UNIVERSITY OF CALIFORNIA, SAN DIEGO

Synthetic gene oscillators and their applications

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Bioengineering

by

Tal Danino

Committee in Charge:

Jeff Hasty, Chair
Susan Golden
Kit Pogliano
Lev Tsimring
Shyni Varghese
Ruth Williams

2011

The dissertation of Tal Danino is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

_____
Chair

University of California, San Diego

2011

DEDICATION

To my friends, family, and bacteria

TABLE OF CONTENTS

Signature Page iii

Dedication iv

Table of Contents v

List of Abbreviations viii

List of Figures ix

List of Tables xi

Acknowledgements xii

Vita xiv

Abstract of the Dissertation xvi

Chapter 1 Introduction 1

Chapter 2 A synchronized quorum of genetic clocks 7
- Introduction 7
- Synchronized genetic oscillators 9
- Spatiotemporal Dynamics 13
- Quantitative modeling 14
- Emergence 18
- Perspective and outlook 18
- Methods 19
  - Strains, growth conditions 19
- Plasmid Construction 20
- Data Analysis 20
  - Space-Time plots 21
- Microscopy and Microfluidics 22
- Modeling 25
- Supplementary Movies 30
- Acknowledgements 31

Chapter 3 Modeling Synthetic Gene Networks 32
- Steady state modeling of genetic networks 32
- Positive feedback gene network 35
- Steady states 37
- Linear stability 39

Negative feedback 40

Chapter 4 In-silico patterning of vascular mesenchymal cells in three-dimensions 43
Introduction 43
Results 47
Discussion 51
Materials and Methods 54
Parameter selection 55
Mathematical derivation of the model 56
Assumptions and Simplifications 57
Linear stability analysis 59
Acknowledgements 61

Chapter 5 Correlated signaling through coupled degradation 63
Introduction 63
Results and Discussion 65
Strains and Flow Cytometry 76
Microfluidics and Microscopy 77
Acknowledgements 91

Chapter 6 Summary and Conclusion 93

Chapter 7 Lab Procedures and Protocols 94
Overview of Molecular Cloning 96
QIAprep Miniprep 99
NanoDrop Spectrophotometer 102
Agarose Gel Electrophoresis 103
Gel Digests 106
Gel Purification 109
Ligations 110
Transformations 111
PCR protocol 113
CPEC Reactions 115
PCR Purification 117
Sequencing DNA (Eton Bio's service) 118
Colony PCR 119
PCR optimization 120
Designing PCR Primers for CPEC 122
Screening colonies from CPEC Reactions 125
Making Competent Cells 129
Making PDMS chips for Microscope Experiments 130
Microscope Experiment Protocol DAW6v2 134
Flow Cytometry 137

Preparing Electrocompetent cells and Electroporations 138

Bibliography 140

LIST OF ABBREVIATIONS

A.U. . . . . arbitrary units

bp . . . . . base pair

DNA . . . . . deoxyribonucleic acid

E. coli . . . . . *Escherichia coli*

Eq. . . . . equation

fig. . . . . figure

FL. . . . . fluorescence

FP . . . . . fluorescent protein

FSC . . . . . forward scatter

GFP . . . . . green fluorescent protein

hrs. . . . . hours

kb . . . . . kilobase

LB . . . . . lysogeny broth

mRNA . . . . . messenger ribonucleic acid

OD . . . . . optical density

ODE . . . . . ordinary differential equation

PCR . . . . . polymerase chain reaction

PDMS . . . . . poly(dimethylsiloxane)

UV . . . . . ultraviolet

YFP . . . . . yellow fluorescent protein

%w/v . . . . . percent weight per volume (concentration)

LIST OF FIGURES

Figure 2.1: Synchronized genetic clocks 10
Figure 2.2: Dynamics of the synchronized oscillator under multiple microfluidic flow conditions 12
Figure 2.3: Spatiotemporal dynamics of the synchronized oscillators 15
Figure 2.4: Modeling of synchronized genetic clocks 17
Figure 2.5: Plasmids for the synchronized oscillator strain TDQS1. 21
Figure 2.6: Stable oscillations in microfluidic device 22
Figure 2.7: Space-time plot of density of cells in Fig3a experiment 23
Figure 2.8: Microfluidic Devices constructed for this study 26
Figure 2.9: Synchronization of oscillations in spatially extended system with diffusion 28
Figure 2.10: Wave propagation in the spatially uniform system with different external AHL diffusion rates 29

Figure 3.1: Modeling biological processes 32
Figure 3.2: Positive feedback gene network 35
Figure 3.3: Steady state value of x as a function of g 38
Figure 3.4: Repression of transcription occurs when dimeric protein x binds to a promoter site 40

Figure 4.1: Diagram showing interactions between BMP-2, MGP, and cells in culture 47
Figure 4.2: 2D steady state patterns of cells 49
Figure 4.3: 3D steady state patterns of cells 50
Figure 4.4: Initial and steady state patterns of cells produced by exogenous BMP-2 52
Figure 4.5: Dispersion relation for k=1.0, c=0.1. The other parameters used here are listed in the supplementary text. 60
Figure 4.6: Phase diagrams for various χ values 61

Figure 5.1: Coupling via common enzymatic machinery: connection to queueing theory 67
Figure 5.2: Coupled enzymatic degradation of yellow and cyan LAA-tagged fluorescent proteins by ClpXP machinery in *E. coli* 71
Figure 5.3: Dynamic behavior of a synthetic signaling network 73
Figure 5.4: Coupled degradation can serve to indirectly couple gene circuits . . . 75
Figure 5.5: Coupled enzymatic degradation of yellow and cerulean LAA-tagged fluorescent proteins by ClpXP machinery in *E. coli* 80
Figure 5.6: Inverted image of Western blot film taken for cellular lysate data . . . 81
Figure 5.7: Cell count correlated with OD 83
Figure 5.8: Mean total protein for the stochastic queueing model with finite Michaelis-Menten constant K 92

Figure 7.1: Overview of CPEC cloning 96
Figure 7.2: Overview of restriction enzyme cloning 97
Figure 7.3: Naming convention of pZ plasmids 98
Figure 7.4: NEB 1kb and 2 Log Ladders 104
Figure 7.5: Overview of Primer Design for CPEC. Primers are designed in two parts, the annealing to template part, and the complementary part, where each have their own but similar design criteria. 122

LIST OF TABLES

Table 4.1: Monomeric BMP-2 (U) is able to dimerize (U_2) and bind to MGP (V) or bind to receptor complexes (R_2, S_2) producing either BMP-2 or MGP, respectively. BMP-2 and MGP degrade at rates c, e. 56

ACKNOWLEDGEMENTS

The last six years have been an amazing journey that I have so many people to thank for. First off, I would like to thank my family and friends who have given me support in every way imaginable and have helped me make it this far. I would like to thank everyone in the lab, who has at some point either taught me a valuable laboratory skill or changed my perspective on science. In particular, I would like to thank Jesse for teaching me molecular cloning and passing off his ad-hoc experimental style to me, and Octavio who has been my close friend and collaborator throughout the years. I am greatly thankful for all of the younger students who I have had the opportunity to work with and pass along my knowledge to (Ivan, Jangir, Arthur, and Phillip), as well as all of the older students who I have enjoyed working with, hanging out with, and making fun of the younger students with over the years (Martin, Natalie, Scott, and Mike). I am also hugely indebted to the undergraduates that have done as much work as me during this thesis (Leo, Cathy, Anna, and Chris). I am grateful to have a great co-advisor, Lev Tsimring, whom I thank for his ability to understand and provide insight to research problems I've worked on for years in just a few minutes. And last but not least, I am greatly indebted to my advisor, Jeff Hasty, for his unfailingly positive support and mentorship throughout the years. His creativity and scientific style have had a tremendous impact on me as a scientist and as a person and I cannot thank him enough.

Chapter 2 contains material originally published as Danino, T.*, Mondragron-Palomino, O.*, Tsimring, L., and Hasty, J. ***Nature*** (2010): A synchronized quorum of genetic clocks. (*equal contribution). Copyright permission to republish here was granted by Nature publishing group.

Chapter 4 contains material originally published as Cookson, N.A.*, Mather, W.H.*, Danino, T., Mondragon-Palomino, O., Williams, R. J., Tsimring, L. S., and Hasty, J. ***Nature*** (in review): Queueing up for enzymatic processing: Correlated signaling through coupled degradation. (*equal contribution). Copyright permission to

republish here was granted by Nature publishing group.

Chapter 5 contains material originally published as Danino,T. Volfson, D., Tsimring, L., and Hasty, J. ***PLoS ONE*** (2011): In-silico pattern formation of vascular mesenchymal stem-cells in three-dimensions. Copyright permission to republish here was granted by Public library of science.

VITA

2005	Bachelor of Science in Physics, Chemistry, Math University of California, Los Angeles
2008	Master of Science in Bioengineering University of California, San Diego
2011	Doctor of Philosophy in Bioengineering University of California, San Diego

PUBLICATIONS

Peer Reviewed Journal Articles

A synchronized quorum of genetic clocks. Danino, T., Mondragron-Palomino, O., Tsimring, L., and Hasty, J. ***Nature*** (2010)

Streaming instability in growing cell populations. Mather, W., Mondragon-Palomino, O., Danino, T., Hasty, J., and Tsimring, L.S. ***Physical Review Letters*** (2010)

In-silico pattern formation of vascular mesenchymal stem-cells in three-dimensions. Danino,T. Volfson, D., Tsimring, L., and Hasty, J. ***PLoS ONE*** (2011)

Buckling instability in ordered bacterial colonies. Boyer, D., Mather, W., Mondragon-Palomino, O., Orozco-Fuentes, S., Danino, T., Hasty, J., and Tsimring, L. S. ***Physical Biology***.

Queueing up for enzymatic processing: Correlated signaling through coupled degradation. Cookson, N.A., Mather, W.H., Danino, T., Mondragon-Palomino, O., Williams, R. J., Tsimring, L. S., and Hasty, J. ***Nature*** (in review)

Entrainment of a synthetic gene oscillator. Mondragon-Palomino, O., Danino, T., Selimkhanov, J., Tsimring, L.S., and Hasty, J. ***Science*** (in review)

FIELDS OF STUDY

Major Field: Bioengineering (Synthetic Biology)

Studies in Biological Dynamics
Professor Jeff Hasty and Dr. Lev S. Tsimring

ABSTRACT OF THE DISSERTATION

Synthetic gene oscillators and their applications

by

Tal Danino

Doctor of Philosophy in Bioengineering

University of California, San Diego, 2011

Jeff Hasty, Chair

Synthetic biology seeks to understand and engineer biological networks that perform a quantitative dynamic function in organisms. Since the original toggle switch (Gardner et al., 2000a) and oscillator designs (Elowitz and Leibler, 2000a), genetic circuits have been constructed that control cellular population growth (You et al., 2004b), detect edges in an image (Friedland et al., 2009), and count discrete cellular events (Friedland et al., 2009). In this thesis, we focus on synthetic gene circuits that produce oscillations. Oscillations are important in a vast range of natural contexts such as circadian rhythms, cardiac function, cell division, and hormonal regulation, as well as key to building synthetic control systems that rely on precise timing. Here we discuss modeling, designing, constructing, and characterizing synthetic gene oscillators. In Chapter One, we give an overview and introduction to the field of synthetic biology and how our research area fits into this discipline. In Chapter Two, I discuss a network design which produces synchronized oscillations in a growing population of cells. In Chapter Three, I discuss an introduction to modeling simple genetic networks. In Chapter Four,

we look at a scenario where synthetic gene circuits that produce an overabundance of tagged components lead to unexpected correlations. In Chapter Five, I discuss modeling genetic networks and further go into detail about spatial modeling of networks that produce patterns in mammalian cells. These parts combine to illustrate how to design, model, construct, and characterize synthetic gene networks in bacteria and mammalian systems.

Chapter 1

Introduction

Our technological ability to rapidly sequence and synthesize DNA has transformed how we study biological systems. Sequencing and high-throughput technologies allow us to characterize organisms at a large scale to understand the connectivity and sequence dependence on genetic networks, emcompassed by the field of ”systems biology.” Complementary to this aim, our ability to construct synthetic gene networks by rearranging basic biological elements such as promoters and genes from various organisms have allowed us to study networks in isolation to gain an understanding of how smaller motif's function. This has resulted in the emergence of ”synthetic biology”, a discipline which seeks to understand small modular gene ”circuits” and systematically increase complexity to understand larger systems.

In the field of synthetic biology, research areas can be roughly parsed into two areas: large scale DNA synthesis and construction of synthetic organisms, and ”bottom-up” DNA construction of small genetic circuits with quantitative and dynamic function. These two efforts differ greatly in the number of genes, from hundreds of genes in the former to only a handful of genes in the latter. They differ in their level of quantitative detail, by producing simple growth/no growth information, to data relaying the amount of proteins in single-cells and their behavior as a function of time. Both aims are complimentary, and represent synthetic biologists' attempts to design life by constructing

DNA circuits to perform various functions.

One of the overarching philosophies of the bottom-up approach is that complex genetic circuits can be dissected into small common regulatory motifs that are easier to understand. Often times large genetic circuits are controlled by core motifs that occur in small circuits, with additional ”bells and whistles” added on. Thus, focusing on the behavior of different small circuits of these systems allows us to test our understanding of nature’s underlying design of gene networks. In particular, knowing how to design the sequences of DNA that produce a desired dynamic function in cells is something synthetic biologists strive for. This necessitates design criteria backed by quantitative modeling to account for transcription, translation, and regulation of these DNA circuits, and to determine other cellular processes affect and are affected by these circuits. Researchers have looked to nature for inspiration on how to design and regulate genetic information, taking useful pieces from various organisms, and on the flipside, building these circuits has shed valuable information on how nature puts these things together through evolution. Being able to predict how a given design of a DNA circuit behaves is key feature of synthetic biology.

The ability to obtain precise measurements of genetic circuits at the single-cell level is crucial for understanding and modeling these systems. Flow cytometry techniques have allowed us to gather statistics on tens of thousands of cells in a population, and thus the ability to look at the dynamics of synthetic circuits and variability between cells in a population. Although these produce a large amount of statistics, one particular cell cannot be tracked as a function of time. The use of measurement technologies such as microfluidics coupled with fluorescence microscopy allows collection of cell trajectories at the single-cell level over long durations of time (Stricker et al., 2008a; Cookson et al., 2005; Danino et al., 2010, PMID: 20090747). These platforms also allow for dynamic changing the environment while keeping cells in a healthy phase of growth (Bennett et al., 2008). In addition, platforms for parallelizing collection of data for multiple inducer concentrations and automated data extraction are necessary to streamline the characterization of these networks(Locke and Elowitz, 2009). As our

ability to construct genetic circuits becomes increasingly fast, constructing new high-throughput technologies such as microfluidics to characterize these circuits will be essential.

We briefly give a survey of recent projects in synthetic biology that give an impression of the field. Scientists working on the first aim recently assembled an entire synthetic Mycoplasma genome of 583 kilobases and showed that it can be inserted into an empty Mycoplasma and grow stably (Gibson et al., 2008a). They constructed this genome by using the ability of yeast to homologously recombine DNA from 101 chemically synthesized cassettes ranging from 5-7kb , and later showing that one can do the entire assembly with 25 longer cassettes (Gibson et al., 2008b) . In the future, this can lead to engineering of entirely new microbes tailored to specific industrial applications, such as to produce biofuels, or other chemicals of interest (Martin et al., 2003)(others) instead of our current efforts to modify and tweak existing microbes to our needs. In the second area, researchers created a wide range of synthetic gene circuits such as switches, filters, sensors, and oscillators in *E.coli* (Elowitz and Leibler, 2000a; Gardner et al., 2000a; Stricker et al., 2008a; Tabor et al., 2009a). One example is creating a gene circuit in *E.coli* which can detect light and produce a color pigment at the edge of a mask. Using computational modeling and cloning techniques the authors were able to produce images of 100Megapixels/square inch. Somewhere in the scale of genes between these two aims are synthetic gene circuits that address important medical and industrial problems with engineered organisms such as bacteria that invade cancer cells, yeast with synthetic pathways to produce antimalarial drug precursors, and bacteria which can digest plant fiber to produce a range of biodiesel components (Anderson et al., 2006; Martin et al., 2003; Steen et al., 2010). In this chapter, we will primarily focus on how to model small genetic circuits dynamically, and focus on oscillators as a model system.

Oscillators are important circuits in the context of synthetic biology as well as in the natural world. In nature, oscillations occur in the daily rhythms of organisms' sleep-wake cycles, cell division, the cell cycle, and in cardiac rhythms. The first synthetic gene oscillator, named the 'repressilator', is one of the circuits that started the field

synthetic biology ten years ago (Elowitz and Leibler, 2000a; Gardner et al., 2000a). Not only could oscillators be used as synthetic timing devices, they represent an important circuit that shows the usefulness of model based design criteria for gene circuits. Since oscillations are periodic, they restrict the structure of equations that can be used to model them and the parameter space which can produce oscillations. For instance, in establishing the design criteria for the repressilator, the authors learned from the model that to get oscillations, repressor protein half lives on the order of *E.coli* mRNA half lives were needed, and thus degradation tags were added to the proteins to decrease their half lives. Modeling how these gene circuits behave also helps us gain intuition about biological systems. For example, in constructing a robust&tunable synthetic oscillator (Stricker et al., 2008a), modeling the many kinetic reactions for the processes of transcription, translation, and protein maturation, led to the insight that an effective form of delay made oscillations robust. This intuition from the modeling led to experimental construction of of a 1 component genetic oscillator based solely on negative feedback and delay. A population based synchronized oscillator was constructed in a similar way, with coupled positive & negative feedback used to produce robust oscillations, and delay differential equations showed that the delay to be an important component.

In modeling the synthetic oscillators we mentioned above, it is important to consider approaches to model the dynamics of the basic (1) positive feedback and (2) negative feedback motifs. Each of these have interesting properties and arise in different natural contexts. Positive feedback systems, systems where a gene X activates its own or another gene's production, is a common regulatory motif found in nature. It can lead to *bistability*, where more than one stable steady state exists. These types of networks act as an important decision making circuit in cells which allows switch between an "ON" state, with high level of gene production, to a "OFF" state, with a low level of gene production, without the needed for persistent external cues. They arise in many natural systems such as the cell cycle, cell differentiation, apoptosis, nutrient utilization in bacteria (Ozbudak et al., 2004), the yeast mating response (Paliwal et al., 2007), and several synthetically constructed circuits (Gardner et al., 2000a; Isaacs et al., 2003). In

many oscillating circuits, bistability can lead to oscillations between these two states. Negative feedback systems, where gene X represses production of another gene is another common motif seen in many natural systems such regulation of body temperature and glucose in humans, and galactose utilization in yeast (Bennett et al., 2008), as well as synthetic circuits such as the repressilator (Elowitz and Leibler, 2000a). Negative feedback systems can also lead to bistability as well. Negative feedback can speed up the response time of gene circuits (Savageau, 1974; Rosenfeld et al., 2002) and function to promote robustness to fluctuations in production rate (Becskei and Serrano, 2000). Most organisms utilize both of these feedback systems together to create common network motifs, logic and sensors to respond and adapt to their environment (Alon, 2007a).

Synthetic oscillators in *E. Coli* use positive and negative feedbacks to produce oscillations, and each feedback as a distinct role. For instance, a single negative feedback of the *lacI* gene could produce and sustain oscillations in *E. coli*. Here the intermediate steps between transcription of a gene X to multimerization causes a ”delay” which allows the system to turn on for some time before shutting itself off. With additional positive feedback added to this circuit, activation & amplitude gets higher, which leads to a longer period of oscillations and a tunable range of periods depending on the strength of the positive feedback. When modeling these circuits it was important to use the fact that proteins are ”tagged” for degradation, i.e., a sequence of DNA is added to the termini of these proteins which is then recognized by a protease(ClpXP) and degraded quickly. In this scenario, enzymes are nearly saturated and the proteins decay at a linear rate independent of their concentration. This mechanism, known as ”degrade and fire”, has the property where the burst amplitude is proportional to period, since waiting for ClpXP to enzymatically decay the activator and repressor is the dominant contribution to the period in many cases. This particular scenario is important to synthetic biologists because it seems to form a fundamental limit to how many tagged components can be used in a particular gene circuit. In Chapter 4, we present work that shows the coupling of tagged proteins from circuits that would normally be expected to function independently.

Not only do simple genetic networks produce a rich variety of dynamic behavior, they can also produce interesting spatial patterns(Murray and Oster, 1984; Gierer and Meinhardt, 1972; Tsimring et al., 1995; Garfinkel et al., 2004; Murray, 2003; Koch and Meinhardt, 1994; Painter et al., 1999; Brenner et al., 1998; Gamba et al., 2003; Sage et al., 2009; Ambrosi and Preziosi, 2006). In these systems, we can use similar modeling approaches starting with mass-action kinetics and add additional diffusion and chemotaxis terms to model the behavior of these networks. In chapter 5 , we present a case study where mammalian cells form ”stripe” patterns in culture and develop a model based on basic biochemical equations(Garfinkel et al., 2004). These spatial patterns are produced by the influence of ”morphogens” that cells move towards and influence gene expression from positive and negative feedbacks. In particular, we were interested in studying how these patterns and interactions might change in three-dimensional situations. Three dimensional simulations were developed and used to make predictions on the pattern formation of these cells. In addition, one common assumption made in these scenarios is the quasi-steady state assumption(QSSA), and is not always a valid approximation. We show how a more accurate ”prefactor” method describes transient dynamics better of monomers more than the QSSA assumption.

In the this thesis, I will go through a particular example of constructing, characterizing, and modeling a specific gene network that produces synchronized oscillations in *E. coli*. Then I will give an introduction to basic modeling of positive and negative feedbacks that occur in networks throughout this thesis. In addition, we will look at how some synthetic gene networks currently have a limited capacity due to crosstalk in the degradation of molecular components and how this is an important feature to model. Finally, we will end with a spatio-temporal modeling section of mammalian cells, where we simulate how these cells can arrange in complex spatial patterns in three dimensions.

Chapter 2

A synchronized quorum of genetic clocks

Introduction

The engineering of genetic circuits with predictive functionality in living cells represents a defining focus of the expanding field of synthetic biology. This focus was elegantly set in motion ten years ago with the design and construction of a genetic toggle switch and oscillator, with subsequent highlights that have included circuits that are capable of generating patterns, shaping intracellular noise, detecting edges in an image, and counting discrete events. Here, we describe an engineered gene network with global intercellular coupling that is capable of generating synchronized oscillations in a growing population of cells. Using microfluidic devices tailored for cellular populations at differing length scales, we investigate the collective synchronization properties along with spatiotemporal waves occurring on millimeter scales. We use computational modeling to quantitatively describe the observed dependence of the period of bulk oscillations on the flow rate and oscillatory amplitude. The synchronized genetic clock sets the stage for the use of microbes in the creation of a macroscopic biosensor with an oscillatory output. In addition, it provides a specific model system for the generation of

a mechanistic description of emergent coordinated behavior at the colony level.

Centralized clocks are of fundamental importance in the coordination of rhythmic behavior among individual elements in a community or a large complex system. In physics and engineering, the Huygens paradigm of coupled pendulum clocks (Mirollo and Strogatz, 1990; Pikovsky et al., 2002; Bennett et al., 2002) has permeated diverse areas from the development of arrays of lasers (Vladimirov et al., 2003) and superconducting junctions (Wiesenfeld et al., 1996) to GPS (Lewandowski et al., 1999) and distributed sensor networks (Li et al., 2002). Perhaps one of the most bizarre (and unintended) examples of synchronization in engineering involved London's Millennium Bridge (GC, 2005), which originally had a resonant frequency that was near the walking frequency of a typical pedestrian. On opening day, out of step pedestrians set the suspension bridge to wobble with a motion that coupled back on the pedestrians and induced synchronized marching which, in turn, further amplified the swaying of the bridge.

In biology, synchronized rhythms are abound, with behavioral examples that include flashing fireflies (Buck and Buck, 1968), swarming locusts (Buhl et al., 2006), and activity waves in ant colonies (Boi et al., 1999). In terms of human physiology, a vast range of intercellular coupling mechanisms lead to synchronized oscillators which govern fundamental processes such as somitogenesis, cardiac function, respiration, insulin secretion, and circadian rhythms (Winfree, 1967; Mirollo and Strogatz, 1990; Elson et al., 1998; Jiang et al., 2000; Glass, 2001; Young and Kay, 2001; Chabot et al., 2007; Kerckhoffs et al., 2009). Typically, synchronization helps stabilize a desired behavior arising from a network of intrinsically noisy and unreliable elements. Sometimes, however, the synchronization of oscillations can lead to a severe malfunction of a biological system, as in epileptic seizures (Grenier et al., 2003).

There is widespread interest in the use of synthetic biology to recreate complex cellular behavior from the underlying biochemical reactions that govern gene regulation and signaling. Synthetic biology can be broadly parsed into efforts aimed at the large-scale synthesis of DNA and the forward engineering of genetic circuits from known

biological components. In the area of DNA synthesis, pathways have been perturbed and replaced (Isalan et al., 2008) in an effort to in an effort to understand the network motifs and transcriptional regulatory mechanisms that control cellular processes and elicit phenotypic responses (Alon, 2007b). On a larger scale, progress has been made towards the creation of entire genomes, providing new insights into what constitutes the minimal set of genes required for microbial life (Gibson et al., 2008a). The genetic circuits approach (Hasty et al., 2002a; Endy, 2005) involves the use of computational modeling in the design of relatively small genetic circuits. Here, the original toggle switch (Gardner et al., 2000a) and oscillator (Elowitz and Leibler, 2000b) have inspired the design and construction of circuits capable of controlling cellular population growth (You et al., 2004b), generating patterns (Basu et al., 2005), triggering biofilm development (Kobayashi et al., 2004), shaping intracellular noise (Austin et al., 2006), detecting edges in an image (Tabor et al., 2009a), and counting discrete cellular events (Friedland et al., 2009). In the context of rhythmic behavior, there have been recent successes in the construction of intracellular oscillators that mimic naturally occurring clocks (Atkinson et al., 2003; Stricker et al., 2008a; Tigges et al., 2009; Fung et al., 2005). Theoretical work has shown how the introduction of an autoinducer in oscillator designs can potentially lead to synchronized oscillations over a population of cells (McMillen et al., 2002; Garcia-Ojalvo et al., 2004).

Synchronized genetic oscillators

The synchronized oscillator design (Fig. 1a) is based on elements of the quorum sensing machineries in *Vibrio fisheri* and *Bacillus Thuringiensis*. We placed the *luxI* (from *V. fischeri*), *aiiA* (from *B. Thurigensis*) and *yemGFP* genes under the control of three identical copies of the *luxI* promoter. The LuxI synthase enzymatically produces an acyl-homoserine lactone (AHL), which is a small molecule that can diffuse between cells and mediates intercellular coupling. It binds intracellularly to the constitutively produced LuxR, and the LuxR-AHL complex is a transcriptional activator for

the *luxI* promoter (Waters and Bassler, 2005). AiiA negatively regulates the promoter through the effective degradation of AHL by catalyzing the degradation AHL (Liu et al., 2008). This network architecture, whereby an activator activates its own protease or repressor, is similar to the motif used in other synthetic oscillator designs (Atkinson et al., 2003; Stricker et al., 2008a; Tigges et al., 2009) and forms the core regulatory module for many circadian clock networks (Glossop et al., 1999; Young and Kay, 2001; Lakin-Thomas and Brody, 2004).

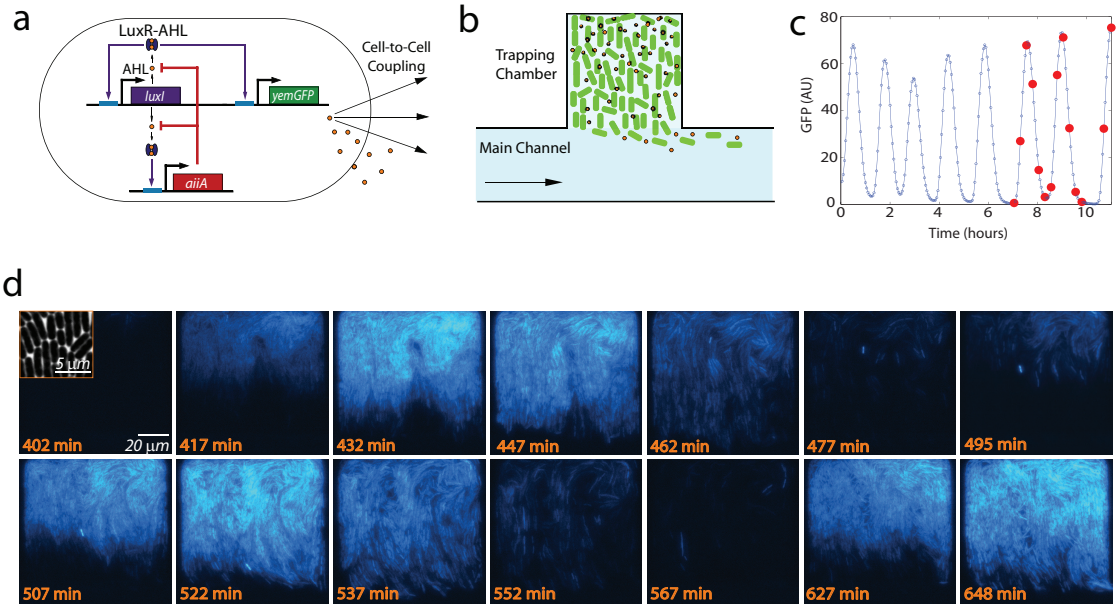


Figure 2.1: Synchronized genetic clocks. **(a)** Network Diagram. The *luxI* promoter drives production of the *luxI*, *aiiA*, and *yemGFP* genes in three identical transcriptional modules. LuxI enzymatically produces a small molecule AHL, which can diffuse outside of the cell membrane and into neighboring cells, activating the *luxI* promoter. AiiA negatively regulates the circuit by acting as an effective protease for AHL. **(b)** Microfluidic device used for maintaining *E. coli* at a constant density. The main channel supplies media to cells in the trapping chamber, and the flow rate can be externally controlled in order to change the effective degradation rate of AHL. **(c)** Bulk fluorescence as a function of time for a typical experiment in the microfludic device. The red circles correspond to the image slices in **(d)**. **(d)** Fluorescence slices of a typical experimental run demonstrate synchronization of oscillations in a population of *E.coli* residing in the microfluidic device (Supplementary Movie 1). Inset in the first snapshot is a 100x zoom of cells.

Most quorum sensing systems require a critical cell density for generation of

coordinated behavior (Reading and Sperandio, 2006). To control cell density, we monitored the synchronized oscillator cells (denoted TDQS1) at the single cell level by time-lapse fluorescence microscopy using microfluidic devices (Cookson et al., 2005). These devices consist of a main nutrient-delivery channel that feeds a rectangular trapping chamber (Fig. 1b). Once seeded, a monolayer of *E. coli* cells grow in the chamber and are eventually pushed into the channel where they then flow to the waste port. This device allows for a constant supply of nutrients or inducers and the maintenance of an exponentially growing colony of cells for more than four days. We found that chamber sizes of 100 x (80-100)μm were ideal for monitoring the intercellular oscillator, as they allowed for sufficient nutrient distribution and optimal cell and AHL densities. In the context of the design parameters, the flow rate can be modulated in order to change the local concentration of AHL. In addition, the device can be scaled up in order to permit the observation of spatial waves over longer length scales (see below).

After an initial transient period, the TDQS1 cells exhibit stable synchronized oscillations which are easily discernible at the colony level (Figs. 1c, 1d, and Supplementary Movies 1-3). The dynamics of the oscillations can be understood as follows. Since AHL diffuses downstream and is degraded by AiiA internally, a small colony of individual cells cannot produce enough inducer to activate expression from the *luxI* promoter. However, once the population reaches a critical density, there is a "burst" of transcription of the *luxI* promoters, resulting in increased levels of LuxI, AiiA, and GFP. As AiiA accumulates, it begins to degrade AHL, and after a sufficient time, the promoters return to their inactivated state. The production of AiiA is then attenuated, which permits another round of AHL accumulation and another burst of the promoters.

In order to determine how the effective AHL diffusion rate affects the period of the oscillations, we conducted a series of experiments at various channel flow rates. At high flow rate, the oscillations stabilize after an initial transient and exhibit a mean period of 90$\pm$6 minutes and mean amplitude of 54$\pm$6 GFP AU (Fig. 2a, Supplementary Movie 3). At low flow rate, we observed a period of 55$\pm$6 minutes and an amplitude of 30$\pm$9 GFP AU. Interestingly, the waveforms were distint, with the slower oscillator

reaching a trough near zero after activation and the faster oscillator decaying to levels above the original baseline (Fig. 2b). We swept the flow rate from 180-280 μm/min and observed an increasing oscillatory period from 55-90 minutes (Fig. 2c). In addition, we found the amplitude to be proportional to the period of the oscillations (Fig. 2d), which is consistent with the "degrade and fire" type oscillations (Mather et al., 2009) observed in a previously reported intracellular oscillator (Stricker et al., 2008a).

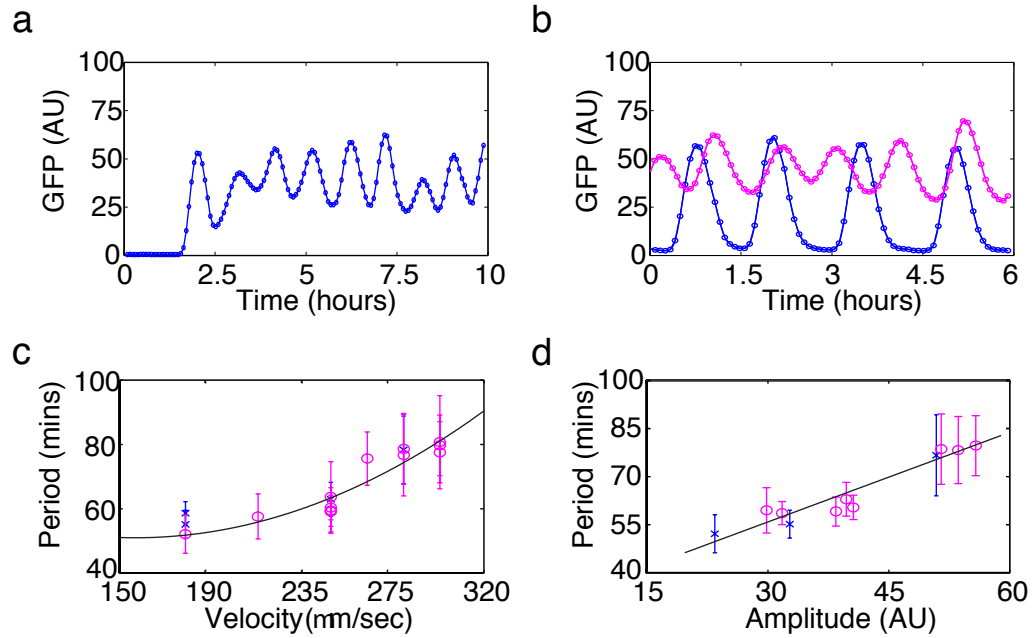


Figure 2.2: Dynamics of the synchronized oscillator under multiple microfluidic flow conditions (Supplementary Movies 2 and 3). **(a)** At around 90 minutes, cells begin to oscillate synchronously after reaching a critical density in the trap. **(b)** The period and amplitude increase for higher flow rates. Magenta curve is at low velocity(240μm/min), blue is at higher velocity(280μm/min). **(c)** Period as a function of velocity in the main channel showing tunability of period between 55-90 minutes. **(d)** Period vs. amplitude for all experiments. Magenta circles **(c,d)** are data from 84 and 90μm traps, blue crosses are 100μm traps.

In experiments conducted at low flow rate, we observed the spatial propagation of the fluorescence signal across the 100 μm chamber. In order to investigate these spatiotemporal dynamics in more detail, we redesigned the microfluidic chip with an

extended $2mm$ trapping chamber (Supplementary Information). Snapshots of a typical experimental run are presented in Fig. 3a (Supplementary Movies 4 and 5). A few isolated colonies begin to grow and subsequently merge into a large monolayer that fills the chamber (Fig. 3a: 66 minutes). At 100 minutes, there is a localized burst of fluorescence that propagates to the left and right in subsequent frames (Fig. 3a: 100-118 minutes). A second burst occurs near the original location and begins to propagate to the left and right as before.

Spatiotemporal Dynamics

To illustrate the spatiotemporal information contained in an entire 460-minute image sequence, we plot the fluorescence intensity as a function of chamber distance and time (Fig. 3b). Note the correspondence of this space-time plot to the images in Fig. 3a. During the first 100 minutes, there is no activity and the space time plot is blue, indicating no fluorescence. Then at 100 minutes, there is an orange spot at around 1350 μm, corresponding to the burst in Fig. 3a. In the space-time plot, propagation of a wave to the left and right appears as an green-yellow line with positive concavity. The larger slope to the left of the burst-origin indicates that the leftward moving wave is traveling slower ($\sim$25μm/min) than the rightward wave ($\sim$35μm/min). Subsequent waves originating from a nearby location arise as additional orange-yellow intensity lines. The second and 3rd intensity lines indicates an "annihilation event", where a leftward moving and rightward moving wave collide and annihilate each other. While these events are striking in the movies (Supplementary Movies 4 and 5), they appear subtly in the space-time plot at locations where positive and negative concavity meet(300-400 μm in 2nd intensity line and on). As the traveling wave gets further from a burst location it breaks off into a packet (170 minutes) which travels leftward at 12.5 μm/min initially, and slows to 8.5μm/min towards the end of the trap where the cell density is lower (between 118-200 minutes). The corresponding cell-density space-time plot shows that a higher density of cells is first reached at the center of the colony and is minimal towards

the left-moving edge (Fig. 3c and Supplementary Movie 4). As a result, the critical cell and AHL densities for wave propagation are reached at different times and spatial locations.

We also investigated how the intercellular oscillator behaves in a three dimensional colony growing in a 400x1000x4.0 μm microfluidic chamber (Figs. 3d, 3e, and Supplementary Movie 6). In this device, the colony grows radially over the course of 180 minutes without fluorescing until it reaches a size of approximately 100 μm. At this time, a large fluorescence burst originates from the center of the colony, with a bright band near the center (Fig. 3d: 228 minutes). During this first burst (273 minutes), the bright band shows that cells at an intermediate cell density have a larger amplitude and longer period than cells near the front or in the interior. As the colony expands an additional 50-100 μm in diameter, a second burst of fluorescence occurs at a similar intermediate cell density. Subsequent oscillations are seen as the cell growth front propagates, while weak oscillations arise and quickly die inside the colony.

Quantitative modeling

In order to quantitatively describe the mechanisms driving bulk synchronization and wave propagation, we developed a computational model using delayed differential equations for protein and AHL concentrations (Supplementary Information). While conceptually the nature of oscillations is reminiscent of the degrade-and-fire oscillations observed in a dual delayed feedback circuit (Stricker et al., 2008a; Mather et al., 2009), an important difference is the coupling among genetic clocks in different cells through extracellular AHL. The modeling of this coupling, and the related cell density dependence, allowed us to explain most of the non-trivial phenomenology of the spatiotemporal quorum clock dynamics.

A broad range of model parameters lead to oscillations (Figs. 4a-d), though there is a distinct absence of oscillations at small and large cell densities for low to medium flow values (Fig. 4c). The qualitative nature of the oscillations can be explained us-

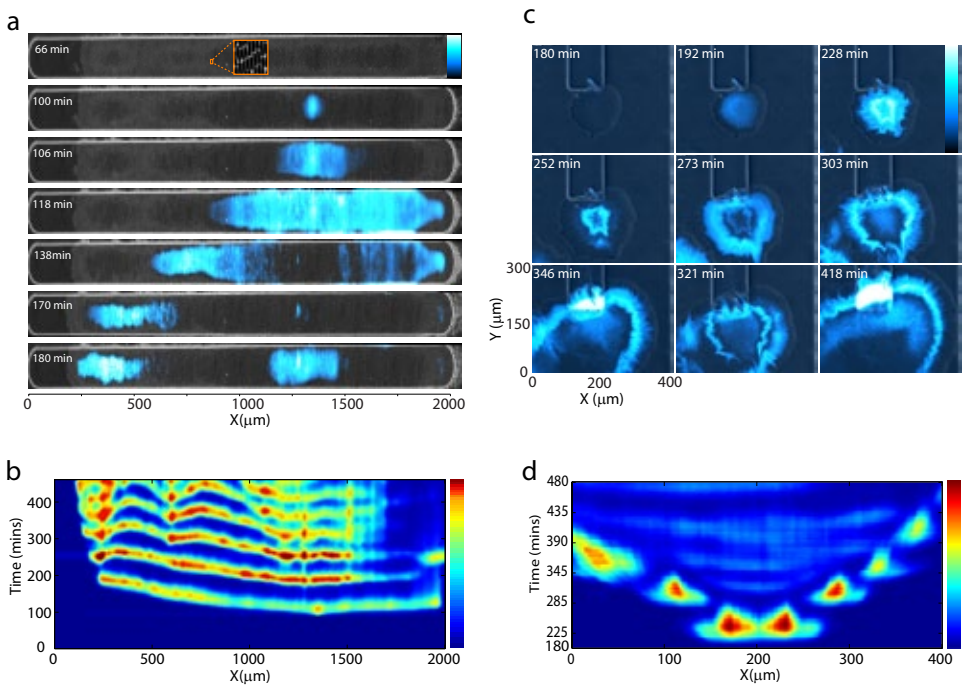


Figure 2.3: Spatiotemporal dynamics of the synchronized oscillators. **(a)** Snapshots of the GFP fluorescence superimposed over brightfield images of a densely packed monolayer of *E. coli* cells are shown at different times after loading (Supplementary Movies 4 and 5). Traveling waves emerge spontaneously in the middle of the colony and propagate outwards with the speed of $\sim$8-35μm/min. At later times waves partially lose coherence due to inhomogeneity in cell population and intrinsic instability of wave propagation (see Modeling Box). **(b)** Corresponding space-time diagram showing the fluorescence of cells along the center of the trap as a function of time. **(c)** Snapshots of the GFP fluorescence superimposed over the brightfield images of a three-dimensional growing colony of *E. coli* cells at different times after loading (Supplementary Movie 6). Bursts of fluorescence begin when the growing colony reaches a critical size of about 100μm. These bursts are primarily localized at the periphery of the growing colony. **(d)** Corresponding space-time diagram showing fluorescence of cells along a horizontal line through the center of the growing colony.

ing Fig. 4a. Each period begins with the latent accumulation of both AiiA and LuxI, which after a delay, burst rapidly to high values. That burst suppresses AHL and further production of AiiA and LuxI which then decay enzymatically, after which the process repeats. As expected, the period of the oscillations is roughly proportional to the enzymatic protein decay time. The period grows with the external AHL flow rate (effective degradation) and the amplitude of the oscillations, in good agreement with the experiments (compare Fig. 4b with Figs. 3c and d).

We modeled the collective spatiotemporal dynamics of the clocks by generalizing the bulk model to include the coupling of individual oscillators through extracellular AHL. The model consists of a one-dimensional array of "cells", each of which is described by the same set of delay-differential equations coupled to a common, spatially nonuniform field of extracellular AHL. The latter is described by a linear diffusion equation with sources and sinks due to AHL diffusion through the cell membrane and dilution. A small AHL perturbation in the middle of the array, initiates waves of LuxI concentration (Fig. 4c), in excellent agreement with the experimental findings (compare Figs. 3b and 4c). The velocity of the front propagation depends on the external AHL diffusion coefficient D_1 (Fig. 4f and Supplementary Information), and for experimentally relevant values of D_1, the simulated front velocity is in good agreement with experimental data. In addition, cell density plays an important role in wave propagation. In order to model the evolution of the three dimensional colony (Figs. 3c and 3d), we set the functional form of the cell density to be an expanding "Mexican hat", as observed in the experiments. Oscillations are then suppressed by the high density of cells in the middle of the colony, and LuxI bursts only occur on the periphery of the growing colony of cells. This phenomenology is also in excellent agreement with our experimental findings (compare Figs. 4d and 3d).

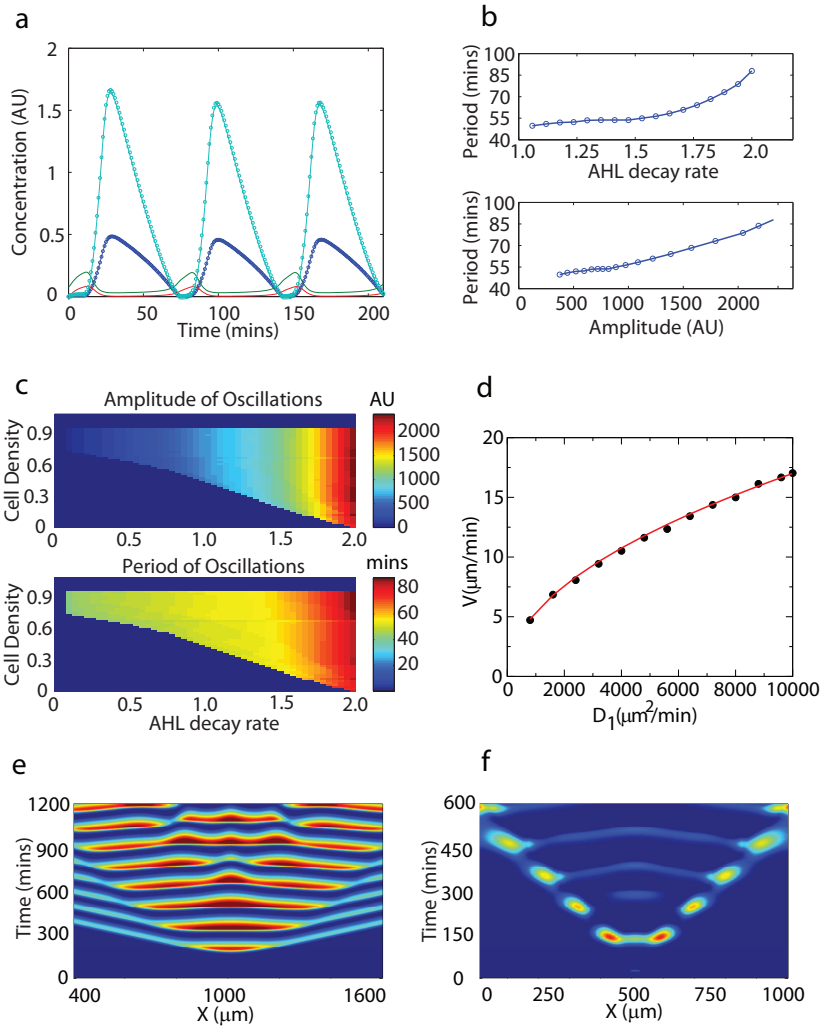


Figure 2.4: Modeling of synchronized genetic clocks. **(a)** A typical time series of concentrations of LuxI (cyan circles), AiiA (blue circles), internal AHL (green line), and external AHL (red line). LuxI and AiiA closely track each other, and are anti-phase with the concentrations of external and internal AHL. **(b)** Period of oscillations as a function of the flow rate μ at cell density $d = 0.5$ (top panel). Period as a function of the amplitude of oscillations for the same cell density (bottom panel). **(c)** Period and amplitude as a function of cell density and AHL decay rate μ. Oscillations occur over a finite range of cell densities, and period increases with μ after the bifurcation line is crossed. The results in **(c)** and **(d)** compare favorably with the experimental results in Figs. 2c and 2d. **(d)** Speed of wave front propagation as a function of the diffusion coefficient D_1. The numerical data scale as $V \sim D_1^{1/2}$ (red line). **(e)** Space-time diagram of traveling waves propagating through a uniform array of cells corresponding to the experiment depicted in Figs. 3a and 3b. **(f)** Space-time diagram of bursting oscillations in a growing cell population corresponding to the experiments in Figs. 3c and 3d.

Emergence

On a fundamental level, the synchronized oscillations represent an emergent property of the colony that can be mechanistically explained in terms of the clock design. Oscillations arise because the small molecule AHL plays a dual role, both enabling activation of the genes necessary for intracellular oscillations and mediating the coupling between cells. Since unbounded growth of the colony leads to an accumulation of AHL that ultimately quenches the bulk oscillations, we used open-flow microfluidic devices to allow for the flow of AHL away from the colony. At low cell densities, oscillations do not occur because intracellular gene activation is decreased as AHL diffuses across the cell membrane and out of the chamber. At intermediate cell densities (i.e. a full chamber), the increased production of AHL in each cell acts to mitigate the outward flow such that activation of the genes can occur in a rhythmic fashion, and colony-wide oscillations emerge in a seemingly spontaneous fashion.

A natural question arises regarding the behavior of individual cells in the absence of coupling. While experimentally we cannot turn off the coupling while maintaining intracellular gene activation, we addressed this question using simulations by artificially setting the AHL diffusion rate across the membrane to zero (with the other parameters fixed). We find that individual cells oscillate independently for any cell density since they are completely decoupled from the environment and each other. This result indicates that the coupling through AHL diffusion provides a means for the synchronization of individual oscillators at intermediate cellular concentrations.

Perspective and outlook

In the mid seventeenth century, Chirstiaan Huygens serendipitously observed that two pendulum clocks oscillated in synchrony when mounted to a common support beam (Bennett et al., 2002). While observations of synchronization in nature surely predate the age of enlightenment, Huygens is credited as the first to systematically char-

acterize the synchronization of oscillators in terms of a known coupling mechanism (which, in the case of the pendulums, he deduced as vibrations in the common support). We have shown how quorum sensing can be used to couple genetic clocks, leading to synchronized oscillations at the colony level. Given the single-cell variability and intrinsic stochasticity of most synthetic gene networks (Ozbudak et al., 2002; Elowitz et al., 2002; Atkinson et al., 2003; Stricker et al., 2008a; Austin et al., 2006), the use of quorum sensing is a promising approach to increasing the sensitivity and robustness of the dynamic response to external signals. Along these lines, our results set the stage for the design of networks that can function as spatially distributed sensors or synthetic machinery for coupling complex dynamical processes across a multicellular population.

Methods

Strains, growth conditions

Three identical transcriptional cassettes for *luxI*, *aiiA*, and *yemGFP* were constructed by replacing a modular pZ plasmid's promoter (Lutz and Bujard, 1997a)(with *yemGFP*) with the lux operon from the native *Vibrio Fischeri* operon *luxR* up to *luxI* stop codon (Dunlap and Greenberg, 1985). *LuxI* and *aiiA* (Thomas et al., 2005) genes were cloned in place of *yemGFP* and a degradation tag was added to the carboxy-terminal of each. A previously used MG1655 strain of Escherichia coli[1] was transformed with plasmids pTD103luxI/GFP which is (colE1,Kan) and pTD103aiiA which is (p15A,Amp) to create strain TDQS1 (Suppl. Info).

Each experiment started with a 1:1000 dilution of overnight culture grown in 50mL LB (10g/L NaCl) with antibiotics 100μg/ml ampicillin and 50μg/ml kanamycin for approximately 2 hours. Cells reached an OD600 of 0.05-0.1 and were spun down and concentrated in 5mL of fresh media with surfactant concentration of 0.075 Tween20 [Sigma-Aldrich, St.Louis,MO] before loading in a device.

Microfluidics and Microscopy

Images were acquired using an epifluorescent inverted microscope (TE2000-U, Nikon Instruments Inc., Tokyo, Japan), and chip temperatures were maintained at 37°C with a plexiglass incubation chamber encompassing the entire microscope. Phase contrast and fluorescent images were taken at 20x or 60x every 2-5 minutes and focus was maintained automatically using Nikon Elements software.

Plasmid Construction

The pTD103 plasmids were constructed by replacing the promoter in a pZ modular plasmid (pZE21yemGFP-LAA) from XhoI to EcoRI restriction sites((Lutz and Bujard, 1997a)) with the *luxR* gene and the *luxI* promoter amplified via PCR from the native *Vibrio Fischeri* operon(pJE202, (Dunlap and Greenberg, 1985)). The pZ plasmid RBS was kept the same, and *luxI* or *aiiA*(from pMAL-t-aiiA, (Thomas et al., 2005)) genes were cloned in place of yemGFP with the TSAANDENYALAA degradation tag on the carboxy-terminal((Stricker et al., 2008a)). The yemGFP reporter module (luxR gene-luxIp-yemGFP-LAA) was then amplified with AvrII and NheI restriction sites and ligated into the AvrII site following the terminator in pTD103luxI-LAA.

Data Analysis

Fluorescence vs. Time curves were obtained by importing fluorescent images into ImageJ and using the 'Intensity vs. Time Monitor' Plug-in to obtain a mean gray value of the entire field of view, and then the background gray value was subtracted (Fig1c 60x magnification, Fig 2a,b 20x magnification). Peak-to-peak values were taken for all period measurements and amplitudes were measured as peak to previous trough values. The data collected in Fig2c,d was obtained from 20x/60x magnification experi-

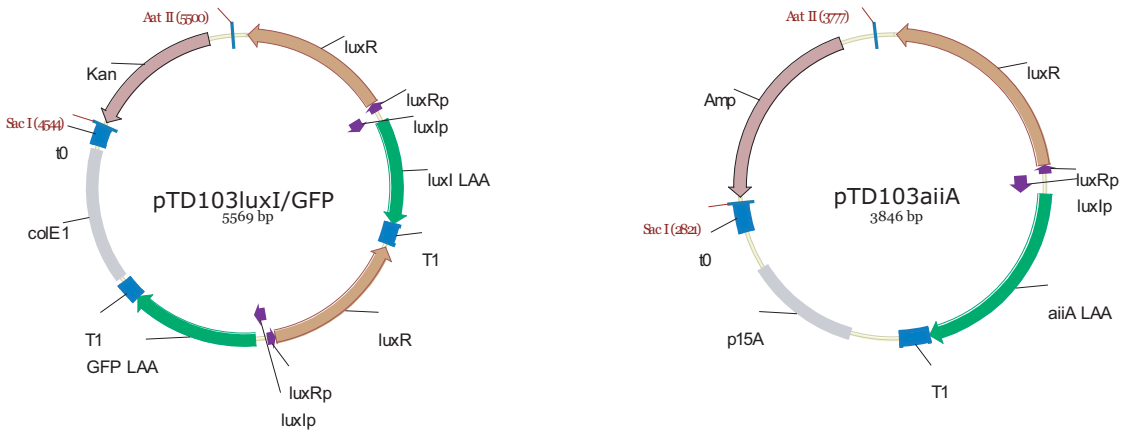


Figure 2.5: Plasmids for the synchronized oscillator strain TDQS1. Construction of the pTD103 plasmids was done in the modular pZ plasmid backbones in three identical transcriptional modules with the same promoter, RBS, and terminator for each.

ments from the parallelized device (Supplement Fig. 3b) in different sized traps. Each data point in Fig. 2c,d represents between 10-40 peak values averaged. We found that traps downstream of each other had simlilar period/amplitude measurements and including them in our averages did not significantly alter the mean values but greatly reduced the errors bar values. This showed that traps downstream of one another were only weakly coupled at our flow rates. In Suppl Fig2, we plot an additional fluorescence trajectory obtained from imaging one of these traps at 60x showing that oscillations exhibit stably over long periods of time.

Space-Time plots

To create the space-time plot in Fig 3b, we averaged a 20 pixel window along the center of the trap (seen in Fig3a) in fluorescent images. To obtain a semi-quantitative measure of cell density we performed the same process on brightfield images. When no cells were present, the mean gray value was darker due to the lighting on the PDMS (polydimethylsiloxane) device, so we manually corrected the blue region in the bottom left of Suppl. Fig3 where no cells were present. Once cells populated the trap, we found

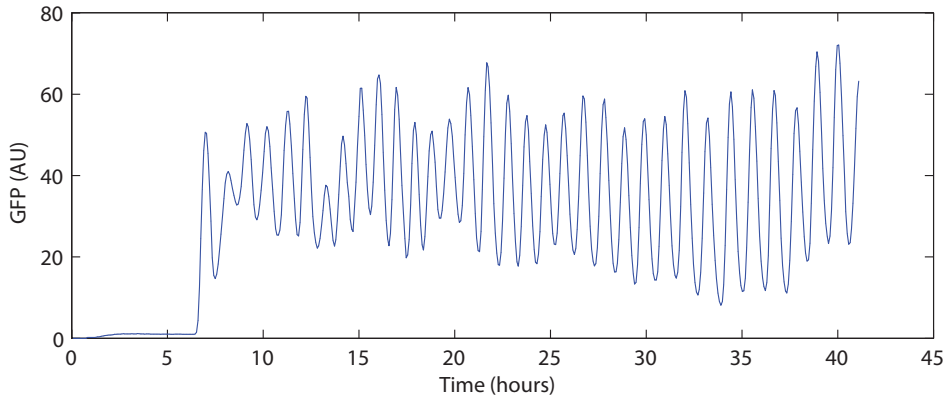


Figure 2.6: Stable oscillations in microfluidic device. Fluorescence vs. Time curve obtained for a 100x84 micron trap over the course of $\sim$ 40 hours.

the gray value to give a measure of the density cells (Suppl. Fig3). The periodicity in the data (apparent at high time values) is an artifact from the stitching of images in the Nikon Elements software (due to the slight difference in focal planes when the camera moves). We obtained the space time plot for Fig3d by averaging the fluorescence (20 pixel window) along the center of the colony. We stitched together 3 continuous image sets with image frequencies of 4 minutes (1-45), 3 minutes (frames 46-99) and 2.25 minutes (frames greater than 100). In the displayed images, another colony growing from bottom left begins to merge with the main one, and slightly influences the fluorescent front on the left but did not affect the front to the right.

Microscopy and Microfluidics

A similar microscope setup was used in (Stricker et al., 2008a), but to maintain temperature at 37°C a plexiglass incubation temperature was used. At 60x, fluorescent images were taken every 3-4.5 minutes which we found to be sufficient to prevent photobleaching(200-500ms exposure,10% lamp setting). At 20x magnification, fluorescent images could be taken more often (every 2 minutes).

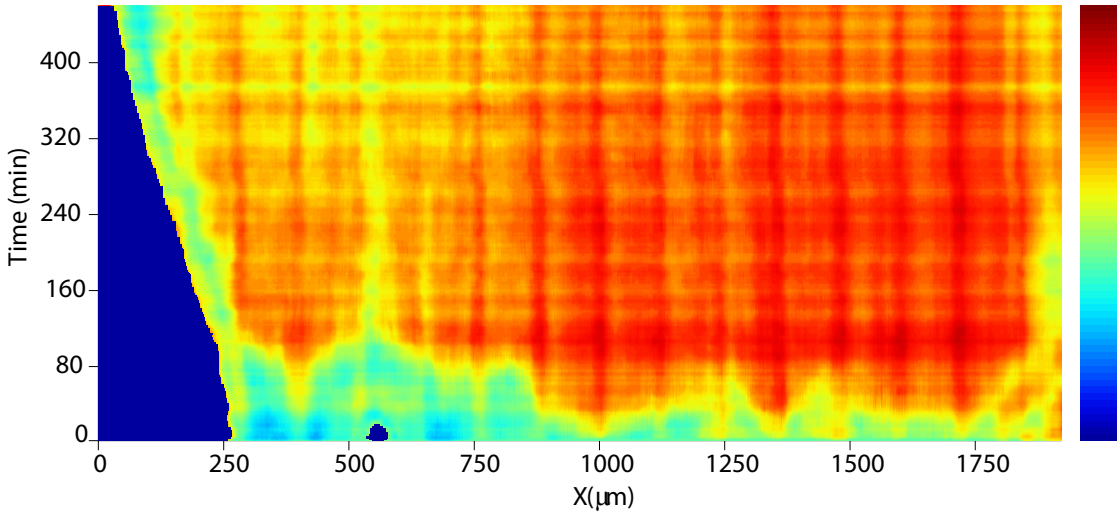


Figure 2.7: Space-time plot of density of cells in Fig3a experiment. Gray value of the brightfield images is plotted as a measure of cell density in the 2000x 100x 0.95μm device. Red indicates higher cell density.

In each device, *E. coli* cells are loaded from the cell port while keeping the media port at sufficiently higher pressure than the waste port below to prevent contamination (Suppl. Fig4). Cells were loaded into the cell traps by manually applying pressure pulses to the lines to induce a momentary flow change. The flow was then reversed and allowed for cells to receive fresh media with 0.075% Tween20 which prevented cells from adhering to the main channels and waste ports. For the yeast device experiments (Fig3c,d), we loaded *E. coli* cells in the main region by not adding surfactant during the loading process.

We initially built the device in Suppl.Fig 4a to test the synchronized oscillator in three different trap sizes with 1.65μm high trap regions. We found the 1.65μm height allowed for better growth of cells presumably because of the additional flow of media into the interstitial spaces (as compared to 0.9-1.0μm high traps). A parallelized version of the chip with 3 channel heights was constructed to generate different flow rates and trap sizes of (70,84,90,100)x100 μm(Suppl. Fig3b). We found that the heights of the

channels did not greatly affect the period measurements presumably since the relevant parameter is only the flow rate in the plane of the trap, and this did not significantly differ between channel heights. Thus, to alter the flow rates we increased the heights of the media reservoir to create different flow rates. To estimate flow rates, we measured the length of traces of fluorescent beads (1.0 μm) upon 100ms exposure to fluorescent light to establish a measure for the average velocity of as a function of height of the media reservoir. We averaged over at least 1000 data points for each to obtain the average velocities in Fig.2c (x-axis), which confirmed that the velocity scales linearly with the pressure difference caused by the height of the media reservoir.

To study spatial temporal behavior of the synchronized oscillator, we designed a microfluidic trap that is 20 times as long (2mm) and 100 μm wide as the original traps (Suppl. Fig4c). Unlike the traps in Suppl. Fig 4*a,b*, the trap is only 0.95 μm high and we found this height optimal for seeding cells in the trap. Since the trap lacks any walls it is open to the flow, it would be difficult to seed cells in a non constraining device. Given the open boundary conditions and the constriction of rod-shaped *E. coli* bacteria to one layer, cells arrange parallel to each other and perpendicular to the edges of the trap. This ordering leads to a very tight packing of a monolayer of cells. Under these conditions, the transport of nutrients, AHL and cell waste happen mainly by diffusion and is less sensitive (but not insensitive) to the flow rate of surrounding media than in the cell traps of devices *a* and *b* below. An example of this is that bursts of fluorescence propagate in both directions of the trap irrespective of the sense of external flow at very high flow rates. In the experiment shown in Suppl. Movie 3, the flow rate was set close to 100 μm/s to counter the increased adherence of cells after long run times, which we believe might be caused by growing them in media with surfactant Tween 20 after long durations.

Modeling

There has been much work on modeling asynchronous, oscillating cells coming into synchrony in the context of synthetic biology ((McMillen et al., 2002; Garcia-Ojalvo et al., 2004)), though less attention has been focused on gene networks that do not oscillate in individual cells but oscillate collectively ((Ma and Yoshikawa, 2009)). Here we constructed a deterministic model of quorum-sensing gene clock. From the biochemical reactions depicted in Fig. 1a, we derived the following set of delay-differential equation model for intracellular concentrations of LuxI (I), AiiA (A), internal AHL (H_i), and external AHL (H_e),

$$\frac{\partial A}{\partial t} = C_A[1-(d/d_0)^4]\,P(\alpha,\tau) - \frac{\gamma_A A}{1+f(A+I)} \tag{2.1}$$

$$\frac{\partial I}{\partial t} = C_I[1-(d/d_0)^4]\,P(\alpha,\tau) - \frac{\gamma_I I}{1+f(A+I)} \tag{2.2}$$

$$\frac{\partial H_i}{\partial t} = \frac{bI}{1+kI} - \frac{\gamma_H A H_i}{1+gA} + D(H_e - H_i) \tag{2.3}$$

$$\frac{\partial H_e}{\partial t} = -\frac{d}{1-d}D(H_e - H_i) - \mu H_e + D_1\frac{\partial^2 H_e}{\partial x^2} \tag{2.4}$$

We did not include an equation for LuxR assuming that it is constitutively produced at a constant level. Previous work found that LuxR is under control of the LuxR-AHL complex to produce a higher concentration of LuxR but we did not find this necessary to capture the essential behavior of the synchronized oscillator((Williams et al., 2008)).

In the first two equations, the Hill function

$$P(\alpha,\tau) = \frac{\delta + \alpha H_\tau^2}{1+k_1 H_\tau^2}$$

describes the delayed production of corresponding proteins, it depends on the past concentration of the internal AHL, $H_\tau(t) = H_i(t-\tau)$. These delayed reactions mimic the complex cascades of processes (transcription, translation, maturation, etc.) leading to formation of functional proteins. The pre-factor $[1-(d/d_0)^4]$ describes slowing down of protein synthesis at high cell density d due to lower nutrient supply and high waste

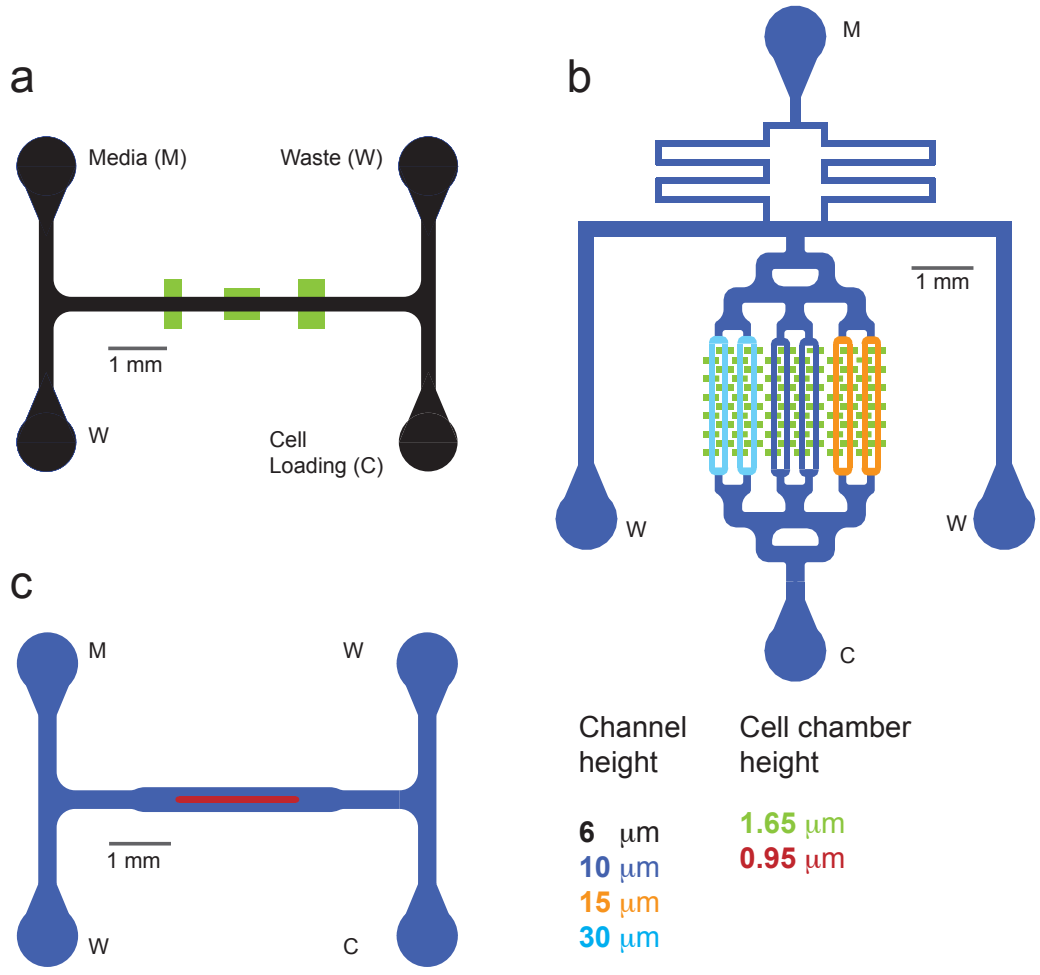


Figure 2.8: Microfluidic Devices constructed for this study. *a)* Device used initially to test the TDSQ1 cells for synchronized oscillations. The dimensions of the traps from left to right are 100x100 μm , 200x50 μm and 150x100 μm, respectively. Traps scaled 300 % in this schematic for visualization. *b)* Parallelized version of Device *a*. Several trap sizes and channel heights could be tested simultaneously. Traps are 100 μm wide and either 70,84,90,or 100 μm deep. *c)* Device used for the wave propagation experiments in Fig3a,b in the main text. The trap is 2000x100 μm wide.

concentration. Terms proportional to γ_x describe enzymatic degradation of proteins and AHL by proteases inside of the cell due to their degradation tags. We model these processes using Michaelis-Menten kinetics. Terms proportional to D describe diffusion of AHL through cell membrane, and the term proportional to μ models dilution of external AHL by external fluid flow. The cell density (d) determines the amount of external AHL and thus affects the AHL decay rate. The factor $d/(1-d)$ follows from the total mass conservation of AHL inside and outside the cells. Since the flow speed ($\sim 100\mu$m/sec) is much higher than the typical wave propagation speed ($\sim 10\mu$m/sec), we neglected the anisotropy imposed by the fluid flow. The last term in equation for H_e describes the diffusion of external AHL.

We used the following experimentally relevant scaled parameters in most of our simulations: $C_A = 1, C_I = 4, \delta = 10^{-3}, \alpha = 2500, \tau = 10, k = 1, k_1 = 0.1, b = 0.06, \gamma_A = 15, \gamma_I = 24, \gamma_H = 0.01, f = 0.3, g = 0.01, d_0 = 0.88, D = 2.5$. We varied the diffusion constant D_1 and the external AHL decay rate (flow rate) μ, as well as the cell density d. For "bulk" simulations we dropped the diffusion term $\sim D_1$ in equation for H_e, and solved the resulting set of ordinary delay-differential equations. For spatio-temporal simulations we replaced the partial delay-differential equations by a one-dimensional array of $N = 200$ systems of ordinary delay-differential equations representing individual "cells" coupled via a second-order discrete diffusion operator $D_1 dx^{-2}[H_{i-1} + H_{i+1} - 2H_i]$ for the external AHL concentration. We used periodic boundary conditions at the ends of the array ($H_1 = H_N$).

In addition to the numerical results presented in the Main Text, we show here the results of additional spatiotemporal simulations. In particular, Suppl. Fig. 5 shows the synchronization of oscillations in cell population with statistically different parameters. As seen from the figure, the coherence of oscillations increases with the diffusion coefficient D_1, as expected. In Suppl. Fig. 6 we show the propagation of waves initiated by a localized initial condition ($I_{N/2} = 1$ while all other $I_i = 0$ and $A_i = 0$) for different diffusion constants. Since parameter δ characterizing the leakiness of the *luxI* promoter is small (10^{-3}), the basal state with $A = I = 0$ is very weakly unstable. Thus, in the

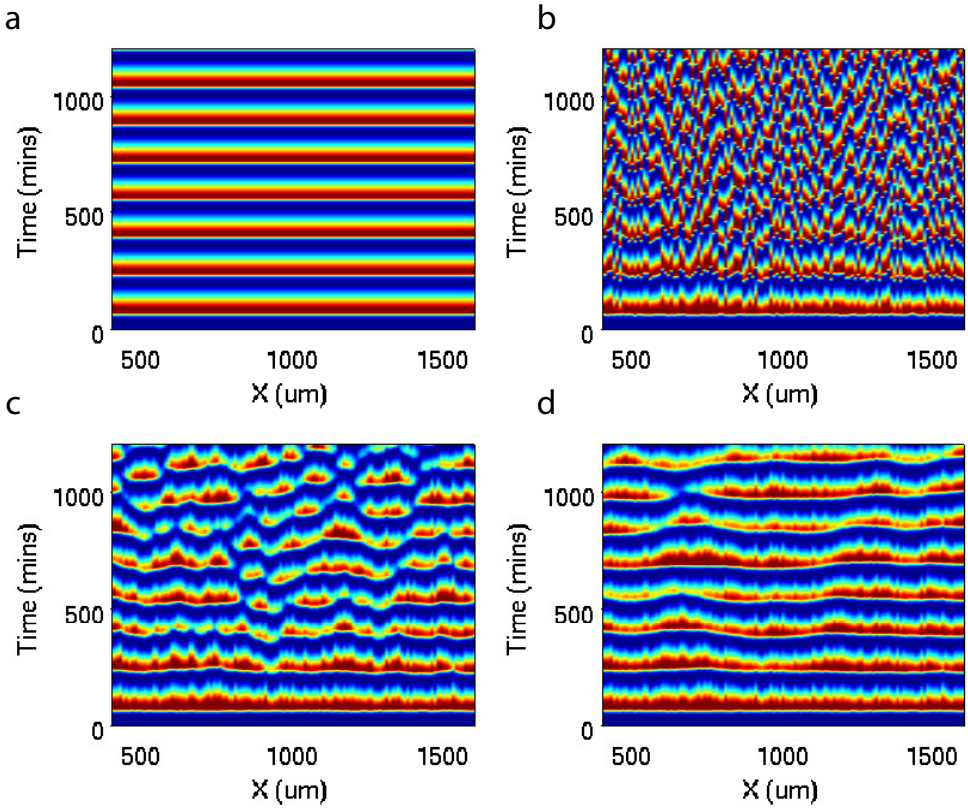


Figure 2.9: Synchronization of oscillations in spatially extended system with diffusion. The parameters (p) of each of 200 oscillators were varied around their nominal values (p_0) as $p = p_0(1+\eta\xi)$ where ξ is a random number uniformly distributed between -0.5 and 0.5, and η characterizes the fluctuations magnitude. To illustrate the role of spatial diffusion in mitigating the stochastic fluctuations, we varied η and D_1: a, $\eta = 0,\ D_1 = 0$, b, $\eta = 0.1,\ D_1 = 0$, c, $\eta = 0.1,\ D_1 = 800\mu\text{m}^2$/sec, d, $\eta = 0.1,\ D_1 = 4000\mu\text{m}^2$/sec,

absence of AHL diffusion ($D_1 = 0$), while the middle cell begins to oscillate immediately, all other cells are still quiescent (Suppl. Fig. 5a). However, when the diffusion is present ($D_1 \neq 0$), cells influence their neighbours and oscillations propagate in the form of traveling waves in both directions (Suppl. Fig . 6b-d). As seen from this set of space-time diagrams, the wave speed increases with D_1. Fig. 4d of the Main text shows that this dependence is well approximated by the formula $V \approx 0.17 d_1^{1/2} \mu$m/sec.

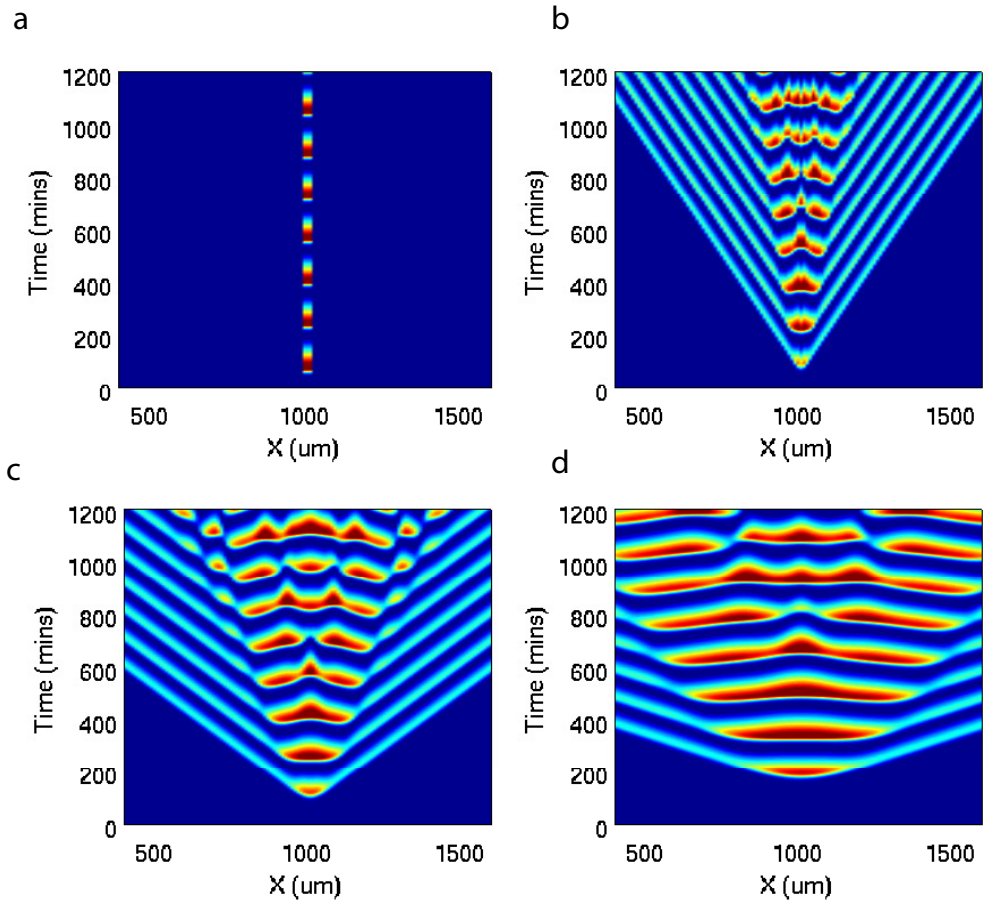


Figure 2.10: Wave propagation in the spatially uniform system with different external AHL diffusion rates: a, $D_1 = 0$, b, $D_1 = 200\mu$m^2/sec, c, $D_1 = 800\mu$m^2/sec, d, $D_1 = 4000\mu$m^2/sec,

Supplementary Movies

Supplementary information, including methods, supplementary figures and tables, and timelapse microscopy movies, is linked to the online version of the paper at www.nature.com/nature.

Supplementary Movie 1. Timelapse fluorescence microscopy of TDQS1 cells at low flow rate in a 100x100μm trap. Fluorescence is shown in cyan hot color table (dark blue low, white high). Total time of movie is 483 min with a sampling rate of one image every 3 min.

Supplementary Movie 2. Timelapse fluorescence microscopy of TDQS1 cells in a 2000 by 100 by 0.95μm open trap showing propagation of AHL at millimeter scale. The brightfield image is shown in gray, and fluorescence is shown in cyan hot color table (dark blue low, white high). Total time of movie is 962 min with a sampling rate of one image every 3 min.

Supplementary Movie 3. Timelapse microscopy of TDQS1 cells at high flow rate in a 100x100μm trap. Fluorescence is shown in cyan hot color table (dark blue low, white high). Total time of movie is 867 min with a sampling rate of one image every 3 min.

Supplementary Movie 4. Zoomed timelapse fluorescence microscopy of TDQS1 cells in a 2000 by 100 by 0.95μm open trap showing close-up of cells and propagation of AHL. The brightfield image is shown in gray, and fluorescence is shown in cyan hot color table (dark blue low, white high). Total time of movie is 962 min with a sampling rate of one image every 3 min.

Supplementary Movie 5. Timelapse fluorescence microscopy of TDQS1 cells in a three dimensional 1000x400x4.0μm trap. The brightfield image is shown in gray, and fluorescence is shown in cyan hot color table (dark blue low, white high). Total time of movie is 636 min with a sampling rate of one image every 2.25-4 min.

Supplementary Movie 6. Simulation of the wave propagation within a uniform population of cells. The oscillations are initiated by a small perturbation in the middle of the chamber. The space-time diagram corresponding to this simulation is shown in Fig. 4e of the Main text.

Supplementary Movie 7. Simulation of the wave propagation within a growing dense cluster of cells. The oscillations are initiated by a small perturbation in the middle of the initially small cluster. The space-time diagram corresponding to this simulation is shown in Fig. 4f of the Main text.

Acknowledgements

We thank Matt Bennett, Kurt Wiesenfeld, and Jim Collins for stimulating discussions during the preparation of the manuscript. This work was supported by the National Institutes of Health and General Medicine (GM69811) and the DOE CSGF fellowship (TD). Chapter 2 contains material originally published as Danino, T.*, Mondragron-Palomino, O.*, Tsimring, L., and Hasty, J. ***Nature*** (2010): A synchronized quorum of genetic clocks. (*equal contribution). Copyright permission to republish here was granted by Nature publishing group.

Chapter 3

Modeling Synthetic Gene Networks

Steady state modeling of genetic networks

In this section, we will go through derivation of a mathematical model for both positive and negative feedbacks. These two types of feedbacks are important in Chapters 2 and 4. First, we model a simple gene network that reaches a steady state and show how approximation methods can be used to get an accurate model for the dynamics of monomers.

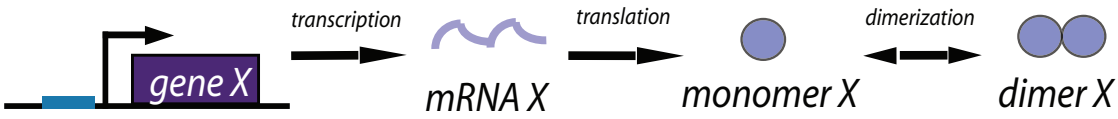


Figure 3.1: Transcription of gene x produces an mRNA , which is translated into a monomeric protein x, then dimerization of protein x produces the biologically active dimer.

We can summarize these processes in the table below.

Using mass-action kinetics, we can write the equations that describe the average concentration of molecules as a function of time. The system can be written as: (Note that the ”2”’s in the first equation are a result of the stoichiometry of x and y)

	reaction		
1	$x + x$	$\xrightarrow{\kappa_+}$	y
1	y	$\xrightarrow{\kappa_-}$	x
2	d_o	$\xrightarrow{\alpha}$	$d_o + m$
2	m	$\xrightarrow{\sigma}$	$m + x$
3	x	$\xrightarrow{\gamma_x}$	$\oslash$
3	m	$\xrightarrow{\gamma_m}$	$\oslash$

$$\dot{x} = 2\kappa_- y - 2\kappa_+ x^2 + \sigma m - \gamma_x x \tag{3.1}$$

$$\dot{y} = \kappa_+ x^2 - \kappa_- y \tag{3.2}$$

$$\dot{m} = \alpha d_o - \gamma_m m \tag{3.3}$$

Since the system is nonlinear and 3 dimensional, analysis can be difficult. To simplify the system, a difference in the time scale of the reactions can be used to reduce the dimensions of the system. Dimerization reactions are typically fast compared to transcription and translation reactions, thus we can make use of this assumption in reducing these equations. This commonly used method is called the quasi-steady state approximation (QSSA), and assumes that the dimerization reaction is fast, so that the variable y can be substituted by kx^2, where k is the ratio of forward and backward reaction rates. The simplified system reduces down to 2 equations:

$$\dot{x} = \sigma m - \gamma_x x \tag{3.4}$$

$$\dot{m} = \alpha d_o - \gamma_m m \tag{3.5}$$

In the context of many gene networks in nature, the mRNA transcription and degradation rate may be slower than protein dimerization as well in which case we can

make the additional assumption that mRNA dynamics approach equilibrium faster than protein dynamics. In this case substitution for the variable *m* above results a single equation (with $A = \frac{d_o \sigma \alpha}{\gamma_m}$) :

$$\dot{x} \quad = \quad Am - \gamma_x x \tag{3.6}$$

In many synthetic gene networks however, proteins are often tagged for faster degradation which is on a similar time scale to that of mRNA degradation. In those cases we must be more careful to make this approximation.

Positive feedback gene network

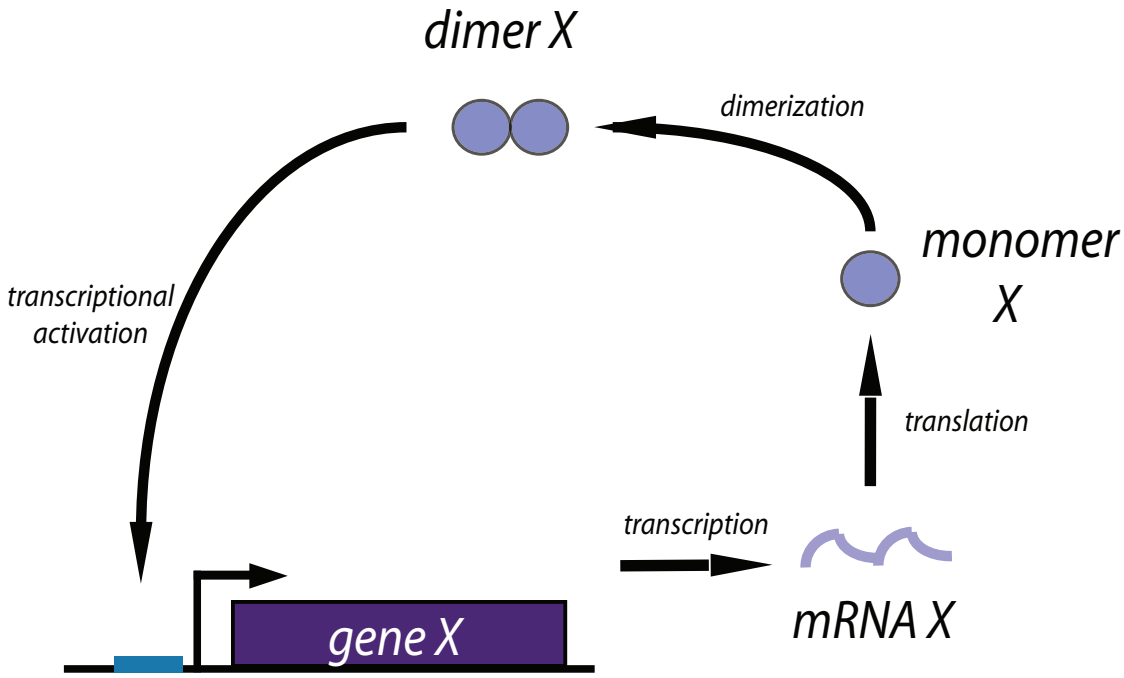


Figure 3.2: Positive feedback gene network. Transcription, translation, and dimerization of gene x produces a functional dimer which positively activates transcription of itself.

In Figure 1, a diagram of a simple gene network with positive feedback is shown. The processes which occur are transcription of gene x into an mRNA, translation into monomeric protein x, reversible binding of protein x into a functional dimer, and transcriptional activation of the promoter by dimeric x. We will assume that mRNAs decay by a linear process and only monomeric protein x decays. To begin modeling these processes we write out the set of mass-action kinetic reactions for these processes. These are:

Using mass-action kinetics, we can write down the full system of differential equations:

	reaction		
1	$x + x$	$\xrightarrow{\kappa_+}$	y
2	y	$\xrightarrow{\kappa_-}$	$x + x$
3	$y + d_o$	$\xrightarrow{k_+}$	d_r
4	d_r	$\xrightarrow{k_-}$	$y + d_o$
5	d_o	$\xrightarrow{\alpha}$	$d_o + m$
6	d_r	$\xrightarrow{\beta}$	$d_r + m$
7	m	$\xrightarrow{\sigma}$	$m + x$
8	x	$\xrightarrow{\gamma_x}$	$\oslash$
9	m	$\xrightarrow{\gamma_m}$	$\oslash$

Full Equations

$$\dot{x} = 2\kappa_- y - 2\kappa_+ x^2 + \sigma m - \gamma_x x \tag{3.7}$$

$$\dot{y} = \kappa_+ x^2 - \kappa_- y + k_- d_r - k_+ d_o y \tag{3.8}$$

$$\dot{d_o} = k_- d_r - k_+ y d_o \tag{3.9}$$

$$\dot{d_r} = k_+ y d_o - k_- d_r \tag{3.10}$$

$$\dot{m} = \alpha d_o + \beta d_r - \gamma_m \tag{3.11}$$

$$\tag{3.12}$$

where d_o is the concentration of promoter sites that are free of the dimer; d_r is the concenctration of promoter sites with bound dimer; m is the concentration of mRNA molecules.

As before, we know that reactions 1-4 typically occur at a much faster time scale than reactions 5-9. Using this information we arrive at :

$$y = c_p x^2; \tag{3.13}$$

$$d_o = d(1 + c_p c_d x^2)^{-1} \tag{3.14}$$

$$d_r = d c_p c_d x^2 (1 + c_p c_d x^2)^{-1} \tag{3.15}$$

where $c_p = \kappa_+/\kappa_-$ and $c_d = k_+/k_-$.

After substitution of these equations into the system, and using the fact that $d = d_o + d_r$ we arrive at the system:

$$\dot{x} = \sigma m - \gamma_x x \tag{3.16}$$

$$\dot{m} = \frac{d}{1 + c_p c_d x^2}[\alpha + \beta c_p c_d x^2] - \gamma_m m \tag{3.17}$$

$$\tag{3.18}$$

For the purposes of showing how to analyze a simple system, we will reduce the system of 2 equations above by assuming that mRNA dynamics come to steady state considerably faster than the protein dynamics. We can think of this scenario as a situation where we model only the dimers in the system. We write this below as a single equation with constants renamed.

$$\dot{x} = \frac{(1 + ax^2)}{1 + x^2} - gx \tag{3.19}$$

Steady states

Now lets examine the steady states of this system. Setting $\dot{x}$=0, we get the cubic equation $gx^3 - ax^2 + gx - 1 = 0$, which results in the 3 roots. Now depending on

the values of a and g, this will yield different types of solutions. For instance, if we look at a=50, and solve for x as a function of g, we get the following curve seen in Figure 3. Notice that for a particular value of g, x can have 1 or 3 positive steady states. For instance, at g=20, x can take on the values 0.1,0.5, or 2. The initial condition of x determines which steady state the system falls upon. For instance if x is initially ¿ 2 (stable branch), then the system will relax onto the higher steady state x=2, and if x¡0.1, than the system relaxes to the lower steady state x=0.1. In between these values is the unstable branch and the system will settle to the stable branch closest. Thus 2 of these steady-states are stable and 1 is unstable. In the next section, we will explore how to determine the stability of these points.

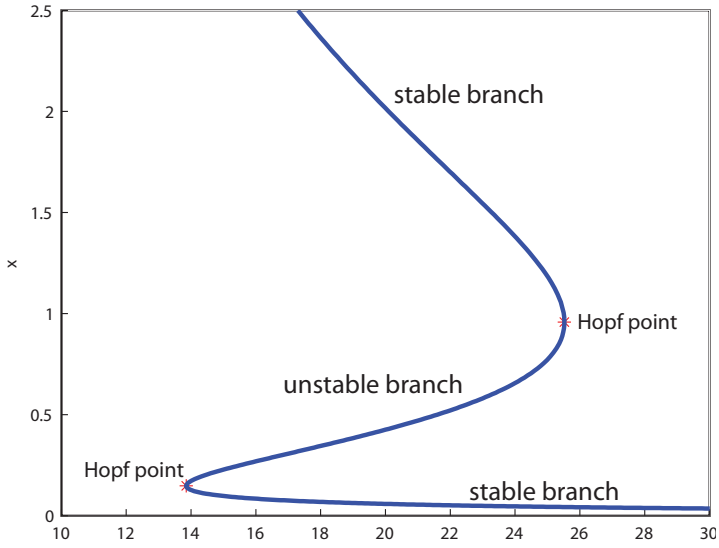


Figure 3.3: *a)* Steady state value of x as a function of g. Shows how 1 or 3 steady states can occur depending on the value of x.

Linear stability

A linear stability analysis can be used on this system to determine which points are bistable or not. The idea behind this technique is that a system near its steady state responds linearly to small perturbations, so we can linearize the system about a steady state and analyze if a perturbation grows away from the steady state or back towards it.

The system we are considering is

$$F(x) = \dot{x} = \frac{(1+ax^2)}{1+x^2} - gx \tag{3.20}$$

$$\tag{3.21}$$

Consider the Taylor expansion of the system below around a steady state x^*:

$$F(x) = F(x*) + F'(x*) * (x - x^*) + F''(x*)(x - x^*)^2 + ... \tag{3.22}$$

where we consider only the first linear term in x as an approximation to the system

$$F(x) \approx F(x*) + F'(x*) * (x - x^*) \tag{3.23}$$

Now consider a time dependent small perturbation $\eta(t)$, such that $x = x^*$ + $\eta(t)$. From this equation we also see that $x' = \eta^{'}(t)$. Substituting this into equation for approx we can show that $\eta'(t) = F'(x^*)\eta(t)$, and integrating this equation, we see that

$$\eta(t) = e^{\lambda t}, where \tag{3.24}$$

$$\lambda = F'(x^*) \tag{3.25}$$

This equation tells us whether the perturbation will grow or decay dynamically in time. The parameter λ is called the *eigenvalue* of the system, and is determined by the value of $F'(x^*)$. If λ is ¿ 0, then the perturbation, $\eta(t)$ will grow as a function of time, and if λ is ¡ 0, the perturbation will decay to the steady state. Thus, evaluating the sign of λ ($F'(x^*)$) will indicate the stability of the system.

Negative feedback

Here we discuss modeling a simple gene network negative feedback, where a dimeric protein represses transcription by physical blocking of the promoter transcribing gene x. As with the previous examples, we begin by summarizing the reactions below:

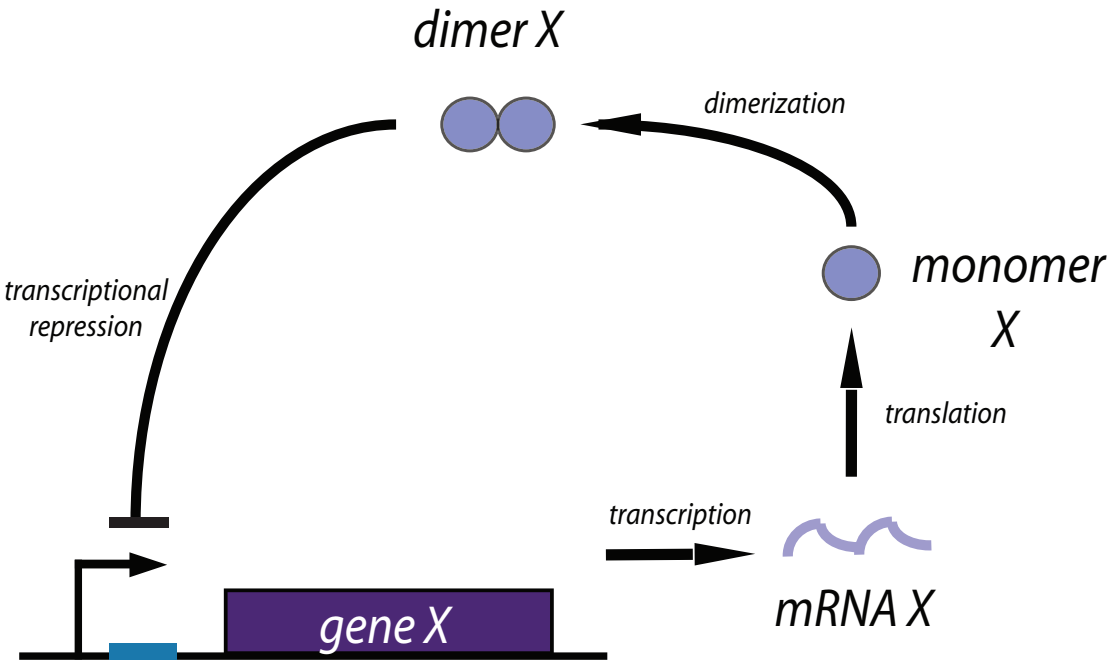


Figure 3.4: Repression of transcription occurs when dimeric protein x binds to a promoter site.

	reaction		
1	$x + x$	$\xrightarrow{\kappa_+}$	y
2	y	$\xrightarrow{\kappa_-}$	x
3	d_o	$\xrightarrow{\alpha}$	$d_o + m$
4	m	$\xrightarrow{\sigma}$	$m + x$
5	x	$\xrightarrow{\gamma_x}$	$\oslash$
6	m	$\xrightarrow{\gamma_m}$	$\oslash$

Note that this differs from the positive feedback example since the reaction

where transcription is increased by the binding of x β is missing. Writing down the mass action equations for the reactions above we arrive at:

Full Equations

$$\dot{x} = 2\kappa_- y - 2\kappa_+ x^2 + \sigma m - \gamma_x x \tag{3.26}$$

$$\dot{y} = \kappa_+ x^2 - \kappa_- y + k_- d_r - k_+ d_o y \tag{3.27}$$

$$\dot{d_o} = k_- d_r - k_+ y d_o \tag{3.28}$$

$$\dot{d_r} = k_+ y d_o - k_- d_r \tag{3.29}$$

$$\dot{m} = \alpha d_o - \gamma_m \tag{3.30}$$

$$\tag{3.31}$$

Again, we assume dimerization and dissociation of the proteins (to themselves and the promoters) are fast compared to other processes.

$$y = c_p x^2; \tag{3.32}$$

$$d_o = d(1 + c_p c_d x^2)^{-1} \tag{3.33}$$

$$\tag{3.34}$$

Substitution of these equations gives

$$f(x)\dot{x} = \sigma m - \gamma_x x \tag{3.35}$$

$$\dot{m} = \frac{\alpha d}{1 + c_p c_d x^2} - \gamma_m m \tag{3.36}$$

$$\tag{3.37}$$

where f(x) is the prefactor for the negative feedback system.

In the new section we extend the use of mass-action kinetic modeling of positive and negative feedbacks to construct a spatial, dynamic model for a mammalian cell system that forms stable spatial patterns.

Chapter 4

In-silico patterning of vascular mesenchymal cells in three-dimensions

Introduction

Cells organize in complex three-dimensional patterns by interacting with proteins along with the surrounding extracellular matrix. This organization provides the mechanical and chemical cues that ultimately influence a cell's differentiation and function. Here, we computationally investigate the pattern formation process of vascular mesenchymal cells arising from their interaction with Bone Morphogenic Protein-2 (BMP-2) and its inhibitor, Matrix Gla Protein (MGP). Using a first-principles approach, we derive a reaction-diffusion model based on the biochemical interactions of BMP-2, MGP and cells. Simulations of the model exhibit a wide variety of three-dimensional patterns not observed in a two-dimensional analysis. We demonstrate the emergence of three types of patterns: spheres, tubes, and sheets, and show that the patterns can be tuned by modifying parameters in the model such as the degradation rates of proteins and chemotactic coefficient of cells. Our model may be useful for improved engineering of three-dimensional tissue structures as well as for understanding three dimensional microenvironments in developmental processes.

The evolution of tissue form in development, wound healing, and regeneration is a dynamic process that involves the integration of local cues on cell fate and function. These cues include interactions with soluble factors (growth factors, morphogens, dissolved gases) and insoluble factors (extracellular matrix, neighboring cells) in a three-dimensional context. A fundamental understanding of how tissue structure evolves is critical to the rational development of engineered tissues for therapeutic applications. There has been increasing evidence that culture of cells in three-dimensions compared to two-dimensions can dramatically impact cellular organization, polarity, and drug responsiveness(Albrecht et al., 2006; Nelson et al., 2006; Zaman et al., 2007; Webb and Horwitz, 2003; Griffith and Schwartz, 2006; Nelson et al., 2005; Cukierman et al., 2002). Here we sought to isolate the role of diffusion/reaction gradients in three dimensions while excluding morphogenetic effects.

Although there have been several modeling efforts to study cell pattern formation and organization in two dimensions(Murray and Oster, 1984; Gierer and Meinhardt, 1972; Tsimring et al., 1995; Garfinkel et al., 2004; Murray, 2003; Koch and Meinhardt, 1994; Painter et al., 1999; Brenner et al., 1998; Gamba et al., 2003; Sage et al., 2009; Ambrosi and Preziosi, 2006), there has not been much attention devoted to three-dimensional systems(Zaman et al., 2006, 2007). Recently, a phenomenological two dimensional reaction-diffusion model with morphogen identified as Bone Morphogenic Protein 2 (BMP-2) and inhibitor Matrix Gla Protein (MGP) was shown to produce the patterning of human vascular mesenchymal cells(Garfinkel et al., 2004). Using a first-principles approach we derive a model based on the underlying biochemical interactions of BMP-2 and MGP and show that our model produces similar patterns as two dimensional experiments. We then perform simulations with our model in three dimensions and explored the types of patterns observed and effect of model parameters. We find that the patterns seen in three dimensions are strikingly different than those seen in two-dimensions and we examine their stability numerically. We discuss these findings in the context of engineering desired tissue structures and also relate to the important differences seen in cell organization between two and three dimensional settings.

The morphogen in the model is Bone Morphogenic Protein 2 (BMP-2), a member of the TGF-β superfamily which to date has over 20 members(Shi and Massague, 2003; Chen et al., 2004). BMP-2 is able to dimerize to its biologically active form [26 kDa for the dimer] and is a potent stimulator of cells to differentiate to an osteoblast-like fate. This occurs through the binding of a BMP-2 dimer to a TGF-β receptor complex, which then functions to phosphorylate the Smad proteins. These proteins then translocate to the nucleus and act as transcription factors for various genes including the gene for BMP-2(Ghosh-Choudhury et al., 1993; Garfinkel et al., 2004). In addition, BMP-2 has been shown to be a strong chemoattractant for these cells and thus is a good candidate for a morphogen in the reaction-diffusion model (Garfinkel et al., 2004; Willette et al., 1999). MGP is a smaller (10.4 kDa) regulatory protein for BMP-2. MGP is thought to inactivate BMP-2 by physical binding to BMP-2 and prevent binding to the receptors (Bostrom, 2000; Bostrom et al., 2001; Zebboudj et al., 2002, 2003; Sweatt et al., 2003; Wallin et al., 2000; Loeser et al., 1992; Price et al., 2002). The presence of BMP-2 also stimulates production of MGP through an unknown mechanism(Garfinkel et al., 2004; Yochelis et al., 2008). In Fig.1, an illustration of the system is shown with the relevant biochemical reactions.

Our simplified model for the reaction-diffusion process of the vascular mesenchymal cell system is derived from the underlying biochemical reactions. The reactions for BMP-2, MGP, and BMP-2 Receptor complexes on the surface of cells are shown schematically in Fig. 1. Transcription, translation, and export out of the cell for BMP-2 and MGP were lumped together for simplicity. We simplified the model using a multiple time scale analysis, which takes advantage of the difference in time scales between the kinetic processes and assumes a local quasi-equilibrium. Below, the model equations are presented in a scaled form with dimensionless concentrations of BMP-2 (U) , MGP (V) , and cells (n) as functions of space (x,y,z) and time (t). The derivation of the model can be found in the Supplementary Info.

$$\frac{\partial U}{\partial t} = D\nabla^2 U + \gamma[\frac{nU^2}{1+kU^2} - cU - KUV] \quad (4.1)$$

$$\frac{\partial V}{\partial t} = \nabla^2 V + \gamma[bnU^2 - eV - KUV] \quad (4.2)$$

$$\frac{\partial n}{\partial t} = q\nabla^2 n - \chi[\nabla \cdot (n\nabla U)] \quad (4.3)$$

In the first equation, the first term on the r.h.s represent diffusion of BMP-2, the second term represents an autocatalytic production of BMP-2 that saturates, the third term is a degradation of BMP-2 at rate *c*, and the fourth is a nonlinear degradation by physical binding of BMP-2 to MGP. The equation for MGP has a similar diffusion term as well as production by BMP-2 term which is known not to saturate(Garfinkel et al., 2004; Zebboudj et al., 2003), degradation of MGP at rate *e*, and nonlinear degradation by physical binding of BMP-2 to MGP. The equation for cell concentration (n) has a diffusion term as well as chemotaxis term that accounts for cells movement toward higher regions of chemoattractant (BMP-2) and also depends on cell density. Parameters D=$\frac{D_U}{D_V}$, $q = \frac{D_n}{D_V}$ are the ratios of diffusion coefficients for BMP-2 to MGP, Cells to MGP, respectively. The coefficient *b* represents the relative production of MGP to BMP-2, *c* and *e* represent the degradation of U and V, and *K* represents the nonlinear degradation of U and V by physical binding. The parameter γ is a scaling parameter for the relation between domain size and chemical kinetics.

The diffusion coefficients, production rate of BMP-2, degradation rates of BMP-2 and MGP were taken from the literature (Garfinkel et al., 2004; DiMilla et al., 1992). The production of MGP is known to be similar to BMP-2 (although its value uncertain) and was set to a value of $b = 1.1$. The nonlinear degradation coefficient, K, can be expressed in terms of kinetic rate parameters but these rates are also unknown, and thus was set to $K = 0.25$ along with $b = 1.1$ to reproduce the stripe patterns seen in previous work(Garfinkel et al., 2004). The mean cell density n_0, which is conserved in the dynamics is set to n_0=1.

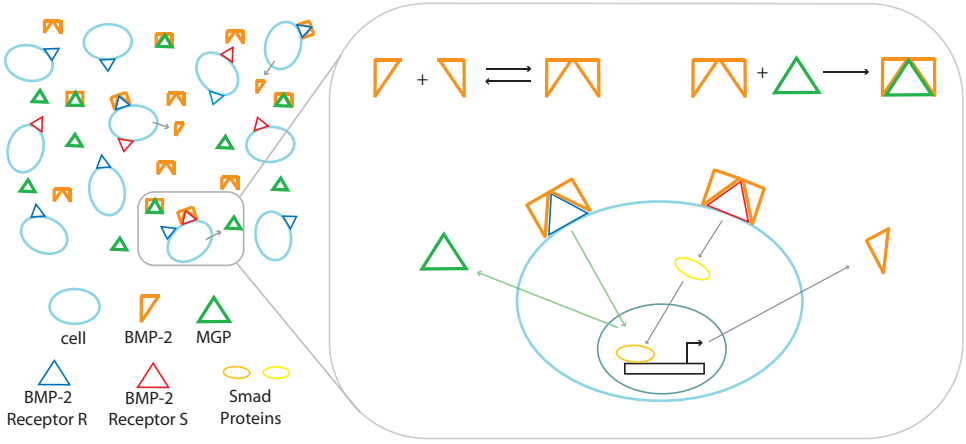


Figure 4.1: Diagram showing interactions between BMP-2, MGP, and cells in culture. The binding of a BMP-2 dimer to receptors R and S stimulates production of BMP-2 and MGP, while the binding of MGP to BMP-2 outside of the cell prevents this process. The production of BMP-2 occurs via the Smad signalling pathway and the production of MGP occurs through an unknown pathway.

Results

The mathematical model admits up to 3 real uniform steady states for the parameter region we explored. Of these, one is always the zero solution$\{U = 0$,$V = 0$,n=1$\}$, the other is low $\{U = 0.1$,$V = 0.2$,n=1$\}$, and the third is high $\{U = 1.0, V = 3.0, n = 1\}$. In the supplementary info, a linear stability analysis was carried out to analyze the stability of these steady states and determine the region where patterns are found. Briefly, the linear stability analysis analyzes a small perturbation from the steady state and determines which modes of the perturbation are unstable, which generally corresponds to the size of the perturbation. Among these states, the zero solution is always stable and the low solution is always unstable. The high state is stable with respect to spatially uniform perturbations, but it can be unstable with respect to spatially non-uniform modes. We performed simulations and analyzed the stability of these steady states (Supplementary Info) and found that only the higher steady state produced patterns that resembled the experiments and is likely the physiologically relevant one. We start with an initial condition at this steady state and add a 1% relative random noise

to model cell variation(Garfinkel et al., 2004). The simulations shown in Figures 2 and 3 are the state distribution of cells with red color indicating high levels of cell density and blue levels indicating low levels of cell density. The lowest values of cell density are made transparent for visuall clarity. The parameters used unless otherwise specified were D=0.005, q=0.003, $\chi = 10^{-5}$, K=0.25, B=1.1, γ=600 and the box length of the simulation is equivalent to 1 cm.

Simulations in two dimensions varying the parameters c (degradation of BMP-2) and k (saturation of production of BMP-2) are shown in Figure 2. Three basic types of steady state patterns emerge from the model (Fig. 2a-c): (a) spots , (b) stripes, and (c) inverse spots. By stripe patterns we mean that cells arrange in higher densities along stripe regions with characteristic thickness. The spot patterns correspond to clusters of cells and the inverse spots show connected structures of cells with gaps of no cells in between. The stripe and spot patterns were previously seen in the experimental two-dimensional setting, although the inverse spot patterns were not. Fig 2(d) shows where the patterns are found in parameter space upon scanning parameters c and k. The solid line between the regions of no patterns and patterns is predicted by our linear stability analysis and matches with our visual inspection of the simulations. We used a 20x20 grid of numerical simulations and visually inspected the simulations to determine their pattern type. In regions that show existence of more than one pattern we labeled the pattern type by the majority of the pattern seen.

In Fig. 3, we show the simulations in three dimensions varying the same parameters c and k. In three dimensions, the steady state patterns produced are (a) spheres of cells, (b) solid tubes, and (c) highly interconnected tubes which have planar surfaces. These three pattern types are somewhat analogous to the 2D patterns of spots, stripes and inverse spots, respectively. Movies for each of these cases can be found in the supplementary info. The distinguishing feature between types (b) and (c) is that the cross section of the sheet like structures resemble stripes while the cross section of the solid tubes resembles spots. Fig 3(d) also shows where the patterns are found in parameter space with a 9x9 grid of numerical simulations.

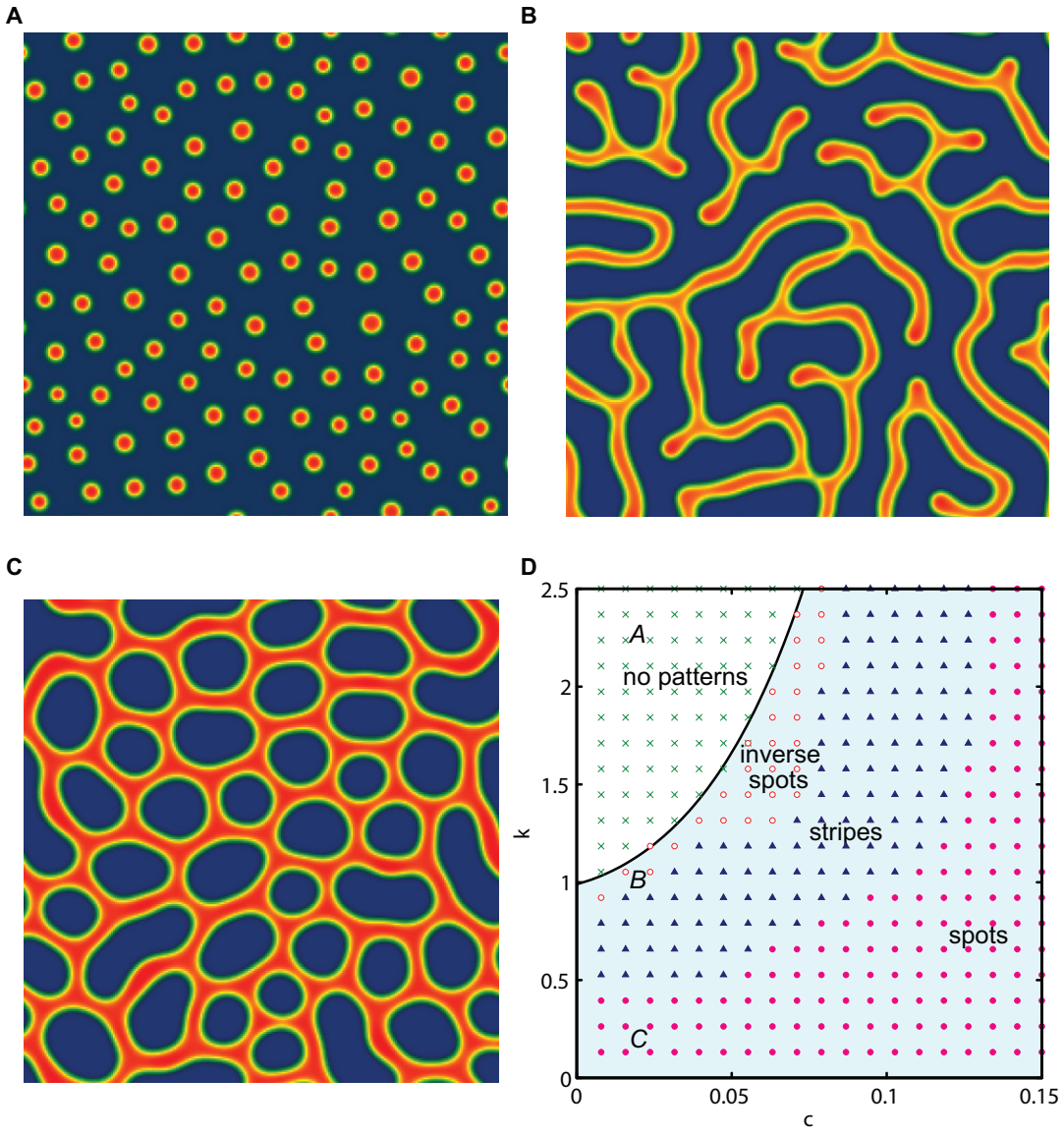


Figure 4.2: 2D steady state patterns of cells. The derived model shows (a)spots(k=0.2,c=0.12), (b)stripes(k=0.7,c=0.04), and (c) inverse spots(k=0.95,c=0.005) by varying k and c. The parameters used were D=0.005, q=0.003, K=0.25, B=1.1, γ=600 and the box length of the simulation is equivalent to 1 cm. Red color indicates higher cell density while blue indicates low.

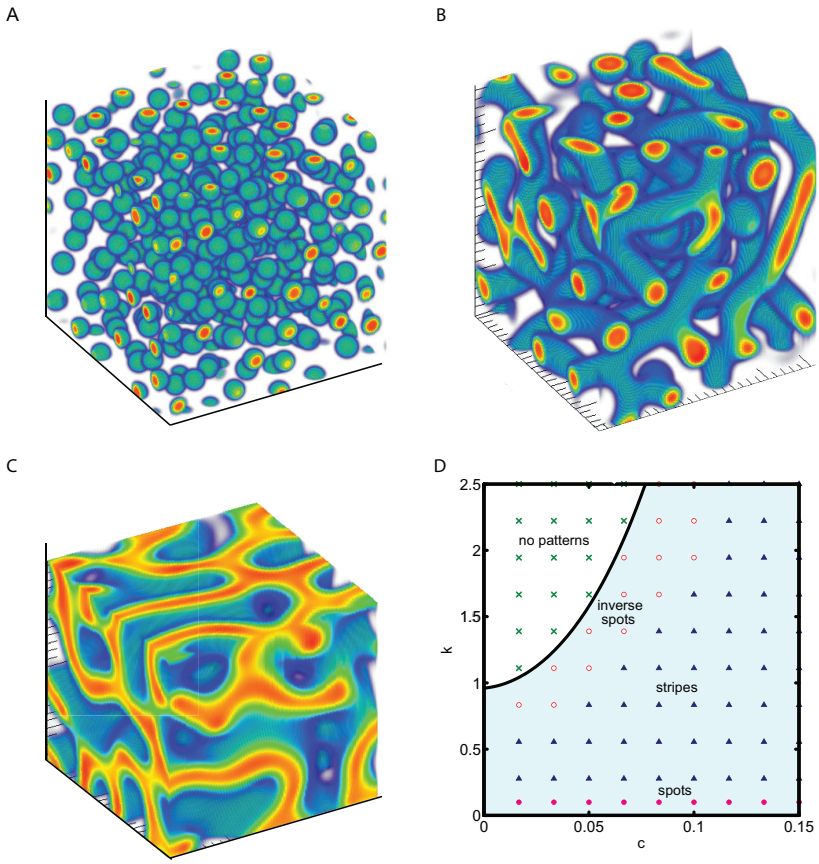


Figure 4.3: 3D steady state patterns of cells. The derived model shows spherical spots(k=0.2,c=0.12), tubes(k=0.2,c=0.04), and sheet-like structures(k=0.8,c=0.04) by varying k and c. The parameters used were D=0.005, q=0.003, K=0.25, B=1.1, γ=600 and the box length of the simulation is equivalent to 1 cm. The lowest values were made transparent for clarity while red color indicates higher cell density while blue indicates low.

Fig. 4 shows the evolution of cells with an initial condition of a (a) spherical or (b) cylindrical region along the center axis containing at 2x higher BMP-2 concentration than the steady state. The surrounding region was set to the zero value. The parameters set for these simulations were those in the stripe pattern regime to mimic the previous experimental setting(Garfinkel et al., 2004).

Discussion

Figures 2(d) and 3(d) show the locations of the types of patterns in two dimensions and three dimensions as a function of parameters c and k. We see that in the two-dimensional case the spot patterns are seen over a wide range of parameters while in three-dimensional case these patterns are only rarely seen. In trying to correlate the 2D pattern region with the 3D pattern region we scaled the diffusion and chemotactic coefficient by 3/2 to reflect the change from 2D to 3D. We found that this did not significantly alter where the patterns are seen in the parameter space. This difference in the pattern location may arise because of the spatial symmetry of the problem. For instance, the tubes which are seen often in three-dimensions can be cut along different axes to form either the spot or stripe patterns seen in two-dimensions. Thus, they occupy a larger region in the parameter space for three-dimensions than in two-dimensions. For an experimental system with fixed parameters, we would predict that the organization of cells in two dimensions greatly differs from that in three dimensions, suggesting a possible reason for the biological differences seen in experimental culture of mammalian cells(Albrecht et al., 2006).

In the parameter space we explored, we found that multiple patterns can coexist for a fixed set of parameters and we examined the stability of each type. We ran a 2D simulation to steady state which showed only spots (point C, Figure 2d), and then increased the parameter k slowly while allowing the system to equilibrate. Doing this from point C to point B in Figure 2d we found that the spot patterns remained stable throughout the region and finally disappeared when reaching the no pattern region(point

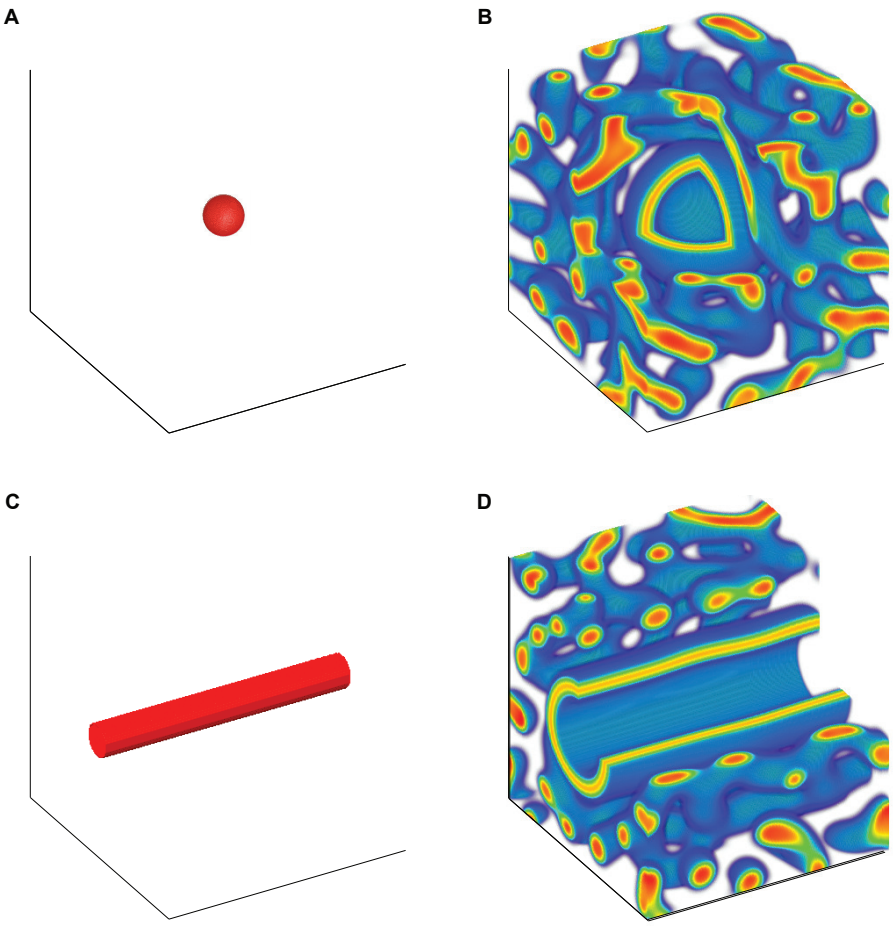


Figure 4.4: Initial and steady state patterns of cells produced by exogenous BMP-2. An initial condition of 2x higher concentration of BMP-2 is placed along the center (a) sphere or (b) cylinder and the cells are allowed to reach steady state. The stripe regime parameters were used and set as D=0.005, q=0.003, K=0.25, B=1.1, k=0.7, c=0.14, γ=600 with simulation box length set to 2cm. The lowest values were made transparent for clarity while red color indicates higher cell density while blue indicates low. A cut of the simulation box in (a) 1/8 of cube and (b) 1/4 of cube was sliced out for easier visualization.

A). In the regions where stripes were found(point B), the spot patterns would temporarily nucleate into stripes and then go back to their spot pattern state. We also performed the opposite case starting at point B and decreasing *k*. In this case we found the patterns to go from the inverse spot pattern type to the stripe pattern, but then we found that at point C the cells remained in the stripe pattern type and did not change into the spot pattern type. This indicates that the inverse spot type of pattern is least stable to perturbations, while the stripe and spot patterns are more stable. Along with the fact that the inverse spot type is seen least in parameter space, this may suggest why this type of pattern has been difficult to realize experimentally(Garfinkel et al., 2004).

We also performed simulations that can be directly tested in three-dimensional experiments. For instance, an experiment where a higher concentration of BMP-2 is produced at the center region can be represented by an analogous initial condition in our simulation. In Fig. 4, simulations were performed with an initial condition set so that a local (a) sphere or (b) cylindrical region of BMP-2 is at a 2x higher concentration than the steady state value(see Supplementary Info for movie). The parameters set for these simulations were those in the stripe pattern regime to mimic previous experimental observations for the vascular mesenchymal cell system. For the spherical case, we found that the morphogen concentration will grow in expanding spheres and the cells will arrange themselves in the same way. For the cylindrical initial condition, we found that the cells will evolve in a hollow cylinder from the initial condition forming a vessel-like shape.

Additionally, we investigated the effect of cell parameters on the patterns observed. The random cell motility, q, and the chemotactic coefficient, χ, both play a role in the stability and pattern selection of cells. We found that by varying the ratio of $\chi/\ q$, it is possible to change the pattern type from one to another and it is possible to end up in a regime where no patterns are formed. This situation occurs for points near the stability border with a change to the nominal value of $\chi = 1 \cdot 10^{-5}$. Whenχ is changed to $\chi = 3 \cdot 10^{-4}$ and then $\chi = 7.5 \cdot 10^{-4}$ the patterns observed are of the inverse spot and stripe pattern type, respectively(Supplementary Info). For the higher ratio of $\chi/\ q$, we

found that the cells are more often found in the spot pattern type, showing that these are most stable types(Supplementary Info).

The simulations we have done here show the importance of three-dimensional modeling of cell organization. In three dimensions we found that the patterns and organization of cells is much richer than in 2D and found that the same model system with fixed parameters in two and three-dimensions can exhibit different steady-state pattern types. Simulations to mimic developmental processes and engineering of three-dimensional tissue structures will thus find these techniques to be useful for predicting cell organization in three dimensions. In addition, we presented simulations that could easily be tested in two- or three- dimensional experiments to validate our model.

Materials and Methods

We performed two- and three- dimensional simulations using a pseudospectral technique as described in(Cross et al., 1994). The method handles the nonlinearities explicitly in real space and diffusion in Fourier space. To simulate the cell equation we kept the zero mode a constant since the total cell mass is conserved. We found that the method shows agreement up to numerical accuracy with solutions to known nonlinear equations (Supplementary info). Furthermore, we saw convergence of our numerical results for a range of timesteps and spatial discretizations. The technique we used assumes periodic boundaries on the spatial domain.

Three-dimensional simulations were parallelized using the Message Passing Interface (MPI 2.0) in conjunction with the FFTW library. We used a 256^3 (a 128^3 for the 9x9 scan in Figure 3) with $dx = 0.5/256$ which typically required about $10^5 - 10^6$ steps to reach steady state at a step size of dt=$2 \cdot 10^{-4}$. For the 256^3 grid, a typical computation time of 120 hours on a single processor or 30 hours on eight processors was needed to perform most simulations. IDL software (ITT Visual Information Solutions) was used for visualizing three-dimensional graphics.

Parameter selection

The degradation rates of BMP-2 (U) and MGP (V) were previously taken to be 1% and 2% of the production rate of BMP-2. Since the previous model had a V in the denominator term for production, and a steady state V=61.1 in their model, we scaled our coefficients by 10 to reflect this change. To obtain a numerical value for γ, we scaled the production rate of 125 ng/mL in 48 hr to get a unitless value of $\gamma = 125$. This gave features for stripe thickness and spacing in our model of 200 and 500 microns. Since these values were largely uncertain, we changed γ by an additional factor of 4 to best fit the experimental stripe thickness of 350 microns. Values for the cell diffusion coefficient were taken from a similar cell type in the literature ($1{\cdot}10^{-9}cm^2/s$(DiMilla et al., 1992). The MGP and BMP2 diffusion coefficients are taken from a previous paper(Garfinkel et al., 2004), where they calculated the values to be $1.5{\cdot}10^{-9}cm^2/s$ and $3{\cdot}10^{-7}cm^2/s$. The chemotaxis coefficient was estimated with a dimensional argument using characteristic values taken from experimental BMP2 chemotaxis data(Willette et al., 1999). We estimated χ from a dimensional argument in a BMP2 cell migration experiment(Willette et al., 1999). There, cells migrate across a BMP-2 gradient, where the number of cells leaving the chamber should be proportional to the gradient, cell concentration, and χ. Thus $\frac{\Delta N}{\Delta t} \approx \frac{N\chi U_0}{L^2}$, which simplifies to $\Rightarrow \chi \approx \frac{L^2}{U_0 t}$. Plugging in characteristic values, L=1μm, U_0=1 μM, t=6 hours, gives $\chi = 4.629 \cdot 10^{-11} \frac{(\mu m)^2}{\mu M s}$. Non-dimensionalizing this with L=4 cm, $t = 3.6x10^3 s$ (typical time and length scales for our problem as used in (Garfinkel et al., 2004)) and choosing $1\mu M$ as our characteristic concentration gives $\chi = 1.04 \cdot 10^{-5}$ used in our simulations. In our simulations, both of these parameters did not seem to change the region of pattern formation by much but changed the selection of patterns (see below for various χ/q values).

Mathematical derivation of the model

The derivation of our model uses the fundamental biochemical reactions to make a full mass action model and reduces to a simplified form because of the large differences in time scales between reactions. The biochemical reactions are listed in Table 1 of the main text.

	reaction		
1	$U + U$	$\xrightarrow{k_1}$	U_2
2	U_2	$\xrightarrow{k_{-1}}$	$U + U$
3	$U_2 + V$	$\xrightarrow{k_2}$	$\oslash$
4	$U_2 + R_2$	$\xrightarrow{k4}$	U_2R_2
5	U_2R_2	$\xrightarrow{k_{-4}}$	$U_2 + R_2$
6	U_2R_2	$\xrightarrow{k_5}$	$U_2R_2 + U$
7	$U_2 + S_2$	$\xrightarrow{k_6}$	U_2S_2
8	U_2S_2	$\xrightarrow{k_{-6}}$	$U_2 + S_2$
9	U_2S_2	$\xrightarrow{k_7}$	$U_2S_2 + V$
10	U_2	$\xrightarrow{c}$	$\oslash$
11	V	$\xrightarrow{e}$	$\oslash$

Table 4.1: Monomeric BMP-2 (U) is able to dimerize (U_2) and bind to MGP (V) or bind to receptor complexes (R_2, S_2) producing either BMP-2 or MGP, respectively. BMP-2 and MGP degrade at rates c, e.

Full Model

$$\dot{U} = 2k_{-1}U_2 - 2k_1U^2 + k_5U_2R_2 \quad (4.4)$$

$$\dot{U}_2 = = -cU - k_2U \cdot Vk_1U_2 - k_{-1}U_2 - k_4U_2 \cdot R_2 + k_{-4}U_2R_2 \quad (4.5)$$

$$\dot{V} = k_7U_2S_2 - k_2U \cdot V - eV \quad (4.6)$$

$$\dot{R}_2 = -k_4U_2 \cdot R_2 + k_{-4}U_2R_2 \quad (4.7)$$

$$\dot{U}_2R_2 = -\dot{R}_2 \quad (4.8)$$

$$\dot{S}_2 = -k_6U_2 \cdot S_2 + k_{-6}U_2S_2 \quad (4.9)$$

$$\dot{U}_2S_2 = -\dot{S}_2 \quad (4.10)$$

Assumptions and Simplifications

We assume transcription, translation, and export out of the cell for BMP-2 and MGP (6,9) are slow in comparison to fast binding binding reactions (1,2,4,5,7,8). Binding and degradation of BMP-2 and MGP (3) are slow as well. We also assume degradation of the monomer to be negligible in comparison to degradation of the monomer because BMP-2 is mostly in dimeric form when transported out of the cell.

Letting R_T and S_T be the total # of receptors gives:

$$U_2R_2 = \frac{K_1K_4[U]^2R_T}{(1 + K_1K_4[U]^2)}; \qquad U_2S_2 = \frac{K_1K_6[U]^2S_T}{(1 + K_1K_6[U]^2)} \quad (4.11)$$

Using these simplifications and adding equations for U and U_2 gives:

$$(1 + \frac{1}{2K_1^{1/2}U_2^{1/2}})\dot{U}_2 = k_5\frac{K_1K_4[U]^2R_T}{(1 + K_1K_4[U]^2)} - cU + -k_2U \cdot V \quad (4.12)$$

$$\dot{V} = k_7\frac{K_1K_6[U]^2S_T}{(1 + K_1K_6[U]^2)} - k_2U \cdot V - eV \quad (4.13)$$

We further assume that K_1 ($K_i = \frac{k_i}{k_{-i}}$) is large since BMP-2 is mostly in dimer form and that saturation by receptor (S_2) is not seen experimentally. Experimentally MGP saturation is not seen(Garfinkel et al., 2004; Zebboudj et al., 2003), which may

be caused by the parameter K_6 to be large, and is removed from our model. We use the relation that $R_T = nR$, where R is the # of receptors per cell, and this is how cell dependence on production comes naturally into the model. We non-dimensionalize the model equations starting from equations that have dimensional variables and parameters as above. We choose $U_0 = V_0$ and $t_0 = \frac{L_0^2}{D_V}$ to simplify the equations. Combining the derived kinetic model with diffusion terms and renaming U_2 to U gives.

$$\frac{\partial U}{\partial t} = D\nabla^2 U + \gamma[\frac{nU^2}{1+kU^2} - cU - KUV] \tag{4.14}$$

$$\frac{\partial V}{\partial t} = \nabla^2 V + \gamma[bnU^2 - eV - KUV] \tag{4.15}$$

$$\frac{\partial n}{\partial t} = q\nabla^2 n - \chi[\nabla \cdot (n\nabla\, U)] \tag{4.16}$$

$$with \quad \gamma = \gamma^\star U_0 n_0 t_0, \qquad D = \frac{D_U}{D_V}, \qquad q = \frac{D_q}{D_V}, \qquad \chi = \frac{\chi^\star}{D_V} \tag{4.17}$$

Linear stability analysis

We perform a linear stability analysis on the system (4.14,4.15,4.16) about a uniform steady state. Letting $w = [U - U_0, V - V_0, n - n_0]$, the system can be written:

$$\frac{\partial w}{\partial t} = \gamma A w + D\nabla^2 w \tag{4.18}$$

$$A = \begin{bmatrix} f_U & f_V & f_n \\ g_U & g_V & g_n \\ h_U & h_V & h_n \end{bmatrix} \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & q \end{bmatrix} \tag{4.19}$$

where f(U,V,n), g(U,V,n) are derived kinetic functions for U, V and $h(U, V, n) = \frac{1}{\gamma}[-\nabla \cdot (n\chi \nabla U)]$ and we linearize this term by hand. We require nontrivial solutions for W_k so the eigenvalues λ are determined by roots of :

$$|\lambda I - \gamma A + Dk^2| = 0$$

We solved this equation numerically to obtain the dispersion relations and hence the stability regions shown in Figures 2(d) and 3(d). One can also use the Routh-Hurwitz criteria(DeJesus and Kaufman, 1987) to reduce the criteria for stability as well, however, the two methods give the same solution and we found it easier to solve this equation directly.

The linear stability analysis showed that the zero steady state was stable for all wave numbers. The small steady state is unstable for a range of wavelengths but is not of the pattern forming type. A typical dispersion relation for the highest steady state and $k = 1.0$ and $c = 0.1$ is shown below. The real parts of the three eigenvalues(λ) are plotted against the wave number k. A finite wavelength instability occurs near $k = 4$ which corresponds to the characteristic size of the pattern. For some of our parameters we found that one of the eigenvalues can become unstable at $k = 0$, but we required that the max nonzero k eigenvalue be larger than the max zero eigenvalue when determining the pattern forming region.

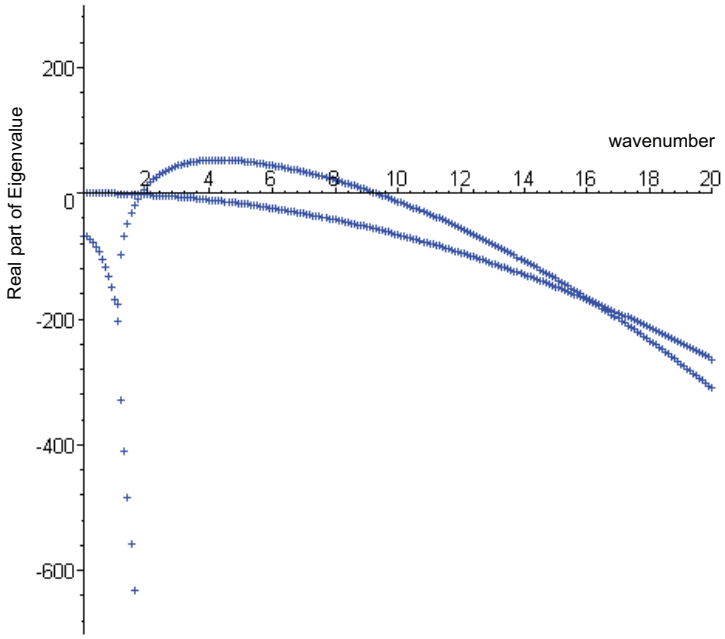


Figure 4.5: Dispersion relation for k=1.0, c=0.1. The other parameters used here are listed in the supplementary text.

We also analyzed the effect of changing the cell parameters, χ and q. Below three plots are shown for different values χ having the value of $q = 0.0033$. The spot patterns are more commonly found in this parameter space when increasing χ.

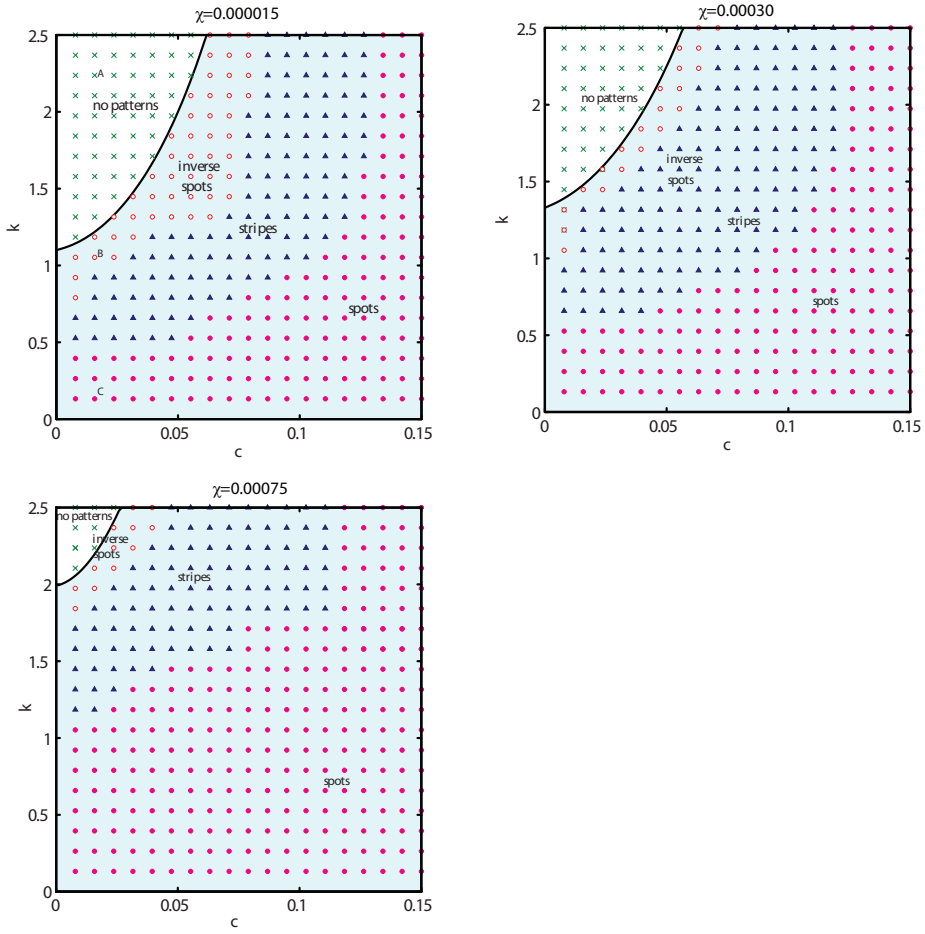


Figure 4.6: Phase diagrams showing where patterns emerge in parameter space for various χ values and other parameters listed above.

Acknowledgements

Chapter 5 contains material originally published as Danino,T. Volfson, D., Tsimring, L., and Hasty, J. ***PLoS ONE*** (2011): In-silico pattern formation of vascular mesenchymal stem-cells in three-dimensions. Copyright permission to republish here was

granted by Public library of science.

Chapter 5

Correlated signaling through coupled degradation

Introduction

High-throughput technologies have led to the generation of complex wiring diagrams as a post-sequencing paradigm for depicting the interactions between vast and diverse cellular species. While these diagrams are useful for analyzing biological systems on a large scale, a detailed understanding of the molecular mechanisms that underlie the observed network connections is critical for the further development of systems and synthetic biology. Here, we use queueing theory to investigate how "waiting lines" can lead to correlations between protein "customers" that are coupled solely through a downstream set of enzymatic "servers". Using the *E. coli* ClpXP degradation machine as a model processing system, we observe significant cross-talk between two networks that are indirectly coupled through a common set of processors. We further illustrate the implications of enzymatic queueing using a synthetic biology application, in which two independent synthetic networks demonstrate synchronized behavior when common ClpXP machinery is overburdened. Our results demonstrate that such post-translational processes can lead to *dynamic* connections in cellular networks and may provide a mech-

anistic understanding of existing but currently inexplicable links.

Evolutionary pressure has driven organisms to develop into energy efficient machines, which conserve resources by minimizing biosynthetic costs and producing only the required amount of certain costly proteins(Warner et al., 2001; Vilaprinyo et al., 2010; Barton et al., 2010). An implication of maintaining only the minimal amount of critical machinery is the potential to overload important pathways during times of stress and to place unanticipated burden on important cellular workhorses. This can lead to the development of "waiting lines" for biochemical processing which may cause apparent correlations between seemingly disconnected components that share the same processing pathways. In that sense, an analogy can be drawn to multiclass queueing theory(Kelly, 1979; Bramson, 1998; Williams, 1998; Bramson and Dai, 2001), which we employ here to provide a unifying model for describing how "waiting lines" for processing by a common enzyme (the "servers") can lead to correlations between two otherwise uncoupled proteins (the "customers").

Generally, when strong correlated behavior is observed between two proteins in response to some perturbation, it is assumed that there is a direct coupling mechanism in place, such as correlated transcription. However, recent studies have revealed a lower degree of correlation between mRNA and protein levels than expected, indicating the need to search for other coupling mechanisms that may lead two protein species to follow similar trends in concentration(Guo et al., 2008; Gygi et al., 1999; Futcher et al., 1999; Greenbaum et al., 2003). Here, we use several experimental approaches along with a new application of queueing theory to reveal that a seemingly minor form of indirect coupling between cellular species can lead to surprisingly strong correlated behavior. As a model system, we use the native *E. coli* protease, ClpXP, as the "server" and impose various static and dynamic conditions of under- and overloading the cells with protein "customers" targeted for destruction by this complex. We demonstrate that the transition from an underloaded to an overloaded regime can manifest itself in significant cross-talk between two independent networks, where the induction level of one protein substantially affects the mean and variability of the other protein.

The correlated behavior that arises from this type of indirect coupling can have many implications, both in the analysis of native cellular networks and in the design and construction of synthetic networks. For example, systems biology employs high throughput technologies to reconstruct cellular networks and generate high level wiring diagrams(Alon et al., 1999; Golub et al., 1999; Ideker et al., 2001; Sauer, 2004; Li et al., 2004). While useful as tools for analyzing and understanding biological networks on a large scale, determining *how* these components are connected is the next critical step in understanding what the underlying interactions are and how they lead to the observed cellular behaviors. Similarly, the field of synthetic biology relies on a fundamental understanding of the relationship between cellular networks and the behaviors that emerge from their complex interactions(Hasty et al., 2002b; Gardner et al., 2000b; Elowitz and Leibler, 2000c; You et al., 2004a; Tabor et al., 2009b; Danino et al., 2010). As many of the emerging studies in synthetic biology aim to develop circuits that exhibit specific, dynamic behaviors, understanding and utilizing both direct and indirect coupling mechanisms will be become essential to designing successful synthetic systems.

Results and Discussion

In order to illustrate how queueing theory can be used to analyze indirect coupling, suppose that two proteins X_1 and X_2 are involved in signaling pathways that do not directly interact, but they are processed downstream by the same enzyme. If the enzymatic "servers" (S) are in abundance relative to the the number of target "customer" molecules (x_1 and x_2), then there are no waiting lines and this corresponds to an *underloaded* system in queueing theory (Figure 1a, left). On the other hand, if the number of servers is small compared to the number of customers, the system becomes *overloaded* as the customers must wait in line to be processed (Figure 1a, right). Such an overloaded system introduces coupling between the different types of customers. For example, consider an increase in the number of X_1 molecules on the right side of Figure 1a. Conceptually, a rise in the number of X_1 molecules will increase the mean waiting time for

the processing of X_2, leading to a decrease in the effective rate at which X_2 leaves the system. In other words, for fixed arrival rate of X_2, the mean number of X_2 will increase as the number of X_1 is increased, even though there is no direct coupling between the two protein species.

If the arrival of the two proteins X_1 and X_2 in Figure 1a is governed by Poissonian statistics, with production rates λ_1 and λ_2, respectively, the enzymatic system can be mapped to a queueing system where analytical formulae for the steady state distributions can be derived (Mather et al., 2010) (see (Mather et al., 2011) for discussion of transients). While the resulting formulae are somewhat complex, they can be used to predict several generic properties that should arise in the context of coupled enzymatic processing (Figs.1b–d). For example, for a fixed production rate λ_2 of X_2, the mean level of X_2 increases as the production rate for X_1 is increased (Figure 1b). Initially, this increase is slow since the system is in the underloaded regime, but as waiting lines begin to lengthen, the system transitions to the overloaded regime and the mean of X_2 rapidly rises as λ_1 is increased. In addition, the transition point between the underloaded and overloaded regimes depends on the (fixed) value of λ_2. We can also investigate the general statistical properties of X_2 fluctuations as the production rate of X_1 is increased (Figure 1c). Interestingly, since the production rate for X_2 is held fixed, such plots isolate the contribution of the degradation process to the overall variability (although it should be noted that fluctuations in the production of X_1 will contribute to such degradation noise in X_2). Finally, in the context of signaling, using a deterministic approximation to the dynamics, we can predict how periodic modulation of the production rate of X_1 will lead to a correlated response of X_2 (Figure 1d).

A common example of shared enzymatic machinery is the superfamily of AAA+ proteases which target for degradation multiple types of proteins that are either damaged or no longer required(Sauer et al., 2004; Levchenko et al., 2000a). As a model enzymatic queueing process, we explored the degradation of ssrA-tagged proteins by the native *E. coli* ClpXP machine. ClpXP is a protease composed of two multimeric subunits, ClpX and ClpP. ClpX is a hexameric ATP-ase which binds and denatures

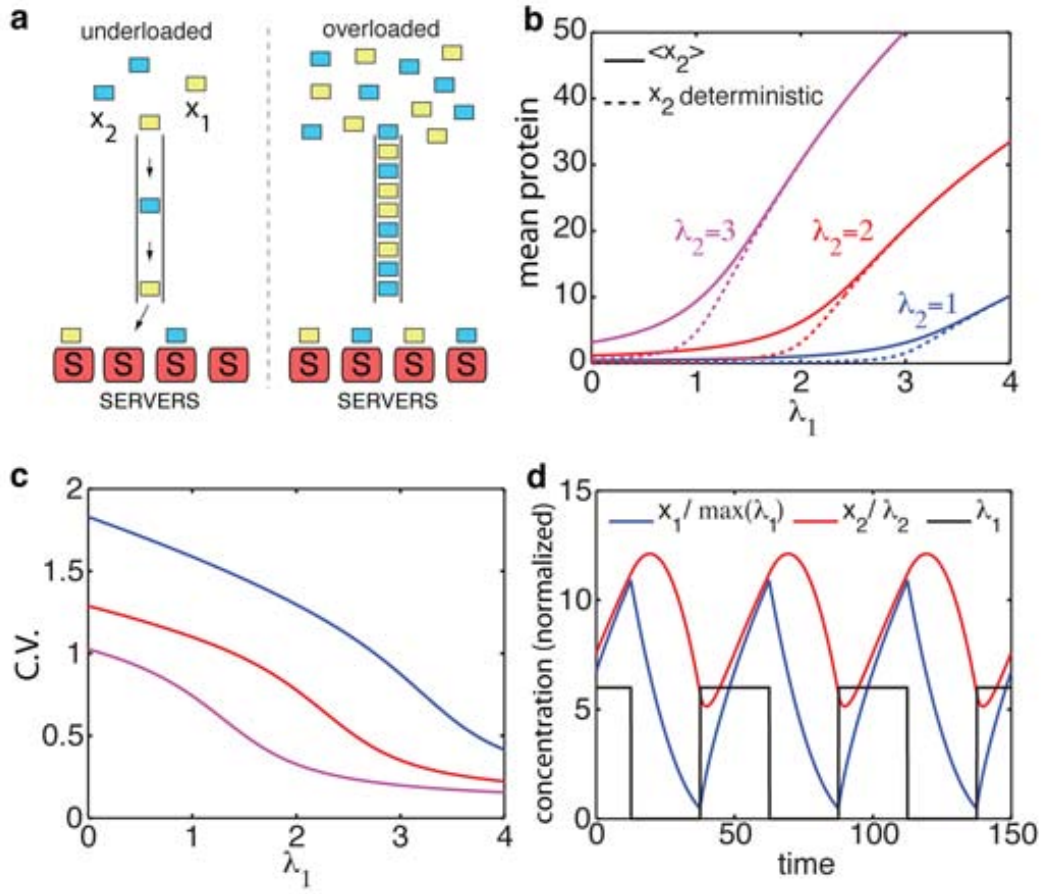


Figure 5.1: Coupling via common enzymatic machinery: connection to queueing theory. (a) Rate-limited processing can couple the numbers of different job types in a queue. (b) Numbers x_1 and x_2 of proteins X_1 and X_2 in the stationary regime depend on the rate of production λ_1 of protein X_1 at fixed values of the production rate of the second protein $\lambda_2 = 1, 2, 3$, the rate constant for enzymatic degradation $\mu = 4$, Michaelis-Menten molar constant $K = 0.2$, and the dilution rate constant $\gamma = 0.02$. Solid lines indicate the mean values of the corresponding concentrations from the exact solution of the full stochastic queueing model, and dashed lines show the deterministic mass action solution (which becomes accurate as the numbers of proteins of each type become large). (c) Coefficients of variation (CV) of x_2 as a function of λ_1 from the queueing model of common enzymatic degradation for the same parameter values as in (b). All three CVs decrease with increasing λ_1. (d) Degradation coupling in a driven system. Here, $\lambda_2 = 2$ is constant, while λ_1 varies between 0 and 6 (indicated by black line). The oscillating signal in x_1 (blue) is propagated into x_2 (red) by the degradation coupling.

proteins that are targeted for rapid removal from the system; these proteins are then translocated into the degradation chamber within ClpP, where they are destroyed. An important role of ClpXP is to target proteins that have stalled in the ribosome; these incomplete proteins are marked with a well-characterized ssrA tag (Keiler et al., 1996). While other proteases, such as ClpAP and Lon (which are implicitly included in our enzymatic degradation model), also contribute to the degradation of ssrA-tagged proteins, ClpXP is responsible for over 90% of this activity (Lies and Maurizi, 2008). In healthy cells, the burden on ClpXP appears to be relatively low, with the capacity to handle an increase in tagged proteins of about 3–fold without overloading the available protease machinery (Moore and Sauer, 2005). However, because ClpXP is involved in increasing degradation rates in response to stress (Damerau and St John, 1993; Schweder et al., 1996; Neher et al., 2006), there is evidence that significant queues are likely to arise when cells must cope with adverse environmental conditions. For example, under DNA damage conditions some critical proteins that are targets of ClpXP can increase in abundance by over 10-fold (Neher et al., 2006). As many of these are already highly abundant proteins, such as the ribosomal subunit proteins *rplJ* and *rplO* (present at over 5K copies each in healthy conditions (Lu et al., 2006)), there are likely tens of thousands of proteins competing to be processed by only 50-100 degradation machines (Farrell et al., 2005).

In order to mimic a cellular environment in which the ClpXP machinery is overburdened, causing target molecules to effectively enter a processing queue, we constructed a synthetic system to over-express two different tagged proteins from separate and uncorrelated promoters (Figure 2a). The $P_{LtetO-1}$ promoter, used to drive expression of YFP (yeast-enhanced yellow fluorescent protein, venus), is tightly repressible by the *Tet* repressor (TetR) and can be regulated over a range of up to 5000-fold by supplying doxycycline to the culture (Lutz and Bujard, 1997b). The hybrid promoter, $P_{lac/ara-1}$, used to drive expression of CFP (yeast-enhanced cerulean fluorescent protein), is tightly repressed by the *Lac* repressor (LacI) and activated by AraC. It can be regulated over a range of up to 1800-fold in the presence of IPTG and arabinose in the

culture (Lutz and Bujard, 1997b). For high expression off of the $P_{lac/ara-1}$ promoter, we used 1mM of IPTG in all samples to fully relieve repression by LacI, and we used various levels of arabinose to tune the induction level of CFP. Both YFP and CFP were tagged on their C terminus with the well-characterized 11-residue "LAA" tag (AANDENYALAA), marking them as targets for rapid degradation by ClpXP (Keiler et al., 1996). In order to ensure stable synthesis and maintenance of the regulatory proteins, the synthetic system was transformed into an *E. coli* DH5αZ1 host that produces constitutive levels of TetR, LacI, and AraC off of the chromosome (Lutz and Bujard, 1997b), and controls were performed to ensure that there was no direct cross-talk between the two promoters (see Supplementary Information).

Our queueing analysis of the enzymatic decay of highly expressed proteins predicts that at certain levels of expression, the mean level of one protein will be significantly coupled to the mean of the other. To test this hypothesis, we planned to hold production of one protein constant, while tuning the level of the other. As a first step, we acquired induction data for the $P_{lac/ara-1}$ promoter, in order to determine several arabinose levels to fix for the constant condition. Interestingly, we found a striking difference in the shapes of the induction curves for tagged and untagged variants of the fluorescent reporter (Figure 2b). While the untagged protein responded in the expected manner (Lutz and Bujard, 1997b), the tagged protein data showed clear evidence of queuing, with a sharp bend in the curve indicating the balance point, where ClpXP transitions from the under- to overburdened state.

We then used two-color flow cytometry to generate induction curves for the two systems, fixing arabinose at three levels and tuning YFP production with doxycycline, and we were surprised by the extent to which the inducer for one species affected the other species. For example, as doxycycline was varied over a full induction range for YFP, we observed up to a 12-fold increase in the mean level of CFP (Figure 2c). Likewise, we observed unambiguous correlations between the two proteins as we swept doxycycline and plotted the mean levels of CFP as a function of the mean levels of YFP (Figure 2d). This result highlights an important point in the general context of

gene-regulatory and signaling networks; were this data collected in the absence of a mechanistic understanding, one would likely conclude that YFP is some form of inducer for CFP.

In order to place this "queueing" phenomenon in the context of native cellular behavior, it is important to develop a quantitative understanding of the system's balance point and the processing capacity of ClpXP in healthy *E. coli* cells. In order to convert mean fluorescence data to a more informative protein count, we used a combination of flow cytometry and quantitative Western blots (see Supplementary Information for details). We used a strain in which tagged YFP and CFP are equally produced from the $P_{LtetO-1}$ promoter to determine a scaling factor with which to directly compare YFP and CFP mean fluorescence levels as reported by the flow cytometer. We then employed Western blots to determine a combined mean number of tagged YFP and CFP molecules per cell, for various induction levels. The combination of these two approaches provides the numerical information necessary to translate the typical "arbitrary units" to a number of proteins and yields a balance point value of 7.6K tagged proteins produced (and degraded) per minute. While it is difficult to ascertain the precise number of tagged proteins in healthy or stressed *E. coli* cells, recent studies have provided lower limits that indicate that our calculated balance point is physiologically relevant. While healthy cells are believed to be in the underloaded regime(Moore and Sauer, 2005), in which they would not possess queues of excess tagged proteins, it is likely that stressful conditions, like DNA damage, will push the degradation system far beyond the balance point(Neher et al., 2006), causing correlations between many proteins as a result. In fact, this may be a beneficial design feature of some stress response networks, where the degradation delay allows the required proteins to build up rapidly for immediate response but to be removed and recycled quickly after the system is repaired.

The favorable comparison between model and experiment indicates that queueing theory provides a quantitative approach for describing coupled enzymatic processing (Figs. 2b–d, solid lines). The model fits were derived through use of a fitting algorithm to determine model parameters μ, γ, K, a Hill function-based parameterization

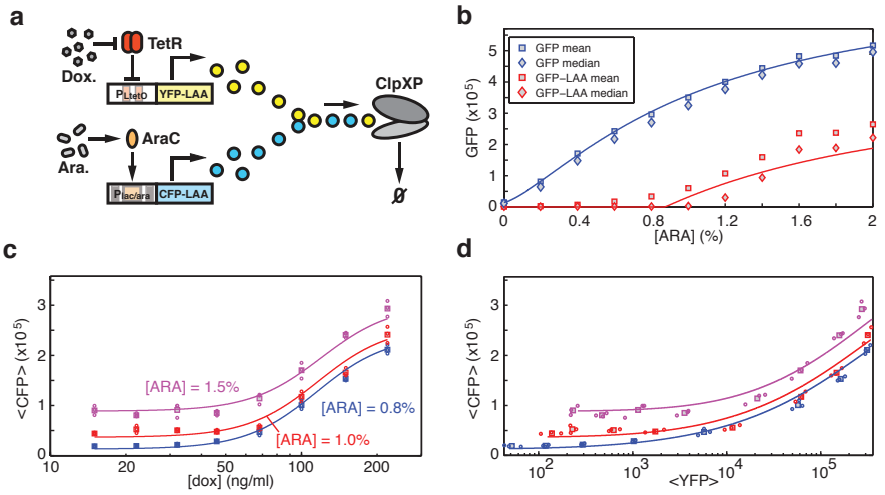


Figure 5.2: Coupled enzymatic degradation of yellow and cyan LAA-tagged fluorescent proteins by ClpXP machinery in *E. coli*. (a) Schematic network diagram of the synthetic circuit. YFP is produced by the $P_{LtetO-1}$ promoter, which is repressed by TetR in the absence of doxycycline. CFP is produced by the $P_{lac/ara-1}$ promoter, which is activated by AraC in the presence of arabinose. Both CFP and YFP molecules are tagged with identical LAA tags and are targeted for degradation by the ClpXP complexes. (b) Induction plots for a single fluorescent protein produced by the $P_{lac/ara-1}$ promoter. Blue and red symbols indicate untagged and LAA-tagged protein, respectively. Squares are mean protein counts, while diamonds are median protein counts. Solid lines are steady-state model fits to the data (including that in Figure c–d). The red line stochastic queueing model prediction for enzymatic protein degradation compares favorably with the data. (c) Mean steady-state expression of CFP as a function of doxycycline concentration at three different levels of arabinose in triplicate flow cytometry measurements. Strong coupling is observed between CFP and YFP. Protein counts are reported using a combination of two-color flow cytometry and Western blots. (d) The means of CFP and YFP increase simultaneously as the doxycycline concentration is increased. The color of the symbols corresponds to panel (c). In both (c) and (d), results for the stationary state of a fitted queueing model are included as solid lines. Supposing a doubling time $\tau_d \approx 30$ min, we find an enzymatic degradation rate $\mu = 4.6 \times 10^4$ min^{-1} for the model provides a good fit to the plotted results. Values for the production rates of YFP at given dox concentrations and for the production rates of CFP at given arabinose concentrations were determined from a best fit to the data. The qualitative result of this fit is that CFP only appears to activate when YFP becomes comparable in magnitude, consistent with a slightly overloaded queue. See Supplementary Information for fitting details and parameters.

for λ_1 (production rate of YFP for a given doxycycline), and the parameterization for λ_2 (production rate of CFP for a given arabinose). Interestingly, although the model involves many parameters, the shape of the fitted curves depends only on a few parameters, particularly the value of μ, the enzymatic degradation rate (see Supplementary Information). This indicates that the model describes the system well and that we are not over-fitting the data.

In order to further investigate the implications of enzymatic queueing, we designed a microfluidic platform to drive and monitor the signaling responses of the two networks at the single-cell level. We drove production of YFP with a periodic square-wave signal of doxycycline (see Methods Summary) and used two-color microscopy to observe the response of both the YFP and CFP signals. Whole-field fluorescence of a population of *E. coli* cells demonstrates how the coupling of the two proteins through the shared degradation pathway serves to drive one system in response to the behavior of the other, as both the YFP (Figure 3a, green) and CFP (Figure 3a, blue) signals oscillate with the driving signal (Figure 3a, red). Similar trends can be observed in the fluorescence trajectories of individual cells (Figure 3b). Correlated behavior between the two reporters is observed in response to the driving signal as well as in long-term expression trends (Figure 3c).

Based on these results, we hypothesized that the effect of queueing through coupled enzymatic processing could have significant implications in the developing field of synthetic biology. As most synthetic biology approaches employ degradation tags in order to enforce the necessarily quick turnover of key network components, waiting times for processing by ClpXP (or similar machinery in other biological systems) may be exploited to generate synchronized behavior amongst seemingly independent circuits. To test this hypothesis, we constructed a synthetic two-color system involving two independently produced fluorescent proteins: one (GFP-LAA) produced periodically by a synthetic oscillator, and the second (CFP-LAA) produced by a separately tunable promoter (Figure 4a). We built a new single-plasmid version of a previously reported synthetic oscillator (Stricker et al., 2008b) with the activator (araC-LAA), re-

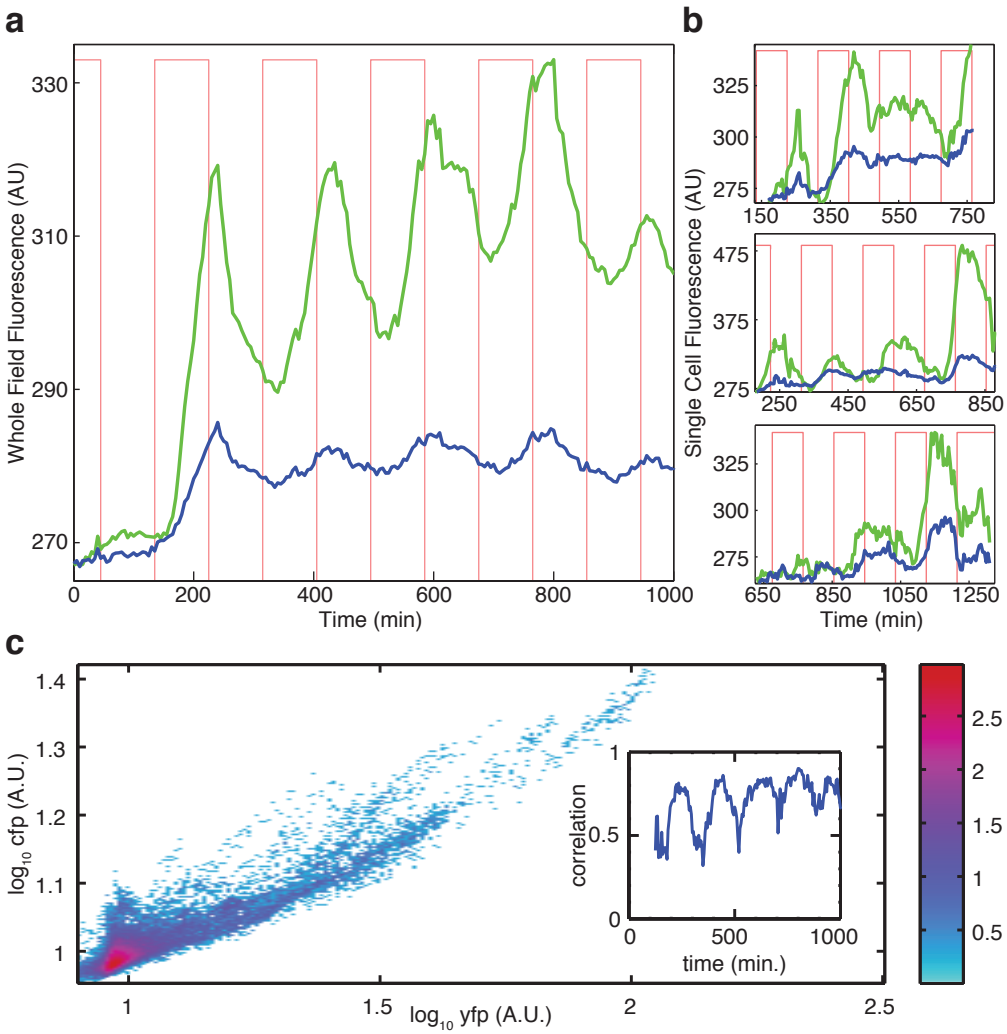


Figure 5.3: Dynamic behavior of a synthetic signaling network. (a) Using a microfluidic platform capable of generating a time-dependent induction signal and multi-color single cell fluorescence measurements, a large population of *E. coli* expressing the synthetic network were subject to a periodic series of doxycycline pulses (red). The total YFP and CFP fluorescence, integrated over the entire colony, demonstrates the direct response of the $P_{LtetO-1}$ promoter, producing YFP (green) as TetR is periodically deactivated, as well as the indirect response of the CFP signal (blue) due to the time-dependent saturation of the ClpXP machinery. (b) Trajectories for several individual cells demonstrate the response at a single-cell level. (c) A two-dimensional histogram depicts correlations between YFP and CFP levels in individual cells throughout the entire experiment duration. The value in each rectangular bin is log scaled (value of $\log_{10}(1+n)$, with n the number of counts in a bin). Inset shows the Pearson correlation coefficient between the two reporters as a function of time.

pressor (LacI-LAA), and reporter (GFP-LAA) all on a p15A plasmid. For the second component, we placed the *LuxI* promoter driving CFP-LAA along with a constitutively produced LuxR on a high copy pUC19 vector, such that the expression of CFP is inducible by the addition of AHL to the medium.

We loaded this system into our microfluidic platform and imaged cells for several hours without AHL, and we found regular oscillations of GFP-LAA with an average period of 28 minutes, while CFP-LAA levels were negligible. However, following induction with 15nM AHL, the mean levels of CFP and GFP both increased, and CFP trajectories rapidly became correlated with GFP trajectories. Some cells produced irregular oscillations of both colors with much longer periods (Figure 4b), while others stopped oscillating altogether. These results indicate that the additional production of tagged proteins by an independent circuit had a dramatic effect on the behavior of the synthetic oscillator, due to an increased burden on the degradation machinery.

The steady-state induction characteristics and dynamic coupling observed in the two synthetic systems, provide unambiguous evidence of how queueing for common enzymatic processing can induce coupling when there is a significant abundance of proteins relative to the number of functional enzymes. If such behavior were observed in a native or uncharacterized system, it would likely be assumed that these two proteins were coupled in a direct way, such as by coordinated gene expression or a protein-protein interaction. As large-scale wiring diagrams have become a post-sequencing paradigm for depicting the interactions between vast and divers cellular species, a major challenge is the deduction of the molecular interactions that underlie the observed correlations. Our results demonstrate that strong correlated behavior can be observed between two components that are only indirectly coupled via an overloaded enzymatic process and suggest that indirect coupling sources should be considered when evaluating native systems or designing and constructing functional synthetic networks.

By taking a quantitative approach to determining the number of tagged proteins observed in our experiments instead of the typical “arbitrary units,” we were able to generate a more precise numerical queueing model and to approximate the true “balance

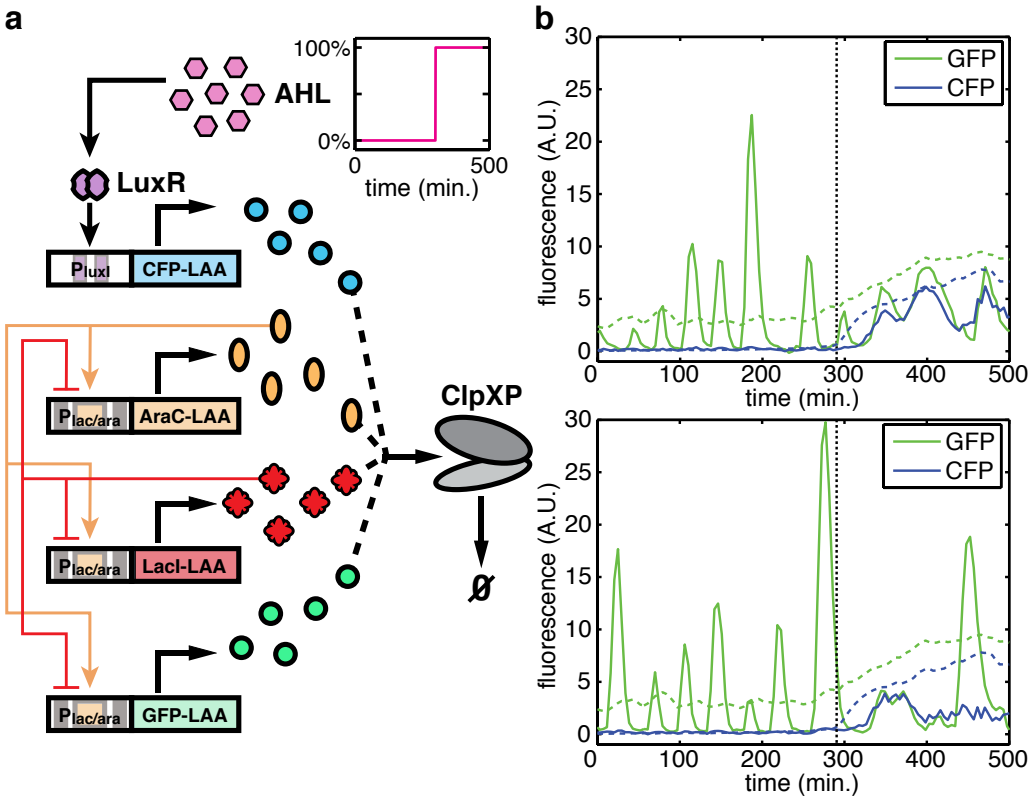


Figure 5.4: Coupled degradation can serve to indirectly couple gene circuits. (a) Circuit diagram for a variant of a synthetic gene oscillator discussed previously(Stricker et al., 2008b), expressed alongside an AHL-inducible, LAA-tagged CFP (see Supplementary Information for details). (b) Cells containing the circuit in (a) were grown in a microfluidic device to test the influence of CFP production on GFP oscillations (see Supplementary Information for details). Two single cell trajectories for GFP and CFP fluorescence (solid lines) show regular oscillations in GFP fluorescence in the absence of external AHL, i.e. at low CFP fluorescence. However, addition of 15 nm AHL (time of induction start is indicated by a vertical dotted line) introduces CFP into the system, causing the GFP oscillations to slow and the CFP signal to oscillate as a result of indirect coupling due to queueing. Coupling is also observed in the mean fluorescence across a region of cells (mean fluorescence as dashed lines), where increasing mean CFP fluorescence is associated with an increase in mean GFP fluorescence.

point" in the natural system, above which the number of tagged proteins will likely lead to correlated behavior. While healthy cells are believed to be in the underloaded regime, in which they would not possess queues of excess tagged proteins, it is likely that stressful conditions, like DNA damage, will push the degradation system far beyond the balance point, causing correlations between many proteins as a result. In fact, this may be a beneficial design feature of some stress response networks, where the degradation delay allows the required proteins to build up rapidly for immediate response but to be removed and recycled quickly after the system is repaired.

Similarly, it may be beneficial to exploit these indirect links when designing new synthetic systems. As many of the interesting studies in synthetic biology focus on the creation of circuits that exhibit precise, dynamical behavior, targeted degradation of key network components has become an almost essential characteristic of synthetic systems. Since our study demonstrates how the use of degradation tags can lead to unexpected correlations, the results will have important implications on efforts to establish a "forward engineering" methodology. That is, it may be both critical to consider coupled degradation in the modeling of genetic circuits as well as potentially interesting to intentionally incorporate waiting lines into the design criteria of novel synthetic systems to enhance the desired behavior.

Strains and Flow Cytometry

The plasmids were constructed using the pZE24-mcs2a cloning plasmid, which has a kanamycin resistance marker and the hybrid $P_{lac/ara-1}$ promoter upstream of a multiple cloning site (mcs)(Lutz and Bujard, 1997b). For the dual-color plasmid, pNO-2cLAA, the sequence for CFP was tagged by PCR with a carboxy-terminal ssrA tag (AANDENYALAA)(Keiler et al., 1996) and inserted between the KpnI and HIndIII sites of the mcs, creating pZE24-CFP-LAA. The YFP fragment was similarly tagged and inserted onto the pZS31-mcs2a cloning plasmid, which contains a chloramphenicol resistance marker and the $P_{LtetO-1}$ promoter upstream of a multiple cloning site.

The fragment of this plasmid containing the marker gene and $P_{LtetO-1}$ driving YFP was copied by PCR and inserted into the SacI site of pZE24-CFP-LAA, creating the final plasmid pNO-2CLAA, containing the independently controlled fluorescent proteins. The plasmids for comparing the tagged and untagged GFP induction curves were constructed similarly on the same plasmid backbone.

For the synthetic oscillator coupled to a constitutive system, the medium copy oscillator (pTDCL8) was constructed by combining oscillator components from pJS167 and pJS169 (Stricker et al., 2008b) onto a single p15A plasmid. The pJS167 module (from SacI up to AvrII) was copied by PCR with Kan resistance and flanking AvrII and NheI sites, and it was inserted via ligation at the AvrII site of the pZA14-LacI vector, maintaining uniqueness of the AvrII site and creating an undigestable AvrII-NheI hybrid site. The pLuxI-CFP plasmid (pZU25-CFP-LAA) was constructed by building pZE25-CFP-LAA and inserting the promoter + CFP module by PCR onto a pUC19 vector via flanking AvrII sites.

Flow cytometry data was taken with a Becton-Dickinson LSR II Cell Analyzer, fitted with 405nm and 488nm lasers. The cells were grown overnight in non-inducing medium: LB plus kanamycin for plasmid selection. The cells were passed in the morning into LB plus kanamycin plus various levels of inducer, either doxycycline, arabinose/IPTG, or both. The cells were grown in a 37°C shaker at 300rpm. After 3 hours, the cells were harvested by centrifugation, resuspended in sterile PBS, and put on ice until they were ready to be sampled. Using the LSR II, 100,000 cells were assayed using MATLAB (The MathWorks, Inc.) and interfacing software.

Microfluidics and Microscopy

Image acquisition was performed on a Nikon Eclipse TI epifluorescent inverted microscope outfitted with fluorescence filter cubes optimized for YFP and CFP imaging and a phase-contrast based autofocus algorithm. Images were acquired using a Photometrics CoolSNAP HQ2 cooled CCD camera, controlled by Nikon Elements software.

For the signaling experiment, images were acquired every 1 minute in phase contrast, in order to provide the optimal temporal coverage to suit the automated tracking program. Fluorescent images in the CFP and YFP channels were acquired every five minutes. The cells were imaged inside a microfluidic device with an upstream switch, with the ability to mix or switch between two different media sources. A custom application written in LabVIEW (National Instruments, Austin, Texas) controlled linear actuators, to which two reservoirs containing inducing and non-inducing medium were attached. Using this algorithm, a square wave of 3 hour period was generated, subjecting the cells to alternating 90 minute terms of 220 ng/ml doxycycline and 0 ng/ml. For the synthetic circuit experiment, CFP-LAA was induced with 15nM AHL (N-3-Oxo-hexanoyl-homoserine lactone) from the *LuxI* promoter as described in the main text. Cells were imaged every 30 seconds in brightfield and every 4.5 minutes in both CFP and YFP channels. Controls were performed to ensure that there was not significant overlap between the CFP and YFP channels, even when imaging CFP and GFP.

The microfluidic experiments were performed as previously described (Danino et al., 2010). Briefly, 50uL of an overnight culture was diluted in 50mL of LB (Difco) + antibiotics the day of the experiment. When cells reached an OD_{600} of 0.1, cells were spun down and resuspended in 5mL of fresh media and loaded into the device. Image processing was performed using a custom application tied into the ImageJ image processing suite. Images were segmented by creating a binary mask to identify individual cells, and cells were tracked from frame to frame using a combination of quantitative characteristics. The output of this process is the single cell trajectory information provided in the main text.

Flow Cytometry

Flow cytometry data was taken with a Becton-Dickinson LSR II Cell Analyzer, fitted with 405nm and 488nm lasers. Because a wide range of fluorescence intensities arose as inducer levels were scanned, the sensitivity (set by voltage) of photomultiplier

tubes in the flow cytometer was varied from sample to sample. We calibrated the photomultiplier tubes by scanning a range of voltages for cells with a constant mean level of fluorescent protein, either YFP or CFP alone. We fit the resulting mean YFP and CFP intensity curves to piecewise-smooth functions of voltage. These functions were used to correct flow cytometry data by scaling all measurements to a common apparent voltage.

In order to be able to directly compare numbers of fluorescent proteins of YFP and CFP, we used a plasmid containing two copies of the $P_{LtetO-1}$ promoter, one driving YFP-LAA and one driving CFP-LAA (pZA11-YC-LAA). This was constructed using similar techniques to those described above. Using this strain and the assumption that the two proteins should be produced in equal mean levels due to their tandem arrangement on the same plasmid, we induced cells at various concentrations of doxycycline and measured mean fluorescence. We scaled CFP fluorescence such that the mean CFP fluorescence values at the selected induction levels were essentially the same as the corresponding YFP mean fluorescence values (difference of 0.5% in the typical mean fluorescence), and we were able to determine a conversion factor to compare YFP "arbitrary units" (AU) to CFP AU. This, combined with the Western blot data (see below) enabled the estimation of the total number of each fluorescent protein in each data set.

As another validation of the queueing theory, we compared this dual-color induction data to an almost identical plasmid, but one in which only YFP-LAA was produced from the $P_{LtetO-1}$ promoter. Interestingly, we found that the overall level of YFP fluorescence was significantly lower in the case of expression of only one of the two colors. This falls in line with queueing theory predictions. That is, in the dual-color case, the total number of tagged proteins is doubled, so ClpXP would be more overloaded than when just a single fluorescent protein type is produced. This is further evidence that over-burdening the protease can lead to coupling between the levels of two different proteins.

As described in the main text, we also compared induction curves of tagged and untagged fluorescent proteins, in an effort to determine if the queueing effect could be directly observed in induction data. This also served to demonstrate that the effects

observed throughout the experiments were in fact due specifically to the abundance of tags, and not simply side effects of general over-expression. To take this data, we created two plasmids very similar to those used for the two-color study. We used the pZE24-mcs2a cloning plasmid (Lutz and Bujard, 1997b), which has a kanamycin resistance marker and the hybrid $P_{lac/ara-1}$ promoter upstream of a multiple cloning site (mcs). We simply replaced the mcs with either GFP or GFP-LAA, and used this to take induction data as described in the main text.

For all flow results, a background subtraction procedure was performed on the raw data (after the voltage correction described above) to arrive at reported YFP and CFP fluorescence statistics. Using data from the experiments discussed in the main text, we defined the background mean of YFP and CFP fluorescence as that derived from cells induced with 1 mM IPTG alone (in the absence of arabinose or doxycycline). The background mean for each color was subtracted from the mean of raw data.

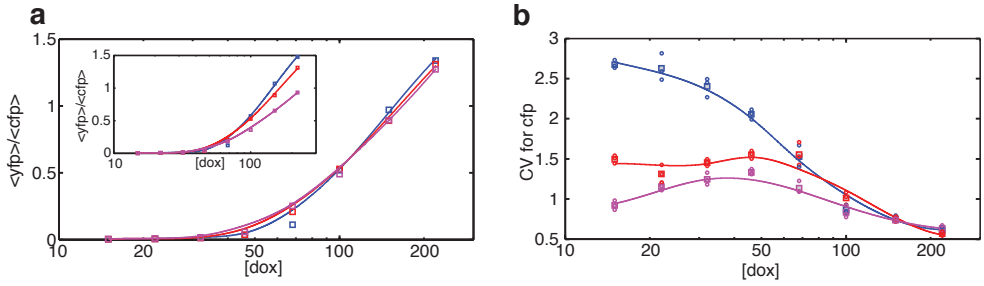


Figure 5.5: Coupled enzymatic degradation of yellow and cerulean LAA-tagged fluorescent proteins by ClpXP machinery in *E. coli*. (a) According to the stochastic queueing model, the ratio of the two mean concentrations $\langle x_1 \rangle / \langle x_2 \rangle$ is equal to the ratio of the corresponding production rates λ_1/λ_2. In accordance with the model, the ratio $\langle \text{yfp} \rangle / \langle \text{cfp} \rangle$ exhibits the same dependence on dox concentration for three different levels of arabinose, which allowed us to collapse all the data to a common curve by normalizing them by the mean value over the whole range of employed dox concentrations. Inset shows the same data without collapsing. (b) Coefficient of variation of CFP concentration decreases with increasing dox in the overloaded regime in qualitative agreement with the queueing theory predictions (different symbols correspond to three levels of arabinose concentration similar to panels c-d of Fig. 2 in the main text). Solid lines represent trend lines through the data.

As one further test of the stochastic queueing model, we used the theoretical results to deduce the scaling relationship between the data sets in Fig. 2c of the main text at

different levels of arabinose. In other words, if correct, the theory can be used to predict how one can plot the data such that it will collapse onto the same curve. The resulting verification of this prediction further confirmed the general validity of the queueing theory approach (Fig. 5.5a). Lastly, we calculated the noise (as measured by the coefficient of variation) of the CFP signal as a function of increasing doxycycline (Fig. 5.5b). The general trends of these curves are also in agreement with the theoretical predictions (compare Fig. 5.5b with Fig. 1c of the main text).

We do not investigate the effect of removing SspB (Levchenko et al., 2000b), a protein associated with increased affinity of tagged proteins to ClpXP. We anticipate a moderate decrease in this Michaelis-Menten affinity would not qualitatively change our conclusions.

Protein Counts

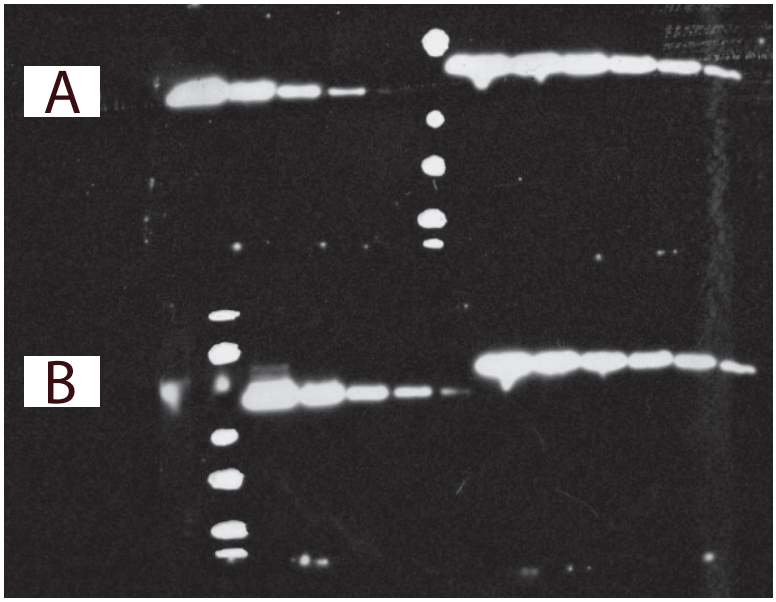


Figure 5.6: Inverted image of Western blot film taken for cellular lysate data from two induction levels, where IPTG and arabinose were held at 1 mM and .8%, respectively, and doxycycline was 32 ng/ml (A) and 68 ng/ml (B). An antibody for GFP variants was used to detect the total amount of CFP and YFP inside these samples, when compared to a purified GFP standard (left five lanes of both A and B).

Western blots were performed using standard techniques in order to quantify the number of tagged fluorescent proteins being measured in our flow cytometry data. As a standard, we used purified Enhanced GFP (BioVision 4999-100), supplied in a 1 mg/ml 100 μl aliquot. We chose to measure protein levels in a sample of cells expressing the pNO-2CLAA plasmid, induced to various levels of dual-color expression. For all samples, we used 1 mM of IPTG and 0.8% arabinose. Samples were induced exactly as done for the flow cytometry experiments. Cells were grown overnight without inducers, and then passed 1:1000 into inducers for 3 hours. In order to obtain enough protein for quantitative detection, 50 ml of each sample was harvested by centrifugation after 3 hours. ODs at 600 nm were taken just before centrifugation, in order to quantify cell number (see below).

After centrifuging the samples and aspirating the inducing medium, the cells were resuspended in 100 μl of SDS sample buffer to aid in cell lysis by boiling. The total volume after resuspension was measured in order to obtain an accurate measurement of cell concentration. Cells were then lysed and proteins denatured by subjecting the samples (both lysates and standards) to boiling water for 5 minutes. A 12 lane 12% Tris-Glycine gel was used in order to have enough lanes for a sufficient dilution series of both the cell lysate sample as well as the standard. The standard was diluted to a concentration of 100 ng/μl, and fives samples were loaded on the gel in subsequent 2-fold dilutions, starting with 500ng. Similarly, the cellular lysate was loaded in subsequent 2-fold dilutions, starting with a volume of 15 μl. The gel was run at 125 V for about 100 minutes, followed by a membrane transfer run at 25 V for 90 minutes.

Standard blocking and probing reactions were set up using a GFP polyclonal rabbit antibody (Cell Signaling 2555S) and an Anti-Rabbit IgG (whole molecule) - Peroxidase antibody. Control experiments were performed to ensure that this antibody binds with equal affinity to GFP, CFP, and YFP. After exposing the membrane to the Chemiluminescent Peroxidase Substrate (Sigma, CPS-1), Kodak BioMax Light Film was exposed to the membranes in a dark room for 60 seconds. Once a satisfactory image was taken, processing was performed using ImageJ. Background correction was

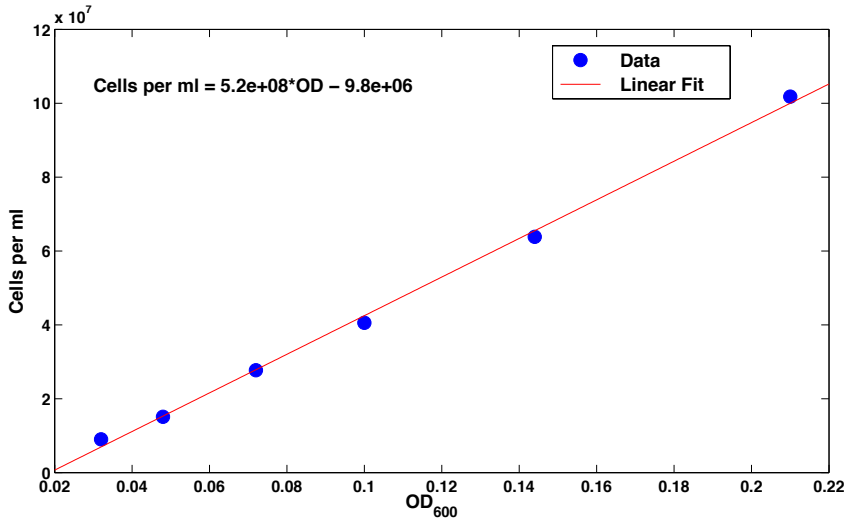


Figure 5.7: Cell count was correlated with OD by taking several measurements throughout logarithmic growth of each. A linear fit of this data was then used to convert the OD of Western blot samples to a particular concentration of cells.

performed to remove some of the background coloration from the image. The image was then inverted so that bands showed as white on a black background (Fig. 5.6), and the freehand selection tool was used to quantify the total intensity of each band. Comparing the total intensity of each standard band to the known protein mass loaded on the gel, we were able to obtain a function to convert band intensity to protein mass, and this was used to quantify several lanes of the cell lysate samples that fell within a linear range (where the known two-fold dilution matched a two-fold drop in band intensity). Finally the protein weight measurement was converted to a total protein count per cell, using the weight of a single protein and the number of cells loaded onto the gel.

Cell counts were done using a hemacytometer. Cells were grown in inducing conditions and sampled every 20 minutes over a 3 hour period around the OD sampled for the Western blot data. Cell count was plotted vs OD over this range and a good linear fit was achieved (Fig. 5.7). Using this linear fit, we were able to calculate cell count for the ODs at which our cultures were sampled for the Western blots.

Experimental Controls

Several control experiments were performed to ensure that the coordinated behavior in our two experimental systems was due to degradation-based coupling and not just an artifact due to some other phenomenon. First, the comparison of the tagged and untagged induction curves in the main text served to demonstrate that the effects we are seeing are due to the tag and not just side effects of over-expression. In the case of the dual-tunable signaling network, another primary concern was ensuring that the two inducers did not interfere with the other promoters (i.e. that there is no cross-talk between the two promoters). To test this, we induced cells with each of the two inducers independently, and ensured that each color was only induced by the appropriate inducer (Fig. **??**). In panel a, it is clear that doxycycline strongly induces YFP (red) and not CFP (blue). When arabinose is introduced in addition to doxycycline, the YFP levels increase (green), as CFP is now being produced as well, causing an increased burden on ClpXP. The reverse is true as well; in panel b, arabinose alone is seen to induce CFP (blue) and not YFP (red), however the addition of doxycycline causes increased levels of CFP. This is further evidence of queueing theory, in addition to a good control for crosstalk between the two promoters.

A similar control was performed for the dual-color synthetic circuit experiment. That is, the oscillator strain was induced with AHL, and we saw no effect on period of the oscillator, indicating that AHL and LuxR do not interfere with the ara/lac promoter. As another general control that the addition of tagged proteins causes the observed effect, as opposed to it being some other artifact of over expression, we tested the synthetic oscillatory system in conjunction with a high level of a general, untagged protein (for this purpose, we used Pn25 driving TetR, untagged on a p15A plasmid). When producing a large amount of untagged TetR (approximately 100,000 copies per cell (Lutz and Bujard, 1997b)) along side the oscillator, we saw no difference between this and the normal behavior of the oscillator. This control provides evidence that there is no apparent effect on period when expressing untagged proteins. In addition, the experi-

mental acquisition of the two induction curves (described above), with the tagged and untagged versions of GFP, served to demonstrate that the effects observed throughout the experiments were in fact due specifically to the abundance of tags, and not simply side effects of general over-expression.

Stochastic Theory - Model

The model considered in this paper involves production and degradation of protein types X_i, where indices $i = 1, 2, ..., m$ identify different protein types. Degradation occurs by the protein binding to a protease P and subsequently being annihilated. Specifically, the model reactions are (rates are rate constants, not including mass action terms)

$$D_i \xrightarrow{\lambda_i} X_i + D_i \tag{5.1}$$

$$X_i \xrightarrow{\gamma} \emptyset \tag{5.2}$$

$$X_i + P \underset{\eta_-}{\overset{\eta_+}{\rightleftharpoons}} X_i P \tag{5.3}$$

$$X_i P \xrightarrow{\mu} P \tag{5.4}$$

$$X_i P \xrightarrow{\gamma} P \tag{5.5}$$

where DNA D_i produces protein X_i with rate constant λ_i, X_i is diluted (due to cell growth and division) with rate constant γ, X_i binds to the protease P with rate constant η_+, X_i unbinds from P with rate constant η_-, and P degrades X_i with rate constant μ. Reactions occur with exponentially distributed times. For simplicity of results, we assume that dilution can act on X_i bound to P, though results can be generalized to when dilution does not act on X_i bound to P. We also assume the presence of a single protease, though the results can be generalized to many proteases. We assume DNA $D_i = 1$ for simplicity.

Using reasonable approximations, we can further simplify Eqs. 5.1–5.5. The simplest approximation is to suppose that $\eta_- \approx 0$ and that η_+ is large, such that the reactions in Eqs. 2 and 5 collapse to Eq. 2 and Eqs. 5.3–5.4 combine into a single degradation reaction, where the protease chooses one particular protein and degrades it at rate μ; the latter has the same steady-state behavior as when Eqs. 3-4 are replaced

by the reaction

$$X_i \xrightarrow{\mu/n} \emptyset \tag{5.6}$$

where $n = \sum_{j=1}^{m} x_j$, and x_i is the count of protein type X_i, as in the main text. Similar results can be derived if both η_+ and η_- are sufficiently large. This leads instead to the approximate degradation reaction

$$X_i \xrightarrow{\mu/(K+n)} \emptyset \tag{5.7}$$

where K is a Michaelis-Menten parameter. In the limit $K \to 0$, the $\eta_- \approx 0$ system is recovered. More details concerning the motivation and derivation of the reduced rates in Eqs. 5.6–5.7 appear in Refs. (**??**).

The reduced system, using either Eq. 5.6 or Eq. 5.7 for the enzymatic degradation reactions Eqs. 5.3–5.4, can be mapped onto a stochastic queueing model. One such queueing model places each new X_i at a random position in a single queue, while P processes (degrades) the protein at the head of the queue. Dilution can be added by allowing "reneging," whereby any member of the queue (including a member being processed by the server) leaves with rate γ.

Stochastic Theory - Results

We have carefully derived several relevant results for the above model in another study (**?**), which applies the theory of multiclass queueing in the context of gene regulation. One key result is that the steady state probability distribution $P(\{x_i\})$ for the set of counts $\{x_i\}$ can be factored into $R(n)$, the probability distribution for the sum, times a multinomial distribution:

$$P(\{x_i\}) = R(n)\, n! \prod_{j=1}^{m} \frac{p_j^{x_j}}{x_j!} \tag{5.8}$$

where $p_i \equiv \lambda_i / \sum_{j=1}^{m} \lambda_j$. From this, it can be shown that moments of x_i are given in terms of moments of n. In particular,

$$\langle x_i \rangle = p_i \langle n \rangle \tag{5.9}$$

$$\sigma_i^2 \equiv \left\langle x_i^2 \right\rangle - \langle x_i \rangle^2 = p_i(1-p_i)\langle n \rangle + p_i^2(\left\langle n^2 \right\rangle - \langle n \rangle^2) \tag{5.10}$$

The moments of n are less general and will depend on the particular model. With the reaction scheme Eqs. 5.1–5.2, 5.7, we find

$$\langle n \rangle = \frac{\alpha\delta M(\alpha+1, \beta+1, \delta)}{\beta M(\alpha, \beta, \delta)} \tag{5.11}$$

$$\left\langle n^2 \right\rangle = \langle n \rangle + \frac{\alpha(\alpha+1)\delta^2}{\beta(\beta+1)} \frac{M(\alpha+2, \beta+2, \delta)}{M(\alpha, \beta, \delta)} \tag{5.12}$$

with $\alpha \equiv K+1$, $\beta \equiv (\mu/\gamma) + \alpha$, $\delta \equiv \Lambda/\gamma$, $\Lambda \equiv \sum_{i=1}^{m} \lambda_i$, and $M(\,\cdot\,,\,\cdot\,,\cdot\,)$ the confluent hypergeometric function of the first kind.

A direct consequence of Eq. 5.9 is

$$\frac{\langle x_i \rangle}{\langle x_j \rangle} = \frac{\lambda_i}{\lambda_j} \tag{5.13}$$

which holds for all levels of expression. Figs. 2c-d of the main text and Fig. S2a test the validity of Eqs. 5.9–5.13 against experimental data.

Deterministic Theory

Though deterministic models do not address the many issues tied to noisy dynamics, e.g. correlations between the counts of the protein species, certain aspects of coupling through degradation can be understood using a deterministic analog of the above stochastic model. In the deterministic model, the concentrations x_i for different proteins X_i obey the ODEs

$$\frac{dx_i}{dt} = \lambda_i - \gamma x_i - \frac{\mu x_i}{K+n} \tag{5.14}$$

where λ_i is a production rate constant, γ is the dilution rate constant, K is a Michaelis-Menten molar constant, μ is the enzymatic processing rate constant, and $n = \sum_{j=1}^{m} x_j$.

Near so-called balance, i.e. when $\Lambda \approx \mu$, the x_i's strongly contract onto a slow manifold. The steady state solutions $x_i^{(ss)}$ to Eqs. 5.14 always satisfy a relation similar to Eq. 5.9

$$x_i^{(ss)} = p_i n^{(ss)} \tag{5.15}$$

where again $p_i \equiv \lambda_i/\Lambda$. For finite K

$$n^{(ss)} = \frac{\Lambda - \mu - K\gamma + \sqrt{(\Lambda-\mu)^2 + K\gamma(2\Lambda + 2\mu + K\gamma)}}{2\gamma} \tag{5.16}$$

which simplifies in the limit $K \to 0$:

$$n^{(ss)} = \frac{\Theta(\zeta)}{\gamma} \tag{5.17}$$

where $\zeta = \Lambda - \mu$, and $\Theta(\cdot)$ is the integrated Heaviside step function: $\Theta(\zeta) = \zeta$ if $\zeta \geq 0$ and $\Theta(\zeta) = 0$ if $\zeta < 0$. The solutions in Eqs. 5.15–5.17 reveal that components strongly interact when at least two λ_i's are simultaneously nonzero and $\Lambda > \mu$. Eqs. 5.15–5.16 are used to plot the deterministic results in Fig. 1b of the main text.

Finally, the deterministic model can be used to investigate signaling by means of varying one of the λ_i's while keeping the other λ_i's constant. Standard numerical integration of the ODEs leads to the solution presented in Fig. 1d of the main text.

Fitting of Steady State Model

Figs. 2b–d include fits of the model to the data. Below, we outline the procedure to obtain fitted model parameters.

The fitted model results presented in Figs. 2c–d of the main text were derived through use of a fitting algorithm to determine model parameters μ, γ, K, a Hill function parameterization for λ_1 (production rate of YFP for a given dox level), and a set of 3 values for λ_2 (production rate of CFP for a given arabinose level). At the end of this section, we revisit these best fit values of λ_2 to find they are in reasonable agreement with single fluorescent protein expression data.

Curve fitting was implemented by a Metropolis algorithm. The energetic penalty used for the algorithm was a weighted sum of the square distances between stationary state model results (ref. Eqs. 5.9 and 5.11) and mean fluorescence data points. Due to the wide range of YFP fluorescence magnitudes, we used linear distance when comparing CFP fluorescence and logarithmic distance when comparing YFP fluorescence.

Parameters γ and K were not important for our fitting. We scale time by the doubling time τ_d (approximately 30 min.), such that the value of the dilution rate is fixed at $\gamma = \ln(2)$ in natural units. We also found when fitting the stochastic queueing model that the system always appears, at least slightly, overloaded, allowing us to set $K = 0$ with little reduction in the goodness of fit.

We found that the deterministic queueing model's stationary state approximates the overloaded stochastic queueing model's mean values well, and so we used the deterministic model's results for rapid fitting of the data. Arbitrary precision calculations in the Maple 11 software package (Waterloo Maple Inc.) confirmed the stochastic model's mean values matched the deterministic model.

Using the data from the dox induction curves in Figs. 2c–d of the main text, λ_1 was fit to a shifted Hill function of the form

$$\lambda_1 = B_1 + D_1 \frac{(([\text{dox}]/C_1)^{H_1})}{(1 + ([\text{dox}]/C_1)^{H_1})}. \tag{5.18}$$

with $H_1 = 3.0782$, $B_1/\mu = 0.0023$, $D_1/\mu = 2.3429$, and $C_1 = 168.2114$ ng/mL. We did not fit λ_2 to a smooth curve, due to a small number of points being available, but we found best fit values $\lambda_2/\mu = 1.0373, 1.1093, 1.2683$.

Other parameters used for Figs. 2c–d in the main text are as follows: Using doubling time τ_d (approximately 30 min. for *E. coli*), $\mu = 7.589 \times 10^3$ min^{-1}, $\gamma = \ln(2)$, $K = 0$.

We tested consistency of the model fit in Figs. 2c–d by comparing results to the data in Fig. 2b. Here, we found a continuous curve fit for the mean untagged (slow degrading) GFP fluorescence multiplied by γ, providing a continuous parameterization of the apparent production rate of GFP. Before fitting, to account for the difference

between single molecule GFP and CFP fluorescence, we scaled GFP fluorescence such the three values of mean CFP fluorescence at low dox (in Figs. 2c–d) were closest in a least squares sense to the corresponding mean GFP fluorescences. We fit the apparent production rate $\lambda_2^* = \gamma < \text{gfp} >$ to the continuous function

$$\lambda_2^* = B_2 + D_2 \frac{(([\text{ARA}]/C_2)^{H_2})}{(1 + ([\text{ARA}]/C_2)^{H_2})}. \tag{5.19}$$

with $H_2 = 1.3660$, $B_2/\mu = 0.0665$, $D_2/\mu = 3.1039$, and $C_2 = 1.0323$ %. The difference $\Delta\lambda_2$ between λ_2 from panels c,d and λ_2^* from panel b are relatively minor, being $\Delta\lambda_2/\lambda_2 = 0.1093, 0.0144, -0.0850$, respectively, suggesting that the fits are consistent.

In Fig. 2b, using the parameterization λ_2^* and the model parameters determined by Fig. 2c–d, we present the prediction for mean protein count as a solid red curve. This prediction is in agreement with the data in Fig. 2b, suggesting that the data in Fig. 2b and in Figs. 2c–d are in agreement.

Though we found a small value of K (e.g. about a thousand or less) was consistent with our model fit, the effect of larger K on the stochastic queueing model at the balance point was considered (see Fig. 5.8).

Acknowledgements

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health (GM07933) and the Department of Energy Computational Science Graduate Fellowship [DE-FG02-97ER25308] (NAC). Research of RJW was supported in part by NSF grants DMS-0825686 and DMS-0906535. LT was supported, in part, by ONR MURI grant N00014-07-0741. We would like to thank Colin Lam, Ivan Razinkov, and Arthur Prindle for assistance in plasmid construction and Jangir Selimkhanov for developing the cell tracking code. Chapter 4 contains material originally published as Cookson, N.A.*, Mather, W.H.*, Danino, T., Mondragon-Palomino, O., Williams, R. J., Tsimring, L. S., and Hasty, J. ***Nature*** (in review): Queueing up for enzymatic processing: Correlated signaling through coupled degradation. (*equal

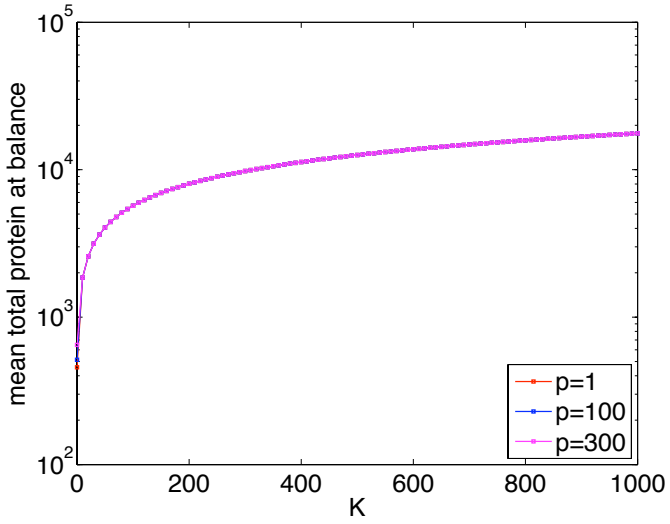


Figure 5.8: At balance, the mean total protein for the stochastic queueing model with finite Michaelis-Menten constant K. Each protease (copy number p) processes at reduced maximum rate μ/p and with Michaelis-Menten constant K. The balance condition is defined when the total protein production rate Λ equals the total processing rate μ, i.e. $\Lambda = \mu$. Other parameters are determined from the model fit in this section. We find the mean total protein at balance depends weakly on p for $K > 0$. A value of $K = 1000$, which leads to approximately 18-thousand proteins at balance, can be considered a moderate perturbation compared to the approximately 100-thousand proteins well above balance in experimental data. For the slightly underloaded condition $\Lambda = 0.9\,\mu$ (not shown), the mean protein level for $K = 1000$ is even less: approximately 7-thousand.

contribution). Copyright permission to republish here was granted by Nature publishing group.

Chapter 6

Summary and Conclusion

Synthetic biology seeks to engineer biological networks to with a quantitative dynamic function. Establishing design criteria for these gene networks backed by mathematical modeling is key to building robust biological networks. And key to building mathematical models is the ability to test on simple gene networks with technologies that allow for single-cell trajectories over time. In this thesis I have described the steps of modeling, constructing, and characterizing simple synthetic gene oscillators as a model system for an engineered biological network. In coupling synthetic gene oscillators to other systems, we also revealed limits to building large biological networks and hidden coupling that arise. Finally we looked at a simple genetic network that produces spatial patterns as well and simulated these in three dimensions. These works highlight the push to understand and simplify engineering of biological circuits for use in applications of synthetic biology. In the future, these robust oscillators will be used to couple to other biological networks for various applications.

Chapter 7

Lab Procedures and Protocols

Tal Danino, Hasty Lab, Bonner 2121

Winter 2009-Spring 2011

Table of Contents:

1. Overview of Molecular Cloning
2. QIAprep Miniprep
3. NanoDrop Spectrophotometer
4. Agarose Gel Electrophoresis
5. Gel Digests
6. Gel Purification
7. Ligations
8. Transformations
9. PCR protocol
10. CPEC Reactions

11. PCR Purification
12. Sequencing DNA (Eton Bio's service)
13. Colony PCR
14. PCR optimization
15. Designing PCR Primers for CPEC
16. Screening colonies from CPEC Reactions
17. Making Competent Cells
18. Making PDMS chips for Microscope Experiments
19. Flow Cytometry
20. Preparing Electrocompetent cells and Electroporations

Overview of Molecular Cloning

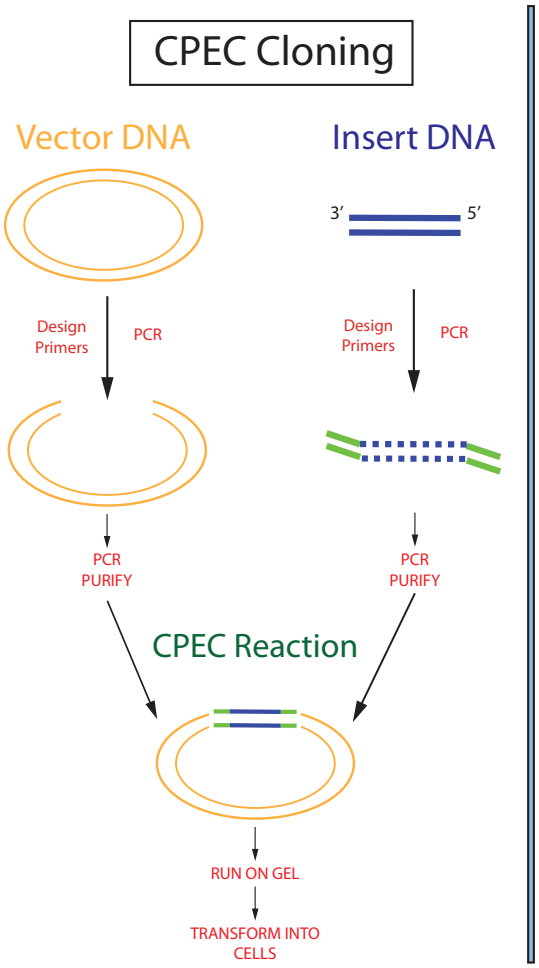


Figure 7.1: Overview of CPEC cloning. The vector and insert pieces are generated by PCR, gel extracted or PCR purified, and then run together in a PCR/CPEC reaction to assemble the full plasmid.

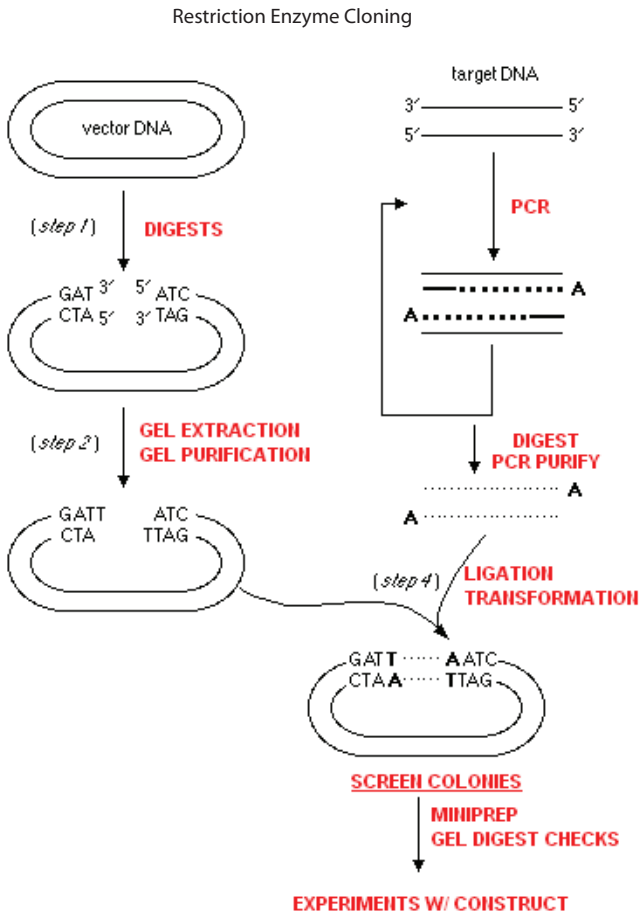
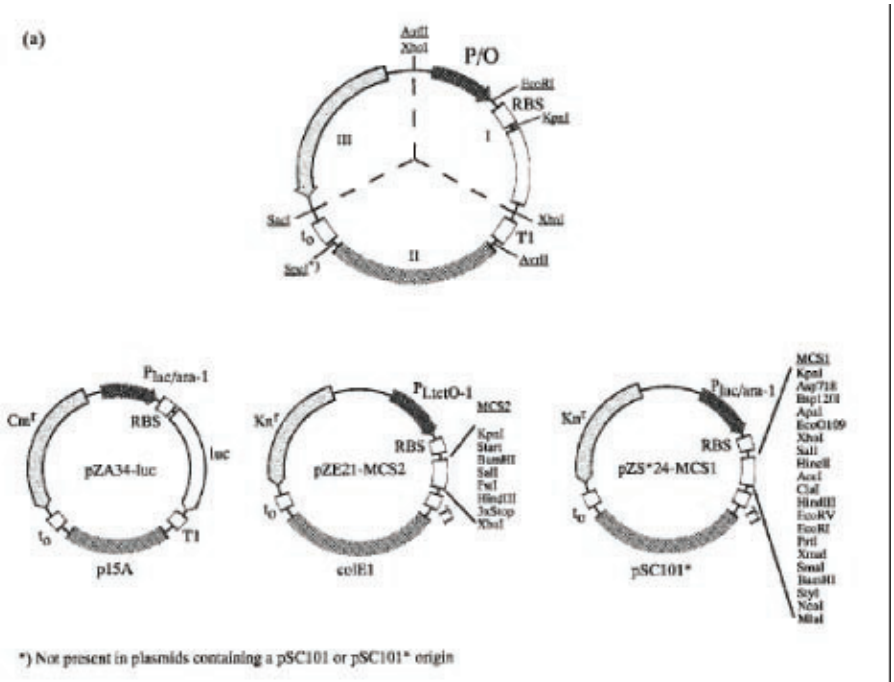


Figure 7.2: Overview of restriction enzyme cloning. Vector piece is generated by digestion with enzymes to create sticky ends, then ligated to a digested PCR/digested insert and assembled together.



(b)

origin of replication		resistance marker		regulatory unit		designation of vector
E	ColE1	1	Ampicillin	1	$P_{LtetO-1}$	pZE11-
A	p15A	2	Kanamycin	2	$P_{LlacO-1}$	pZA22-
S	pSC101	3	Chloramphenicol	3	$P_{A1lacO-1}$	pZS33-
S*	pSC101*	4	Spectinomycin	4	$P_{lac/ara-1}$	pZS*44-
		5	Zeocin	5	pluxI (w/luxR)	pZS*35sspB
				6	pN25	pN25tetR

Figure 7.3: Naming convention of pZ plasmids. Figure taken from Lutz and Bujard 1997 explaining the pZ naming convention

Expanded Naming Set of pZ plasmids (Tal Danino, 4/27/2011) [pZ Plasmid System]pZ Plasmid System Nucleic Acids Res. 1997 Mar 15;25(6):1203-10.

Sm = pSC101m (Jesse's higher copy variation , 10-15 copies)(note this looks more like pSC101 than pSC101*)

7=Parsenic promoter

8=PL promoter (lambda)

QIAprep Miniprep

Designed for the purification of up to 20 µg high-copy plasmid DNA from 1-5 ml overnight *E.coli* culture in LB medium.

14mL culture tube protocol: (Convenient when doing many minipreps)

1. Spin down culture tubes in big centrifuge for 5 mins. [14mL culture tubes can be spun down in big centrifuge up to 3g]
2. Pour supernatant (&tip) into bleach flask, being careful to remove all liquid. If liquid remains in tube then use pipette to remove excess.
3. Add 250 uL buffer P1 to each tube.
4. Place tubes on rack and vortex entire rack slowly till resuspended (watch that fluid doesn't reach top of tube, usually setting 4-6)
5. Transfer 300uL to 1.5 mL microcentrifuge tube.

This replaces steps 1-4 below.

Original Protocol:

· **NOTE**: Use Blue spin columns; centrifuge rpm 14000

1. Pipette 1.5ml from the test tube used for growth and transfer to clean 1.5ml microcentrifuge tube.
2. Spin down cells for 9-10 seconds @ 14000 rpm and remove LB medium.
3. Add 1.5 mL of culture again & repeat.
4. Re-suspend pelleted bacterial cells (via vortexing) in 250 µl Buffer P1 (located in right-most fridge)
5. Add 250 µl Buffer P2 and mix GENTLY by inverting tube 4-6 times. Solution should turn blue.

6. Add 350 µl Buffer N3 and mix immediately and thoroughly by inverting the tube 4-6 times. Solution should turn colorless.

7. Centrifuge for 10 min.

8. Pipette the supernatant into the QIAprep spin column.

9. Centrifuge for 60 sec. Discard the flow-through.

10. Wash QIAprep by adding 0.70 ml (700 µl) Buffer PE and centrifuge 60 sec.

11. Discard the flow-through, and centrifuge for an additional 2 min to remove residual wash buffer. Wait 2-3 mins to dry (removes excess PE buffer and gives better sequencing results).

12. To elute DNA, place QIAprep column in a clean microcentrifuge tube. Add 30 µl Buffer EB to center of each spin column, let stand for 1 min, and centrifuge for 1 min.

13. NanoDrop to find concentration.

High volume minipreps: If plasmid is low copy (i.e. SC101) or a high amount or concentration is desired, grow 50 mL of cell cultures and use this modified miniprep protocol: I typically only do 10mL of culture per tube to allow for clean minipreps. In the past I've had trouble sequencing preps with too much culture volume, or I see very high bands of DNA on my gels (probably genomic DNA).

--

For 50mL's culture

A) Spin down in big centrifuge for 5 mins, remove media thoroughly

B) Add 7.5x amount of P1 (1875 uL)

C) Resuspend, and aliquot to 5 tubes (400 uL each)

D) Add 1.5x amount of P2 (375uL), invert gently 4-6 times

E) Add 1.5x amount of P3 (525uL), invert throughly

F) Spin Down for 10 mins

G) Run supernatant through 3-5 columns

H) Elute with 50-100uL

Notes: Also, after step (F), I sometimes collect the supernatant of microcentrifuge tubes into a big tube to prevent contamination with cell debris. Sometimes I do PE buffer wash twice when a high culture volume is used as well.

NanoDrop Spectrophotometer

Designed to find the concentration (in ng/µl) of DNA.

Procedure:

- **NOTE**: Use 2 µl of liquid for each measurement

1. Blank the instrument using Buffer EB.
2. Drop liquid onto meter and press “measure.”
3. Write concentration on side of microcentrifuge tube – “[concentration]”.
4. Wipe liquid off of bottom- and top-half of meter using Kimwipe.
5. Repeat steps 2-4 for each solution.

Agarose Gel Electrophoresis

Designed to separate DNA, RNA, or protein molecules using an electric current applied to a gel matrix.

"Creating the Gel" Procedure:

· **CAUTION**: Ethidium Bromide (EtBr) is a carcinogen. When handling, be sure to wear gloves and clean any spills thoroughly.

1. Weigh 0.35 g of agarose powder and pour into "EtBr" flask (For a standard 0.7% w/v gel)
2. Measure 50 ml of Buffer TAE and pour into "EtBr" flask.
3. Heat "EtBr" flask in lab microwave for 1 min 15 sec, or until all agarose particles dissolve.
4. Remove flask from microwave using "hot hands" and swirl (careful for vigorous bubbling).
5. Take plastic gel tray and ensure it is tight on gel caster. Insert desired gel comb into grooves.
6. Pipette 5 µl of GelRed dye into flask and swirl to ensure uniformity.
7. Pour contents of flask into plastic gel tray and let stand for ˜30 min to solidify.

"Running the Gel" Procedure:

· **NOTE**: Loading dye is "6x," meaning if there is 5 µl of DNA to be tested, 1 µl of dye should be added. Ladder used is 1kb (1000 base pairs). Both dye and ladder are located in third-to-right fridge.

· **NOTE**: Gel may be run @ 110 volts for 40 min(gel extractions, more careful gels) or @ 150 volts for 20 min (diagnostics/quick checks)

1. Remove tape and well placers from gel tray.
2. Place gel tray into electrophoresis apparatus, ensuring that Buffer TAE

covers the gel entirely.

3. Mix liquid DNA solution and dye. Solution should turn blue/violet.

4. Pipette liquid (generally 10-15 µl) into respective wells. Begin and end with ladder.

5. Place cover onto electrophoresis apparatus, matching red to red, black to black, and press "run."

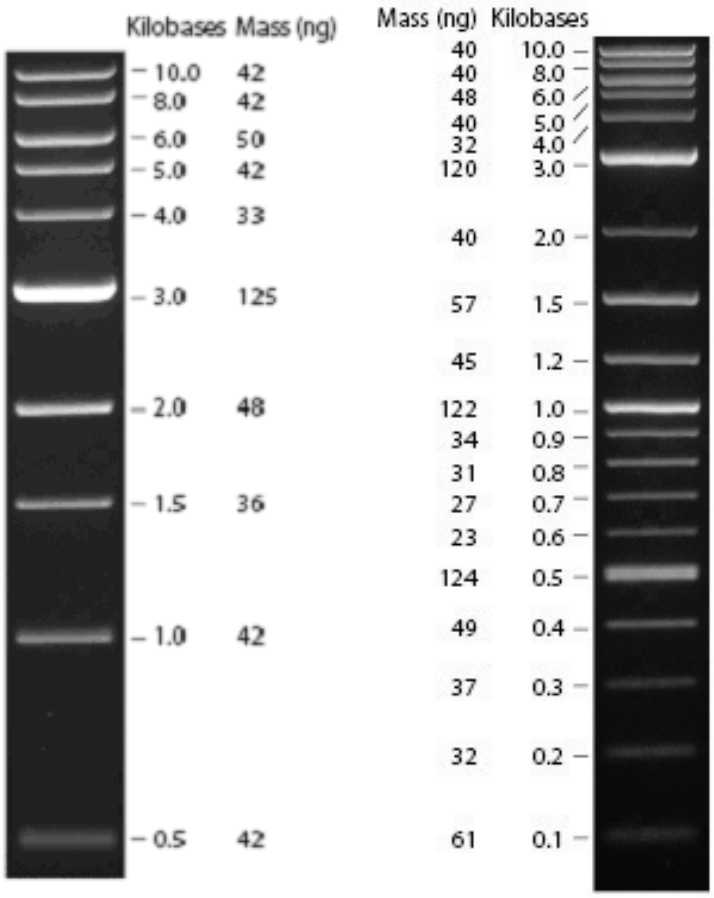


Figure 7.4: NEB 1kb and 2 Log Ladders. Taken from New England Biolabs website. We typically use the 1kb ladder at 180ng/uL concentration

"Assessing the Results" Procedure:

· **CAUTION**: Be sure to turn off UV lamp when done with photo-capture machine. If gel is to be saved for extraction, limit gel exposure to UV light.

· **NOTE**: Gel will be prone to slide off tray when wet.

1. Remove gel tray from electrophoresis apparatus, letting buffer drip off.
2. Wipe bottom of gel tray and place in photo-capture machine (located near lab microwave).
3. Focus and adjust light (white) of photo.
4. Turn on UV lamp and capture photo.
5. Press save and dump gel into designated bucket located to right of machine.
6. Access and print the photo through server online. Label wells and compare results with the ladder in order to designate size of DNA fragments.

Special Considerations:

For high separation of small fragments in the range of 0-1000bp, use a 1% gel. For separation of large fragments use a 0.5% gel.

Gel Digests

To cut DNA at specific sequences and often to leave sticky ends for ligating pieces together.

There are two types of Digests we do with restriction enzymes:

(1) screening/diagnostic of colonies

(2) digesting PCRs/plasmids for gel extractions&ligations.

(1) For diagnostic gels, we use between 50-200ng of DNA and digest for 1/2-1 hour before running on a gel. This is just to check if our DNA has the correct fragments

(2) For gel extractions, we use between 1500-3000ng of DNA and digest for 2-3 hours before running on a gel. The gel extraction procedure has low yield thus a lot is needed to start with.

Type (1) Diagnostic Gels

· MIX: Typically 1.0-4.0 microliters DNA, 1.0 of each 10x buffer (check chart), 0.25-0.5 of each enzyme, fill up to 10.0 microliters total with water.

· Leave at 37 incubator for 1/2-1 hour

· Make gel with thin comb

· By the time gel solidifies, probably ready to run gel

Typical Mix is

2.0 DNA

1.0 Buf 2

1.0 BSA

0.5 KpnI

5.5 H20

10.0 TOTAL

Type (2) Gel Extractions

· Digest desired amount of DNA (will proba bly be around 30-60 microliters) for 2-3 hours at 37.

· Make wide gel comb

· Run gel till bands are well separated and cut out gel piece, trying to minimize amount of agarose

Typical Mix

46.0 DNA

6.0 Buf 2

6.0 BSA

1.0 KpnI

1.0 MluI

60.0 TOTAL

Note: After creating mix, vortex briefly, then spin down briefly to ensure consistency.

Note: Addition of BSA does not affect digests so if 1 enzyme requires it, just add it.

Creating Master Mixes

Often times a large number of samples are digested with the same enzymes & buffers, etc., so a mix is created for all of them, then aliquoted into the tubes. This is to ensure uniformity across samples and allow for ease of pipetting. For instance if your gel digest mix is:

2.5 DNA

1.0 Buf 2

0.25 KpnI

6.25 H20

and you need this for 9 different samples of DNA, then your master mix is created for 10 samples without the DNA in it (always +1 or +2 the # of samples) :

10.0 Buf 2

2.5 KpnI

62.5 H20

=75 / 10=7.5 microliters/reaction

Before using, vortex & mini-centrifuge your mix to get the liquid to the bottom & ensure it is well mixed.

Keep on ice until aliquoted.

The same technique is done for PCR mixes.

Gel Purification

Designed to purify nucleic acids from an excised gel fragment

1. Weigh excised DNA (TARE with blank tube)
2. Add 3 gel volumes buffer QG(i.e. if piece is 100mg, add 300uL QG).
3. Allow to dissolve at 50 degrees for ˜10 mins, vortexing every few mins
4. If gel piece is <500bp or >4000bp, add 1 gel volume of isopropanol & mix
5. To bind DNA, apply the sample to the QIAquick column and centrifuge for 60 sec. Discard flow-through and place QIAquick column back into the same tube.
6. To wash, add 0.7 ml (700 µl) Buffer PE to the column and centrifuge for 60 sec. Discard flow-through and place the column back into the same tube.
7. Centrifuge for the column for an additional for 2 min.
8. Place each QIAquick column in a clean 1.5 ml microcentrifuge tube
9. Wait ˜5 min to dry.
10. To elute DNA, add 30 µl Buffer EB (water if no buffer) to the center of the QIAquick membrane, let stand 1 min, and then centrifuge 1 min.
11. NanoDrop to find concentration.

Ligations

To ligate fragments of DNA together.

Typically do a 10.0 microliter mix, with 30-100ng of vector. It is important to do the ligation at a 3:1 or 6:1 insert: vector ratio.

Ligation mix:

- 1.0 T4 ligase buf (make sure thawed, and no flakes at bottom of tube)
- 0.5 T4 ligase
- 2.5-8.5 microliters of DNA pieces at 3:1 molar ratio
- fill up to 10.0 microliters of water

Allow reaction to go on bench top for 30mins-3 hours. Alternatively, leave reaction in 16C water bath for >2-3 hours or overnight.

Calculating vector ratio example:

vector = 15ng / uL, 3 kb size

insert =30 ng/ uL, 1 kb

vector molarity=15/3=5

insert molarity=30/1=30

Here we can do 6:1 insert: vector and add equal amounts of vector & insert. It is probably easiest to add 4.25 microliters of both vector and insert to the ligation mix.

Transformations

The genetic alteration of a cell resulting from the uptake, genomic incorporation, and expression of foreign DNA.

· Competence: the ability of a cell to take up extracellular DNA from its environment.

· **NOTE**: Two kinds of cells employed: Bought ($15/each, stronger comp.) and Made-in-Lab ($0.15/each, weaker comp.). Located in -80 C freezer, Made Mach1 and DH5alpha cells are in "DH5alphaZ1" box, unlabelled tubes with 70-100 µl of cells in each tube. Bought Mach1 cells are in red Invitrogen box, unlabelled tubes with 25 µl of cells in each tube.e

Procedure:

1. Thaw cells on ice for ˜5 min. Ice located down the hall near -80C freezer.
2. Pipette ligated DNA into cell tubes. Add no more than 10% of total volume (e.g. if cell volume is 50 µl, add no more than 5 µl ligated DNA). Swirl solution with tip, DO NOT pipette up and down as it will damage cells.
3. Leave on ice for 30 min.
4. Heat shock @ 42 C for 90sec (30 sec for supercomps).
5. Leave on ice for 2 min.
6. Add 500 µl SOC.
7. Grow for 1 hour by placing tubes within flask and incubating inside shifting 37 C incubator.
8. While growing, heat appropriate resistance growth plates in 37 C incubator.
9. When hour is done, spin down cells @ 14000 rpm for 9-10 sec.
10. Remove 400 µl of SOC and re-suspend cells (via vortexing).
11. Plate ˜75 µl of cells.
12. Leave rest of cells on bench (in case plated on wrong resistance, can replate in the morning)

13. Put plate in 37C incubator O/N.

PCR protocol

To amplify fragments of DNA

New PCR Mix: This mix is a little better and allows you to keep primers undiluted and gives more product for CPEC's

Phusion: Keep in -20 holder, add to mix last

1. PCR mix (amounts in microlters.

a. 20.0 HF buffer

b. 2.0 dNTPS

c. 0.5 Phusion (polymerase)

d. 0.5 primer –S (undiluted, 100uM)

e. 0.5 primer –AS (undiluted, 100uM)

f. ˜1.0-2.0 template DNA (need 20-40ng of DNA)

i) If template is at X concentration, make a dilution to get it to 20-40ng/uL.

g. 75.0 qH20

Old PCR Mix: PCR primers should be at concentration 10.0 uM

Before Making Mix Below: Prepare primers

We buy primers at 100.0 uM (pmol/microliter) concentration. For the mix below they need to be at 10.0 uM (10x dilution).

- Label a new 0.6 mL microcentrifuge tube with the primer name.
- Add 90 microliters buffer EB
- Add 10.0 microliters of primer
- Vortex briefly

Phusion: Keep in -20 holder, add to mix last

1. PCR mix (amounts in microlters.

a. 10.0 HF buffer

b. 1.0 dNTPS

c. 0.5 Phusion (polymerase)

d. 2.5 primer –S

e. 2.5 primer –AS

f. 0.5-1.0 template DNA (need 10-20ng of DNA)

i) If template is at X concentration, make a dilution to get it to 10-20ng/uL.

g. 33.0 qH20

Common protocol 98 2mins,

32 x

98 10 sec

50-65 15-30 sec

72 for 15-20sec/kb

72 for 40sec/kb.

-Typically we do 2 reactions at Ta=56, 60 for our primers designed at Tm of 55. For primers designed at 57 , I typically find Ta=58 & 62 are good Ta's to use. This varies depending on the reaction.

-For plasmid PCR's, I find 15sec/kb is good, for genomic PCRs, I will typically use 20-30secs.

-Annealing time of 20sec is what I start with standardly, but if lower/higher specificity is needed you can change this.=

PCR Digests

Before using PCR products in a ligation, products must be digested. If PCR template is same resistance as final construct, digest with DpnI with the additional enzymes as well. DpnI digests methylated template DNA and reduces background.

CPEC Reactions

Cloning method for assembling PCR fragments together. I typically set my overlapping regions to a Tm of ˜60-65C and my PCR annealing region to a Tm of ˜57-60C.

OLD: PCR purify each piece , then setup a PCR reaction with these requirements:

NEW: Gel Purify each piece , then setup a PCR reaction with these requirements:

- about ˜300ng of vector
- about ˜200ng of insert
- 1:1 molar ratio of pieces
- PCR mix (amounts in microliters)

a. 10.0 HF buffer

b. 1.0 dNTPS

c. 0.5 Phusion (polymerase)

d. PCR fragments

g. fill up to 50 with qH20

CPEC protocol is:

98C 30sect

30 cycles of:

98C 10sec

55C 30sec

72C 20sec/kb

72C 5mins

Before running CPEC protocol, save 10 uL on ice(freezer) as a control.

After CPEC is done, run 10 microliters of the CPEC reaction with the control in the next lane. Check that you can see your assembled reaction on the gel, or that your inserts are less visible than the control (they assembled). If so, then transform 10 microliters into regular competent cells.

Note: In some situations it is difficult to fulfill the first 2 requirements (200 ng insert & 300 ng vector). In those cases I would just make sure there is enough vector, and add enough insert at a insert:vector of (1:1-3:1)

Note2: I've found recently that Gel Purifying the vectors gives much higher efficiency and much cleaner results when running the CPEC on a gel. This seems to reduces smears in some CPEC reactions or strange sized pieces(also what they do in paper). I would recommend running PCRs of 50-100 microliters and then gel extracting directly after the PCR instead of gel checking & PCR purifying.

PCR Purification

Designed to quickly purify nucleic acids.

Procedure:

· **NOTE**: Use Pink spin columns; centrifuge rpm 14000.

1. Add 5 volumes of Buffer PBI to 1 volume of the PCR reaction and mix.
2. To bind DNA, apply the sample to the QIAquick column and centrifuge for 60 sec. Discard flow-through and place QIAquick column back into the same tube.
3. To wash, add 0.7 ml (700 µl) Buffer PE to the column and centrifuge for 60 sec. Discard flow-through and place the column back into the same tube.
4. Centrifuge for the column for an additional for 2 min.
5. Place each QIAquick column in a clean 1.5 ml microcentrifuge tube. Wait ˜3-5 min to dry.
6. To elute DNA, add 30 µl Buffer EB (water if no buffer) to the center of the QIAquick membrane, let stand 1 min, and then centrifuge 1 min.
7. NanoDrop to find concentration.

Sequencing DNA (Eton Bio's service)

To sequence a piece of the plasmid DNA starting from a primer binding site. Sequence generated is typically 1000 bp if successful, 400-500 bp if mildly successful. I typically design my sequencing primers with Tm=55-60C and about 50-100bp upstream of desired sequence. Results should arrive next day.

Protocol

-Aliquot 250-500ng of DNA to be sequenced in a new microcentrifuge tube and label

-Aliquot 1 microliter of 100uM primer to a new microcentrifugetube, add 19 microliters of qH20, and label (only need to give them 3-4 microliters of this mix)

-Put both tubes in -20 freezer in the ”Eton” box

-Fill out online order form at http://etonbio.com/plslogin.php with U/N: hasty-sequence@gmail.com

--Check box for primer & DNA template separate

--Fill out concentrations of DNA and primer (primer concentration should now be 5pmol/microliter)

--Fill in your email

Colony PCR

To screen many colonies by PCR'ing a part of the inserted piece

PCR Mix

0.2 dNTPs

0.2 Taq

2 PCR 10x buffer

0.2 -S primer (1o um conc)

0.2 -AS primer

17 water

---

20 total

1) Aliquot mixture to PCR strips

2) Touch colony with Red pipette tip, then let tip stand in tube for ~1 minute

2a) For higher copy plasmids, you might want to dilute into 20 microliters water, and add 1 microliter as template

3) Create master plate with tip for each colony , place at 37 C

4) Run standard PCR protocol but (98 30sec, 50 30sec, 72 1min/kb)x30

5) Run 15 microliters on gel

Alternatively use Taq Master Mix:

8.5 microliters mix

0.2 -S primer (10 um conc)

0.2 -AS primer

1 cell colony diluted by 20-100x

---

10 total

PCR optimization

Below is a guide for troubleshooting PCR reactions in various situations. My standard PCRs are with the mixes described above.

Scenario: No product

Troubleshooting steps (in order of importance, but many can be done at same time):

-Check that primers sequences bind sequence file at correct location. Check primer Tm's are <3C apart.

-Redo PCR(sanity check). Use original sources, check DNA template concentration(should use 10-20ng/50uL).

-Do a gradient of temps from the Tm (or a few degrees below) to about 5 degrees above.

-Increase annealing time to 30 sec, & do gradient of temps

-Increase/Decrease template concentration by factor of 2-3 (up to 50ng), important for longer PCRs where molarity changes. Or increase/decrease primer concentrations.

-If you have a control plasmid that primers should bind use this as a test

-Use a test primer for the -S & -AS and see which primer isn't working (or order a new one)

-Optimize Mg2+ concentration. 1) Decrease Phusion to 0.25 uL. or 2) Increase Mg2+ in 0.2mM steps

-Sequence template to determine if primer site is there.

-Try DMSO addition to 0-5%

Scenario: Unspecific bands (longer than desired product)

-Check for primer binding sequences that are similar in Vector NTI

-If band is longer than desired product, shorten PCR extension till band goes away

-Increase annealing temperature, search for optimal Ta

-Shorten annealing time

-Use touchdown PCR technique (Start at Ta˜+10C of your initial Ta, then decrease by 1-2C/cycle on machine, do 35x cycles of PCR)

Scenario: Unspecific bands (shorter than desired product)

-Check for primer binding sequences that are similar in Vector NTI

-Increase annealing temperature, search for optimal Ta

-Shorten annealing time

-Use touchdown PCR technique (Start at Ta˜+10C of your initial Ta, then decrease by 1-2C/cycle on machine, do 35x cycles of PCR)

Scenario: Primer dimers

-optimize as in ”No product” section

-I typically find for this situation Mg2+ concentration & primer/template ratio are most important factors

Scenario: Smears

-typically poor reaction conditions (get new stocks of buffers, dNTPs, etc.)

A good reference from NEB: http://www.neb.com/nebecomm/products/protocol222.asp

Designing PCR Primers for CPEC

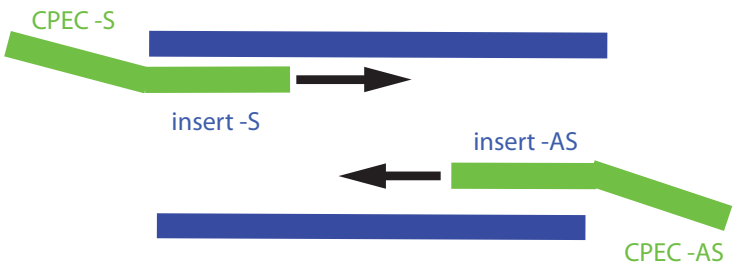


Figure 7.5: Overview of Primer Design for CPEC. Primers are designed in two parts, the annealing to template part, and the complementary part, where each have their own but similar design criteria.

For my CPEC reactions I design CPEC ends on the inserts only and PCR the vector w/o CPEC ends. The pZ vectors that are PCR'd can then be added to any insert interchangeably.

[Warning: Image not found]

I typically follow 3 requirements for designing primers:

(1) Primer Tm about 60C

(2) Primer pair $<$ 3C apart

(3) GC content 40-60% (roughly, though can be stretched)

(4) CPEC ends Tm 60-65C

t

Step 1: Build desired sequence in Vector NTI

Step 2: Build insert -S primer

(a1) copy first ˜25 bp of insert sequence into Finnzymes Tm caclculator (link) - should start with ATG/GTG

(b1) remove bases from the **end** of the sequence till Tm is about ˜60C

(c1) save part of sequence

Build CPEC -S overhang

(a2) copy 25 bp of sequence upstream of insert

(b2) remove bases from **front** of sequence till Tm is about ˜60-65C

(c2) save sequence

Combine sequences in order of (c2) then (c1) for complete -S primer

Step 3: Build insert -AS primer

(a3) copy last ˜25 bp of insert sequence into Finnzymes Tm caclculator (link) - should end with TAA/TAG/TGA

(b3) remove bases from the **front** of the sequence till Tm is about ˜60C

(c3) reverse complement sequence (link)

(d3)save reverse complemented sequence

Build CPEC -AS overhang

(a4) copy 25 bp of sequence directly after insert

(b4) remove bases from **end** of sequence till Tm is about ˜60-65C

(c4) reverse complement sequence (link)

(d4) save reverse complemented sequence

Combine sequences in order of (d4) then (d3) for complete -AS primer

Step 4: Error checking

(1) make sure insert binding sequences binds to separate insert plasmid source within a few mismatches

(2) make sure full primer sequences bind assembled Vector NTI sequence

Put sequences into ordering site (set Normalization YES, 100uM):

Valugene: https://valuegene.com/process.cgi

pZ plasmid Construction Notes

For simple gene insertion into pZ plasmid with original RBS , here are the CPEC ends:

CPEC -S: attaaagaggagaaaggtacc (add to beginning of insert -S primer)

for LAA tagged inserts:

CPEC -AS: TCGTCGTTCGCTGC (add to beginning of insert -AS primer)

for non-LAA tagged inserts:

CPEC -AS: gcctctagcacgcgt (add to beginning of insert -AS primer)

Vector Primers

LAA vector primers: ZE24EcoRBS orig vect -AS , pZE25vect -S

noLAA vector primers: ZE24EcoRBS orig vect -AS , pZE25vectnoLAA -S

Combining pZ plasmids

I use the following set of primers with a synthetic linker following the AvrII site which seems to combine these plasmids efficiently. I have similar primers for various origins & promoters as well:

Pa/l vect -S tgtatagtcacgactggtcg cctagggcgttcgg

Pa/l vect -AS CCAGTTGATCGACGATTC tctagggcggcgg

pLtet cfp -s GAATCGTCGATCAACTGG CTAA GAAACCATTAT TATCAT-GACA

pLtet cfp -as cgaccagtcgtgactataca gTCTAGGG CGGCG

Screening colonies from CPEC Reactions

After miniprepping several colonies from your CPEC reaction, choosing the proper restriction enzymes to distinguish between correct and incorrect colonies is crucial. The most important considerations are:

1. What are the possible incorrect plasmids? Where did insert & vector pieces come from and if it is possible that they were transformed (i.e., they have the same resistance, etc.) ?

2. How big are the fragments being generated with restriction enzymes used? Will it help distinguish between incorrect/correct plasmids? Keep in mind larger fragments are harder to differentiate.

3. Will the fragments generated be visible on the gel? Keep in mind the lower limit of detection is 25-50ng of DNA with GelRed.

Below is an example screening for a typical CPEC reaction. Here is the chart I usually setup before doing a CPEC:

Consideration 1: Determining incorrect / correct plasmids

The final correct plasmid here (pZA11sfGFP) is Amp resistant. Hypothetically, the templates from fragments S1 & S2 can be transformed at a low level, but in this case we see that piece S2 (template pZE24sfGFP-LAA) cannot be transformed because it is Kan resistant. Thus the only possible incorrect template that could be transformed is pZA11yfp. In addition, CPEC reactions often times give a background of closed vector, i.e., the S1 PCR closes on itself and is transformed. Since this is Amp resistant as well this can be a possible incorrect plasmid (call it pZA11empty).

This gives us three possible cases of plasmids that could be transformed:

1. Correct plasmid -- pZA11sfGFP
2. Background plasmid from S1 -- pZA11yfp
3. Background closed vector -- pZA11empty

We need to choose a restriction enzyme(s) that will help us distinguish between these cases. The best way to do this is to use or construct sequences for each of the three cases and then write down the size of the fragments from Vector NTI.

Constructing Sequences

Use primer sequences from above to search for where primers bind on the initial template sequence of vector & insert, S1 & S2(only use binding part not CPEC overhangs). For the vector, delete the rest of the bases and save the sequence (this will be your pZA11empty sequence). Then use Copy & Paste commands to put the insert sequence into your vector and save as final plasmid (pZA11tetR). You should now have the three sequence files to digest with.

Common Digest types for screening

Single cut: If we choose an enzyme which cuts once on each of these vectors we can often screen this way. In this situation, the sizes of the fragments would be (1) 3.01, (2) 3.02 (3)2.4 kb. We can differentiate here between (1)&(2) vs. (3) but not between (1) & (2) on a gel, thus this is not an ideal screen for this situation.

Insert cut: One of the simplest options is usually to try and cut out the insert (sfGFP, yfp, or empty) and check the corresponding size of the fragment. Here sfGFP and YFP both have sizes of 750bp, thus it is not a good option for this situation.

Interior cut(s): Another way would be to find an enzyme cutting inside of either tetR or yfp, and screen this way.

I typically don't like to see a situation where 1 of the plasmids is uncut because this does not control against the reaction failing altogether. However, it may be possible to find an enzyme which cuts once in yfp/sfGFP, and twice in the other, allowing us to distinguish the two situations. Looking at both sequences in Vector NTI, and adding common ”Restriction Sites” we can see that there are two XhoI cuts in sfGFP while only one in the YFP plasmid. This can be used as a screen between the two cases.

Restriction Summary of Multiple Plasmids: We can use this site and input the plasmids to highlight the differences in the restriction sites between 2 or 3 plasmids. Enter them in the form:

>Plasmid 1

atatatggg

>Plasmid 2

atatatttt

>Plasmid 3

Consideration 2: Distinguishing fragments

Digests of the 3 plasmids with XhoI gives the following fragments

1. 500bp , 2500bp
2. 3000bp
3. 2200bp

These can be distinguished on a 0.7% gel pretty easily.

Consideration 3: Visualizing fragments

Typically we do something between 100-200 ng of DNA for gel digest checks. To be more precise, what we mean is that we need each fragment cut out to be > 50 ng to be visualized on a gel, and usually 100-200 ng total DNA covers this requirement. Since smaller pieces of DNA are harder to visualize (less mass & less dye bound), we need to be careful about being able to visualize the 500 bp piece. In addition, since our screen is based mostly on being able to see the 500bp or not, it is important to make sure we get this right.

If we want the 500 bp piece to be approximately 50 ng's, and it is about 1/6 of the total mass of the plasmid, then for this digest it would be safe to start with about ˜300 ng's of starting DNA. If DNA is concentration of minipreps is in the range of 40-50 ng/uL for instance, we could do something like 6.5-7.5 uL of DNA for each. This will also make the master mix simpler since no water will be added.

Notes:

Need to find a website which can do a Restriction Summary on multiple plasmids at once, highlighting difference in cuts.

Making Competent Cells

0. Start an O/N culture of your strain in 1-5mL LB media.

1. 1/100x dilution of overnight culture in 25 mL LB (makes 25 competent cells). Grow up for ~2-3 hours(depending on strain), OD needs to be between 0.4-0.6 for significant competency.

2. Spin down cells at 8000 for 5 min. Resuspend in 2.5 mL TSS solution(1/10x original). Aliquot to prechilled microcentrifuge tubes (prechill at -20C). Freeze at -80C for storage (retain competency for about ~6 months)

Making PDMS chips for Microscope Experiments

Tal Danino

Overall, during this procedure you want to be as CLEAN and CAREFUL as possible with making chips.

-Wafers are very fragile and should not be bent. Best to hold them on sides away from features, and kept in a clean petri dish when not in use.

-During mixing & the rest of chip making you want to be careful about getting dust in your mix or on your chips, and if your gloves get dirty you will want to change them.

Making foil dish for wafer: Put on clean gloves. Cut out about a 8 inch by 8inch piece of foil. Take non-useful wafer and trace the circular shape onto the foil. After this draw a 2 inch circular margin around and cut out with scissors. Flatten and straighten the foil so that there is less creases in it.

Making PDMS mix: Take a plastic hexagonal dish (not the top one in the bag), and blow dry with airgun. Tare scale w/dish and 30g of Silicone, then 3 g of curing agent. Clean a glass rod with a kimwipe and clean with blow gun. Under the fume hood, stir mixture for 5 minutes. The goal is to get the mixture as uniform as possible.

Removing bubbles from PDMS mix:

NEW: Pour entire mixture into 50mL Falcon tube. Spin down for 3-5 minutes to remove bubbles.

OLD:Place tray w/ mix in the dessicator on the RIGHT and put on lid. Turn on vacuum pump in the hood while valves are all closed. Make sure back valve(air release) is closed and SLOWLY open up front valve. The lid of the dessicator should be sticking to the blue part, and you will see the gauge go to about ˜25 or so. Allow air

bubbles to come to surface, and every 2-3 minutes OPEN back valve for a few seconds to release pressure. Do this about 5-10 times or until you see no bubbles.

Pouring PDMS mixture over wafer: Place your desired wafer in the foil dish you made and raise up the sides around it so that you can pour in the PDMS mix. Its best to try and have the wafer lie FLAT on the foil so that PDMS doesnt get in under it. Place tray w/ wafer in the dessicator on the RIGHT. Very slowly pour over the PDMS mix w/o bubbles onto the wafer. Best to pour at center of the wafer and allow it to spread to the sides. REPEAT procedure to remove bubbles from the wafer. There should already be much less so you can leave it going for 5-15 minutes and come back and check if bubbles are gone.

Baking PDMS chips: To harden the chips, bake in the 80 degree oven for 1 hour. Carry over your foil tray w/ wafer & mix to the oven very carefully and try not to introduce more bubbles.

Removing PDMS layer from the chip: After baking, leave dish at room temperature for about ˜5minutes. This part you have to be the most careful with since there is a chance of breaking the mask in half or damaging the surface. SLOWLY peel off foil from the PDMS. If you can, try and remove some of the PDMS from the bottom of the wafer at the same time. First you'll want to remove all PDMS from bottom of the wafer. To do this, take a razor blade and slowly cut off the thin layer up until the outside of the circle. Make sure all of the PDMS from the bottom is removed before trying to lift off the PDMS from the top. VERY VERY SLOWLY lift up the PDMS layer from one side of the chip, and try and do this evenly from all sides. When you peel off the PDMS from the feature part of the chip do this really carefully and allow PDMS to peel off by itself w/o applying a lot of pressure. Make sure that you do not bend the wafer at all, because this can cause it to snap in half. When separated place wafer into a clean dish and set aside.

Punching holes in chips: Cut out excess PDMS and save a smaller square around the chips you want to use so that its easier to work with. Take YELLOW puncher, glass dish w/ rubber bottoms, and tweezers to light scope. Focus on the circular holes and place the puncher directly above the circular part. Make sure you are holding it as vertical as possible, and then press down hard to punch through. Slowly lift up puncher(easier if you turn it back and forth), and remove PMDS from puncher with tweezer. Continue doing this for ALL holes on the chips.

Cleaning holes: (use syringe filled with water to clean out excess PDMS from holes.)

Use the razor to cut out each chip individually. Place chip in clean Petri dish. Plug in port with syringe and apply pressure till water leaves from other side. Do this for all holes and on both sides of the chip.

Cleaning chips: This part is very important so that no dust enters in chips. Place chips in the glass dish (wash w/ water and blow first) and add 70% Ethanol to submerge them. Swirl the chips around in the dish w/ the lid and pour off the excess ethanol. Repeat this rinse with water from milliQ system. After water rinse, individually blow dry each chip on both sides and put in a new Petri dish. For each chip use scotch tape to remove dust from the chip. **Important: Run fingernail(or razor) over features of the chip several times to remove dust(this seems to be most revelant step for clean chips)**. Leave chip sealed with a fresh piece of scotch tape on the feature side.

Cleaning coverslips:

Be gentle/careful with coverslips, they can break easily and are very sharp. Take a new coverslip and spray on both sides with Heptane. Wipe clean on both sides with a Kimwipe and make sure there are no liquid spots or residues on both sides. Do the same for Methanol. Now wash with water on both sides and then air dry with blow gun. Absolutely make sure there are no spots or dust on both sides (if there are redo water wash), and place clean in a Petri dish.

Bonding chips to coverslips:

Open O2 valve of plasma bonder, and make sure O2 level is between 0.4 to 0.6. Turn on plasma bonder and run for 5 minutes to warm up. Place 1 chip(make SURE feature side is facing up!) and 1 coverslip in bonder tray (use tweezers for coverslips). Run for 3 minutes. When done, open up tray and flip over chip onto coverslip to bond. You'll want to do this as quick as possible for best seal. Place in 80 degree oven overnight.

Microscope Experiment Protocol DAW6v2

Tal Danino 8.26.2009

1. Grow O/N culture from a -80 stock or plate.

2. **Grow up cells.** 2 hours before setting up experiment, do a 1/1000x dilution of cells in appropriate inducers/antibiotics in 50mL media. Grow cells for ~2 hours till OD=0.05-1.0 (I shoot for 0.08), and have chip wetted and setup before spinning down cells.

3. **Prep syringes** (During 2 hours) (2 water, 2media, 2cells). Make sure to use media & waters filtered + 0.075% Tween20. Add dye to one of the media's (1uL per 5mL).

4. **Inspect chips** (During 2 hours) Go to microscope and check out chips at 4x magnification(PhL condenser setting), don't need to screw down chips. Look around for pieces of dust blocking channels that could be a problem. And check to make sure traps are not collapsed (they look collapsed if color of posts look same as traps). Tape the additional device on the chip to prevent from getting contaminated or wetted. Set the temperature of the scope box at this point to with fan at max speed. Screw tight all 4 points on the coverslip holder with red rubber slips evenly.

5. **Wet the chip** from 1 of the 2 media ports and raise the reservoir to the highest position to speed up wetting. Once a port has become wet (looks like no fluid movement on punched hole, or droplet on surface), then plug that port in and raise it up to the highest as well (at same point as other one is ideal). I prefer wetting the chip with the media with dye in it.

6. **Tape lines.** Once all ports are wet, tape down each of the lines to the square microscope insert. Be very gentle when touch lines because bonding of chip can get ruined.

7. Set ports to the appropriate heights. Medias should be a DAW height of 500-550. Shunt (water) at 21-22", Junction (water) at 8".

8. **Spin down cells** for 4 mins in big centrifuge. Dump excess media back into 50mL flask as backup cells and put in 37 shaker. Add 3-5 fresh media and resuspend by vortexing briefly. Prepare syringe as before and set up at proper height (7") and plug in. Set both cell ports at identical heights.

9. Slowly raise the cell ports or lower the junction about 1-2". Watch the cell ports at 20x mag (Ph1 condenser). After approximately ˜2-5 mins you'll see cells start to come down from ports and towards the waste. Check the junction and make sure cells are going in there and Mcherry dye is present. Also make sure cell reservoirs are not mixing and both going directly to the junction.

10. **Load cells** Set a slow speed for cells going by the traps and proceed to flicking. The goal is to get at least 1-2 traps loaded for each lane (4 total), and best if they are loaded at traps closest to junction (cause they will fill all the ones downstream). Before flicking, tape down lines. Hold lines taut, and give a few hard flicks, then checking to see what got loaded.

11. If cells are loaded then reverse the flow by raising the heights to previous values. Cells should be seen zooming by traps at a rate of 25-200 um/sec.

12. **Setup DAW.** Go to DAW setup and adjust heights so that its at 50% level. Jog+ one of the heights till it's at the 100% level. This should be extremely close the boundary, but not allow the other media to flow through. Set the 100 % level on software. Do the same for 0 level. And hit Calibrate on DAW software. Test 0, 50, and 100% levels with slider. Create and load DAW run table file.

13. Allow cells to grow inside trap for 3-4 doublings(1-2 hours). Make sure you give them the appropriate media before starting the experiment. In our case, we want to give them 0% Arabinose so first step will induce them.

14. Setup 100x objective. CAREFULLY remove stage insert and put a big drop of oil on 100x objective(Ph3 condenser). Make sure it snugly inserted. Slowly bring up objective till oil touches coverslip. Bring condenser down to level which produces highest amount of lights exposure and hit Auto Exposure.

15. **Set up Scope software**. Select your XY points, time frequency for camera acquisition, and set brightfield / GFP / Mcherry wavelengths. For DAW6v2, shoot for 30-60 seconds for a round of brightfield and fluorescent images (about 3-5 xy points) and set GFP every 6 to 12 frames. Set up autofocus and make sure global settings match for Advanced Brightfield Phase only and not GFP. Use red box for autofocusing over cells. Make sure fluorescent lamp is off & set GFP exposure to desired (1-3seconds). Test a single loop for autofocus to see how long it takes and that it catches the right focus. If focal planes are far away in Z direction adjust allen wrench tightness on one side or carefully push down on stage insert to even setup. Mcherry exposure setting should be set to 200-300 ms and every 12 mins.

16. **Make sure fluorescent lamp is set to 10%. Start the DAW and scope run simultaneously.**

Flow Cytometry

Settings are what I use for typical tagged and untagged FP's in E.coli

1. Turn flow cytometer ON, **then** turn ON Computer (Make sure computer is off first , and I usually wait 30 seconds to turn on the computer)

2. Login: Bridget, Leave PW blank

3. **Setup Software:** Apple-> CellQuest

4. Acquire->Connect to Cytometer

5. File->Open->Istanbul->FACS Calibur Users->Hasty->Tal->"Acquisition Template"

6. Threshold: FSC-H -> 0; FL1 Log 750; FSC Log 350; SSC Log E00; Compensation all to 0

7. Acquire->Parameter Description, Hit Folder and make new folder with date. Select.

8. Acquire->Counters

9. Hit RUN on flow cytometer.

10. Acquire while in Setup mode (checked) and adjust flow LOW or MED or HIGH so that events are not more than 5000events/sec

11. To acquire data, uncheck Setup, and just go through samples.

12. **Shutdown:** Put in a tube with bleach (follow instructions on paper printout above cytometer)

13. **Transfer Files:** Apple->Recent Servers->Images3. Password is normal complex one.

14. Make new folder in your directory. Copy files over from Istanbul(On desktop)->FACS Calibur Users->Hasty->Tal->Directory. Copying directories doesn't work as well

Preparing Electrocompetent cells and Electroporations

T.D adapted from

Lambda mediated gene replacement on openwetware
Makes ˜12-15 aliquots of e- comp cells.

- Grow 5mL strain of interest with pKD46 at 30C overnight.
- Prepare two flasks with 1/100x dilution of overnight in 250 mL LB and grow at 30C.
- Label 1 flask + and the other - L-arabinose (control)
- When OD_{600} of cells(+pKD46) reaches 0.1 (˜2 hours), add L-arabinose to concentration of 0.15% to induce pKD46 λ-red expression
- – add ˜2mL of 25% L-arabinose to 250 mL + culture, none to - culture
- Continue to grow at 30°C to OD_{600} = 0.4 (˜2-3 hours)
- Chill cells in ice-water bath 10 minutes
- Centrifuge 10 min at 4000rcf 4°C in 35mL nalgene centrifuge tubes
- Pipette off supernatant and resuspend pellets in 1-5 mL ice-cold dH_2O(filtered)
- Centrifuge 10 min at 4000rcf 4°C
- Pipette off all dH20 carefully
- (Optional: another spin wash step in ice-cold dH20)
- Resuspend pellet in 1000 μL dH_20 +10-15% glycerol

- Aliquot 50uL per tube (prechilled)

- **Electroporation**
- For electroporation step, include 2 conditions: +/- PCR fragment
- Chill electroporation cuvettes for 5 minutes on ice(or don't need if kept in -20C)
- Add 5 pg to 0.5 μg PCR amplified DNA to cells (For genomic insertions: Typically I add 50-100ng (of 50-100ng/uL))
- Set electroporation apparatus to ”Bacteria”
- Prepare 1mL SOC in pipette. Take cuvette off ice, wipe metal electrodes with kimwipe.
- Place the cuvette into the sample chamber(sort of quickly so no condensation on electrodes happens)
- Apply the pulse by pushing the button
- Remove the cuvette. Immediately add 1 mL LB medium and transfer to a sterile culture tube
- Incubate 60-120 min with moderate shaking at 37°C
- Plate at 37C (for genomic insertions)
- If transformation doesn't work, replate in the morning (as in Datsenko)

Bibliography

Albrecht, D. R., Underhill, G. H., Wassermann, T. B., Sah, R. L., and Bhatia, S. N., 2006: Probing the role of multicellular organization in three-dimensional microenvironments. *Nat. Methods*, **3**(5), 369–375.

Alon, U., 2007a: *An introduction to systems biology: design principles of biological circuits*. CRC Press.

Alon, U., 2007b: Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, **8**(6), 450–461.

Alon, U., Barkai, N., Notterman, D., Gish, K., Ybarra, S., Mack, D., and Levine, A., 1999: Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proceedings of the National Academy of Sciences*, **96**(12), 6745.

Ambrosi, D., and Preziosi, L., 2006: Mechanics and chemotaxis in the morphogenesis of vascular networks. *Bulletin of mathematical biology*.

Anderson, J., Clarke, E., Arkin, A., and Voigt, C., 2006: Environmentally controlled invasion of cancer cells by engineered bacteria. *Journal of molecular biology*, **355**(4), 619–627.

Atkinson, M., Savageau, M., Myers, J., and Ninfa, A., 2003: Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli. *Cell*, **113**(5), 597–607.

Austin, D., Allen, M., McCollum, J., Dar, R., Wilgus, J., Sayler, G., Samatova, N., Cox, C., and Simpson, M., 2006: Gene network shaping of inherent noise spectra. *Nature*, **439**(7076), 608–611.

Barton, M., Delneri, D., Oliver, S., Rattray, M., and Bergman, C., 2010: Evolutionary Systems Biology of Amino Acid Biosynthetic Cost in Yeast. *PloS one*, **5**(8), e11935.

Basu, S., Gerchman, Y., Collins, C., Arnold, F., and Weiss, R., 2005: A synthetic multicellular system for programmed pattern formation. *Nature*, **434**(7037), 1130–1134.

Becskei, A., and Serrano, L., 2000: Engineering stability in gene networks by autoregulation. *Nature*, **405**(6786), 590–593.

Bennett, M., Pang, W., Ostroff, N., Baumgartner, B., Nayak, S., Tsimring, L., and Hasty, J., 2008: Metabolic gene regulation in a dynamically changing environment. *Nature*, **454**(7208), 1119.

Bennett, M., Schatz, M., Rockwood, H., and Wiesenfeld, K., 2002: Huygens's clocks. *Proceedings: Mathematical, Physical and Engineering Sciences*, **458**(2019), 563–579.

Boi, S., Couzin, I., Del Buono, N., Franks, N., and Britton, N., 1999: Coupled oscillators and activity waves in ant colonies. *Proceedings of the Royal Society B: Biological Sciences*, **266**(1417), 371.

Bostrom, K., Tsao, D., Shen, S., Wang, Y., and Demer, L. L., 2001: Matrix gla protein modulates differentiation induced by bone morphogenetic protein-2 in c3h10t1/2 cells. *J. Biol. Chem.*, **276**(17), 14044–52.

Bostrom, K. I., 2000: Cell differentiation in vascular calcification. *Z. Kardiol*, **89**, 69–74.

Bramson, M., 1998: State space collapse with application to heavy traffic limits for multiclass queueing networks. *Queueing Systems*, **30**(1-2), 89–148.

Bramson, M., and Dai, J. G., 2001: Heavy traffic limits for some queueing networks. *Annals of Applied Probability*, **11**(1), 49–90.

Brenner, M., Levitov, L., and Budrene, E. O., 1998: Physical mechanisms for chemotactic pattern formation by bacteria. *Biophys. J.*, **74**, 1677–1693.

Buck, J., and Buck, E., 1968: Mechanism of Rhythmic Synchronous Flashing of Fireflies: Fireflies of Southeast Asia may use anticipatory time-measuring in synchronizing their flashing. *Science*, **159**(3821), 1319.

Buhl, J., Sumpter, D., Couzin, I., Hale, J., Despland, E., Miller, E., and Simpson, S., 2006: From disorder to order in marching locusts. *Science*, **312**(5778), 1402.

Chabot, J., Pedraza, J., Luitel, P., and van Oudenaarden, A., 2007: Stochastic gene expression out-of-steady-state in the cyanobacterial circadian clock. *Nature*.

Chen, D., Zhao, M., and Mundy, G. R., 2004: Bone morphogenetic proteins. *Growth Factors*, **22**(4), 233–41.

Cookson, S., Ostroff, N., Pang, W., Volfson, D., and Hasty, J., 2005: Monitoring dynamics of single-cell gene expression over multiple cell cycles. *Molecular Systems Biology*, **1**(1).

Cross, M., Meiron, D., and Tu, Y., 1994: Chaotic domains: A numerical investigation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **4**, 607.

Cukierman, E., Pankov, R., and Yamada, K. M., 2002: Cell interactions with three-dimensional matrices. *Current Opinion in Cell Biology*, **14**(5), 633 – 640.

Damerau, K., and St John, A., 1993: Role of clp protease subunits in degradation of carbon starvation proteins in escherichia coli. *Journal of bacteriology*, **175**(1), 53.

Danino, T., Mondragon-Palomino, O., Tsimring, L., and Hasty, J., 2010: A synchronized quorum of genetic clocks. *Nature*, **463**, 326–330.

Danino, T., Mondragón-Palomino, O., Tsimring, L., and Hasty, J., 2010, PMID: 20090747: A synchronized quorum of genetic clocks. *Nature*, **463**(7279), 326–330.

DeJesus, E. X., and Kaufman, C., 1987: Routh-hurwitz criterion in the examination of eigenvalues of a system of nonlinear ordinary differential equations. *Phys. Rev. A.*, **35**(12), 5288–5290.

DiMilla, P. A., Quinn, J. A., Albeida, S. M., and Lauffenburger, D. A., 1992: Measurement of individual cell migration parameters for human tissue cells. *AIChE J.*, **38**(7), 1092–1104.

Dunlap, P., and Greenberg, E., 1985: Control of Vibrio fischeri luminescence gene expression in Escherichia coli by cyclic AMP and cyclic AMP receptor protein. *Journal of bacteriology*, **164**(1), 45–50.

Elowitz, M., and Leibler, S., 2000a: A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**(6767), 335–338.

Elowitz, M., Levine, A., Siggia, E., and Swain, P., 2002: Stochastic gene expression in a single cell. *Science*, **297**(5584), 1183–1186.

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

Elowitz, M. B., and Leibler, S., 2000c: A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**(6767), 335–8.

Elson, R., Selverston, A., Huerta, R., Rulkov, N., Rabinovich, M., and Abarbanel, H., 1998: Synchronous behavior of two coupled biological neurons. *Physical Review Letters*, **81**(25), 5692–5695.

Endy, D., 2005: Foundations for engineering biology. *Nature*, **438**(7067), 449–453.

Farrell, C., Grossman, A., and Sauer, R., 2005: Cytoplasmic degradation of ssrA-tagged proteins. *Molecular microbiology*, **57**(6), 1750–1761.

Friedland, A., Lu, T., Wang, X., Shi, D., Church, G., and Collins, J., 2009: Synthetic gene networks that count. *Science*, **324**(5931), 1199.

Fung, E., Wong, W., Suen, J., Bulter, T., Lee, S., and Liao, J., 2005: A synthetic gene–metabolic oscillator. *Nature*, **435**(7038), 118–122.

Futcher, B., Latter, G., Monardo, P., McLaughlin, C., and Garrels, J., 1999: A sampling of the yeast proteome. *Molecular and cellular biology*, **19**(11), 7357.

Gamba, A., Ambrosi, D., Coniglio, A., and Candia..., A. D., 2003: Percolation, morphogenesis, and burgers dynamics in blood vessels formation. *Physical review letters*.

Garcia-Ojalvo, J., Elowitz, M., and Strogatz, S., 2004: Modeling a synthetic multicellular clock: Repressilators coupled by quorum sensing. *Proceedings of the National Academy of Sciences*, **101**(30), 10955–10960.

Gardner, T., Cantor, C., and Collins, J., 2000a: Construction of a genetic toggle switch inEscherichia coli. *Nature*, **403**(6767), 339–342.

Gardner, T. S., Cantor, C. R., and Collins, J. J., 2000b: Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, **403**(6767), 339–42.

Garfinkel, A., Tintut, Y., Petrasek, D., Bostrom, K., and Demer, L. L., 2004: Pattern formation by vascular mesenchymal cells. *Proc. Natl. Acad. USA*, **101**(25), 9247–9250.

GC, P., 2005: Crowd synchrony on the Millennium Bridge. *Nature*, **438**, 3.

Ghosh-Choudhury, N., Choudhury, G., Harris, M. A., Wozney, J., Mundy, G. R., Abboud, S. L., and Harris, S. E., 1993: Autoregulation of mouse bmp-2 gene transcription is directed by the proximal promoter element. *Biochem. Biophys. Res. Commun.*, **286**(1), 101–8.

Gibson, D., Benders, G., Andrews-Pfannkoch, C., Denisova, E., Baden-Tillson, H., Zaveri, J., Stockwell, T., Brownley, A., Thomas, D., Algire, M., et al., 2008a: Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome. *Science's STKE*, **319**(5867), 1215.

Gibson, D., Benders, G., Axelrod, K., Zaveri, J., Algire, M., Moodie, M., Montague, M., Venter, J., Smith, H., and Hutchison, C., 2008b: One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic Mycoplasma genitalium genome. *Proceedings of the National Academy of Sciences*, **105**(51), 20404.

Gierer, A., and Meinhardt, H., 1972: A theory of biological pattern formation. *Kybernetik*, **12**, 30–39.

Glass, L., 2001: Synchronization and rhythmic processes in physiology. *Nature*, **410**(6825), 277–284.

Glossop, N., Lyons, L., and Hardin, P., 1999: Interlocked feedback loops within the Drosophila circadian oscillator. *Science*, **286**(5440), 766.

Golub, T., Slonim, D., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J., Coller, H., Loh, M., Downing, J., Caligiuri, M., et al., 1999: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, **286**(5439), 531.

Greenbaum, D., Colangelo, C., Williams, K., and Gerstein, M., 2003: Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol*, **4**(9), 117.

Grenier, F., Timofeev, I., and Steriade, M., 2003: Neocortical very fast oscillations (ripples, 80-200 Hz) during seizures: intracellular correlates. *Journal of Neurophysiology*, **89**(2), 841.

Griffith, L., and Schwartz, M. A., 2006: Capturing complex 3d tissue physiology in vitro. *Nat. Rev. Mol. Cell Biol.*, **7**, 211–223.

Guo, Y., Xiao, P., Lei, S., Deng, F., Xiao, G., Liu, Y., Chen, X., Li, L., Wu, S., and Chen, Y., 2008: How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes. *Acta biochimica et biophysica Sinica*, **40**(5), 426–436.

Gygi, S., Rochon, Y., Franza, B., and Aebersold, R., 1999: Correlation between protein and mRNA abundance in yeast. *Molecular and cellular biology*, **19**(3), 1720.

Hasty, J., McMillen, D., and Collins, J., 2002a: Engineered gene circuits. *Nature*, **420**(6912), 224–230.

Hasty, J., McMillen, D., and Collins, J. J., 2002b: Engineered gene circuits. *Nature*, **420**(6912), 224–30.

Ideker, T., Thorsson, V., Ranish, J., Christmas, R., Buhler, J., Eng, J., Bumgarner, R., Goodlett, D., Aebersold, R., and Hood, L., 2001: Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science*, **292**(5518), 929.

Isaacs, F., Hasty, J., Cantor, C., and Collins, J., 2003: Prediction and measurement of an autoregulatory genetic module. *PNAS*, **100**(13), 7714.

Isalan, M., Lemerle, C., Michalodimitrakis, K., Beltrao, P., Horn, C., Raineri, E., Garriga-Canut, M., and Serrano, L., 2008: Evolvability and hierarchy in rewired bacterial gene networks. *Nature*, **452**(7189), 840.

Jiang, Y., Aerne, B., Smithers, L., Haddon, C., Ish-Horowicz, D., and Lewis, J., 2000: Notch signalling and the synchronization of the somite segmentation clock. *Nature*, **408**(6811), 475–478.

Keiler, K., Waller, P., and Sauer, R., 1996: Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger rna. *Science*, **271**(5251), 990.

Kelly, F. P., 1979: *Reversibility and Stochastic Networks*. John Wiley and Sons, Chichester.

Kerckhoffs, R., McCulloch, A., Omens, J., and Mulligan, L., 2009: Effects of biventricular pacing and scar size in a computational model of the failing heart with left bundle branch block. *Medical Image Analysis*, **13**(2), 362–369.

Kobayashi, H., Kærn, M., Araki, M., Chung, K., Gardner, T., Cantor, C., and Collins, J., 2004: Programmable cells: interfacing natural and engineered gene networks. *Proceedings of the National Academy of Sciences*, **101**(22), 8414–8419.

Koch, A. J., and Meinhardt, H., 1994: Biological pattern formation: from basic mechanisms to complex structures. *Rev. Mod. Phys.*, **66**(4), 1481–1507.

Lakin-Thomas, P., and Brody, S., 2004: Circadian rhythms in microorganisms: new complexities. *Annu. Rev. Microbiol.*

Levchenko, I., Seidel, M., Sauer, R., and Baker, T., 2000a: A specificity-enhancing factor for the ClpXP degradation machine. *Science*, **289**(5488), 2354.

Levchenko, I., Seidel, M., Sauer, R. T., and Baker, T. A., 2000b: A specificity-enhancing factor for the ClpXP degradation machine. *Science*, **289**(5488), 2354–6.

Lewandowski, W., Azoubib, J., and Klepczynski, W., 1999: GPS: primary tool for time transfer. *Proceedings of the IEEE*, **87**(1), 163–172.

Li, D., Wong, K., Hu, Y., and Sayeed, A., 2002: Detection, classification and tracking of targets in distributed sensor networks. *IEEE Signal Processing Magazine*, **19**(2), 17–29.

Li, S., Armstrong, C., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P., Han, J., Chesneau, A., Hao, T., et al., 2004: A map of the interactome network of the metazoan C. elegans. *Science*, **303**(5657), 540.

Lies, M., and Maurizi, M., 2008: Turnover of endogenous SsrA-tagged proteins mediated by ATP-dependent proteases in Escherichia coli. *Journal of Biological Chemistry*, **283**(34), 22918.

Liu, D., Momb, J., Thomas, P., Moulin, A., Petsko, G., Fast, W., and Ringe, D., 2008: Mechanism of the Quorum-Quenching Lactonase (AiiA) from Bacillus thuringiensis. 1. Product-Bound Structures. *Biochemistry*, **47**(29), 7706–7714.

Locke, J., and Elowitz, M., 2009: Using movies to analyse gene circuit dynamics in single cells. *Nature Reviews Microbiology*, **7**(5), 383–392.

Loeser, R., Carlson, C. S., Tulli, H., Jerome, W. G., Miller, L., and Wallin, R., 1992: Articular-cartilage matrix gamma-carboxyglutamic acid-containing protein. characterization and immunolocalization. *Biochem. J.*, **282 (Pt 1)**, 1–6.

Lu, P., Vogel, C., Wang, R., Yao, X., and Marcotte, E., 2006: Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nature biotechnology*, **25**(1), 117–124.

Lutz, R., and Bujard, H., 1997a: Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic acids research*, **25**(6), 1203.

Lutz, R., and Bujard, H., 1997b: Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I_1-I_2 regulatory elements. *Nucleic Acids Res.*, **25**(6), 1203–10.

Ma, Y., and Yoshikawa, K., 2009: Self-sustained collective oscillation generated in an array of nonoscillatory cells. *Physical Review E*, **79**(4), 46217.

Martin, V., Pitera, D., Withers, S., Newman, J., and Keasling, J., 2003: Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nature biotechnology*, **21**(7), 796–802.

Mather, W., Bennett, M., Hasty, J., and Tsimring, L., 2009: Delay-Induced Degrade-and-Fire Oscillations in Small Genetic Circuits. *Physical Review Letters*, **102**(6), 68105.

Mather, W., Cookson, N., Hasty, J., Tsimring, L., and Williams, R., 2010: Correlation Resonance Generated by Coupled Enzymatic Processing. *Biophysical Journal*, **99**(10), 3172–3181.

Mather, W., Hasty, J., Tsimring, L., and Williams, R., 2011: Factorized time-dependent distributions for certain multiclass queueing networks and an application to enzymatic processing networks. *Queueing Systems*, 1–16.

McMillen, D., Kopell, N., Hasty, J., and Collins, J., 2002: Synchronizing genetic relaxation oscillators by intercell signaling. *Proceedings of the National Academy of Sciences*, **99**(2), 679–684.

Mirollo, R., and Strogatz, S., 1990: Synchronization of pulse-coupled biological oscillators. *SIAM Journal on Applied Mathematics*, 1645–1662.

Moore, S., and Sauer, R., 2005: Ribosome rescue: tmRNA tagging activity and capacity in Escherichia coli. *Molecular microbiology*, **58**(2), 456–466.

Murray, J., and Oster, G., 1984: Cell traction models for generating pattern and form in morphogenesis. *J. Math. Biol.*, **19**, 265–279.

Murray, J. D., 2003: On the mechanochemical theory of biological pattern formation with application to vasculogenesis. *C. R. Biol.*, **326**, 239–252.

Neher, S., Villén, J., Oakes, E., Bakalarski, C., Sauer, R., Gygi, S., and Baker, T., 2006: Proteomic profiling of ClpXP substrates after DNA damage reveals extensive instability within SOS regulon. *Molecular cell*, **22**(2), 193–204.

Nelson, C., Jean, R., Tan, J., Liu, W., Sniadecki, N., Spector, A., and Chen, C., 2005: From the Cover: Emergent patterns of growth controlled by multicellular form and mechanics. *Science's STKE*, **102**(33), 11594.

Nelson, C. M., VanDuijn, M. M., Inman, J. L., Fletcher, D. A., and Bissell, M. J., 2006: Tissue geometry determines sites of mammary branching morphogenesis in organotypic cultures. *Science*, **314**(5797), 298–300.

Ozbudak, E., Thattai, M., Kurtser, I., Grossman, A., and van Oudenaarden, A., 2002: Regulation of noise in the expression of a single gene. *nature genetics*, **31**(1), 69–73.

Ozbudak, E., Thattai, M., Lim, H., Shraiman, B., and Van Oudenaarden, A., 2004: Multistability in the lactose utilization network of Escherichia coli. *Nature*, **427**(6976), 737–740.

Painter, K. J., Maini, P. K., and Othmer, H. G., 1999: Stripe formation in juvenile Pomacanthus explained by a generalized Turing mechanism with chemotaxis. *PNAS*, **96**(10), 5549–5554.

Paliwal, S., Iglesias, P., Campbell, K., Hilioti, Z., Groisman, A., and Levchenko, A., 2007: MAPK-mediated bimodal gene expression and adaptive gradient sensing in yeast. *Nature*, **446**(7131), 46–51.

Pikovsky, A., Rosenblum, M., Kurths, J., and Hilborn, R., 2002: Synchronization: A universal concept in nonlinear science. *American Journal of Physics*, **70**, 655.

Price, P. A., Thomas, G. R., Pardini, A. W., Figueira, W. F., Caputo, J. M., and Williamson, M. K., 2002: Discovery of a high molecular weight complex of calcium, phosphate, fetuin, and matrix gla protein in the serum of etidronate-treated rats. *J. Biol. Chem*, **277**(6), 3926–3934.

Reading, N., and Sperandio, V., 2006: Quorum sensing: the many languages of bacteria. *FEMS microbiology letters*, **254**(1), 1–11.

Rosenfeld, N., Elowitz, M., and Alon, U., 2002: Negative autoregulation speeds the response times of transcription networks. *Journal of molecular biology*, **323**(5), 785–793.

Sage, A., Tintut, Y., Garfinkel, A., and Demer, L., 2009: Systems biology of vascular calcification. *Trends Cardiovasc Med*, **19**(4), 118–23.

Sauer, R., Bolon, D., Burton, B., Burton, R., Flynn, J., Grant, R., Hersch, G., Joshi, S., Kenniston, J., Levchenko, I., et al., 2004: Sculpting the proteome with AAA+ proteases and disassembly machines. *Cell*, **119**(1), 9–18.

Sauer, U., 2004: High-throughput phenomics: experimental methods for mapping fluxomes. *Current Opinion in Biotechnology*, **15**(1), 58–63.

Savageau, M., 1974: Comparison of classical and autogenous systems of regulation in inducible operons. *Nature*, **252**(5484), 546–549.

Schweder, T., Lee, K., Lomovskaya, O., and Matin, A., 1996: Regulation of escherichia coli starvation sigma factor (sigma s) by clpxp protease. *Journal of bacteriology*, **178**(2), 470.

Shi, Y., and Massague, J., 2003: Mechanisms of tgf-β signaling from cell membrane to the nucleus. *Cell*, **113**, 685–700.

Steen, E., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., del Cardayre, S., and Keasling, J., 2010: Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature*, **463**(7280), 559–562.

Stricker, J., Cookson, S., Bennett, M., Mather, W., Tsimring, L., and Hasty, J., 2008a: A fast, robust and tunable synthetic gene oscillator. *Nature*, **456**(7221), 516–519.

Stricker, J., Cookson, S., Bennett, M. R., Mather, W. H., Tsimring, L. S., and Hasty, J., 2008b: A fast, robust and tunable synthetic gene oscillator. *Nature*, **456**(7221), 516–9.

Sweatt, A., Sane, D. C., Hutson, S. M., and Wallin, R., 2003: Matrix gla protein (mgp) and bone morphogenetic protein-2 in aortic calcified lesions of aging rats. *J. Thromb. Haemost.*, **1**(1), 178–85.

Tabor, J., Salis, H., Simpson, Z., Chevalier, A., Levskaya, A., Marcotte, E., Voigt, C., and Ellington, A., 2009a: A synthetic genetic edge detection program. *Cell*, **137**(7), 1272–1281.

Tabor, J., Salis, H., Simpson, Z., Chevalier, A., Levskaya, A., Marcotte, E., Voigt, C., and Ellington, A., 2009b: A synthetic genetic edge detection program. *Cell*, **137**(7), 1272–1281.

Thomas, P., Stone, E., Costello, A., Tierney, D., Fast, W., et al., 2005: The quorum-quenching lactonase from Bacillus thuringiensis is a metalloprotein. *Biochemistry(Washington)*, **44**(20), 7559–7569.

Tigges, M., Marquez-Lago, T., Stelling, J., and Fussenegger, M., 2009: A tunable synthetic mammalian oscillator. *Nature*, **457**(7227), 309–312.

Tsimring, L., Levine, H., Aranson, I., Ben-Jacob, E., Cohen, I., Shochet, O., and Reynolds, W. N., 1995: Aggregation patterns in stressed bacteria. *Phys. Rev. Lett.*, **75**(9), 1859–1862.

Vilaprinyo, E., Alves, R., and Sorribas, A., 2010: Minimization of biosynthetic costs in adaptive gene expression responses of yeast to environmental changes. *PLoS Computational Biology*, **6**(2), 184–1037.

Vladimirov, A., Kozyreff, G., and Mandel, P., 2003: Synchronization of weakly stable oscillators and semiconductor laser arrays. *Europhysics Letters*, **61**(5), 613–619.

Wallin, R., Cain, D., Hutson, S. M., Sane, D. C., and Loeser, R., 2000: Modulation of the binding of matrix gla protein (mgp) to bone morphogenetic protein-2 (bmp-2). *J. Thromb. Haemost.*, **84**(6), 1039–44.

Warner, J., Vilardell, J., and Sohn, J., 2001: Economics of ribosome biosynthesis. In *Cold Spring Harbor symposia on quantitative biology*, volume 66, 567. Cold Spring Harbor Laboratory Press.

Waters, C., and Bassler, B., 2005: Quorum sensing: cell-to-cell communication in bacteria. *Annual Review of Cell and Developmental Biology*.

Webb, D. J., and Horwitz, A. F., 2003: New dimensions in cell migration. *Nat. Cell Biol.*, **5**, 690–692.

Wiesenfeld, K., Colet, P., and Strogatz, S., 1996: Synchronization transitions in a disordered Josephson series array. *Physical Review Letters*, **76**(3), 404–407.

Willette, R., Gu, J. L., Lysko, P. G., Anderson, K. M., Minehart, H., and Yue, T., 1999: Bmp-2 gene expression and effects on human vascular smooth muscle cells. *J. Vasc. Res.*, **36**, 120–125.

Williams, J., Cui, X., Levchenko, A., and Stevens, A., 2008: Robust and sensitive control of a quorum-sensing circuit by two interlocked feedback loops.

Williams, R. J., 1998: Diffusion approximations for open multiclass queueing networks: sufficient conditions involving state space collapse. *Queueing Systems*, **30**(1-2), 27–88.

Winfree, A., 1967: Biological rhythms and the behavior of populations of coupled oscillators. *Journal of theoretical biology*, **16**(1), 15.

Yochelis, A., Tintut, Y., and Demer. . . , L., 2008: The formation of labyrinths, spots and stripe patterns in a biochemical approach to cardiovascular calcification. *New Journal of* . . . .

You, L., Cox, I., R. S., Weiss, R., and Arnold, F. H., 2004a: Programmed population control by cell-cell communication and regulated killing. *Nature*, **428**(6985), 868–71.

You, L., Cox, R., Weiss, R., and Arnold, F., 2004b: Programmed population control by cell–cell communication and regulated killing. *Nature*, **428**(6985), 868–871.

Young, M., and Kay, S., 2001: Time zones: a comparative genetics of circadian clocks. *Nature Reviews Genetics*, **2**(9), 702–715.

Zaman, M., Matsudaira, P., and Lauffenburger, D. A., 2007: Understanding effects of matrix protease and matrix organization on directional persistence and translational speed in three-dimensional cell migration. *Ann. Biomed. Eng.*, **35**(1), 91–100.

Zaman, M. H., Trapani, L. M., Sieminski, A. L., MacKellar, D., Gong, H., Kamm, R. D., Wells, A., Lauffenburger, D. A., and Matsudaira, P., 2006: Migration of tumor cells in 3d matrices is governed by matrix stiffness along with cell-matrix adhesion and proteolysis. *Proc. Natl. Acad. Sci. USA*, **103**(29), 10889–10894.

Zebboudj, A. F., Imura, M., and Bostrom, K., 2002: Matrix gla protein, a regulatory protein for bone morphogenetic protein-2. *J. Biol. Chem.*, **277**(6), 4388–94.

Zebboudj, A. F., Shin, V., and Bostrom, K., 2003: Matrix gla protein and bmp-2 regulate osteoinduction in calcifying vascular cells. *J. Cell Biochem.*, **90**(4), 756–65.