodynamic calculations and measurements can easily be obtained for different nucleic acid species, quantitative models for their dynamic behavior is not available.

For instance, Bayer and Smolke developed variants with both lower and higher thresholds for theophylline by changing the stem stability. Although the threshold shift can be qualitatively explained, it is unclear how one can estimate the switching threshold from thermodynamic considerations.

An *in vitro* reconstruction with known components offers a unique opportunity to investigate how system behavior derives from reaction mechanisms. The first non-trivial system behavior created by an *in vitro* chemical system was the Belousov-Zhabotinsky oscillator (Zaikin and Zhabotinsky 1970), but it is difficult to see how these reaction mechanisms could support a wide variety of chemical logic, as is found in biochemistry. An excellent example of *in vitro* reconstruction using biochemical components is the cyanobacterial circadian clock, the operation of which has been shown to be independent of transcription and translation (Nakajima et al. 2005). However, operating and characterizing biochemical circuits outside the cell remains a challenge. A reconstituted cell-free transcription-translation system requires almost one hundred purified components (Shimizu et al. 2001) or poorly characterized cell extracts (Noireaux et al. 2003). A variety of interesting circuits can be constructed within cell-free transcription-translation systems. For example, Noireaux et al. constructed transcriptional activation and repression cascades, where the protein product of each stage activates or inhibits the following stage. Their construction used wheat germ extract and transcription is performed by bacteriophage T7, SP6, and *E. coli* RNA polymerase. The construction of different cascades is straightforward and potentially confers advantages because gene and polymerase concentrations can be controlled, reporter measurements are quantitative, and a large parameter space can be studied. However, substantial time delays and dramatic decrease in output production are incurred with each additional stage because of a bottleneck at the translation machinery, which hampered the development of quantitative models for this study. An *in vitro* transcription-translation network that emulates *Drosophila* embryonic patterns was constructed (Isalan et al. 2005). Utilizing regulatory interaction mediated by previously characterized zinc-finger proteins, networks of various connectivity were tested. The patterning behavior was qualitatively correct, and more mutual repression lead to overall lower activity yet sharper patterns. Moreover, the addition of protease stabilized the pattern over time, although the overall protein level was too low to be easily confirmed. However, the model did not achieve quantitative agreement with experiment and the model predicted unphysiologically slow diffusion for pattern generation in the *Drosophila* embryo.

Nucleic-acid-based networks greatly reduce the complexity of the production machinery. For example, feedback circuits modeled after predator-prey dynamics have been constructed as a much

simpler *in vitro* system containing only three enzymes, T7 RNA polymerase, M-MLV reverse transcriptase, and *E. coli* RNase H (Wlotzka and McCaskill 1997; Ackermann et al. 1998). The reaction scheme is based on self-sustained sequence replication, an isothermal amplification scheme for the coupled amplification of both DNA and RNA oligomers (Guatelli et al. 1990). Mathematical modeling suggests that coupling prey and predator cycles, where prey cycle provides a primer for predator cycle, with an appropriate flow rate in a chemostat can lead to oscillation. Yet, quantitative agreement of models and experiments was not achieved, possibly because of unmodeled dead-end side reactions or coupling of reaction rates by common use of enzymes. Even if the model assumptions are correct, the hybridization rates required for oscillation were too close to diffusion limit (10^8/M/s) to be achieved experimentally. Furthermore, it is unclear how to construct more complex circuits using this approach.

On a larger scale of synthetic efforts, assembling some type of cell (i.e., a self-replicating, membrane-encapsulated collection of biomolecules) would be a next major challenge (Forster and Church 2006). However small, a cellular gene set has to be self-sufficient in the sense that cells generally import metabolites but not functional macromolecules. *Mycoplasma genitalium*, a parasitic bacteria with small genome size, is an attractive model in search of minimal genome. One of the early estimates of the minimal genome set suggested 256 genes including 234 *M. genetalium* genes that have orthologs among the genes of *Haemophilus influenzae* and 22 nonorthologous displacements (Mushegian and Koonin 1996). Most of the proteins encoded by the genes from the minimal set suggested by Mushegian and Koonin (1996), have eukaryotic or archaeal homologs but the key proteins of DNA replication do not, which led authors to speculate that the last common ancestor had an RNA genome. The estimated gene number could be further reduced by eliminating cofactors, regulatory genes, and applying the parsimony principle (Benner et al. 1989).

A recent estimate suggests 151 genes, 38 RNAs and 113 proteins, for a minimal genome (Forster and Church 2006). Lipids alone have been shown to be sufficient for formation of rudimentary membranous compartments capable of both transmembrane transport of small molecules and autocatalytic fission (Szostak et al. 2001). A bare-bone genome will perform basic DNA replication, transcription and translation processes. A surprisingly large fraction (96%) of the minimal gene set is devoted to translation mechanisms: ribosome components, a set of tRNA, a set of translational initiation, elongation, release factors, and a few chaperones. Different approaches for necessary mechanisms such as adaptation of rolling circle amplification for DNA replication were employed to reduce the number of genes. It is uncertain whether the minimal genome can function with-

out catabolism (nucleases and proteases), active conversion or removal of waste products (energy regenerating enzymes and membrane transporters), and regulatory feedback.

At any rate, a much simpler purified system based on a real cell will be easier to model and understand. It could certainly answer questions that cannot be answered *in vivo*, such as which set of macromolecules is sufficient for a functional cellular subsystem (Forster and Church 2006). The process can be iterative. If an *in vitro* system is a poor model for the *in vivo* system, then more complex analogs of the *in vivo* system can be included in the *in vitro* system. The simplest approach for creating a minimal cell may be by evolving an RNA polymerase made exclusively of RNA to replace all protein components of *in vitro* replicating and evolving systems (Szostak et al. 2001). One of the exciting development towards this direction is the templated assembly of RNA products catalyzed by ribozymes (Johnston et al. 2001). Their ribozyme used nucleoside triphosphates and the coding information of an RNA template to extend an RNA primer by the successive addition of up to 14 nucleotides with high accuracy. Their finding supports the RNA-world hypothesis regarding the early evolution of life, the main appeal of which is that ribozymes would have been far easier to duplicate than proteinaceous enzymes. Given that most of the minimal gene set is devoted to translation, a nucleic-acid-based artificial cell is certainly attractive. It remains to be seen whether suitable set of ribozymes can be found by *in vitro* evolution to replace protein enzymes.

Taken together, the synthetic approaches have successfully demonstrated several interesting networks *in vivo* and *in vitro*. Yet, quantitative modeling and prediction remain a challenge for synthetic networks. For *in vivo* circuits, the sheer complexity of production and degradation machinery and other cellular components interacting with the synthetic network makes quantitative modeling and prediction difficult. On the other hand, for *in vitro* circuits, the lack of complex feedback regulation for production and degradation machinery can lead to high variability and sensitivity in their performance. A far greater challenge is constructing a minimal cell, which would require both detailed characterization of individual macromolecules and step-by-step analysis as we construct and combine subsystems of greater complexity.

1.3 Contribution

The development of a simple *in vitro* biochemical regulatory network will have far-reaching consequences for understanding biochemical network design principles and constructing artificial cells

or subsystems therein.

Therefore, we address the following questions:

1. How can we construct biochemical circuits that are simple enough for quantitative analysis?

2. What is the computational power of biochemical circuits thus created?

Question (2) was first addressed by constructing an ideal biochemical reaction network model that mimics genetic regulatory circuits and by demonstrating its correspondence to a general class of neural networks. Our biochemical network model makes use of only RNA polymerase and ribonucleases in addition to synthetic DNA templates regulated by RNA transcripts. Here, we treat the transcriptionally controlled DNA switches as synapses and the concentration of RNA species as the states of neurons. Despite the simplicity of our system compared to other *in vitro* approaches (Noireaux et al. 2003; Isalan et al. 2005) that use protein signals, we show theoretically that arbitrary logic circuits and abstract neural network computations can be implemented (Kim et al. 2004). Moreover, the weights and thresholds of corresponding neural networks are represented by continuously adjustable concentrations of DNA molecules in our *in vitro* circuits (Kim et al. 2004). Whereas in previous formal models that treat genetic regulatory circuits as neural networks (Mjolsness et al. 1991; Buchler et al. 2003), thresholds are encoded as binding constants of regulatory proteins, and therefore tuning the circuit requires modifying protein structure through natural or directed evolution (Yokobayashi et al. 2002). An interesting possibility arises when the enzyme reactions are treated as Michaelis–Menten reactions rather than first-order reactions: a global feedback regulation by shared use of enzyme can be exploited for winner-take-all computation.

Question (1) was then addressed by experimentally verifying postulated model reactions and constructing various transcriptional regulatory circuits. We developed an experimental analogue of genetic regulatory circuits that makes use of only two proteins, T7 RNA polymerase and *E. coli* ribonuclease H, in addition to synthetic DNA templates regulated by RNA transcripts. This system meets our goal of dramatically reducing the chemical complexity by removing the irrelevant genes and regulatory processes of the whole organism, which includes removing protein production and degradation machinery. We chose nucleic acids as our signal molecules because it is most suitable to apply "independence properties" useful for synthesis. Nucleic acid regulatory molecules have the advantage that the structures are well defined and that interactions governed by Watson–Crick base-paring rules can be easily programmed, allowing for modular designs. First, we demonstrate that switches and feedforward circuits exhibit sigmoidal transfer curves with sharp and adjustable thresholds. The threshold is established by a competitive hybridization mechanism analogous to

the “inhibitor ultrasensitivity” mechanism (Ferrell 1996). Second, we demonstrate that our synthetic switch design is modular and programmable. We constructed several alternative networks by straightforward rewiring of switch connections. The performance of networks reasonably agree with individual switch characterization results. Third, we achieve dynamic behavior and steady-states in our *in vitro* circuit by balancing enzyme-controlled production and degradation mechanisms. This contrasts with many previous studies (Atkinson et al. 2003; Rosenfeld et al. 2005) that treat degradation and dilution of signal molecules (as occurs in exponentially growing cells and in chemostats) as a first-order process. We find that the degradation by RNase H could achieve steady-state switch activities in feedforward circuits and dynamically change switch states by removing transient RNA signals in feedback circuits. Finally, as a test of our understanding of reaction mechanisms, we construct simple mathematical models to explain various aspects of the circuit dynamics. The models reproduce the transfer curves for individual switches, the bifurcation diagram for the bistable circuits, the phaseplane dynamics for the bistable circuits, and the time-courses for the oscillators.

1.4 Publication list

This thesis contains material form one conference publication, one published journal article and two journal articles in preparation. The creative ideas and the actual work for all the results presented here are primarily due to me, except where explicitly noted. Some of the text and figures in this thesis come directly from those articles, although most of it has undergone revision for incorporation into this thesis. I am solely reponsible for any mistakes herein.

Chapter 2 uses material from:

Jongmin Kim, John J. Hopfield, Erik Winfree, “Neural network computation by *in vitro* transcriptional circuits” (Kim et al. 2004). The original model is due to Winfree and Hopfield. My contribution is working out details in consideration of my experimental system (such as leaky expression and Michaelis–Menten enzyme reactions) and case studies. I included more details and derivations in the appendix.

Chapter 3 uses material from:

Jongmin Kim, Kristin S. White, Erik Winfree, “Construction of an *in vitro* bistable circuit from synthetic transcriptional switches” (Kim et al. 2006). Winfree outlined some experiments. White performed early fluorescence experiments. I developed transcriptional switches and carried out all further experiments and analysis.

Chapter 4 uses material from:

Pakpoom Subsoontorn, Jongmin Kim, Erik Winfree, "Bistability in an *in vitro* positive feedback circuit" (in preparation). Winfree outlined some experiments. The actual experimental work is carried out by Pakpoom Subsoontorn. The original idea is due to me. I guided the whole process including DNA sequence design, experimental design, and mathematical modeling.

Chapter 5 uses material from:

Jongmin Kim, Erik Winfree, "*In vitro* synthetic oscillators" (in preparation). The creative idea is due to Winfree and me. All experimental work and data analysis is carried out by me.

1.5 Support

My work at Caltech was supported by NSF NIRT grant no. 0103002 to N. Seeman, ONR YIP grant no. N000140110813, and NSF ITR grant no. 0113443.

Chapter 2 A Theoretical Framework for *in Vitro* Transcriptional Circuits

2.1 Construction of the transcriptional network

The structural similarity of neural networks and genetic regulatory networks to digital circuits, and hence to each other, was noted from the very beginning of their study (McCulloch and Pitts 1943; Monod and Jacob 1961). In this work, we propose a simple biochemical system whose architecture mimics that of genetic regulation and whose components allow for *in vitro* implementation of arbitrary circuits. We use only two enzymes in addition to DNA and RNA molecules: RNA polymerase (RNAP) and ribonuclease (RNase). We develop a rate equation for *in vitro* transcriptional networks, and derive a correspondence with general neural network rate equations (Hopfield 1984). As proof-of-principle demonstrations, an associative memory task and a feedforward network computation are shown by simulation. A difference between the neural network and biochemical models is also highlighted: global coupling of rate equations through enzyme saturation can lead to global feedback regulation, thus allowing a simple network without explicit mutual inhibition to perform the winner-take-all computation. Thus, the full complexity of the cell is not necessary for biochemical computation: a wide range of functional behaviors can be achieved with a small set of biochemical components.

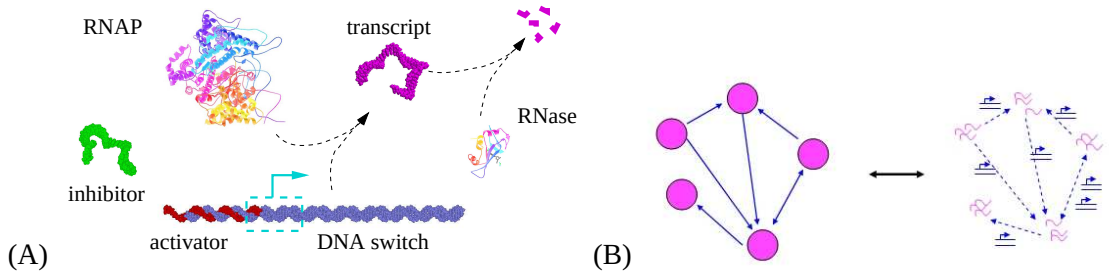
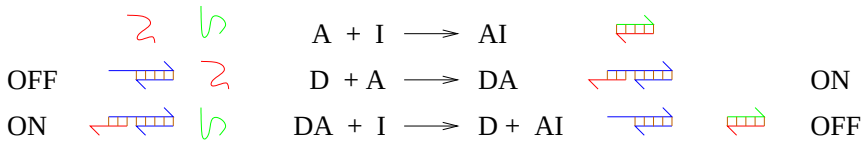


Figure 2.1. **(A)** The components of an *in vitro* circuit. The switch template (blue) is shown with the activator (red) attached. The dotted box indicates the promoter sequence and the downstream direction. **(B)** The correspondence between a neural network and an *in vitro* biochemical network. Neuron activity corresponds to RNA transcript concentration, while synaptic connections correspond to DNA switches with specified input and output.

2.1.1 The DNA transcriptional switch

The elementary unit of our networks will be a DNA switch, which serves the role of a gene in a genetic regulatory circuit. The basic requirements for a DNA switch are to have separate input and output domains, to transcribe poorly by itself (Martin and Coleman 1987), and to transcribe efficiently when an activator is bound to it. A possible mechanism of activation is the complementation of an incomplete promoter region, allowing more favorable binding of RNAP to the DNA template. Figure 2.1A illustrates our proposed design for DNA transcriptional switches and circuits. We model a single DNA switch with the following binding reactions:



where D (blue) is a DNA template with an incomplete promoter region, A (red) is an activator that complements the incomplete promoter region, and I (green) is an inhibitor complementary to A. Thus, I can bind free A. Furthermore, activator A contains a "toehold" region (Yurke and Mills 2003) that overhangs past the end of D, allowing inhibitor I to strip off A from the DA complex. D is considered OFF and DA is considered ON, based on their efficiency as templates for transcription. This set of binding reactions provides a means to choose the threshold of the sigmoidal activation function, as will be explained later.

RNAP and RNase drive changes in RNA transcript concentration; their activity is modeled using a first-order approximation for enzyme kinetics. For the moment, we assume that the input species (activator and inhibitor) are held at constant levels by external control.

By RNA polymerase	*By RNase*
$DA \xrightarrow{k_p} DA + R$	$R \xrightarrow{k_d} \phi$
$D \xrightarrow{\alpha k_p} D + R$	

For OFF-state switch D, the transcription rate is multiplied by α ($0 < \alpha < 1$) due to lack of activation. The complete degradation of RNA products by RNase is represented by ϕ . The concentration of enzymes set k_d and k_p.

In general, a set of chemical reactions obeying mass action have dynamics described by

$$\frac{d[X_i]}{dt} = \sum_\beta k_\beta \prod_j [X_j]^{r_j^\beta} (p_i^\beta - r_i^\beta),$$

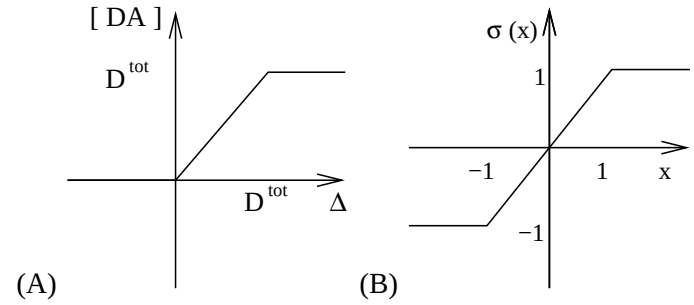


Figure 2.2. **(A)** $[DA]$ as a function of Δ. **(B)** The sigmoid $\sigma(x)$.

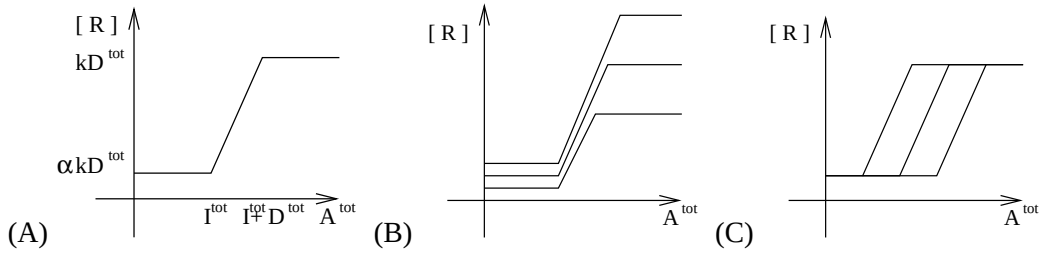


Figure 2.3. **(A)** $[R]$ as a function of A^{tot} **(B, C)** Tuneable sigmoid for three values of D^{tot} and I^{tot}, respectively.

where k_β is the rate constant, r_i^β is the stoichiometry of species X_i as a reactant (typically 0 or 1), and p_i^β is the stoichiometry of X_i as a product in reaction β. Analysis of our system is greatly simplified by the assumption that the binding reactions are fast and go to completion. We define D^{tot} as the sum of free and bound species: $D^{tot} = [D] + [DA]$. Similarly, $I^{tot} = [I] + [AI]$ and $A^{tot} = [A] + [DA] + [AI]$. Then, $[DA]$ depends on D^{tot} and Δ, where $\Delta = A^{tot} - I^{tot}$. Because I can scavenge A whether the latter is free or bound to D, A can activate D only when $\Delta > 0$. The amount of $[DA]$ is proportional to Δ when $0 < \Delta < D^{tot}$, as shown in figure 2.2A. It is convenient to represent this nonlinearity using a piecewise-linear approximation of a sigmoidal function, specifically, $\sigma(x) = \frac{|x+1|-|x-1|}{2}$ (figure 2.2B). Thus, we can represent $[DA]$ using σ and a rescaled Δ: $[DA] = \frac{1}{2}D^{tot}(1 + \sigma(\hat{\Delta}))$, where $\hat{\Delta} = \frac{2\Delta}{D^{tot}} - 1$ is called the signal activity. At steady-state, $k_d[R] = k_p[DA] + \alpha k_p[D]$; thus,

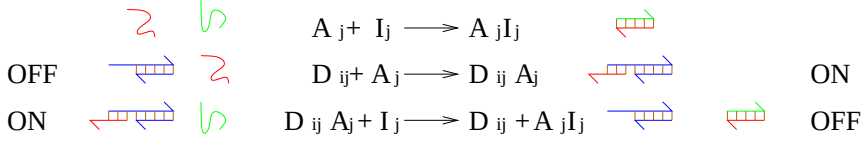
$$[R] = \frac{1}{2}\frac{k_p}{k_d}D^{tot}((1-\alpha)\sigma(\hat{\Delta}) + 1 + \alpha) \ .$$

If we consider the activator concentration as an input and the steady-state transcript concentration as an output, then the (presumed constant) inhibitor concentration, I^{tot}, sets the threshold, and the function assumes a sigmoidal shape (figure 2.3A). Adjusting the amount of template, D^{tot},

sets the magnitude of the output signal and the width of the transition region (figure 2.3B). We can adjust the width of the transition region independent of the threshold such that a step function would be achieved in the limit. Thus, we have a sigmoidal function with an adjustable threshold, without reliance on cooperative binding of transcription factors as is common in biological systems (Shea and Ackers 1985).

2.1.2 Networks of transcriptional switches

The input domain of a DNA switch is upstream of the promoter region; the output domain is downstream of the promoter region. This separation of domains allows us to design DNA switches that have any desired connectivity. We assume that distinct signals in the network are represented as distinct RNA sequences that have negligible crosstalk (undesired binding of two molecules representing different signals). The set of legitimate binding reactions is as follows:



where D_{ij} is the DNA template that has the jth input domain and ith output domain, the activator A_j complements the incomplete promoter region of D_{ij}, and the inhibitor I_j is complementary to A_j. Note that I_j can strip off A_j from the $D_{ij}A_j$ complex, thus imposing a sharp threshold as before. Again, we assume fast and complete binding reactions.

The set of enzyme reactions for the transcriptional network is as follows:

By RNA polymerase		*By RNase*
$D_{ij}A_j \xrightarrow{k_p} D_{ij}A_j + A_i$	if $s_{ij} = 1$	$I_j \xrightarrow{k_d} \phi$
$D_{ij} \xrightarrow{\alpha k_p} D_{ij} + A_i$		$A_j \xrightarrow{k_d} \phi$
$D_{ij}A_j \xrightarrow{k_p} D_{ij}A_j + I_i$	if $s_{ij} = -1$	$A_jI_j \xrightarrow{k_d} \phi$
$D_{ij} \xrightarrow{\alpha k_p} D_{ij} + I_i$		$D_{ij}A_j \xrightarrow{k_d} D_{ij},$

where $s_{ij} \in \{+1, -1\}$ indicates whether switch ij will produce an activator or an inhibitor. In fact, no single RNase can degrade all of the single-stranded RNA, double-stranded RNA, and DNA-RNA hybrid substrates. We assume that a suitable combination of RNases is able to achieve the desired degradation kinetics. This notation reflects that the production of I_i is equivalent to the consumption

of A_i. The change of RNA concentrations over time is easy to express with $\Delta_i = A_i^{tot} - I_i^{tot}$:

$$\frac{d\Delta_i}{dt} = -k_d \cdot \Delta_i + k_p \sum_j s_{ij}([D_{ij}A_j] + \alpha[D_{ij}]) \ . \tag{2.1}$$

2.1.3 Network equivalence

We show next that the time evolution of this biochemical network model is equivalent to that of a general Hopfield neural network model (1984):

$$\tau \frac{dx_i}{dt} = -x_i + \sum_j w_{ij}\sigma(x_j) + \theta_i \ . \tag{2.2}$$

Equation (2.1) can be rewritten to use the same nonlinear activation function σ defined earlier. Let $\hat{\Delta}_i = \frac{2\Delta_i}{D_{*i}^{tot}} - 1$ be a rescaled difference between activator and inhibitor concentrations, where D_{*i}^{tot} is the load on A_i, i.e., the total concentration of all switches that bind to A_i: $D_{*i}^{tot} = \sum_j D_{ji}^{tot}$ and $D_{ij}^{tot} = [D_{ij}A_j] + [D_{ij}]$. Then, we can derive the following rate equation, where $\hat{\Delta}_i$ plays the role of unit i's activity x_i (appendix 2.4.1):

$$\frac{1}{k_d}\frac{d\hat{\Delta}_i}{dt} = -\hat{\Delta}_i + \sum_j \left(\frac{k_p}{k_d}(1-\alpha)s_{ij}\frac{D_{ij}^{tot}}{D_{*i}^{tot}} \right) \sigma(\hat{\Delta}_j) + \left(\sum_j \frac{k_p}{k_d}(1+\alpha)s_{ij}\frac{D_{ij}^{tot}}{D_{*i}^{tot}} - 1 \right). \tag{2.3}$$

Given the set of constants describing an arbitrary transcriptional network, the constants for an equivalent neural network can be obtained immediately by comparing equations (2.2) and (2.3). The time constant τ is the inverse of the RNase degradation rate: fast turnover of RNA molecules leads to fast response of the network. The synaptic weight w_{ij} is proportional to the concentration of switch template ij, attenuated by the load on A_i. However, the threshold θ_i is dependent on the weights, perhaps implying a lack of generality. To implement an arbitrary neural network, we must introduce two new types of switches to the transcriptional network. To achieve arbitrary thresholds, we introduce bias switches D_{iB} which have no input domain and thus produce outputs constitutively; this adds an adjustable constant to the right-hand side of equation (2.3). When we translate arbitrary neural network constants to transcriptional network constants, it is convenient to have D_{*i}^{tot}s as constant. To balance the load on A_i, we add null switches D_{0i} which bind to A_i but have no output domain; this allows us to ensure that all D_{*i}^{tot}s are equal. Consequently, given any neural network with weights w_{ij} and thresholds θ_i, we can specify concentrations D_{ij}^{tot} such that the biochemical network has identical dynamics, for some τ (appendix 2.4.2).

2.1.4 Michaelis–Menten enzyme reactions

Next, we explore the validity of our assumption that enzyme kinetics are first-order reactions. A basic but more realistic model is the Michaelis–Menten mechanism (Hammes 2000), in which the enzyme and substrate bind to form an enzyme-substrate complex. For example, if E is RNAP,

$$E + D_{ij}A_j \underset{k_-}{\overset{k_+}{\rightleftharpoons}} ED_{ij}A_j \overset{k_{cat}}{\rightarrow} E + D_{ij}A_j + I_i \text{ or } A_i \ .$$

An important ramification of Michaelis–Menten reactions is that there is competition for the enzyme by the substrates, because the concentration of available enzymes is reduced as they bind to substrates, leading to saturation when the enzyme concentration is limiting. Using the steady-state assumption for Michaelis–Menten reactions, we establish the following relations to the rate constants of first-order reactions (appendix 2.4.3):

$$k_p = \frac{E^{tot}}{1+L} \cdot \frac{k_{cat}}{K_M} \qquad \alpha \cdot k_p = \frac{E^{tot}}{1+L} \cdot \frac{k'_{cat}}{K'_M} \qquad k_d = \frac{E_d^{tot}}{1+L_d} \cdot \frac{k_{d,cat}}{K_{d,M}}, \tag{2.4}$$

where k_{cat} and $K_M = (k_- + k_{cat})/k_+$ are the catalytic constant (enzyme's speed) and Michaelis constant (inverse of enzyme's affinity to target) of RNAP for the ON state switch, k'_{cat} and K'_M are for the OFF state switch, and $k_{d,cat}$ and $K_{d,M}$ are the constants of RNase. E^{tot} and E_d^{tot} are the concentrations of RNAP and RNase, respectively. We define the load on enzymes as the total concentration of binding targets divided by the Michaelis constants of the enzymes. The load on RNAP is calculated as $L = \sum_{i,j} \frac{[D_{ij}A_j]}{K_M} + \sum_{i,j} \frac{[D_{ij}]}{K'_M}$ and the load on RNase is calculated as $L_d = \sum_{i,j} \frac{[A_j]+[I_j]+[A_jI_j]+[D_{ij}A_j]}{K_{d,M}}$, both of which may be time varying. To make the first-order approximation valid, we must keep L and L_d constant. Introduction of a new type of switch with different Michaelis constants can make L constant by balancing the load on the enzyme. A scheme to keep L_d constant is not obvious, so we set reaction conditions such that $L_d \ll 1$.

2.2 Example computations by transcriptional networks

2.2.1 Feedforward networks

We first consider a feed-forward network to compute $f(x, y, z) = \bar{x}yz + \bar{y}z + x$. From the Boolean circuit shown in figure 2.4A, we can construct an equivalent neural network. We label units 1 through 6: units 1, 2, 3 correspond to inputs x, y, z whereas units 4, 5, 6 are computation units.

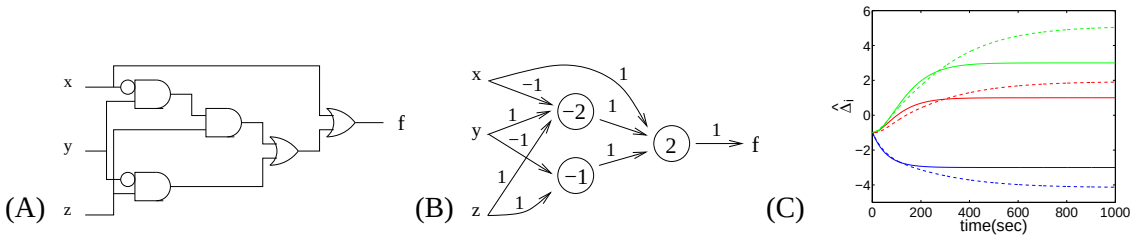


Figure 2.4. **(A, B)** A Boolean circuit and a neural network to compute $f(x, y, z) = \bar{x}yz + \bar{y}z + x$. **(C)** The activity of computation units (first-order approximation: solid lines; Michaelis-Menten reaction: dotted lines) for x = True = 1, y = False = -1, z = True = 1.

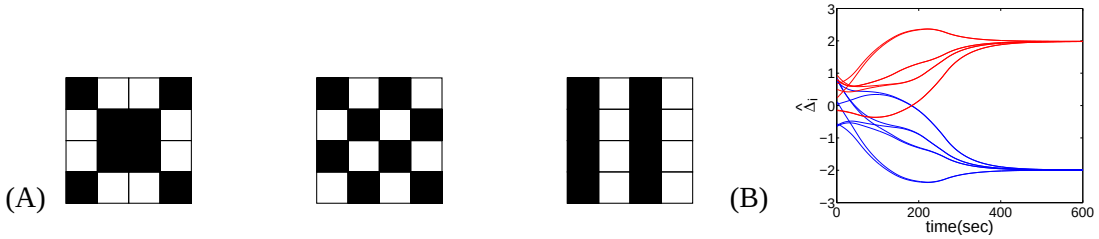


Figure 2.5. **(A)** The three patterns to be memorized. **(B)** Time-course for the transcriptional network recovery of the third pattern (odd columns: blue lines; even columns: red lines).

Using the conversion rule discussed in the network equivalence section, we can calculate the parameters of the transcriptional network. Under the first-order approximation of equation 2.3, the simulation result is exact (figure 2.4C). For comparison, we also explicitly simulated mass action dynamics for the full set of chemical equations without fast and complete hybridization and first-order enzyme reaction assumptions. We used the Michaelis–Menten enzyme reactions with biologically plausible rate constants and used E^{tot} and E_d^{tot} calculated from equation 2.4 with estimated values of L and L_d. The full model performs the correct calculation of f for all eight 3-bit inputs, although the magnitude of signals is exaggerated due to an underestimate of RNase load (figure 2.4C).

2.2.2 Associative memories

Figure 2.5A shows three 4-by-4 patterns to be memorized in a continuous neural network (Hopfield 1984). We chose orthogonal patterns because a 16 neuron network has limited capacity. Our training algorithm is gradient descent combined with the perceptron learning rule. After training, the parameters of the neural network are converted to the parameters of the transcriptional network as previously described. Starting from a random initial state, a typical response of the transcriptional network (with the first-order approximation of equation 2.3) is shown in figure 2.5B. Thus, our *in vitro* transcriptional networks can support complex sets of stable steady-states.

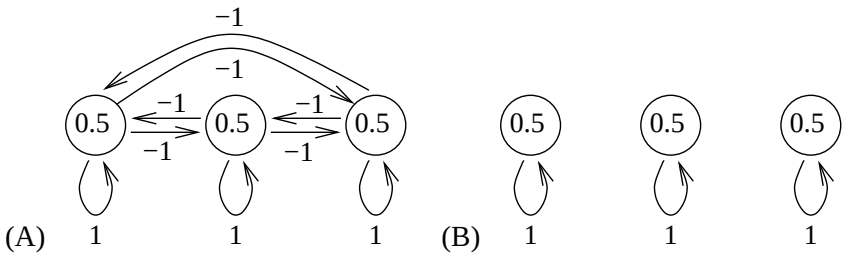


Figure 2.6. **(A)** A 3-unit WTA network with explicit mutual inhibition. **(B)** An equivalent biochemical network.

2.2.3 A winner-take-all network

Instead of trying to compensate for the saturation phenomena of Michaelis–Menten reactions, we can make use of it for computation. As an example, consider the winner-take-all computation (Yuille and Geiger 1995), which is commonly implemented as a neural network with $O(N^2)$ mutually inhibitory connections (figure 2.6A), but which can also be implemented as an electrical circuit with $O(N)$ interconnections by using a single global inhibitory feedback gate (Tank and Hopfield 1986). In a biochemical system, a limited global resource, such as RNAP, can act to regulate all the DNA switches and thus similarly produce global inhibition. This effect is exploited by the simple transcriptional network shown in figure 2.6B, in which the output from each DNA switch activates the same DNA switch itself, and mutual inhibition is achieved by competition for RNAP. Specifically, we have switch templates D_{ii} with fixed thresholds set by DNA I_i, and D_{ii} produces A_i as its output RNA. With the instant binding assumption, we then derive the following equation (appendix 2.4.4):

$$\frac{dA_i^{tot}}{dt} = -\frac{E_d^{tot}}{1+L_d} \cdot \frac{k_{d,cat}}{K_{d,M}} A_i^{tot} + \frac{E^{tot}}{1+L} \left(\frac{k_{cat}}{K_M}[D_{ii}A_i] + \frac{k'_{cat}}{K'_M}[D_{ii}] \right) . \tag{2.5}$$

The production rate of A_i depends on A_i^{tot} and on L, while the degradation rate of A_i depends on A_i^{tot} and on L_d, as shown in figure 2.7. For a winner-take-all network, an ON state switch draws more RNAP than an OFF state switch (because of the smaller Michaelis constant for the ON state). Thus, if the other switches are turned OFF, the load on RNAP (L) becomes small, leading to faster production of the remaining ON switches. When the production rate curve and the degradation rate curve have three intersections, bistability is achieved such that the switches remain ON or OFF, depending on their current states.

For a simple, 2-switch system, a vector field on the phase plane illustrates that Michaelis–Menten type RNA polymerase reaction is essential for winner-take-all network implementation.

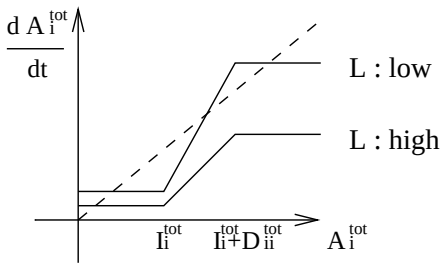


Figure 2.7. For WTA networks: Production rates (solid lines) for two different L's, compared to a linear degradation rate (dotted line).

Assume that the two switches are equivalent. The switch D_{ii} turns on as the amount of RNA signal A_i^{tot} increases form I_i^{tot} to $I_i^{tot} + D_{ii}^{tot}$. Because D_{ii}^{tot} is small compared to the threshold, we plot the transition region as lines in the phase plane. For a first-order RNAP reaction, one switch's production rate is independent of the other switch's production rate so that the vector is a mere sum of two switch's production rates (figure 2.8A). Thus, if individual switch being ON is stable, both switches being ON is also stable. For a Michaelis–menten RNAP reaction, however, the production rates are coupled through the load on RNA polymerase (L) such that one switch produces RNA faster when the other switch is OFF (figure 2.8B). Thus, it is possible that one switch being ON is stable, but both switches being ON is unstable (figure 2.8E).

For a first-order RNase reaction, the RNA activator's degradation rate linearly depends on its own concentration. Notice that the degradation vector always points to the origin, and the magnitude of the degradation vectors are the same at a fixed radius from the origin (figure 2.8C). For a Michaelis–Menten RNase reaction, there is a coupling of degradation rates through RNase load (L_d). If the load L_d is high, the RNase is at its maximum capacity such that the total degradation rate cannot increase even if the RNA concentration increases (figure 2.8D).

Enzyme saturation occurs when the concentration of substrates exceeds its Michaelis constant. Depending on the choice of template concentration (substrate for RNAP) and threshold (RNA signal, substrate for RNase, has to exceed this level for signal propagation), we can induce the saturation of RNA polymerase or RNase or both. The saturation of RNA polymerase helps the implementation of winner-take-all network. But the saturation of RNase does not help. On the contrary, it can stabilize the state of both switches being ON because the saturated RNase works at a similar rate as when only one switch is ON. An interesting possibility arises in a situation when RNA polymerase is not saturated but RNase is saturated. This can lead to an "all-or-none" circuit where both switches being ON is stable but one switch being ON is unstable.

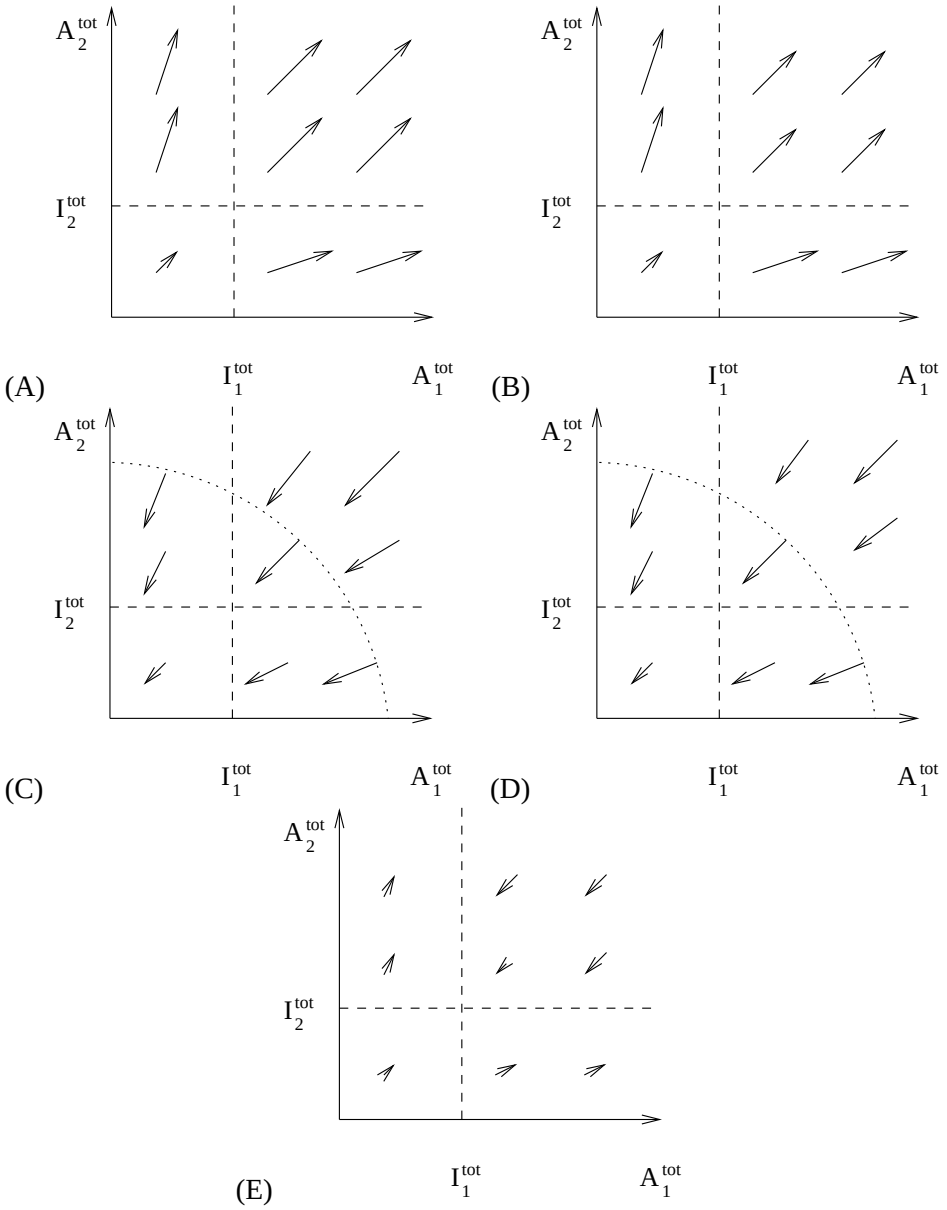


Figure 2.8. **(A)** Production rates by RNAP with a first-order approximation **(B)** Production rates by RNAP with the Michaelis–Menten reaction **(C)** Degradation rates by RNase with the first-order approximation **(D)** Degradation rates by RNase with the Michaelis–Menten reaction **(E)** Production rates minus degradation rates with the Michaelis–Menten RNAP reaction and a first-order RNase reaction

Consider n equivalent switches starting with initial activator concentrations above the threshold, and with the highest concentration at least δ above the rest (as a percentage). Analysis indicates that a less leaky system (small α) and sufficient differences in initial activator concentrations (large δ) can guarantee the existence of a unique winner (appendix 2.4.4). Simulations of a 10-switch winner-take-all network confirm this analysis, although we do not see perfect behavior (figure 2.9A).

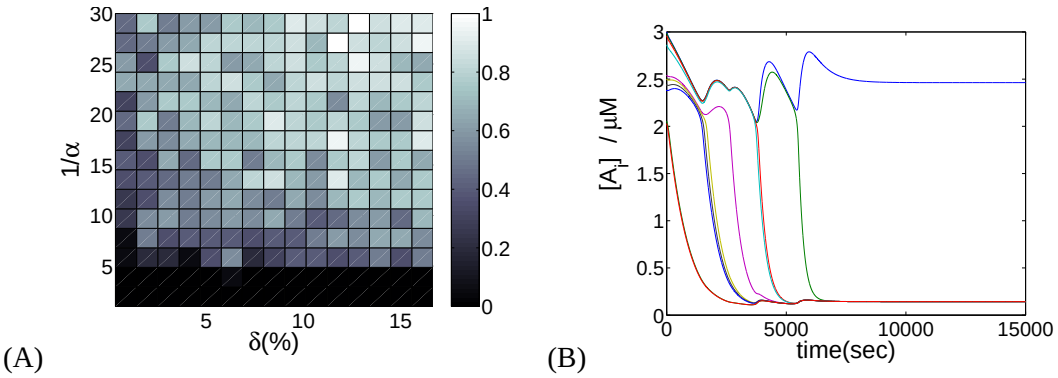


Figure 2.9. For WTA networks: **(A)** Empirical probability of correct output as a function of α and δ. **(B)** Time-course with $\delta = 0.33\%$ and $\alpha = 0.04$.

Figure 2.9B shows a time-course of a unique winner situation. Switches get turned OFF one by one whenever the activator level approaches the threshold, until only one switch remains ON.

Similarly, we can consider a k-WTA network where k winners persist. If we set the parameters appropriately such that k winners are stable but $k + 1$ winners are unstable, the simulation result recovered k winners most of the time. Even a single k-WTA gate can provide impressive computational power (Maass 2000).

2.3 Discussion

We have shown that if we treat transcriptionally controlled DNA switches as synapses and the concentrations of RNA species as the states of neurons, then the *in vitro* transcriptional circuit is equivalent to the neural network model and therefore can be programmed to carry out a wide variety of tasks. The structure of our biochemical networks differs from that of previous formal models of genetic regulatory circuits (Glass and Kauffman 1973; Mjolsness et al. 1991; Buchler et al. 2003). For example, consider the work of Buchler et al. (2003), which established a connection to the class of Boltzmann machines. There, the occupancy of regulatory binding sites corresponds to the state of neurons, the weights are set by the cooperative interaction among transcription factors, and the thresholds are the effective dissociation constants at a binding site. Thus, implementing a general N-unit neural network requires only $O(N)$ biochemical species, but up to $O(N^2)$ significant binding interactions must be encoded in the molecular sequences. Changing or tuning a network is therefore non-trivial. In contrast, in our transcriptional networks, each weight and threshold is represented by the continuously adjustable concentration of a distinct species, and the introduction or deletion of

any node is straightforward.

Each synapse is represented by a DNA switch with a single input–output specification, so the number of DNA switches grows as $O(N^2)$ for a fully recurrent neural network with N neurons (unlike the circuits of Buchler et al. (2003) which are linear). This constraint may be relieved because, in many networks of interest, most nodes have a small number of connections (Bray 2003; Reed 1993). The time for computation will increase as $O(N)$ due to finite hybridization rates because if the total concentration of all RNA signals is capped, the concentration of any given species will decrease as $1/N$. The weights are capped by the maximum gain of the system, which is the production rate divided by the degradation rate. Since the time constant of the network is the inverse of the degradation rate, if we wish to implement a network with large weights, we must increase the time constant.

We can analyze the cost of computing by considering basic physical chemistry. The energy consumption is about 20 kT (=10^{-19} J) per nucleotide incorporated, and 1 bit of information is encoded by a sequence containing tens of nucleotides. The encoding energy is large, since the molecule for each bit must contain specific instructions for connectivity, unlike spatially arranged digital circuits where a uniform physical signal carrier can be used. Furthermore, many copies (e.g., 10^{13} for a 1 μM signal in 20 μL) of a given species must be produced to change the concentration in a bulk sample. Worse yet, because degradation is not modulated in the transcriptional network, switching relies on selective change of production rates, thus continually using energy to maintain an ON state. Devising a scheme to minimize maintenance energy costs, such as in CMOS technology for electrical circuits, is an important problem.

The theory presented here is meant to serve as a guide for the construction of real biochemical computing networks. Naturally, real systems will deviate considerably from the idealized model (although perhaps less so than do neural-network models from real neurons). For example, hybridization is neither instantaneous nor irreversible, strands can have undesired conformations and crosstalk, and enzyme reactions depend on the sequence and are subject to side reactions that generate incomplete products. Some problems, such as hybridization speed and crosstalk, can be reduced by slowing the enzyme reactions and using proper sequence design (Dirks et al. 2004). Ultimately, some form of fault tolerance will be necessary at the circuit level. Restoration of outputs to digital values, achieved by any sufficiently high-gain sigmoidal activation function, provides some level of immunity to noise at the gate level, and attractor dynamics can provide restoration at the network level. A full understanding of fault tolerance in biochemical computing remains an important open

question.

Future directions include utilizing the versatility of active RNA molecules (such as aptamers, ribozymes, and riboswitches (Lilley 2003; Nudler and Mironov 2004)) for more general chemical input and output, devising a biochemical learning scheme analogous to neural network training algorithms (Mills Jr. et al. 1999), and studying the stochastic behavior of the transcriptional network when a very small number of molecules are involved in small volumes (Elowitz and Leibler 2000).

2.4 Appendix

2.4.1 Network equivalence

Analogous to a single switch case, we can derive the following with a fast and complete binding assumption:

$$\sum_i [D_{ij}A_j] = \frac{1}{2} D^{tot}_{*j}(\sigma(\hat{\Delta}_j) + 1).$$

The concentration of active and inactive switches will be proportional to the total concentration of the switches,

$$[D_{ij}A_j] = \frac{D^{tot}_{ij}}{D^{tot}_{*j}} \sum_i [D_{ij}A_j] = \frac{1}{2} D^{tot}_{ij}(\sigma(\hat{\Delta}_j) + 1) \qquad [D_{ij}] = \frac{1}{2} D^{tot}_{ij}(1 - \sigma(\hat{\Delta}_j)).$$

Substituting $\hat{\Delta}_i$ for Δ_i and applying the relation above to equation (2.1),

$$\frac{d\Delta_i}{dt} = \frac{1}{2} D^{tot}_{*i} \frac{d\hat{\Delta}_i}{dt} = -k_d \frac{1}{2} D^{tot}_{*i}(\hat{\Delta}_i + 1) + k_p \sum_j s_{ij} \frac{1}{2} D^{tot}_{ij}(1 + \sigma(\hat{\Delta}_j) + \alpha(1 - \sigma(\hat{\Delta}_j))).$$

Dividing both sides by $\frac{1}{2} k_d \cdot D^{tot}_{*i}$,

$$\frac{1}{k_d} \frac{d\hat{\Delta}_i}{dt} = -\hat{\Delta}_i + \frac{k_p}{k_d}(1 - \alpha) \sum_j s_{ij} \frac{D^{tot}_{ij}}{D^{tot}_{*i}} \sigma(\hat{\Delta}_j) + \frac{k_p}{k_d}(1 + \alpha) \sum_j s_{ij} \frac{D^{tot}_{ij}}{D^{tot}_{*i}} - 1.$$

2.4.2 Calculating biochemical constants from a neural network: Bias and load balance

We introduce switches D_{iB}'s without binding domains for A's such that they act as a bias unit. For example, D_{iB} has a full duplex promoter so that its production is constitutive. From equation (2.1),

we have an additional term for a bias unit

$$\frac{d\Delta_i}{dt} = -k_d \cdot \Delta_i + k_p \sum_j s_{ij}([D_{ij}A_j] + \alpha[D_{ij}]) + k_p s_{iB}[D_{iB}],$$

where $s_{iB} = 1$ if D_{iB} produces A_i and $s_{iB} = -1$ if D_{iB} produces I_i. Following similar derivation, θ_i becomes

$$\theta_i = -1 + \frac{k_p}{k_d}(1+\alpha)\sum_j s_{ij}\frac{D_{ij}^{tot}}{D_{*i}^{tot}} + 2\frac{k_p}{k_d}s_{iB}\frac{[D_{iB}]}{D_{*i}^{tot}}.$$

Assume that D_{*i}^{tot} is a constant for all i's with the introduction of null switch D_{0i}. Then,

$$\sum_i |w_{ij}| = \frac{k_p}{k_d}(1-\alpha)\sum_i \frac{D_{ij}^{tot}}{D_{*i}^{tot}} = \frac{k_p}{k_d}(1-\alpha)\frac{D_{*j}^{tot}}{D_{*i}^{tot}} = \frac{k_p}{k_d}(1-\alpha).$$

Thus, we can set $D_{ij}^{tot} = \beta|w_{ij}|$ for a β of our choice, and $D_{*i}^{tot} = D_{*j}^{tot} = \sum_i D_{ij}^{tot} = \beta\frac{k_p}{k_d}(1-\alpha)$. With α having a characteristic value depending on the system, we can determine k_p.

In summary,

$$\begin{aligned} k_d &= \frac{1}{\tau}, \\ k_p &= \frac{1}{\tau(1-\alpha)} max(\sum_i |w_{ij}|), \\ D_{ij}^{tot} &= \beta|w_{ij}|, \\ D_{*i}^{tot} &= \beta \cdot max(\sum_i |w_{ij}|), \\ s_{ij} &= sgn(w_{ij}). \end{aligned}$$

2.4.3 Michaelis–Menten enzyme reaction

We investigate the Michaelis–Menten enzyme reaction for the RNA polymerase. Consider the following reaction system,

$$\begin{aligned} E + D_{ij}A_j \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ED_{ij}A_j &\overset{k_{cat}}{\rightarrow} E + D_{ij}A_j + I_i/A_i, \\ E + D_{ij} \underset{k'_{-1}}{\overset{k'_1}{\rightleftharpoons}} ED_{ij} &\overset{k'_{cat}}{\rightarrow} E + D_{ij} + I_i/A_i, \end{aligned}$$

where the steady-state assumption requires $k_1[E][D_{ij}A_j] = (k_{-1} + k_{cat})[ED_{ij}A_j],\ k'_1[E][D_{ij}] = (k'_{-1} + k'_{cat})[ED_{ij}]$ so that,

$$\begin{aligned}[ED_{ij}A_j] &= \frac{[E][D_{ij}A_j]}{K_M} \quad K_M = \frac{k_{-1} + k_{cat}}{k_1}, \\ [ED_{ij}] &= \frac{[E][D_{ij}]}{K'_M} \quad K'_M = \frac{k'_{-1} + k'_{cat}}{k'_1}.\end{aligned}$$

Assume that all $D_{ij}A_j$ have K_M, k_{cat}, while all D_{ij} have K'_M, k'_{cat}. Then,

$$E^{tot} = [E] + \sum_{i,j}[ED_{ij}A_j] + \sum_{i,j}[ED_{ij}] = [E](1 + \sum_{i,j}\frac{[D_{ij}A_j]}{K_M} + \sum_{i,j}\frac{[D_{ij}]}{K'_M}).$$

Thus,

$$\begin{aligned}\frac{d\Delta_i}{dt} &= -k_d \cdot \Delta_i + \sum_j s_{ij}(k_{cat}[ED_{ij}A_j] + k'_{cat}[ED_{ij}]) \\ &= -k_d \cdot \Delta_i + [E]\sum_j s_{ij}(\frac{k_{cat}}{K_M}[D_{ij}A_j] + \frac{k'_{cat}}{K'_M}[D_{ij}]) \\ &= -k_d \cdot \Delta_i + \frac{E^{tot}}{1 + \sum_{i,j}\frac{[D_{ij}A_j]}{K_M} + \sum_{i,j}\frac{[D_{ij}]}{K'_M}}\sum_j s_{ij}(\frac{k_{cat}}{K_M}[D_{ij}A_j] + \frac{k'_{cat}}{K'_M}[D_{ij}]).\end{aligned}$$

This is equivalent to equation (2.1) if the load on RNA polymerase $L = \sum_{i,j}\frac{[D_{ij}A_j]}{K_M} + \sum_{i,j}\frac{[D_{ij}]}{K'_M}$ is held constant. The following relations to the rate constants of first-order reactions are obtained:

$$k_p = \frac{E^{tot}}{1 + L}\frac{k_{cat}}{K_M} \qquad \alpha \cdot k_p = \frac{E^{tot}}{1 + L}\frac{k'_{cat}}{K'_M}.$$

If $K_M \simeq K'_M$, L can be considered to be constant: $L = \frac{1}{K_M}\sum_{i,j}([D_{ij}A_j] + [D_{ij}]) \simeq \frac{D^{tot}}{K_M}$. However, this is unlikely if the binding of A to D invokes a change in the promoter region. Then, introduction of a new kind of switch $\bar{D}_{ij}$ is necessary with a different set of Michaelis constants. The idea is that $\bar{D}$ draws more RNA polymerase than $\bar{D}A$ such that the the load on RNA polymerase is the same whether the switches are ON or OFF. The transcripts from $\bar{D}_{ij}$ are unrelated to the signal sequences in the system.

$$[E\bar{D}_{ij}A_j] = \frac{[E][\bar{D}_{ij}A_j]}{\bar{K}_M} \quad [E\bar{D}_{ij}] = \frac{[E][\bar{D}_{ij}]}{\bar{K}'_M}$$

Let $\bar{D}_{ij}^{tot} = rD_{ij}^{tot}$. Then, we can derive

$$L = \sum_{i,j}(\frac{[D_{ij}A_j]}{K_M} + \frac{[\bar{D}_{ij}A_j]}{\bar{K}_M} + \frac{[D_{ij}]}{K'_M} + \frac{[\bar{D}_{ij}]}{\bar{K}'_M}) = \sum_{i,j}((\frac{1}{K_M} + \frac{r}{\bar{K}_M})[D_{ij}A_j] + (\frac{1}{K'_M} + \frac{r}{\bar{K}'_M})[D_{ij}]).$$

If there exists an $r > 0$ such that $\frac{1}{K_M} + \frac{r}{\bar{K}_M} = \frac{1}{K'_M} + \frac{r}{\bar{K}'_M} = C$, then L is a constant by the following approximation: $L = C\sum_{i,j}([D_{ij}A_j] + [D_{ij}]) \simeq C \cdot D^{tot}$. Since $r = \frac{\frac{1}{K_M} - \frac{1}{K'_M}}{\frac{1}{\bar{K}'_M} - \frac{1}{\bar{K}_M}}$, the existence of r is guaranteed, if $K_M - K'_M$ and $\bar{K}_M - \bar{K}'_M$ are of the opposite sign. Note that, with the introduction of $\bar{D}_{ij}$, we effectively increased the amount D_{*i}^{tot} by a factor of $1 + r$.

Next, we investigate the Michaelis–Menten enzyme reaction for the RNase. Let RNase have Michaelis constant $K_{d,M}$ and catalytic constant $k_{d,cat}$ for I_j, A_j, A_jI_j, and $D_{ij}A_j$. For example,

$$\begin{aligned} E_d + A_i &\underset{k_{d,-1}}{\overset{k_{d,1}}{\rightleftharpoons}} E_dA_i \overset{k_{d,cat}}{\rightarrow} E_d, \\ [E_dA_i] &= \frac{[E_d][A_i]}{K_{d,M}} \quad K_{d,M} = \frac{k_{d,-1} + k_{d,cat}}{k_{d,1}}. \end{aligned}$$

Following similar derivation, we get

$$k_d = \frac{E_d^{tot}}{1 + L_d}\frac{k_{d,cat}}{K_{d,M}},$$

where $L_d = \sum_{i,j}\frac{[A_j]+[I_j]+[A_jI_j]+[D_{ij}A_j]}{K_{d,M}}$ is the load on RNase.

2.4.4 Winner-take-all network

Consider a reaction system where each switch activates itself by producing its own activator and all switches compete for RNA polymerase. We have D_{ii} and I_i as DNA, and $D_{ii}A_i$ or D_{ii} produces A_i as its output RNA. The only RNA species being produced and degraded are A_i's. With a first-order approximation, we can easily derive

$$\frac{dA_i^{tot}}{dt} = -k_d \cdot A_i^{tot} + k_p([D_{ii}A_i] + \alpha[D_{ii}]).$$

Applying the relation in Equation 4 for a Michaelis–Menten enzyme reaction, we get

$$\frac{dA_i^{tot}}{dt} = -\frac{E_d^{tot}}{1 + L_d}\frac{k_{d,cat}}{K_{d,M}}A_i^{tot} + \frac{E^{tot}}{1 + L}(\frac{k_{cat}}{K_M}[D_{ii}A_i] + \frac{k'_{cat}}{K'_M}[D_{ii}]),$$

where $L = \sum_i \frac{[D_{ii}A_i]}{K_M} + \sum_i \frac{[D_{ii}]}{K'_M}$ is the load on RNA polymerase and $L_d = \sum_i \frac{A_i^{tot}}{K_{d,M}}$ is the load on RNase.

Analysis of k-WTA network

For n switches, we set $D_{ii}^{tot} = Dt$ and $I_i^{tot} = It$ as constants. Without loss of generality, assume $A_1^{tot} > A_2^{tot} > A_3^{tot} > \cdots > A_n^{tot}$. We are interested in finding a condition where k winners $A_1, A_2, \ldots, A_k$ persists. Then, a WTA network is a special case with $k = 1$. Starting from all switches being ON, the system will reach a state where some of the activator levels fall below the threshold, turning off those switches.

We proceed by showing the following three properties of the biochemical WTA network.

1. The ordering of A_i^{tot} is preserved.

2. k winners are stable, but $k + 1$ winners are unstable.

3. The derivative of A_k^{tot} is non-negative when A_k^{tot} is at the threshold.

Let $k_{mp} = \frac{E^{tot}}{1+L}\frac{k_{cat}}{K_M}$, $k_{md} = \frac{E_d^{tot}}{1+L_d}\frac{k_{d,cat}}{K_{d,M}}$. Unlike k_p and k_d in the first-order reactions, these global parameters are time-varying as a function of network state. With a simplifying assumption $k_{cat} = k'_{cat}$, it is shown that $\frac{1}{K'_M} = \alpha\frac{1}{K_M}$. Thus,

$$\frac{dA_i^{tot}}{dt} = -k_{md} \cdot A_i^{tot} + k_{mp}(Dt + (\alpha - 1)[D_{ii}]).$$

It is easy to derive property 1 from this equation. For every ON switch, the ordering determined by initial activator concentrations cannot change because the activator levels exponentially decay to the same steady-state value $A_{ON,ss} = \frac{k_{mp}}{k_m d}Dt$. In case $A_{ON,ss} < It + Dt$, the activator levels fall below the threshold in the same order. The switch whose activator level falls below the threshold has decreased production rate than ON switches. These OFF switches will approach the steady-state value $A_{OFF,ss} = \alpha\frac{k_{mp}}{k_m d}Dt$.

We need to set the concentrations properly for property 2. If there are k winners and $n - k$ losers, at steady-state, $A_1^{tot} = A_2^{tot} = \cdots = A_k^{tot} \geq It + Dt$ and $A_{k+1}^{tot} = A_{k+2}^{tot} = \cdots = A_n^{tot} = \alpha A_1^{tot} \leq It$. Thus, considering the load on enzymes,

$$A_1^{tot} = \frac{k_{mp}}{k_{md}}Dt = \frac{1}{\gamma} \cdot \frac{K_{d,M} + kA_1^{tot} + (n-k)\alpha A_1^{tot}}{K_M + kDt + (n-k)\alpha Dt}Dt,$$

where $\frac{1}{\gamma} = \frac{k_{cat}E^{tot}}{k_{d,cat}E_d^{tot}}$. Equivalently,

$$\frac{A_1^{tot}}{K_{d,M}} = \frac{1}{\gamma} \cdot \frac{1 + k\frac{A_1^{tot}}{K_{d,M}} + (n-k)\alpha\frac{A_1^{tot}}{K_{d,M}}}{\frac{K_M}{Dt} + k + (n-k)\alpha}.$$

There are two unitless parameters $t_k = \frac{A_1^{tot}}{K_{d,M}}$, and $x = \frac{K_M}{Dt}$. The former is the ratio of the steady-state of winners to the Michaelis constant of RNase and the latter is the ratio of the Michaelis constant of RNAP to the amount of DNA templates. The subscript of t denotes the number of winners. Substituting $\frac{1}{\gamma}$ and rearranging terms,

$$\frac{t_{k+1}}{t_k} = \frac{x + k + (n-k)\alpha}{x + k + (n-k)\alpha + (1-\alpha)(1 - x \cdot t_k)}.$$

A necessary condition is $\frac{t_{k+1}}{t_k} < 1$, thus $x \cdot t_k < 1$. By choosing proper γ, Dt, and It, let $t_k = \frac{It+Dt}{K_{d,M}}$ such that the steady-states for k winners are at the threshold $It + Dt$. For $k+1$ winners, the steady-states are less than the threshold $It + Dt$, contradicting the assumption that the switch $k+1$ is completely ON at steady-state. From the necessary condition, we see that Dt must be relatively large compared to K_M, while $It + Dt$ must be relatively small compared to $K_{d,M}$. That is, RNAP should be more saturated than RNase. If the gain is high ($\alpha \simeq 0$) and x is sufficiently small, the above equation reduces to $\frac{t_{k+1}}{t_k} = \frac{k}{k+1} < 1$. Consequently, an arbitrary choice of k is possible.

To show property 3, we want $\frac{dA_k^{tot}}{dt} \geq 0$ when $A_k^{tot} = It + Dt$. Then, we need

$$f = \frac{k_{mp} \cdot Dt}{k_{md}(It + Dt)} = \frac{1}{\gamma} \cdot \frac{K_{d,M} + \sum_i A_i^{tot}}{K_M + \sum_i (Dt + (\alpha - 1)D_{ii})} \frac{Dt}{It + Dt} \geq 1.$$

Assume that the concentrations satisfy property 2. A worst-case scenario is when switches $1, \ldots, k+1$ have similar activator levels quite far from the rest, switches $k+2, \ldots, n$. For example, switches $k+2, \ldots, n$ have activator levels at OFF steady-state value and switches $1, \ldots, k$ have activator levels at the threshold, minimizing the numerator. Let $A_{k+1}^{tot} = It + Dt - \epsilon \cdot Dt$ ($0 < \epsilon < 1$) such that the switch $k+1$ is a fraction of ϵ OFF.

$$f(\epsilon) = \frac{1}{\gamma} \frac{\frac{1}{t_k} + n + (\alpha - 1)(n - k - 1 + \epsilon\frac{Dt}{It+Dt})}{x + n + (\alpha - 1)(n - k - 1 + \epsilon)} \geq 1$$

When $\epsilon = 0$, it is impossible to differentiate switch $k+1$ from switch k, and consequently, the k-WTA network fails. Solving the above inequlity, we get $\epsilon \geq \frac{\gamma - 1}{\gamma - \frac{Dt}{It+Dt}}$ in the worst case. Usually,

$\gamma \simeq 1$ and $\frac{Dt}{It+Dt} \ll 1$ such that this condition is satisfied given enough differences in initial activator concentrations.

Chapter 3 A Bistable Circuit

3.1 Results

The synthetic DNA template design is modular with easily programmable connectivity dictated by Watson–Crick base-pairing rules. The regulatory domain is upstream of the promoter region; the output domain is downstream of the promoter region. This separation of domains allows us to design DNA templates that have any desired connectivity. Regulated DNA templates are called switches ("Sw"), whereas unregulated DNA templates are called sources ("So"). A switch can assume two different conformations with different transcription efficiency: ON or OFF (Figure 1A). The OFF state of the switch consists of a double-stranded DNA template ("T") with a partially single-stranded (ss) and thus incomplete promoter region. Similar templates are known to transcribe poorly (Martin and Coleman 1987). The switch is turned on by the addition of a ssDNA activator ("A") that completes the promoter region. Templates with nicked promoters ("T·A") have been found to transcribe well, approximately half as efficiently as fully double-stranded sources (data not shown, a promoter structure with a nick have been studied (Jiang et al. 2001)). The activator contains a "toehold," a single-stranded overhang beyond the helical domain it forms with the DNA template, where an inhibitor can bind to initiate a toehold-mediated strand displacement reaction (Yurke and Mills 2003). Thus, the switch can be turned off upon addition of an inhibitor strand (either ssRNA, "I," or ssDNA, "dI"). An ON state source template has a complete promoter sequence with a nick and an OFF state source template is missing 5 bases of the promoter sequence on the template side. Source templates do not interact with activators due to the hairpin stem permanently covering the branch migration sequence, and therefore maintain their transcription efficiency in the presence of inhibitors. Due to the identical structures in the promoter region (17 bases colored blue in figure 3.1A), an ON or OFF source has the same transcription speed as an ON or OFF switch with less than 10% deviation (data not shown). In a typical reaction network, the RNA inhibitor strands will be produced by RNAP from upstream templates using NTP as fuel and will be degraded by RNase H.

An important goal for our circuits is to obtain switches with an ultrasensitive response, i.e., a sharp threshold. Several alternative mechanisms can give rise to the ultrasensitive response in biological circuits, for example, cooperative binding of regulatory proteins and multistep phospho-

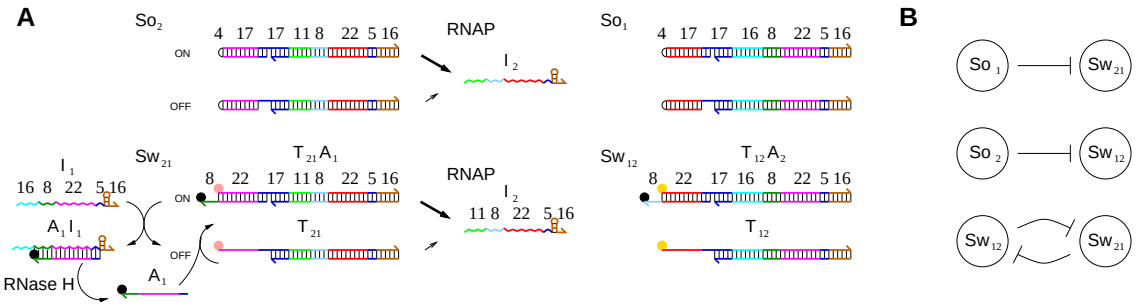


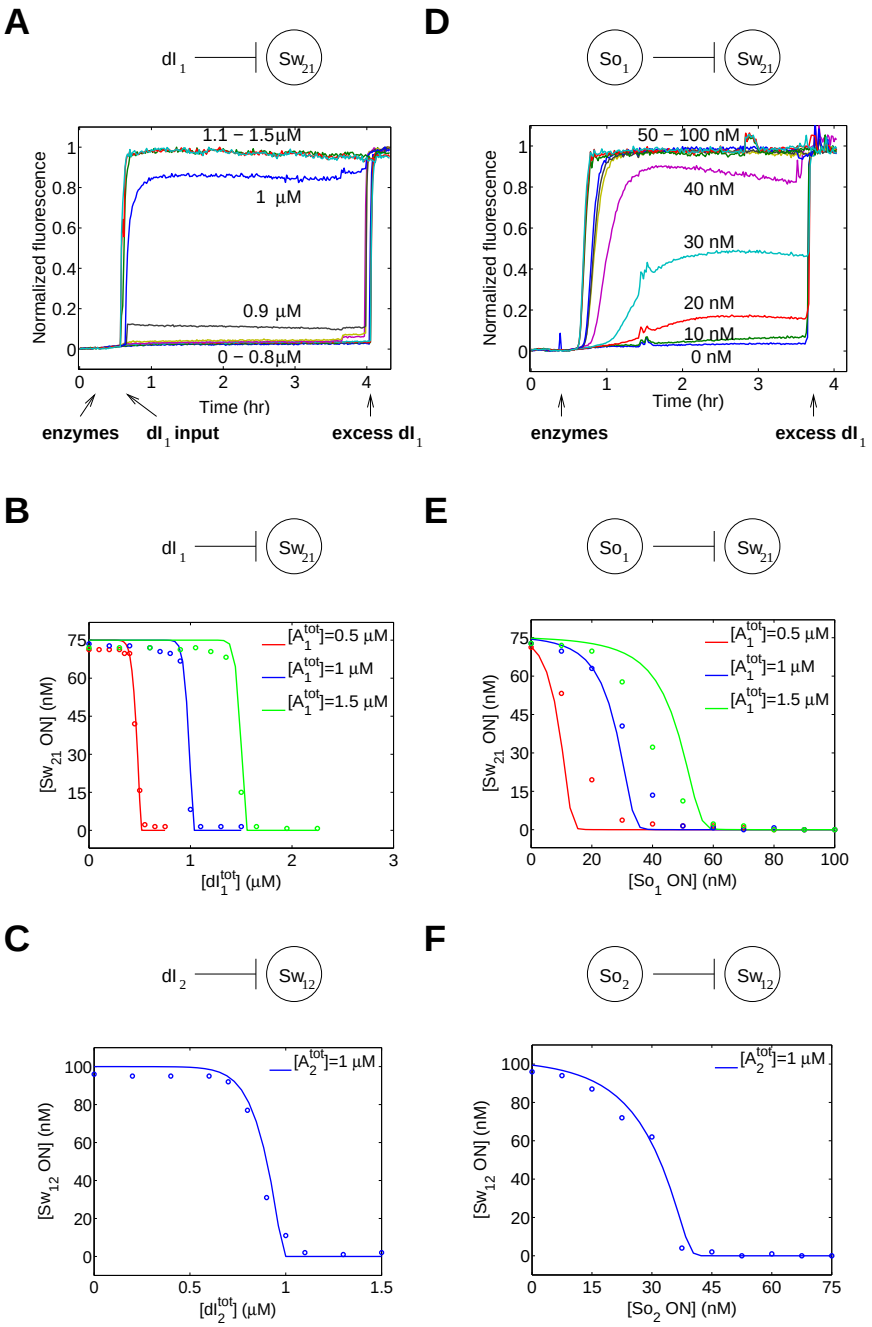
Figure 3.1. Design of synthetic switches and circuits. **(A)** Design of DNA sequences and reaction mechanisms. Each sequence subdomain is color-coded with the number of bases marked above. For example, 62 base long inhibitor I_2 consists of five parts: spacer sequence including 6 base initiation sequence (11 bases, green), toehold-binding sequence (8 bases, light blue), branch migration sequence (22 bases, red), 5' end of the promoter sequence (5 bases, blue) and 3' end hairpin structure (16 bases, brown). Switch templates have two distinct states, ON or OFF, with different transciption speed. The ON state switch template (T·A complex) has complete promoter sequence (17 bases, blue) with a nick and the OFF state switch template (T) is missing 5 bases of the promoter sequence on the template side. The two single strands that form the switch templates are the longer non-template side strand, T-nt, and the shorter template side strand, T-t (materials and methods). The templates T's are labeled with fluorophores (pink circle, Texas Red; yellow circle, TAMRA) and activators A's are labeled with quenchers (black circles) such that the state of each switch can be monitored by measuring the fluorescence quenching efficiency (Marras et al. 2002). The ON state source template has a complete promoter sequence with a nick (like the ON state switch) and the OFF state source template is missing 5 bases of the promoter sequence on the template side (like the OFF state switch). Unlike the switch templates, source templates do not interact with activators, due to the hairpin stem permanently covering the branch migration sequence, and therefore maintain their transcription efficiency in the presence of inhibitors. The source activity can be controlled by preparing a mixture of ON and OFF source templates. The two single strands that form the source templates are the longer non-template side strand, So-nt (ON or OFF), and the shorter template side strand, T-t; the template side strands are the same for the switch and source templates that encode the same outputs (materials and methods). For any given transcriptional circuit, we use either source template So_j (with some desired mixture of ON and OFF templates) or switch template Sw_{ji} to produce RNA inhibitor I_j. RNAP produces RNA inhibitors from DNA templates, while RNase H degrades RNA inhibitors bound to DNA activators. Detailed reaction mechanisms are listed in appendix 3.4. **(B)** Two feedforward circuits where a source So_j controls a switch Sw_{ij} by supplying inhibitor I_j and a bistable system where two switches, Sw_{12} and Sw_{21}, inhibit each other.

rylation of target protein by kinases (Ferrell 1996). We use competitive binding of nucleic acid species rather than cooperative binding to achieve ultrasensitivity. Our approach is closely related to "inhibitor ultrasensitivity," where a stoichiometric inhibitor to the activating enzyme is used. Similar mechanisms have been suggested for regulation of mitosis (Thron 1994) and sporulation (Voigt et al. 2005). The threshold in our transcriptional circuit derives from three types of strong DNA and RNA hybridization reactions (appendix 3.4), which we call activation, annihilation, and inhibition. An activator binds to an OFF switch template to turn the switch on (activation); an activator binds

to an inhibitor and is not available for the switch template (annihilation); an inhibitor displaces an activator from an ON switch template, T·A complex, to turn the switch off (inhibition). The key requirements for the inhibition mechanism are that the activator-inhibitor binding is thermodynamically more favorable than the template-activator binding, and that there is a fast kinetic pathway to the lowest energy state (in our case, toehold-mediated branch migration (Yurke and Mills 2003)). Since the activator and inhibitor annihilate each other, the difference of total activator and inhibitor concentrations is the most important determinant of the state of switch: an excess of inhibitor will turn the switch off while an excess of activator will turn the switch on. All three mechanisms are needed for fast switch response.

Sequences of the synthetic DNA templates are chosen to minimize alternative folding (Flamm et al. 2000) and spurious interactions (Seeman 1982). Various domain lengths have been experimentally tested for functionality (figure 3.1A); for example, the binding domains of an OFF switch template to an activator (27 bases) and of an activator to an inhibitor (35 bases) are long enough to ensure the activation and annihilation mechanisms, while the toehold of an activator (8 bases) is long enough to facilitate the inhibition mechanism without being so long as to reduce ON-state transcription efficiency (appendix 3.4). The 3' end hairpin structure (16 bases) increases copy number and also decreases self-coded extension of RNA transcripts by RNAP (Triana-Alonso et al. 1995). An OFF switch template has only 5 bases missing in the promoter region, which permits leaky expression. However, increasing the extent of activator binding to the promoter domain can cause a spurious binding between non-matching activator and template pairs (appendix 3.4). The specific, strong, and repeatable hybridization of template-activator pairs and activator-inhibitor pairs has been confirmed (figures 3.6 and 3.7).

To determine whether the proposed hybridization mechanism leads to a sharp threshold, the transfer curves of individual switches were measured with the total concentration of DNA inhibitors as inputs (figure 3.2B and C). For the total concentration of certain species, we consider both isolated species and complexes containing that species. Thus, for the DNA inhibitor, [dI^{tot}] = [dI] + [dI·A]. The DNA inhibitor is a permanent input signal because it is not degraded by RNase H. The switch activity, defined as the concentration of ON switch template, [T·A], is measured in real-time using fluorescence: an OFF switch template T is labeled with a fluorophore and an activator A is labeled with a quencher such that the fluorescence of OFF switch T is high but the fluorescence of ON switch T·A is low due to fluorescence quenching (Marras et al. 2002). The time-course of Sw_{21} inhibited by DNA inhibitor dI_1 with 1 μM activator A_1 is shown (figure 3.2A). The fluorescence signal is stable



in the presence of enzymes. When the DNA inhibitor dI_1 input is less than 0.9 μM, the fluorescence signal remains low. On the other hand, the fluorescence signal quickly reaches maximal value upon addition of more than 1 μM dI_1 input. The transfer curves are constructed by measuring normalized fluorescence signal, which is taken as the proportion of OFF state switch among the total switch template. We have shown that the correspondence of normalized fluorescence to the switch state

Figure 3.2. Characterization of switches and feed-forward circuits. The total concentration of the source So_1 or the switch template T_{12} is 100 nM, and the total concentration of the source So_2 or the switch template T_{21} is 75 nM. Downstream activator concentration is 1 μM. The ratio of maximum to minimum fluorescence signals before normalization is greater than 10 for both TAMRA and Texas Red dyes. **(A)** Normalized fluorescence time-courses for the DNA inhibitor dI_1 inhibiting switch Sw_{21} with the total concentration of activator A_1 at 1 μM. The enzymes are added at 10 minutes and different amounts of DNA inhibitor dI_1 inputs are added at 35 minutes. Additional DNA inhibitor dI_1 is added at 4 hours to generate maximal fluorescence levels. The concentrations of DNA inhibitor dI_1 inputs are marked on the time-courses. **(B and C)** The normalized fluorescence signals immediately prior to the addition of excess dI as shown in (A) are used for the construction of tranfer curves. Experimental data points are plotted as circles and the model fits (see appendix 3.4 for model and parameters) are plotted as lines. **(B)** The transfer curves of switch Sw_{21} with the total concentration of DNA inhibitor dI_1 as inputs. **(C)** The transfer curve of switch Sw_{12} with the total concentration of DNA inhibitor dI_2 as inputs. **(D)** Normalized fluorescence time-courses for the feedforward circuit of source So_1 inhibiting switch Sw_{21} with the total concentration of activator A_1 at 1 μM. The enzymes are added at 20 minutes and excess DNA inhibitor dI_1 is added at 210 minutes to generate maximal fluorescence levels. The concentrations of ON state source So_1 are marked on the time-courses. **(E and F)** The normalized fluorescence signals immediately prior to the addition of excess dI as shown in (D) are used for the construction of transfer curves. Experimental data points are plotted as circles and the model fits are plotted as lines. **(E)** The transfer curves of switch Sw_{21} with the source So_1 activity as inputs. **(F)** The transfer curve of switch Sw_{12} with the source So_2 activity as inputs.

is quantitative, as discussed below. The role of the activator as an adjustable threshold is verified by the transfer curves of switch Sw_{21} with different switching thresholds dependent on the total concentration of activator A_1 (figure 3.2B). To test programmability of our synthetic switch design, we swap the input and output domain of switch Sw_{21} to create the switch Sw_{12}. The response of switch Sw_{12} is similar to that of switch Sw_{21}, flat at low and high levels of DNA inhibitor dI_2, yet sensitive when the total concentration of DNA inhibitor dI_2 is close to the total concentration of activator A_2, 1 μM (figure 3.2C).

This tunable sigmoidal curve has the piecewise-linear shape predicted by the model (appendix 3.4), which uses a single parameter set for all model fitting results shown in this paper. To compare the sharpness of the transition achieved by our competitive inhibition mechanism with that of other biological mechanisms such as binding cooperativity, the transfer curves were also fit to the following Hill equation:

$$y = y_{min} + \frac{y_{max} - y_{min}}{1 + \left(\frac{x}{K}\right)^n}, \tag{3.1}$$

where y_{min} is the minimum switch activity, y_{max} is the maximum switch activity, x is the total concentration of DNA inhibitor, n is the Hill coefficient, and K is the total concentration of DNA

inhibitor required for half repression. The Hill coefficient for the switch Sw_{12} transfer curve is 14.1, and those for the switch Sw_{21} transfer curves are 18.6, 30.0, and 32.9, respectively. The transfer curve of the switch Sw_{12} is not as sharp as the transfer curve of the switch Sw_{21}, possibly due to an unexpected secondary structure in activator A_2 or in the single-stranded region of switch template T_{12}, which could interfere with the binding of A_2 to T_{12}. Nonetheless, the Hill coefficients are much higher than most biological repressors and are adjustable by changing the total concentration of activators. We can understand the change of Hill coefficients as follows: increasing total activator concentration increases the threshold, i.e., K, yet the switching width in terms of DNA inhibitor concentration change remains constant because it depends on the total concentration of switch template (appendix 3.4). Consequently, the switching width becomes narrower relative to the threshold as we increase the threshold, resulting in a higher Hill coefficient, n. The change of sharpness is clear when the inputs are scaled by the half-repression points (figure 3.8A).

Next, to determine whether the switches have a sharp threshold with RNA input signals, the transfer curves of switch Sw_{ij} driven by an RNA inhibitor I_j were measured (figure 3.2E and F). Unlike DNA inhibitors which serve as permanent input signals, the RNA inhibitors are continuously being produced and degraded by the enzymes. This dynamic control of regulatory signals is necessary to achieve dynamic behavior within circuits where individual switches change their states in response to inputs and states of other switches. An upstream source So_j is used to achieve a steady-state RNA inhibitor I_j level and the source So_j activity is controlled by preparing a mixture of ON and OFF source templates. We keep the total concentration of ON and OFF source templates the same, while tuning the source activity, to simulate the continuous tuning of an upstream switch activity. Note that when source template concentrations are sufficiently high, RNA inhibitor production by RNAP will exceed RNase H's capacity for degradation, and RNA inhibitor levels will increase without bound rather than achieve a steady-state. In such cases, however, switch activity nonetheless approaches an asymptotic "steady-state" value. An additional complicating factor is that over the course of several hours, both RNAP and RNase H activities decrease due to depletion of NTPs, change of buffer composition, and other effects, thus altering the instantaneous "steady-state" level. In this paper, we use the term in this loose sense, acknowledging such effects.

The time-course of Sw_{21} inhibited by So_1 with 1 μM activator A_1 is shown (figure 3.2D). Unlike the dI-triggered switch state changes, here the fluorescence signal changes do not occur immediately after enzyme addition because the production of RNA inhibitor I_1 takes time and inhibitor I_1 is mostly consumed by the free activator A_1 initially. For low source activities, no detectable flu-

orescence change ensued. As the source So_1 activity increases, the steady-state concentration of inhibitor I_1 increases and turns the switch Sw_{21} off. As with the DNA input, thresholds are determined by the the total concentration of activator; three thresholds were demonstrated for switch Sw_{21} (figure 3.2E). The transitions are not as sharp as with DNA inhibitors due to constant turnover of RNA inhibitors bound to activators by RNase H. This breakdown pathway of the activator-inhibitor complex partially reverses the annihilation and inhibition mechanisms necessary to establish a sharp transition (appendix 3.4). The transfer curves were fit to the Hill equation (3.1) where x is the upstream source activity, and K is the upstream source activity required for half repression. The Hill coefficient for the switch Sw_{12} transfer curve is 5.17, and those for the switch Sw_{21} transfer curves are 3.09, 5.40, and 5.96, respectively. Although lower than for DNA inhibitors, the Hill coefficients are comparable to that of a two-stage synthetic biological repressor cascade (Hooshangi et al. 2005). The change of sharpness is clear when the inputs are scaled by the half-repression points (figure 3.8B).

To confirm that the fluorescence read-out reflects the actual molecular state of the system, gel-based experiments were performed for switch Sw_{21} regulated by source So_1 with the total concentration of activator A_1 at 1 μM. In this experiment only, fluorophore- (rather than quencher-) labeled activators were used for easy identification in the gel. In the denaturing gel (figure 3.3C), all DNA and RNA species migrate as single strands such that major bands can be identified based on length. We measured the total concentration of inhibitors, the sum of the concentration of free floating inhibitor, [I], and the concentration of activator-inhibitor complex, [A·I], as a function of source So_1 activity (figure 3.3B). In the non-denaturing gel (figure 3.3D), the concentrations of activator-inhibitor complexes (figure 3.3B) and the concentration of OFF switch template, $[T_{21}]$, are measured as a function of source So_1 activity. For the total concentration of RNA inhibitor I_j less than the total concentration of A_j, $[A_jI_j]$ is the same as $[I_j^{tot}]$ because of the strong binding interactions of activator-inhibitor pairs. The ON switch template, $T_{21}{\cdot}A_1$, had low fluorescence and was not clearly identified in lanes 1 to 3, presumably due to fluorescence resonance energy transfer (Marras et al. 2002) from Texas Red on T_{21} to Cy5 on A_1 (Figure 9). We used $[T_{21}^{tot}]$-$[T_{21}]$ in place of $[T_{21}A_1]$ in figure 3.3A. The switch states measured in the non-denaturing gel agree with the switch states measured by fluorescence quenching in the fluorometer. Interestingly, we observed activator bands that migrate slower than free activators but faster than the activator-inhibitor complexes in control lanes (bracket in figure 3.3D). We interpret this as activators binding to a mixture of incomplete inhibitors (abortive transcripts or incomplete degradation products, figure 3.10) shown

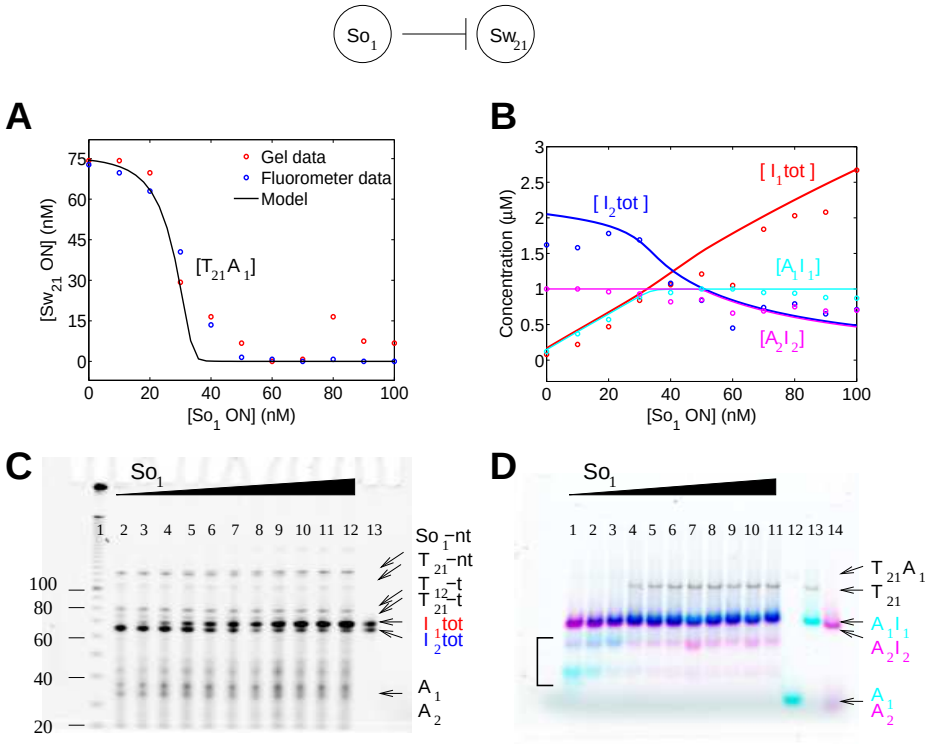


Figure 3.3. Characterization of a feed-forward circuit with gel. The feedforward circuit of source So_1 inhibiting switch Sw_{21} is characterized with both denaturing and non-denaturing gels. The total concentration of the source is 100 nM, and the total concentration of the switch template T_{21} is 75 nM. The total concentrations of activator A_1 and A_2 are 1 μM. **(A)** The non-denaturing gel data (red circles) and fluorometer data (blue circles) for the switch Sw_{21} states are compared. The gel data is quantitated by measuring the fluorescence of OFF switch T_{21} band: $[T_{21}A_1] = [T_{21}^{tot}] - [T_{21}]$. The fluorometer data and the model fit (black line) are the same as that of blue transfer curve in figure 3.2E. **(B)** The total concentrations of inhibitors are measured from the denaturing gel and the concentrations of activator-inhibitor complexes are measured from the native gel. Experimental data are plotted as circles and model fits are plotted as lines. **(C)** Denaturing gel stained with SYBR gold. Lane 1 contains a 10-base ladder, lanes 2 through 12 contain samples from the feedforward circuit with the concentration of ON source So_1 template increasing from left to right. Lane 13 contains purified inhibitors, I_1 and I_2. Two DNA strands, So_1-nt (either ON or OFF) and T_{12}-t, form the source So_1 template (either ON or OFF), while two DNA strands, T_{21}-nt and T_{21}-t, form the switch template, T_{21}. **(D)** For the non-denaturing gel, three distinct excitation and emission scan results are overlayed (Texas Red-labeled T_{21}: gray; Cy5-labeled A_1: cyan; and FAM-labeled A_2: magenta). Individual scan results are shown in figure 3.9. Lanes 1 through 11 contain samples from the feedforward circuit with the concentration of ON source So_1 template increasing from left to right. Lane 12 contains the ON switch template $T_{21}A_1$ and the activator A_1, lane 13 contains the OFF switch template T_{21} and the activator-inhibitor complex A_1I_1, and lane 14 contains the activator-inhibitor complex A_2I_2 and the activator A_2.

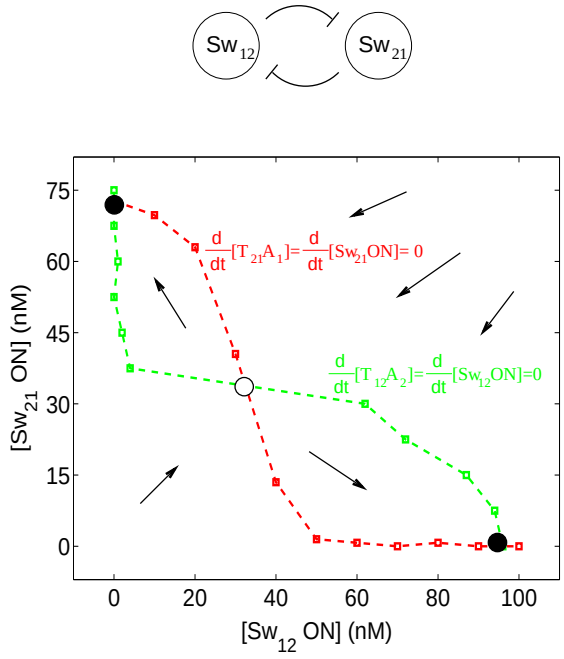


Figure 3.4. Nullclines of the mutually inhibitory circuit as inferred from the transfer curves of feedforward circuits. It is assumed that the source So_i and the switch Sw_{ij} activities are equivalent and that the steady-state behavior is achieved instantaneously. With these assumptions, the nullclines of mutually inhibitory circuit are the same as the empirical transfer curves of feedforward circuits in figure 3.2E and F. Here, two blue curves of figure 3.2E and F are used for the case of both activators at 1 μM. The experimental data points are plotted as squares with dotted lines drawn to guide the eye. The nullclines ($\frac{d}{dt}[\mathrm{T_{12}A_2}] = 0$ and $\frac{d}{dt}[\mathrm{T_{21}A_1}] = 0$) intersect at three points because of the sigmoidal shape. Possible vector flow directions are shown as arrows. Open circle: unstable fixed point; dark circles: stable fixed points.

as smearing less than 40 nucleotide long in the denaturing gel. We counted these bands as free activators and excluded incomplete inhibitor bands for our simple model. In the non-denaturing gel, we did not observe spurious binding complexes containing switch template $\mathrm{T_{21}}$, but a short-lived interaction might occur among the switch template, activators and incomplete inhibitors. From the gel data analysis, we could account for the concentrations of all the major species present in the system: [T], [T·A], [A], [A·I], and [I]. The gel data are consistent with switching behavior observed by fluorescence readout and with the model fits (figure 3.3A and B).

We further constructed a mutually inhibitory circuit (figure 3.1B, bottom) where the switch $\mathrm{Sw_{12}}$ and the switch $\mathrm{Sw_{21}}$ inhibit each other. The behavior of the mutually inhibitory circuit can be understood in terms of the characterized feedforward circuits (figure 3.1B, top). For the moment, consider that the mutually inhibitory circuit is essentially a two-dimensional dynamical system where the two switch activities give a complete description of the state of system because other variables, the

concentrations of activators, inhibitors, and the enzyme-substrate complexes in Michaelis–Menten enzyme reactions, relax to their steady-states much more rapidly. Then, the mutually inhibitory circuit behavior can be described as $\frac{d}{dt}[\mathrm{T}_{12}\mathrm{A}_2] = f([\mathrm{T}_{21}\mathrm{A}_1])$ and $\frac{d}{dt}[\mathrm{T}_{21}\mathrm{A}_1] = g([\mathrm{T}_{12}\mathrm{A}_2])$. We replace the input source So_i activities in the transfer curves of figure 3.2E and F with equivalent switch Sw_{ij} activities and interpret them as the nullclines of the mutually inhibitory circuit given by $f([\mathrm{T}_{21}\mathrm{A}_1]) = 0$ and $g([\mathrm{T}_{12}\mathrm{A}_2]) = 0$ (figure 3.4). The sigmoidal shape of transfer curves results in three fixed points, two stable and one unstable, to the extent that the approximations are valid. Thus, we expect that the two mutually inhibiting switches will show bistability with two stable attractors.

Because activator concentrations set the switching thresholds, we systematically varied activator concentrations to probe the conditions for multistability and to test the robustness of our system to parameter variation. A convenient experimental way to probe for multistability is to subject the network to different initial conditions and explore whether the network gets locked in different stable expression states. We therefore started the reaction either in the presence of excess RNA inhibitor I_1 (switch Sw_{21} OFF, switch Sw_{12} ON) or in the presence of excess RNA inhibitor I_2 (switch Sw_{21} ON, switch Sw_{12} OFF), the expected stable attractors from figure 3.4. If the mutually inhibitory circuit is bistable, the steady-state switch activities are locked in different states depending on the initial RNA inputs (figure 3.5A, right), while the switch activities converge irrespective of initial RNA inputs in the monostable parameter regime (figure 3.5A, left). A large part of activator parameter space displays persistent memory bordered by monostable regions (figures 3.5B and 3.11). In the latter cases, the switch Sw_{21} activity (M2), switch Sw_{12} activity (M1), or both switch activities (M0) were completely inhibited independent of the history. The activator concentrations can be shifted from the bistable regime to one of the monostable regimes (by way of adding DNA inhibitor dI_j to eliminate activator A_j) and back to the bistable regime (by way of adding activator A_j). The circuit maintained the state acquired in the monostable parameter regime when returned to the bistable parameter regime (figure 3.12).

We singled out the case with both activators at 1 μM to probe detailed dynamics in the phase plane. By initiating the reaction with various amounts of externally supplied RNA inhibitors, we can reset the system to various locations in the phase plane. The amount of RNA inhibitors determine the initial conditions yet do not determine multistability as shown in figure 3.5A. We chose a 15% variation of RNA inhibitor concentrations spanning the high-gain region of switch. The dynamic responses of both switches are simultaneously shown as trajectories in the switch activity phase plane with arrows to indicate directions (figure 3.5C) and as time-courses (figure 3.13A). Both switches

start by inhibiting each other and the trajectories move towards the corner where both switches are OFF. In all 12 cases, one of the switches recovers its activity and stabilizes. Two initial conditions are highlighted, where a 5% difference of RNA inhibitor I_2 concentrations leads trajectories to different stable attractors on opposite corners: (switch Sw_{21} activity, switch Sw_{12} activity) = (64nM, 3nM) and (4nM, 63nM). The experimental trajectories agree with the simulation results shown as trajectories in the switch activity phase plane (figure 3.5D) and as time-courses (figure 3.13B). The location of two attractors (circles with crosses inside, figure 3.5C), although not perfect, agrees with the location of attractors determined by the nullcline analysis (black circles, figure 3.4). Some experimental trajectories cross themselves and each other, and both experimental and simulation trajectories cross the separatrix for initial switch activity constructed from the model (blue line, figure 3.5D). This indicates that the state of dynamical system cannot be completely described by the switch activities alone and is influenced by unmeasured variables and possibly by unmodeled effects such as "bursting" enzyme kinetics (Jia and Patel 1997; Kuzmine and Martin 2001). Furthermore, the model does not accurately reproduce the kinetics of the system (figure 3.13B) although it gives the correct qualitative behavior. The recovery process is especially slow and incomplete for the switch Sw_{12}, which also showed less ideal behavior in the feedforward circuit. The system could maintain its memory for up to 11 hours, after which loss of NTP or loss of RNAP activity lead to decrease in inhibitor levels, turning both switches on.

3.2 Discussion

Surprisingly, controlling the degradation pathway turned out to be more difficult than controlling the production pathway. RNase H can only degrade the signal part of the transcript where hybridization to DNA activator occurs (Lima and Crooke 1997) (figure 3.10). Thus, unlike the full length transcripts (active signal) which turn over continuously, the shorter degradation products (inactive signal) accumulate during the circuit operation. This may help explain the slowness when switching multiple times between ON and OFF states. We have implemented the transcriptional circuit in combination with other ribonucleases to clean up inactive signals. Commercially available ribonucleases, of which we tested RNase I, A, III, and V, are endoribonucleases. Their activities quickly saturate because more RNA substrates are generated as a result of RNA substrates being cut internally. Also, they degrade long RNA substrates faster than short RNA substrates, accelerating the build-up of inactive signal. Following the observation that the degradation of RNA in *E. coli* is

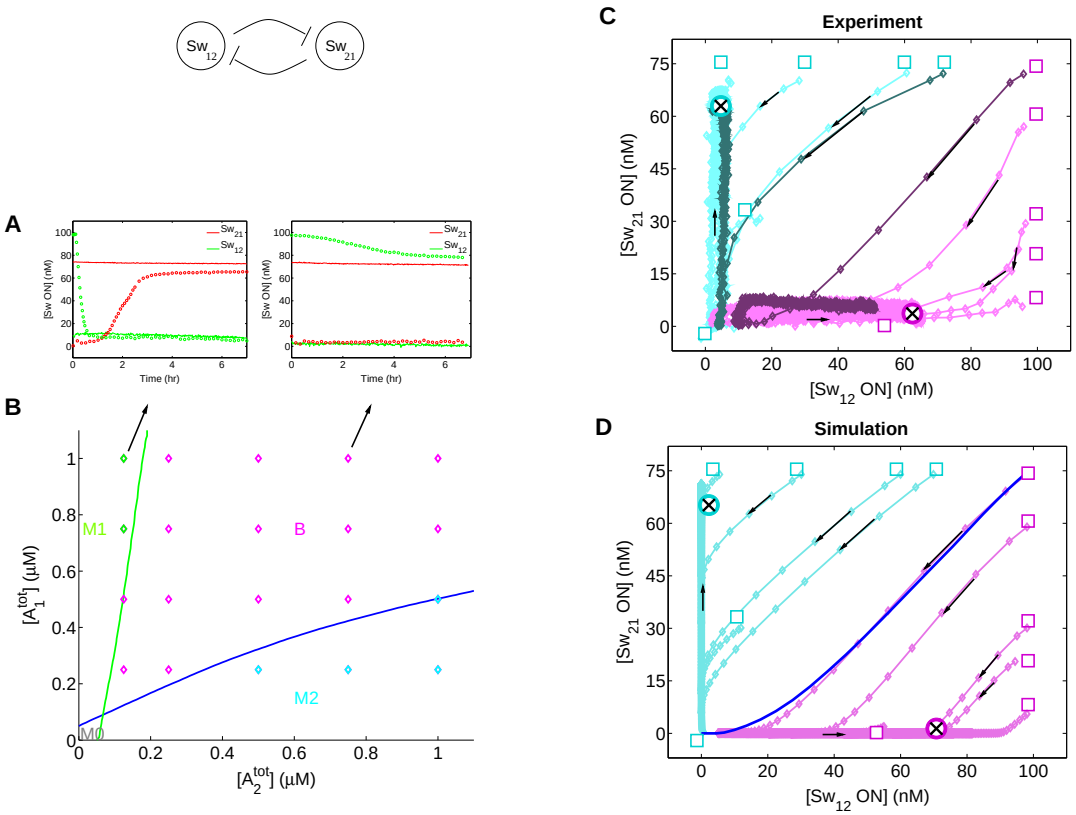


Figure 3.5. Characterization of bistable circuit. **(A)** Switch activity time-courses with high initial concentration of RNA inhibitor I_1 (circles) or high initial concentration of RNA inhibitor I_2 (lines). In the monostable case (left), the switch activities converge irrespective of the initial RNA inhibitor concentrations, while they stay apart in the bistable case (right). **(B)** Bifurcation diagram. Both the experiments and simulations are initiated with two switch states that are expected stable attractors if the circuit is bistable: (switch Sw_{21} ON, switch Sw_{12} OFF) and (switch Sw_{21} OFF, switch Sw_{12} ON). The model generates the following four domains. M0: both initial conditions are unstable, M1: only (switch Sw_{21} ON, switch Sw_{12} OFF) is stable, M2: only (switch Sw_{21} OFF, switch Sw_{12} ON) is stable, and B: both initial conditions are stable. Color-coded diamonds are experimentally determined stability results (Figure 11), two of which are shown as time-courses in (A). **(C)** Switch activity phase plane with both activators at 1 μM. Initial points are marked by rectangles and their attractors are marked by circles of the same color with black crosses inside. The black arrows indicate the direction of trajectories and diamonds are experimental data points measured at 1 minute intervals with lines drawn to guide the eye. Two trajectories with the following initial conditions are highlighted: the initial concentration of RNA inhibitor I_1 is 0.91 μM for both, and the initial concentration of RNA inhibitor I_2 is 0.89 μM (purple), or 0.93 μM (green). Despite small difference in initial RNA inputs, the two trajectories are attracted to the opposite corners of the phase plane. **(D)** Model fits of the switch activity phase plane. Diamonds are simulation results of 1 minute intervals as in the experiments with lines drawn to guide the eye. The black arrows indicate the direction of trajectories. The separatrix for initial conditions is generated from the model (blue line).

completed by the concerted action of an endoribonuclease, a processive exoribonuclease and a helicase in the degradosome (Grunberg-Manago 1999), we tested two 3' processive exoribonucleases.

RNase R (gift of Dr. Deutscher) (Cheng and Deutscher 2002) proved to be an excellent cleaner, but it also degrades full-length transcripts, consequently lowering the switch gain. Bistability was lost when a significant amount of RNase R was used, because inhibitor I_2 was a better substrate for RNase R than inhibitor I_1 in competing situations. OligoRNase (gift of Dr. Malhotra) (Fiedler et al. 2004) specifically degrades short single-stranded species, thus it would attack only incomplete RNA products. Unfortunately, the high NTP concentration in our transcriptional circuit inhibits the activity of oligoRNase (Datta and Niyogi 1975). Apparently, "all-or-none" degradation of RNA signals would require simultaneous activity of multiple ribonucleases with different substrate specificities, a challenge for our *in vitro* experiments.

Reproducing the bifurcation diagram (figure 3.5B) with our model is a stringent test on the validity of the model over a wide range of parameters, and demonstrates that Michaelis–Menten saturation is essential to the circuit behavior. When the circuit is modeled with first-order enzyme reactions, it is bistable at low levels of activators and monostable above a certain level of activators, quite contrary to experimental observation. Because in a first-order model the degradation rates of inhibitors in activator-inhibitor complexes increase linearly with the total concentration of activators, at high activator levels, the degradation rates exceed the production rates of inhibitors and repression of the target switch is lost. However, with Michaelis–Menten enzyme reaction equations, the degradation capacity of RNase H is shared by two activator-inhibitor substrates. This means that the relative abundance of an activator-inhibitor complex determines the probability of the activator-inhibitor complex being associated with RNase H, which in turn determines the degradation rate of inhibitor within that complex. Consequently, the bistability is maintained if the degradation rates of both inhibitors are relatively balanced despite high activator concentrations.

For engineering purposes, switches are most suitable when they contribute independently to the whole circuit. However, global coupling of rate equations through enzyme saturation can lead to global feedback regulation in the Michaelis–Menten enzyme reactions (Noireaux et al. 2003; Ackermann et al. 1998; Kim et al. 2004). First-order enzyme reactions are justified in the following cases. First, when the substrate concentrations are well below the Michaelis constants of enzymes: the enzyme reactions are limited by substrate, and consequently become first-order. However, both RNAP and RNase H have low Michaelis constants such that slow operation limited by DNA and RNA hybridization speed is inevitable in this parameter regime. Second, when there are many competing and compensatory species, the enzyme is at a similar saturation level even if a few substrate concentrations fluctuate, thus free enzyme concentration would be roughly a constant.

This might be true for genetic regulatory circuits in a cell (McClure 1985), but not for our bistable circuit with only four substrates for RNAP and two for RNase H. The concentration change in one substrate has a significant impact on the effective enzyme rate on another substrate.

Our circuit construction is much simpler than other approaches using protein signals (Noireaux et al. 2003; Isalan et al. 2005), yet general in computational power (Kim et al. 2004) and quantitatively explained better than other nucleic-acid-based feedback circuits (Wlotzka and McCaskill 1997; Ackermann et al. 1998), although accurately predicting kinetics rather than steady-states still remains a challenge. We did not model known enzyme activities such as "bursting" (Jia and Patel 1997; Kuzmine and Martin 2001) or side reactions (Cazenave and Uhlenbeck 1994; Zaher and Unrau 2004). Nonetheless, the characterization of feedforward circuits lead us to expect bistability in our feedback circuit, attesting modularity and programmability of the components. Thus, the synthetic switches in principle can be assembled to implement different logical networks with increasing complexity and offer a testbed for probing the design space of biochemical networks. For example, systematic exploration of parameter space in small feedback circuits and feedforward circuits, such as alternative implementations of oscillators and cascades, could elucidate principles for biochemical circuit design. A theoretical correspondence to neural network architecture would allow implementing networks of arbitrary complexity (Hopfield 1984). *In vitro* transcriptional circuits are suitable both for studying continuous mass action dynamics and, in principle, stochastic dynamics in small volumes (McAdams and Arkin 1997). Because of its simplicity, the characterization of noise source and propagation in small biochemical circuits should be facilitated. As discussed in Ackermann et al. (1998), dead-end side reactions can be important in *in vitro* systems that lack sophisticated control mechanisms found in the cell. Some of the known side reactions can be suppressed by experimental design, as in this work, and other constraints such as the exhaust of fuel (Klungsøyr et al. 1968) and build-up of degradation products could be relaxed in a chemostat (Atkinson et al. 2003), a dialysis bag (Madin et al. 2000), or vesicles (Noireaux and Libchaber 2004). The *in vitro* transcriptional circuit could be generalized to utilize active RNA signals (such as aptamers, ribozymes, and riboswitches (Mandal and Breaker 2004)) and could provide logical control of nanoscale devices (Dittmer and Simmel 2004) and artificial cells (Noireaux and Libchaber 2004). Although the current synthetic switch design is not suitable for *in vivo* implementation, the *in vitro* transcriptional circuit can serve as a tool for characterizing various biochemical circuit designs and studying generic problems such as composability, performance, robustness, and efficiency.

3.3 Materials and methods

DNA oligonucleotides and enzymes

The sequences of all DNA molecules and expected RNA transcripts were chosen to minimize the occurrence of alternative secondary structures, checked by the Vienna group's DNA and RNA folding program (Flamm et al. 2000). All DNA oligonucleotides were purchased (Integrated DNA Technologies, Coralville, Iowa, United States). T_{21}-nt is labeled with Texas Red at the 5' end, T_{12}-nt is labeled with TAMRA at the 5' end, A_1 is labeled with Cy5 or Blackhole Quencher-2 at the 3' end, and A_2 is labeled with FAM or Blackhole Quencher-2 at the 3' end. The T7 RNA polymerase (enzyme mix), transcription buffer, and NTP were purchased as part of the T7 Megashortscript kit (Ambion, Austin, Texas, United States; #1354). DNase I, RNase H, A, I, and V (Ambion; #1906, #2293, #2270, #2294, and #2275) and RNase III (Epicenter, Madison, Wisconsin, United States; #RN02950) were purchased. RNase R was a gift from Dr. Deutscher and OligoRNase was a gift from Dr. Malhotra, both at the University of Miami School of Medicine.

Transcription

Switch templates (T-nt and T-t strands) or source templates (So-ON/OFF-nt and T-t strands) were annealed with 10% (v/v) 10x transcription buffer from 90°C to 37°C over 1 hour at 5 times the final concentration used. To the annealed templates, activators and DNA or RNA inhibitors from a high concentration stock (50 μM), 7.5 mM each NTP, 8% (v/v) 10x transcription buffer, 3% (v/v) T7 RNA polymerase, and 0.35% (v/v) *E. coli* RNase H were added. Transcription reactions for spectrofluorometer experiments were prepared as a total volume of 70 μL. Transcription reactions for gel studies were prepared as a total volume of 50 μL and were stopped by phenol-chloroform extraction. For the purification of RNA inhibitors, I_1 and I_2, the full-length template side strands (the complement of T-nt rather than T-t) were used to prepare fully duplex DNA templates. The transcription reaction was prepared as a total volume of 60 μL with 0.2 μM fully duplex DNA templates. The transcription condition was the same as above except that 20% (v/v) T7 RNA polymerase was used and RNase H was omitted. After 6 hour incubation at 37°C, the reaction mixture was treated with 2.5 μL DNase I for 30 minutes to remove DNA templates and stopped by phenol-chloroform extraction. The reaction mixture is run on 8% denaturing gel, RNA inhibitor bands were excised and eluted from gel by crush and soak method and ethanol precipitated.

Data acquisition

For spectrofluorometer experiments, excitation and emission for TAMRA-labeled T_{12} were at

559 nm and 580 nm, while excitation and emission for Texas Red-labeled T_{21} were at 597 nm and 615 nm. The fluorescence was recorded every minute using a SPEX Fluorolog-3 (Jobin Yvon, Edison, New Jersey, United States) and converted to switch activity by normalizing against minimum fluorescence (measured before the addition of enzymes with excess quencher labeled activators) and maximum fluorescence (measured at the end of reaction with excess DNA inhibitors to displace activators). Denaturing polyacrylamide gels (8% 19:1 acrylamide:bis and 7 M urea in TBE buffer) were allowed to run for 50 min with 10 V/cm at 65°C in TBE buffer (100 mM Tris, 90 mM Boric Acid, 1 mM EDTA). The 10-base DNA ladder (Invitrogen, Carlsbad, California, United States; #10821-015) was used in the control lane and the denaturing gel was stained with SYBR gold (Molecular Probes, Eugene, Oregon, United States; #S-11494) for quantitation. The non-denaturing gels (10% 19:1 acrylamide:bis in TAE buffer) were allowed to run for 100 min with 13 V/cm at 35°C in TAE buffer containing 12.5 mM Mg^{2+} (40 mM Tris-Acetate, 1 mM EDTA, 12.5 mM Mg-Acetate, pH 8.3). The gel data was quantitated using the Molecular imager FX (Biorad, Hercules, California, United States). The total concentrations of inhibitors in the denaturing gel were measured with respect to 1 μM purified RNA inhibitors run in a control lane. The concentrations of labeled species in the non-denaturing gel were measured with respect to the maximum fluorescence of the corresponding bands.

Hill coefficients

The transfer curves of single switches and feedforward circuits were fit to the Hill equation (3.1). The best parameter for the Hill equation was determined by a linear regression of log-log plots using MATLAB (The MathWorks).

Model simulation

The kinetic simulations and parameter fittings were implemented in MATLAB. Differential equations were solved using the *ode23s* routine, while mean squared deviation of model fits on experimental data was minimized using the *fmincon* routine. During the fit, the parameters are constrained within plausible range spanning about two orders of magnitude. Other constraints are that $K_{M,ON}$ is smaller than $K_{M,OFF}$ and that k_{TAI} is not faster than k_{AI} since inhibition mechanism involves branch migration step in addition to simple hybridization. Two parameters are at the limit of range after fitting: $K_{M,ON,12}$ and $K_{M,H,2}$. The transfer curves and the time-courses of the phase plane and bifurcation diagram were used for parameter fitting. Three additional parameters were required to fit whole data set simultaneously: Rv, Rhv, and Dv. Separate data sets of transfer curves, phase plane, or bifurcation diagram fit well without these additional parameters.

Rv and Rhv indicate the relative activities of RNAP and RNase H (respectively) for two batches of enzymes; one batch was used for the switch and feedforward circuit characterization, while the other batch was used for the bistable circuit experiments. For the bistable circuit, the simulations use [RNAP^{tot}] = 30 nM and [RNase H^{tot}] = 4.4 nM, while for the feedforward circuits, we used [RNAP^{tot}] = Rv * 30 nM and [RNase H^{tot}] = Rhv * 4.4 nM. To account for sample loss during additional manipulations used during sample preparation for the bifurcation diagram measurements, we include a parameter Dv that indicates the remaining fraction of DNA template for switch Sw_{21}. Thus, for the bifurcation diagram, we used [T_{21}^{tot}] = Dv * 75 nM in the simulation.

DNA sequences

So_2ON-nt (122mer), 5'-TATTAGTGTGTAGTAGTAGTTCAAAAGAACTACTACTACACACT-AATACGACTCACTATAGGGAGAAGGAGAGGCGAAGATTGAGGTAAGAAAGGTAAGGAT-AATACTGACAAAGTCAGAAA-3'.

So_2OFF-nt (117mer), 5'-GTGTGTAGTAGTAGTTCAAAAGAACTACTACTACACACTAATAC-GACTCACTATAGGGAGAAGGAGAGGCGAAGATTGAGGTAAGAAAGGTAAGGATAATACT-GACAAAGTCAGAAA-3'.

T_{21}-nt (101mer), 5'-CTAATGAACTACTACTACACACTAATACGACTCACTATAGGGAGAAG-GAGAGGCGAAGATTGAGGTAAGAAAGGTAAGGATAATACTGACAAAGTCAGAAA-3'.

T_{21}-t (74mer), 5'-TTTCTGACTTTGTCAGTATTATCCTTACCTTTCTTACCTCAATCTTCGCC-TCTCCTTCTCCCTATAGTGAGTCG-3'.

A_1 (35mer), 5'-TATTAGTGTGTAGTAGTAGTTCATTAGTGTCGTTC-3'.

So_1ON-nt (127mer), 5'-TATTATCCTTACCTTTCTTACCAAAAGGTAAGAAAGGTAAGGATA-ATACGACTCACTATAGGGAGAAACAAAGAACGAACGACACTAATGAACTACTACTACACA-CTAATACTGACAAAGTCAGAAA-3'.

So_1OFF-nt (122mer), 5'-TCCTTACCTTTCTTACCAAAAGGTAAGAAAGGTAAGGATAATACG-ACTCACTATAGGGAGAAACAAAGAACGAACGACACTAATGAACTACTACTACACACTAAT-ACTGACAAAGTCAGAAA-3'.

T_{12}-nt (106mer), 5'-ATTGAGGTAAGAAAGGTAAGGATAATACGACTCACTATAGGGAGAA-ACAAAGAACGAACGACACTAATGAACTACTACTACACACTAATACTGACAAAGTCAGAA-A-3'.

T_{12}-t (79mer), 5'-TTTCTGACTTTGTCAGTATTAGTGTGTAGTAGTAGTTCATTAGTGTCGT-TCGTTCTTTGTTTCTCCCTATAGTGAGTCG-3'.

A_2 (35mer), 5'-TATTATCCTTACCTTTCTTACCTCAATCTTCGCCT-3'.

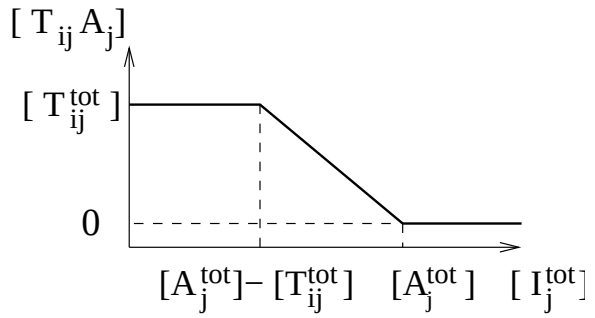
3.4 Appendix

Model equations

We present a simple model for DNA and RNA hybridization reactions, branch migration reactions, and Michaelis–Menten enzyme reactions in the transcriptional circuit: (i,j)∈{(1,2),(2,1)}.

DNA/RNA hybridization and branch migration reactions

$$\begin{aligned} \mathrm{T}_{ij} + \mathrm{A}_j &\overset{k_{TAj}}{\rightarrow} \mathrm{T}_{ij}\mathrm{A}_j && \text{(Activation)} \\ \mathrm{A}_j + \mathrm{I}_j &\overset{k_{AIj}}{\rightarrow} \mathrm{A}_j\mathrm{I}_j && \text{(Annihilation)} \\ \mathrm{T}_{ij}\mathrm{A}_j + \mathrm{I}_j &\overset{k_{TAIj}}{\rightarrow} \mathrm{T}_{ij} + \mathrm{A}_j\mathrm{I}_j && \text{(Inhibition)} \end{aligned}$$



The superscript *tot* indicates that all complexes containing that species are considered, e.g., $[\mathrm{T}_{ij}^{tot}]$=$[\mathrm{T}_{ij}]$+$[\mathrm{T}_{ij}\cdot\mathrm{A}_j]$ and $[\mathrm{A}_j^{tot}]$=$[\mathrm{A}_j]$+$[\mathrm{T}_{ij}\cdot\mathrm{A}_j]$+$[\mathrm{A}_j\cdot\mathrm{I}_j]$. The DNA/RNA hybridization reactions lead to the above transfer curve with the total concentration of I_j as an input and the concentration of active switch $\mathrm{T}_{ij}\mathrm{A}_j$ as an output. (1) $[\mathrm{I}_j^{tot}] <[\mathrm{A}_j^{tot}]$-$[\mathrm{T}_{ij}^{tot}]$, the inhibitor is consumed upon binding to free activator and does not inhibit the switch. (2) $[\mathrm{A}_j^{tot}]$-$[\mathrm{T}_{ij}^{tot}] < [\mathrm{I}_j^{tot}] < [\mathrm{A}_j^{tot}]$, the inhibitor is enough to consume all free activator and strips off activator bound to the switch stoichiometrically. (3) $[\mathrm{I}_j^{tot}] > [\mathrm{A}_j^{tot}]$, the inhibitor consumes all activator, free or bound to the switch, and the switch is completely OFF.

Michaelis–Menten enzyme reactions

$$\begin{aligned} \mathrm{RNAP} + \mathrm{T}_{ij}\mathrm{A}_j &\underset{k_{-,ON,ij}}{\overset{k_+}{\rightleftharpoons}} \mathrm{RNAP}\cdot\mathrm{T}_{ij}\mathrm{A}_j \overset{k_{cat,ON,ij}}{\rightarrow} \mathrm{RNAP} + \mathrm{T}_{ij}\mathrm{A}_j + \mathrm{I}_i \\ \mathrm{RNAP} + \mathrm{T}_{ij} &\underset{k_{-,OFF,ij}}{\overset{k_+}{\rightleftharpoons}} \mathrm{RNAP}\cdot\mathrm{T}_{ij} \overset{k_{cat,OFF,ij}}{\rightarrow} \mathrm{RNAP} + \mathrm{T}_{ij} + \mathrm{I}_i \\ \mathrm{RNaseH} + \mathrm{A}_j\mathrm{I}_j &\underset{k_{-,H,j}}{\overset{k_{+,H}}{\rightleftharpoons}} \mathrm{RNaseH}\cdot\mathrm{A}_j\mathrm{I}_j \overset{k_{cat,H,j}}{\rightarrow} \mathrm{RNaseH} + \mathrm{A}_j. \end{aligned}$$

We do not consider side-reactions or incomplete production and degradation products. The Michalis–

Menten enzyme reactions are further simplified by the steady-state assumption for the enzyme-substrate complexes. Since k_+'s are presumed to be fast, we express the available enzyme concentrations using the standard steady-state derivation:

$$[\mathrm{RNAP}] = \frac{[\mathrm{RNAP}^{tot}]}{1 + \sum \frac{[\mathrm{T}\cdot\mathrm{A}]}{K_{M,ON}} + \sum \frac{[\mathrm{T}]}{K_{M,OFF}}}, \qquad [\mathrm{RNaseH}] = \frac{[\mathrm{RNaseH}^{tot}]}{1 + \sum \frac{[\mathrm{A}\cdot\mathrm{I}]}{K_{M,H}}},$$

where the Michaelis constants are calculated as $K_M = \frac{k_- + k_{cat}}{k_+}$ to determine the affinity of substrates to the enzymes. From mass balance, $[\mathrm{T}_{ij}^{tot}]$ and $[\mathrm{A}_j^{tot}]$ are preserved such that $[\mathrm{T}_{ij}\cdot\mathrm{A}_j]$ and $[\mathrm{A}_j\cdot\mathrm{I}_j]$ can be calculated from $[\mathrm{T}_{ij}]$ and $[\mathrm{A}_j]$, under the assumption that the enzyme bound complexes are negligible, which is approximately valid because enzyme concentrations are low compared to substrate concentrations. Thus, the dynamics of each switch is described by the following three ordinary differential equations:

$$\begin{aligned}
\frac{d[\mathrm{T}_{ij}]}{dt} &= -k_{TAj}[\mathrm{T}_{ij}][\mathrm{A}_j] + k_{TAIj}[\mathrm{T}_{ij}\cdot\mathrm{A}_j][\mathrm{I}_j], \\
\frac{d[\mathrm{A}_j]}{dt} &= -k_{AIj}[\mathrm{A}_j][\mathrm{I}_j] - k_{TAj}[\mathrm{T}_{ij}][\mathrm{A}_j] + \frac{k_{cat,H,j}}{K_{M,H,j}}[\mathrm{RNaseH}][\mathrm{A}_j\cdot\mathrm{I}_j], \\
\frac{d[\mathrm{I}_j]}{dt} &= -k_{AIj}[\mathrm{A}_j][\mathrm{I}_j] - k_{TAIj}[\mathrm{T}_{ij}\cdot\mathrm{A}_j][\mathrm{I}_j] + \frac{k_{cat,ON,ji}}{K_{M,ON,ji}}[\mathrm{RNAP}][\mathrm{T}_{ji}\cdot\mathrm{A}_i] \\
&\quad + \frac{k_{cat,OFF,ji}}{K_{M,OFF,ji}}[\mathrm{RNAP}][\mathrm{T}_{ji}].
\end{aligned}$$

The kinetic simulations and parameter fittings were implemented in MATLAB. Three additional parameters were required to fit whole data set simultaneously: Rv, Rhv, and Dv (materials and methods). These parameters are not unique; similar fits can be achieved with some parameters changing by more than a factor of 10 when appropriate trade-offs are made. However, our choice of parameters shows that the model we present here is quantitatively plausible. For comparison, the parameter values and enzyme constants from other biochemical studies are listed below. The T7 RNA polymerase parameters were measured on synthetic DNA templates that have a complete promoter sequence (different from our ON state template by a nick at -12) or a promoter sequence with 5 bases missing on the template side (identical to our OFF state template) (Martin and Coleman 1987). Higher K_M for our ON state template may be attributed to the presence of nick. Because the transcript was very short (5 bases) in (Martin and Coleman 1987), only the initiation rate constant was measured as k_{cat}. The initiation rate was 30 times faster than the steady-state transcription rate ("bursting") in another study (Jia and Patel 1997), which may explain our small k_{cat} values.

The RNase H parameters were measured on RNA-DNA hybrid stems of molecular beacon (Rizzo et al. 2002). Since our substrates (activator-inhibitor complexes) are longer than those of Rizzo et al. (2002), slower k_{cat}'s are plausible. Thus, our enzyme parameters are reasonable compared to other biochemical studies. Hybridization rate constants (k_{TA}, k_{AI}, k_{TAI}) are expected to be on the order of 10^5/M/s in the absence of enzymes; in our fits, k_{TA} is consistently slow, suggesting that enzyme binding or interaction with degradation products is interfering in the reaction.

Table 3.1. Model parameters

	i=2, j=1	i=1, j=2	Other studies
$K_{M,ON,ij}$ (nM)	259	316	15-37
$k_{cat,ON,ij}$ (/s)	0.064	0.105	0.73-1.12
$K_{M,OFF,ij}$ (μM)	1.05	1.27	0.1-1.1
$k_{cat,OFF,ij}$ (/s)	0.007	0.023	0.11-0.18
$K_{M,H,j}$ (nM)	91	10	16-130
$k_{cat,H,j}$ (/s)	0.176	0.004	0.02-0.6
$k_{TA,j}$ (/M/s)	3.94*10^3	1.20*10^3	-
$k_{AI,j}$ (/M/s)	6.96*10^4	1.52*10^5	-
$k_{TAI,j}$ (/M/s)	6.96*10^4	1.52*10^5	-
Rv	0.50	-	-
Rhv	0.75	-	-
Dv	0.80	-	-

DNA switch design The choice of domain lengths on the synthetic transcriptional switches are based on earlier experiments which used different DNA sequences and transcription reagents. Here, we describe the earlier switch designs and the changes we made for improved switch functionality.

1. **Toehold**: We tested different toehold lengths for effective implementation of the inhibition mechanism. A toehold length of 6, 8, or 10 bases showed fast kinetics in initiation of branch migration, however a toehold length of 10 decreased the transcription rate from an ON state switch template. Thus, we kept the toehold length at 8 bases for increased programmability compared to 6 bases and for increased transcription efficiency compared to 10 bases.

2. **Branch migration region**: Initially, we used 17 base branch migration regions upstream of the T7 RNAP promoter sequence. However, the ON state switch template transcription rate dropped dramatically after a few hours. We interpreted this as the binding between the

template and activator being relatively weak and sensitive to the change of buffer condition as the transcription reaction progressed. When we increased the branch migration region to 27 bases, the transcription rates of ON state templates became more stable for up to 12 hours.

3. **Position of nick**: We tested different nick positions from -8 to -14 in the T7 RNAP promoter sequence. The ON state transcription rates were similar for various nick positions, but the OFF state transcription rates were higher as the nick position moved from -8 to -14. To test crosstalk, we used a DNA activator with a different branch migration sequence together with the OFF state template, and tested whether the leaky transcription from the OFF state template increased. The nick position -10 showed increased leaky transcription from an OFF state template when the unrelated DNA activator is used in excess, while the nick position -12 had no detectable change in the leaky transcription. To prevent crosstalk, we kept the nick position at -12 of the promoter region.

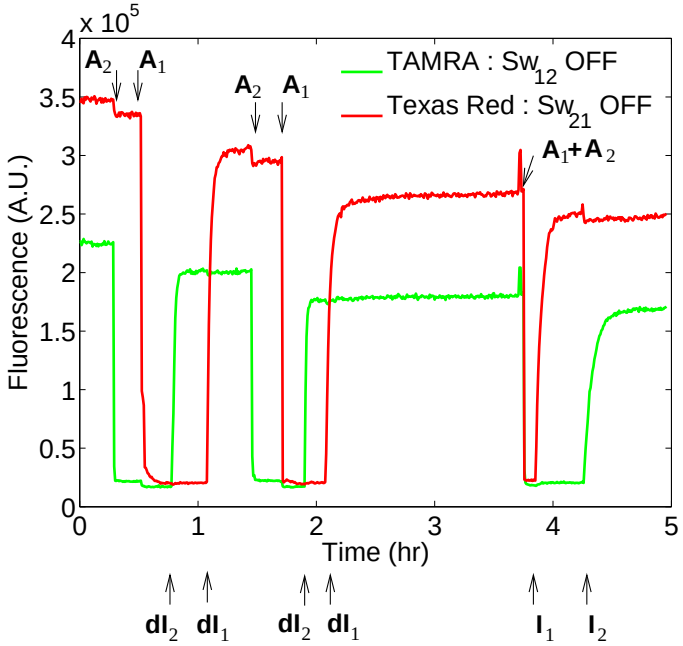


Figure 3.6. Specificity of template-activator interactions. Initially, the 70 μL reaction mixture contains 50 nM Texas Red-labeled switch template T_{21}, 50 nM TAMRA-labeled switch template T_{12}, and 10% (v/v) transcription buffer. Quencher-labeled activators and DNA/RNA inhibitors are added in the following order: (1) 0.3 μL of 50 μM activator A_2 at 15 minutes, (2) 0.3 μL of 50 μM activator A_1 at 30 minutes, (3) 0.5 μL of 50 μM DNA inhibitor dI_2 at 45 minutes, (4) 0.5 μL of 50 μM DNA inhibitor dI_1 at 65 minutes, (5) 0.5 μL of activator A_2 at 85 minutes, (6) 0.6 μL of activator A_1 at 100 minutes, (7) 0.5 μL of DNA inhibitor dI_2 at 115 minutes, (8) 0.5 μL of DNA inhibitor dI_1 at 125 minutes, (9) 0.5 μL of activator A_1 and 0.5 μL of activator A_2 at 225 minutes, (10) 0.4 μL of 80 μM RNA inhibitor I_1 at 230 minutes, (11) 0.4 μL of 70 μM RNA inhibitor I_2 at 255 minutes. The TAMRA signal changes only upon the introduction of A_2, dI_2, and I_2, while the Texas Red signal changes only upon the introduction of A_1, dI_1, and I_1, demonstrating the specificity of interactions. The fluorescence crosstalk between two channels is about 2% of the total fluorescence signals. The maximum fluorescence level dropped about 30% by the end of repeated hybridization reactions. One-third of the signal drop can be explained by dilution. Other sources of signal loss include absorption and loss of fluorophore-labeled DNA on pipet tips during mixing (data not shown).

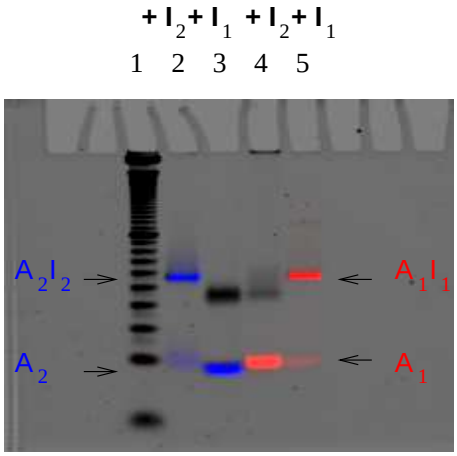


Figure 3.7. Specificity of activator-inhibitor interactions analyzed in a 10% non-denaturing gel. The gel is scanned for FAM fluorescence (blue, excitation: 488 nm, emission: 500–560 nm) and Cy5 fluorescence (red, excitation: 645 nm, emission: 670–720 nm). After staining with SYBRgold, the gel is scanned for SYBRgold signal (black). These three images were digitally aligned and superimposed. Lane 1 contains 10-base ladder. Lane 2 contains 1 μM FAM-labeled activator A_2 and 0.6 μM RNA inhibitor I_2. Lane 3 contains 1 μM FAM-labeled activator A_2 and 2 μM RNA inhibitor I_1. Lane 4 contains 1 μM Cy5-labeled activator A_1 and 2 μM RNA inhibitor I_2. Lane 5 contains 1 μM Cy5-labeled activator A_1 and 0.6 μM RNA inhibitor I_1. Activator A_2 and A_2I_2 complex (blue) can be identified in lane 2, but activator A_2 (blue) and RNA inhibitor I_1 (black) migrate separately and no A_2I_1 complex is identified in lane 3. Activator A_1 and A_1I_1 complex (red) can be identified in lane 5, but activator A_1 (red) and RNA inhibitor I_2 (black) migrate separately and no A_1I_2 complex is identified in lane 4.

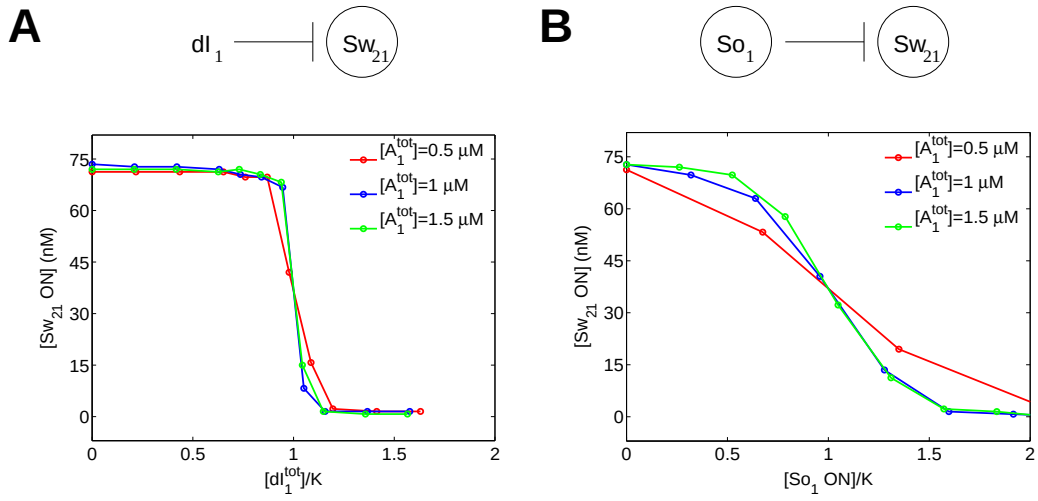


Figure 3.8. Rescaled transfer curves of switch Sw_{21} and a feedforward circuit. The data sets are the same as those used in figures 3.2B and E. The inputs are normalized with respect to K's (the amount of inputs required for half repression). The transfer curves with higher total activator concentrations show sharper transitions. Experimental data are plotted as circles with lines drawn to guide the eye.

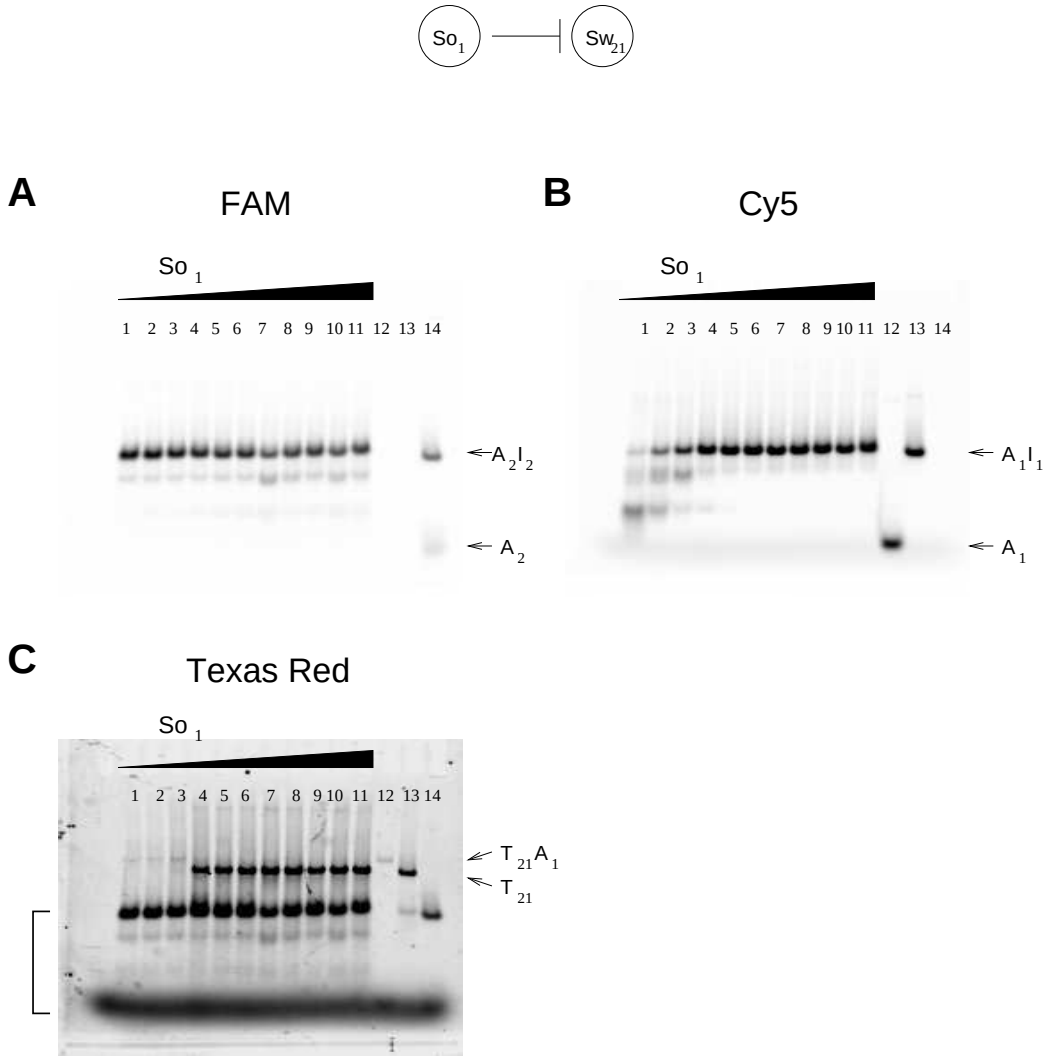


Figure 3.9. Feedforward circuit analyzed in a 10% non-denaturing gel. The gel is scanned for FAM fluorescence (excitation: 488 nm, emission: 500–560 nm), Cy5 fluorescence (excitation: 645 nm, emission: 670–720 nm) and Texas Red fluorescence (excitation: 532 nm, emission: 580–630 nm). These three images are digitally aligned and superimposed in figure 3.3D. **(A)** FAM–labeled activator A_2 and A_2I_2 complex can be identified. The A_2 band in lane 14 is faint because almost stochiometric amount of I_2 is used. The unidentified bands may be A_2 bound to a partially degraded I_2. **(B)** Cy5–labeled activator A_1 and A_1I_1 complex can be identified. The unidentified bands may be A_1 bound to a partially degraded I_1. **(C)** Texas Red–labeled switch template T_{21} and $T_{21}A_1$ can be identified. The ON switch template $T_{21}A_1$ has low fluorescence (it is barely discernable in lanes 1, 2, 3, and 12), presumably due to fluorescence resonance energy transfer from Texas Red on T_{21} to Cy5 on A_1. Because the Texas Red fluorescence is relatively weak, emission from FAM-labeled species and XCFF loading dye show up as well (bracket).

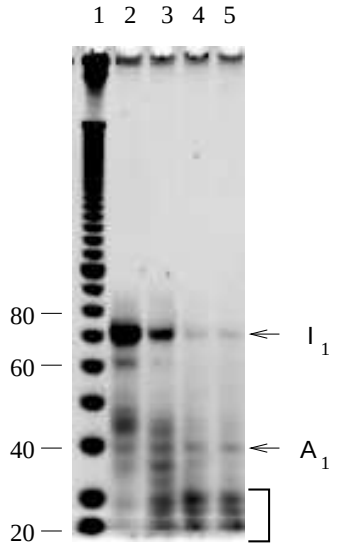


Figure 3.10. Ribonuclease H activity. The reaction mixture contains 4 μM RNA inhibitor I_1, 100 nM activator A_1, 100 nM RNAP, 20 nM RNase H, 10% (v/v) transcription buffer, and 40% (v/v) NTP. The reaction mixture is incubated at 37°C and samples are taken at different times to be analyzed in a 8% denaturing gel. Lane 1 contains 10-base ladder, lane 2 to 5 contain samples taken at 0, 20, 40, and 60 minutes, respectively. About 3 μM of inhibitor I_1 is processed during the first 20 minutes. Note that smaller RNA species are not degraded after 40 minutes (bracket). These are presumably the 5' overhang and 3' hairpin structures in inhibitor I_1, which are not complementary to activator A_1.

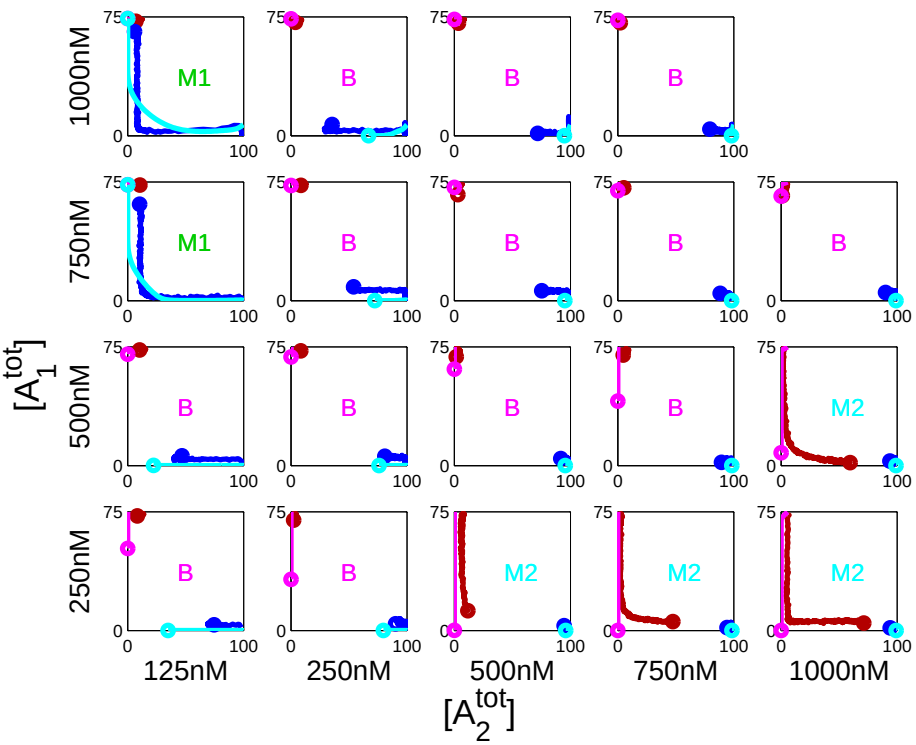


Figure 3.11. Kinetic trajectories of the bistable circuit used for the bifurcation diagram (figure 3.5B). Switch activity phase planes with the switch Sw_{12} activity as x-axes and the switch Sw_{21} activity as y-axes are embedded in the bifurcation diagram for the activator concentrations. Each phase plane contains trajectories starting from the opposite corners: (switch Sw_{21} ON, switch Sw_{12} OFF) or (switch Sw_{21} OFF, switch Sw_{12} ON). Both experimental trajectories (red and blue) and simulation trajectories (magenta and cyan) are shown with final points as circles. Bistability is achieved when the two trajectories do not converge. The stability assessment for experimental results are shown as letters: only (switch Sw_{21} ON, switch Sw_{12} OFF) is stable (M1), only (switch Sw_{21} OFF, switch Sw_{12} ON) is stable (M2), or both initial conditions are stable (B).

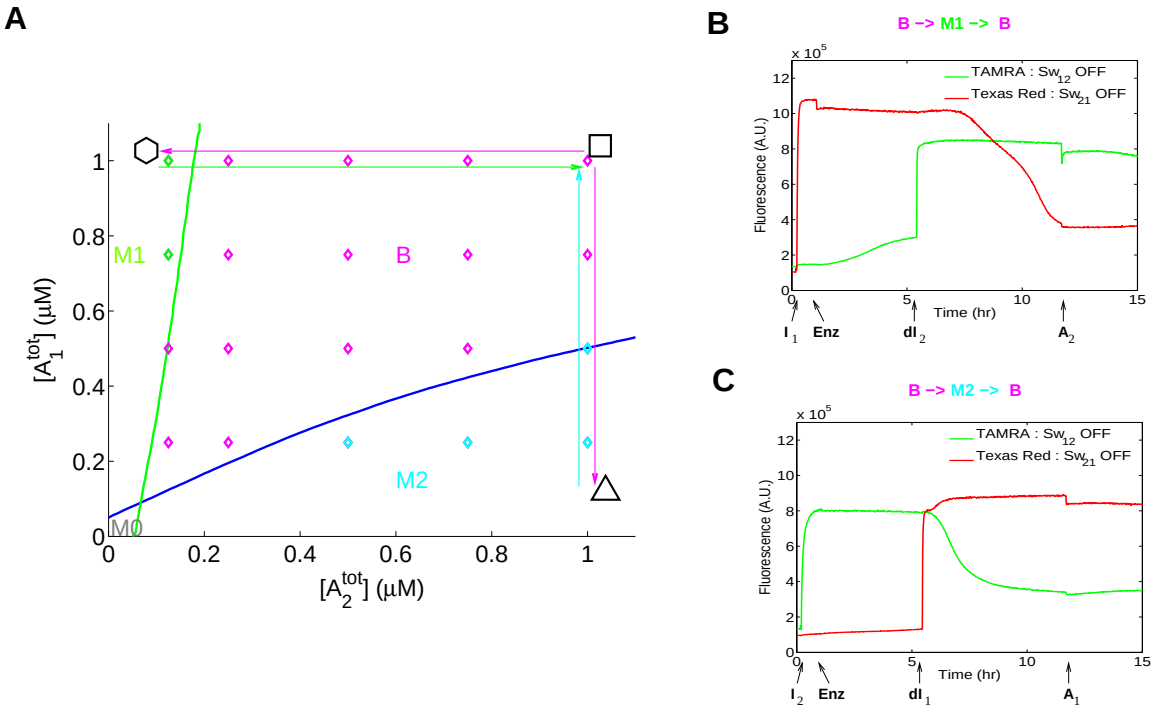


Figure 3.12. Hysteresis of the bistable circuit. **(A)** Experimental design. The bistable circuit is perturbed in two different ways to test hysteresis. The activator concentrations were shifted from the bistable regime (square) to one of the monostable regimes (hexagon and triangle) and back to the bistable regime. The circuit maintained the state acquired in the monostable parameter regime when returned to the bistable parameter regime. **(B and C)** Fluorescence time-courses. Initially, the 70 μL reaction mixture contains 75 nM Texas Red–labeled switch template T_{21}, 100 nM TAMRA-labeled switch template T_{12}, 1 μM quencher-labeled activator A_1 and A_2 with transcription buffer and NTP. The reaction condition is the same as that of Figure 5C except for the amount of initial RNA inhibitors. Initial fluorescence time-courses of (B) and (C) closely match those of figure 3.5A right when converted to the switch activities. **(B)** Other reagents are added in the following order: (1) 3.5 μL of 20 μM inhibitor I_1 at 10 minutes, (2) 2.1 μL RNAP and 0.16 μL RNase H (an equivalent amount of this RNase H batch as 0.24 μL used in other reactions of bifurcation diagram results (figure 3.11)) at 60 minutes, (3) 1.25 μL of 50 μM DNA inhibitor dI_2 at 320 minutes, (4) 1.25 μL of 50 μM quencher-labeled activator A_2 at 700 minutes. The TAMRA signal stays low (switch Sw_{12} ON) and the Texas Red signal stays high (switch Sw_{21} OFF) initially, but the signal state switches after the addition of DNA inhibitor dI_2 because the circuit has moved to a monostable parameter regime (hexagon in (A): switch Sw_{12} OFF, switch Sw_{21} ON). The fluorescence signals stay at the same level even after the addition of activator A_2, which brings the circuit back to the bistable regime. **(C)** Other reagents were added in the following order: (1) 3.5 μL of 20 μM inhibitor I_2 at 10 minutes, (2) 2.1 μL RNAP and 0.16 μL RNase H at 60 minutes, (3) 1.25 μL of 50 μM DNA inhibitor dI_1 at 320 minutes, (4) 1.25 μL of 50 μM quencher-labeled activator A_1 at 700 minutes. The TAMRA signal stays high (switch Sw_{12} OFF) and the Texas Red signal stays low (switch Sw_{21} ON) initially, but the signal state switches after the addition of DNA inhibitor dI_1 because the circuit has moved to a monostable parameter regime (Triangle in (A): switch Sw_{12} ON, switch Sw_{21} OFF). The fluorescence signals stay at the same level even after the addition of activator A_1, which brings the circuit back to the bistable regime.

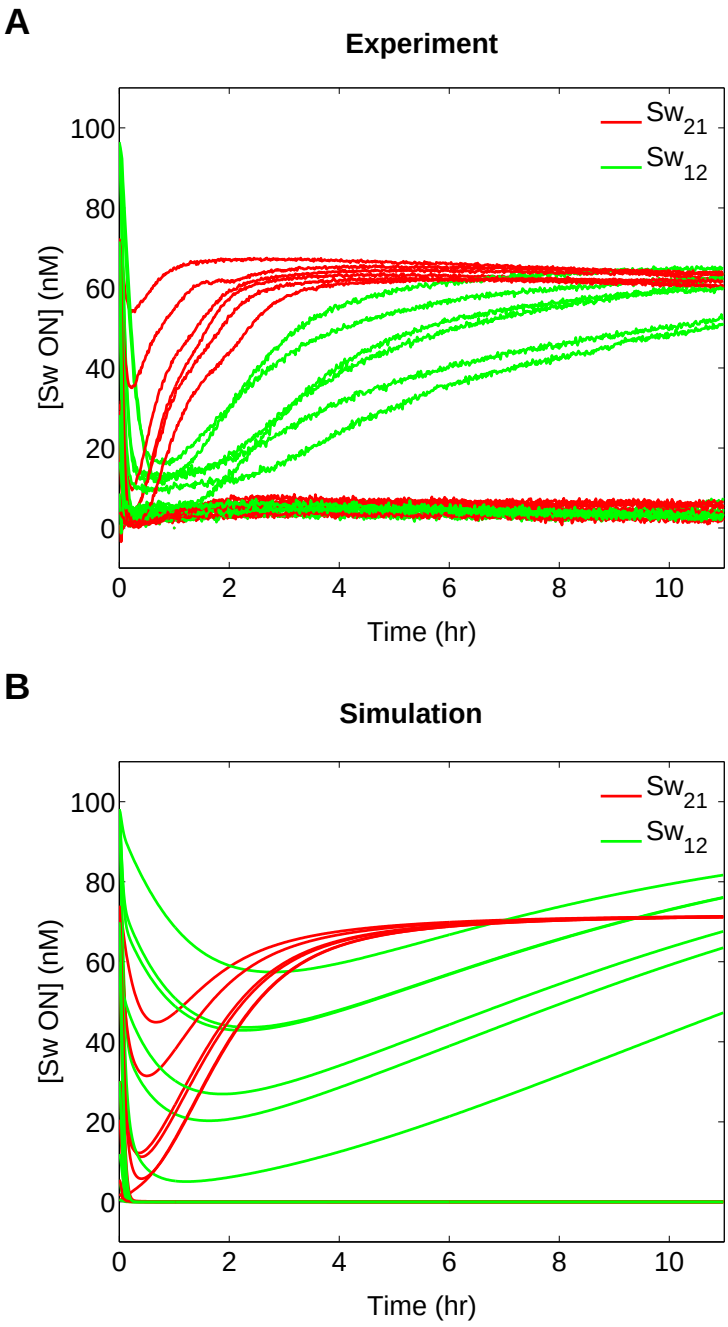


Figure 3.13. Kinetic trajectories of the bistable circuit used for the phaseplane. **(A)** Experimental switch activity time-courses of the bistable circuit (figure 3.5C). **(B)** Switch activity time-courses of the bistable circuit generated from the model (figure 3.5D).

Chapter 4 A Positive Feedback Circuit

4.1 Results

A positive feedback circuit can amplify input signals and sharpen the switching threshold in biological systems. The strength of the positive feedback determines the behavior of the switch, for example, whether it is monostable or bistable (Wolf and Arkin 2003). A memoryless switch could serve as a signal thresholding component. For example, the positive feedback loop of M-Cdk and M-cyclin genes generates a clear distinction between mitosis phase and G1 phase of cell cycles. The former is marked by high M-Cdk concentration and the latter is marked by low M-Cdk concentration. Bistable switches can be used for controlling development. For example, bacteriophage lambda uses intricate positive feedback loops to commit itself to either the lytic or lysogenic phase (Ptashne 1992). Synthetic positive feedback circuits have been constructed and studied *in vivo* (Becskei et al. 2001; Isaacs et al. 2003).

Here, we investigate an autoregulatory transcriptional switch with positive feedback in a greatly simplified *in vitro* setting. The synthetic DNA template design is modular with easily programmable connectivity dictated by Watson–Crick base-pairing rules. The regulatory domain is upstream of the promoter region; the output domain is downstream of the promoter region. We exploit the modularity of the DNA switch template to modify only the regulatory region of a previously characterized transcriptional switch (Kim et al. 2006). Although the RNA signal sequence is identical to a previously used inhibitory RNA signal, the RNA signal relays an excitatory regulatory signal to the modified regulatory domain of the new switch. As before, regulated DNA templates are called switches (“Sw”) with two different conformations with different transcription efficiency: ON or OFF (figure 4.1). The OFF state of the switch consists of a double-stranded DNA template (“T”) with a partially single-stranded (ss) and thus incomplete promoter region. The switch is turned on by the addition of a ssDNA activator (“dA”) that completes the promoter region. The activator contains a “toehold,” a single-stranded overhang beyond the helical domain it forms with the DNA template, where an inhibitor can bind to initiate a toehold-mediated strand displacement reaction (Yurke and Mills 2003). Thus, the switch can be turned off upon addition of a ssDNA inhibitor strand (“dI”). The DNA inhibitor dI, in turn, has its toehold region where the RNA activator strand (“rA”) can bind

to initiate a toehold-mediated strand displacement reaction, releasing DNA activator dA from the dA·dI complex. In the positive feedback circuit, the RNA activator strand is produced from switch templates by RNAP using NTP as fuel and is degraded by RNase H or RNase R.

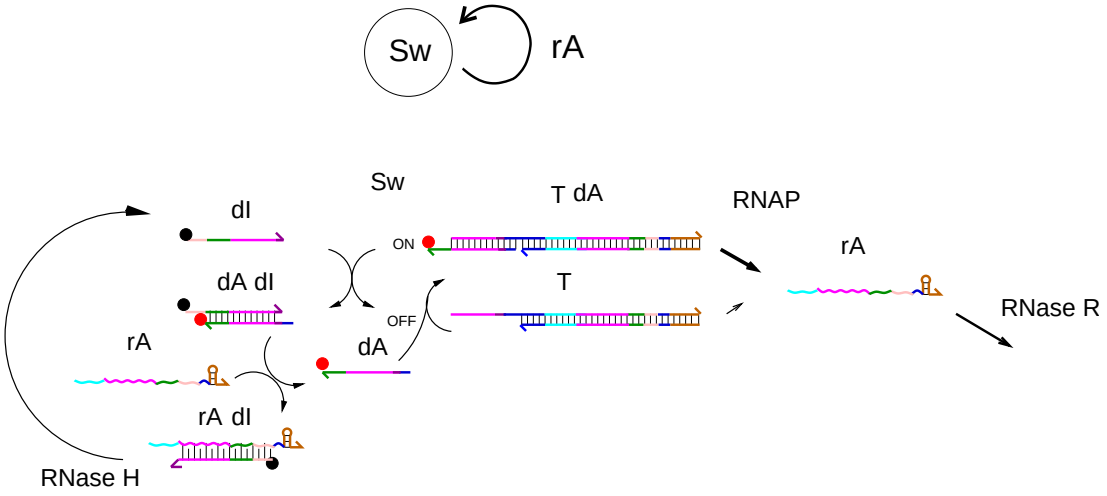


Figure 4.1. Design of synthetic autoregulatory switch with positive feedback. Each sequence subdomain is color-coded. Switch templates have two distinct states, ON or OFF, with different transciption speed. The ON-state switch template (T·dA complex) has complete promoter sequence (17 bases, blue) with a nick and the OFF-state switch template (T) is missing 5 bases of the promoter sequence on the template side. The two single strands that form the switch templates are the longer non-template side strand, T-nt, and the shorter template side strand, T-t (materials and methods). The DNA activator dA is labeled with fluorophore Cy5 (red circle) and DNA inhibitor dI is labeled with quencher (black circle) such that the state of switch can be monitored by measuring the fluorescence quenching efficiency (Marras et al. 2002). RNAP produces RNA activator rA from DNA templates (either ON or OFF), while RNase H degrades RNA activator bound to DNA inhibitor and RNase R degrades isolated RNA activator. Detailed reaction mechanisms are listed in the appendix 4.4.

We use competitive binding of nucleic acid species rather than cooperative binding to achieve ultrasensitivity. The threshold in our transcriptional circuit derives from four types of strong DNA and RNA hybridization reactions (appendix 4.4), which we call activation, annihilation, inhibition, and release. An activator binds to an OFF switch template to turn the switch on (activation); a DNA or RNA activator binds to a DNA inhibitor and is not available for the switch template (annihilation); a DNA inhibitor displaces a DNA activator from an ON switch template, the T·dA complex, to turn the switch off (inhibition); an RNA activator displaces a DNA activator from the dA·dI complex (release). The key requirements for the inhibition and release mechanism are that the resulting complex is thermodynamically more favorable than the starting complex and that there is a fast kinetic pathway to the lowest energy state (in our case, toehold-mediated strand displacement (Yurke

and Mills 2003)). Since the activator and inhibitor annihilate each other, the difference of total activator (DNA or RNA) and inhibitor concentrations is the most important determinant of the state of switch: an excess of inhibitor will turn the switch off while an excess of activator will turn the switch on. All four mechanisms are needed for fast switch response. We have verified individual reaction mechanisms by gel studies (appendix 4.4.1).

Depending on the gain and the threshold of switch response, we expect that the autoregulatory switch can show either bistable or monostable behavior. First, we use the simple mathematical model (appendix 4.4.2) to study the effect of each parameter variation. Assuming fast and almost complete hybridization among T, dA, dI, and rA species, we plot the production rate of rA as a function of total concentration of rA as shown in figure 4.2A. The production rate of rA is dependent on several parameters. For example, changing dA, dI, or T concentrations will shift the threshold of switching, while changing T or RNAP concentrations will shift the overall speed of production. The degradation rate of rA is dependent on the concentrations of RNase H and dI. Because RNase H degrades RNA activator rA in an rA·dI complex, the degradation rate of rA cannot continue to increase after the total concentration of rA exceeds the total concentration of dI. Also, the degradation rate curve shows saturation effect because RNase H has small Michaelis constant.

In case the production rate curve stays above the degradation rate curve (figure 4.2B), the RNA activator level will increase irrespective of initial RNA activator concentration. On the other hand, in case the production rate curve stays below the degradation rate curve except at very low RNA levels (figure 4.2C), the concentration of RNA activator will approach a low steady-state irrespective of initial RNA concentrations. The former is ON-state monostable behavior and the latter is OFF-state monostable behavior. In figure 4.2A, the production and degradation rate curves have two crossings, a stable fixed point and an unstable fixed point, and therefore the system shows bistable behavior. When the switch is ON with enough RNAP, RNA activator production by RNAP can exceed RNase H's capacity for degradation, the concentration of RNA activator will increase without bound rather than achieve a steady-state. In such cases, however, switch activity nonetheless approaches an asymptotic "steady-state" value. Thus, although the RNA activator level may not reach a steady-state, we will use ON steady-state or bistability in this loose sense.

We determined the effect of each parameter variation experimentally (figure 4.3). A convenient experimental way to probe for multistability is to subject the network to different initial conditions and explore whether the network gets locked in different stable expression states. For each parameter set, we initiate the reaction mixture with either a low amount of RNA activator (0 nM rA) or

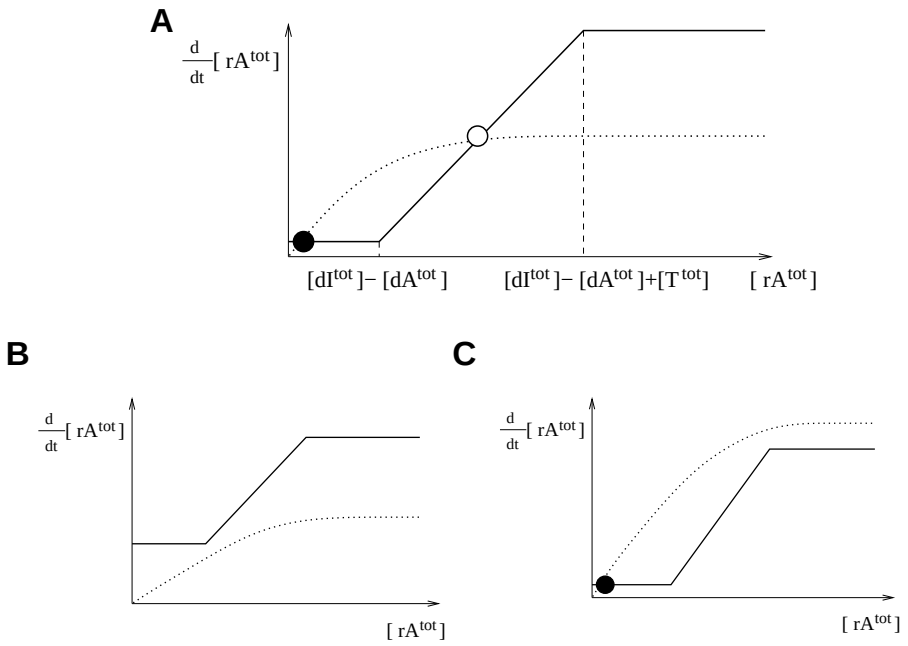


Figure 4.2. The production and degradation rates of RNA activator as a function of RNA activator concentration. The production rate curves are drawn as solid lines and the degradation rate curves are drawn as dotted lines. The filled circle indicates a stable steady-state, while the open circle indicates an unstable steady-state. **(A)** With fast and irreversible hybridization assumptions, the switch states hence the production rate of RNA activator depends on the total RNA activator concentration piecewise-linearly. The switch is completely OFF when [rA^{tot}]<[dI^{tot}]-[dA^{tot}] and completely ON when [rA^{tot}]>[dI^{tot}]-[dA^{tot}]+[T^{tot}]. Both ON and OFF switch states are stable and the final state of switch will be determined by the initial RNA activator concentrations. **(B)** A monostable ON switch state will be achieved irrespective of initial RNA activator concentrations if the production rate exceeds the degradation rate of the RNA activator. **(C)** A monostable OFF switch state will be achieved irrespective of initial RNA activator concentrations if the degradation rate exceeds the production rate of the RNA activator when the switch is ON.

a high amount of RNA activator (800 nM rA) and check whether the final difference of rA has increased. The autoregulatory switch reached the OFF state under conditions with low T, low dA, low RNAP, high dI, or high RNase H, irrespective of initial conditions. Conversely, the autoregulatory switch reached the ON state with high T, high dA, high RNAP, low dI, or low RNase H, irrespective of initial conditions. When the autoregulatory switch is in a monostable ON state, the rA concentrations do not converge completely because the production rate of rA exceeds the degradation capacity of RNase H. Consequently, the difference of final rA concentrations will be similar to the difference of initial rA concentrations. The model could capture the dependence of system behavior on each parameter (figure 4.3).

Next, we probed the phasediagram in more detail with different thresholds determined by dI.

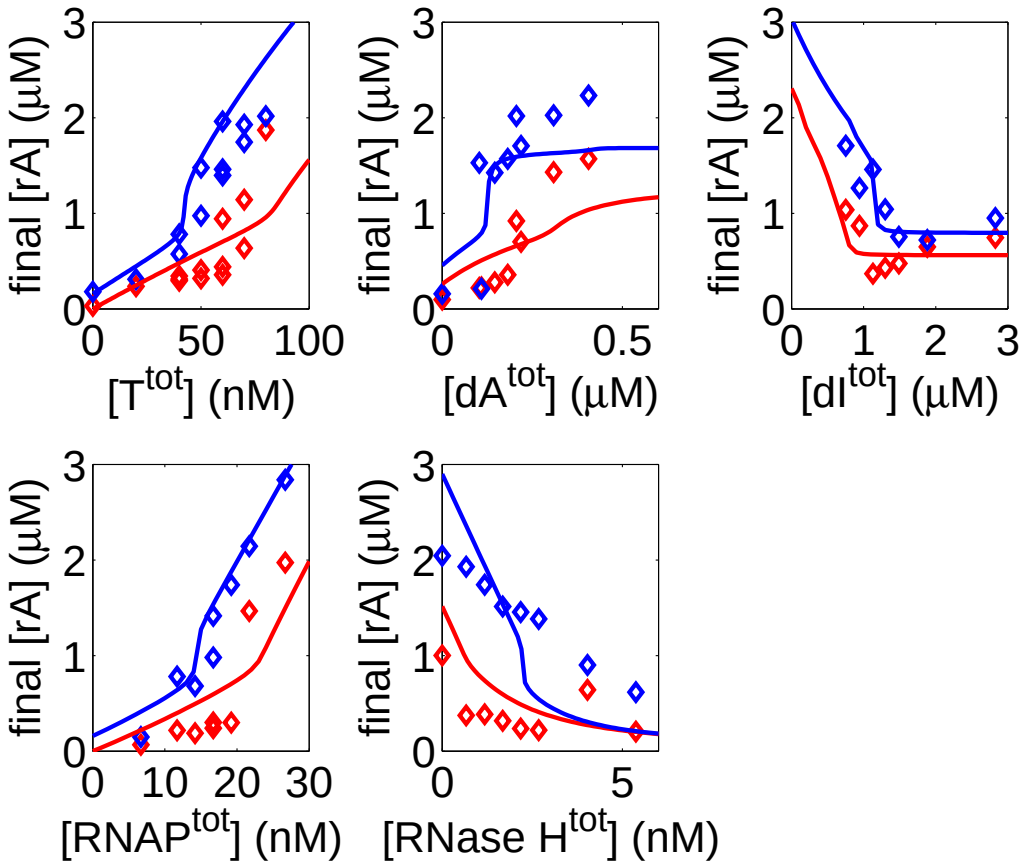


Figure 4.3. Various parameter variations and fits. We systematically vary each parameter ([T^{tot}], [dA^{tot}], [dI^{tot}], [$RNAP^{tot}$], or [RNase H^{tot}]) and measure the final RNA activator concentrations for two different initial RNA concentrations: 0 nM (red) and 800 nM (blue). The experimental measurements are plotted as diamonds and simulation results are plotted as lines.

We expect that the required amount of initial RNA activator to turn on the switch will increase as the threshold set by DNA inhibitor dI increases. Three different thresholds have been demonstrated in duplicate experiments (figure 4.4). Note the high variance in the medium threshold experiment (blue), which may be explained by sensitive dependence on initial conditions around the switching threshold.

Finally, we monitor the autoregulatory switch behavior real-time in a fluorometer. The DNA activator dA is labeled with fluorophore (Cy5) and the DNA inhibitor dI is labeled with quencher. The fluorescence is low when dA·dI complex is formed due to fluorescence quenching (Marras et al. 2002), while the fluorescence is high when dA is free or bound to T. We initate the reaction with four different amounts of RNA activators and ask whether the system approaches different steady-states. For the two cases where the initial rA concentration was below the threshold, the

fluorescence stayed low, indicating that dA is bound to dI and the switch is OFF. For the other two cases where the initial rA concentration exceeded the threshold, the fluorescence quickly increased to the maximum value, indicating that dA is freed from the dA·dI complex and the switch turns on completely (figure 4.5A). Gel analysis of fluorometer reaction samples verified the bistable response in terms of RNA activator concentrations (figure 4.5B).

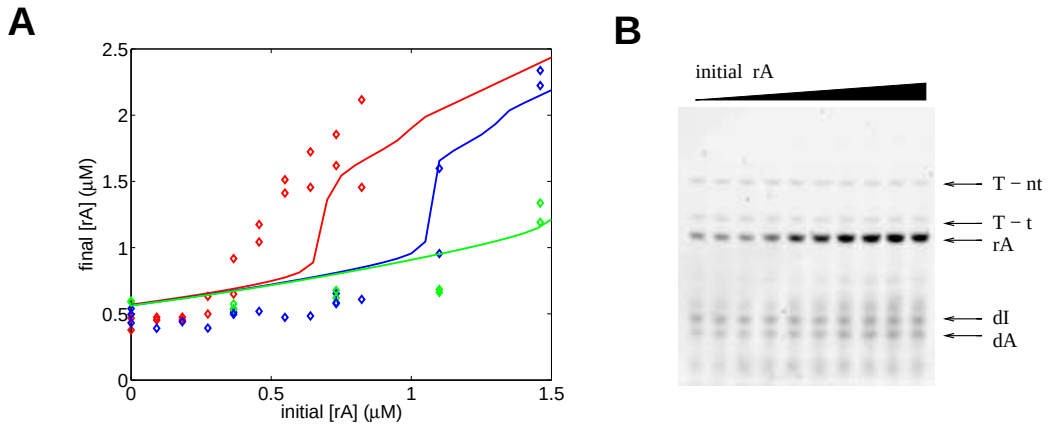


Figure 4.4. Threshold shift by adjusting dI. **(A)** Three thresholds have been demonstrated for [dI^{tot}]= 1.13 μM (red), 1.48 μM (blue), and 1.87 μM (green). The experimental results are plotted as diamonds and the simulation results are plotted as lines. **(B)** One of the duplicate gel results for [dI^{tot}]= 1.13 μM in (A). We could identify all single-stranded DNA and RNA species from this denaturing gel: T-nt and T-t together forms the OFF-state switch T; DNA activator dA; DNA inhibitor dI; and RNA activator rA. The initial RNA activator rA concentration increases from 0 nM to 900 nM (100 nM increase per lane).

Previous studies indicate that the accumulation of incomplete degradation product could interfere with proper hybridization reactions (Kim et al. 2006). We tested RNase R (gift of Dr. Deutscher), a 3' processive exoribonuclease to clean up the incomplete degradation products. We verified that RNase R degrades only isolated RNA activator rA (appendix 4.4.1). We expect that the ON-state switch can now achieve steady-state RNA activator concentration due to additional degradation of RNase R. The gel studies verified that the ON-state rA concentration approaches a steady-state and the difference of ON and OFF-state rA is significant (figure 4.6A). Thus, RNase R could clean up the incomplete degradation product as well as improving the system performance. Similarly, the maintenance of either ON or OFF state is verifed in the fluorometer (figure 4.6B) .

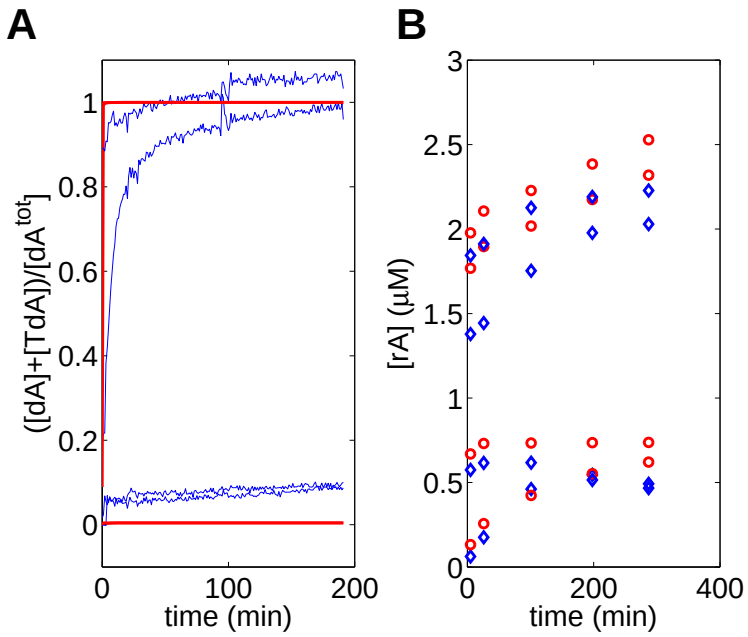


Figure 4.5. Fluorometer result. **(A)** Normalized fluorescence measures the fraction of dA that is not in a dA·dI complex. Blue lines are experimental measurements and red lines are simulation results. **(B)** Gel analysis of rA concentration sampled from fluorometer experiment at different times. Blue diamonds are experimental measurements and red circles are simulation results at corresponding time points.

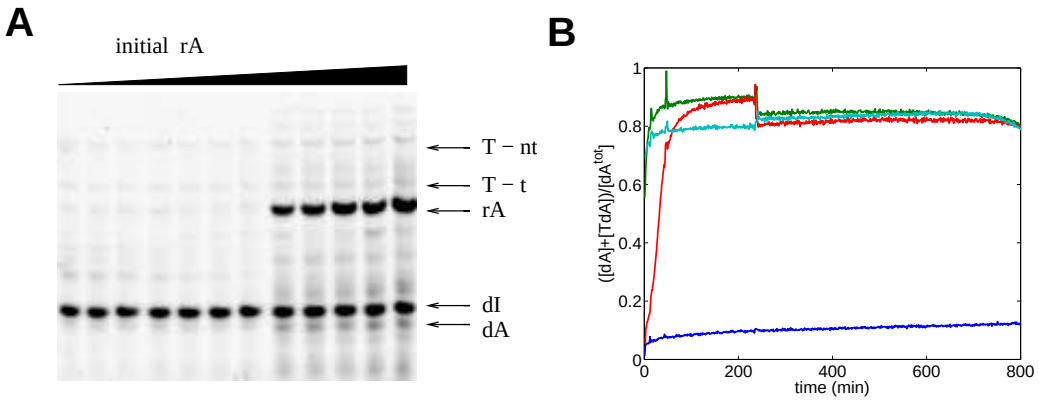


Figure 4.6. Gel and fluorometer results with RNase R. **(A)** Gel experiment. The dI concentration was 1000 nM, while initial rA concentrations were increased by 100 nM steps from left to right. The final rA concentrations were high if the initial rA concentrations were higher than 700 nM. **(B)** Fluorometer experiment. The dI concentration was 1000 nM, while initial rA concentrations were 700, 800, 900, and 1050 nM. The fluorescence remained low when the initial rA concentration was 700 nM (blue) and quickly reached the high signal for higher initial rA concentrations. The abrupt signal changes around 200 minutes were caused by disturbances due to sampling for gel studies.

4.2 Discussion

The simple mathematical model could reproduce the system behavior with reasonable accuracy. However, the experimentally determined switching threshold was typically lower than the expected

switching threshold from the simple model. We took two approaches to explain this phenomenon (appendix 4.4.2). First, we consider "initial burst" of RNAP activity (Jia and Patel 1997). With the burst effect, the RNA activator amount can increase quickly to overcome the threshold set by dI. However, the fitted burst parameters were typically a few fold larger than what we observed experimentally. Second, we consider additional sources of leaky expression. A plausible mechanism is activation of the OFF-state switch template by the dA·dI complex. Since the dA·dI complex has a weak 5-base toehold that can initiate branch migration reaction with an OFF-state template, T, we assume that a complex T·dA·dI can be formed. A T·dA·dI complex cannot be very stable because the enthalpy gain is only 5 additional base pairs while the entropy is lost upon binding of T and dA·dI. We modeled the binding reaction between T and dA·dI as a reversible reaction and also assumed that the complex T·dA·dI can transcribe RNA activator as efficiently as an ON switch T·dA. This fits nicely with the observation that leaky expression increases with more DNA activator dA, which cannot be explained by the simple model or with initial burst.

In summary, we verified that an autoregulatory switch with positive feedback could be constructed, exhibiting both bistable and monostable behavior. A simple mathematical model could capture most of the switch behavior. However, quantitative agreement to experimental data could not be obtained for certain parameter variations. Additional mechanisms such as the initial burst and an alternative source of leaky expression were included in the mathematical model to explain the experimental data and this extended model was used to generate figures. Even for our simplified biochemical circuit, identification of reaction mechanisms outside the originally postulated reaction mechanisms is a significant challenge. Systematic parameter variation and quantitative analysis is certainly helpful in identifying such missing links in the reaction network. Due to its simplicity, the *in vitro* transcriptional circuit can serve as a tool for characterizing various biochemical circuit designs and studying generic problems such as composability, performance, robustness, and efficiency.

4.3 Materials and methods

DNA oligonucleotides and enzymes

The sequence of all DNA molecules and expected RNA transcript sequences were chosen to minimize the occurrence of alternative secondary structures, checked by the Vienna group's DNA- and RNA-folding program (Flamm et al. 2000). All DNA oligonucleotides were purchased (Inte-

grated DNA Technologies, Coralville, Iowa, United States). The DNA activator, dA, is labeled with Cy5 at the 3' end, while the DNA inhibitor, dI, is labeled with IowaBlack-RQ at the 5' end. The T7 RNA polymerase (enzyme mix), transcription buffer, and NTP were purchased as part of the T7 Megashortscript kit (Ambion, Austin, Texas, United States; #1354). RNase H (Ambion; #2293) was purchased.

Transcription

Switch template (T-nt and T-t strands) were annealed with 10% (v/v) 10x transcription buffer from 90°C to 20°C over 1 hour at 5 times the final concentration used. To the annealed templates, activators and DNA or RNA inhibitors from a high concentration stock (50 μM), 7.5 mM each NTP, and 8 % (v/v) 10x transcription buffer were added. Transcription reactions for spectrofluorometer experiments were prepared as a total volume of 60 μL. Transcription reactions for gel studies were prepared as a total volume of 10 μL and were stopped by denaturing dye (80 % formamide, 10 mM EDTA, 0.01g XCFF).

Data acquisition

For spectrofluorometer experiments, excitation and emission for Cy5-labeled dA were at 648 nm and 665 nm. The fluorescence was recorded every minute using a SPEX Fluorolog-3 (Jobin Yvon, Edison, New Jersey, United States) and converted to dA fraction that is not bound with dI by normalizing against minimum fluorescence (measured before the addition of enzymes with excess quencher-labeled dI) and maximum fluorescence (measured prior to the addition of dI). Denaturing polyacrylamide gels (8% 19:1 acrylamide:bis and 7 M urea in TBE buffer) were allowed to run for 50 min with 10 V/cm at 65°C in TBE buffer (100 mM Tris, 90 mM Boric Acid, 1 mM EDTA). The 10-base DNA ladder (Invitrogen, Carlsbad, California, United States; #10821-015) was used in the control lane and the denaturing gel was stained with SYBR gold (Molecular Probes, Eugene, Oregon, United States; #S-11494) for quantitation. The gel data were quantitated using the Molecular imager FX (Biorad, Hercules, California, United States).

Model simulation

The kinetic simulations and parameter fittings were implemented in MATLAB. Differential equations were solved using the *ode23s* routine, while mean squared deviation of model fits on experimental data was minimized using the *fmincon* routine.

DNA sequences

T-nt (106mer), 5'-CATTAGTGTCGTTCGTTCACAGTAATACGACTCACTATAGGGAGAAACA-AAGAACGAACGACACTAATGAACTACTACTACACACTAATACTGACAAAGTCAGAAA-3'.

T-t (79mer), 5'-TTTCTGACTTTGTCAGTATTAGTGTGTAGTAGTAGTTCATTAGTGTCGTTC-GTTCTTTGTTTCTCCCTATAGTGAGTCG-3'.

dA (36mer), 5'-TATTACTGTGAACGAACGACACTAATGAACTACTAC-3'.

dI (38mer), 5'-GTGTGTAGTAGTAGTTCATTAGTGTCGTTCGTTCACAG-3'.

4.4 Appendix

4.4.1 Detailed study of reaction mechanisms

We have verified individual reaction mechanisms by gel: DNA/RNA hybridization reactions and branch migration reactions, and enzyme reactions.

First, DNA and RNA hybridization reactions and branch migration reactions are characterized by gel (figure 4.7). Comparing lanes 2 and 3 shows that activator dA binds to an OFF-state template T (activation). Comparing lanes 3 and 4 shows that inhibitor dI strips off an activator dA from an ON-state template T·dA (inhibition). Although an RNA activator rA can bind to an OFF-state template T (lane 5), it prefers binding to the DNA inhibitor dI and allows the DNA activator dA to bind to an OFF-state template T (lane 6). This verified that our design criterion of thermodynamic stability dictated the reaction mechanisms: rA·dI $>$ dA·dI $>$ T·dA $>$ T·rA.

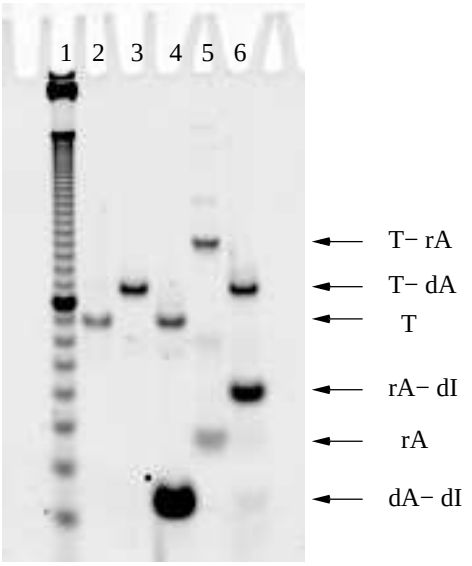


Figure 4.7. Nondenaturing gel characterization of hybridization reactions. Lane 1 is size marker, lane 2 contains 50 nM OFF-state switch T, lane 3 contains 50 nM T and 500 nM DNA activator dA, lane 4 contains 50 nM T, 500 nM dA, and 700 nM DNA inhibitor dI, lane 5 contains 50 nM T and 500 nM RNA activator rA, lane 6 contains 50 nM T, 500 nM rA, 700 nM dI, and 500 nM rA.

Second, the enzyme reactions are characterized for RNAP, RNase H, and RNase R, respectively (figure 4.8). RNAP transcribed RNA activator rA from the ON switch, T·dA, more than ten times faster than from the OFF switch, T. RNase H degraded purified RNA activator rA successfully when provided with DNA inhibitor dI. However, no degradation of rA by RNase H was observed when no dI was provided. In contrast, RNase R could degrade purified RNA activator rA completely when no DNA inhibitor dI was provided. Yet, rA·dI complex was apparently protected from degradation by RNase R. Taken together, the enzyme reactions used in our model are plausible without significant side reactions.

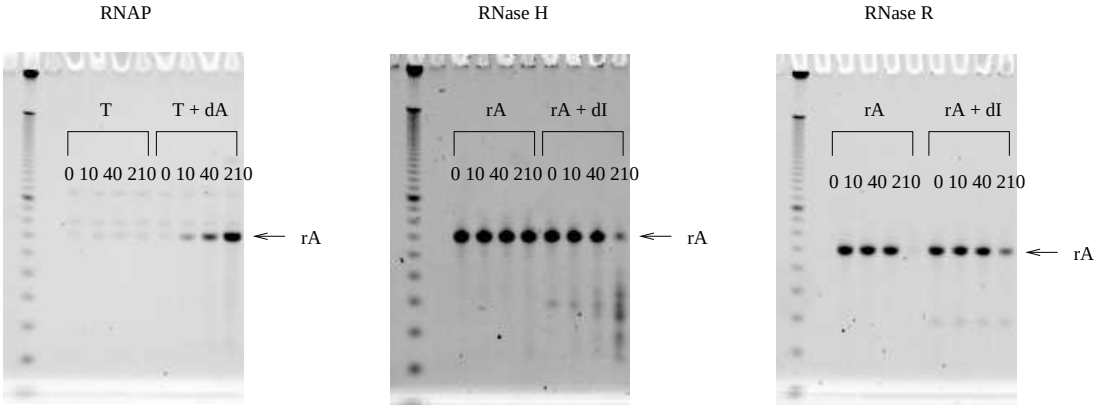


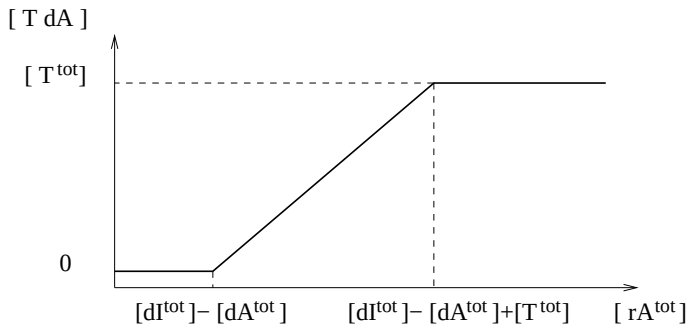
Figure 4.8. Denaturing gel characterization of enzyme reactions. For each enzyme reactions, samples are taken at 0, 10, 40 and 210 minutes to monitor the reaction time-courses. For the RNAP reactions, 60 nM OFF-state switch T alone and 60 nM OFF-state switch T with 200 nM DNA activator dA were used as templates for rA production. For RNase H and RNase R reactions, 3 μM RNA activator rA alone and 3 μM RNA activator rA with 1 μM DNA inhibitor dI were used as substrates.

4.4.2 Model equations

Based on the experimentally verified elementary reactions, we construct a simple mathematical model to explain the behavior of autoregulatory switch.

DNA/RNA hybridization and branch migration reactions

$$
\begin{aligned}
\mathrm{T} + \mathrm{dA} &\xrightarrow{k_{TA}} \mathrm{TdA} && \text{(Activation)} \\
\mathrm{dA} + \mathrm{dI} &\xrightarrow{k_{AI}} \mathrm{dAdI} && \text{(Annihilation)} \\
\mathrm{TdA} + \mathrm{dI} &\xrightarrow{k_{TAI}} \mathrm{T} + \mathrm{AdI} && \text{(Inhibition)} \\
\mathrm{rA} + \mathrm{dI} &\xrightarrow{k_{rAI}} \mathrm{rAdI} && \text{(Annihilation)} \\
\mathrm{rA} + \mathrm{dAdI} &\xrightarrow{k_{AI_rA}} \mathrm{rAdI} + \mathrm{dA} && \text{(Release)}
\end{aligned}
$$



The superscript *tot* indicates that all complexes containing that species are considered, e.g., [T^{tot}]=[T]+[T·dA] and [dA^{tot}]=[dA]+[T·dA]+[dA·dI]. With fast and irreversible hybridization assumptions, the DNA/RNA hybridization reactions lead to the above transfer curve with the total concentration of rA as an input and the concentration of active switch T·dA as an output. (1) [rA^{tot}]<[dI^{tot}]-[dA^{tot}], the RNA activator is consumed upon binding to free DNA inhibitor and does not activate the switch. (2) [dI^{tot}]-[dA^{tot}]<[rA^{tot}]<[dI^{tot}]-[dA^{tot}]+[T^{tot}], the RNA activator is enough to consume all free DNA inhibitor and release DNA activator bound to DNA inhibitor stoichiometrically. The DNA activator thus released will bind to and activate the switch. (3) [rA^{tot}]>[dI^{tot}]-[dA^{tot}]+[T^{tot}], the DNA inhibitor is all consumed upon binding to RNA activator, and the switch is completely ON.

Michaelis–Menten enzyme reactions

$$\begin{aligned}
\mathrm{RNAP} + \mathrm{TdA} &\underset{k_{-,ON}}{\overset{k_+}{\rightleftharpoons}} \mathrm{RNAP \cdot TdA} \overset{k_{cat,ON}}{\rightarrow} \mathrm{RNAP} + \mathrm{TdA} + \mathrm{rA} \\
\mathrm{RNAP} + \mathrm{T} &\underset{k_{-,OFF}}{\overset{k_+}{\rightleftharpoons}} \mathrm{RNAP \cdot T} \overset{k_{cat,OFF}}{\rightarrow} \mathrm{RNAP} + \mathrm{T} + \mathrm{rA} \\
\mathrm{RNaseH} + \mathrm{rAdI} &\underset{k_{-,H}}{\overset{k_{+,H}}{\rightleftharpoons}} \mathrm{RNaseH \cdot rAdI} \overset{k_{cat,H}}{\rightarrow} \mathrm{RNaseH} + \mathrm{dI} \\
\mathrm{RNaseR} + \mathrm{rA} &\underset{k_{-,R}}{\overset{k_{+,R}}{\rightleftharpoons}} \mathrm{RNaseR \cdot rA} \overset{k_{cat,R}}{\rightarrow} \mathrm{RNaseR}
\end{aligned}$$

We do not consider side-reactions or incomplete production and degradation products. The Michalis–Menten enzyme reactions are further simplified by the steady-state assumption of enzyme-substrate complex, which is approximately valid when enzyme concentrations are low compared to substrate concentrations. We express the available enzyme concentrations as follows:

$$[\mathrm{RNAP}] = \frac{[\mathrm{RNAP}^{tot}]}{1 + \sum \frac{[\mathrm{T \cdot dA}]}{K_{M,ON}} + \sum \frac{[\mathrm{T}]}{K_{M,OFF}}}, \qquad [\mathrm{RNaseH}] = \frac{[\mathrm{RNaseH}^{tot}]}{1 + \frac{[\mathrm{rA \cdot dI}]}{K_{M,H}}},$$

$$[\mathrm{RNaseR}] = \frac{[\mathrm{RNaseR}^{tot}]}{1 + \frac{[\mathrm{rA}]}{K_{M,R}}},$$

where the Michaelis constants are calculated as $K_M = \frac{k_- + k_{cat}}{k_+}$ to determine the affinity of substrates to the enzymes.

Thus, the dynamics of the system is described by the following four ordinary differential equations:

$$\begin{aligned}
\frac{d[\mathrm{T}]}{dt} &= -k_{TA}[\mathrm{T}][\mathrm{dA}] + k_{TAI}[\mathrm{TdA}][\mathrm{dI}], \\
\frac{d[\mathrm{dA}]}{dt} &= -k_{AI}[\mathrm{A}][\mathrm{dI}] - k_{TA}[\mathrm{T}][\mathrm{dA}] + k_{AIrA}[\mathrm{dAdI}][\mathrm{rA}], \\
\frac{d[\mathrm{dI}]}{dt} &= -k_{AI}[\mathrm{A}][\mathrm{dI}] - k_{rAI}[\mathrm{rA}][\mathrm{dI}] - k_{TAI}[\mathrm{TdA}][\mathrm{dI}] + \frac{k_{cat,H}}{K_{M,H}}[\mathrm{RNaseH}][\mathrm{rAdI}], \\
\frac{d[\mathrm{rA}]}{dt} &= -k_{rAI}[\mathrm{rA}][\mathrm{dI}] - k_{AIrA}[\mathrm{dAdI}][\mathrm{rA}] + \frac{k_{cat,ON}}{K_{M,ON}}[\mathrm{RNAP}][\mathrm{TdA}] + \frac{k_{cat,OFF}}{K_{M,OFF}}[\mathrm{RNAP}][\mathrm{T}] \\
&\quad - \frac{k_{cat,R}}{K_{M,R}}[\mathrm{RNaseR}][\mathrm{rA}].
\end{aligned}$$

The remaining variables, [T·dA], [dA·dI], and [rA·dI] are calculated from mass conservation.

The model equations are further modified to include other mechanisms to explain the experi-

mental data more quantitatively. For the initial burst mechanism, the RNAP catalytic rate constants (for both ON or OFF state templates) are modified as follows:

$$k_{cat,burst} = k_{cat}(1 + Ae^{-t/\tau}),$$

where A is the burst amplitude and τ is the decay constant. We estimated A and τ experimentally by following the time-course of RNA transcript accumulation in the absence of degradation. We observed A <10 and τ <10 min. These are used as upper bounds of model parameters.

We observed increased leaky expression with increased amount of DNA activator dA, which could not be explained by the basic model with initial burst. Thus, the inhibition mechanism is modified to include the complex T·dA·dI as an intermediate species whose dissociation to T and dA·dI is considered reversible.

$$\mathrm{T \cdot dA + dI} \overset{k_{TAI}}{\rightarrow} \mathrm{T \cdot dA \cdot dI} \underset{k_{ass}}{\overset{k_{dis}}{\rightleftharpoons}} \mathrm{T + dA \cdot dI}$$

This requires the addition of T·dA·dI species and modification of T and rA rate calculations as follows:

$$\begin{aligned}
\frac{d[\mathrm{TdAdI}]}{dt} &= k_{TAI}[\mathrm{TdA}][\mathrm{dI}] + k_{ass}[\mathrm{T}][\mathrm{dAdI}] - k_{dis}[\mathrm{TdAdI}], \\
\frac{d[\mathrm{T}]}{dt} &= -k_{TA}[\mathrm{T}][\mathrm{dA}] + k_{dis}[\mathrm{TdAdI}] - k_{ass}[\mathrm{T}][\mathrm{dAdI}], \\
\frac{d[\mathrm{rA}]}{dt} &= -k_{rAI}[\mathrm{rA}][\mathrm{dI}] - k_{AIrA}[\mathrm{dAdI}][\mathrm{rA}] + \frac{k_{cat,ON}}{K_{M,ON}}[\mathrm{RNAP}]([\mathrm{TdA}] + [\mathrm{TdAdI}]) \\
&\quad + \frac{k_{cat,OFF}}{K_{M,OFF}}[\mathrm{RNAP}][\mathrm{T}] - \frac{k_{cat,R}}{K_{M,R}}[\mathrm{RNaseR}][\mathrm{rA}].
\end{aligned}$$

Table 4.1. Model parameters

	i=2, j=1	Other studies
$K_{M,ON}$ (nM)	235	15-37
$k_{cat,ON}$ (/s)	0.056	0.73-1.12
$K_{M,OFF}$ (nM)	370	0.1-1.1
$k_{cat,OFF}$ (/s)	0.019	0.11-0.18
$K_{M,H}$ (μM)	2.53	16-130
$k_{cat,H}$ (/s)	0.21	0.02-0.6
k_{TA} (/M/s)	$5.73*10^5$	-
k_{AI} (/M/s)	$3.60*10^4$	-
k_{rAdI} (/M/s)	$1.29*10^5$	-
k_{TAI} (/M/s)	$3.34*10^4$	-
k_{AIrA} (/M/s)	$1.29*10^5$	-
A_{ON}	9.9	-
τ_{ON} (s)	200	-
A_{OFF}	9.8	-
τ_{OFF} (s)	266	-
k_{ass} (/M/s)	$1.92*10^5$	-
k_{dis} (/s)	0.05	-

Chapter 5 Oscillators

5.1 Results

Most living organisms, from cyanobacteria to plants, insects, and mammals, are capable of diplaying spontaneous sustained oscillations with a period close to 24 hr (Young and Kay 2001; Dunlap 2006; Reppert and Weaver 2002). In all cases, the molecular mechanism of circadian oscillations relies on negative autoregulation of gene expression. For example, in *Drosophila*, the PER (period) and TIM (timeless) proteins form a complex that indirectly represses the activation of the *per* and *tim* genes. A positive regulation is also found. The PER-TIM complex derepresses the expression of the *clock* gene and CLOCK in turn activates the expression of the *per* and *tim* genes. Minimal gene circuitry required for oscillation in biochemical systems can be quite simple. The circadian clock of cyanobacteria requires assembly of only three proteins and ATP fuel (Nakajima et al. 2005). A synthetic ring oscillator has been demonstrated in *E. coli* with just three regulatory genes (Elowitz and Leibler 2000). However, high variability of the synthetic ring oscillator with purely inhibitory connections suggests that additional form of control may be necessary for noise reduction (Barkai and Leibler 2000). A synthetic design closer to that of Barkai and Leibler (2000) showed a longer-period damped oscillation at the population level (Atkinson et al. 2003). A direct comparison between different synthetic oscillator designs is challenging due to differences in their regulatory components as well as potential inteference with the cellular network.

Here, we use our *in vitro* transcription network approach to construct and characterize alternative oscillator designs. The construction of model oscillators with different connectivity opens up the possibility of quantitative analysis and model evaluation to reveal design principles in biochemical networks. We have demonstrated sigmoidal inhibitory regulation (Kim et al. 2006) and excitatory regulation (chapter 4) with adjustable thresholds. The construction of a synthetic oscillator demonstrates composability of simple programmable switches. We use competitive binding of nucleic acid species rather than cooperative binding to achieve ultrasensitivity. The threshold in our transcriptional circuit derives from four types of strong DNA and RNA hybridization reactions (appendix 5.4.1), which we call activation, annihilation, inhibition, and release. An activator binds to an OFF switch template to turn the switch on (activation); a DNA or RNA activator binds to an

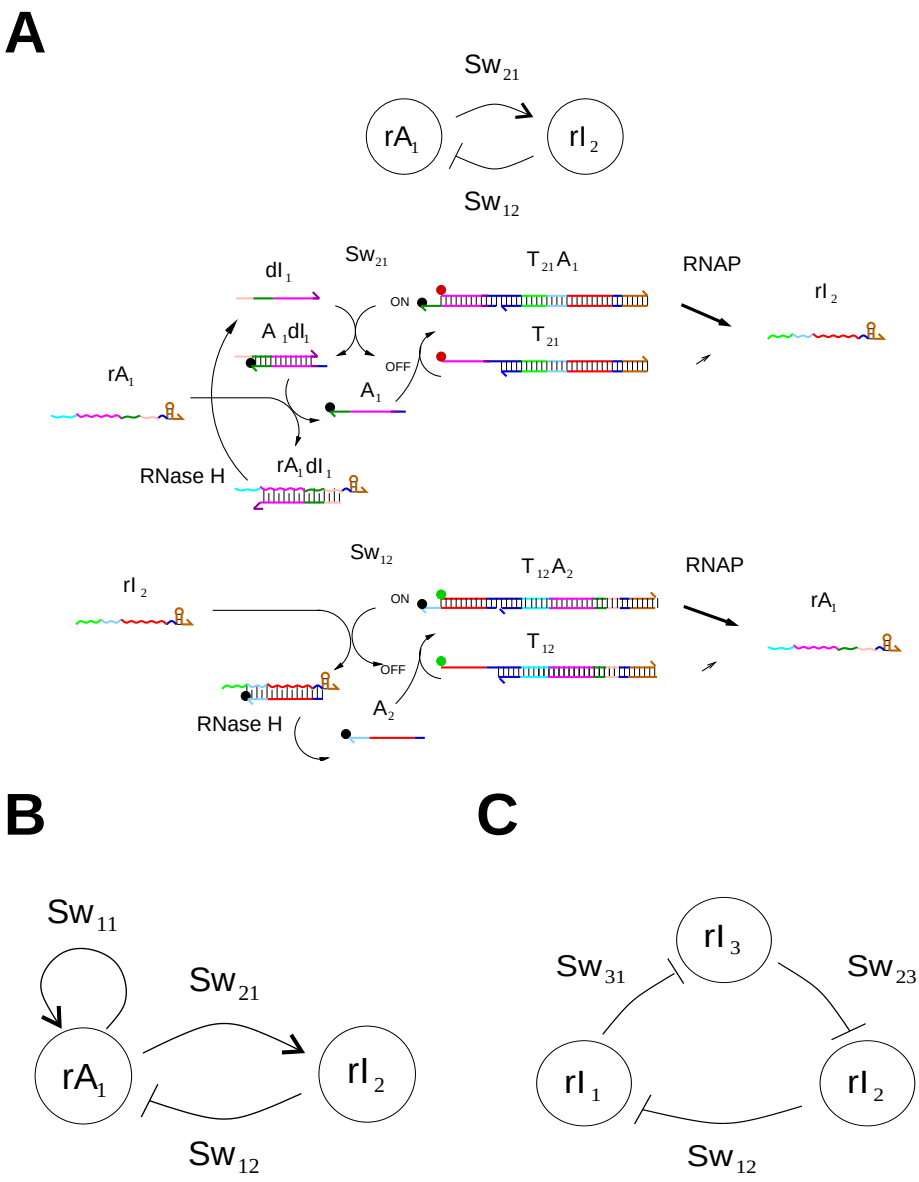
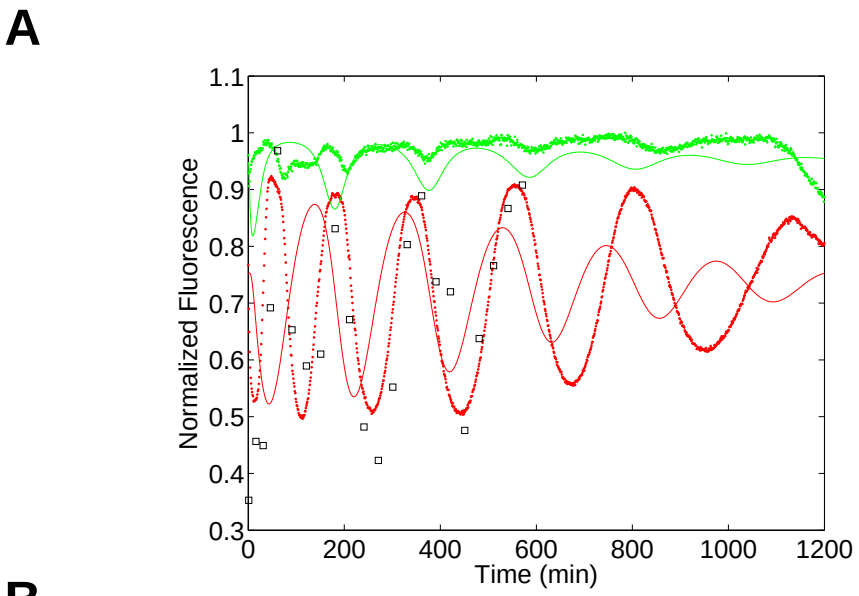


Figure 5.1. Design of synthetic oscillators. **(A)** A two-node oscillator. The *trans*-activating synthetic switch, Sw_{12}, and the *trans*-inhibiting synthetic switch, Sw_{21}, are connected to form a delayed negative feedback loop. T_{21} is labeled with Texas Red (red circle) at 5' end and A_1 is labeled with Iowablack-RQ (black circle) at 3' end such that the OFF state of Sw_{21} emits high fluorescence signal in red channel. T_{12} is labeled with TAMRA (green circle) at 5' end and A_2 is labeled with Iowablack-RQ (black circle) at 3' end such that the OFF state of Sw_{12} emits high fluorescence signal in green channel. The switch states will be monitored real-time as fluorescence change. **(B)** A two-node oscillator with positive feedback. We modified the two-node oscillator by adding of a positive feedback on rA_1. **(C)** A ring oscillator. We modified the two-node oscillator by replacing excitatory connection formed by Sw_{12} with two inhibitory connections formed by Sw_{13} and Sw_{32}.

inhibitor and is not available for the switch template (annihilation); an inhibitor displaces an activator from an ON switch template, the T·A complex, to turn the switch off (inhibition); an RNA



activator displaces a DNA activator from an A·dI complex (release). The key requirements for the inhibition and release mechanisms are that the resulting complex is thermodynamically more favorable than the starting complex and that there is a fast kinetic pathway to the lowest energy state (in our case, toehold-mediated strand displacement (Yurke and Mills 2003)). For example, we use an excitatory connection by switch Sw_{21} (the subscripts on switch Sw_{ij} means that it is controlled by

Figure 5.2. Detailed characterization of two-node oscillator **(A)** The fluorescence time-courses measured in fluorometer. TAMRA signal measures the fraction of switch Sw_{12} that is in the OFF state. Texas Red signal measures the fraction of switch Sw_{21} that is in the OFF state. The experimental results are plotted as lines and the model fits are plotted as dotted lines. The nondenaturing gel results in (B) are plotted as squares. **(B)** Denaturing and nondenaturing gel of the fluorometer experiment in (A). We measured the total RNA concentrations in the denaturing gel. The RNA activator rA_1 concentration was not much higher than the background signal and could not be measured reliably. We measured the OFF-state switch concentrations in the nondenaturing gel. The switch Sw_{12} showed little variation compared to the switch Sw_{21}, as expected from the fluorescence data in (A). The nondenaturing gel data for later time-courses could not be measured reliably due to high background. **(C)** We assume that the band of 35 nucleotides in the denaturing gel is representative of accumulating incomplete degradation products. If the short degradation products interfere with the proper hybridization of rI_2, it effectively nullifies part of inhibitory signals. Subtracting a fraction of incomplete degradation product from the total rI_2 product can be a better representation of effective RNA inhibitor concentration.

RNA species j and produces RNA species i) and an inhibitory connection by switch Sw_{12} for our two-node oscillator design. For this design, RNA activator rA_1 activates the production of RNA inhibitor rI_2, which in turn inhibits the production of rA_1. For the excitatory connection mediated by switch Sw_{21}, the initial amount of DNA inhibitor dI_1 sets the threshold for activation. On the other hand, for the inhibitory connection mediated by switch Sw_{12}, the initial amount of DNA activator dA_2 sets the threshold for inhibition. Since the activator and the inhibitor annihilate each other, the difference of total activator (DNA or RNA) and inhibitor concentrations is the most important determinant of the state of switch: an excess of inhibitor will turn the switch off while an excess of activator will turn the switch on. We can build a simple model to explain the system behavior (appendix 5.4.1).

5.1.1 Two-node oscillator

Negative feedback with proper delay can lead to oscillation in genetic circuits (Elowitz and Leibler 2000). We constructed a two-node oscillator where RNA activator rA_1 activates the production of RNA inhibitor rI_2 by modulating switch Sw_{21} and RNA inhibitor rI_2 in turn inhibits the production of RNA activator rA_1 by modulating switch Sw_{12}. For a negative-feedback oscillator, more than three independent reaction components are required for oscillation (Griffith 1968). Although we have only two RNA species as dynamic signals, they hybridize with other DNA species and influence switch states which determine the strength of excitatory and inhibitory connections. Additional delays are introduced because the enzymatically-controlled RNA signal production and degradation are slow. Prior to experimental characterization, we used a mathematical model to search for rea-

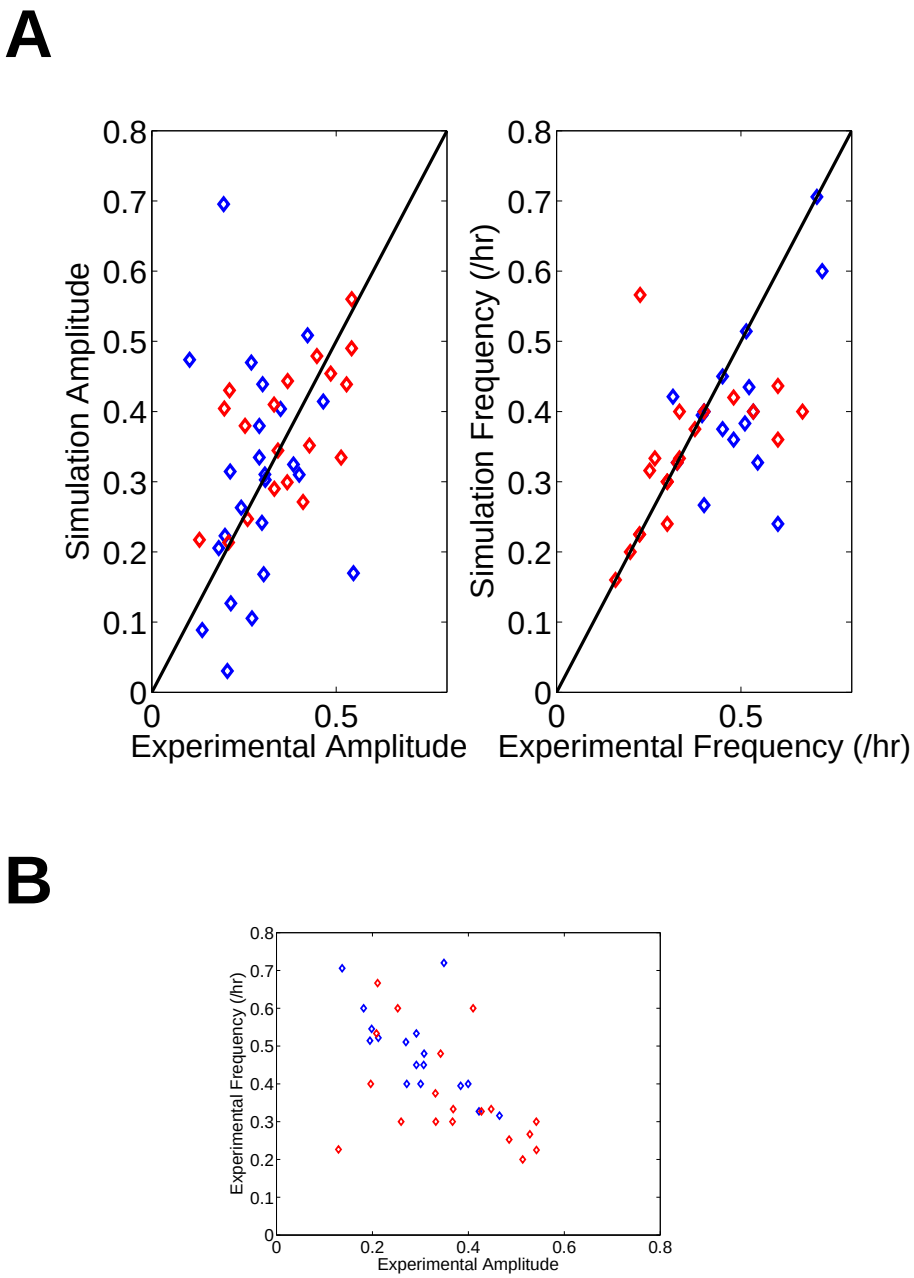


Figure 5.3. Model fits of two-node oscillator. **(A)** Amplitude and frequency in experimental and simulation results. Compared to data set 1 (red), data set 2 (blue) shows larger variation because it was not used for model fitting (figures 5.6 and 5.7). Black lines indicate ideal one-to-one correspondence between experiment and simulation. **(B)** Amplitude and frequency trade-off in experimental results. Although amplitude and frequency of oscillation can be tuned for various experimental conditions, high amplitude and high frequency oscillation is difficult to achieve.

sonable parameter space and found a sustained oscillatory region. The model analysis indicates that the system is not sensitive to most hybridization parameters. The initial choice of experimental parameters guided by the mathematical model resulted in a damped oscillation. Further exploration

of experimental parameter space lead to a sustained oscillation. We monitor the state of switches by fluorescence measurement because the ON-state switch has low fluorescence and OFF-state switch has high fluorescence (figure 5.2A). At the same time, the state of switches and concentration of RNA signals were measured by gel (figure 5.2B). The nondenaturing gel studies of switch states agree with the fluorescence signal change quantitatively. The system showed up to seven distinct oscillations before the production rate cannot be sustained due to exhaustion of NTP fuel and buffer. Unlike other synthetic oscillators, we can change amplitude and frequency relatively easily with trade-offs (figure 5.3B). For example, increasing RNase H concentration increases oscillation frequency but decreases oscillation amplitude. The model fit captures the dependence of oscillation amplitude and frequency with reasonable accuracy (figure 5.3A). We initiated the reaction with different combinations of RNA signals and ask whether the oscillations approach the same limit cycle (figure 5.6). Initiated with no RNA signals or RNA signals that brings the switch states to various locations in the phaseplane, the system approaches similar oscillation trajectories. Interestingly, the RNA inhibitor rI_2 concentration builds up over time as the reaction progresses. We hypothesize that the short fragments of rI_2 generated by degradation process may interfere with proper hybridization reaction of rI_2 signals to its regulatory target switch Sw_{12} and hence more signals are needed to overcome the interference. We could estimate the concentration of short degradation products from the denaturing gel which shows linear increase over time. The model could reproduce similar behavior when we took this reaction mechanism into account (figure 5.2C).

5.1.2 Alternative oscillator designs – autoregulatory positive feedback and purely negative feedback

The flexible architecture of our synthetic transcriptional network allowed us to characterize alternative oscillator designs. We constructed an oscillator analogous to a core circadian oscillator proposed by Barkai and Leibler (2000) where RNA activator promotes its own production through a positive feedback in addition to the negative feedback implemented in the two-node oscillator. The oscillator design with both positive and negative feedback is prevalent in nature and is theoretically shown to be resistant to noise (Vilar et al. 2002). For example, the bistable feature of a positive feedback loop is important for sustained oscillation in cell cycles of *Xenopus laevis* embryo (Pomerening et al. 2005). An analogous synthetic relaxation oscillator constructed in *E. coli*, however, showed damped oscillation (Atkinson et al. 2003). The addition of the positive feedback element reinforces the excitatory signal of rA_1 such that sustained oscillations are feasible with weaker activation mod-

ule (figure 5.4). A stronger positive feedback element increases both the oscillation period and amplitude. However, too much emphasis on the positive feedback will potentially lock the system in a permanant excited state because the swich Sw_{11}, which mediates the positive feedback, cannot turn itself off in our design.

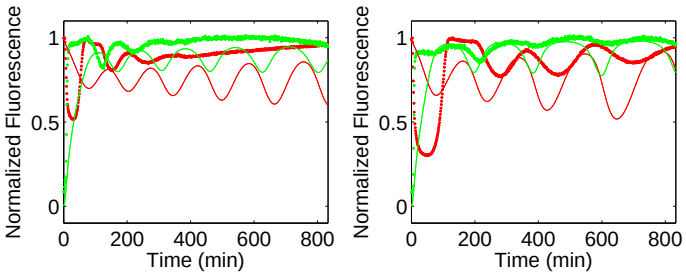


Figure 5.4. Two-node oscillator with positive feedback. Fluorescence time-courses and model fits. We monitor the states of switch Sw_{12} and switch Sw_{21} as before. It is expected that the state of switch Sw_{11} is similar to the state of switch Sw_{21} because they have the identical regulatory domain. The addtion of positive feedback shifted a damped oscillation parameter region to sustained oscillation parameter region. However, the model failed to reproduce the damped oscillation.

We replaced an excitatory connection by two inhibitory connections to increase delay. This oscillator design is analogous to the synthetic ring oscillator (Elowitz and Leibler 2000) that showed sustained oscillation at the single cell level. Initially, we assumed that all three inhibitory connections between switches are equivalent. The choice of experimental parameters with this assumption corresponded to a strongly damped oscillation. Characterization of individual inhibitory connections showed that the inhibitory connection formed by Sw_{12} is relatively strong, possibly due to fluorophore-quencher interaction stabilizing the ON state of switch Sw_{12} (Marras et al. 2002). Thus, we decreased the concentration of activator A_2 and increased the concentration of activator A_3 to balance the strength of inhibitory connections to achieve a sustained oscillation. When initiated with no externally supplied RNA inhibitors, the first oscillation amplitude is small. Yet, the second and third oscillations become stable cycles with larger amplitudes (figure 5.5A). We could observe four cycles before the system ran out of fuel. Although we monitor only one switch state, the gel studies confirm that all three RNA signals vary periodically (figure 5.5B and C). Model studies indicate that our ring oscillator design admits stable limit cycle oscillations in a much narrower parameter space compared to our two-node oscillator design. This observation is possibly due to the differences of degradation processes in our model and in other synthetic network designs. Unlike the *in vivo* synthetic oscillators where bacteria are in an exponentially growing phase or in chemostats (Elowitz and Leibler 2000; Atkinson et al. 2003), our degradation mechanisms mediated by RNase H operate

in a closed system. Thus, the degradation machinery shows saturation when the amount of RNA product increases. Consequently, to maintain a stable amount of RNA signals, it is helpful to have a resting stage where all RNA signal productions are temporarily shut off, as in the two-node oscillator design. The limited degradation capacity in our ring oscillator design can cause exponential slowdown of oscillation due to saturation of the degradation machinery.

5.2 Discussion

The simple mathematical model could capture most of oscillator behavior quantitatively. The comparison of the model prediction on parameter dependence of oscillation amplitude and period (figure 5.3A) is reasonable for many parameter variations. Qualitative information on reaction kinetics was sufficient to obtain damped oscillation in the experiment and further exploration of parameter dependence lead to sustained oscillations in our oscillator designs. Adjusting weights and thresholds encoded as continuously tunable concentrations of nucleic-acid species rather than affinity of protein molecules greatly facilitates probing parameter space. The construction of alternative oscillator designs with similar constraints such as limited amount of fuel (NTP), gene (DNA switches), and regulatory target (DNA activator or inhibitor thresholds) facilitates direct comparison of performance. For example, given similar constraints, the ring oscillator has a narrower parameter region that admits stable limit cycle oscillations and shows slower and larger oscillation compared to the two-node oscillator. The detailed characterization of oscillators revealed the importance of waste product management. It is surprising that our oscillator can operate in the presence of the waste products of the worst kind, the subsequence of effective signals. The addition of waste product interaction mechanisms to the mathematical model could explain the observed experimental behavior where the RNA signal increased throughout the reaction.

A hallmark of circadian rhythms, temperature compensation, is implemented at the molecular level for the cyanobacteria circadian clock by KaiC autophosphorylation and autodephosphorulation rates that are insensitive to temperature change (Tomita et al. 2005). For our synthetic oscillators, the RNAP catalytic rate depends sensitively on temperature (Maslak and Martin 1993) and hence temperature compensation was not achieved. An interesting question is how to achieve regulatory compensation of temperature changes, hence kinetic parameter changes, by network designs. The current implementation of our transcriptional network have finite operation time due to the decrease of enzyme activity and NTP fuel as well as the accumulation of waste products. There is a trade-off

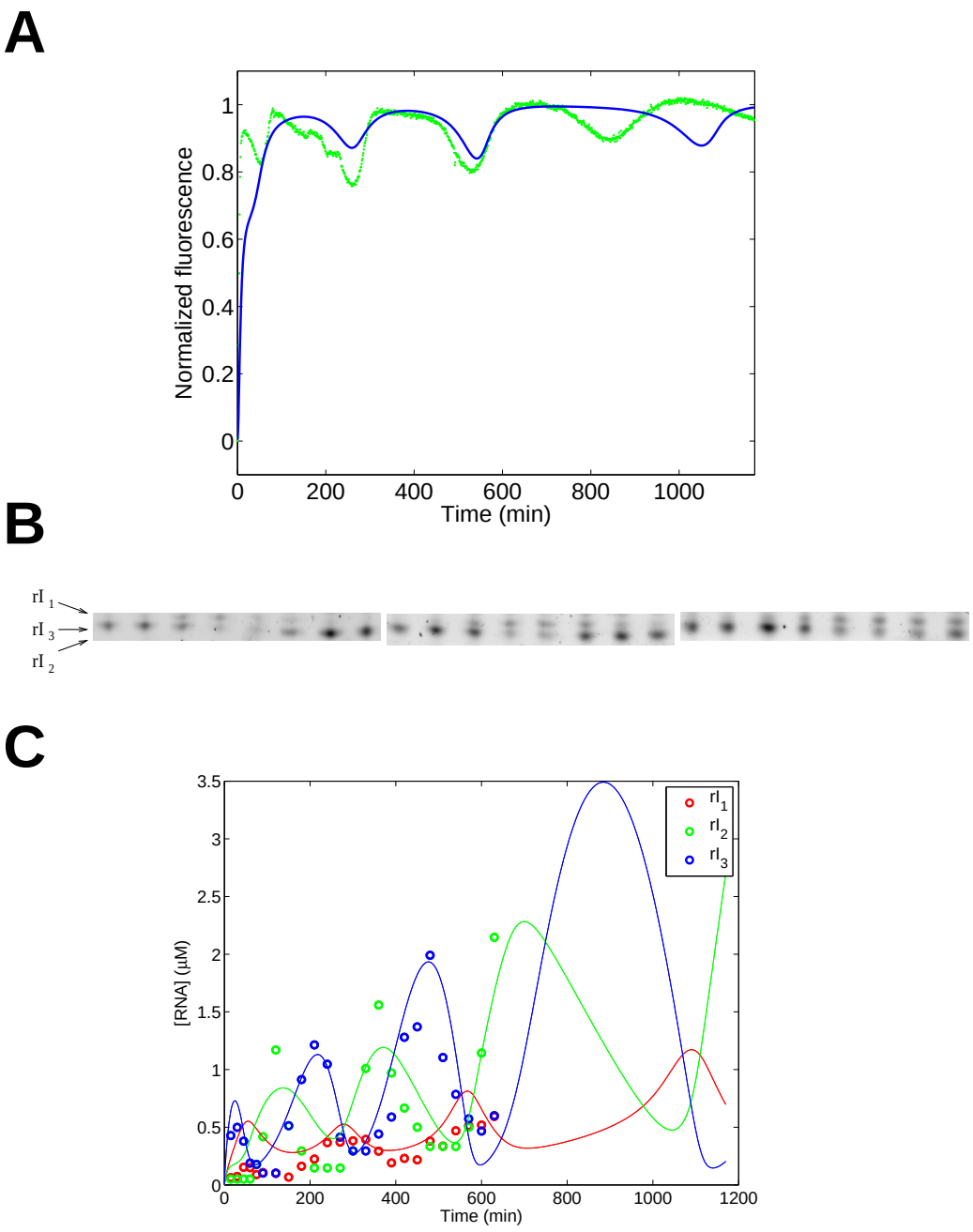


Figure 5.5. Ring oscillator. **(A)** Fluorescence time-courses and model fits. Experimental time-course is plotted as green line and simulation time-course is plotted as blue line. We monitor only the state of switch Sw_{12} since the other two switches are unlabeled. **(B)** Denaturing gel of the fluorometer experiment in (A). We measured the total RNA concentrations in the denaturing gel. The RNA inhibitor rI_1, rI_2, and rI_3 concentrations change periodically in the order we expect from their regulatory connections. The RNA inhibitors, rI_2 and rI_3, are only two nucleotides apart such that their peaks did not separate clearly for lanes 8, 15, and 16. **(C)** Gel data (circles) and model fits (lines).

between the speed of operation and the total operation time. With more genes and enzymes, NTP fuel exhaustion and waste product accumulation will have significant impact on reaction kinetics in a relatively short time. Some of the known side reactions can be suppressed by experimental design, as in this work, and other constraints such as the exhaustion of fuel (Klungsøyr et al. 1968) and build-up of degradation products could be relaxed in a chemostat (Atkinson et al. 2003), a dialysis bag (Madin et al. 2000), or vesicles (Noireaux and Libchaber 2004). In principle, the *in vitro* transcriptional network are suitable for stochatic noise characterization in small volumes. It would be interesting to decrease the molecular numbers of different oscillator designs to those typically found in cells and ask whether the stochastic noise of analogous *in vivo* circuits is intrinsic to the system or extrinsic to the system (Elowitz et al. 2002). It is challenging to ask a similar question by scaling up the *in vivo* synthetic circuits. Analyzing the stochastic noise contribution is important because the stochastic noise *in vivo* can drive the biochemical system out of stable attractors and cause oscillation (Vilar et al. 2002). The *in vitro* oscillator can be generalized to explore biochemical design principles and used to drive nucleic-acid-based nanomachines (Dittmer and Simmel 2004) and artificial cells (Noireaux and Libchaber 2004).

5.3 Materials and methods

DNA oligonucleotides and enzymes

The sequence of all DNA molecules and expected RNA transcript sequences were chosen to minimize the occurrence of alternative secondary structures, checked by the Vienna group's DNA and RNA folding program (Flamm et al. 2000). All DNA oligonucleotides were purchased (Integrated DNA Technologies, Coralville, Iowa, United States). T_{21}-nt is labeled with Texas Red at the 5' end, T_{12}-nt is labeled with TAMRA at the 5' end, A_1 and A_2 are labeled with Iowablack-RQ at the 3' end. The T7 RNA polymerase (enzyme mix), transcription buffer, and NTP were purchased as part of the T7 Megashortscript kit (Ambion, Austin, Texas, United States; #1354). RNase H (Ambion; #2293) was purchased.

Transcription

Switch templates (T-nt and T-t strands) were annealed with 10 % (v/v) 10x transcription buffer from 90°C to 20°C over 1 hour at 5 times the final concentration used. To the annealed templates, activators and DNA or RNA inhibitors from a high concentration stock (50 μM), 10.75 mM each NTP (1.5 times the suggested amount of the kit), 8 % (v/v) 10x transcription buffer, 5 % (v/v)

300 mM $MgCl_2$, were added. Transcription reactions for spectrofluorometer experiments were prepared as a total volume of 60 μL. Samples for gel studies were stopped by denaturing dye (80 % formamide, 10 mM EDTA, 0.01g XCFF).

Data acquisition

For spectrofluorometer experiments, excitation and emission for TAMRA-labeled T_{12} were at 559 nm and 580 nm, while excitation and emission for Texas Red–labeled T_{21} were at 597 nm and 615 nm. The fluorescence was recorded every minute using a SPEX Fluorolog-3 (Jobin Yvon, Edison, New Jersey, United States) and converted to switch activity by normalizing against minimum fluorescence (measured before the addition of enzymes with excess quencher labeled activators) and maximum fluorescence (measured at the end of reaction with excess DNA inhibitors to displace activators). Denaturing polyacrylamide gels (8% 19:1 acrylamide:bis and 7 M urea in TBE buffer) were allowed to run for 50 min with 10 V/cm at 65°C in TBE buffer (100 mM Tris, 90 mM Boric Acid, 1 mM EDTA). The 10-base DNA ladder (Invitrogen, Carlsbad, California, United States; #10821-015) was used in the control lane and the denaturing gel was stained with SYBR gold (Molecular Probes, Eugene, Oregon, United States; #S-11494) for quantitation. The gel data were quantitated using the Molecular imager FX (Biorad, Hercules, California, United States).

Model simulation

The kinetic simulations and parameter fittings were implemented in MATLAB. Differential equations were solved using the *ode23s* routine, while mean squared deviation of model fits on experimental data was minimized using the *fmincon* routine. The fluorescence time-courses and gel data were used for parameter fitting.

DNA sequences

T_{21}-nt (101mer), 5'-CATTAGTGTCGTTCGTTCACAGTAATACGACTCACTATAGGGAGAGTAAAACGGATTGAAGCAAGGGTAAGATGGAATGATAATACTGACAAAGTCAGAAA-3'.

T_{21}-t (74mer), 5'-TTTCTGACTTTGTCAGTATTATCATTCCATCTTACCCTTGCTTCAATCCGTTTTACTCTCCCTATAGTGAGTCG-3'.

A_1 (36mer), 5'-TATTACTGTGAACGAACGACACTAATGAACTACTAC-3'.

dI_1 (38mer), 5'-GTGTGTAGTAGTAGTTCATTAGTGTCGTTCGTTCACAG-3'.

T_{12}-nt (106mer), 5'-AAGCAAGGGTAAGATGGAATGATAATACGACTCACTATAGGGAGAAACAAAGAACGAACGACACTAATGAACTACTACTACACACTAATACTGACAAAGTCAGAAA-3'.

T_{12}-t (79mer), 5'-TTTCTGACTTTGTCAGTATTAGTGTGTAGTAGTAGTTCATTAGTGTCGT-

TCGTTCTTTGTTTCTCCCTATAGTGAGTCG-3'.

A_2 (35mer), 5'-TATTATCATTCCATCTTACCCTTGCTTCAATCCGT-3'.

T_{31}-nt (103mer), 5'-CTAATGAACTACTACTACACACTAATACGACTCACTATAGGGAGATCA-AATTTACAACGCAACTAACATATAATCGAAGACTTAATACTGACAAAGTCAGAAA-3'.

T_{31}-t (76mer), 5'-TTTCTGACTTTGTCAGTATTAAGTCTTCGATTATATGTTAGTTGCGTTGTA-AATTTGATCTCCCTATAGTGAGTCG-3'.

A_3 (35mer), 5'-TATTAAGTCTTCGATTATATGTTAGTTGCGTTGTA-3'.

T_{23}-nt (101mer), 5'-AACTAACATATAATCGAAGACTTAATACGACTCACTATAGGGAGAGTA-AAACGGATTGAAGCAAGGGTAAGATGGAATGATAATACTGACAAAGTCAGAAA-3'.

5.4 Appendix

5.4.1 Model equations

Two-node oscillator

We present a simple model for DNA and RNA hybridization reactions, branch migration reactions, and Michaelis–Menten enzyme reactions in the transcriptional circuit (i,j)=(1,2),(2,1):

DNA/RNA hybridization and branch migration reactions

$$\begin{aligned}
\mathrm{T}_{ij} + \mathrm{A}_j &\overset{k_{TAj}}{\rightarrow} \mathrm{T}_{ij}\mathrm{A}_j && (\text{Activation}),\\
\mathrm{A}_j + (\mathrm{d},\mathrm{r})\mathrm{I}_j &\overset{k_{AIj}}{\rightarrow} \mathrm{A}_j(\mathrm{d},\mathrm{r})\mathrm{I}_j && (\text{Annihilation}),\\
\mathrm{T}_{ij}\mathrm{A}_j + (\mathrm{d},\mathrm{r})\mathrm{I}_j &\overset{k_{TAIj}}{\rightarrow} \mathrm{T}_{ij} + \mathrm{A}_j(\mathrm{d},\mathrm{r})\mathrm{I}_j && (\text{Inhibition}),\\
\mathrm{rA}_1 + \mathrm{dI}_1 &\overset{k_{rAI1}}{\rightarrow} \mathrm{rA}_1\mathrm{dI}_1 && (\text{Annihilation}),\\
\mathrm{rA}_1 + \mathrm{A}_1\mathrm{dI}_1 &\overset{k_{AIrA1}}{\rightarrow} \mathrm{rA}_1\mathrm{dI}_1 + \mathrm{A}_1 && (\text{Release}),
\end{aligned}$$

where (d, r) is used to indicate that the reaction schemes are applicable for both dI_1 and rI_2.

Michaelis–Menten enzyme reactions

$$\begin{aligned}
\mathrm{RNAP} + \mathrm{T}_{ij}\mathrm{A}_j \quad &\underset{k_{-,ON,ij}}{\overset{k_+}{\rightleftharpoons}} \quad \mathrm{RNAP}\cdot\mathrm{T}_{ij}\mathrm{A}_j \quad \overset{k_{cat,ON,ij}}{\rightarrow} \quad \mathrm{RNAP} + \mathrm{T}_{ij}\mathrm{A}_j + \mathrm{rA}_1 \text{ or } \mathrm{rI}_2 \\
\mathrm{RNAP} + \mathrm{T}_{ij} \quad &\underset{k_{-,OFF,ij}}{\overset{k_+}{\rightleftharpoons}} \quad \mathrm{RNAP}\cdot\mathrm{T}_{ij} \quad \overset{k_{cat,OFF,ij}}{\rightarrow} \quad \mathrm{RNAP} + \mathrm{T}_{ij} + \mathrm{rA}_1 \text{ or } \mathrm{rI}_2 \\
\mathrm{RNaseH} + \mathrm{rA}_1\mathrm{dI}_1 \text{ or } \mathrm{A}_2\mathrm{rI}_2 \quad &\underset{k_{-,H,j}}{\overset{k_{+,H}}{\rightleftharpoons}} \quad \mathrm{RNaseH}\cdot\mathrm{A}_j\mathrm{I}_j \quad \overset{k_{cat,H,j}}{\rightarrow} \quad \mathrm{RNaseH} + \mathrm{dI}_1 \text{ or } \mathrm{A}_2
\end{aligned}$$

We do not consider side-reactions or incomplete production and degradation products. The Michalis–Menten enzyme reactions are further simplified by the steady-state assumption for the enzyme-substrate complex, which is approximately valid when enzyme concentrations are low compared to substrate concentrations. We express the available enzyme concentrations as follows:

$$[\mathrm{RNAP}] = \frac{[\mathrm{RNAP}^{tot}]}{1 + \sum \frac{[\mathrm{T}\cdot\mathrm{A}]}{K_{M,ON}} + \sum \frac{[\mathrm{T}]}{K_{M,OFF}}}, \qquad [\mathrm{RNaseH}] = \frac{[\mathrm{RNaseH}^{tot}]}{1 + \frac{[\mathrm{rA}_1\cdot\mathrm{dI}_1]}{K_{M,H}} + \frac{[\mathrm{A}_2\cdot\mathrm{rI}_2]}{K_{M,H}}},$$

where the Michaelis constants are calculated as $K_M = \frac{k_- + k_{cat}}{k_+}$ to determine the affinity of substrates to the enzymes. From mass balance, $[\mathrm{T}_{ij}^{tot}]$, $[\mathrm{A}_j^{tot}]$, and $[\mathrm{dI}_1^{tot}]$ are preserved such that $[\mathrm{T}_{ij}\cdot\mathrm{A}_j]$, $[\mathrm{A}_1\cdot\mathrm{dI}_1]$, $[\mathrm{rA}_1\cdot\mathrm{dI}_1]$, and $[\mathrm{A}_2\cdot\mathrm{rI}_2]$ can be calculated from other variables.

For RNAP, the enzyme concentration is comparable to the switch concentrations and the above approximation may not be valid. To obtain a better approximation of available RNAP concentrations, we use the following method. Consider the [TA] used above as the sum of enzyme-bound species [RNAP· TA] and isolated species $[\mathrm{TA}]^F$. Then, at steady-state,

$$\begin{aligned}
[\mathrm{RNAP}\cdot\mathrm{TA}] &= \frac{[\mathrm{RNAP}][\mathrm{TA}]^F}{K_{M,ON}} = \frac{[\mathrm{RNAP}]([\mathrm{TA}] - [\mathrm{RNAP}\cdot\mathrm{TA}])}{K_{M,ON}}, \\
[\mathrm{RNAP}\cdot\mathrm{TA}] &= \frac{[\mathrm{RNAP}][\mathrm{TA}]}{K_{M,ON} + [\mathrm{RNAP}]}.
\end{aligned}$$

Similarly,

$$[\mathrm{RNAP}\cdot\mathrm{T}] = \frac{[\mathrm{RNAP}][\mathrm{T}]}{K_{M,OFF} + [\mathrm{RNAP}]}.$$

So that

$$\begin{aligned}
\mathrm{RNAP}^{tot} &= [\mathrm{RNAP}] + \sum([\mathrm{RNAP}\cdot\mathrm{TA}] + [\mathrm{RNAP}\cdot\mathrm{T}]) \\
&= [\mathrm{RNAP}](1 + \frac{[\mathrm{T_{21}A_1}]}{K_{M,ON,1} + [\mathrm{RNAP}]} + \frac{[\mathrm{T_{12}A_2}]}{K_{M,ON,2} + [\mathrm{RNAP}]} \\
&\quad + \frac{[\mathrm{T_{21}}]}{K_{M,OFF,1} + [\mathrm{RNAP}]} + \frac{[\mathrm{T_{12}}]}{K_{M,OFF,2} + [\mathrm{RNAP}]}).
\end{aligned}$$

The free enzyme concentration was not solved analytically but was estimated numerically by Newton's method (c.f. Martin and Coleman (1987) where [RNAP· TA] was solved analytically).

Thus, the dynamics of the system is described by the following seven ordinary differential equations:

$$\begin{aligned}
\frac{d[\mathrm{T_{21}}]}{dt} &= -k_{TA1}[\mathrm{T_{21}}][\mathrm{A_1}] + k_{TAI1}[\mathrm{T_{21}A_1}][\mathrm{dI_1}], \\
\frac{d[\mathrm{A_1}]}{dt} &= -k_{AI1}[\mathrm{A_1}][\mathrm{dI_1}] - k_{TA1}[\mathrm{T_{21}}][\mathrm{A_1}] + k_{AIrA1}[\mathrm{A_1dI_1}][\mathrm{rA_1}], \\
\frac{d[\mathrm{dI_1}]}{dt} &= -k_{AI1}[\mathrm{A_1}][\mathrm{dI_1}] - k_{rAI1}[\mathrm{rA_1}][\mathrm{dI_1}] - k_{TAI1}[\mathrm{T_{21}A_1}][\mathrm{dI_1}] + \frac{k_{cat,H,1}}{K_{M,H,1}}[\mathrm{RNaseH}][\mathrm{rA_1dI_1}], \\
\frac{d[\mathrm{rA_1}]}{dt} &= -k_{rAI1}[\mathrm{rA_1}][\mathrm{dI_1}] - k_{AIrA1}[\mathrm{A_1dI_1}][\mathrm{rA_1}] + \frac{k_{cat,ON,12}}{K_{M,ON,12}}[\mathrm{RNAP}][\mathrm{T_{12}A_2}] \\
&\quad + \frac{k_{cat,OFF,12}}{K_{M,OFF,12}}[\mathrm{RNAP}][\mathrm{T_{12}}], \\
\frac{d[\mathrm{T_{12}}]}{dt} &= -k_{TA2}[\mathrm{T_{12}}][\mathrm{A_2}] + k_{TAI2}[\mathrm{T_{12}A_2}][\mathrm{rI_2}], \\
\frac{d[\mathrm{A_2}]}{dt} &= -k_{AI2}[\mathrm{A_2}][\mathrm{rI_2}] - k_{TA2}[\mathrm{T_{12}}][\mathrm{A_2}] + \frac{k_{cat,H,2}}{K_{M,H,2}}[\mathrm{RNaseH}][\mathrm{A_2rI_2}], \\
\frac{d[\mathrm{rI_2}]}{dt} &= -k_{AI2}[\mathrm{A_2}][\mathrm{rI_2}] - k_{TAI2}[\mathrm{T_{12}A_2}][\mathrm{rI_2}] + \frac{k_{cat,ON,21}}{K_{M,ON,21}}[\mathrm{RNAP}][\mathrm{T_{21}A_1}] \\
&\quad + \frac{k_{cat,OFF,21}}{K_{M,OFF,21}}[\mathrm{RNAP}][\mathrm{T_{21}}].
\end{aligned}$$

The remaining variables, $[\mathrm{T_{21}A_1}]$, $[\mathrm{A_1dI_1}]$, $[\mathrm{rA_1dI_1}]$, $[\mathrm{T_{12}A_2}]$ and $[\mathrm{A_2rI_2}]$, are calculated from mass conservation.

Two-node oscillator with positive feedback

Here, we introduce a new template $\mathrm{T_{11}}$ to the two-node oscillator which performs positive autoregulation on the amount of $\mathrm{rA_1}$. The system can now be described by eight ordinary differential equations. One more ordinary differential equation for $[\mathrm{T_{11}}]$ is introduced and the differential equa-

tions of $[A_1]$, $[dI_1]$, and $[rA_1]$ are updated as follows:

$$\begin{aligned}
\frac{d[T_{11}]}{dt} &= -k_{TA11}[T_{11}][A_1] + k_{TAI11}[T_{11}A_1][dI_1], \\
\frac{d[A_1]}{dt} &= -k_{AI1}[A_1][dI_1] - k_{TA1}[T_{21}][A_1] - k_{TA11}[T_{11}][A_1] + k_{AIrA1}[A_1dI_1][rA_1], \\
\frac{d[dI_1]}{dt} &= -k_{AI1}[A_1][dI_1] - k_{rAI1}[rA_1][dI_1] - k_{TAI1}[T_{21}A_1][dI_1] - k_{TAI11}[T_{11}A_1][dI_1] \\
&\quad + \frac{k_{cat,H,1}}{K_{M,H,1}}[\mathrm{RNaseH}][rA_1dI_1], \\
\frac{d[rA_1]}{dt} &= -k_{rAI1}[rA_1][dI_1] - k_{AIrA1}[A_1dI_1][rA_1] + \frac{k_{cat,ON,12}}{K_{M,ON,12}}[\mathrm{RNAP}][T_{12}A_2] \\
&\quad + \frac{k_{cat,OFF,12}}{K_{M,OFF,12}}[\mathrm{RNAP}][T_{12}] + \frac{k_{cat,ON,11}}{K_{M,ON,11}}[\mathrm{RNAP}][T_{11}A_1] + \frac{k_{cat,OFF,11}}{K_{M,OFF,11}}[\mathrm{RNAP}][T_{11}].
\end{aligned}$$

The Michaelis constant of switch Sw_{11} should be similar to that of switch Sw_{21} due to identical promoter structures and the catalytic constant of switch Sw_{11} should be similar to that of switch Sw_{12} due to idential output sequences. However, the hybridization kinetics of switch Sw_{11} was not assumed to be identical to that of Sw_{21} because the fluorescence label on T_{21} strand can potentially stabilize the $T_{21}A_1$ complex compared to tha $T_{11}A_1$ complex (Marras et al. 2002). The differential equations of $[T_{21}]$, $[T_{12}]$, $[A_2]$ and $[rI_2]$ are unaffected. The remaining variables, $[T_{21}A_1]$, $[A_1dI_1]$, $[rA_1dI_1]$, $[T_{11}A_1]$, $[T_{12}A_2]$ and $[A_2rI_2]$, are calculated from mass conservation.

Ring oscillator

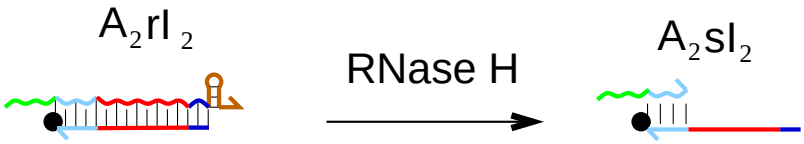
For the three-node oscillator, individual switch reactions are described by three differential equations such that the system can be described by nine ordinary differential equations. Take the switch Sw_{12} which was used as a component of the two-node oscillator. The only difference here is that the RNA signal rI_2 is produced by switch Sw_{23} rather than switch Sw_{21}.

$$\begin{aligned}
\frac{d[T_{12}]}{dt} &= -k_{TA2}[T_{12}][A_2] + k_{TAI2}[T_{12}A_2][rI_2] \\
\frac{d[A_2]}{dt} &= -k_{AI2}[A_2][rI_2] - k_{TA2}[T_{12}][A_2] + \frac{k_{cat,H,2}}{K_{M,H,2}}[\mathrm{RNaseH}][A_2rI_2] \\
\frac{d[rI_2]}{dt} &= -k_{AI2}[A_2][rI_2] - k_{TAI2}[T_{12}A_2][rI_2] + \frac{k_{cat,ON,23}}{K_{M,ON,23}}[\mathrm{RNAP}][T_{23}A_3] \\
&\quad + \frac{k_{cat,OFF,23}}{K_{M,OFF,23}}[\mathrm{RNAP}][T_{23}]
\end{aligned}$$

The equations for the switch Sw_{31} and switch Sw_{23} are derived similarly.

Consideration of incomplete degradation products

To explain the experimental observation that the amount of RNA inhibitor in the two-node oscillator increased over time, we developed an extended model where we also consider incomplete degradation products. We propose that the consideration of incomplete degradation products can potentially explain both the slowdown of oscillation frequency and the accumulation of RNA signals. We define an incomplete degradation product sI_2 as the 5' partial sequence of rI_2 which is produced upon degradation of A_2rI_2. Due to the binding requirement of RNase H, five to seven bases of the toehold binding sequence within RNA inhibitor rI_2 cannot be degraded and remain as part of the incomplete degradation product sI_2 (Lima and Crooke 1997). Consequently, sI_2 can reversibly bind to activator A_2 or an ON switch $T_{12}A_2$ and we expect that a $T_{12}A_2sI_2$ complex will not be efficiently inhibited by rI_2 because the toehold binding sequence is not accessible.



Specifically, the following hybridization reactions are included in the model:

$$\begin{aligned}
\mathrm{A_2 + sI_2} &\underset{k_{dis}}{\overset{k_{AI2}}{\rightleftharpoons}} \mathrm{A_2sI_2}, \\
\mathrm{T_{12} + A_2sI_2} &\overset{k_{TA2}}{\rightarrow} \mathrm{T_{12}A_2sI_2}, \\
\mathrm{T_{12}A_2 + sI_2} &\underset{k_{dis}}{\overset{k_{AI2}}{\rightleftharpoons}} \mathrm{T_{12}A_2sI_2}, \\
\mathrm{rI_2 + A_2sI_2} &\overset{k_{AI2}}{\rightarrow} \mathrm{A_2rI_2 + sI_2}.
\end{aligned}$$

We assumed that while the incomplete degradation product remains bound to an ON switch $T_{12}A_2$, the production of output rA_1 is as fast as an ON switch by itself, yet inhibition by rI_2 does not take place. This change requires expanding the equations for switch Sw_{12} to six dimensions as follows:

$$
\begin{aligned}
\frac{d[\mathrm{T}_{12}]}{dt} &= -k_{TA2}[\mathrm{T}_{12}][\mathrm{A}_2] + k_{TAI2}[\mathrm{T}_{12}\mathrm{A}_2][\mathrm{rI}_2] - k_{TA2}[\mathrm{T}_{12}][\mathrm{A}_2\mathrm{sI}_2], \\
\frac{d[\mathrm{A}_2]}{dt} &= -k_{AI2}[\mathrm{A}_2][\mathrm{rI}_2] - k_{TA2}[\mathrm{T}_{12}][\mathrm{A}_2] - k_{AI2}[\mathrm{A}_2][\mathrm{sI}_2] + k_{dis}[\mathrm{A}_2\mathrm{sI}_2], \\
\frac{d[\mathrm{rI}_2]}{dt} &= -k_{AI2}[\mathrm{A}_2][\mathrm{rI}_2] - k_{AI2}[\mathrm{A}_2\mathrm{sI}_2][\mathrm{rI}_2] - k_{TAI2}[\mathrm{T}_{12}\mathrm{A}_2][\mathrm{rI}_2] \\
&\quad + \frac{k_{cat,ON,21}}{K_{M,ON,21}}[\mathrm{RNAP}][\mathrm{T}_{21}\mathrm{A}_1] + \frac{k_{cat,OFF,21}}{K_{M,OFF,21}}[\mathrm{RNAP}][\mathrm{T}_{21}], \\
\frac{d[\mathrm{sI}_2]}{dt} &= -k_{AI2}[\mathrm{A}_2][\mathrm{sI}_2] + k_{dis}[\mathrm{A}_2\mathrm{sI}_2] + k_{AI2}[\mathrm{A}_2\mathrm{sI}_2][\mathrm{rI}_2] - k_{AI2}[\mathrm{T}_{12}\mathrm{A}_2][\mathrm{sI}_2] \\
&\quad + k_{dis}[\mathrm{T}_{12}\mathrm{A}_2\mathrm{sI}_2], \\
\frac{d[\mathrm{A}_2\mathrm{sI}_2]}{dt} &= k_{AI2}[\mathrm{A}_2][\mathrm{sI}_2] - k_{dis}[\mathrm{A}_2\mathrm{sI}_2] - k_{AI2}[\mathrm{A}_2\mathrm{sI}_2][\mathrm{rI}_2] - k_{TA2}[\mathrm{T}_{12}][\mathrm{A}_2\mathrm{sI}_2] \\
&\quad + \frac{k_{cat,H,2}}{K_{M,H,2}}[\mathrm{RNaseH}][\mathrm{A}_2\mathrm{rI}_2], \\
\frac{d[\mathrm{T}_{12}\mathrm{A}_2\mathrm{sI}_2]}{dt} &= k_{AI2}[\mathrm{T}_{12}\mathrm{A}_2][\mathrm{sI}_2] - k_{dis}[\mathrm{T}_{12}\mathrm{A}_2\mathrm{sI}_2] + k_{TA2}[\mathrm{T}_{12}][\mathrm{A}_2\mathrm{sI}_2].
\end{aligned}
$$

The remaining variables for switch Sw_{12}, $[\mathrm{T}_{12}\mathrm{A}_2]$ and $[\mathrm{A}_2\mathrm{rI}_2]$ are calculated from mass balance.

Interestingly, the toehold of rA_1 sequence lies close to its 3' end where degradation of toehold-bound RNA sequence can be complete. It is possible that the lack of interfering signal (short rA_1 that hides the toehold sequence) may keep the excitatory connection effective with small amount of RNA activator rA_1 throughout the reaction, although further experimental verification is necessary. A similar transformation was implemented for the ring oscillator where all three inhibitory connections are potentially affected by the build-up of incomplete degradation products.

5.4.2 Data and tables

Table 5.1. Reaction conditions for the two-node oscillator (dataset 1)

Reaction #	T_{21} (nM)	A_1 (nM)	dI_1 (nM)	T_{12} (nM)	A_2 (nM)	rA_1 (nM)	rI_2 (nM)	RNAP (μL)	RNase H (μL)
1	250	250	700	120	350	0	0	4	0.44
2	250	250	700	120	350	0	0	4	0.54
3	250	250	1000	120	500	0	0	4	0.44
4	250	250	1000	120	500	0	0	4	0.54
5	250	250	700	120	350	0	0	6	0.72
6	250	250	700	120	350	0	0	6	0.84
7	250	250	1000	120	500	0	0	6	0.72
8	250	250	1000	120	500	0	0	6	0.84
9	250	250	700	120	350	740	0	6	0.84
10	250	250	700	120	350	0	370	6	0.84
11	250	250	700	120	350	740	370	6	0.84
12	250	250	700	120	350	500	370	6	0.84
13	250	250	1000	120	500	800	550	6	0.84

Table 5.2. Reaction conditions for the two-node oscillator with different enzyme batches (dataset 2)

Reaction #	T_{21} (nM)	A_1 (nM)	dI_1 (nM)	T_{12} (nM)	A_2 (nM)	RNAP (μL)	RNase H (μL)
1	150	150	500	70	250	2	0.2
2	150	150	500	70	250	3	0.3
3	200	200	500	100	250	2	0.2
4	200	200	500	100	250	3	0.3
5	250	250	500	120	250	3	0.3
6	250	250	500	120	250	4	0.4
7	250	250	700	120	350	3	0.3
8	250	250	700	120	350	4	0.4
9	250	250	700	120	350	4	0.5
10	250	250	700	120	450	4	0.5
11	250	250	700	120	350	6	0.75
12	250	250	700	120	450	6	0.75
13	250	250	1000	80	500	6	0.72
14	250	250	1500	80	750	6	0.72
15	250	250	1000	100	500	6	0.72
16	250	250	1500	100	750	6	0.72
17	250	250	1000	120	350	6	0.72
18	250	250	1300	120	350	6	0.72
19	250	250	1000	180	350	6	0.72
20	250	250	1200	180	450	6	0.72
21	150	150	1000	180	350	6	0.72
22	150	150	1000	180	350	6	0.9
23	200	200	1000	180	350	6	0.72
24	200	200	1000	180	350	6	0.9

Table 5.3. Reaction conditions for the two-node oscillator with a positive feedback (dataset 3)

Reaction #	T_{21} (nM)	A_1 (nM)	dI_1 (nM)	T_{12} (nM)	A_2 (nM)	T_{11} (nM)	RNAP (μL)	RNase H (μL)
1	250	250	1000	100	500	0	6	0.84
2	250	250	1000	80	500	20	6	0.84
3	250	250	1000	80	500	40	6	0.84
4	250	250	1000	100	500	20	6	0.84
5	250	250	1000	60	500	0	6	0.84
6	250	250	1000	60	500	30	6	0.84
7	250	250	1000	60	500	60	6	0.84
8	250	250	1000	60	500	90	6	0.84

Table 5.4. Model parameters for two-node oscillators

	i=2, j=1	i=1, j=2	i=1, j=1	Other studies
$K_{M,ON,ij}$ (nM)	60	57	46	15-37
$k_{cat,ON,ij}$ (/s)	0.072	0.033	0.02	0.73-1.12
$K_{M,OFF,ij}$ (μM)	1.96	1.96	1.41	0.1-1.1
$k_{cat,OFF,ij}$ (/s)	0.011	0.0074	0.0096	0.11-0.18
$K_{M,H,j}$ (nM)	160	19	-	16-130
$k_{cat,H,j}$ (/s)	0.028	0.111	-	0.02-0.6
$k_{TA,ij}$ (/M/s)	1.30*10^4	7.28*10^3	9.18*10^3	-
$k_{AI,j}$ (/M/s)	1.61*10^4	9.09*10^3	-	-
$k_{TAI,ij}$ (/M/s)	9.06*10^3	1.61*10^4	3.76*10^4	-
$k_{rAI,j}$ (/M/s)	1.32*10^4	-	-	-
$k_{AIrA,j}$ (/M/s)	2.11*10^4	-	-	-
Rv	1.54	-	-	-
Rhv	1.99	-	-	-
k_J (/s)	0.08	-	-	-

Table 5.5. Reaction conditions for the ring oscillator (dataset 4)

Reaction #	T_{31} (nM)	A_1 (nM)	T_{12} (nM)	A_2 (nM)	T_{23} (nM)	A_3 (nM)	rI_2 (nM)	RNAP (μL)	RNase H (μL)
1	80	300	100	200	120	400	500	4	0.5
2	80	300	100	250	120	500	500	4	0.5
3	100	300	100	160	120	400	400	4	0.5
4	100	300	100	160	120	400	400	4	0.6
5	100	300	100	200	120	400	400	4	0.5
6	100	300	100	200	120	400	400	4	0.6
7	100	300	100	200	120	400	0	4	0.5
8	100	300	100	200	120	400	200	4	0.5
9	100	300	100	200	120	400	0	4	0.5
10	100	300	100	200	120	400	0	5	0.5
11	100	300	100	200	120	400	0	4	0.5

Table 5.6. Model parameters for the ring-oscillator

	i=3, j=1	i=1, j=2	i=2, j=3	Other studies
$K_{M,ON,ij}$ (nM)	43	160	92	15-37
$k_{cat,ON,ij}$ (/s)	0.088	0.068	0.045	0.73-1.12
$K_{M,OFF,ij}$ (μM)	1.41	0.52	0.90	0.1-1.1
$k_{cat,OFF,ij}$ (/s)	0.001	0.022	0.008	0.11-0.18
$K_{M,H,j}$ (nM)	152	79	41	16-130
$k_{cat,H,j}$ (/s)	0.087	0.148	0.157	0.02-0.6
$k_{TA,j}$ (/M/s)	5.45*10^3	1.31*10^4	1.32*10^4	-
$k_{AI,j}$ (/M/s)	5.13*10^3	1.26*10^4	8.90*10^3	-
$k_{TAI,j}$ (/M/s)	9.05*10^3	2.24*10^4	1.51*10^4	-
k_J (/s)	0.32	-	-	-

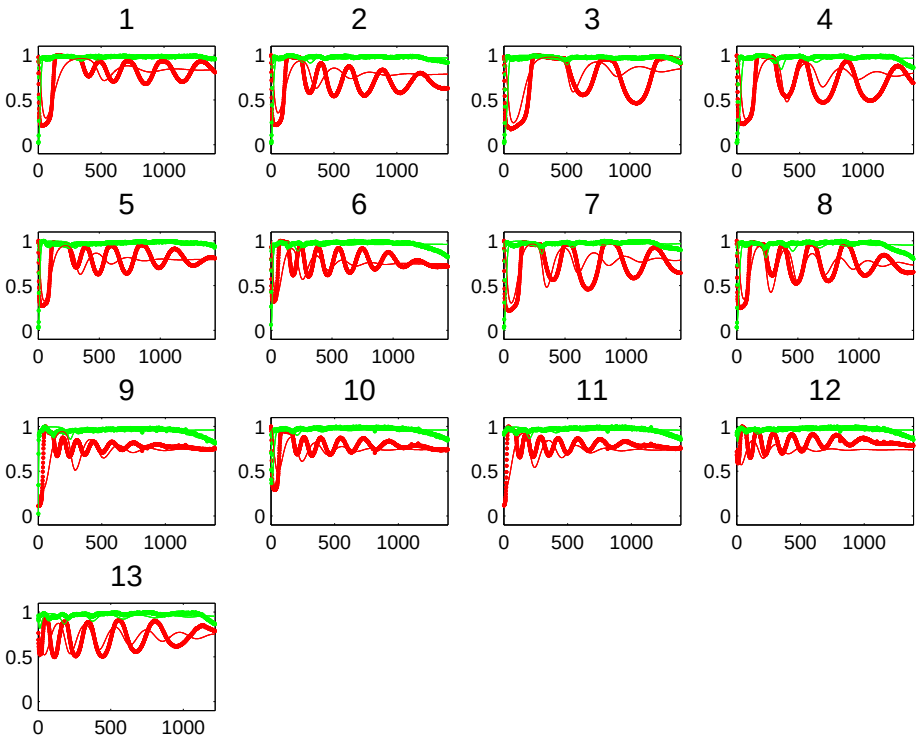


Figure 5.6. Experimental result and model fittings of the two-node oscillator. X-axes are time in minutes, y-axes are normalized fluorescence signal. The experimental time-courses are plotted as dots (which appear as thick lines) and simulation time-courses are plotted as lines: Texas Red (red) and TAMRA (green). The experimental conditions are listed in table 5.1.

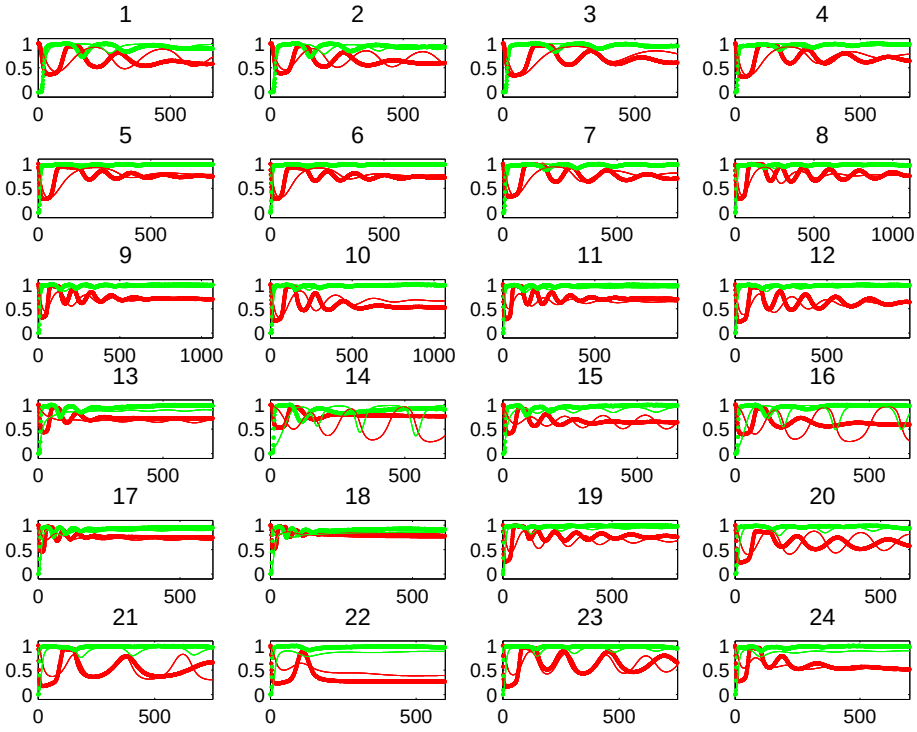


Figure 5.7. Experimental result and model fittings of the two-node oscillator with different enzyme batch. The experimental time-courses are plotted as dots (which appear as thick lines) and simulation time-courses are plotted as lines: Texas Red (red) and TAMRA (green). We use the same parameter set as in figure 5.6. We use two additional parameters to account for enzyme batch difference, Rv and Rhv, assuming that the difference of enzyme batches are only in their concentrations. For this data set, RNAP and RNase H concentrations are multiplied by Rv and Rhv, respectively, compared to figure 5.6. The performance is still reasonable except for a couple of cases considering that this data set was never used to fit the parameters. Experimental conditions are listed in table 5.2. No external RNA is added initially.

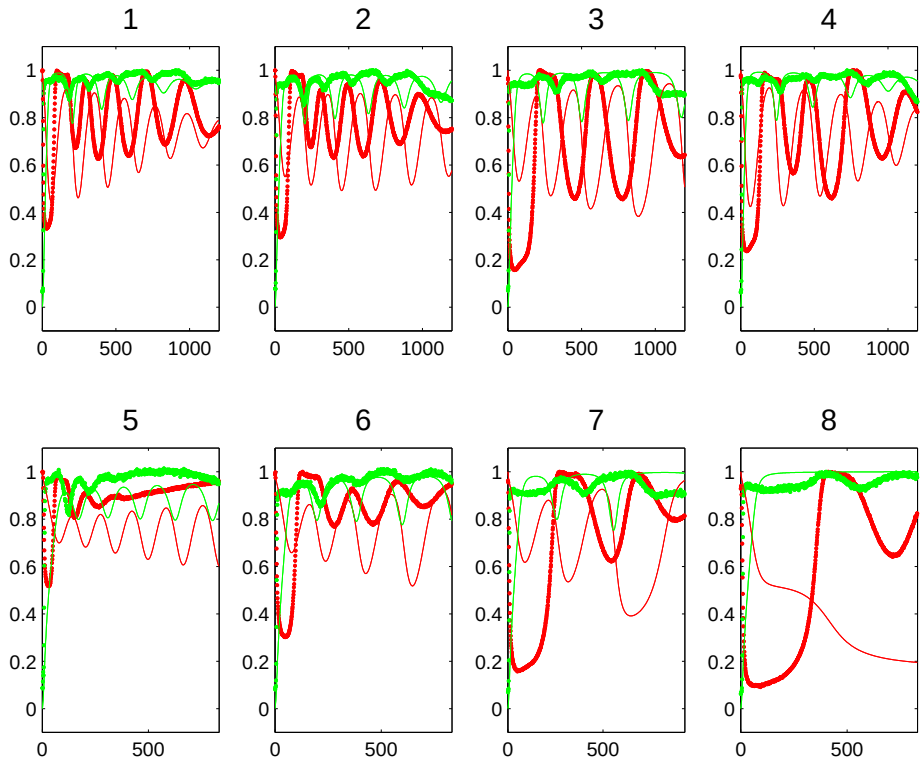


Figure 5.8. Experimental result and model fittings of the two-node oscillator with positive feedback. The experimental time-courses are plotted as dots (which appear as thick lines) and simulation time-courses are plotted as lines: Texas Red (red) and TAMRA (green). The experimental conditions are listed in table 5.3. No external RNA is added initially.

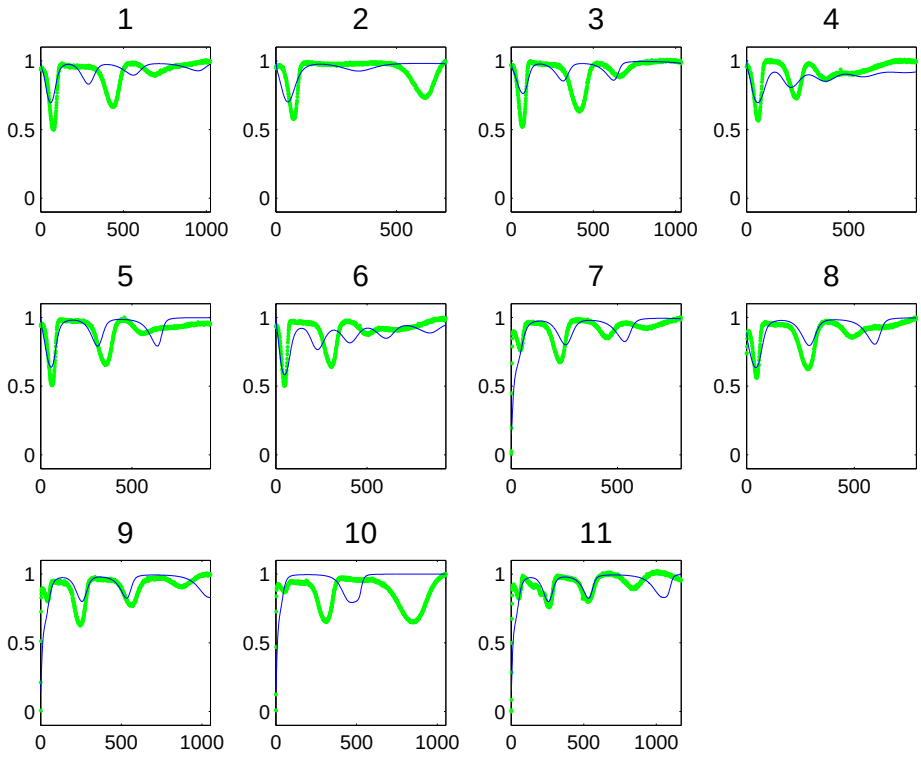


Figure 5.9. Experimental result and model fittings of the ring oscillator. The experimental time-courses are plotted as dots (green) and simulation time-courses are plotted as lines (blue). TAMRA signal from switch Sw_{12} is monitored. The experimental conditions are listed in table 5.5.

Chapter 6 Discussion and Future Directions

The simplicity of our *in vitro* transcriptional regulatory circuit makes it a useful model system to study complex biochemical circuit designs. The modular switch design allows easy rewiring of connections to achieve alternative circuit designs with both excitatory and inhibitory connections with sharp and adjustable thresholds. However, the current switch design has but a single regulatory domain such that combinatorial regulation would require multiple switches. On the other hand, there is no limitation to how many different outputs can be encoded by a single regulatory switch. To test composability of synthetic switches, we have constructed various elementary circuits: feedforward circuits, autoregulatory switches, bistable circuits, and oscillators. These exercises demonstrate that basic composition works in our transcriptional circuits. However, the composition is nonlinear: global feedback through shared use of enzyme affects the entire system dynamics when a new component is added. Especially for dynamical systems such as oscillators, proper kinetic modeling becomes challenging due to intricate and time-varying nature of kinetic parameters.

Designability is one of the key issues since well-defined secondary structure is essential for functional DNA switches and RNA signals. The strength of interaction among different species can be evaluated with standard models, thus algorithmic optimization is feasible. However, the sequence dependence of production rates and degradation rates is rarely documented, which obstructs *a priori* selection of "good" set of sequences. The asymmetry among different components interferes with the composition of high-gain regions required for signal restoration and multilayered architecture. A closely related issue of crosstalk may be worse than genetic regulatory networks with reversible protein binding due to the requirement of irreversible binding to correct partners. Low concentrations of molecular species alleviate the crosstalk problem at the cost of slowing down the overall network operation. This opens up an interesting question of speed and error rate trade-offs in the biochemical network, and why stoichiometric irreversible binding reactions are not often observed *in vivo* (Ferrell 1996). An irreversible binding mechanism may require the destruction of binding partners to recover an important reusable component, which is potentially wasteful especially with multiple binding partners. Even though the irreversible binding mechanism can lead to a sharp transition with a linear input change compared to a logarithmic input change as in the cooperative binding mechanism, the narrow transition regions can be problematic to match when multiple

cellular gates are chained.

Another great challenge is understanding enzyme reactions in detail. Some aspects of the enzyme reactions are quite complex. T7 RNA polymerase, although simple and efficient, has some side reactions and abortive transcription activity (Biebricher and Luce 1996; Cazenave and Uhlenbeck 1994; Diaz et al. 1996). This leads to the accumulation of potentially problematic short RNA products and double-stranded RNA species. RNase H cannot degrade the 5' end of RNA strands on the RNA-DNA hybrid (Lima and Crooke 1997). This lack of complete degradation leads to the accumulation of toehold binding short RNA products, slowing down the circuit operation. RNase R, albeit useful for clean-up purposes, has not been fully characterized (Cheng and Deutscher 2002). The difference of RNase R affinity to different RNA strands could disrupt bistability in our mutually inhibitory circuit. Understanding how enzymes differ from each other and how they spend time on the DNA and RNA species involved are crucial for a large-scale implementation.

One of the experimental difficulty we encountered is that the circuit operation is sensitive to enzyme concentrations. The batch-to-batch variation in enzyme stock solutions required calibration of enzyme activities to achieve desired circuit operation. As shown in our formal model in chapter 2, the gain of switch depends on the enzyme reaction rate constants. Negative autoregulation is commonly used in nature to maintain stable level of metabolic flux and has also been demonstrated by synthetic approaches (Becskei and Serrano 2000). We explored negative autoregulation in our *in vitro* network. The regulatory domain of switch was redesigned to accept its own transcript as an inhibitory input. Preliminary characterization demonstrates that the negative autoregulation indeed confers robustness to parameter variation such as the template concentration. The regulated switch produces the desired amount of output, the amount of activator, irrespective of the template concentrations (figure 6.1A). Fluorometer experiments show that the response time of negative autoregulation is reasonably fast, on the order of 10 minutes (figure 6.1B). Although negative feedback with delay can lead to oscillation, we did not observe oscillatory behavior in the negative autoregulatory switch possibly due to the fast hybridization of inhibitory input. The stable output achieved with a negative autoregulation can be an important mechanism to maintain network performance when the parameters fluctuate. For instance, we can engineer circuits with negative autoregulation check points for a robust control of the output level. The robustness to parameter variation may lead to temperature compensation in our oscillator design, a hallmark of circadian clocks.

The simplicity of our *in vitro* circuit makes it an appealing model system for broader research topics. For instance, we can explore stochastic noise in biochemical circuits. Stochasticity in gene

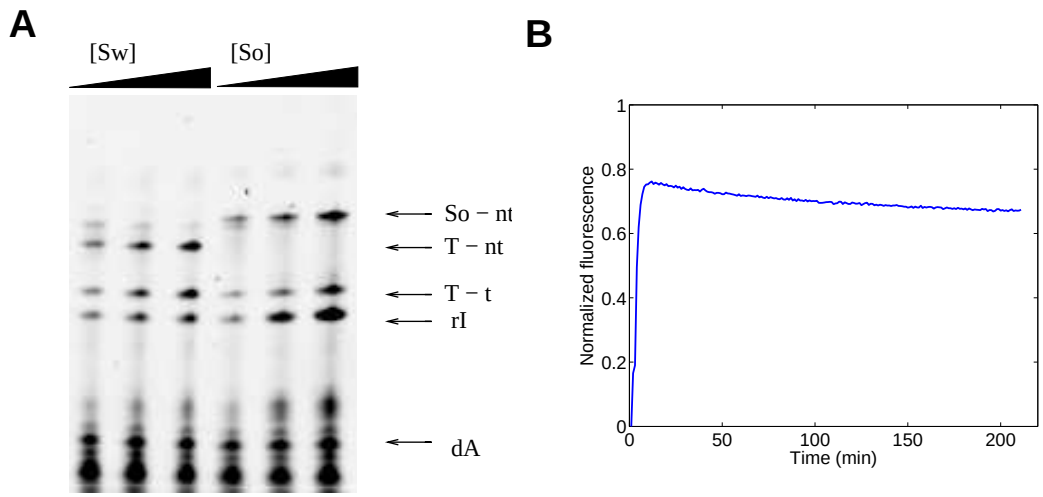


Figure 6.1. Negative autoregulatory switch **(A)** RNA outputs from negative autoregulatory switch Sw and unregulated source So measured in denaturing gel. The RNA output from unregulated source So show almost linear response to source template concentration change. However, the RNA output from negative autoregulatory switch Sw does not increase further once it reaches the activator level irrespective of switch template concentration. **(B)** Negative autoregulatory switch response measured in fluorometer. The switch state quickly reaches a stable steady-state where it produces enough RNA inhibitory signal to consume most of activator.

expression arises from fluctuations in transcription and translation, despite constant environmental conditions (Kaern et al. 2005). Both stochasticity inherent in the biochemical process of gene expression (intrinsic noise) and fluctuation in other cellular components (extrinsic noise) contribute to overall variation (Elowitz et al. 2002). Intrinsic noise in gene expression has three contributions: finite-number effect, translational bursting, and transcriptional bursting (Kaern et al. 2005). Translational bursting mechanisms are experimentally verified in *B. subtilis* (Ozbudak et al. 2002) and *S. cerevisiae* (Blake et al. 2003). Transcriptional bursting mechanisms are shown to be important in *S. cerevisiae* possibly due to slow chromatin remodeling in gene activation (Raser and O'Shea 2004). Characterization of stochastic noise in our *in vitro* transcriptional system would explore the question of the noise floor in transcription reactions where only the finite-number effect contributes significantly. The potential advantages are as follows: First, the parameters are known and free of external fluctuations in the *in vitro* system, allowing accurate measurement of the finite-number effect, Second, the *in vitro* system is scalable such that the same biochemical reactions can be studied in volumes changing several orders of magnitude, Third, the *in vitro* system can be monitored with high temporal resolution. We explored our *in vitro* transcription network in small volumes to test whether it is a suitable model system for stochastic noise characterization. We used a microfluidic chamber of 100 fL reaction volume prepared by Dr. Seungyong Jung in Collier group at Caltech.

We verified that the transcription of RNA output and regulation of switch have similar timescales as in the bulk fluorometer measurement. We measured the fluctuation of fluorescence signal once the system reached a steady state. Because the fluorescence signal was low, the background noise was significant and we observed only a small variation in the switch fluorescence (figure 6.2). Further measurements in smaller volumes would be necessary to test whether the finite-number effect constributes significant noise in our *in vitro* circuit.

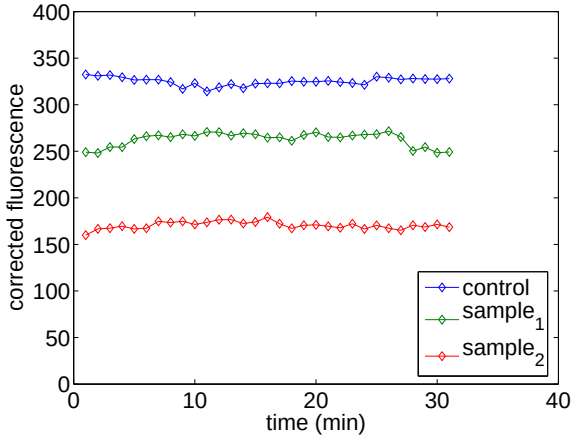


Figure 6.2. Experimental measurement of stochastic noise in a negative autoregulatory switch operating in a 100 fL volume. Two different samples were prepared, the coefficients of variation (standard deviation divided by mean) after subtracting background noise were 0.024 (sample_1) and 0.014 (sample_2).

Pattern formation and self-organization are one of the fascinating phenomena observed in physical, chemical, and biological systems. Moreover, pattern formation in very different systems may have common features (Muratov and Osipov 1996). Intriguing patterns have been achieved in a chemical reaction: replication of spot growth and labyrinthian patterns (Lee et al. 1994). Recently, an *in vitro* transcription-translation network that emulate *Drosophila* embryonic pattern was constructed (Isalan et al. 2005). To test whether our *in vitro* circuit can exhibit interesting pattern formation, we subject our bistable circuit in a confined reaction chamber shaped as a thin film (20 mm $\cdot$ 20 mm $\cdot$ 0.25 mm). We expect that both the diffusion and reaction timescales of RNA signals are much slower than the typical diffusion and reaction timescales of small molecules used in chemical pattern formation. Model studies indicate that our circuit should be able to maintain spatial patterns if the diffusion rates are not too fast compared to the reaction rates. The initially dispersed pattern in our bistable circuit morphed into a circular shape as sharp edges were trimmed and convex holes were filled (figure 6.3). Our observation agrees with a theoretical treatment of bistable

reaction-diffusion system (Meerson and Sasorov 1996), only an equilibrium planar interface and a perfect circular interface are stable with respect to deformations in 2-dimensional space. The next step would be characterizing spatial pattern formation by oscillator circuits.

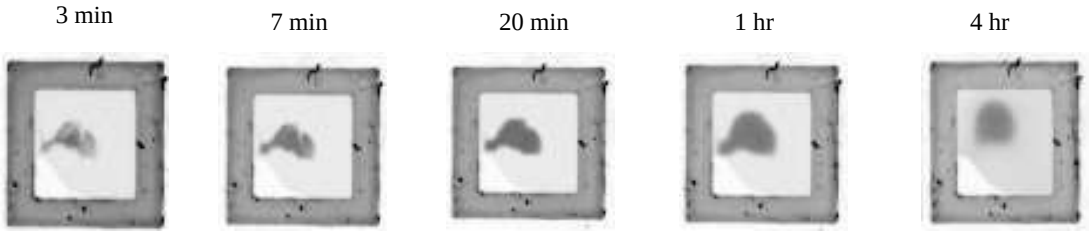


Figure 6.3. Pattern formation by a mutually inhibitory circuit where switch Sw_{12} and switch Sw_{21} inhibit each other. Initially, RNA inhibitor rI_1 was externally provided to generate spatial heterogeneity. Switch Sw_{21} state is monitored by fluorescence. Both high and low signals are stable states in homogeneous reaction mixtures, yet the diffusion of RNA inhibitors moves the boundary of different states. The domain boundary remains distinct throughout the reaction and develops into a circular shape. In the diffusion experiment without enzymes (data not shown), the boundary spread out about 4 mm in each direction after 5 hours.

Recently, a number of DNA-based nanomachines have been demonstrated *in vitro* (Yurke et al. 2000; Yan et al. 2002). These molecular machines make use of the recognition specificities of DNA strands to switch between different conformations and potentially perform mechanical work once embedded in a suitable platform. So far, the operation of these nanomachines have been demonstrated by repeated addition of different trigger DNA strands by the experimenter. Our *in vitro* oscillator can potentially control these DNA nanomachines autonomously by clocked "fuelling" of desired RNA strands. However, if the oscillator regulatory signals are used to drive DNA nanomachines as well as the oscillator itself, the oscillation frequency and amplitude would change due to an increased demand for regulatory signals. Also, the oscillator regulatory signals need to be redesigned for different DNA nanomachines. Thus, it is desirable to include transcriptional switches that are regulated by oscillator signals and produce output RNA strands for the target DNA nanomachines so that the core oscillator regulatory signals will not be diverted and arbitrary sequences for driving nanomachines can be encoded.

Our *in vitro* circuit can be useful for construction of an artificial cell. In one of the recent estimates of necessary genes for a minimal cell (Forster and Church 2006), a surprisingly large fraction (96%) of minimal gene set is devoted to translation mechanisms: ribosome components, a set of tRNA, a set of translational initiation, elongation, release factors, and a few chaperones. If a suitable set of ribozymes can be found for the production and degradation of nucleic acid species,

our *in vitro* circuit can be adapted to perform arbitrary regulatory function in the absence of protein machinery. This would greatly decrease the genome size required for a minimal cell, which in turn could provide a valuable framework for testing essential regulatory functions for a minimal cell. For instance, our oscillator realized a premitive cell-cycle for an artificial cell. How we can link the clocked operation of oscillator to the replication of DNA genome and other machinery and eventual duplication of cell volume requires further investigation. We envision that using periodic RNA outputs from our oscillator as primers for a reverse transcriptase would be enough for regulated duplication of DNA genome (Guatelli et al. 1990). For the minimal genome considered in Forster and Church (2006), catabolism (nucleases and proteases), active conversion or removal of waste products (energy regenerating enzymes and membrane transporters), and regulatory feedbacks were omitted. However, our experience with *in vitro* circuit suggests that waste product management and regulatory feedback would be essential for prolonged operation of an artificial cell. For example, the build-up of incomplete RNA degradation products seemed to interfere with our oscillator operation and resulted in an increased demand for regulatory signals over time. One approach to regulate short RNA products was including OligoRNase in the system. However, the activity of OligoRNase was very low due to high NTP concentration in our experiments. Thus, by using OligoRNase with low NTP concentration, we should be able to obtain better waste product management at the cost of slow operation. We hope that the reiterative nature of design process as we grapple with the trade-offs required to enable system function will teach us valuable lessons for how to build a functional artificial cell.

Bibliography

J. Ackermann, B. Wlotzka, and J. S. McCaskill. *In vitro* DNA-based predator-prey system with oscillatory kinetics. *Bull. Math. Biol.*, 60:329–353, 1998.

U. Alon. Biological networks: The tinkerer as an engineer. *Science*, 301:1866–1867, 2003.

M. R. Atkinson, M. A. Savageau, J. T. Myers, and A. J. Ninfa. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell*, 113:597–607, 2003.

F. K. Balagadde, L. You, C. L. Hansen, F. H. Arnold, and S. R. Quake. Long-term monitoring of bacteria undergoing programmed population control in a microchemostat. *Science*, 309:137–140, 2005.

N. Barkai and S. Leibler. Circadian clocks limited by noise. *Nature*, 403:267–268, 2000.

S. Basu, R. Mehreja, S. Thiberge, M.-T. Chen, and R. Weiss. Spatiotemporal control of gene expression with pulse-generating networks. *Proc. Natl. Acad. Sci. USA*, 101(17):6355–6350, 2004.

T. S. Bayer and C. D. Smolke. Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nature Biotech.*, 23(3):337–343, 2005.

A. Becskei, B. Seraphin, and L. Serrano. Positive feedback in eukaryotic gene networks: Cell differentiation by graded to binary response conversion. *EMBO J.*, 20(10):2528–2535, 2001.

A. Becskei and L. Serrano. Engineering stability in gene networks by autoregulation. *Nature*, 405: 590–593, 2000.

S. A. Benner, A. D. Ellington, and A. Tauer. Modern metabolism as a palimpsest of the RNA world. *Proc. Natl. Acad. Sci. USA*, 86:7054–7058, 1989.

S. A. Benner and A. M. Sismour. Synthetic biology. *Nature Rev. Genet.*, 6:533–543, 2005.

C. K. Biebricher and R. Luce. Template-free generation of RNA species that replicate with bacteriophage T7 RNA polymerase. *EMBO J.*, 15:3458–3465, 1996.

W. J. Blake, M. Kaern, C. R. Cantor, and J. J. Collins. Noise in eukaryotic gene expression. *Nature*, 422:633–637, 2003.

D. Bray. Molecular networks: The top-down view. *Science*, 301:1864–1865, 2003.

N. E. Buchler, U. Gerland, and T. Hwa. On schemes of combinatorial transcription logic. *Proc. Natl. Acad. Sci. USA*, 100:5136–5141, 2003.

J. C. Carrington and V. Ambros. Role of microRNAs in plant and animal development. *Science*, 301:336–338, 2003.

C. Cazenave and O. C. Uhlenbeck. RNA template-directed RNA synthesis by T7 RNA polymerase. *Proc. Natl. Acad. Sci. USA*, 91:6972–6976, 1994.

Z.-F. Cheng and M. P. Deutscher. Purification and characterization of the *Escherichia coli* exoribonuclease RNase R. *J. Biol. Chem.*, 277:21,624–21,629, 2002.

A. K. Datta and S. K. Niyogi. A novel oligoribonuclease of *Escherichia coli*. *J. Biol. Chem.*, 250: 7313–7319, 1975.

G. A. Diaz, M. Rong, W. T. McAllister, and R. K. Durbin. The stability of abortive cycling T7 RNA polymerase complexes depends upon template conformation. *Biochemistry*, 35:10,837–10,843, 1996.

R. Dirks, M. Lin, E. Winfree, and N. A. Pierce. Paradigms for computational nucleic acid design. *Nucleic Acids Research*, 32:1392, 2004.

W. U. Dittmer and F. C. Simmel. Transcriptional control of DNA-based nanomachines. *Nano Lett.*, 4(4):689–691, 2004.

J. C. Dunlap. Running a clock requires quality time together. *Science*, 311:184–186, 2006.

M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403:335–338, 2000.

M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain. Stochastic gene expression in a single cell. *Science*, 297:1183–1186, 2002.

J. E. Ferrell, Jr. Tripping the switch fantastic: How a protein kinase cascade can convert graded inputs into switch-like outputs. *Trends Biochem. Sci.*, 21:460–466, 1996.

T. J. Fiedler, H. A. Vincent, Y. Zuo, O. Gavrialov, and A. Malhotra. Purification and crystallization of *Escherichia coli* oligoribonuclease. *Acta Cryst.*, D60:736–739, 2004.

C. Flamm, W. Fontana, I. Hofacker, and P. Schuster. RNA folding at elementary step resolution. *RNA*, 6:325–338, 2000.

A. C. Forster and G. M. Church. Towards synthesis of a minimal cell. *Mol. Syst. Biol.*, 2:45, 2006.

T. S. Gardner, C. R. Cantor, and J. J. Collins. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 403:339–342, 2000.

L. Glass and S. A. Kauffman. The logical analysis of continuous, non-linear biochemical control networks. *J. Theor. Biol.*, 39:103–129, 1973.

J. S. Griffith. Mathematics of cellular control processes. I. Negative feedback to one gene. *J. Theor. Biol.*, 20:202–208, 1968.

M. Grunberg-Manago. Messenger RNA stability and its role in control of gene expression in bacteria and phages. *Annu. Rev. Genet.*, 33:193–227, 1999.

J. C. Guatelli, K. M. Whitefield, D. Y. Kwoh, K. J. Barringer, D. D. Richman, and T. R. Gingeras. Isothermal, *in vitro* amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. *Proc. Natl. Acad. Sci. USA*, 87:1874–1878, 1990.

C. C. Guet, M. B. Elowitz, W. Hsing, and S. Leibler. Combinatorial synthesis of genetic networks. *Science*, 296:1466–1470, 2002.

G. G. Hammes. *Thermodynamics and Kinetics for the Biological Sciences*. Wiley, 2000.

L. H. Hartwell, J. J. Hopfield, S. Leibler, and A. W. Murray. From molecular to modular cell biology. *Nature*, 402:C47–C52, 1999.

S. Hooshangi, S. Thiberge, and R. Weiss. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc. Natl. Acad. Sci. USA*, 102(10):3581–3586, 2005.

J. J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. USA*, 81(10):3088–3092, 1984.

F. J. Isaacs, D. J. Dwyer, and J. J. Collins. RNA synthetic biology. *Nature Biotech.*, 24:545–554, 2006.

F. J. Isaacs, D. J. Dwyer, C. Ding, D. D. Pervouchine, C. R. Cantor, and J. J. Collins. Engineered riboregulators enable post-transcriptional control of gene expression. *Nature Biotech.*, 22(7): 841–847, 2004.

F. J. Isaacs, J. Hasty, C. R. Cantor, and J. J. Collins. Prediction and measurement of an autoregulatory genetic module. *Proc. Natl. Acad. Sci. USA*, 100:7714–7719, 2003.

M. Isalan, C. Lemerle, and L. Serrano. Engineering gene networks to emulate *Drosophila* embryonic pattern formation. *PLoS Biol.*, 3(3):e64, 2005.

Y. Jia and S. S. Patel. Kinetic mechanism of transcription initiaion by bacteriophage T7 RNA polymerase. *Biochemistry*, 36:4223–4232, 1997.

M. Jiang, M. Rong, C. Martin, and W. T. McAllister. Interrupting the template strand of the T7 promoter facilitates translocation of the DNA during initiation, reducing transcript slippage and the release of abortive products. *J. Mol. Biol.*, 310(3):509–522, 2001.

W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner, and D. P. Bartel. RNA-catalyzed RNA polymerization: Accurate and general RNA-templated primer extension. *Science*, 292: 1319–1325, 2001.

M. Kaern, T. C. Elston, W. J. Blake, and J. J. Collins. Stochasticity in gene expression: From theories to phenotypes. *Nature Rev. Genet.*, 6:451–464, 2005.

N. Kashtan and U. Alon. Spontaneous evolution of modularity and network motifs. *Proc. Natl. Acad. Sci. USA*, 102:13773–13778, 2005.

J. Kim, J. J. Hopfield, and E. Winfree. Neural network computation by *in vitro* transcriptional circuits. In *Advances in Neural Information Processing Systems (NIPS)*, volume 17, pages 681–688, 2004.

J. Kim, K. S. White, and E. Winfree. Construction of an *in vitro* bistable circuit from synthetic transcriptional switches. *Mol. Syst. Biol.*, 2:68, 2006.

P. M. Kim and B. Tidor. Limitations of quantitative gene regulation models: A case study. *Genome Res.*, 13:2391–2395, 2003.

L. Klungsøyr, J. H. Hagemen, L. Fall, and D. E. Atkinson. Interaction between energy charge and product feedback in the regulation of biosynthetic enzymes. Aspartokinase, phosphoribosyladenosine triphosphate synthetase, and phosphoribosyl pyrophosphate synthetase. *Biochemistry*, 7(11):4035–4040, 1968.

C. Kramer, J. J. Loros, J. C. Dunlap, and S. K. Crosthwaite. Role for antisense RNA in regulating circadian clock function in *Neurospora crassa*. *Nature*, 421:948–952, 2003.

I. Kuzmine and C. T. Martin. Pre-steady-state kinetics of initiation of transcription by T7 RNA polymerase: A new kinetic model. *J. Mol. Biol.*, 305:559–566, 2001.

S. B. Laughlin and T. J. Sejnowski. Communication in neuronal networks. *Science*, 301:1870–1874, 2003.

K. Lee, W. D. McCormick, J. E. Pearson, and H. L. Swinney. Experimental observation of self-replicating spots in a reaction-diffusion system. *Nature*, 369:215–218, 1994.

D. M. Lilley. The origins of RNA catalysis in ribozymes. *Trends Biochem. Sci.*, 28:495–501, 2003.

W. F. Lima and S. T. Crooke. Cleavage of single strand RNA adjacent to RNA-DNA duplex regions by *Escherichia coli* RNase H1. *J. Biol. Chem.*, 272:27,513–27,516, 1997.

W. Maass. On the computational power of winner-take-all. *Neural Computation*, 12:2519–2535, 2000.

K. Madin, T. Sawasaki, T. Ogasawara, and Y. Endo. A highly efficient and robust cell-free protein synthesis system prepared from wheat germ embryos: Plants apparently contain a suicide system directed at ribosomes. *Proc. Natl. Acad. Sci. USA*, 97(2):559–564, 2000.

M. Mandal and R. R. Breaker. Gene regulation by riboswitches. *Nature Rev. Mol. Cell. Biol.*, 5: 451–463, 2004.

S. A. E. Marras, F. R. Kramer, and S. Tyagi. Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic Acids Res.*, 30(21):e122, 2002.

C. T. Martin and J. E. Coleman. Kinetic analysis of T7 RNA polymerase-promoter interactions with small synthetic promoters. *Biochemistry*, 26:2690–2696, 1987.

M. Maslak and C. T. Martin. Kinetic analysis of T7 RNA polymerase transcription initiation from promoters containing single-stranded regions. *Biochemistry*, 32:4281–4285, 1993.

J. S. Mattick and M. J. Gagen. Accelerating networks. *Science*, 307:856–858, 2005.

J. S. Mattick and I. V. Makunin. Non-coding RNA. *Human Mol. Genet.*, 15:R17–R29, 2006.

H. H. McAdams and A. Arkin. Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. USA*, 94:814–819, 1997.

W. R. McClure. Mechanism and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.*, 54:171–204, 1985.

W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.*, 5:115, 1943.

B. Meerson and P. V. Sasorov. Domain stability, competition, growth, and selection in globally constrained bistable systems. *Phys. Rev. E*, 53:3491–3494, 1996.

A. P. Mills Jr., B. Yurke, and P. M. Platzman. Article for analog vector algebra computation. *Biosystems*, 52:175–180, 1999.

E. Mjolsness, D. H. Sharp, and J. Reinitz. A connectionist model of development. *J. Theor. Biol.*, 152:429–453, 1991.

J. Monod and F. Jacob. General conclusions: Teleonomic mechanisms in cellular metabolism, growth and differentiation. *Cold Spring Harb. Symp. Quant. Biol.*, 26:389–401, 1961.

C. B. Muratov and V. V. Osipov. Scenarios of domain pattern formation in a reaction-diffusion system. *Phys. Rev. E*, 54:4860–4879, 1996.

A. R. Mushegian and E. V. Koonin. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc. Natl. Acad. Sci. USA*, 93:10,268–10,273, 1996.

M. Nakajima, K. Imai, H. Ito, T. Nishiwaki, Y. Murayama, H. Iwasaki, T. Oyama, and T. Kondo. Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation *in vitro*. *Science*, 308:414–415, 2005.

V. Noireaux, R. Bar-Ziv, and A. Libchaber. Principles of cell-free genetic circuit assembly. *Proc. Natl. Acad. Sci. USA*, 100(22):12,672–12,677, 2003.

V. Noireaux and A. Libchaber. A vesicle bioreactor as a step toward an artificial cell assembly. *Proc. Natl. Acad. Sci. USA*, 101(51):17,669–17,674, 2004.

E. Nudler and A. S. Mironov. The riboswitch control of bacterial metabolism. *Trends Biochem. Sci.*, 29:11–17, 2004.

E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden. Regulation of noise in the expression of a single gene. *Nature Genet.*, 31:69–73, 2002.

E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. van Oudenaarden. Multistability in the lactose utilization network of *Escherichia coli*. *Nature*, 427:737–740, 2004.

J. R. Pomerening, S. Y. Kim, and J. E. Ferrell, Jr. Systems-level dissection of the cell-cycle oscillator: Bypassing positive feedback produces damped oscillations. *Cell*, 122:565–578, 2005.

M. Ptashne. *A Genetic Switch, 2nd ed.* Cell Press & Blackwell, 1992.

J. M. Raser and E. K. O'Shea. Control of stochasticity in eukaryotic gene expression. *Science*, 304: 1811–1814, 2004.

R. Reed. Pruning algorithms – a survey. *IEEE Trans. Neural Net.*, 4:740–747, 1993.

S. M. Reppert and D. R. Weaver. Coordination of circadian timing in mammals. *Nature*, 418: 935–941, 2002.

J. Rizzo, L. K. Gifford, X. Zhang, A. M. Gewirtz, and P. Lu. Chimeric RNA-DNA molecular beacon assay for ribonuclease H activity. *Molecular and Cellular Probes*, 16:277–283, 2002.

N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, and M. B. Elowitz. Gene regulation at the single-cell level. *Science*, 307:1962–1965, 2005.

J. SantaLucia, Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. USA*, 95:1460–1465, 1998.

M. A. Savageau. Parameter sensitivity as a criterion for evaluating and comparing the performance of biochemical systems. *Nature*, 229:542–544, 1971.

N. C. Seeman. Nucleic-acid junctions and lattices. *J. Theor. Biol.*, 99:237–247, 1982.

M. A. Shea and G. K. Ackers. The o_r control system of bacteriophage lambda. *J. Mol. Biol.*, 181: 211–230, 1985.

S. S. Shen-Orr, R. Milo, S. Mangan, and U. Alon. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genet.*, 31:64–68, 2002.

Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, and T. Ueda. Cell-free translation reconstituted with purified components. *Nature Biotech.*, 19:751–755, 2001.

S. H. Strogatz. Exploring complex networks. *Nature*, 410:268–276, 2001.

J. W. Szostak, D. P. Bartel, and P. L. Luisi. Synthesizing life. *Nature*, 409:387–390, 2001.

D. W. Tank and J. J. Hopfield. Simple 'neural' optimization networks: An A/D converter, signal decision circuit, and a linear programming circuit. *IEEE Trans. Circ. Syst.*, 33:533–541, 1986.

C. D. Thron. Theoretical dynamics of the cyclin B-MPF system: A possible role for p13^{suc1}. *Biosystems*, 32:97–109, 1994.

J. Tomita, M. Nakajima, T. Kondo, and H. Iwasaki. No transcription-translation feedback in circadian rhythm of kaiC phosphorylation. *Science*, 307:251–254, 2005.

F. J. Triana-Alonso, M. Dabrowski, J. Wadzack, and K. H. Nierhaus. Self-coded 3'-extension of run-off transcripts produces abberant products during *in vitro* transcription with T7 RNA polymerase. *J. Biol. Chem.*, 270:6298–6307, 1995.

E. van Nimwegen. Scaling laws in the functional content of genomes. *Trends Genet.*, 19:479–484, 2003.

B. D. Ventura, C. Lemerle, K. Michalodimitrakis, and L. Serrano. From *in vivo* to *in silico* biology and back. *Nature*, 443:527–533, 2006.

J. M. G. Vilar, H. Y. Kueh, N. Barkai, and S. Leibler. Mechanisms of noise-resistance in genetic oscillators. *Proc. Natl. Acad. Sci. USA*, 99:5988–5992, 2002.

C. A. Voigt, D. M. Wolf, and A. P. Arkin. The *Bacillus subtilis sin* operon: An evolvable network motif. *Genetics*, 169(3):1187–1202, 2005.

B. Wlotzka and J. S. McCaskill. A molecular predatory and its prey: Coupled isothermal amplification of nucleic acids. *Chemistry and Biology*, 4:25–33, 1997.

D. M. Wolf and A. P. Arkin. Motifs, modules and games in bacteria. *Curr. Opin. Microbiol.*, 6: 125–134, 2003.

W. W. Wong and J. C. Liao. The design of intracellular oscillators that interact with metabolism. *Cell Mol. Life Sci.*, 63:1215–1220, 2006.

H. Yan, X. Zhang, Z. Shen, and N. C. Seeman. A robust DNA mechanical device controlled by hybridization topology. *Nature*, 415:62–65, 2002.

Y. Yokobayashi, R. Weiss, and F. H. Arnold. Directed evolution of a genetic circuit. *Proc. Natl. Acad. Sci. USA*, 99(26):16,587–16,591, 2002.

L. You, R. S. Cox III, R. Weiss, and F. H. Arnold. Programmed population control by cell-cell communication and regulated killing. *Nature*, 428:868–871, 2004.

M. W. Young and S. A. Kay. Time zones: A comparative genetics of circadian clocks. *Nature Rev. Genet.*, 2:702–715, 2001.

A. L. Yuille and D. Geiger. Winner-take-all mechanisms. In *The Handbook of Brain Theory and Neural Networks*. MIT Press, 1995.

B. Yurke and A. P. Mills, Jr. Using DNA to power nanostructures. *Genet. Program. and Evol. Mach.*, 4:111–122, 2003.

B. Yurke, A. J. Turberfield, A. P. Mills, Jr., F. C. Simmel, and J. L. Nuemann. A DNA-fuelled molecular machine made of DNA. *Nature*, 406:605–608, 2000.

H. S. Zaher and P. J. Unrau. T7 RNA polymerase mediates fast promoter-independent extension of unstable nucleic acid complexes. *Biochemistry*, 43:7873–7880, 2004.

A. N. Zaikin and A. M. Zhabotinsky. Concentration wave propagation in 2-dimensional liquid-phase self-oscillating system. *Nature*, 225:535–537, 1970.