Title
Applications of synthetic biology inspired by natural bacterial functionality /

Author
Samayoa, Phillip R.

Publication Date
2014

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Applications of synthetic biology inspired by natural bacterial functionality

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

in

Bioinformatics and Systems Biology

by

Phillip R. Samayoa

Committee in charge:

Jeff Hasty, Chair
Ruth Williams, Co-Chair
Steve Briggs
Alex Hoffmann
Kit Pogliano
Scott Rifkin
Lev Tsimring

2014

The dissertation of Phillip R. Samayoa is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego

2014

DEDICATION

To my friends and family for their love and support. To my educators for inspiring me.
To science for intriguing and challenging me.

EPIGRAPH

Nature does nothing uselessly.

Aristotle

TABLE OF CONTENTS

Signature Page iii

Dedication iv

Epigraph v

Table of Contents vi

List of Abbreviations ix

List of Figures x

List of Tables xii

Acknowledgements xiii

Vita xv

Abstract of the Dissertation xvi

Chapter 1 Introduction 1

Chapter 2 A sensing array of radically coupled genetic biopixels 7
- Introduction 7
- Synergistic Synchronization 9
- Sensing Array of Biopixels 13
- Scaling Up Synthetic Biology 17
- Quantitative Modeling 19
- Methods 21
  - Strains and Plasmids 21
  - Microfluidics and Microscopy 21
- Plasmid Construction 22
- Additional Experimental Results 24
- Data Analysis 31
- Microscopy and Microfluidics 31
- Modeling for Biosensor Array 34
- Supplementary Movies 40
- Acknowledgements 41

Chapter 3 Applications of Quorum Sensing Biology 42
  A fast and robust macroscopic toggle switch using synergistic coupling . . . 42
    Introduction 42
    Toggle Design 43
    Computational Modeling 43
    Results 44
  Redox Induced Toggling 46
  A population-controlled therapeutic oscillator 51

Chapter 4 Genetic circuits in *Salmonella typhimurium* 54
  Introduction 54
  Experimental Results 57
  Conclusions and outlook 62
  Plasmid Characterization 64
  Microscopy and Microfluidics 66
  Degradation and Production Rate Quantification 67
  Modeling 67
  Acknowledgements 70

Chapter 5 Programming cyanobacteria to produce protein nutrients at supra-agricultural efficiencies 72
  Introduction 72
  A Photosynthetic Nutrient Catalyst 74
  Methods 84
    Strains and plasmids 84
    Heterologous protein production 84
    Analytics 85
  Model Construction and Simulations 85
    Metabolic reconstruction for photosynthetic nutrient production 85
    Simulation constraints and objective function 87
  Experimental Methods 97
    High Throughput Cultivation 97
    Strain Selection and Growth Rate Testing 97
    Media Optimization 99
    Salinity Tolerance 99
    Promoter and Strain Construction 101
    Flask Experiments 105
  Nutrient Production 112
    Screening for Ovalbumin Production 112
    Lysis and Purification 117
  In silico Prediction of Cyanobacteria Sec-type Leader Sequences 118
    In-silico prediction model 118
    Screening predicted leader sequences 122

Acknowledgments 125

Chapter 6 Summary and Conclusion 126

Bibliography 127

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
GMO	genetically modified organism
hrs.	hours
kb	kilobase
LB	lysogeny broth
lbs	pounds
mRNA	messenger ribonucleic acid
OD	optical density
ODE	ordinary differential equation
PCC	Pasteur Culture Collection of Cyanobacteria
PCR	polymerase chain reaction
PDMS	poly(dimethylsiloxane)
s	second
UV	ultraviolet
YFP	yellow fluorescent protein
%w/v	percent weight per volume (concentration)

LIST OF FIGURES

Figure 2.1: Sensing array of radically coupled genetic biopixels 10
Figure 2.2: Frequency modulated genetic biosensor 14
Figure 2.3: Computational modeling of radical synchronization and biosensing . 16
Figure 2.4: Radical synchronization on a macroscopic scale 18
Figure 2.5: Plasmids for the coupled sensing array strain 23
Figure 2.6: Biopixels with NDH-2 engineered synchronization 24
Figure 2.7: Catalase degradation prevent colony communication 25
Figure 2.8: SodA produces excess H_2O_2, switching the cellular redox state 26
Figure 2.9: Synchronized oscillations across fluidically isolated devices 27
Figure 2.10: Heatmap of low fluorescence biopixel trajectories 28
Figure 2.11: Thioeurea effect on synchronized oscillations 29
Figure 2.12: Ampicillin effect on synchronized oscillations 30
Figure 2.13: Primary microfluidic device for biopixel array 33
Figure 2.14: Computational description of biopixel synchronicity 37
Figure 2.15: Computational demonstration of Antiphase behavior 38
Figure 2.16: Computational demonstration of alternating large and small amplitude oscillations 39

Figure 3.1: Circuit diagram for the macroscopic toggle switch system 44
Figure 3.2: Space-time diagrams of uncoupled and coupled toggle systems 45
Figure 3.3: Unbalanced coupled toggle switch 47
Figure 3.4: Redox-triggered toggle split media array 48
Figure 3.5: Split media chip design 49
Figure 3.6: H_2O_2 -induced toggling 50
Figure 3.7: Pattern of H_2O_2 mediated colony stimulation 50
Figure 3.8: Population-controlled oscillators 52
Figure 3.9: *E. coli* stripes formed by a population-controlled oscillator 53

Figure 4.1: A fast, robust, and tunable genetic oscillator in *S. typhimurium* 59
Figure 4.2: A synchronized quorum of genetic clocks in *S. typhimurium* 60
Figure 4.3: A genetic toggle switch in *S. typhimurium* 61
Figure 4.4: Computational modeling of *S. typhimurium* genetic circuits 63
Figure 4.5: Properties of plasmid-based circuits in *S. typhimurium* 65

Figure 5.1: Direct nutrient catalyst 76
Figure 5.2: Cyanobacterial host characterization 80
Figure 5.3: Optimal PAR flux per cell 82
Figure 5.4: Protein secretion 83
Figure 5.5: Light path characterization 88
Figure 5.6: Maximal translation capacity 90
Figure 5.7: Quantum efficiency 91

Figure 5.8: Culture density dependence 92
Figure 5.9: Carbon metabolism isocline 93
Figure 5.10: High density growth capacity 94
Figure 5.11: Areal protein production 95
Figure 5.12: High-light protein production efficiency 96
Figure 5.13: Long term PCC7002 cultivation 100
Figure 5.14: PCC7002 salt tolerance 101
Figure 5.15: pAQ1 P*cpc*-yfp* plasmid map 103
Figure 5.16: pAQ3 P*cpc*-cymO-yfp* plasmid map 105
Figure 5.17: P*cpc*-cymO-yfp* and P*cpc-cymO-yfp* expression 106
Figure 5.18: Comparison of P*cpc*-cymO-yfp* and P*cpc-cymO-yfp* expression strength in pAQ3 constructs 107
Figure 5.19: YFP standard curve 108
Figure 5.20: pAQ1 biomass efficiency 109
Figure 5.21: pAQ3 carbon metabolism 110
Figure 5.22: pAQ3 mRNA levels 113
Figure 5.23: Ovalbumin dot plot expression screen 115
Figure 5.24: pAQ3 P*cpc-ova* plasmid map 115
Figure 5.25: Ovalbumin caliper detection 116
Figure 5.26: Signal peptide neural network scoring 120
Figure 5.27: PCC7002 SignalP network score 121
Figure 5.28: Glucose standard curve 124

LIST OF TABLES

Table 4.1: Mean Degradation and Net Production Rate during *S. typhimurium* oscillation 67

Table 5.1: Protein production reaction set 87
Table 5.2: RT-PCR transcript quantification primers 112
Table 5.3: Ovalbumin expression constructs 114
Table 5.4: Prediction statistics for trained S and C networks 119

ACKNOWLEDGEMENTS

Reflecting on my journey to complete this PhD, I'm very much stricken by the number of people who allowed me to have such an incredible experience. It's hard not to first think back to everyone in my life who prepared me for this opportunity. I owe so much to my family and friends for their love, support, and pride through every ambition I've ever chased. I'm thankful for my wonderful teachers who challenged me (Mrs. Cuello), mentored me (Prof. Manalis and Natalie Kuldell), and nurtured my love of science (Mrs. McDaniel). I would like to thank my close friends (Neil, Eric, Dom, Josh, Pall, and of course Maggie) who made my time in San Diego so special through timely distractions and their continuous motivation. I will never forget 121 30th street. I want to thank Lauren for making me so happy through some of the most challenging periods of this process. A questionable thank you to Parker for pulling me away from work to take him on walks. I would like to thank everyone in the lab who helped me along the way, from taking the time to teach me lab techniques from scratch (Bridget) to navigating the requirements of a graduate department. In particular, I want to thank Tal for patiently passing on all of his wisdom on molecular cloning and inspiring my projects through his incredibly creative approach to research. I'm thankful for my peers and friends in the lab (Arthur, Brooks, and Jangir), who made the daily life of research something I looked forward to. I feel extremely lucky to have had the opportunity to collaborate with Arthur so closely throughout the years. I'm indebted to the older students (Mike, Natalie, Martin, Ivan) who paved the way for my research through their incredible work to build out the capabilities of our lab. I also want to thank the undergraduates who I had the chance to work with during my thesis (Omar and Anna). I'm so happy to have had a co-advisor in Lev Tsimring for expanding the way I think about biological systems. His ability to distill complex dynamical phenomena that arise in biology into insightful principles amazes me. Finally, I'm deeply grateful for my advisor, Jeff Hasty, for his mentorship, flexibility, and tremendous leadership in pushing us to achieve breakthrough science. His creativity and unique ability to powerfully communicate science has truly had an impact on me. I cannot thank him enough for this experience.

Chapter 2 contains material originally published as Prindle, A.*, Samayoa, P.*, Razinkov, I., Danino, T., Tsimring, L. and Hasty, J. ***Nature*** (2012): A sensing array of radically coupled genetic 'biopixels'. (*equal contribution). Copyright permission to republish here was granted by Nature publishing group.

Chapter 4 contains material originally published as Prindle, A., Selimkhanov, J., Danino, T., Samayoa, P., Goldberg, A., Bhatia, S., and Hasty, J. ***ACSSynBio*** (2012): Genetic circuits in *Salmonella typhimurium*. Copyright permission to republish here was granted by ACS Publications.

Chapter 5, in part, is currently being prepared for submission for publication of the material. Samayoa, P. Von Maltzahn, G., Boghigian, B., Basu, S., Young, D., Silver, N., Hamill, M. and Berry, D: Programming cyanobacteria to produce protein nutrients at supra-agricultural efficiencies.

VITA

2009	Bachelor of Science in Physics Massachusetts Institute of Technology
2009	Bachelor of Science in Biological Engineering Massachusetts Institute of Technology
2014	Doctor of Philosophy in Bioinformatics and Systems Biology University of California, San Diego

PUBLICATIONS

Peer Reviewed Journal Articles

A sensing array of radically coupled genetic ’biopixels’. Prindle, A.*, Samayoa, P.*, Razinkov, I., Danino, T., Tsimring, L., and Hasty, J. ***Nature*** (2012)

Genetic circuits in *Salmonella typhimurium*. Prindle, A., Selimkhanov, J., Danino, T., Samayoa, P., Goldberg, A., Bhatia, S., and Hasty, J. ***ACSSynBio*** (2012)

Programming cyanobacteria to produce protein nutrients at supra-agricultural efficiencies. Samayoa, P. Von Maltzahn, G., Boghigian, B., Basu, S., Young, D., Silver, N., Hamill, M. and Berry, D. (in preparation)

FIELDS OF STUDY

Major Field: Bioinformatics and Systems Biology (Synthetic Biology)

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

ABSTRACT OF THE DISSERTATION

Applications of synthetic biology inspired by natural bacterial functionality

by

Phillip R. Samayoa

Doctor of Philosophy in Bioinformatics and Systems Biology

University of California, San Diego, 2014

Jeff Hasty, Chair
Ruth Williams, Co-Chair

Synthetic biology has enabled the engineering of biological networks capable of producing quantitatively predictable, dynamic function in a host organism. Advanced synthetic toolkits have been developed and applied within model bacterial systems to construct functional assemblies of biological parts that toggle (Gardner et al., 2000), oscillate (Elowitz and Leibler, 2000; Stricker et al., 2008; Danino et al., 2010), control cellular populations (You et al., 2004) and trigger host responses (Lu and Collins,

2007), form logical gates (Tamsir et al., 2010; Moon et al., 2012), perform arithmetic operations (Friedland et al., 2009), filter signals (Sohka et al., 2009) and sense complex environmental cues (Tabor et al., 2009; Kobayashi et al., 2004). This diverse set of genetic circuit designs, varying in complexity and host, have further been applied to systematically engineer biological systems towards applications in biosynthesis, environmental remediation, intracellular diagnostics, and therapeutics. Amidst this rapid expansion of tools and extensive demonstration of complex functional capabilities, the lack of commercially translatable technologies thus far reflects a looming challenge for the field. Specifically, an integrated synthetic design must go beyond matching signal levels between interacting circuits in model organisms. It must also function effectively across relevant spatial scales and among industrial hosts, for which the genetic circuit must predictably integrate with the host's regulatory and metabolic biology. In this thesis, we focus on constructing scalable synthetic circuits that are inspired by the natural functionality of bacteria and translatable into commercially relevant hosts. Here we discuss host selection, modeling, design, construction, and characterization of gene circuits towards applications in biosensing and protein production. In Chapter One, we give an introduction to the field of synthetic biology and how our research area fits into this discipline. In Chapter Two, I describe a circuit design that leverages native regulatory components of bacteria to produce a scalable frequency modulated biosensor. In Chapter Three, I demonstrate further functionalities of quorum sensing signaling in bacteria. In Chapter Four, we discuss the translation of synthetic circuits to a therapeutically relevant host. In Chapter Five, I present an integrated approach to engineer industrially relevant cyanobacteria for efficient protein production. These sections combine to expand our approach in engineering natural biological systems towards directed, translatable applications.

Chapter 1

Introduction

The study of biology has transformed amid a wave of technological innovation that enabled rapid sequencing of genetic material and provided the tools with which to rapidly synthesize libraries of biological elements, including entire microbial genomes. The field of systems biology emerged in a push to direct rich sequencing datasets toward unraveling patterns of connectivity that enable complex genetic networks to exhibit functional behavior. Moreover, a desire to demonstrate control of such networks led to the emergence of synthetic biology. The ability to rewire endogenous gene regulatory systems afforded by synthesis technologies allows rapid swapping of functional biological parts (not limited to genes, promoters, repressors, enhancers) within and between organisms to explore network design principles.

Targeting the predictable and modular design of gene circuits, synthetic biology has sought to rigorously characterize the dynamics of smaller motifs in isolation before scaling complexity towards larger networks. This is reflected by the fact that the majority of existing examples of synthetic gene circuits consist of fewer than 10 genes. The bottom-up approach has brought forward powerful design principles that occur repeatedly across natural gene networks. An ability to further reconstruct circuits with externally tunable parts derived from a variety of organisms has allowed synthetic biologists to build small regulatory motifs, inspired by nature, that allow for its quantitative characterization captured by computational modeling. This work has fostered a next generation of circuit design, supported by a careful understanding of the dynamics spanning gene expression through regulation, to construct novel genetic functions that switch, perform logical operations, and trigger host responses.

Yet this reductionist approach reaches a stark discontinuity with our technological ability to mine rich sequencing information across a diversity of organisms to access novel, transferable biological elements. Furthermore, advancement in synthesis capabilities have yielded demonstrations of reconstructing artificial genomes in excess of 1,000 genes (Gibson et al., 2010). The difficulty in transitioning the principles elucidated by bottom-up synthetic biology research into more complex, industrially applicable synthetic systems represents the lack of well-characterized genetic components which can seamlessly be integrated into larger devices as well as the ability to predictably design according the host's regulatory and metabolic biology to construct robust systems.

A significant amount of consideration in the field has been directed towards more efficient "part" screening, characterization, and classification to support the construction of more complex genetic devices (Wang et al., 2013). Minimizing crosstalk between components has been a particular focus, as increasing the scale of designed genetic interactions will require a precise understanding of pairwise regulatory dynamics across the system to achieve the predicted behaviors. Screened libraries of orthogonal genetic elements has provided more specific pairing, with the mathematical ability to predict off-target interactions within a multilayer genetic circuit (Mutalik et al., 2012). Properly matching elements such that their input and output levels operate within a consistent dynamic range is also necessary to engineer complex functionality and represents a significant design bottleneck. Poor input-output alignment can manifest in differences in concentration, stoichiometry, enzymatic activities, or expression levels and lead to genetic instability, protein mislocalization, accumulation of toxic intermediates, and growth inhibition.

More fundamentally, inconsistent output levels between parts as well as variability in measurement techniques can lead to poor device sensitivity. In the case of applications for synthetic circuits to detect input stimuli, optical or chemical reporters require careful instrument calibration and signal normalization. In Chapter 2 we address these design challenges by constructing an oscillatory sensor that digitizes the output of a genetic sensor by converting input simuli into the frequency space. This enables the design of multiple sensing strategies, including a threshold sensor with an On/Off output as well as a frequency modulated sensor.

The host's regulatory and metabolic biology also plays an important role in the construction of more advanced genetic circuits. As the complexity of genetic elements is scaled, the need to reduce crosstalk and level-match ”parts” is joined by the challenge of designing around host interference. This represents both the effect of a synthetic network on the host's native functionality as well as the host's unintended impact on the gene circuit dynamics. Traditionally work has aimed to introduce orthogonal parts, including transcriptional machinery via novel polymerases or translational machinery by introducing unnatural amino acid biology via mutant tRNAs (Rackham and Chin, 2005; Chin et al., 2003; Chatterjee et al., 2012). Further obstacles include designing for evolutionary stability, whereby the circuit is not subject to host adjustments of circuit expression level that may cause unprogrammed changes in synthetic circuit activity or performance over time.

Another approach to build complexity into a synthetic circuit is to specifically harness the native regulatory networks of the host in its design. In Chapter 2, we demonstrate the ability to utilize native redox-sensing machinery of *E. coli* to couple the behavior of genetic circuits across individual cells as well as between colonies. This work further demonstrates the ability to scale synthetic circuits spatially, generating precise functionality between cells across large populations and distances. As cellular activity is governed by stochastic bursts of transcription and translation, reducing intercellular variability in circuit behavior represents yet another challenge to designing complex networks more robustly.

Similarly, design considerations for transitioning an engineered circuit into the relevant environment for which it was designed has not carefully been addressed in

”bottom-up” synthetic biology research. Recent work has demonstrated that even simple circuits can have difficulties translating towards industrial scenarios (Moser et al., 2012). Designing circuits that function in hosts that are more amenable to industrial scaling and applicable media conditions is crucial to translating synthetic biology. In the therapeutics space, for example, a trigger circuit has recently been demonstrated in-vivo using hosts that reside naturally in the human gut (Kotula et al., 2014). In Chapter 4, we demonstrate the ability to construct quorum-mediated circuits in the therapeutically relevant microbe, *S. typhimurium*. Attenuated strains of this bacteria have exhibited safety in several human clinical trials and specifically grow in tumor environments, providing a chassis capable of facilitating a functional synthetic behavior at a relevant point of intervention.

To realize the potential of novel circuit designs, synthetic biology must aim to harness the natural functionality exhibited in biological organisms towards useful applications. Bacteria have adapted sophisticated modalities for survival across nearly all environmental conditions via metabolic and regulatory evolution. Across this phylogenetic diversity, the scope of native bacterial phenotypes may provide useful frameworks for synthetic design. Highly sensitive regulatory modules have evolved in bacteria based on pressures to sample and respond to environmental cues. Cells have developed precise intercellular communication strategies to preserve energy for host colonization (Fuqua et al., 1994), metabolic functionalities as both individual cells and larger communities to achieve symbiosis with complex host biologies (Kiers et al., 2003; Lundberg et al., 2012; Human Microbiome Project Consortium and others, 2012), as well as community strategies to overcome threats to the microbial ecosystem (Lee et al., 2010). Autotrophic bacteria have specifically optimized their photosystems to utilize incident light for efficiently producing chemical energy (Nogales et al., 2012). Finding the intersection between a significant unmet need and the evolutionary niche of a biological organism will likely yield the most profound results in this field. Likewise the development of synthetic biology toolkits that can be implemented across host types to utilize these diverse functionalities will significantly expand the field. In Chapter 5, we present an integrated approach to engineer an industrially relevant cyanobacteria for efficient protein production, including the characterization of new synthetic biology components for engineering marine autotrophs. This work demonstrates the capacity to re-direct a specific, highly complex biological functionality towards the following crucial problem:

The world's population, and its demand for food, is growing at a rate that cannot be sustained by our current agricultural infrastructure. With food demand expected to rise between 70-100% by 2050, driven by increased urbanization of a growing population projected to reach 9.1 billion, annual cereal and meat production will need to double (rom, 2009). While the green revolution provided a dramatic rise in productivity of our major wheat and rice food crops, resulting in a doubling of grain production over the past 50 years, arable agricultural land has only increased 9% globally (Godfray et al., 2010). Looking forward, 90% of the required increase in crop production is expected to come from higher cropping intensity and productivity, closing yield gaps globally, with only 5% from arable land expansion (rom, 2009). Yet wheat and rice annual yield increases are now slowing to below 1%, including in developing countries that suffer from the highest yield deficiencies (Fischer et al., 2009). Moreover, the impact of climate change on cropping regions worldwide is beginning to drive a measurable decline in yield and arable land (Lobell et al., 2011; Zhang and Cai, 2011).

To ensure the security of future food production, meeting increased demand requires a sustainable solution. A rapidly increasing demand for sources of concentrated protein and vitamins from meat and dairy poses unique challenges. Low production efficiencies (10% plant to animal conversion) occupies 33% of the world's cereal, while leaving a serious environmental footprint due to a freshwater demand of $>$100L and methane emissions surpassing 115 kg CH_4 every day per head of cattle (Godfray et al., 2010; Goodland et al., 2009). Using a single cell marine cyanobacteria, we leverage synthetic biology to design and characterize a sustainable alternative to crop and animal-based protein production. Streamlining their highly optimized photosynthetic capacity towards direct protein production enables a significant leap in productivity over conventional practices.

As the field of synthetic biology seeks to increase complexity through assembling large sets of orthogonal genetic elements, building off of the complex biology already perfected by natural organisms may far outpace our ability to design from scratch. In this thesis, I will implement such an approach to engineer natural biological systems toward directed, translatable applications.

Chapter 2

A sensing array of radically coupled genetic biopixels

Introduction

While there has been significant progress in the development of engineering principles for synthetic biology, a substantial challenge is the construction of robust circuits in a noisy cellular environment. Such an environment leads to considerable intercellular variability in circuit behavior, which can hinder functionality at the colony level. Here, we engineer the synchronization of thousands of oscillating colony ”biopixels” over centimetre length scales through the use of synergistic intercellular coupling involving quorum sensing within a colony and gas-phase redox signaling between colonies. We use this platform to construct an LCD-like macroscopic clock that can be used to sense arsenic via modulation of the oscillatory period. Given the repertoire of sensing capabilities of bacteria such as *E. coli*, the ability to coordinate their behavior over large length scales sets the stage for the construction of low cost genetic biosensors that are capable of detecting heavy metals and pathogens in the field.

Synthetic biology can be broadly parsed into the ”top-down” synthesis of genomes (Gibson et al., 2010) and the ”bottom-up” engineering of relatively small genetic circuits (Hasty et al., 2002; Sprinzak and Elowitz, 2005; Endy, 2005; Ellis et al., 2009; Kobayashi et al., 2004; You et al., 2004; Basu et al., 2005; Mukherji and Van Oudenaarden, 2009; Grilly et al., 2007). In the genetic circuits arena, toggle switches (Gardner et al., 2000) and oscillators (Elowitz and Leibler, 2000) have progressed into triggers (Lu and Collins, 2007), counters (Friedland et al., 2009) and synchronized clocks (Danino et al., 2010). Sensors have arisen as a major focus in the context of biotechnology (Tamsir et al., 2010; Tabor et al., 2009; Kobayashi et al., 2004), while oscillators have provided insights into the basic-science functionality of cyclic regulatory processes (Stricker et al., 2008; Mondragón-Palomino et al., 2011; Tigges et al., 2009). A common theme is the concurrent development of mathematical modeling that can be used for experimental design and characterization, as in physics and the engineering disciplines.

The synchronization of genetic clocks provides a particularly attractive avenue for synthetic biology applications. Oscillations permeate science and technology in a number of disciplines, with familiar examples including AC power (Westinghouse, 1887), GPS (Lewandowski et al., 1999), and lasers (Vladimirov et al., 2003). These technologies have demonstrated that operating in the frequency domain can offer significant advantages over steady-state designs in terms of information gathering and transmission. In particular, oscillatory sensors confer a number of advantages to traditional ones (Gast, 1985), since frequency is easily digitized and can be quickly updated with repeated measurements. For sensors that use optical reporters, measurements of frequency are less sensitive to experimental factors such as beam power and exposure time than intensity measurements which must be normalized and calibrated.

While the bottom-up approach to synthetic biology is increasingly benefiting from DNA synthesis technologies, the general design principles are still evolving. Within this context, a substantial challenge is the construction of robust circuits in a cellular environment that is governed by noisy processes such as random bursts of transcription and translation (Ozbudak et al., 2002; Elowitz et al., 2002; Golding et al., 2005; Blake et al., 2006; Austin et al., 2006). Such an environment leads to considerable intercellular variability in circuit behavior, which can hinder their functionality at the colony level.

An ideal design strategy for reducing variability across a cellular population would involve both strong and long-range coupling that would instantaneously synchronize the response of millions of cells. Quorum sensing typically involves strong intercellular coupling over tens of microns (Waters and Bassler, 2005; Basu et al., 2005; Danino et al., 2010), yet the relatively slow diffusion time of molecular communication through cellular media leads to signaling delays over millimetre scales. Faster communication mechanisms, such as those mediated in the gas phase, may increase the length scale for instantaneous communication, but are comparatively weak and short-lived since the vapor species more readily disperse.

Synergistic Synchronization

In order to develop a frequency modulated biosensor, we designed a gene network capable of synchronizing genetic oscillations across multiple scales (Fig. 2.1a and Fig. 2.5). We constructed an LCD-like microfluidic (Ferry et al., 2011) array that allows many separate colonies of sensing bacteria to grow and communicate rapidly by gas exchange (Fig. 2.1b and Fig. 2.13). Since previous work (Danino et al., 2010) has demonstrated that coupling through quorum sensing leads to incoherent oscillations at the millimetre scale, this mode of cellular communication is too slow for the generation of synchronized oscillations at the macroscopic scale. However, the slower quorum sensing can be used to synchronize small local colonies, provided there is a second level of design that involves faster communication for coordination between the colonies. Therefore rather than attempting to engineer a sensor from a single large-colony oscillator, we wired together thousands of small oscillating colonies, or ”biopixels”, in a microfluidic array. Coupling between biopixels involves redox signaling by hydrogen peroxide (H_2O_2) and the native redox sensing machineries of *E. coli*. The two coupling mechanisms act synergistically in the sense that the stronger, yet short-range, quorum sensing is necessary to coherently synchronize the weaker, yet long-range, redox signaling. Using this method we demonstrate synchronization of approximately 2.5

million cells across a distance of 5 mm, over 1,000 times the length of an individual cell (Fig. 2.1c-d and Supplementary Movies 1 and 2). This degree of synchronization yields extremely consistent oscillations, with a temporal accuracy of about 2 minutes compared to 5 - 10 minutes for a single oscillator (Danino et al., 2010) (Fig. 2.1d).

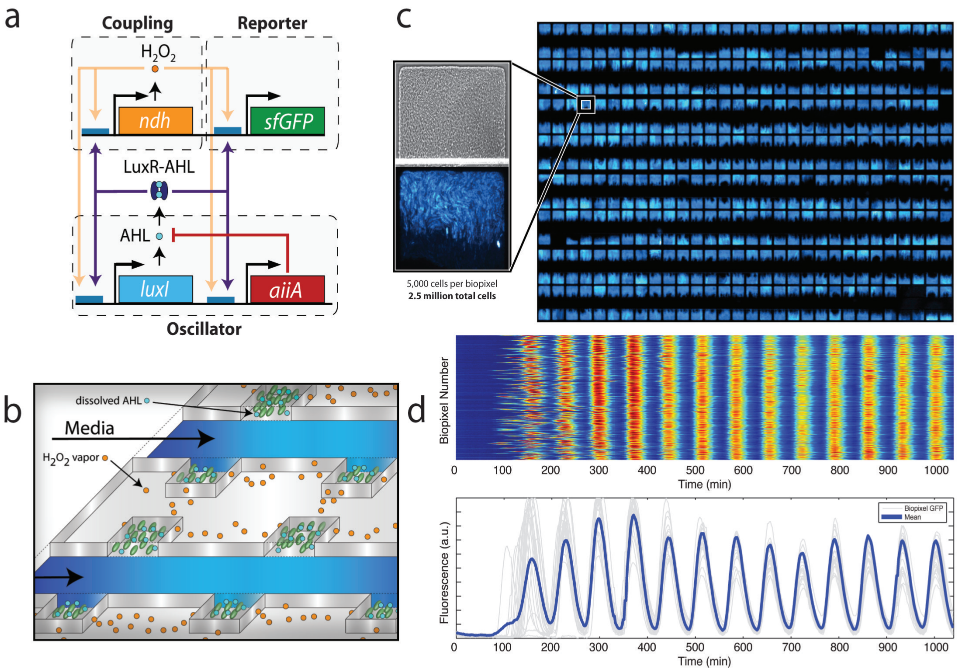


Figure 2.1: Sensing array of radically coupled genetic biopixels. (a) Network diagram. The *luxI* promoter drives expression of *luxI*, *aiiA*, *ndh*, and *sfGFP* in four identical transcription modules. The quorum-sensing genes *luxI* and *aiiA* generate synchronized oscillations within a colony via AHL. The *ndh* gene codes for NDH-2, an enzyme that generates H_2O_2 vapor which is an additional activator of the *luxI* promoter. H_2O_2 is capable of migrating between colonies and synchronizing them. **(b)** Conceptual design of the sensing array. AHL diffuses within colonies while H_2O_2 migrates between adjacent colonies through the PDMS. Arsenite-containing media is passed in through the parallel feeding channels. **(c)** Fluorescent image of an array of 500 *E. coli* biopixels containing about 2.5 million cells. Inset: brightfield and fluorescent images display a biopixel of 5,000 cells. **(d)** Heatmap and trajectories depicting time-lapse output of 500 individual biopixels undergoing rapid synchronization. Sampling time is 2 minutes.

The global synchronization mechanism is comprised of two modes of communication that work on different scales. The quorum-sensing machinery (LuxI, AiiA) uses an acyl-homoserine lactone (AHL) to mediate intracolony synchronization. In our device, the degree to which neighboring colonies are able to influence each other via AHL diffusion is negligible owing to the high media channel flow rates. Instead, we engineered the cells to communicate via gas exchange by placing a copy of the gene coding for NADH dehydrogenase II (NDH-2) under the control of an additional *lux* promoter. NDH-2 is a membrane-bound respiratory enzyme that produces low levels of H_2O_2 and superoxide (O_2^-) (Messner and Imlay, 1999). Since H_2O_2 vapor is able to pass through the 25 μm oxygen-permeable PDMS walls that separate adjacent colonies, periodic production of NDH-2 yields periodic exchange of H_2O_2 between biopixels. When H_2O_2 enters the cell, it transiently changes its redox state, interacting with our synthetic circuit through the native aerobic response control systems, including arcAB which has a binding site in the lux promoter region (Bose et al., 2007; Georgellis et al., 2001). Under normal conditions, ArcAB is partially active so *lux* is partially repressed. In contrast, oxidizing conditions triggered by H_2O_2 inactivate ArcAB, relieving this repression. Each oscillatory burst promotes firing in neighboring colonies by relieving repression on the *lux* promoter. This constitutes an additional positive feedback that rapidly synchronizes the population (Fig. 2.6 and Supplementary Movie 1).

We investigated the effects of catalase and superoxide dismutase (SOD) to probe the nature of H_2O_2 communication. When a population of synchronized colonies was exposed to a step increase of 200 U/ml catalase, an enzyme that rapidly degrades extracellular H_2O_2 (Seaver and Imlay, 2001), synchronization was broken and colonies continued to oscillate individually (Fig. 2.7). Since the cell membrane is impermeable to catalase, this confirms that communication between colonies depends on external H_2O_2 while oscillations within a colony do not. Conversely, when we enhanced the rate of superoxide conversion to H_2O_2 by expressing *sodA* (Fridovich, 1978; McCord and Fridovich, 1969) from an additional *lux* promoter, colonies quickly fired in a spatial wave and failed to oscillate further despite no changes to growth rate or cell viability (Fig. 2.8). Since H_2O_2 is produced internal to the cell, this confirms that H_2O_2 is capable of escaping the cell and activating lux-controlled genes in neighboring colonies via diffusion. The apparent higher output of H_2O_2 by SOD as compared to NDH-2 is likely due to its very high catalytic efficiency (Berg et al., 2006). Finally, we observed synchronization between arrays of traps even when they were fluidically isolated but held in close proximity (Fig. 2.9). These devices share no common fluid sources or channels, making communication by dissolved molecules like AHL impossible. Taken together, these results confirm that gaseous H_2O_2 is the primary mode of communication between oscillating colonies.

Based on our understanding of the mechanism for global synchronization, we expected that we could simplify the circuitry by eliminating *ndh* and achieve the same effect with intermittent bursts of high-intensity blue light. In this design, the GFP molecule acts as a photosensitizer, releasing free radicals upon exposure that produce oxygen species (ROS) including H_2O_2 (Remington, 2006). At the peak of oscillation, significant vapor-phase H_2O_2 is produced by exposing GFP-containing cells to fluorescent light. Conversely, at the trough of oscillation, cells contain almost no GFP, and therefore produce very little H_2O_2 upon fluorescing. Bursts of light thus generate bursts of H_2O_2 vapor whose concentration depends on the oscillating GFP level, just as periodic production of NDH-2 did previously. Indeed, this strategy was similarly able to synchronize our sensor array (Fig. 2.1d and Supplementary Movie 2). Numerous controls were performed to ensure that synchronized oscillations did not occur at low fluorescence intensities (Fig. 2.10 and Supplementary Movie 9).

To probe this mode of synchronization, we investigated the effects of thiourea and the antibiotics ampicillin and kanamycin. When a synchronized population of colonies was exposed to 35 mM thiourea, a potent radical quencher (Kelner et al., 1990; Touati et al., 1995), we observed sharply decaying synchronized oscillations while growth rate and cell viability were unaffected (Fig. 2.11). This suggests that without O_2^-, oscillations cannot be produced. Next, we ran a series of experiments switching the antibiotic resistance genes on our plasmids. We noted that radical-producing antibiotics (Kohanski et al., 2010), particularly Ampicillin, significantly reduced the degree of synchronization, showing that an excess of radical species such as O_2^- also hinders communication (Fig. 2.12). Since our final constructs included a plasmid with kanamycin resistance, which was also found to produce some radicals, we used full (50 μg/ml) selection when growing up the cells but very low (5 μg/ml) selection during the experimental run. Persistence of oscillations, sequencing, and subsequent growth in full selection following the run confirmed the presence of all 3 plasmids despite this low experimental selection. Catalase and sodA results were identical to those with NDH-2 synchronization. These results show that fluorescence-mediated synchronization involves the production of radical species following fluorescence exposure and communication via H_2O_2.

Sensing Array of Biopixels

With a platform for generating consistent and readily-detectable oscillations, we sought to use the circuit to engineer an arsenic-sensing macroscopic biosensor. We rewired the network to include an extra copy of the positive-feedback element, the AHL-synthase LuxI, under the control of a native arsenite-responsive promoter which is repressed by ArsR in the absence of arsenite (Fig. 2.2a). When arsenite is not present in the media, supplemental *luxI* is not transcribed and the circuit functions normally, generating baseline oscillations. However, the addition of trace amounts of arsenite relieves this repression and allows supplemental *luxI* to be produced, increasing the oscillatory amplitude and period. Tuning the level of LuxI by varying arsenite concentration results in clear changes to the oscillatory period (Fig. 2.2b and Supplementary Movie 2). To determine the range of detection, we swept arsenite concentrations from 0 to 1 μM and measured the oscillatory period (Fig. 2.2c, top). Using statistical methods (Supplementary Data Analysis), we generated a sensor calibration curve (Fig. 2.2c, bottom) that depicts the maximum possible arsenite concentration present (α = 95%) for a given measured period. This curve is an illustration of how data generated by our array would be used to measure arsenite concentrations in an unknown sample using our device. Our system was able to reliably quantify arsenite levels as low as 0.2 μM, below the 0.5 μM WHO-recommended level for developing nations (Nordstrom, 2002).

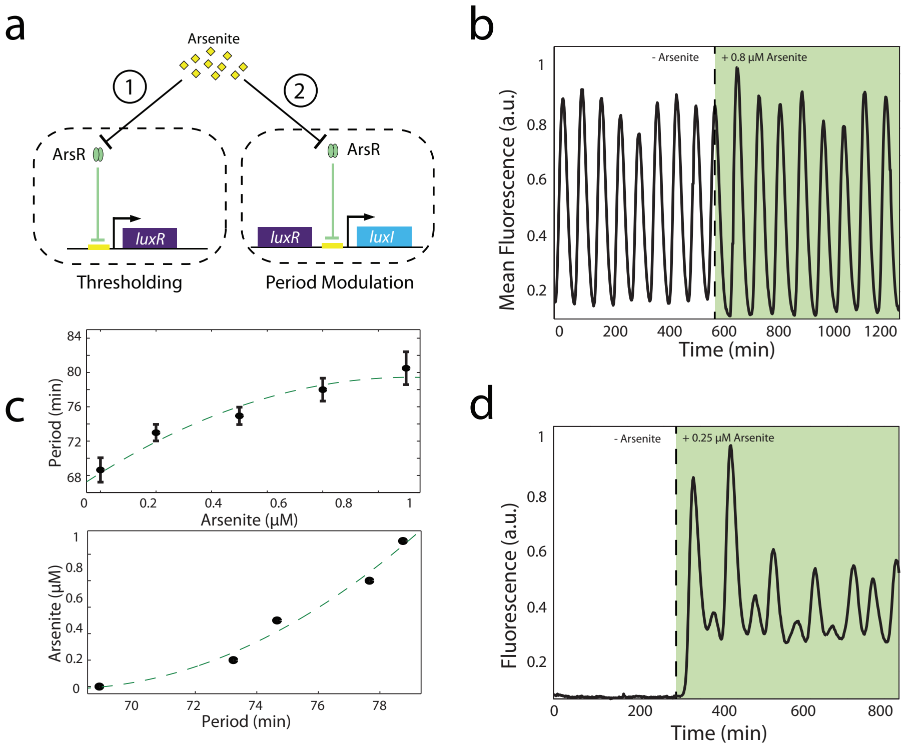


Figure 2.2: Frequency modulated genetic biosensor. **(a)** Network diagrams depicting two constructed sensing modules. In thresholding (1), the *luxR* gene is removed from the oscillator network and supplemented by a new copy driven by an arsenic-responsive promoter. In period modulation (2), a supplemental *luxI* gene tagged for increased degradation is driven by the arsenic-responsive promoter which affects the period of oscillation. **(b)** A sample period modulation sensor output following a step increase of 0.8 μM arsenite. Oscillatory period increases from 69 minutes to 79 minutes. **(c)** (Top) Period versus arsenite concentration for the sensor array. Error bars indicate ± 1 standard deviation averaged over 500 biopixel trajectories. Dotted line represents model-predicted curve. (Bottom) Sensor calibration curve generated from experimental data. Points indicate the maximum arsenite level with 95% certainty for a given measured period as determined statistically from experimental data. **(d)** Thresholder output following a step increase of 0.25 μM arsenite. A dramatic shift from rest to oscillatory behavior is observed within 20 minutes following the addition of arsenite.

As an alternative sensing strategy, we rewired the network to include a copy of the *luxR* gene controlled by an arsenic-responsive promoter while removing it from the rest of the circuit (Fig. 2.2a). Since the LuxR-AHL complex must be present to activate the *lux* promoter (Waters and Bassler, 2005), cells produce no LuxR when the media is free of arsenite, generating no fluorescence or oscillations. The addition of arsenite stimulates the production of LuxR, restoring circuit function and producing clear, synchronized oscillations (Fig. 2.2d and Supplementary Movie 3). This ON/OFF detection system has a threshold of 0.25 μM, a detection limit that can be adjusted by changing the copy number, ribosome binding site (RBS) strength, or promoter strength of the sensing plasmid (Supplementary Methods).

The sensing array is also capable of producing complex behaviors arising from the dynamic interaction of cellular colonies. By making modifications to the size, number, and arrangement of biopixels in the device, we are able to dramatically alter the output waveforms. For example, when we constructed a device in which trap separation distance is increased (45 μm versus 25 μm), we observed local anti-phase synchronization between neighboring colonies (Fig. 2.3d, top right). To explore this phenomenon on a larger scale, we constructed a device that contains an array of 416 traps constructed according to the specifications above. In these experiments, we observe initial global synchronization that gradually falls into local anti-phase synchronization across the array (Fig. 2.3d, middle and Supplementary Movie 4). Phase alignment is maintained over at least 48 hours, with patches of synchronization typically 3-6 colonies in size. Alternatively, by changing dimensions such that the array contains traps of two slightly different sizes, we observe a 1:2 resonance synchronization where larger traps pulse at double the frequency of smaller traps while maintaining synchronization (Fig. 2.3d, top left and Supplementary Movie 7). Lastly, when LuxR is limited as in the thresholding scheme, we observe synchronized oscillations of alternating large and small peaks in both experiment and model (Fig. 2.16). Our computational model (see Modeling Box) captures these effects (Fig. 2.3d, bottom and Figs 2.15 and 2.16) and indicates that further array manipulation will yield new, richer dynamics that could not be produced directly by changing circuit structure.

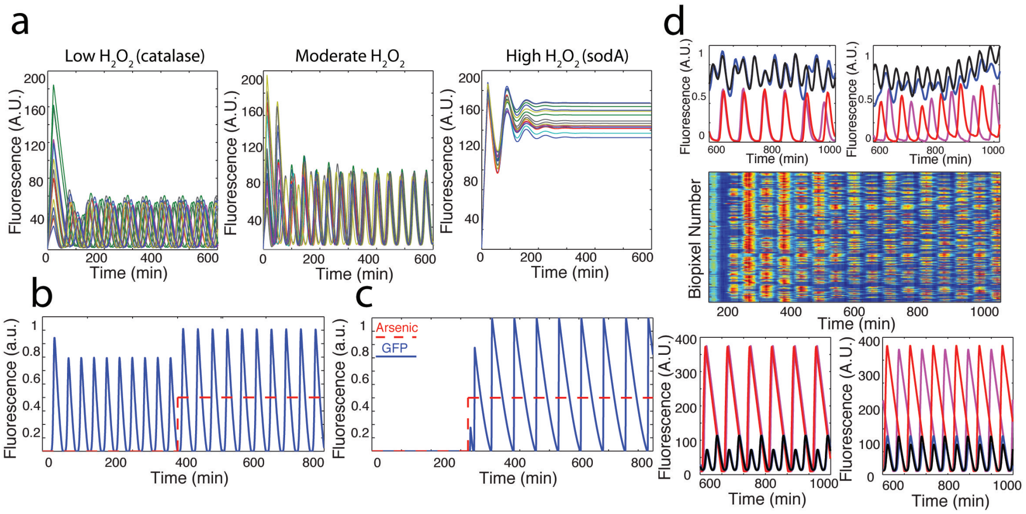


Figure 2.3: Computational modeling of radical synchronization and biosensing. (a) Time series of a population of biopixels producing varying amounts of H_2O_2 vapor. Synchronization occurs only for moderate levels while high levels lock ON and low levels oscillate asynchronously. **(b)** A typical time series for our period modulation sensor undergoing a step increase of arsenite. Oscillations increase in both amplitude and period. **(c)** A typical time series output for the thresholding sensor. Oscillations arise following the addition of arsenite. **(d)** Experimental and computational output depicting complex dynamic behaviors between neighboring traps. (Top 2 panels) 1:2 resonance and anti phase synchronization observed when trap size (left, black/blue = 95 μm depth and red/magenta = 85 μm depth) and separation distance (right, same colors) are modified experimentally, (Middle) Scaled-up array experimental data for increased trap separation experiments demonstrating anti phase synchronization, (Bottom) Computational model trajectories depicting 1:2 resonance and anti phase synchronization when trap size (same colors as experimental data) and separation distance are changed.

While our sensor array is capable of performing a variety of complex functions in the laboratory, adapting this technology to a real-world device will require the elimination of the expensive and bulky microscopy equipment. However, measuring genetic oscillations in the absence of any magnification or powerful illumination will require even further increased signal. Using this mechanism of global synchronization, we were able to scale up to a 24 mm x 12 mm array that houses over 12,000 communicating biopixels (Fig. 2.4a). Synchronization is maintained across the entire array, a distance over 5,000 times the length of an individual cell, using an inexpensive light-emitting diode, LED (Fig. 2.4b,c and Supplementary Movie 5). The signal strength generated by the large number of cells in the array (about 50 million) will allow us to adapt the device to function as a handheld sensor. In our conceptual design (Fig. 2.4d), the sensor will continuously read the oscillatory frequency using off-the-shelf electronic components costing less than \$50.

There have been many examples of bacteria-based biosensors (van der Meer and Belkin, 2010; Daunert et al., 2000; Leveau and Lindow, 2002), usually involving an optical reporter driven by a single promoter. Since optical intensity readings are sensitive to imaging conditions like beam power and exposure time, measurements must typically be normalized and calibrated. Measuring period of oscillation allows us to avoid these issues since peak-to-peak time does not depend on individual peak intensity. Secondly, oscillations produced at the colony level effectively decouple the signal from the growth state of individual cells, which can also affect fluorescence intensity. By using a dynamic readout that depends on communication between biopixels, we scan and tune potential output signals by changing device parameters rather than redesigning the underlying circuit. For example, we might design a new sensing scheme in which oscillations synchronize with the addition of some toxin and shift to anti phase or resonant synchronization when critical toxin levels are present.

Scaling Up Synthetic Biology

By nesting two modes of communication we are able to expand the scale over which individual cells are coordinated and increase the complexity of their interaction. Indeed, there are many familiar examples of hierarchical systems. Airline routes are often designed such that small airports are connected locally to larger hubs that are connected internationally. It would neither be feasible nor desirable to connect every airport together. Similarly, individual cells communicate locally by one method, generating impulses large enough to enable colonies to communicate globally by another.

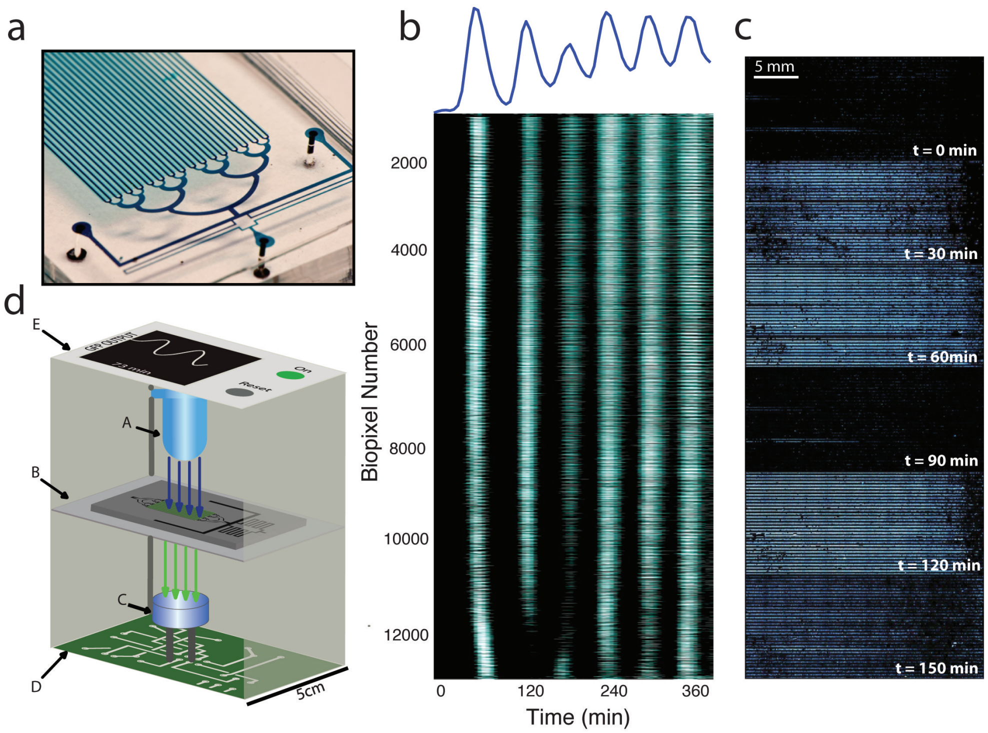


Figure 2.4: Radical synchronization on a macroscopic scale. (**a**) The scaled-up array is 24 mm x 12 mm and houses over 12,000 biopixels that contain approximately 50 million total cells when filled. (**b**) Global synchronization is maintained across the array. Heatmap of individual trajectories of all 12,224 oscillating biopixels. (**c**) Image series depicting global synchronization and oscillation for the macroscopic array. Each image is produced by stitching 72 fields of view imaged at 4X magnification. (**d**) Schematic diagram illustrating our design for a handheld device utilizing the sensing array. An LED (A) excites the array (B) and emitted light is collected by a photodetector (C), analyzed by an onboard processor (D), and displayed graphically (E).

Nesting communication mechanisms in this way may allow us to better scale up synthetic circuits of different types, such as switches and logic gates, paving the way for the next generation of synthetic biology pursuits.

Quantitative Modeling

Our model of the frequency modulated (FM) biosensor is based on a published model for the quorum-sensing synchronized oscillator (Danino et al., 2010). In addition to the reactions reflected in that model, we include the arsenite-induced production and degradation of LuxI and/or LuxR. From the biochemical reactions, we derived a set of delay-differential equations to be used as our model. These delayed reactions mimic the complex cascade of processes (transcription, translation, maturation, etc.) leading to formation of functional proteins. As expected, our model predicts oscillations that change frequency when changes in arsenite occur (Fig. 2.2c and 2.3b). The amplitude and period of the oscillations both depend on the concentrations of the toxin. We then modified the model to describe the LuxR-based detection system. Our model predicts a marked transition from rest to oscillations upon addition of arsenite, consistent with experimental observations (Fig. 2.3c).

The multi-scale nature of communication in our array allows us to treat colony and array-level dynamics separately, where arsenite affects the quorum-sensing machinery of a colony, producing changes to oscillatory period that propagate between biopixels in the array. To quantitatively describe the mechanisms driving synchronization at the array-level, we treat each colony as a single oscillator that acts according to degrade-and-fire kinetics (Mather et al., 2009). We also include the production of H_2O_2 and its interaction with neighboring colonies by two-dimensional diffusion. Using this model we identified three regimes that correlate well with experimental observations (Fig. 2.3a). When the effective production of H_2O_2 is low, as with catalase, we observe unsynchronized oscillations owing to constant, mild repression of the *lux* promoter via ArcA (Fig. 2.3a, left). In contrast, when H_2O_2 production is very high, neighboring colonies rapidly fire in succession and remain on due to the permanent activation of the *lux* promoter, consistent with the SOD experiment (Fig. 2.3a, right). Finally, at intermediate H_2O_2, we observe globally synchronized oscillations (Fig. 2.3a, middle). As colonies are moved further apart, synchronicity breaks due to slowed migration of H_2O_2 (Fig. 2.14).

Methods

Strains and Plasmids

The plasmids were constructed using a CPEC (Quan and Tian, 2009) cloning strategy in which the origin of replication, antibiotic resistance, and circuit genes were assembled in different combinations using PCR reactions. The *ndh* and *sodA* genes were amplified directly from the native *E. coli* genome by PCR. Various arsenic-responsive promoters were tested, including a recently reported synthetic version (Stocker et al., 2003), but the final design uses the native *E. coli* version. Promoter output was tuned by changing the RBS sequence and quantified using flow cytometry. All circuit components except *luxR* were tagged by PCR with a carboxy-terminal ssrA tag (AANDENYALAA) (Keiler et al., 1996) for fast degradation.

Microfluidics and Microscopy

Image acquisition was performed on a Nikon Eclipse TI epifluorescent inverted microscope outfitted with fluorescence filter cubes optimized for GFP imaging and a phase-contrast based autofocus algorithm. Images were acquired using an Andor Clara cooled CCD camera or Andor DU-897 EMCCD camera, both controlled by Nikon Elements software. Images were acquired every 2 minutes in phase contrast and fluorescence. 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 of arsenite-containing and pure medium were attached. Using this algorithm, arsenite concentration was dynamically varied to probe sensor output.

The microfluidic experiments were performed as previously described (Danino et al., 2010). Briefly, 50 μL 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.

Plasmid Construction

The oscillator plasmids were constructed by modifying the constructs used in a previous study (Danino et al., 2010). The antibiotic resistance genes of pTD103AiiA was switched to chloramphenicol. The reporter protein on pTD103LuxI/GFP was switched to a recently reported superfolding green fluorescent protein, sfGFP (Pedelacq et al., 2006). The ndh and sodA genes were amplified directly from the native *E. coli* genome by PCR. Promoter output was tuned by changing the RBS sequence and quantified using flow cytometry. We initially constructed the sensing plasmid with a published synthetic background-reduced version that contains additional ArsR operator sites (Stocker et al., 2003) but failed to produce enough LuxR. To increase LuxR output, we reverted to the native promoter sequence, switched the RBS to that of pZ plasmids (Lutz and Bujard, 1997), and increased the copy number by a factor of 5 by switching to a mutated SC101 origin of replication. All circuit components except LuxR were tagged by PCR with a carboxy-terminal ssrA tag (AANDENYALAA) (Keiler et al., 1996) for fast degradation. Modular pieces (resistance genes, promoters, origins, and ORFs) were assembled using a PCR-based cloning scheme named CPEC (Quan and Tian, 2009).

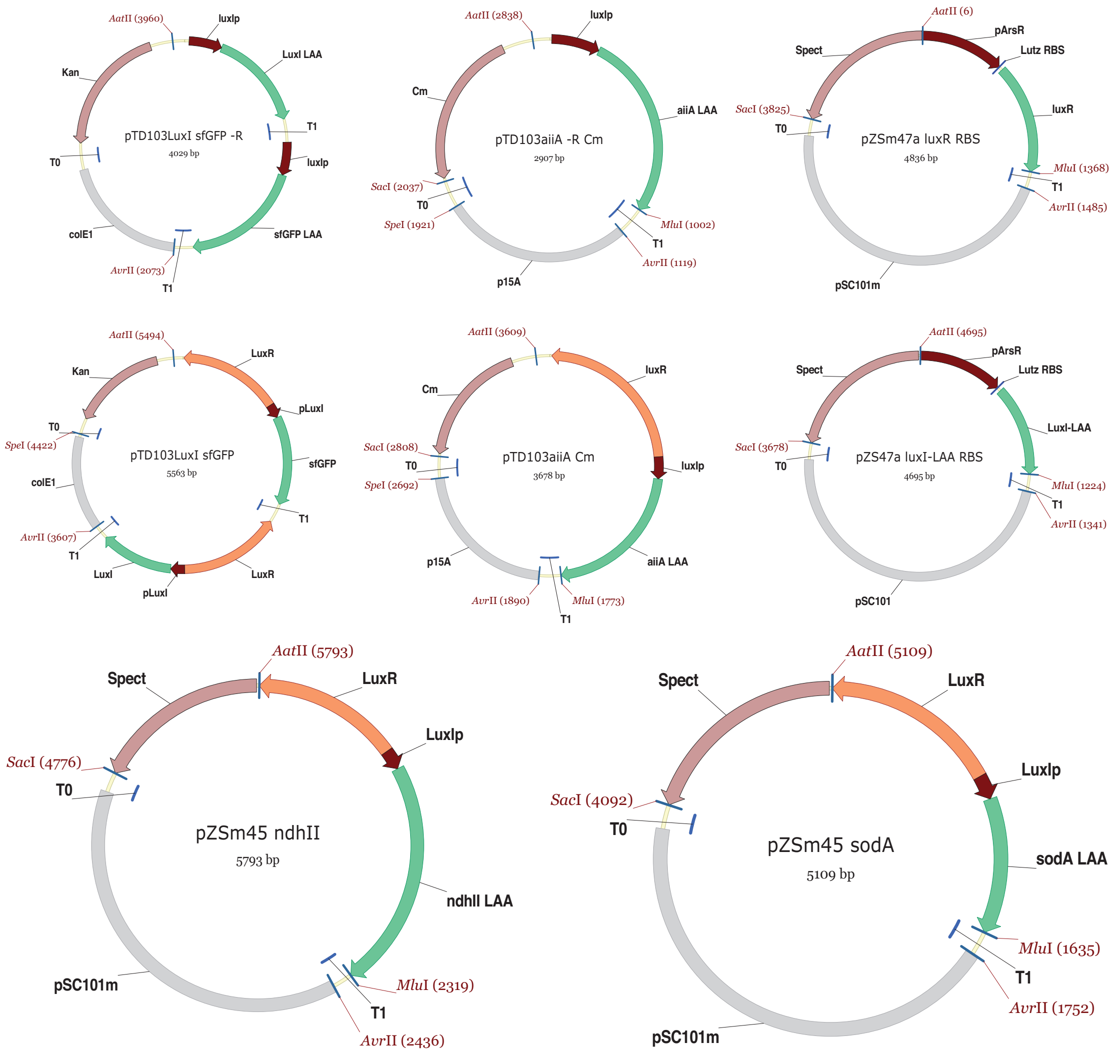


Figure 2.5: Plasmids for the coupled sensing array strain. Top row is the thresholding sensor: 2 oscillator plasmids with luxR genes removed and a plasmid containing pArs::luxR. Middle row is the period modulator: 2 oscillator plasmids and a plasmid containing pArs::luxI-laa. Bottom row contains 2 plasmids used to study H_2O_2production and synchronization: pLux::ndh and pLux::sodA. NDH-2 synchronization strain is the oscillator plasmids with pZSm45 ndhII.

Additional Experimental Results

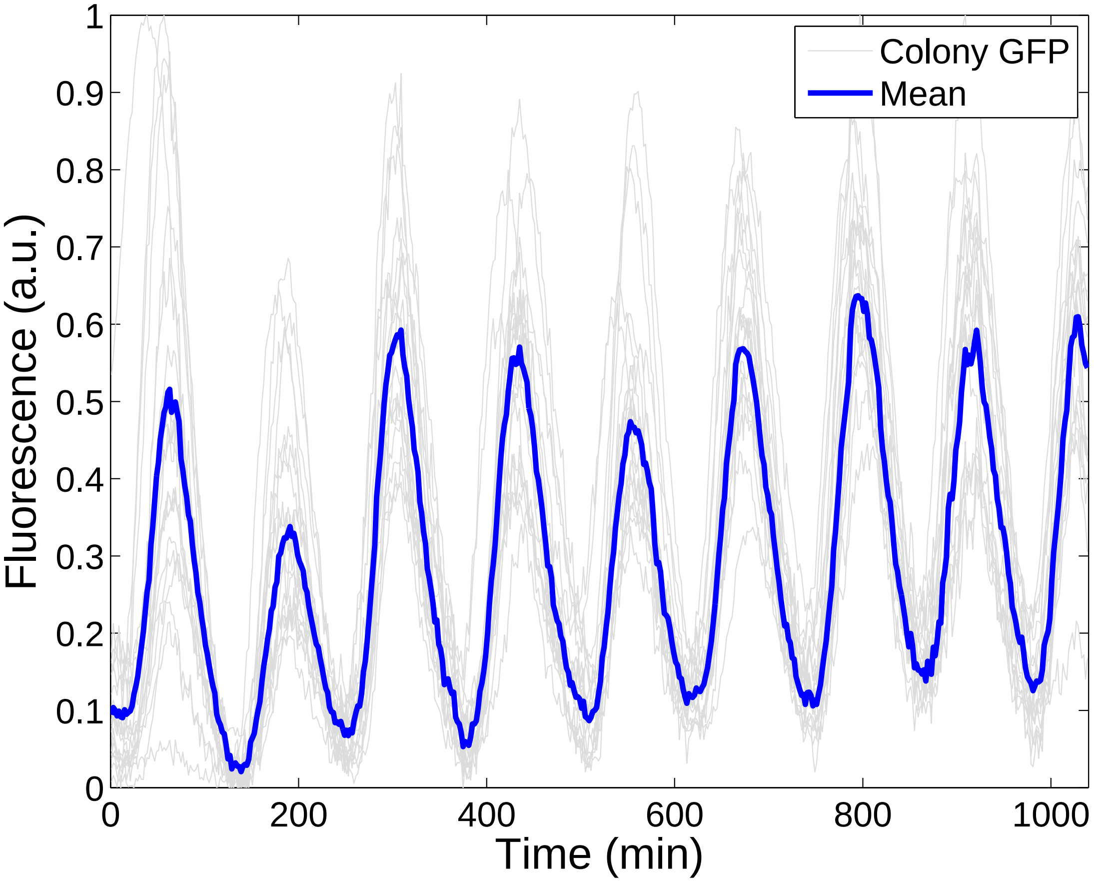


Figure 2.6: Biopixels with NDH-2 engineered synchronization observed at ultra-low fluorescence (4X, 20ms exposure, 3% power) using an EMCCD camera to ensure no fluorescence interaction. Synchronized oscillations are maintained across the array for the length of the experiment (14 hours).

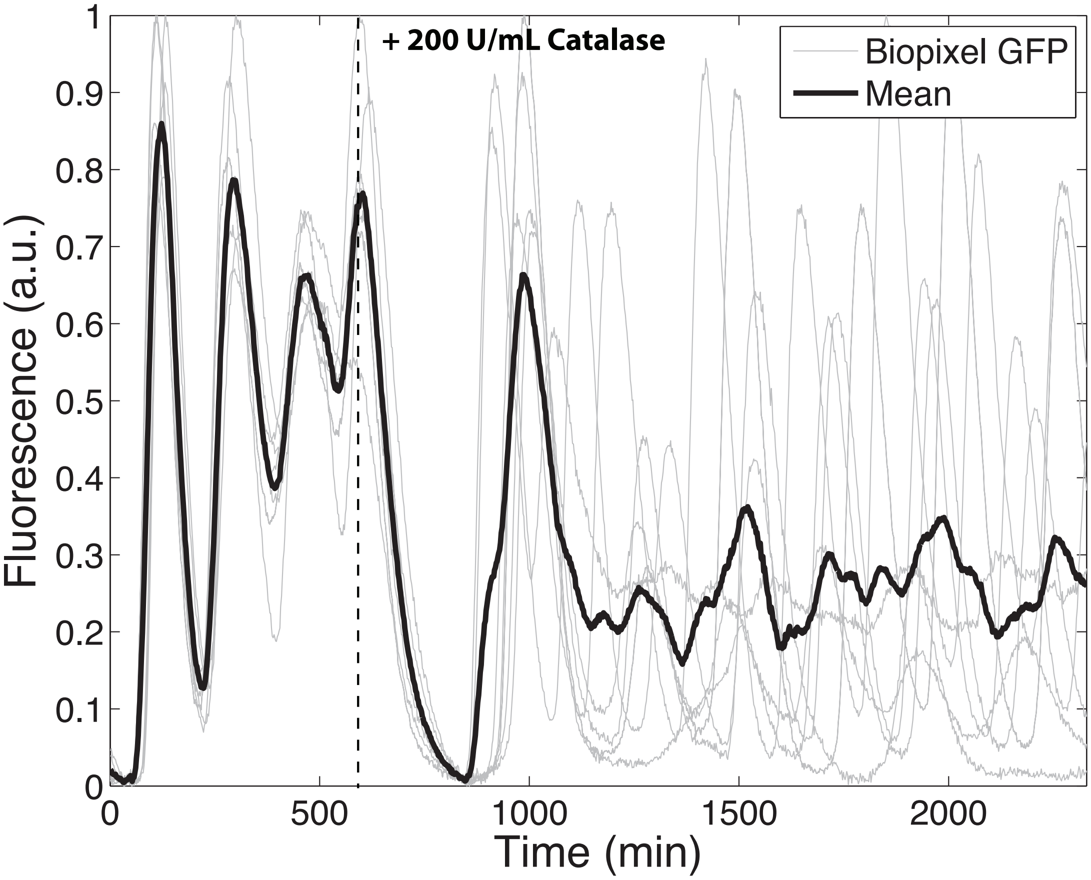


Figure 2.7: Catalase degrades external H_2O_2 and prevents communication between colonies. When a synchronized population of biopixels was exposed to a step increase of 200 U/ml catalase, synchronization was broken and biopixels continued to oscillate individually. Since catalase can't cross the cell membrane, this shows that synchronization between colonies depends on H_2O_2but oscillations with a colony do not.

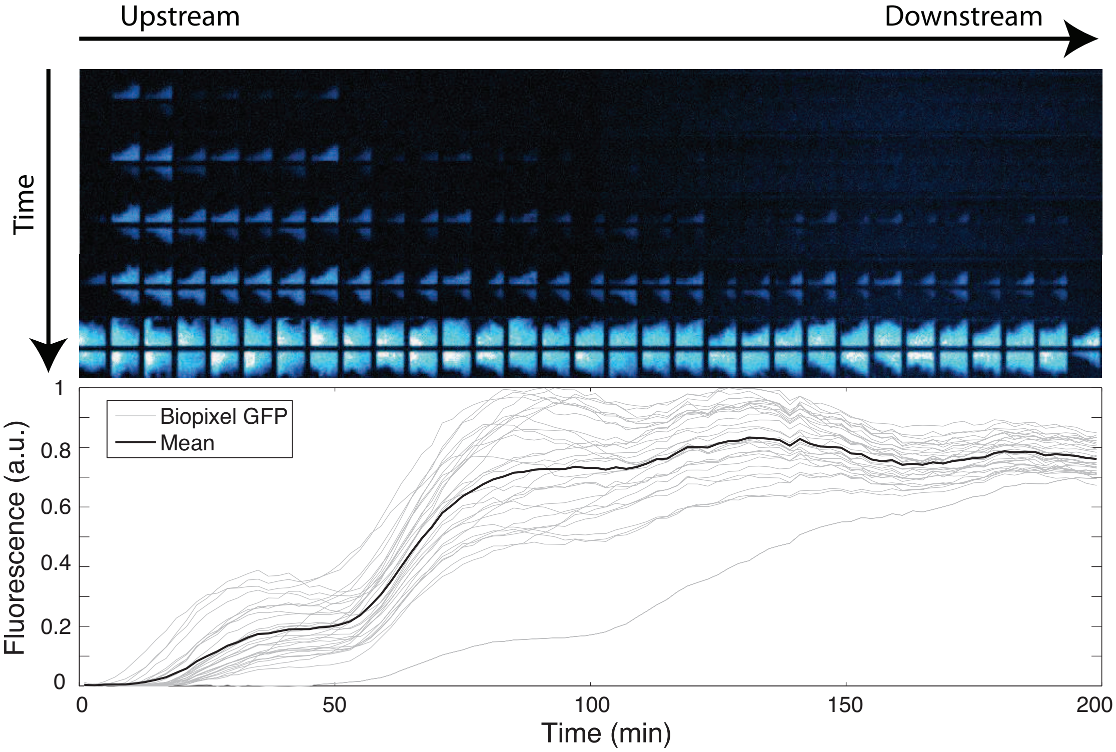


Figure 2.8: SodA produces H_2O_2internal to the cell, permanently switching the cellular redox state (oxidizing) thereby activating lux-controlled genes. Biopixels rapidly fire and lock on in a spatial wave, far earlier than is typical for colonies of this size. The propagation of ON biopixels suggests that colonies are capable of activating those nearby via migrating H_2O_2species.

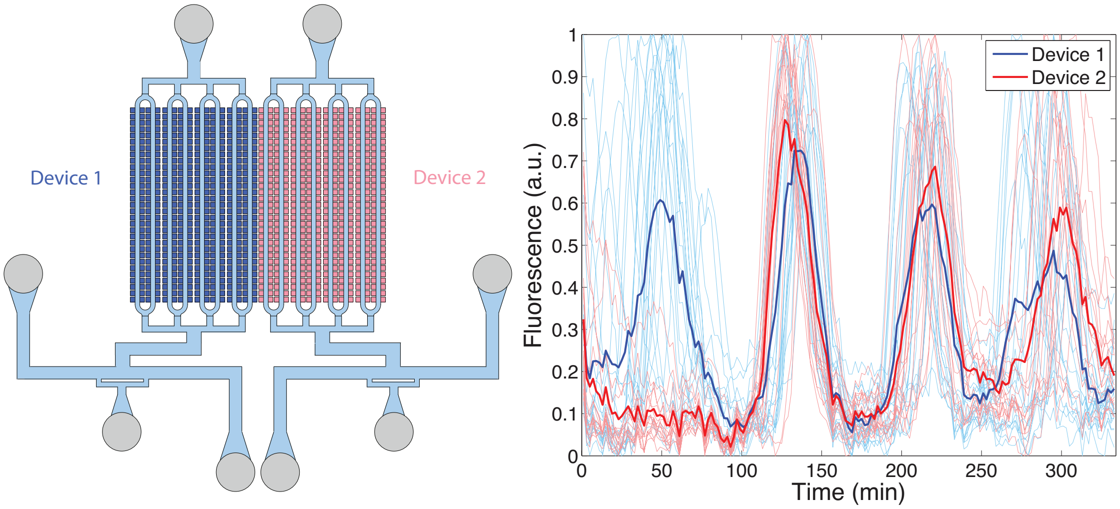


Figure 2.9: Synchronized oscillations occur across 2 fluidically isolated devices held in close proximity. In this experiment, the devices were started at different times yet become synchronized. Since these devices share no common fluid sources or sinks, this confirms that synchronization is mediated by vapor species.

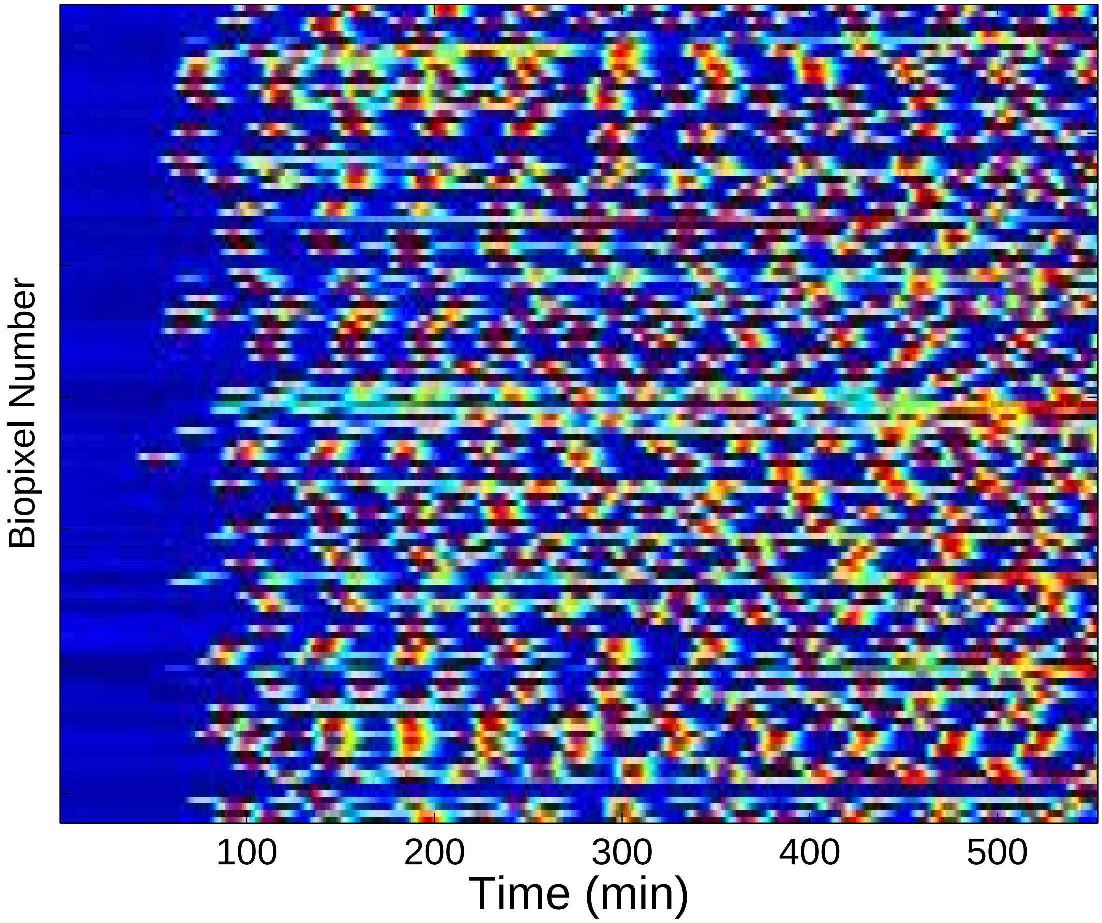


Figure 2.10: Heatmap of trajectories extracted from low fluorescence intensity control (Suppl. Movie 9) when NDH-2 plasmid is not present. Biopixels oscillate individually but fail to synchronize.

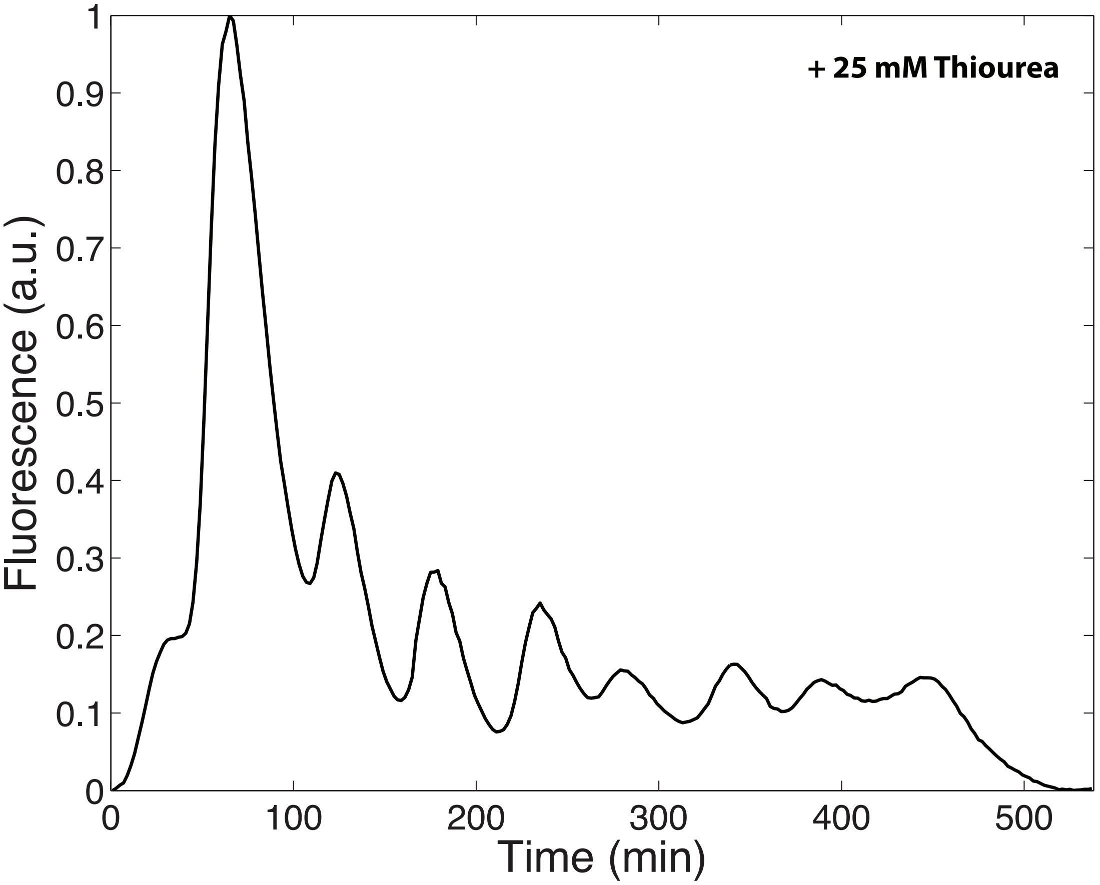


Figure 2.11: The introduction thiourea, a potent radical quencher, produces decaying synchronized oscillations across a population of biopixels. Because radical species are precursors for H_2O_2, eliminating them lowers the production of $H_2O_2$and therefore dampens the oscillations. Colonies are still able to synchronize because, while thiourea eliminates radicals within cells, it does not prevent $H_2O_2$from diffusing between cells.

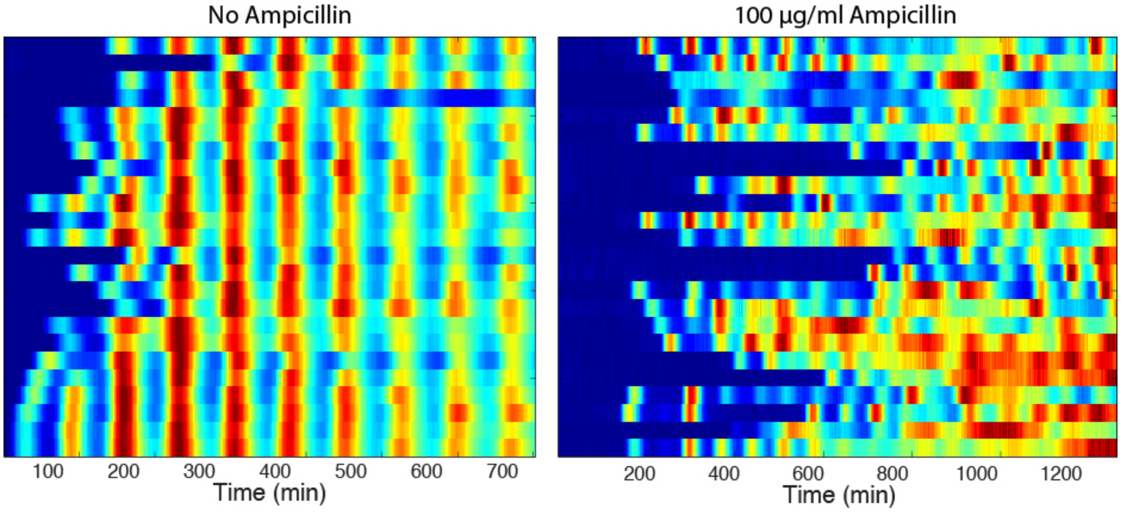


Figure 2.12: Synchronization is prevented when 100 μg/ml Ampicillin is used in the media. The constructs, strains, and experimental conditions are otherwise identical.

Data Analysis

Fluorescence data was obtained by importing fluorescent images into ImageJ and subtracting cell signal from background signal. Oscillatory period was taken to be the average of peak-to-peak and trough-to-trough distance, calculated using a MATLAB script. The data represented in Fig. 2.1d and 2.2b-d were collected by stitching 4 images taken at 4X magnification. The mean trajectory in Fig. 2.1d was found by averaging 373 individual biopixel trajectories, of which 20 are shown. Biopixel trajectories were extracted from image series using a MATLAB script, where a bright field image of the corresponding array was used to generate a mask. The data shown in Fig. 2c was measured over 4 separate experiments using 10-30 oscillatory periods per data point.

Sensor calibration curve (Fig. 2.2c, bottom) was generated using a series of 2-population ttests comparing the experimental datasets to randomly generated new sample sets. The mean of generated sets was decremented until the ttest failed with $\alpha = 95\%$, indicating the lowest period that could be associated with that arsenite concentration. We repeated this process for each arsenite level and fit the points with a quadratic since we expected it to take the inverse shape of the period vs. arsenite measurements.

Microscopy and Microfluidics

We used a microscopy system similar to our recent studies (Danino et al., 2010), with the addition of a high-sensitivity Andor DU-897 EMCCD camera. Fluorescent images were taken at 4X every 30 seconds using the EMCCD camera (20ms exposure, 97% attentuation) or 2 minutes (2s exposure, 90% attenuation) using a standard CCD camera to prevent photobleaching or phototoxicity.

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 (Fig. 2.13). 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% Tween which prevented cells from adhering to the main channels and waste ports.

To measure fluid flow rate before each experiment, we measured the streak length of fluorescent beads (1.0 μm) upon 100 ms exposure to fluorescent light. We averaged at least 1,000 data points for each.

We constructed several microfluidic devices over the course of the study. The trap dimensions were always 100 μm x 85 μm x 1.65 μm high, which we previously found to be optimal for oscillator function, except when size was varied to study dynamic interactions. Spacing between traps was 25 μm, except in devices designed to study the effects of increasing separation distance between traps. For sensor array devices, we constructed 500 and 12,000 trap arrays as well as a tandem device which holds two 150 trap arrays in close proximity (25 μm) without sharing fluid sources or sinks.

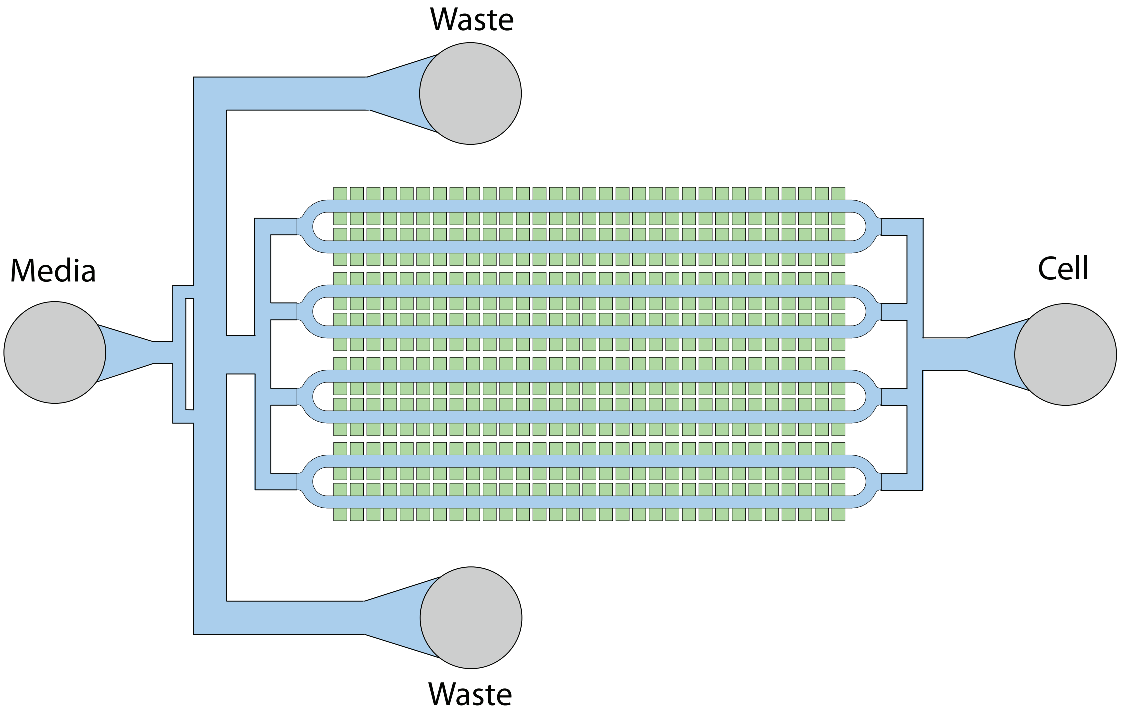


Figure 2.13: Primary microfluidic device used for this study. Media containing variable arsenite concentration is fed through the cell port, flowing past the biopixel array into the cell and waste ports. During loading, pressure is increased at the cell port and decreased at the waste ports to reverse the flow, allowing cells to pass by the trapping regions. Other microfluidic devices used have the same layout with trap number, separation, and size varied.

Modeling for Biosensor Array

To model the dynamics of the quorum-sensing oscillator, we used our previously described model for intracellular concentrations of LuxI (I), AiiA (A), internal AHL (H_i), and external AHL (H_e) (Danino et al., 2010),

$$\begin{aligned} \frac{\partial A}{\partial t} &= C_A[1-(d/d_0)^4]\,G(\alpha,\tau) - \frac{\gamma_A A}{1+f(A+I)} && (2.1)\\ \frac{\partial I}{\partial t} &= C_I[1-(d/d_0)^4]\,G(\alpha,\tau) - \frac{\gamma_I I}{1+f(A+I)} && (2.2)\\ \frac{\partial H_i}{\partial t} &= \frac{bI}{1+kI} - \frac{\gamma_H A H_i}{1+gA} + D(H_e - H_i) && (2.3)\\ \frac{\partial H_e}{\partial t} &= -\frac{d}{1-d}D(H_e-H_i) - \mu H_e + D_1\frac{\partial^2 H_e}{\partial x^2} && (2.4) \end{aligned}$$

In the original model, the concentration of the constitutively produced LuxR protein R was assumed constant. In the ON/OFF threshold arsenic biosensor circuit, LuxR production is induced by arsenic, which we model by the equation

$$\dot{R} = \frac{\alpha_c A}{(A_0+A)} - \gamma_R R \tag{2.5}$$

in which the LuxR expression from the arsenic promoter follows a standard saturating function of the arsenic concentration A. Accordingly, we modified the Hill function for Lux promoter to include the explicit dependence on R:

$$G(\alpha,\tau) = \frac{\delta + \alpha(R_\tau H_\tau)^2}{1+k_1(R_\tau H_\tau)^2} \tag{2.6}$$

For modeling the period-modulating sensor, we modified the equation for LuxI (4.5) to include additional production from the arsenic promoter,

$$\dot{I} = C_I[1-(d/d_0)^4]G(\alpha,\tau) + \frac{\alpha_c A}{(A_0+A)} - \frac{\gamma_I I}{(1+f(A+I)} \tag{2.7}$$

The following additional parameters were used for the biosensor simulations: $\alpha_c = 50$, $A_0 = 2$, $\gamma_R = .1$.

Arsenic levels were swept across the dynamic range of the arsenic promoter to produce the curve in Fig. 2.2c. The period for each arsenic level was calculated from the peak-to-peak average of 15 oscillatory periods. Adjusting the flow rate to match the baseline period ($A = 0$) with that of our large array, the resulting curve fits the experimental data closely.

To model the spatial synchronization of oscillating colonies across a microfluidic array, we generalized a simplified "degrade-and-fire" model (Mather et al., 2009). The delay-differential equation

$$\dot{X}_{i,j} = \frac{\alpha(1 + \nu P_{i,j,\tau_2})}{(1 + \frac{X_{i,j,\tau_1}}{C_0})^2} - \frac{\gamma X_{i,j}}{k + X_{i,j}} \tag{2.8}$$

describes oscillations of individual biopixel $\{i, j\}$ as a combined effect of production and delayed autorepression (first term in the r.h.s.) of the colony-averaged LuxI concentration $X_{i,j}$ and its enzymatic degradation by ClpXP (second term). Unlike (Mather et al., 2009), the first (production) term in Eq. 5.4 describes both delayed auto-repression of LuxI and its delayed activation by H_2O_2 proportional to its local concentration $P_{i,j}$. Subscripts τ_1 and τ_2 indicate the delayed concentrations, $X_{i,j,\tau_1}(t)$ = $X_{i,j}(t - \tau_1)$ and $P_{i,j,\tau_2}(t)$ = $P_{i,j}(t - \tau_2)$. The dynamics of $P_{i,j}$ is described by the equation

$$\dot{P}_{i,j} = \mu + \alpha_p X_{i,j} - \gamma_p P_{i,j} + \hat{S}\{P_{i,j}\} \tag{2.9}$$

where the first three terms describe the basal and induced production and degradation of H_2O_2. The last term models the spatial coupling of neighboring biopixels via the H_2O_2 exchange. For a square $N \times N$ array of traps, we used the following discrete diffusion form of the spatial operator,

$$\hat{S}\{P_{i,j}\} = D\Delta^{-2}[P_{i-1,j} + P_{i+1,j} + P_{i,j-1} + P_{i,j+1} - 4P_{i,j}] \tag{2.10}$$

Each colony is affected by the H_2O_2 produced in four neighboring colonies, two in each dimension of the array, separated by the equal distance Δ. We used the boundary condition $P_{i,j} = 0$ for the edges of the array $i, j = 0, N + 1$. This represents the infinite external sink of H_2O_2 diffusing out of the microfluidic chip. The diffusion operator above can be generalized if the row spacing differs from the column spacing, or for other spatial arrangements of colonies within the biosensor.

Since we were specifically interested in the dynamics underlying the spatial coupling of independently oscillating colonies, this simplified representation of the quorum sensing oscillator was appropriate. The equation for LuxI (X) takes the observed degrade-and-fire oscillatory waveform by including enzymatic degradation proportional to γ, due to its ClpXP degradation tag, and delayed negative feedback, caused by the repressor (AiiA) tracking bursts in LuxI (Danino et al., 2010). The production of LuxI depends inversely on its own past concentration, $X_{\tau_1}(t) = X(t - \tau_1)$. Activated coupling is mediated by production of H_2O_2 within a trap. The level of H_2O_2 response to oscillatory bursts is controlled by α_p, which is tuned according to whether H_2O_2 is being produced by ndhII/sodA enzymes, and tracking the the Lux promoter in the genetic circuit, or by fluorescence, and tracking the level of sfGFP. H_2O_2 then feeds back into the oscillator by increasing the production of the LuxI, which follows from the mechanism of H_2O_2 relieving the repression by sodA on the Lux promoter. To account for the redox signaling involved in this response, we use the delayed level of H_2O_2 , $P_{\tau_2}(t) = X(t - \tau_2)$, to upregulate the Lux promoter.

We introduced variability among different traps by randomizing oscillator parameters for individual traps in each simulation. Specifically, LuxI (X) activation and degradation parameters ($p = \{\alpha, \gamma\}$) of each of the oscillators in the array were varied around their nominal values (p_0) as $p = p_0 + \delta$ where δ is a random number uniformly distributed between -0.25 and 0.25. We used the following dimensionless parameters for most of our simulations: $\alpha_0 = 8.25$, $\gamma_0 = 5.75$, $\nu = 1$, $\tau_1 = 10$, $\tau_2 = 20$, $C_0 = 6$, $k = 10$, $\mu = 20$, $\alpha_p = 1$, $\gamma_p = 10$, $D = 7$, $\Delta = 1$.

For the characterization of various regimes of array synchronization, 16 colonies were modeled in the 4×4 array. Scaling up the simulation with larger numbers of colonies produced equivalent results. Overproduction of H_2O_2 by expressing sodA was captured by increasing α_p 20-fold. This is consistent with expression from a pSC101m plasmid with a copy number of 20-30. Depletion of external H_2O_2 by catalase was modeled by increasing H_2O_2 degradation (γ_p) and decreasing H_2O_2 diffusion, D. In (Fig. 2.14) we show the variance of the concentrations $X_{i,j}$ within the array averaged over time and parameter variations. This plot demonstrates that the synchronicity among the biopixels decreases with increase of spacing among them, and for $\Delta > 5$ is completely lost.

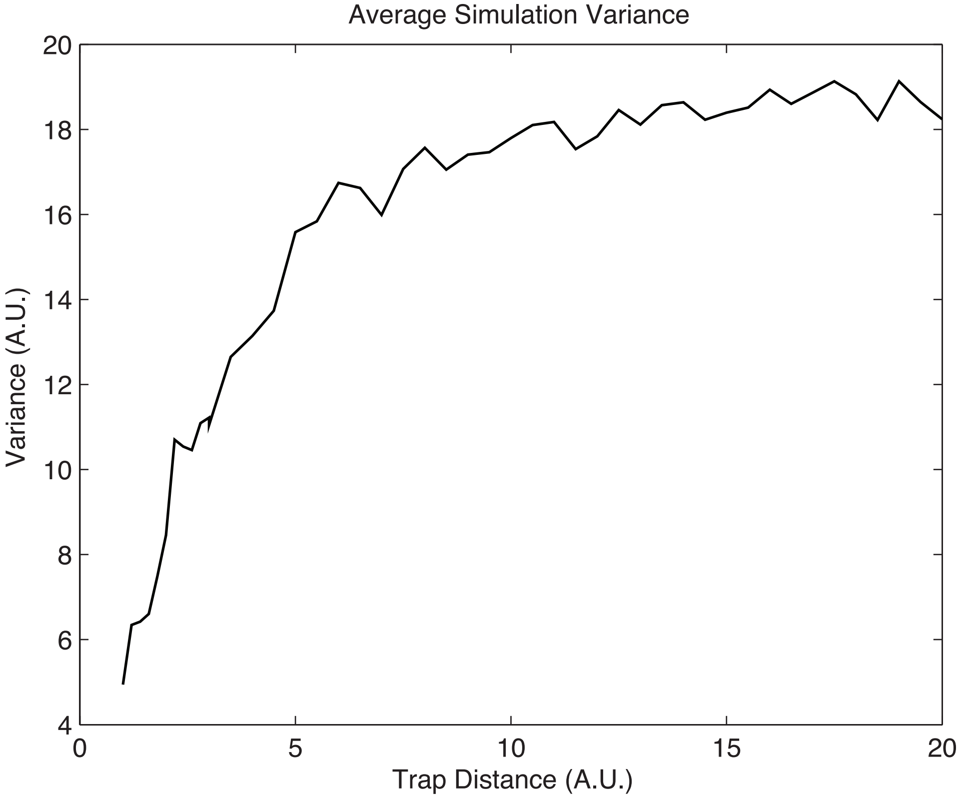


Figure 2.14: Computational results depicting biopixel synchronicity as a function of trap separation distance. As biopixels are moved farther apart, the entropy increases due to decreased effective migration of H_2O_2 between colonies.

Increasing the trap spacing Δ 2-fold while simultaneously decreasing k 4-fold allowed us to reproduce the more complex waveforms observed experimentally in our arrays. Note that changing *k* models the change of the trap depth. As the size of the trap decreases, the flow of media is able to more rapidly sweep away AHL and increase the effective degradation for the colony. Simulating smaller and more sparse trap sizes recovered antiphase behavior for neighboring biopixels (Fig. 2.15). We also simulated the arrays with traps of two different sizes in different rows and recovered the experimental 2:1 biopixel resonance or 2:1 + antiphase behavior depending on the trap spacing (Fig. 2.3d, bottom).

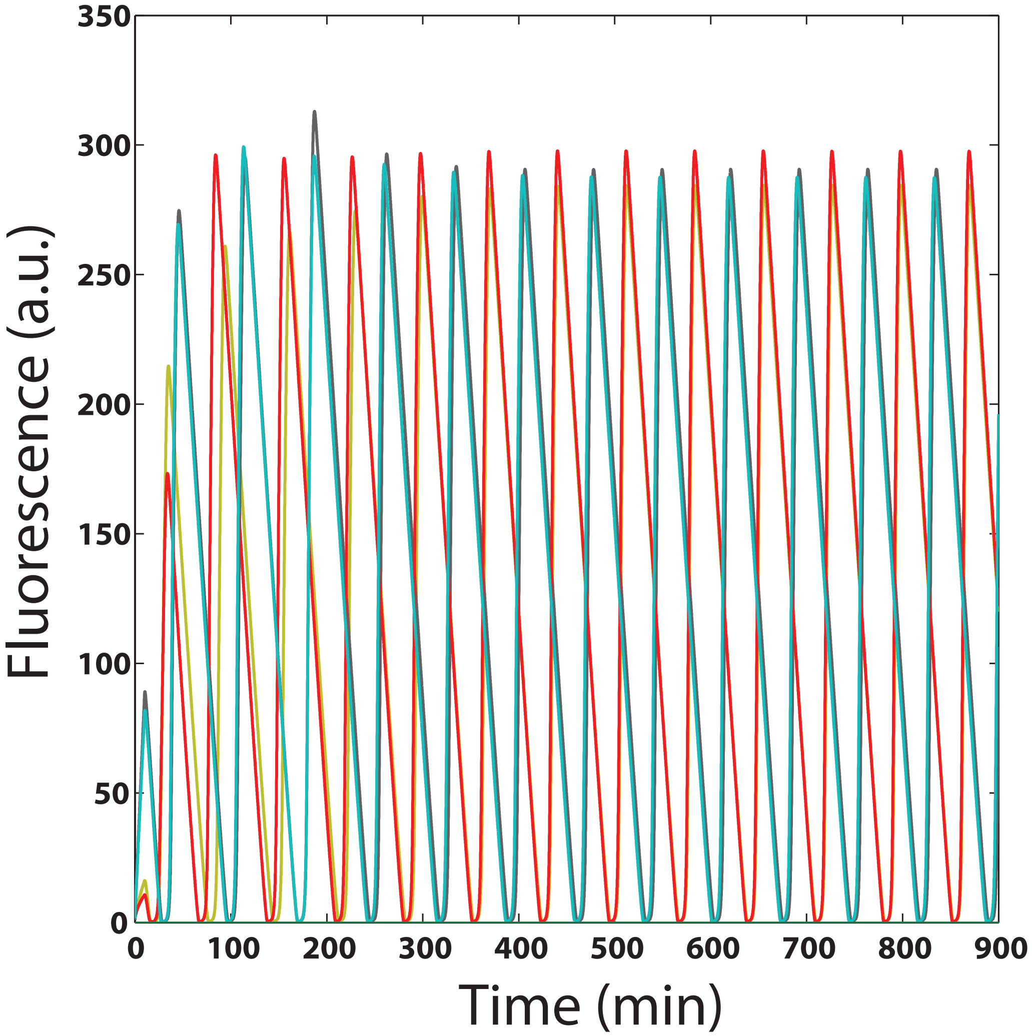


Figure 2.15: Antiphase behavior of 4 neighboring biopixels having equal trap sizes and spacing $\Delta = 3$.

The model was also able to capture the alternating large and small amplitude oscillations observed in the ON/OFF biosensor (Fig. 2.16). This behavior was seen when C_0 was increased 2-fold, capturing the decreased level of LuxR in ON/OFF experiments where it was the limiting factor for oscillations.

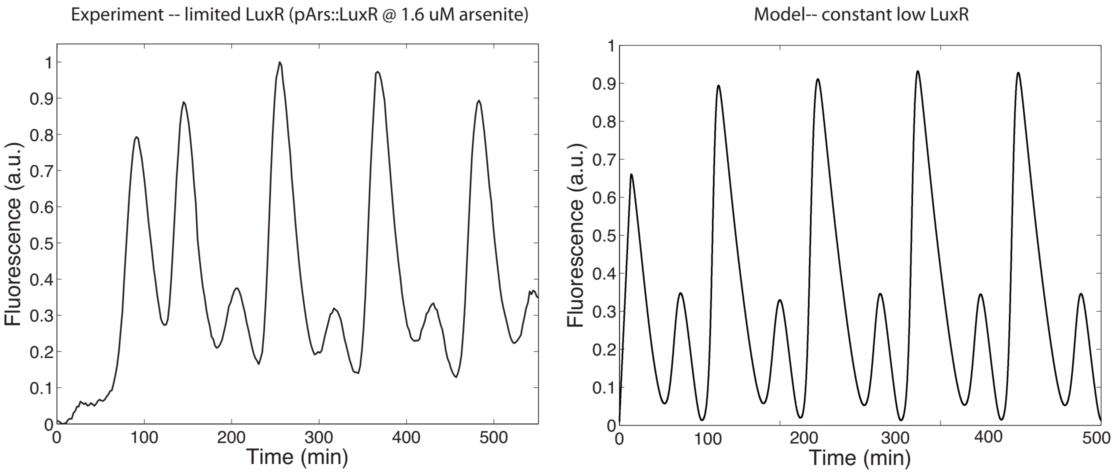


Figure 2.16: Oscillations of alternating large and small amplitude when LuxR is limited in experiments and simulations. The alternating oscillations vanish when LuxR is restored to its normal level in the model. Experimentally, we were unable to build a system in which LuxR is tunable between big/small and normal amplitude regimes. This is probably due to the small dynamic range of arsenite promoter-driven output of LuxR compared to the level produced by 3 constitutively expressed copies in the original circuit.

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 a 200 trap sensor array displaying NDH-2 engineered synchronization. An EMCCD camera was used to keep exposure times extremely low (4X magnification, 20ms, 95% attenuation) to ensure no fluorescence interaction, hence the appearance of lower signal.

Supplementary Movie 2. Timelapse fluorescence microscopy of the 500 trap biosensor array showing the onset of synchronization from disparate initial conditions using period modulator circuit. Flashes indicate changes in arsenite concentration which result in changes in the oscillatory period.

Supplementary Movie 3. Timelapse fluorescence microscopy of a sensor array containing thresholding circuit. Red color indicates addition of 0.25 μM arsenite that initiates oscillations in blue.

Supplementary Movie 4. Timelapse fluorescence microscopy of a modified 500 trap sensor array in which traps are farther apart. This increased separation results in anti phase oscillations, where a biopixel and its nearest neighbors alternate bursts.

Supplementary Movie 5. Timelapse fluorescence microscopy of the 12,000 trap scaled up array showing oscillation and synchronization maintained over a maximum distance of 27 mm.

Supplementary Movie 6. Real time microscopy depicting the loading of our microfluidic device. Cells flow in from the cell port and fill the trapping regions.

Supplementary Movie 7. Timelapse fluorescence microscopy of a modified 500 trap sensor array in which traps of 2 sizes are present. This results in 2:1 resonant oscillations where larger traps oscillate at twice the frequency of smaller traps.

Supplementary Movie 8. Timelapse fluorescence microscopy of a 500 trap sensor array showing unsynchronized oscillations when NDH-2 is not present and high-intensity fluorescence bursts are not used.

Supplementary Movies can be found at http://www.nature.com/nature/journal/v481/n7379/full/nature10722.html#/supplementary-information

Acknowledgements

This work was supported by the National Institutes of Health and General Medicine (R01GM69811), the San Diego Center for Systems Biology (P50GM085764), the DoD NDSEG (AP) and NSF Graduate Research (PS) Fellowship Programs. LT was supported, in part, by the ONR (MURI N00014-07-0741). We would like to thank Jim Imlay, Sameh Ali, and Jim Collins for helpful discussions and Jen Hickman, Brooks Taylor, and Katie Lomax for their help with illustrations.

Chapter 2 contains material originally published as Prindle, A.*, Samayoa, P.*, Razinkov, I., Danino, T., Tsimring, L. and Hasty, J. ***Nature*** (2012): A sensing array of radically coupled genetic ’biopixels’. (*equal contribution). Copyright permission to republish here was granted by Nature publishing group.

Chapter 3

Applications of Quorum Sensing Biology

A fast and robust macroscopic toggle switch using synergistic coupling

Introduction

A synthetic gene toggle switch circuit was first implemented in *E. coli* over a decade ago when synthetic biology was still in its infancy (Gardner et al., 2000). Several other designs of bistable and hysteretic circuits were proposed and implemented more recently (Atkinson et al., 2003; Kramer and Fussenegger, 2005; Ham et al., 2008). Most of the existing designs suffer from the intrinsic stochasticity of the underlying biochemical processes and cell-cell variability that leads to spontaneous transitions between states and a lack of robust and yet sensitive switching behavior. Here we improve on the previous switch designs by building on the synergistic coupling mechanisms described in Chapter 1 of quorum-sensing and redox signaling (Mondragón-Palomino et al., 2011; Prindle et al., 2011). Specifically, by synchronizing the switching behavior of millions of cells, we aimed to increase the sharpness of state transitions while reducing erroneous stochastic switching. This required balancing the strength of the two states experimentally by modifying plasmid copy number, ribosome binding site strength, or promoter strength. To guide this process, we constructed and fit a computational model based on the underlying genetic interactions.

Toggle Design

We designed a symmetric macroscopic toggle switch; each state stabilizes itself through two positive feedbacks and destabilizes the opposing state through a single negative feedback (Fig. 3.1). In state 1, a hybrid RHL/Lac promoter is induced by the quorum sensing molecule RHL while a hybrid AHL/Tet promoter is repressed by the TetR dimer. In state 2, AHL/Tet promoter is induced by AHL while RHL/Lac promoter is repressed by the LacI tetramer. The second positive feedback through the long-range redox coupling is achieved by fluorescence induced production of H_2O_2 that in turn activates the promoters via the native aerobic response system arcAB similar to the biosensor circuit described above. The negative feedback elements TetR and LacI can be modulated through the chemical inducers Dox and IPTG, respectively, providing a chemical way of switching between the two states.

Computational Modeling

Our preliminary computational model is based on a system of six stochastic differential equations describing a four-gene system, LacI (L), TetR (T), LuxI (U), RhlI (H) and two small molecules, AHL (A), and RHL (R). For simplicity we neglected the dynamics of two reporter proteins and redox signaling here. The model equations can be written in the following abbreviated form,

$$\frac{\partial X}{\partial t} = P(Y,Z) - \frac{\gamma_X X}{1+fT} + \left[P(Y,Z) + \frac{\gamma_X X}{1+fT}\right]^{1/2} \xi \tag{3.1}$$

$$\frac{\partial V}{\partial t} = \frac{bW}{1+kW} - \gamma A + D(\langle V \rangle - V) \tag{3.2}$$

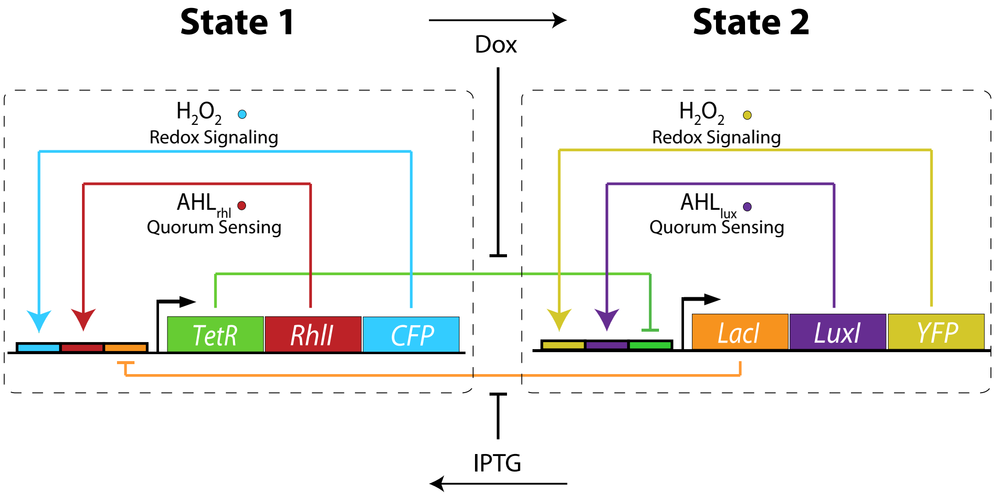


Figure 3.1: Circuit diagram for the macroscopic toggle switch system The circuit can be switched between the two states with the introduction of the chemical inducers IPTG and Dox. The activity of each half is independently visualized with the fluorescent proteins CFP and YFP.

where $P(Y,Z) = (\delta + \alpha Z^2)[(1+k_1Y^2)(1+k_2Z^2)]^{-1}$, the triplet (X,Y,Z) stands for (L,T,A), (U,T,A), (T,L,R), (H,L,R), the pair (V,W) stands for (A,U) and (R,H), and T stands for the sum of concentrations of all proteins $L+T+U+H$. The last term in Eq. 3.1 denotes multiplicative noise due to intrinsic fluctuations in the underlying biochemical reactions, and the last term in Eq. 3.2 describes global coupling through the external fields of corresponding quorum-sensing molecules. We find that without coupling across cells ($D = 0$), individual cells spontaneously switch from one metastable state to another (see Fig. 3.2a-b). However, global coupling between the cells via quorum sensing molecules ($D = 1$) locks the system in one of the two states, as set by inducers (see Fig. 3.2b).

Results

We constructed various plasmid systems to test the macroscopic toggle switch. Because neither of the hybrid promoters previously existed, our first iteration of the circuit had unbalanced feedback, where only one state, ”ON”, has stabilizing positive feed-

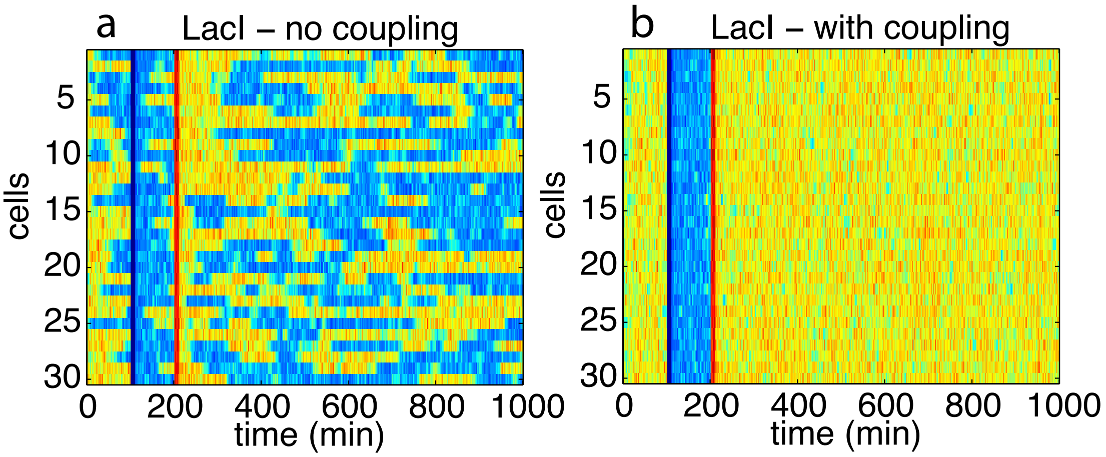


Figure 3.2: Space-time diagrams of uncoupled and coupled toggle systems Concentrations of a reporter protein (LacI) in uncoupled **(a)** and coupled **(b)** systems. External inducers are imposed (indicated by shaded vertical lines) to temporarily bias the toggle into a particular polarity. Coupling extends the persistence of polarity.

back (Fig. 3.3a). This version included a new hybrid AHL/Tet promoter that was built by adding a Tet operator site (Lutz and Bujard, 1997) following the transcription initiation site of the *lux* promoter (Bose et al., 2007). The AHL/Tet promoter exhibited the predicted expression in response to various combinations of chemical inducers AHL and Dox when tested in an *E. coli* strain expressing constitutive TetR. The TetR repression is dominant over AHL induction, which is necessary to ensure that fluctuations in global AHL cannot incorrectly trigger the ”ON” state during the high TetR ”OFF” state. Following relief of TetR repression by Dox, the promoter exhibits a large range of induction for increasing concentrations of AHL (Fig. 3.3b). We employed a two plasmid circuit to enable the relative strength of the two states to be easily tuned by plasmid copy number. The ON plasmid consists of the hybrid AHL/Tet promoter expressing LacI, LuxI, and the reporter, sfGFP, while the OFF plasmid is the $P_{\mathrm{LlacO\text{-}1}}$ promoter (Lutz and Bujard, 1997) expressing TetR. Using the same *E. coli* strain that constitutively expressed TetR, the ON plasmid was characterized to ensure that the positive AHL feedback leads to stabilization of the ”ON” state (Fig. 3.3c). Interestingly for high levels of Dox (300 ng/ul), the expression level decreases slightly from full induction, which is likely due to toxicity from the chemical inducer.

We co-transformed the two plasmids into our JS006 experimental strain, which expresses no genomic LacI or TetR (Stricker et al., 2008). Various combinations of replication origins were tested by flow cytometry for the ability of the strain to maintain each of the ”OFF” and ”ON” states following the introduction and subsequent removal of the corresponding chemical inducers (IPTG and Dox respectively). This ensures the strain is capable of stably toggling to two states, as opposed to transiently switching and remaining dependent on the chemical inducer. After testing all combinations of ColE1 (100-300 copies), p15A (50-100 copies), Sc101m (30-50 copies), and Sc101* (3-5 copies), we found one combination that exhibited coupled toggling. Placing the OFF plasmid on the p15A origin and the ON plasmid on Sc101m exhibited maintenance of each state by flow cytometry, and was confirmed by testing in the microfluidic array described previously (Fig. 3.3d). Individual colonies within the array behaved identically, with no traps remaining OFF or switching OFF after the removal of Dox. This was expected based on the coupled positive feedback design that includes AHL quorum sensing between cells and long-range fluorescence-mediated redox coupling between colonies. Interestingly, the transition to the ON state was bi-phasic, presumably consisting of the rapid quorum-switching followed by the slower ramping up of signal as TetR undergoes enzymatic degradation. Confirmation of this would require the addition of a reporter to the OFF plasmid to track the level of TetR during the switch from ”OFF” to ”ON”.

Redox Induced Toggling

Using the synergistically coupled toggle with the microfluidic array platform, we investigated the scale of redox mediated communication between colonies. We constructed a split media microfluidic array, where individual channels of side traps receive their own supply of media. This enabled us to trigger certain sets of traps directly by the chemical inducer Dox, and observe neighboring channels of traps for transitions to the

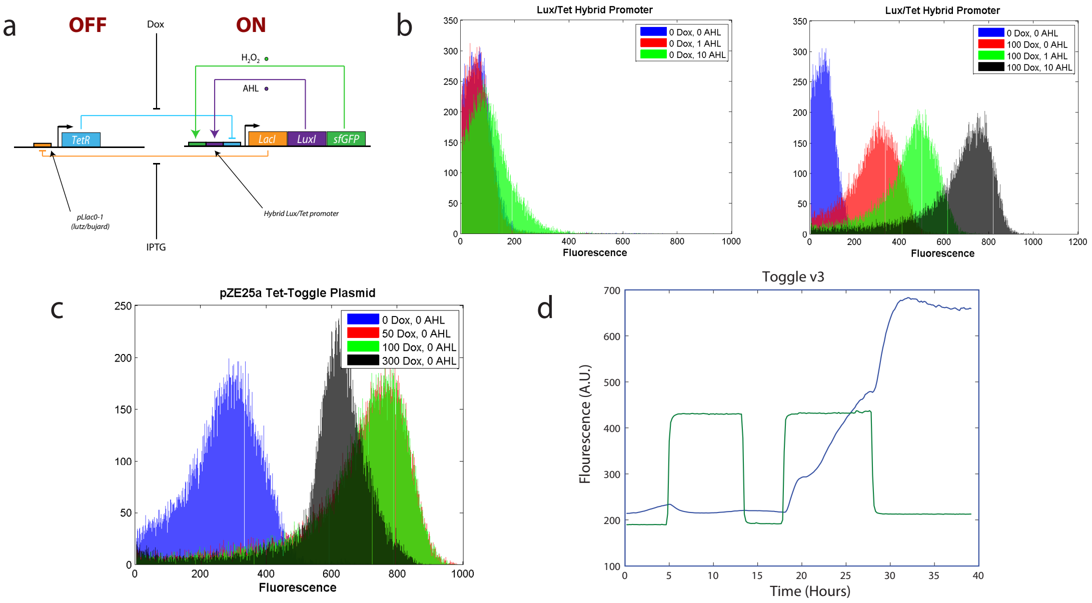


Figure 3.3: Unbalanced coupled toggle switch. **(a)** Network diagram. The high LacI ”ON” state is stabilized by AHL coupling through the hybrid AHL/Tet promoter. **(b)** Hybrid AHL/Tet promoter FACS characterization. The hybrid promoter exhibits AHL induced expression only when TetR repression is relieved by the addition of the chemical inducer Dox. **(c)** ON plasmid FACS characterization. The ON plasmid reaches the ”ON” state, without externally added AHL, in the presence of Dox. **(d)** Coupled toggle in a microfluidic array. The mean fluorescence of all traps (blue) stably maintains both the ”OFF” and ”ON” states following the removal of the chemical inducer (green, first pulse = 1mM IPTG; second pulse = 50 ng/ul Dox).

”ON” state (Fig. 3.4). This constrained the triggering event in channels not exposed to Dox to a H_2O_2 related response. By observing the timing of redox-mediated switches we were be able to deduce the range of H_2O_2 diffusion.

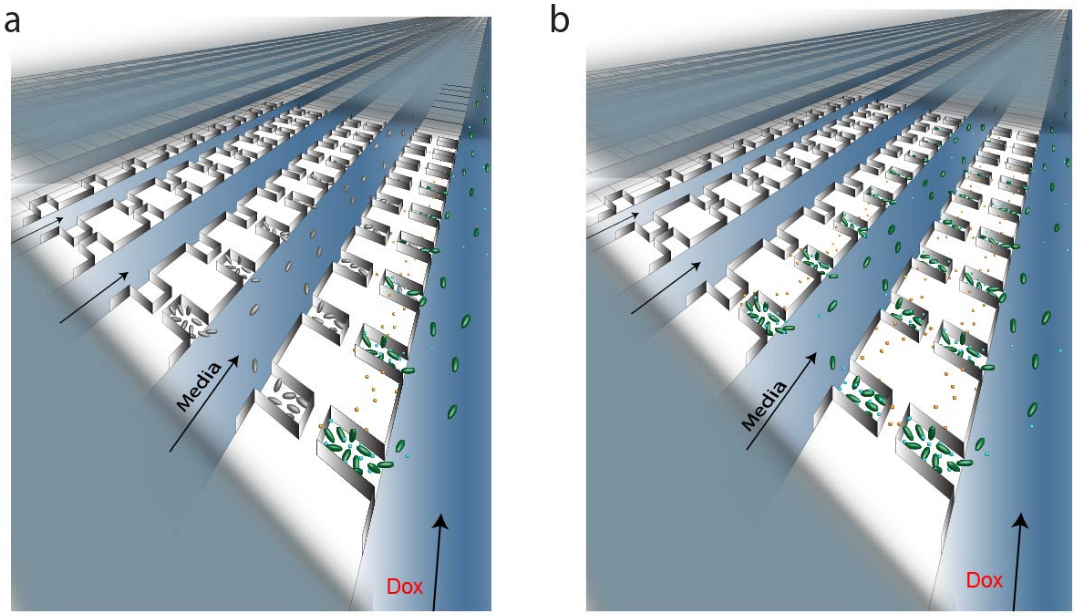


Figure 3.4: Redox-triggered toggle split media array. Traps along a channel containing the Dox inducer **(a)** are directly switched on, while traps along neighboring channels **(b)** are later triggered by H_2O_2 diffusion that upregulates the basal *lux* promoter activity.

The microfluidic device was designed to have full channels of side traps receiving a chemical inducer along the bottom, with adjacent channels on the upper portion of the chip exposed to a separate media supply that does not trigger side trapped colonies. Transitions of the upper channels can only be induced as a result of an H_2O_2 response to the triggered lower channel. The upper portion is sparsely filled with side traps to generate a directed triggering path, exposed to a more even left-right distribution of H_2O_2 from the chemically induced region (Fig. 3.5).

Triggering the lower channel with Dox initially only directly activates the chemically exposed colonies along the bottom of the chip (Fig. 3.6, top). The upper redox triggered lanes were then measured over the next 10 hours as they were exposed to the H_2O_2 release from the activated lower channel. This proved to be sufficient stimula-

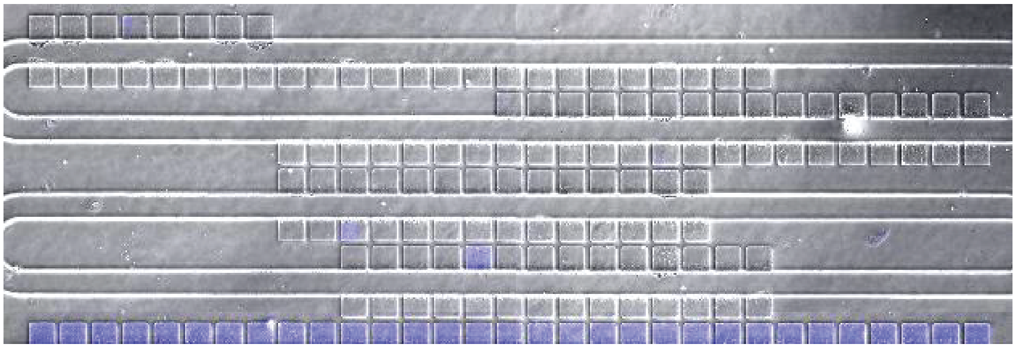


Figure 3.5: Split Media Chip Design. Microfluidic device with channels that either receive chemical inducer (blue channel, bottom) or are only triggered via H_2O_2 diffusion (clear channels).

tion to trigger colonies that weren't directly modulated by Dox, with the majority of the redox-modulated colonies toggling to the on state within 10 hours (Fig. 3.6, bottom).

The pattern of individual channel stimulation further supports the outward H_2O_2 diffusion from chemically triggered colonies, as channels proximal to the dox-induced lane appear to trigger earliest (Fig. 3.7). While this experiment describes the transmission of H_2O_2 between colonies and implies a diminishing signal strength with distance from the original H_2O_2 source, it is limited in its ability to discern the absolute range of H_2O_2 interaction between colonies of various distances. Future chip designs must implement single channels of varying distances from the chemically-induced set of colonies to determine the range of H_2O_2 diffusion across which an induced colony can trigger an un-induced one.

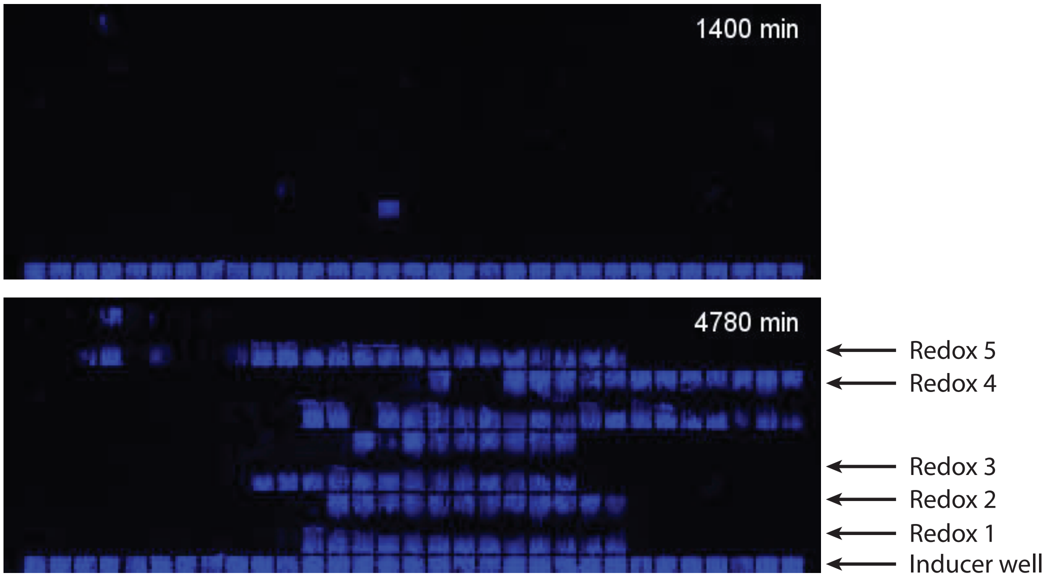


Figure 3.6: H_2O_2 -induced Toggling. Colony fluoresence following initial exposure of the lower channel to Dox (top) and long-term activation of all microfluidic channels via H_2O_2 (bottom).

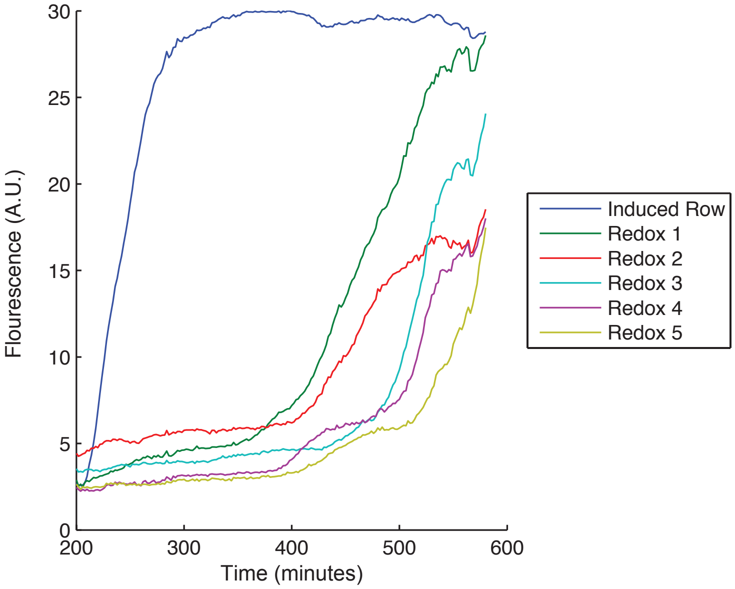


Figure 3.7: Pattern of H_2O_2 mediated colony stimulation. Time profile of the dox-induced channel and dox-free channels (Redox 1-5), ranging from closest to furthest from the chemically induced bottom channel.

A population-controlled therapeutic oscillator

Since the number of bacteria in organs such as the spleen is low (approximately 10e3), while much greater inside tumors (>10e6), a quorum based-trigger can provide a useful strategy to trigger therapeutic production and delivery only in high density populations of bacteria in tumors. Quorum sensing was previously used in synthetic biology to build synchronized oscillators, biosensors, edge detectors, and predator prey models (Danino et al., 2010; Prindle et al., 2011; Tabor et al., 2009; Balagadde et al., 2008; Basu et al., 2005) and is a well studied system for engineering gene circuits. We sought to engineer a network capable of producing quorum-mediated synchronized oscillations that maintain cell density at a prescribed level. This strain would accomplish the dual-purpose of delivering therapeutic periodically and controlling the bacterial population. Additionally, in limiting the rate at which tumor cells are destroyed, it may provide a clinical platform for reducing the risk of tumor necrosis and host sepsis.

We cloned an enhanced version of the kill-gene, ccdB (Balagadde et al., 2008), in place of AiiA in the previously published synchronized oscillator (Danino et al., 2010), where cell death takes the place of AHL degradation as a negative feedback (Fig. 3.8a). To test this design, we transformed *E. coli* with this network and observed their behavior *in vitro* in a large microfluidic chamber (Fig. 3.8b-c). The colony initially grew, followed by a burst in the *lux* promoter when it reached a critical density, and consequently terminated itself via gyrase inhibition. A small number of surviving cells at the perimeter grew to form a second colony which similarly died, and the colony proceeded in this pattern. Testing our population-controlled oscillator in multiple microfluidic designs, we were able to achieve remarkable patterns of colony proliferation and death, such as the stripes of *E. coli* in Fig. 3.9. We additionally observed, in the mcherry channel, fluorescent dye entering the cells which had recently expressed ccdB, indicating that the membrane became permeable. This will be a particularly useful property if further engineering the bacterial delivery device to express specific therapeutic agents, with the goal of releasing it proximal to the disease site.

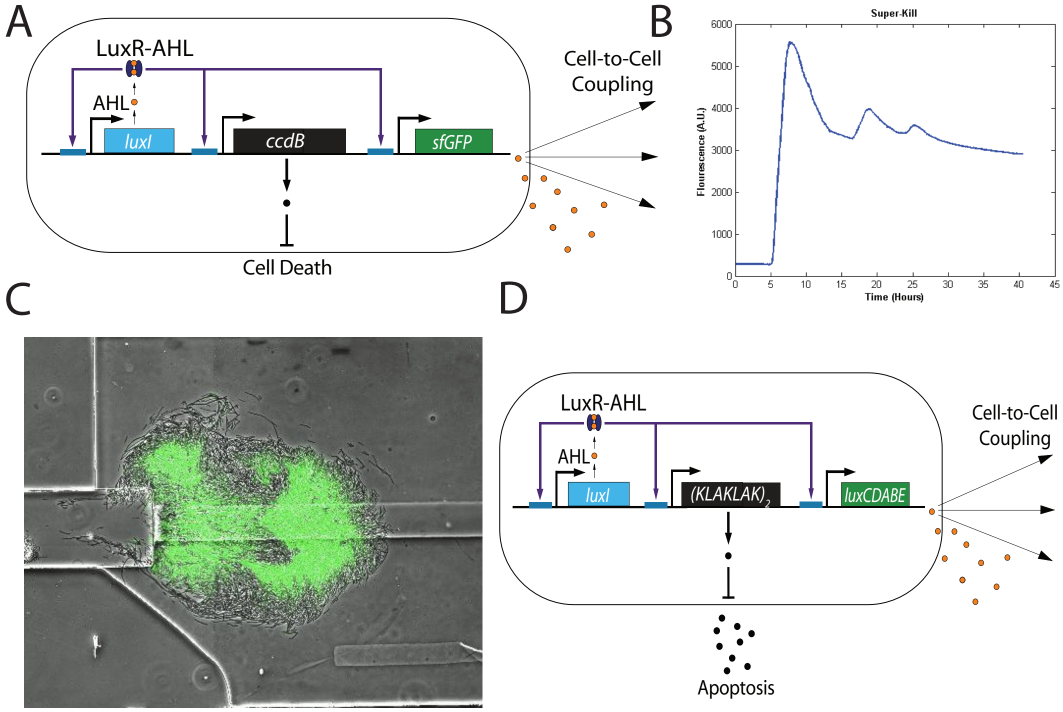


Figure 3.8: Population-controlled oscillators. **(a)** The AiiA gene is replaced with ccdB, a gyrase inhibitor that has been previously shown to be highly toxic to *E. coli*. **(b)** Mean output from the first experimental test of the circuit in *E. coli* showing 3 periods of oscillation. **(c)** Microscope image depicting the 3D colony just after the 2nd round of quorum sensing fires. **(d)** A proposed circuit in which ccdB is replaced with (KLAKLAK)$_2$, a peptide toxic to both microbes and mammalian cells.

In an integrated design (Fig. 3.8d), we could replace ccdB with a kill gene that is toxic to both microbes and mammalian cells, such as $(KLAKLAK)_2$ which has been recently used for cancer therapy (Agemy et al., 2011). Thus, every time the colony kills itself, it also kills neighboring tumor cells. This strategy could solve multiple problems: bacterial overgrowth, tumor necrosis, and tumor cell drug resistance.

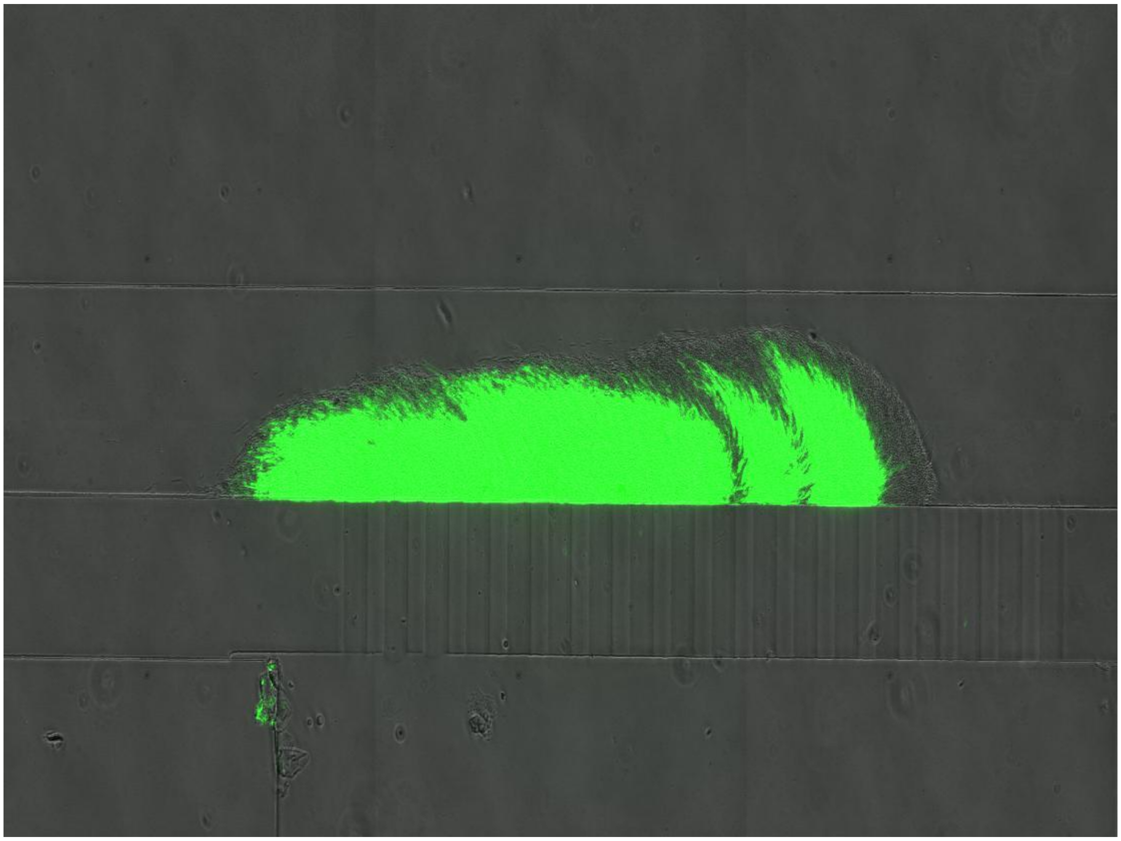


Figure 3.9: *E. coli* stripes formed by a population-controlled oscillator. Generations of *E. coli* arrange from left-to-right as populations proliferate and terminate sequentially, with new cell populations accumulating upstream of the stagnant cells from the preceeding generation.

Chapter 4

Genetic circuits in *Salmonella typhimurium*

Introduction

Synthetic biology has rapidly progressed over the last decade and is now positioned to impact important problems in health and energy. In the clinical arena, the field has thus far focused primarily on the use of bacteria and bacteriophages to overexpress therapeutic gene products. The next generation of multi-gene circuits will control the triggering, amplitude, and duration of therapeutic activity *in vivo*. This will require a host organism that is easy to genetically modify, leverages existing successful circuit designs, and has the potential for use in humans. Here, we show that gene circuits that were originally constructed and tested in *E. coli* translate to *Salmonella typhimurium*, a therapeutically relevant microbe with attenuated strains that have exhibited safety in several human clinical trials. These strains are essentially non-virulent, easy to genetically program, and specifically grow in tumor environments. Developing gene circuits on this platform could enhance our ability to bring sophisticated genetic programming to cancer therapy, setting the stage for a new generation of synthetic biology in clinically relevant microbes.

An explosion of DNA sequencing (Jasny and Zahn, 2011), synthesis (Matzas et al., 2010), and manipulation (Gibson et al., 2010) technologies has driven the development of synthetic genetic programs of increasing complexity in living cells (Weber and Fussenegger, 2011; Nandagopal and Elowitz, 2011; Khalil and Collins, 2010). Underlying this work is the hope that engineered biological systems will be used to solve important problems in energy and health over the coming years. Initially inspired by electronic circuits, researchers began by designing small transcriptional switches (Gardner et al., 2000) and oscillators (Elowitz and Leibler, 2000). These early successes fostered a growing population of physicists, computer scientists, and engineers that aimed to apply an engineering-based methodology to the design of biological systems. In the past decade, substantial success has been achieved using this genetic circuits approach termed *synthetic biology* (Weber and Fussenegger, 2011; Nandagopal and Elowitz, 2011; Hasty et al., 2002).

Multi-gene logic gates capable of integrating environmental signals have been constructed in bacteria (Tamsir et al., 2010), yeast (Regot et al., 2010), and mammalian cells (Guido et al., 2006). Electronics-inspired networks have included counters (Friedland et al., 2009), pulse generators (Basu et al., 2004), filters (Sohka et al., 2009; Hooshangi et al., 2005), and communication modules (You et al., 2004; Basu et al., 2004). Sophisticated circuits can now be controlled by light, yielding genetic programs readily tunable both *in vitro* (Tabor et al., 2009) and *in vivo* in live animals (Ye et al., 2011). Dynamic genetic clocks have been constructed that function at the single-cell (Stricker et al., 2008), colony (Danino et al., 2010), and multi-colony (Prindle et al., 2011) level in growing bacterial populations, and even in mammalian cells (Tigges et al., 2009). In a recent study, redox signaling mediated by H_2O_2vapor permitted the synchronization of millions of oscillating bacteria across an LCD-like sensor array (Prindle et al., 2011).

Early efforts toward clinical applications have utilized bacteria (Anderson et al., 2006; Xiang et al., 2006; Duan et al., 2008; Rao et al., 2005; Duan and March, 2010) and bacteriophages (Lu and Collins, 2007, 2009) (viruses that infect bacteria) to perform therapeutic functions in vivo. Commensal bacteria have been engineered to fight diabetes (Duan et al., 2008), HIV (Rao et al., 2005), and cholera (Duan and March, 2010) by producing and delivering therapeutic agents directly in the human microbiome. Because certain bacteria grow preferentially in hypoxic environments, a number of studies have engineered cancer-fighting bacteria to selectively attack tumors (Anderson et al., 2006; Xiang et al., 2006). Toward still another application, a pair of studies has engineered phages to produce foreign enzymes, making them far more potent than their unmodified counterparts at dispersing bacterial biofilms (Lu and Collins, 2007, 2009).

In most of these cases, the genetic programs involved were responsible for overexpressing target genes, similar to traditional genetic engineering where genes are added, removed, or modified one at a time in a stepwise fashion. To truly achieve its clinical potential, synthetic biology must continue to do what has made it successful: engineer progressively more complex, multi-input networks in which the triggering, amplitude, and duration of therapeutic activity is controllable. This will require using hosts that are easy to genetically modify and compatible with the clinical requirements regarding safety, immunogenicity, and drug resistance. While bacteriophage and adenovirus have their advantages, viruses have smaller genomes and therefore have a narrower range of genetic modifications, frequently induce host resistance, and are highly cell-type specific (Merril et al., 2003; Projan, 2004).

As one potential bridge between organisms such as *E. coli* and clinically relevant microbes, *Salmonella typhimurium* is a bacterial anti-cancer platform that is closely related to *E. coli* , has been extensively studied in vivo for therapeutic applications (Forbes, 2010; Pawelek et al., 1997; Hoffman, 2011; Lm et al., 2001; Forbes et al., 2003; Low et al., 1999) and has been shown to be safe in human clinical trials (Heimann and Rosenberg, 2003; Toso et al., 2002; Nemunaitis et al., 2003). The development of attenuated strains has utilized auxotrophy and phoPQ deletions to suppress virulence cell invasion and virulence (Hoffman, 2011). Lipid A mutations have been generated to reduce immunogenicity, stimulating a much weaker immune response than wild-type strains (Low et al., 1999). Despite this reduced potency, systemically injected *S. ty-*

phimurium cells retain their ability to target and selectively replicate within tumors, displaying a thousand-fold growth preference relative to other organs (Forbes, 2010; Pawelek et al., 1997; Forbes et al., 2003; Low et al., 1999; Hoffman, 2011). Their motility allows them to follow chemical gradients and penetrate deep into the tumor vasculature (Kasinskas and Forbes, 2006, 2007), much further than passively diffusing small molecules (Forbes, 2010). And many of these strains also display innate oncolytic activity, regressing tumors simply by growing in them (Pawelek et al., 1997; Nagakura et al., 2009; Low et al., 1999; Jia et al., 2007).

Perhaps the most important property of *S. typhimurium* for synthetic biology is the ease of genetic modification. It is a model organism whose genome is sequenced (McClelland et al., 2001), has knockout collections, and the genetic tools are almost identical to *E. coli* . *S. typhimurium* is capable of stably expressing recombinant DNA from plasmid-based circuits *in vivo*. This approach has already been used to produce a number of therapeutic compounds directly within tumors, but most often via "always on" expression of well-established genes (Hoffman, 2011; Guo et al., 2011; Nguyen et al., 2010). This work has laid the foundation for more sophisticated functionality, such as programmed delivery profiles that take advantage of plasmid instability (Danino et al., 2012). Such a focus will merge the dynamic sensing, production, and delivery capabilities of genetic circuits with the native tumor seeking and penetration of *S. typhimurium*.

Experimental Results

In order to test the degree to which existing synthetic circuits function in *S. typhimurium*, we transformed the attenuated strain ELH430 (SL1344 ΔphoPQ, gift of Elizabeth Hohmann, MGH) (Hohmann et al., 1996) with several genetic oscillator constructs. First, we tested a single-plasmid variant of a published single-cell gene oscillator (Stricker et al., 2008). Using our microfluidic platform (Ferry et al., 2011; Cookson et al., 2005), we observed robust oscillations for all *S. typhimurium* cells over many generations (Fig. 4.1a,b). While the qualitative period-inducer relationship was similar to *E. coli* , the curve was shifted toward faster periods as compared to *E. coli* strain JS006 (MG1655 ΔaraC,lacI) (Fig. 4.1c). In contrast, we initially expected *S. typhimurium* to oscillate slower since longer division times generally result in period lengthening (Stricker et al., 2008). When we measured the dependence of oscillatory period on temperature in *S. typhimurium*, we found the trend qualitatively similar to *E. coli* , where lower temperatures (and therefore longer doubling times) resulted in longer oscillatory periods (Fig. 4.1d). We therefore hypothesized that the faster oscillations in *S. typhimurium* are not due to growth rate differences, but rather a strain-dependent factor such as mean promoter level, transcription rate, or enzymatic degradation rate.

To explore this quantitatively, we used automated single-cell tracking using a previously developed algorithm (Mondragón-Palomino et al., 2011) to compare a large number of single-cell time courses from *S. typhimurium* and *E. coli* (Supplementary Information). Oscillators are an ideal circuit to quantify strain-specific parameters such as transcription and degradation rates since they allow for hundreds of measurements in a single experiment. For each oscillatory period, the trough-to-peak and peak-to-trough slopes were measured. Since the ClpXP degradation machinery is likely saturated (Cookson et al., 2011), the peak-to-trough slope yields an estimate for the zeroth-order enzymatic degradation rate in degrade-and-fire oscillators (Mather et al., 2009). Interestingly, we found that the apparent enzymatic degradation rate in *S. typhimurium* was roughly 1.5-fold that of *E. coli* (Fig. 4.4a). In our computational model of the oscillator, this increase reproduced the experimentally observed period-inducer relationship (Fig. 4.4b).

Next, we transformed *S. typhimurium* with a quorum-sensing oscillator that had been previously characterized in *E. coli* (Danino et al., 2010), and observed coherent, colony-level oscillations for more than 48 hours (Fig. 4.2a,b). Here, we found that the period-flow rate dependence was markedly different in *S. typhimurium* than in the original study, where oscillatory period was much longer and changed very little across

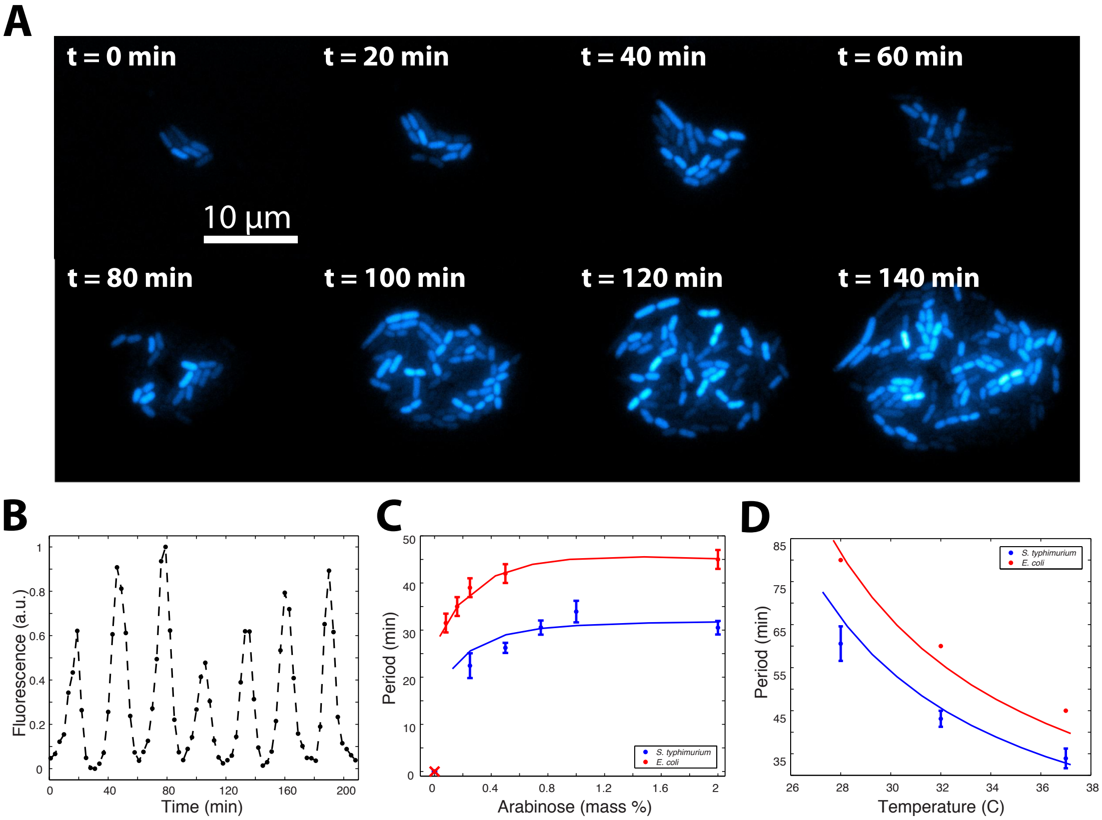


Figure 4.1: A fast, robust, and tunable genetic oscillator in *S. typhimurium*. (a) Timelapse fluorescence microscopy depicting asynchronous oscillations in a growing colony of *S. typhimurium*. (b) A single-cell trajectory extracted from image data. (c) Period vs. inducer concentration for *S. typhimurium* compared to original data taken in *E. coli* . The trends are qualitatively similar yet *S. typhimurium* is shifted toward shorter periods. Points are experimental measurements fit to a line generated by computational modeling. (d) Period vs. temperature for *S. typhimurium* compared to original data taken in *E. coli* with similar trends.

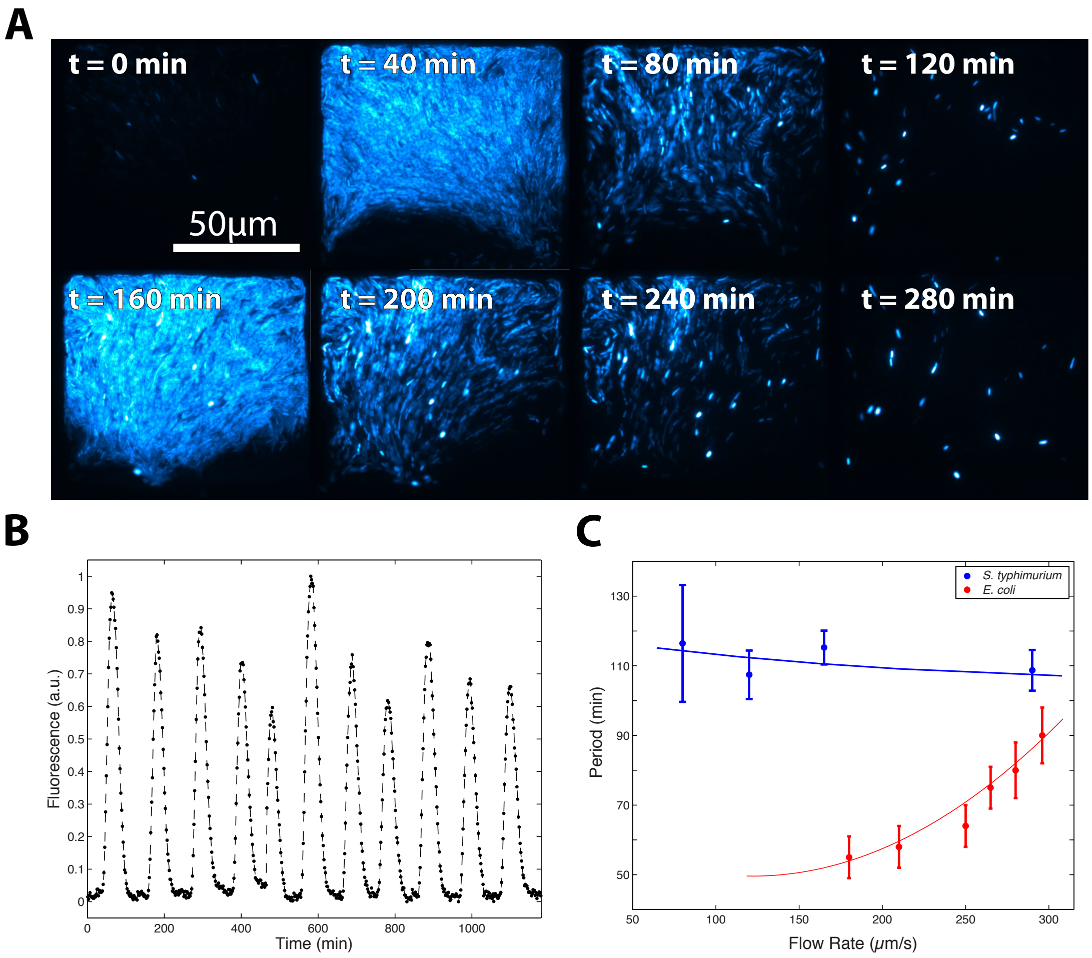


Figure 4.2: A synchronized quorum of genetic clocks in *S. typhimurium*. (a) Timelapse fluorescence microscopy depicting coherent oscillations at the colony-level for a growing colony of *S. typhimurium*. (b) A colony trajectory extracted from image data that illustrates the regularity of oscillations over time. (c) Period vs. flow rate for *S. typhimurium* compared to original data taken in *E. coli* . *S. typhimurium* displays much higher periods that appear to be independent of flow rate.

a wide range of flow rates (Fig. 4.2c). Interestingly, while increased degradation rate resulted in faster oscillations for single-cells (Fig. 4.4b), our computational model correctly predicts the opposite trend for the quorum-sensing oscillator when degradation is increased (Fig. 4.4c).

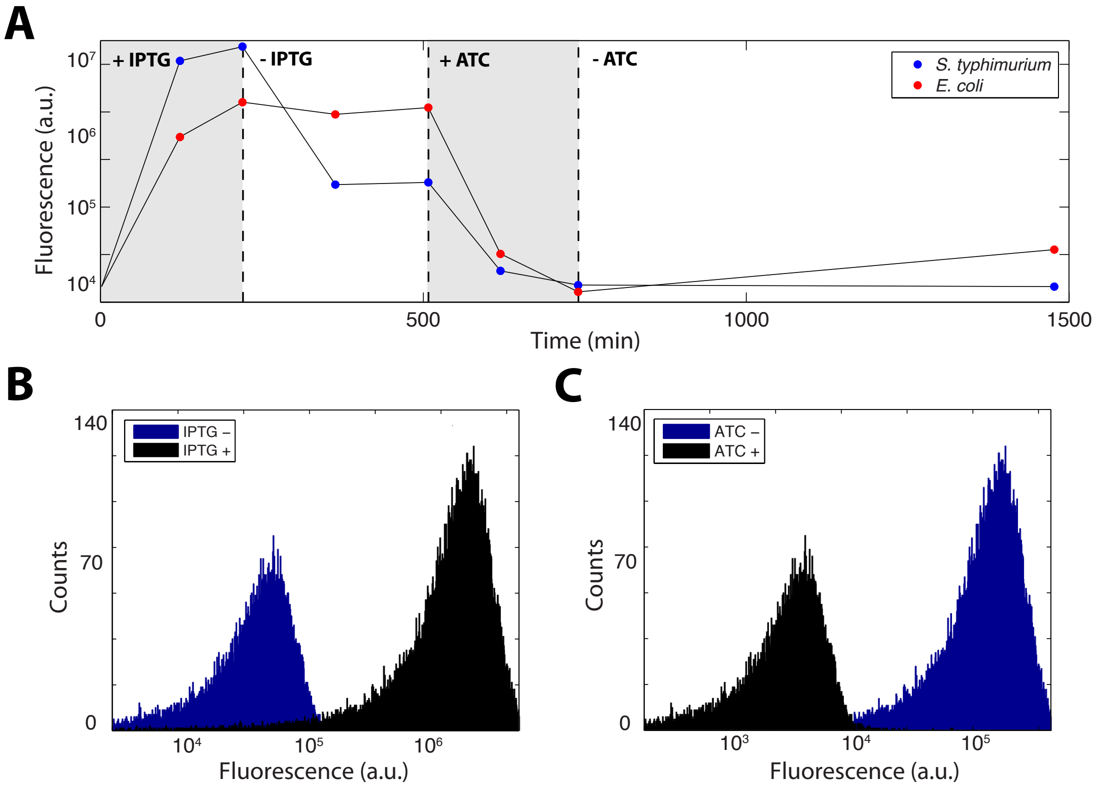


Figure 4.3: A genetic toggle switch in *S. typhimurium*. (a) A time course of fluorescence output that illustrates switching by both IPTG and ATC quantified by flow cytometry in periodically diluted batch culture experiments. (b) Raw flow cytometer data illustrating switching by 2mM IPTG and (c) 500 ng/μl ATC.

Finally, we tested the original genetic toggle switch, plasmid pIKE107 (Gardner et al., 2000). In this circuit, a transient pulse of IPTG inducer turns the switch ON and reporter expression is maintained at a high level. A second pulse of ATC inducer turns the switch OFF, dropping reporter expression indefinitely. In periodically diluted batch culture experiments similar to the original study, we used flow cytometry to observe robust switching and bistability when inducing with either 2 mM IPTG or 500 ng/ul dox in cultures growing at 37 C (Fig. 4.3a-c). Interestingly, the fluorescence level at which *S. typhimurium* settled after we removed IPTG was lower than the same circuit in *E. coli* (Fig. 4.3a). We suspected that the differences in apparent degradation and expression rates (Fig. 4.4a) might explain this change, since the steady-state repressor balance would be adjusted.

To test this hypothesis, we used the original computational model of the toggle switch (Gardner et al., 2000) and quantified the steady-state expression level over time for strain parameters measured in *E. coli* and *S. typhimurium*. We found that the *S. typhimurium* parameters reproduced the experimentally observed curves, where expression rises to a higher level when switched ON then decays to a lower steady-state when IPTG is removed (Fig. 4.4d). While these parameters are particularly important for dynamic circuits, they can also impact the performance of stable switches since repressors are continuously being produced and degraded.

Conclusions and outlook

A central issue in the design of genetic circuits is the degree to which native and engineered networks should be integrated. Synthetic biology began by fully isolating itself from the strain background, using it solely to supply energy, enzymatic machinery, and a cellular volume in which to function. In contrast, industrial applications in medicine and energy have commonly utilized a variety of microbes for their native networks (Ruder et al., 2011; Weber and Fussenegger, 2011; Alper and Stephanopoulos, 2009). As our biological knowledge of native networks and our ability to engineer new circuits has improved, it has become increasingly possible to blend these two strategies (Nandagopal and Elowitz, 2011).

S. typhimurium is an ideal strain for clinical synthetic biology since it is closely related to *E. coli* , well studied *in vivo*, has safety precedence for clinical trials in humans, and displays a thousand-fold growth preference for tumor environments (Forbes, 2010; Pawelek et al., 1997; Hoffman, 2011; Lm et al., 2001). Moving to other microbes for clinical and industrial purposes will require the determination of the critical

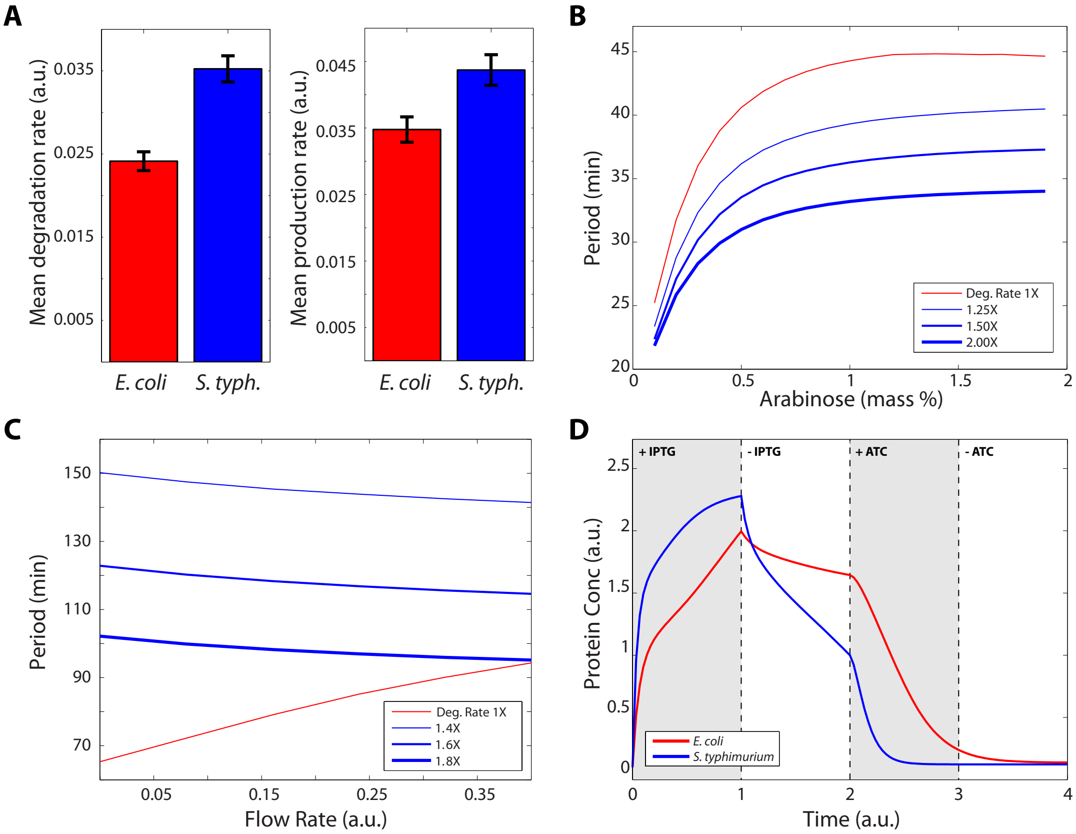
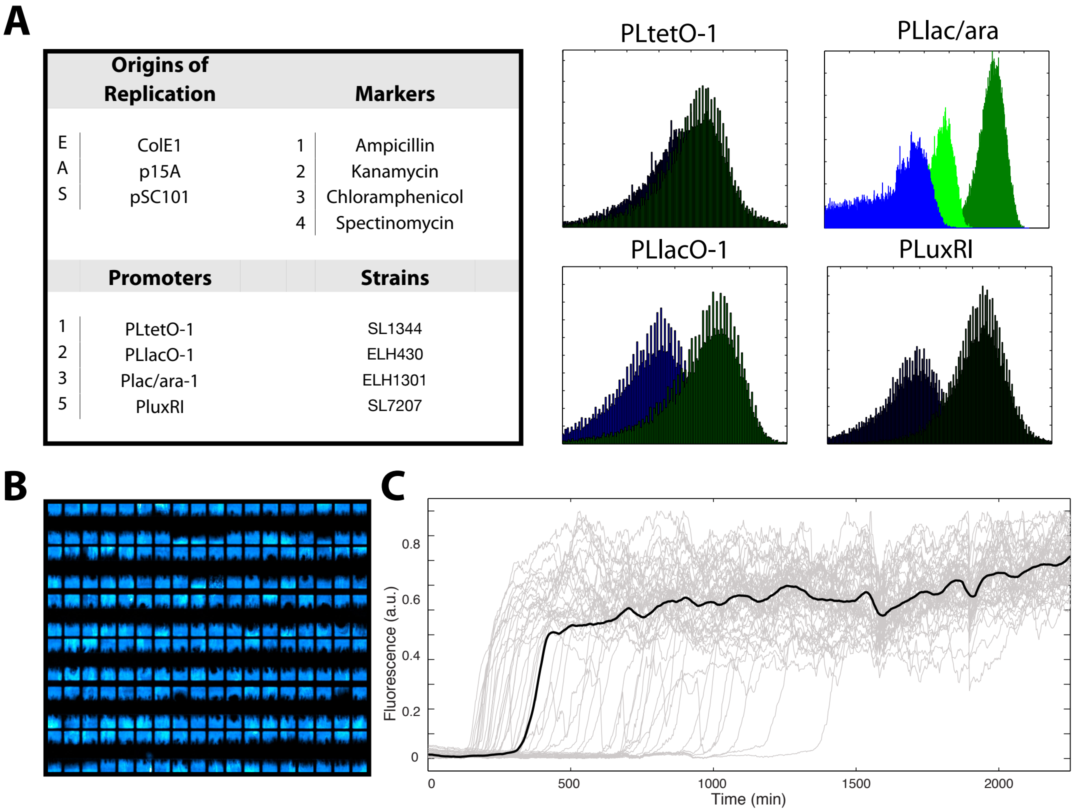


Figure 4.4: Computational modeling of *S. typhimurium* genetic circuits. (a) Comparison of enzymatic degradation rate between *S. typhimurium* and *E. coli* generated from automated single-cell tracking. Degradation rate is approximately 1.5X higher in *S. typhimurium*. (b) A higher degradation rate results in the shorter periods observed experimentally for the single-cell oscillator. (c) In contrast, increased degradation rate results in longer periods for the quorum-sensing oscillator that are comparatively unchanged with flow rate. (d) Increased degradation and expression rates produce the experimentally observed behavior for the *S. typhimurium* toggle switch.

strain parameters that define the space of bacteria capable of hosting genetic circuits. Next steps will involve measurement of these parameters and testing circuits in strains of interest that are further removed in the phylogenetic tree (Wu et al., 2009). One such roadmap would begin with more distantly related gamma proteobacteria like *Pseudomonas aeruginosa* before moving outside the phylum to alpha proteobacteria such as *Calubacter crescentus*. Additionally, individual components and modules can also receive a "portability" score that estimate the degree to which they translate to other hosts. For example, while lacI- and tetR-based circuits are nearly universal, more generally the function of other components are likely to be more sensitive to strain-specific parameters. This work will enable synthetic biology to move beyond *E. coli* into a diverse range of microbes for clinical and industrial applications.

Plasmid Characterization

As a preliminary test, we systematically characterized the popular pZ plasmid system (Lutz and Bujard, 1997) (origins, markers, and promoters) in *S. typhimurium*, finding that they all worked similarly to *E. coli* (Fig. 4.5a). While plasmid maintenance is a commonly reported problem for *S. typhimurium* (Bauer et al., 2004; Gahan et al., 2007), our plasmid-based circuits appeared to be highly stable. To examine this further, we designed an experiment in which hundreds of independent colonies of *S. typhimurium* were transformed with a vector containing constitutively-expressed GFP and grown in a microfluidic device without selective pressure for many generations (Fig. 4.5b). For the duration of the experiment (48 hours, or about 100 divisions), all colonies maintained the plasmid circuit, expressing GFP at a nearly constant level (Fig. 4.5c). To verify that GFP signal was a measure of plasmid maintenance, we returned kanamycin to the media and observed that all colonies continued to grow and fluoresce (not shown). Thus, plasmid-based circuits could be reliably studied *in vitro* in microfluidic devices without stability concerns.



	Origins of Replication		Markers
E	ColE1	1	Ampicillin
A	p15A	2	Kanamycin
S	pSC101	3	Chloramphenicol
		4	Spectinomycin

	Promoters	Strains
1	PLtetO-1	SL1344
2	PLlacO-1	ELH430
3	Plac/ara-1	ELH1301
5	PluxRI	SL7207

Figure 4.5: Properties of plasmid-based circuits in *S. typhimurium*. (a) Table of strains, promoters, markers, and origins used in this study. Flow cytometer data depicting induction of a number of popular promoters in *S. typhimurium*. (b) Fluorescence microscopy image of an array of colonies used to study plasmid maintenance in *S. typhimurium*in the absence of selection. (c) Individual trajectories extracted from longterm time-lapse data.

Microscopy and Microfluidics

We used a microscopy system similar to our recent studies (Danino et al., 2010). Fluorescent images were taken at 4X every 30 seconds using the EMCCD camera (20ms exposure, 97% attentuation) or 2 minutes (2s exposure, 90% attenuation) using a standard CCD camera to prevent photobleaching or phototoxicity.

In each device, *S. typhimurium*cells are loaded from the cell port while keeping the media port at sufficiently higher pressure than the waste port below to prevent contamination. 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% Tween which prevented cells from adhering to the main channels and waste ports.

To measure fluid flow rate before each experiment, we measured the streak length of fluorescent beads (1.0 μm) upon 100 ms exposure to fluorescent light. We averaged at least 1,000 data points for each.

We used several microfluidic devices over the course of the study. For single-cell oscillators (Fig. 4.1), we used a previously described device consisting of a trapping region and a dynamic switch (Mondragón-Palomino et al., 2011). Traps have dimensions 40 μm wide x 50 μm long x 0.95 μm high, with the long sides open to media flow. Since *E. coli* and *S. typhimurium* cells have a 1 μm diameter, the trap maintains growing cells in a monolayer. For colony oscillators (Fig. 4.2), we used a previously described device consisting of arrays of square trapping regions (Danino et al., 2010; Prindle et al., 2011). Trap dimensions were always 100 μm x 85 μm x 1.65 μm high and spacing between traps was 25 μm. This size allows cells to grow in a colony arrangement rather than a monolayer, while still allowing quantitative measurement of colony fluorescence.

Degradation and Production Rate Quantification

Single cell fluorescence trajectories were obtained from time-lapse movies using custom software previously developed in MATLAB (Mondragón-Palomino et al., 2011). Each cell fluorescence trajectory represents the median GFP fluorescence signal inside that cell over time. Using built-in MATLAB functions we identified the peaks and troughs for each trajectory. The degradation rate was calculated by taking the amplitude change from peak to the successive trough and dividing by the time change between the peak and the trough. These peak-to-trough sections of the trajectory represent the time when the production of GFP is repressed and the observed dynamics are solely driven by degradation of GFP. Similarly we calculated the net production rate, by calculating the amplitude change from trough to successive peak and dividing by the time change between the trough and the peak. The measurement gives the net production rate, which includes the degradation of the protein.

Table 4.1: Mean Degradation and Net Production Rate during *S. typhimurium* oscillation. Calculation performed based on rate of amplitude change between peaks and troughs for each trajectory.

	Mean Degradation Rate (SE)	**Mean Net Production Rate (SE)**
E. coli	0.024 (0.001)	0.035 (0.002)
S. typhimurium	0.035 (0.002)	0.044 (0.002)

Modeling

To generate the plot in Figure 4.4d, we used previously described genetic toggle switch model (Gardner et al., 2000). We included three additional parameters to model the effects of IPTG (C_{IPTG}), ATC (C_{ATC}), and dilution (D) on the synthesis and degradation of proteins:

$$\frac{\partial u}{\partial t} = \frac{C_{IPTG_{(0,1)}}\alpha_u}{1+v^n} - (\gamma_u + D)u \tag{4.1}$$

$$\frac{\partial v}{\partial t} = \frac{C_{ATC_{(0,1)}}\alpha_v}{1+u^n} - (\gamma_v + D)v \tag{4.2}$$

In this model, we set n=2 to allow for cooperativity of repression of both promoters. C_{IPTG_0} and C_{ATC_0} were set to 1 for the case of no inducers present. Next, we used metropolis algorithm to find the rest of the parameters to fit the qualitative nature of the curves from Figure 4.1a. The parameters found to generate the *E. coli* curve were: $C_{IPTG_1} = 1.25$,$C_{ATC_1} = 1.68$,$\alpha_u = 4.28$,$\alpha_v = 5.80$,$\gamma_u = 1.76$,$\gamma_v = 2.37$,$D = 0.11$. The parameters found to generate the *S. typhimurium*curve were: $C_{IPTG_1} = 1.25$, $C_{ATC_1} = 1.68$,$\alpha_u = 11.00$,$\alpha_v = 8.36$,$\gamma_u = 4.86$,$\gamma_v = 3.21$,$D = 0.08$. It is interesting to note that the optimized parameters show higher production and degradation as well as lower dilution for *S. typhimurium*curve relative to *E. coli* curve, which correlates well with our experimental measurements.

The dynamics of single cell oscillator were modeled using previously described model for activator (a_2) and repressor (r_4) proteins (Stricker et al., 2008). The production and degradation of these proteins is described by the following set of reactions:

$$
\begin{aligned}
P_{0,0}^{a/r} &\xrightarrow{b_{a/r}} P_{0,0}^{a/r} + m_{a/r} \\
P_{1,0}^{a/r} &\xrightarrow{\alpha b_{a/r}} P_{1,0}^{a/r} + m_{a/r} \\
m_a &\xrightarrow{t_a} m_a + a_{uf} \\
m_r &\xrightarrow{t_r} m_r + r_{uf} \\
a_{uf} &\xrightarrow{k_{fa}} a \\
r_{uf} &\xrightarrow{k_{fr}} r \\
a + a &\xrightleftharpoons[k_{-da}]{k_{da}} a_2 \\
r + r &\xrightleftharpoons[k_{-dr}]{k_{dr}} r_2 \\
r_2 + r_2 &\xrightleftharpoons[k_{-t}]{k_t} r_4 \\
a_{uf} &\xrightarrow{\lambda f(X)} \emptyset \\
r_{uf} &\xrightarrow{f(X)} \emptyset \\
a &\xrightarrow{\lambda f(X)} \emptyset \\
r &\xrightarrow{f(X)} \emptyset \\
a_2 &\xrightarrow{\lambda f(X)} \emptyset \\
r_2 &\xrightarrow{f(X)} \emptyset \\
r_4 &\xrightarrow{f(X)} \emptyset
\end{aligned}
$$

We updated the degradation function $F(X)$ to include dilution as follows:

$$f(X) = \frac{\gamma}{c_e + X} + DX \tag{4.3}$$

Here, X is the total number of $ssrA$ tags in the system (one for each monomeric version, two for dimers, and four for tetramers, including proteins bound to operator sites). We varied the parameter γ from 1x to 2x to evaluate the effect of degradation difference between *E. coli* and *S. typhimurium*on the period of oscillation calculated from single cell model simulations. Dilution rate was calculated from experimentally measured cell half life as $\frac{ln(2)}{T_{\frac{1}{2}}}$.

To model the dynamics of the quorum-sensing oscillator, we used our previously described model for intracellular concentrations of LuxI (I), AiiA (A), internal AHL (H_i), and external AHL (H_e) (Danino et al., 2010),

$$\begin{aligned}
\frac{\partial A}{\partial t} &= C_A[1-(d/d_0)^4]\,G(\alpha,\tau) - \frac{\gamma_A A}{1+f(A+I)} - DA && (4.4)\\
\frac{\partial I}{\partial t} &= C_I[1-(d/d_0)^4]\,G(\alpha,\tau) - \frac{\gamma_I I}{1+f(A+I)} - DI && (4.5)\\
\frac{\partial H_i}{\partial t} &= \frac{bI}{1+kI} - \frac{\gamma_H A H_i}{1+gA} + D(H_e - H_i) - DH_i && (4.6)\\
\frac{\partial H_e}{\partial t} &= -\frac{d}{1-d}D(H_e - H_i) - \mu H_e + D_1\frac{\partial^2 H_e}{\partial x^2} && (4.7)
\end{aligned}$$

To model the difference in periods of oscillation between *E. coli* and *S. typhimurium* we varied the degradation parameters γ_A and γ_I. We looked at the changes in the period over different values of the flow rate parameter μ, while varying the degradation prapameters from 1x to 2x of the original model value. To account for the difference in doubling time between the two strains, we introduce exponential decay terms into the model to account for dilution in addition to the enzymatic degradation terms. We add terms $-DI$, $-DH_i$, and $-DH$ to the first three equations respectively, with $D = \frac{ln(2)}{T_{\frac{1}{2}}}$. We then looked at how the change in doubling time affected the period of both strains shown in Figure 4.2d.

Supplementary information, including methods, supplementary figures and tables are available free of charge via the Internet at http://pubs.acs.org/.

Acknowledgements

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health (GM069811), the Department of Defense National Defense Science and Engineering Graduate Fellowship (AP), the National Science Foundation Graduate Research Fellowship (PS), and the Misrock Postdoctoral fellowship (TD). We would like to thank K. Pogliano for helpful discussions.

Chapter 4 contains material originally published as Prindle, A., Selimkhanov, J., Danino, T., Samayoa, P., Goldberg, A., Bhatia, S., and Hasty, J. ***ACSSynBio*** (2012): Genetic circuits in *Salmonella typhimurium*. Copyright permission to republish here was granted by ACS Publications.

Chapter 5

Programming cyanobacteria to produce protein nutrients at supra-agricultural efficiencies

Introduction

Agriculture occupies nearly a third of all terrestrial land, making it by far the most widely deployed solar energy technology globally. Despite this prevalence, current agricultural and livestock production methods suffer from multiple energetic and resource inefficiencies, resulting in annual solar energy conversion into nutrients of $>$100-1,000-times lower than photovoltaic efficiencies. Here, we apply synthetic biology and systems modeling to create streamlined cyanobacterial systems for the direct photosynthetic production of high-quality protein nutrients at supra-agricultural photonic efficiencies. Our modeling shows that these systems provide the potential to produce nutrient proteins at $>$100-times higher annual photonic efficiencies than common plant- and livestock-based systems. We develop a foundation biological tools to rationally program the marine cyanobacteria *Synechococcus sp. PCC 7002* and demonstrate the selective production and secretion of model protein nutrients, including production of the agricultural protein ovalbumin at 8-fold the annualized photosynthetic efficiency of animal production. In concert with advances in photobioreactor design, our streamlined approach motivates the potential for continuous agricultural nutrient production at dramatically improved areal and resource efficiencies compared with conventional agriculture.

Agriculture serves as the primary conduit through which solar energy is converted into essential nutrients for human consumption. Plant- and animal-based agriculture is the single largest user of arable land and fresh water globally and is estimated to produce a third of global greenhouse gas emissions (Vermeulen et al., 2012; Food and Agriculture Organization of the United Nations , FAO). Despite thousands of years of improvement, the fundamental photosynthetic efficiency of agricultural nutrient production has changed relatively little over time and today's agricultural productivities are estimated to be only $\sim$4-6-fold higher than the yields of conserved crops thousands of years ago (Araus et al., 2003; Butzer, 1976). The rising global demand for high-quality protein sources, increasing population, and limited supply of phosphate and other key inputs are expected to further intensify pressures on food production (rom, 2009; Fischer et al., 2009; Godfray et al., 2010; Goodland et al., 2009).

The integration of synthetic biology with systems-level mathematical modeling has provided a framework for rewiring biological systems and introducing entirely new functionalities, including toggle switches (Gardner et al., 2000), logical operators (Friedland et al., 2009; Tamsir et al., 2010; Moon et al., 2012), and synchronized community behaviors (You et al., 2004; Tabor et al., 2009; Danino et al., 2010; Prindle et al., 2011; Danino et al., 2012). However, relatively little attention has been directed toward applications of synthetic biology to the central challenges facing agriculture and sustainable protein nutrient production. Here, through computational and experimental approaches, we find that streamlined nutrient production systems relying on single cell photosynthetic organisms offer the potential to produce high-quality protein nutrients with $>$100-fold higher annual photonic efficiencies than traditional agricultural practices. We extend the toolkit of synthetic biology in marine autotrophic strains and provide a systematic demonstration that single-cell photosynthetic cyanobacteria can be engineered to serve as novel conduits for 'sunlight-to-protein' conversion.

A Photosynthetic Nutrient Catalyst

Chloroplasts, theorized to have arisen from cyanobacterial endosymbionts, are the photosynthetic engines of higher plants, harnessing solar energy to drive the assimilation of CO_2 and subsequent generation of biomass and nutrients throughout the plant (Fig. 5.1a). While the chloroplast's short-term photosynthetic energy efficiency can approach that of photovoltaic-driven electrolysis in the same light spectra, agricultural crops direct much of this energy toward plant growth and only store a small portion of their captured energy in harvestable seeds and other food products (Blankenship et al., 2011; Fischer and Edmeades, 2010; Zhu et al., 2008, 2010). To overcome this limitation of agricultural nutrient production, we sought to generate synthetic biological systems that could approach the chloroplast's theoretical photosynthetic energy efficiency by directly converting solar energy into desired nutrients. Specifically, we hypothesized that engineered single cell cyanobacterial autotrophs could act as ultra-efficient nutrient production chassis by dramatically reducing the biomass requirements of agricultural crops and allowing nutritive proteins to be photosynthetically produced at supra-agricultural efficiencies. Using a combination of metabolic modeling and synthetic biology, we present a systematic framework for generating cyanobacterial hosts that realize significant improvements in photosynthetic protein production efficiencies relative to traditional agriculture processes.

We initially sought to characterize the theoretical photosynthetic efficiency of protein production in agricultural and cyanobacterial systems. The instant photosynthetic efficiency of agriculture is governed by thermodynamic limits of chemical conversion (Blankenship et al., 2011; Zhu et al., 2008) and is bounded by the plant's requirement to divert fixed carbon towards biomass. We determined the average photosynthetic efficiency of agricultural production for each crop based on USDA reported harvestable yield averages and NREL reported incident solar radiation. Theoretical crop efficiencies were defined as their maximally observed short-term rate of harvestable biomass production (Beadle and Long, 1985; Zhu et al., 2010). In order to assess the theoretical efficiency of direct cyanobacterial nutrient production, we modeled the photonic efficiency for nutritive protein production with a customized metabolic flux model for cyanobacterial strains producing heterologous proteins (Fig. 5.1b). Building on metabolic reconstructions of photosynthetic performance in cyanobacteria, including nine cooperative electron flow pathways that retain quantum efficiency of photon capture under high light conditions (Nogales et al., 2012), we generated a model for analyzing the photosynthetic conversion efficiency of sequence-specific protein biosynthesis and secretion by cyanobacteria (Allen and Palsson, 2003). Assuming a steady state metabolic demand for transcription and mRNA turnover, conversion efficiency of incident photosynthetically active radiation (PAR) was constrained by the translation rates facilitated by the expression construct. Compared to the most efficiently produced commodity crop proteins, measured during their optimally efficient window of growth, we find that an autotroph has the potential to directly catalyze protein production at $>$20 times higher photosynthetic efficiency (Fig. 5.1c). Because a cyanobacterial nutrient production system offers the potential for year-round production, we assessed the annual areal protein production potential for such a system. Even in projecting the peak efficiency of protein biomass production in corn across an entire growing season, cyanobacterial production systems would provide a 50-fold advantage in annual protein production. A further benchmark is the capacity to directly produce those proteins that serve as the standard for nutritive quality (Schaafsma, 2000). The theoretical efficiency for cyanobacterial production of ovalbumin and beta-lactoglobulin, the major constituents of egg and whey protein respectively, approaches three orders of magnitude improvement upon what is typically achieved through animal production of these proteins.

We next sought to apply computational modeling of photosynthetic energy capture and carbon metabolism to guide cyanobacterial host selection and genetic optimization for the production of agricultural proteins. We first modeled the tradeoff

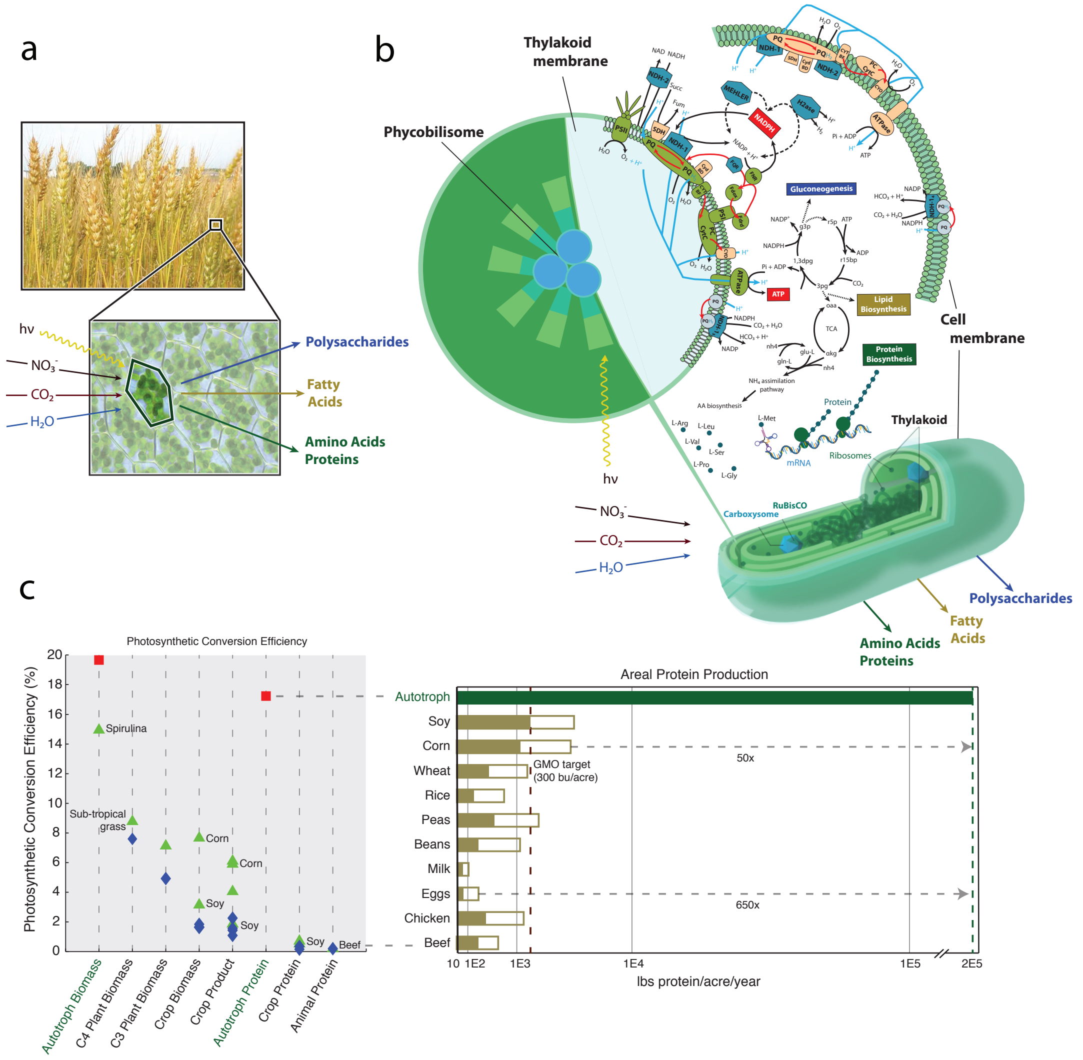


Figure 5.1: Direct nutrient catalyst. **(a)** Conceptual diagram illustrating a photosynthetic unit capable of directly utilizing incident radiation to fix carbon for nutrient production, including polypeptides, fatty acids, and polysaccharides. **(b)** Network diagram depicting the flux-based-analysis (FBA) photosynthetic model of nutrient production. Modeled metabolic units include the photosynthetic apparatus, central metabolic pathways, and protein biosynthesis reactions. **(c)** Comparison of photosynthetic conversion efficiency of incident PAR between a simulated single cell autotroph (red squares) and agricultural production at peak (green triangles) and average growing season (blue diamonds) efficiencies. Conversion efficiency represents the ratio of final energy content of biomass, edible product, or protein to energy from incident PAR. Efficiency improvements over crops by direct protein production translate to log-scale differences in areal yield. Areal agricultural production capacity is calculated as peak photosynthetic conversion efficiency applied to total incident PAR trough a growing season of a top-yielding region in the United States (open bars) or current agricultural yields (filled bars).

in carbon metabolism between cellular growth and heterologous protein production in metabolically active cells. Simulations of photosynthetic conversion efficiency to protein were conducted across strains of varying maximal doubling time, testing protein production over the range of viable growth rates for each. Metabolic efficiencies were bounded computationally such that the simulated growth rate could not exceed its maximal measured growth rate by capping the photosynthetic capacity of a model strain. Our simulations predicted that cyanobacterial hosts with rapid doubling rates could offer higher photosynthetic conversion efficiencies for protein production than slower growing strains due to the direct shift in metabolism from biomass generation to protein accumulation that is mediated by heterologous genetic constructs (Fig. 5.2a). To identify a strain with such rapid biomass accretion, we experimentally examined growth rates of a diversity of cyanobacterial strains, while assessing their resilience to salt water culture conditions and long-term viability in the absence media replenishment, and nominated *Synechococcus sp. PCC 7002* as a preferred chassis for production (Fig. 5.2b and Fig. 5.13).

While the synthetic biology toolkit in well-studied heterotrophic organisms has benefited from decades of development of foundational parts and some fresh water cyanobacterial hosts have been characterized and modified for the production of lipid-based fuels and specialty chemicals (Weyer et al., 2010; Wijffels et al., 2010; Robertson et al., 2011), *Synechococcus sp. PCC 7002* and other marine cyanobacteria have remained relatively unexplored as synthetic biology chassis (Wang et al., 2012; Berla et al., 2013; Gronenberg et al., 2013). In order to divert metabolic energy from biomass to heterologous protein production, we sought to develop expression systems with the capacity to efficiently produce heterologous proteins in this host. In distinction to classical heterotrophic hosts in synthetic biology, PCC7002 and other marine cyanobacteria partition genomic information across their chromosome and a large number of plasmids that can asymmetrically segregate during host division (Xu et al., 2011). In order to generate strains with the capacity to efficiently express nutritional protein transcripts, we screened a set of promoters derived from multiple cyanobacterial strains with insertion targets in the host choromosome, as well as the full suite of plasmids, while verifying stable propagation of each expression cassette through extensive plasmid segregation. This yielded two promoters, P*cpc** from *Thermosynechococcus elongatus BP-1* and P*psbaII* from *Synechocystis elongatus PCC 7942*, that when inserted into host plasmid PAQ1 enabled protein photosynthetic conversion efficiencies throughout the linear growth phase that far exceeds ($\sim$5-fold) crop and animal protein efficiencies (Fig. 5.2c). Validating the model behavior, increases in promoter strength shifted protein photosynthetic conversion efficiency along the simulated carbon metabolism isocline, with more efficient protein production occurring at the expense of reduced cellular biomass production (Fig. 5.2d). Such a carbon partitioning effect has been effective in engineered fermentative systems by separating growth and production phases to achieve high maximal productivities, and proposed for direct production of diesel-like alkanes in cyanobacteria (Robertson et al., 2011). Recognizing this demonstrated efficiency tradeoff for carbon metabolism, our model predicts that an optimized rate of translation for a strain with a doubling time of 4 hours at the tested culturing densities would allow approximately 10-fold and 140-fold improvements of protein photosynthetic efficiency beyond current crop and animal production methods, respectively.

We next sought to apply these tools to efficiently produce agricultural proteins in cyanobacterial hosts. As heterologous protein production often exhibits variation in the compatibility between expression constructs and desired heterologous proteins (Punt et al., 2002; van Dijl et al., 2013) (Fig. 5.18), we screened expression vectors at multiple plasmid sites in PCC7002 to test for photosynthetic production of ovalbumin, the major protein constituent of chicken eggs. At lab-scale, we were able to observe ovalbumin production efficiencies that surpassed the estimated annual photosynthetic efficiency of agricultural-based ovalbumin production (Fig. 5.2e, Fig. 5.25). This is the first time a sustainable means of producing an equivalent protein has been validated and has been done at 8x the limit of nature for producing the same. Moreover, this exceeds the solar efficiency of all animal-based protein production, demonstrating the potential for cyanobacteria to act as conduits for the improved production of resource-intensive agricultural proteins.

We further applied metabolic modeling to identify bottlenecks to photosynthetic conversion efficiency and inform strain engineering and culture optimization. By simulating the efficiency of protein production across cell densities and translation rates, we observed the importance of balancing the PAR flux distributed per cell with the rate of carbon assimilation into protein allowed by the expression construct. This balancing of PAR flux and translation capacity ensures minimal photon energy dissipation through ATP-consuming futile cycles (Fig. 5.7). We identified an optimality for protein production in which the PAR flux available per cell for carbon fixation matches the strain's translation capacity, thereby balancing cellular carbon capacity with nutrient protein production (Fig. 5.3a). This effect has been empirically described previously by increasing culture density to maintain photoautotrophic growth rates at high photon flux density irradiance (Qiang et al., 1998; Richmond et al., 2003; Grobbelaar, 1994; Grobbelaar et al., 1996; Janssen et al., 2003). A loss in efficiency is due to the PAR flux per cell both exceeding the rate of photosystem utilization as well as its subsequent photoinhibitory effects.

Our computational modeling of the tradeoff between protein and cell biomass production indicated that an inducible promoter system could allow biomass expedited expansion to higher cell-densities prior to initiating protein production. Such a system could circumvent the carbon metabolism bottleneck in which stronger protein expression results in a lower final culture cell-density and provide improved per-cell PAR conversation efficiency to protein by operating at higher culture density (Fig. 5.3a and Fig. 5.20). To evaluate this, we built a novel set of cumate inducible promoters and constructed genetic cassettes for insertion across multiple plasmid loci (Fig. 5.3b). We observed the system to provide tight regulation over protein expression, allowing the initiation of protein production at variable optical densities through the linear growth phase of PCC7002. In practice, increasing cell density in a fixed volume culture also slows the rate of biomass production as cyanobacteria approach their carrying capacity (Fig. 5.20). To experimentally decouple the subsequent shift towards a higher protein

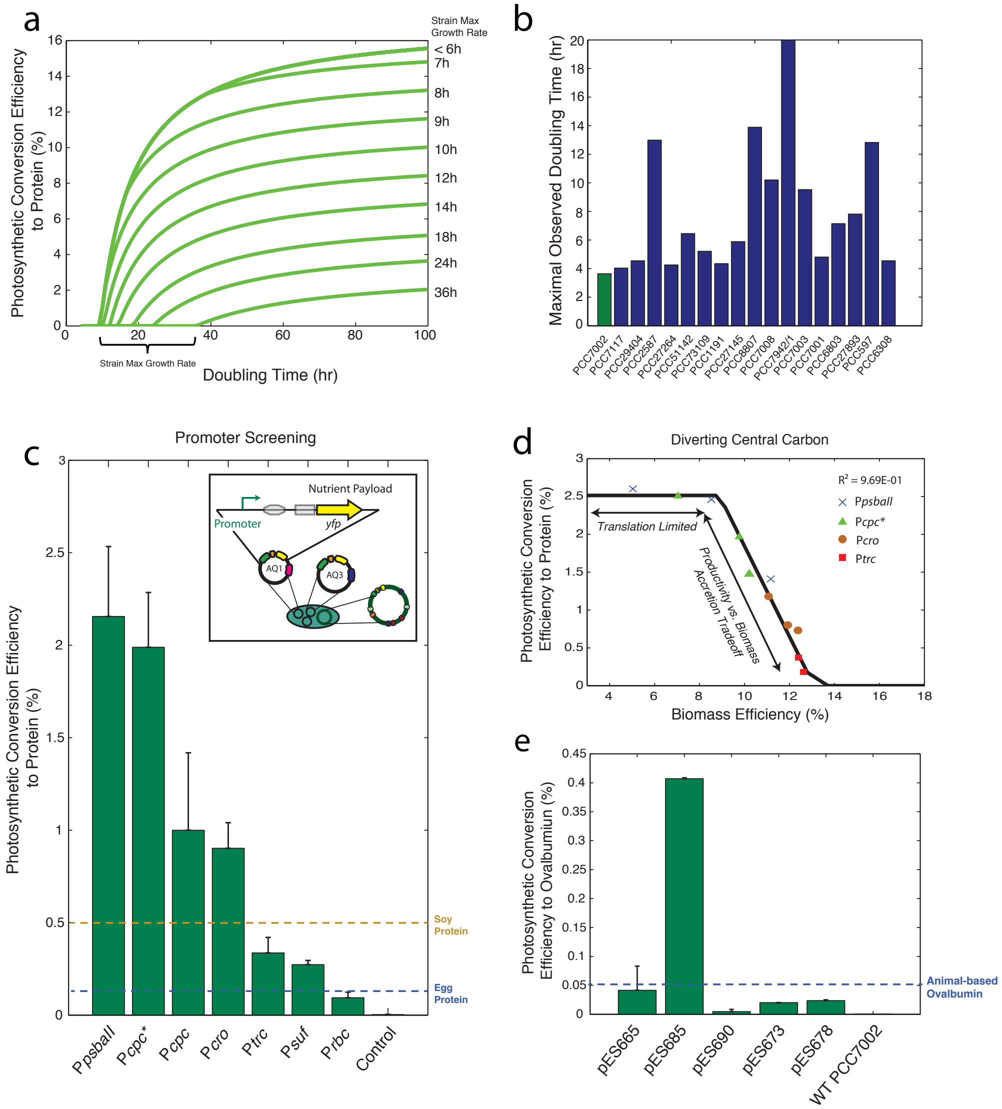


Figure 5.2: Cyanobacterial host characterization. **(a)** Simulated photosynthetic capacity for protein production scales with strain maximal growth rate. Curves represent individual strains across their capable doubling times, demonstrating increasing production efficiencies when metabolism is shifted away from cellular growth. **(b)** Growth rate screen of cyanobacterial strains, identifying a fast growing strain (PCC7002) for characterization of single-cell protein production efficiency. **(c)** YFP production efficiency in PCC7002 across a set of expression constructs, demonstrating experimental PAR conversion efficiencies to protein that exceeds agricultural capacity. **(d)** Relationship between efficiency of PAR conversion to biomass and protein for individual expression constructs. The experimental trade off between metabolism for cellular growth and protein production matches closely to the predicted carbon metabolism isocline of the photosynthetic model. **(e)** Ovalbumin production efficiency in PCC7002. Successful expression is observed from pAQ3-P*cpc* at an efficiency of 0.41%.

production efficiency as demonstrated by the cell's carbon partitioning above, we used a low copy expression vector to bring protein production out of a carbon metabolism limited regime (Fig. 5.21). In this range of promoter strength, efficiency is limited by the number of messages per cell. By inducing protein expression at increased cell densities, we observed conversion efficiency to increase 3-fold as PAR flux per cell was reduced to match the message-limited translation capacity (Fig. 5.3c). This finding suggests that further improvements in production can be achieved by moving to higher light intensities, utilizing stronger expression systems, and inducing expression at even higher cell culture densities (Fig. 5.12).

Nutrient Secretion

We next probed the potential for cyanobacterial chassis to selectively secrete protein nutrients and provide a potential conduit for continuous nutrient production and facile purification. Under conditions where biomass accretion is limited to maintain a high conversion efficiency to protein, our model shows that heterologous protein production must retain a metabolic balance by programming the cell to secrete the protein into the culture medium. Secretion of heterologous protein has been demonstrated in PCC6803 using a native leader sequence, but these sequences do not directly translate to the PCC7002 host (Sergeyenko and Los, 2003). To test for *Synechococcus sp. PCC 7002* capacity for heterologous protein secretion, we used a neural network strategy, trained on a SignalP training set (Bendtsen et al., 2004), to screen >3800 genes in the synechococcal protein database for identification of N-terminal Sec-type signal sequences that are functional in PCC7002. From a ranking of all putative Sec-type leader sequences, a set of 48 N-terminal leader-sequences were selected and fused to the N-terminus of heterologous lichenase to test for secretion from PCC7002 (Fig. 5.4a). We found that the addition of specific N-terminal Sec-type signal sequences enabled up to an 80-fold increase in secretion rate over that of native lichenase, demonstrating that marine autotrophs have the capacity to direct extracellular localization of heterologous

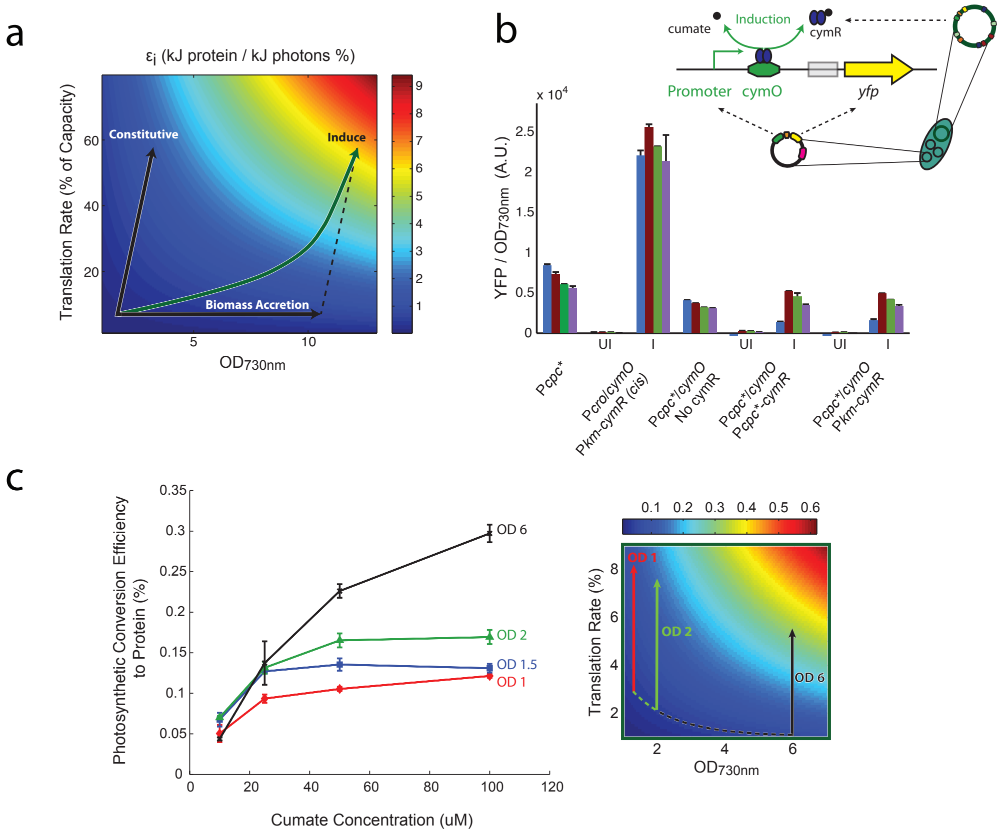


Figure 5.3: Optimal PAR flux per cell. **(a)** Computational scan of culture density and translation rate. Photosynthetic efficiency is optimized when the photon flux per cell is scaled to match the strain's translation capacity. Arrows depict protein conversion efficiency under constitutive production, pure cellular growth, or inducible production. **(b)** Cumate-inducible promoter expression range, normalized to cell density. Both P*cro-cymO* and P*cpc*-cymO* exhibit tight repression without induction (UI, 0 μM cumate), while P*cro-cymO* achieves a higher range of induction in the induced state (I, 50 μM cumate). **(c)** YFP production efficiency per-cell increases upon induction at higher cell-densities. Cultures were tested in constant light (80 μE/m^2/s), varying the PAR flux per cell. Photosynthetic efficiency measurements for the characterized expression constructs are mapped to the model's profile.

protein nutrients (Fig. 5.4b-c) and providing a path toward continuous production over prolonged time periods.

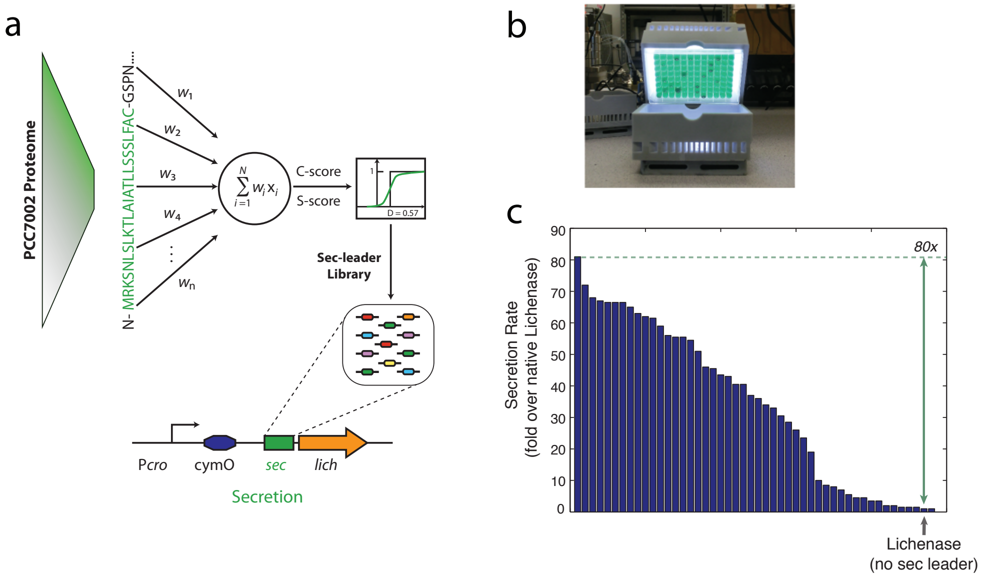


Figure 5.4: Protein secretion. **(a)** Conceptual diagram of the neural network implemented to predict natural signal peptides within the synechococcal library of polypeptide sequences. A D-score above 0.57 indicates prediction of a sec leader on a queried protein by the SignalP v4.0 network. **(b)** Schematic diagram of the cyanobacterial high-throughput screening platform, enabling use of 96-well plates for parallel characterization of strains and culture conditions. **(c)** Secretion screening in PCC7002 of a set of 48 network-predicted Sec-type leader sequences fused to the heterologous Lichenase protein. Addition of an N-terminal leader peptide sequence increases heterologous protein secretion 80-fold over that of the native protein sequence.

The technology that we set forth provides a programmable conduit to convert solar energy into agricultural proteins at supra-agricultural efficiencies. We computationally and experimentally illustrate this system's potential to produce nutritional proteins at efficiencies that surpass peak agricultural conversion rates, including our production of the model protein ovalbumin at 6-fold the annualized photosynthetic efficiency of animal production. The ability to divert large proportions of incident energy towards carbon assimilation into nutrients rather than supporting growth of unharvested biomass within plants, overcomes a significant barrier to achieving yield improvements in traditional agricultural systems. Moreover, the ability to utilize marine strains of cyanobacteria, capable of operating in salt conditions, opens the possibility for this type of platform to reduce fresh water requirements for nutrient production. Our experimentally-validated model of photoautotrophic nutrient production further provides a tool to rationally guide the improvement of protein production through iterative strain engineering and systems design of customized bioreactors. Given the accelerated genome engineering and testing cycles of single-cell autotrophs compared with higher plants, the tools and approach described here offer the potential to rapidly realize systems with radically improved areal efficiencies of nutrient production over current agricultural methods.

Methods

Strains and plasmids

All parent strains used in this study were saltwater or freshwater cyanobacteria from the The Pasteur Culture Collection of Cyanobacteria (PCC) at the Institut Pasteur (Paris, France) or the American Type Culture Collection (ATCC; Manassas, VA). All engineered strains in this study were derivatives of the saltwater cyanobacteria *Synechococcus sp. PCC 7002*. Promoters screening was chosen based on strong and/or inducible promoters from various cyanobacteria (P*cpc*, P*cpc**, P*suf*, P*rbc*), synthetic (P*trc*), Pseudomonas putida (P*cym*), or the lambda phage (P*cro*). Insertion sites for heterologous constructs within PCC7002 included extrachromosomally replicating plasmids pAQ1, pAQ3, or chromosomal loci.

Heterologous protein production

All PCC7002 derivatives were cultivated in A+ medium (18 g/l sodium chloride, 0.6 g/l potassium chloride, 1 g/l sodium nitrate, 5 g/l magnesium sulfate heptahydrate, 50 mg/l monobasic potassium phosphate, 266.4 mg/l calcium chloride, 30 mg/l tetrasodium EDTA dehydrate, 3.89 mg/l iron (III) chloride hexahydrate, 8.20 mM Tris pH 8.2, 4 mg/l vitamin B12, and 1 ml/l P1 metals solution (1000x P1 metals: 34.26 g/l H_3BO_3, 4.32 g/l MnC_{12}-$4H_2O$, 315 mg/l ZnCl, 30 mg/l MoO_3, 3mg/l $CuSO_4$-$5H_2O$, 12.15 mg/l CoC_{12}-$6H_2O$), or a derivative containing a higher amount of nitrogen, phosphate, iron, and antifoam. All growth conditions were undertaken at 35°C and 2% (v/v) CO_2. 96DWB experiments were undertaken with LED lights at $\sim$100 μE/m^2/s at 800rpm, while shake flask experiments were undertaken with fluorescent lights at $\sim$80 μE/m^2/s at 150rpm. Samples were taken intermittently from between 1 and 10 days.

Analytics

For intracellular protein expression experiments, analysis was undertaken using either: 1) in situ fluorescence at 505/550nm for YFP against a purified standard curve, 2) anti-FLAG antibody and dot-blot platform, against a FLAG-BAP standard curve, or 3) the Caliper LabChip GXII system with automatic peak integration. For the latter two methods, cyanobacteria pellets were lysed by bead milling using the QIAGEN TissueLyser II. For protein secretion with lichenase, an activity measurement based on glucose release by the dinitrosalicylic acid assay was used to quantify productivity.

Model Construction and Simulations

Metabolic reconstruction for photosynthetic nutrient production

The *Synechocystis* network was constructed using the COBRA toolbox (Schellenberger et al., 2011) to specifically assess photosynthetic nutrient conversion efficiency. The previously published metabolic model of *Synechocystis sp. PCC 6803* (Nogales et al., 2012) was expanded to include sequence specific demands for heterologous protein synthesis based on the model described by Allen and Palsson (Allen and Palsson, 2003). It was likewise expanded to include 2-oxoglutarate decarboxylase, identified by Zhang and Bryant (Zhang and Bryant, 2011) to functionally complete the

TCA cycle. While a number of detailed genome-scale reconstructions exist for *Synechocystis sp. PCC 6803* (Shastri and Morgan, 2005; Hong and Lee, 2007; Fu, 2009; Navarro et al., 2009; Knoop et al., 2010; Montagud et al., 2010, 2011; Saha et al., 2012) as well as a recent reconstruction of our experimental host, *Synechococcus sp. PCC 7002* (Hamilton and Reed, 2012), we constructed the nutrient production model based on the *Synechocystis sp. PCC 6803* model by Nogales et al. due to the specific treatment of the photosynthetic alternate flow pathways in *i*JN678 as well as its detailed biomass objective function. The ability to capture robust photoautotrophic metabolism throughout the high-light, carbon-limited regime provides a more accurate model of the cyanobacterial capacity for nutrient production under outdoor conditions.

To model heterologous protein production, we added reactions to include transcription, mRNA decay, translation, and secretion, accounting for metabolic demands on the existing *Synechocystis* reconstruction based on the specific sequence of an experimental construct. For tRNA charging, we used existing metabolites included as dead-end metabolites in the *i*JN678 reconstruction. The protein production reaction set (Table 5.1) included additional metabolites for the expression construct (G), RNA polymerase (Rp), initiated RNA polymerase (G^*), mRNA (M), unfolded mRNA for degradation (M_o), ribosome (R), initiated ribosome (R^*), cytoplasmic and periplasmic protein (P_c and P_p respectively).

The mRNA unfolding reaction assumes a transcript secondary structure with 50% of the nucleotides engaged in a base pair. To ensure a metabolic demand for transcription and mRNA turnover, both upper and lower reaction bounds for mRNA degradation (equation 4) were fixed to the turnover rate ($-\dot{m_c} = \lambda * m_c$) that is defined by transcript half life (λ), set to 2 minutes (Rott et al., 2003), and the steady-state mRNA level ($mRNA_c$). An example YFP construct for characterizing photosynthetic efficiency is implemented in Supplementary Dataset S1.

Table 5.1: A transcript containing n nucleotides, representing g gmp, a amp, u ump, and c cmp; a coding sequence of length N, where AA^{cod} represents the coding sequence amino acid distribution; a signal peptide sequence of length S, where AA^{sig} represents the coding sequence of the signal peptide. ($N^{+} = N + 1$)

$$\begin{aligned}
G + Rp &\rightarrow G^{*} \\
G^{*} + nNTP &\rightarrow (N-1)pp_i + M + G + Rp \\
M + \frac{n}{4}(atp + h_2o) &\rightarrow \frac{n}{4}(adp + h + p_i) + M_o \\
M_o + (n-1)h_20 &\rightarrow nNMP + (n-1)h \\
M + R + f_{met}tRNA + gtp + h_2o &\rightarrow R^{*} + tRNA_{met} + gdp + h + p_i \\
R^{*} + 2Ngtp + N^{+}h_2o + \sum_{i=1}^{20} AA_i^{cod} &\rightarrow R + M + 2N(pi + gdp) + N^{+}h + P_c \\
P_c + S(atp + h_2o) &\rightarrow P_p + S(adp + h + p_i) + \sum_{j=1}^{20} AA_j^{sig}
\end{aligned}$$

Simulation constraints and objective function

Simulations were performed using constraints for BG-11 minimal medium as described in (Nogales et al., 2012) and a conversion to dry cell weight of 0.6E-12 g/cell. Purely autotrophic metabolism was simulated using HCO_3 uptake constrained to 3.7 mmol/gDW/h, representing the maximum CO_2 uptake rate described in (Nogales et al., 2012; Shastri and Morgan, 2005). The photon exchange flux was defined based on the maximum irradiation available to the culture media, ranging from 175 μE/m^2/s in flask experiments to 7300 MJ/m^2/y, representing the average outdoor ground radiation in high light-intensity regions (Robertson et al., 2011). Available photosynthetically active radiation (PAR) flux was calculated with a PAR radiation fraction of 0.487, a culture reflection loss of 15%, and using an average PAR photon energy of 226E3 J/mol as in (Robertson et al., 2011).

For a given level of light-intensity, photon exchange per cell (or dry cell mass) was calculated based on the media depth and culture density (0.28 gDW/L*OD$_{730nm}$ for PCC7002). For flask experiments, the media depth of 0.31cm was derived from the flask dimensions (63.2mm inner diameter) and 10 ml culture volume. Modeling of the system under high-light intensity (as in Fig. 5.3d) used a light path of 1.5cm based on experimental characterization of optimal light path on areal output rate under high light intensity (Qiang et al., 1998; Richmond et al., 2003; Grobbelaar, 1994; Grobbelaar et al., 1996). Photon uptake rate assumes a well-mixed system where each cell experiences an average distribution of available photon flux. The dependence on light path of a cell's maximal growth rate is shown in Fig. 5.5a. Longer paths distribute photon flux (per m^2) over more cells for a given culture density, limiting the growth rate. Multiple strains of varying photosynthetic capacity were modeled by scanning the proton exchange flux constraint to affect the simulated cell's capacity to make ATP via the ATP synthase reaction. (Fig. 5.5b).

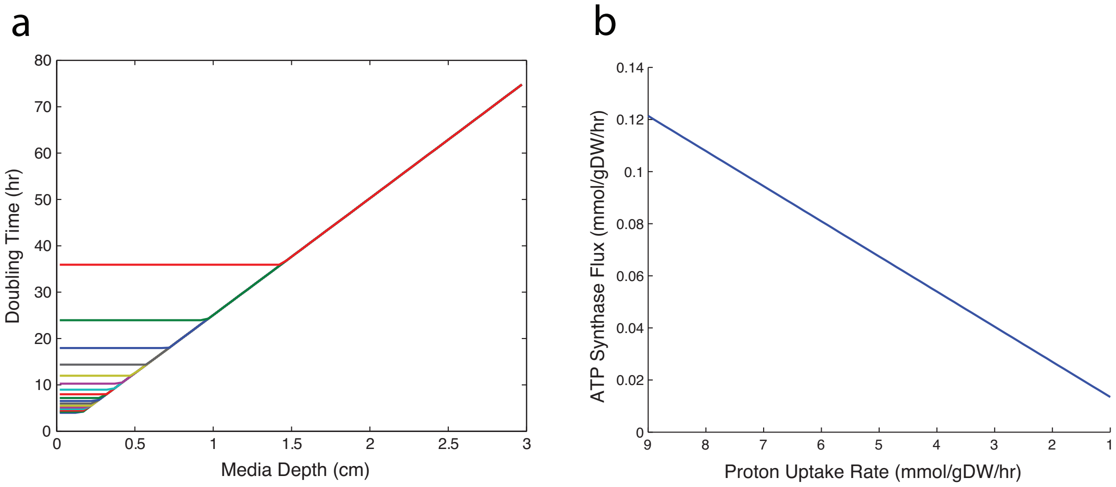


Figure 5.5: Light path characterization. Strain growth rate depends on light path for a given culture-cell density and light intensity. (a) Shorter media depth allows an elevated photon distribution per cell, increasing the culture's growth rate. Multiple curves represent a scan of the proton exchange flux constraint, and it's effect on cell growth capacity. (b) Varying the proton exchange flux affects the rate of ATP production via ATP synthase, shifting maximal growth rate.

Protein production efficiency was calculated as the per-cell energy conversion rate:

$$\varepsilon_i = \frac{kJprotein/gDW/hr}{kJincidentphoton/gDW/hr} \tag{5.1}$$

The protein energy density was calculated based on the metabolizable energy content of amino acids bound in a polypeptide (Ferrer-Lorente et al., 2007), calculating amino acid-specific densities of 23.32 KJ/g YFP and 23.62 kJ/g Ovalbumin. The biomass energy density for PCC7002 was calculated to be 26.5 kJ/g, based on the biomass composition outlined in Nogales et. al. (Nogales et al., 2012) and component energy densities of 24 kJ/g protein, 39.8 kJ/g lipid, and 17.7 kJ/g glycogen. In modeling heterologous protein production, the objective function was set to the translation reaction (equation 6). The biomass function was either fixed to a specific growth rate (as in Fig. 5.2a) or unconstrained to allow for maximal protein production, with the highest protein production efficiencies occurring when all carbon metabolism is shifted away from cellular growth. We captured the varying expression strength of promoters by applying a constraint to the translation flux (equation 6). Changing the steady state level of mRNA, we scaled the upper bound on translation rate as a function of ribosome processing rate (T_{el} = 16 AA/s) and minimal ribosome spacing (D_{min} = 17 AA),

$$v6 = \frac{mRNA_c * T_{el}}{D_{min}} \tag{5.2}$$

Scanning the translation rate upper bound, we found the rate limit for per-cell heterologous protein production was determined by the strains' photosynthetic capacity (Fig. 5.6). This rate limit was used as the strain 100% translation capacity for Fig. 5.3a. For light intensities that provide photon flux above which the translation rate can support, the protein production efficiency diminishes as excess photons are instead dissipated via futile cycles (Fig. 5.7)). Increasing the culture density, however, can reduce the photon flux per cell to match its cellular translation rate, increasing the overall efficiency of the culture (per m^2). This is demonstrated across a range of cell-densities and light paths in Fig. 5.8.

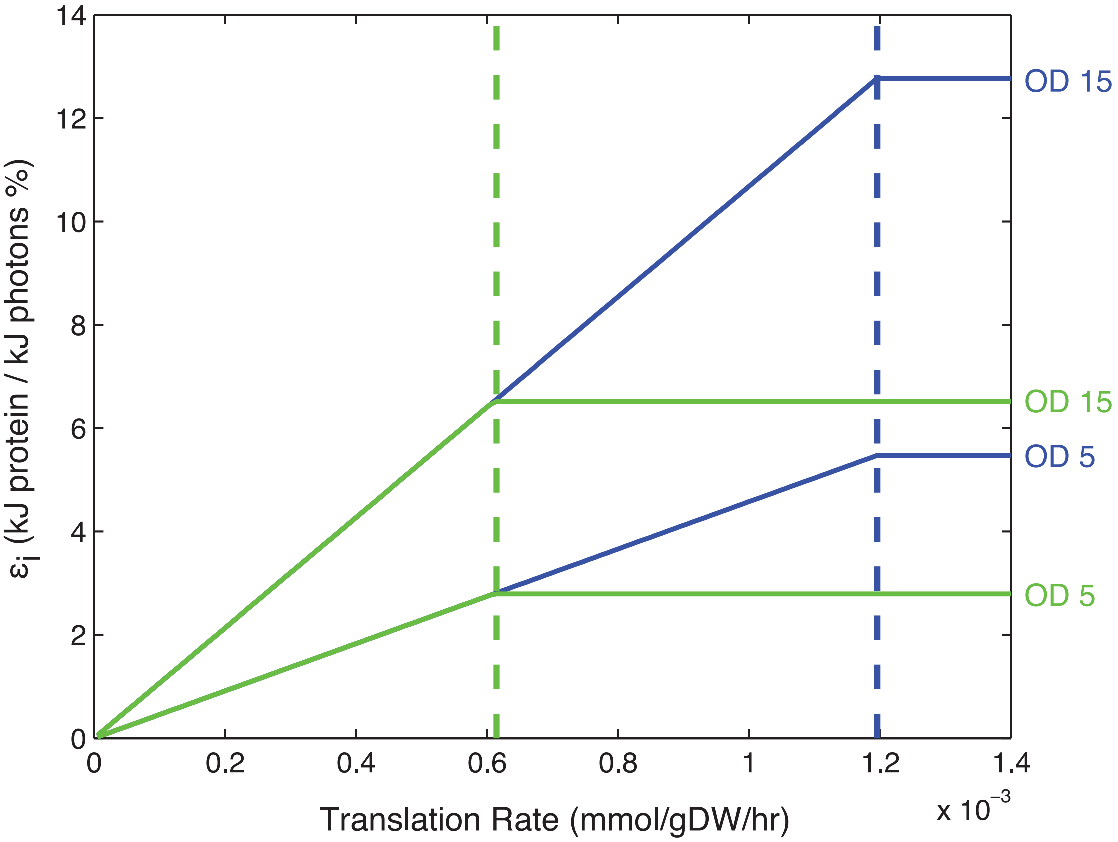


Figure 5.6: Maximal Translation Capacity. x-axis is the upper bound on translation rate applied in the simulation. Proton uptake rates of 1.25 and 2.5 mmol/gDW/hr were characterized (green and blue lines respectively). Dashed lines are the maximum translation flux the cells reach, mirroring their modeled photosynthetic capacity.

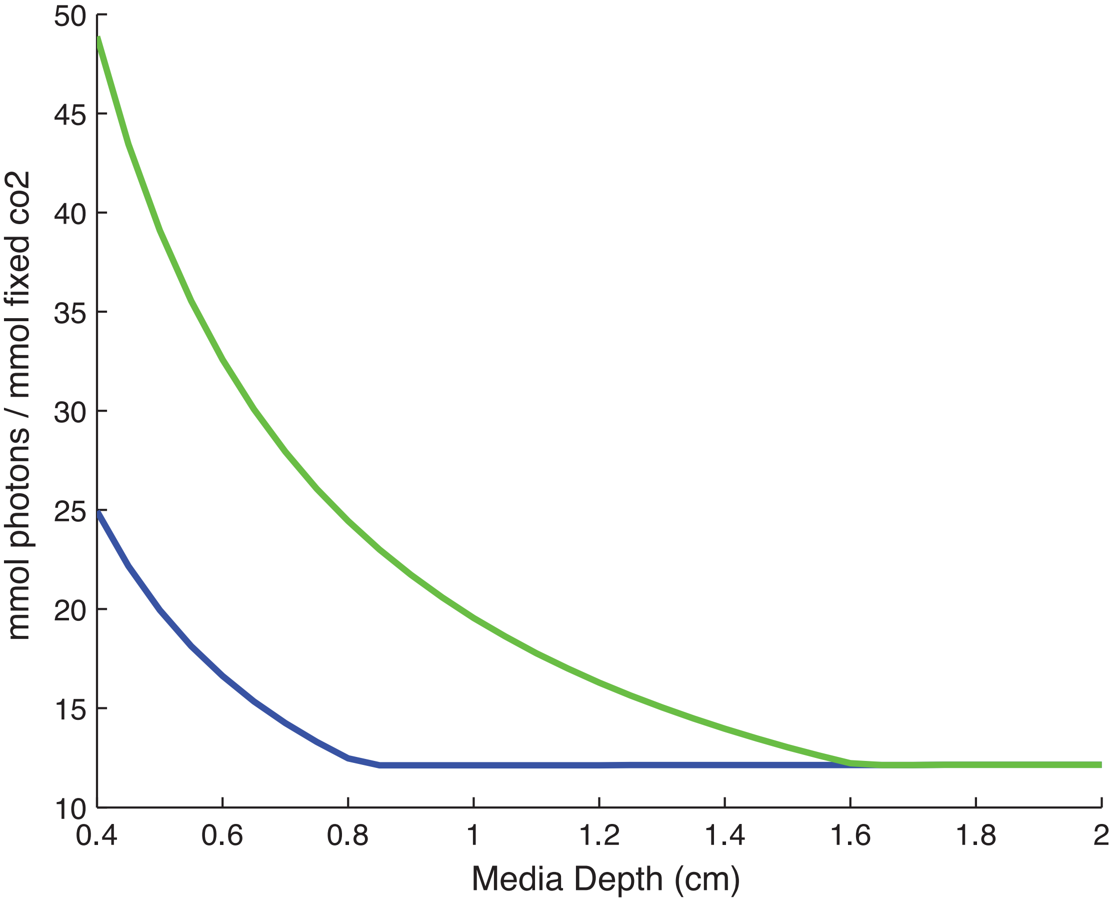


Figure 5.7: Quantum Efficiency. Photon flux per cell exceeding the stain translation capacity leads to decreasing quantum efficiency (mmol photon / mmol fixed CO^2). Proton uptake rates of 1.25 and 2.5 mmol/gDW/hr were characterized (green and blue lines respectively) across media depth.

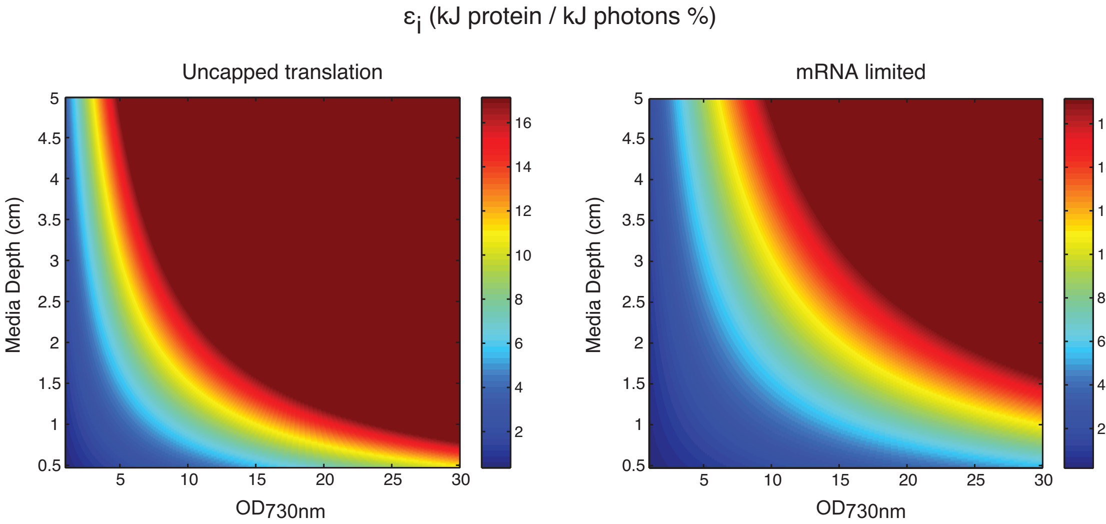


Figure 5.8: Culture Density Dependence. Protein production efficiency depends on culture-cell density and media depth. Increased efficiencies occur at higher translation rates for a given cell-density and media depth. The production efficiency was tested at translation rates of (a) 100% of strain capacity and (b) 50% of strain capacity.

The constitutive promoter screen (Fig. 5.2c-d) was fit to the production model by scanning photosynthetic capacity and translation rate upper bound. Model conditions matched experimental at OD$_{730nm}$ 8 and 85 μE/m^2/s PAR light intensity. The fit gave a Pearson linear correlation coefficient (r) of 9.845E-01. The theoretical strain protein photosynthetic efficiency (at low growth rate and with translation rate upper bound at 100% strain capacity) is 7.6% (Fig. 5.9). The inducible promoter characterization was similarly modeled by varying OD$_{730nm}$ and translation rate upper bound (induction range). A constitutive promoter producing during the early growth phase as in Fig. 5.9 will be less efficient then a promoter which allows biomass expansion to higher cell-densities prior to production (demonstrated in Fig. 5.3), capable of approaching a theoretical protein production efficiency of 16%. Final productivity estimates at high light (Fig. 5.1) were done using a dense cell culture, where photon uptake rate per cell becomes the limiting constraint for (per m^2) heterologous protein production. In this regime, differences in maximal growth rate collapse onto one curve, as the light available to each cell is below that which can support faster growth rates (Fig. 5.10). As such, protein production levels per cell are below their strain photosynthetic capacity, and thus not limited by the translation rate limit. The areal yield per m^2, however, yields optimal protein production since the culture is efficiently utilizing all incident radiation (Fig. 5.11 and 5.12).

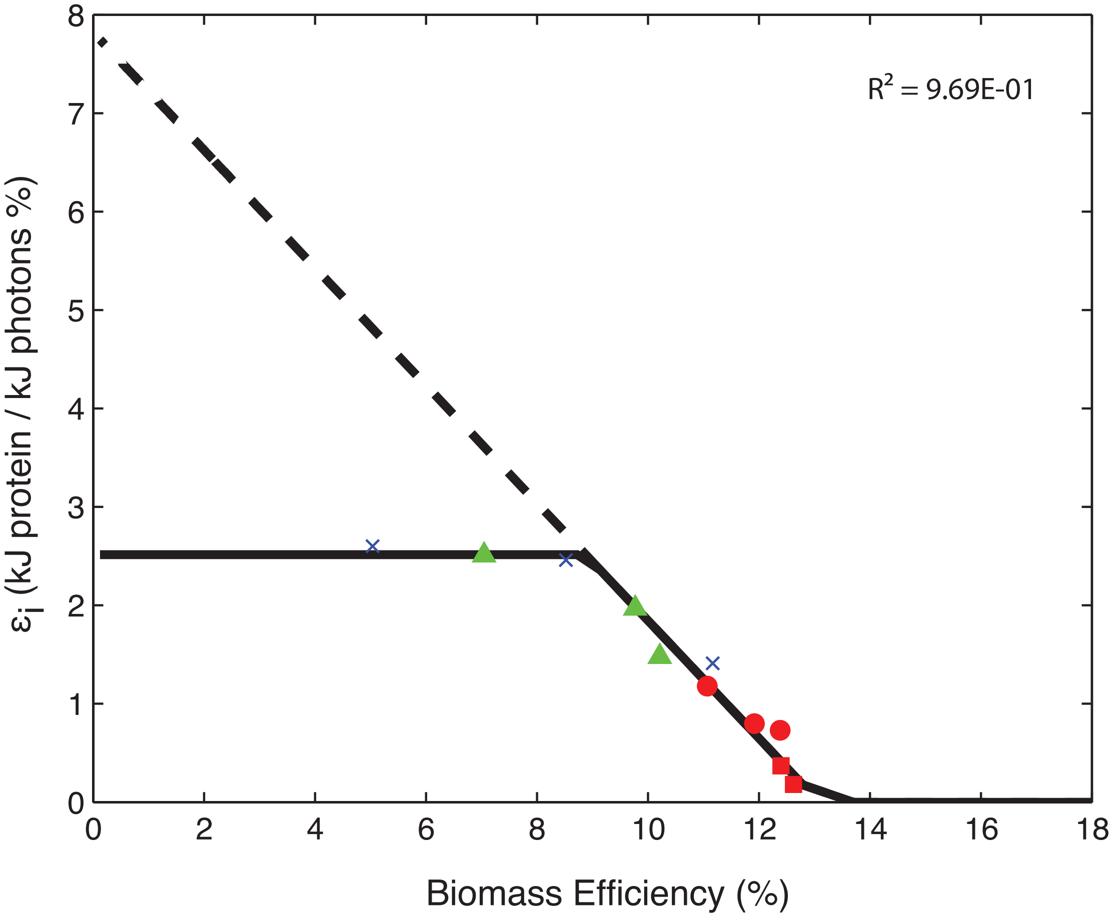


Figure 5.9: Carbon Metabolism Isocline. Photosynthetic efficiency of YFP production in PCC7002 across a range of promoters. Strains expressing high-levels of protein exhibit decreasing efficiencies of biomass production. PCC7002 photosynthetic capacity is predicted by FBA model fit for the pAQ1 promoter scan. The y-intercept describes the strain's protein production capacity when carbon metabolism is fully shifted away from growth, with the translation rate upper bound at 100% capacity.

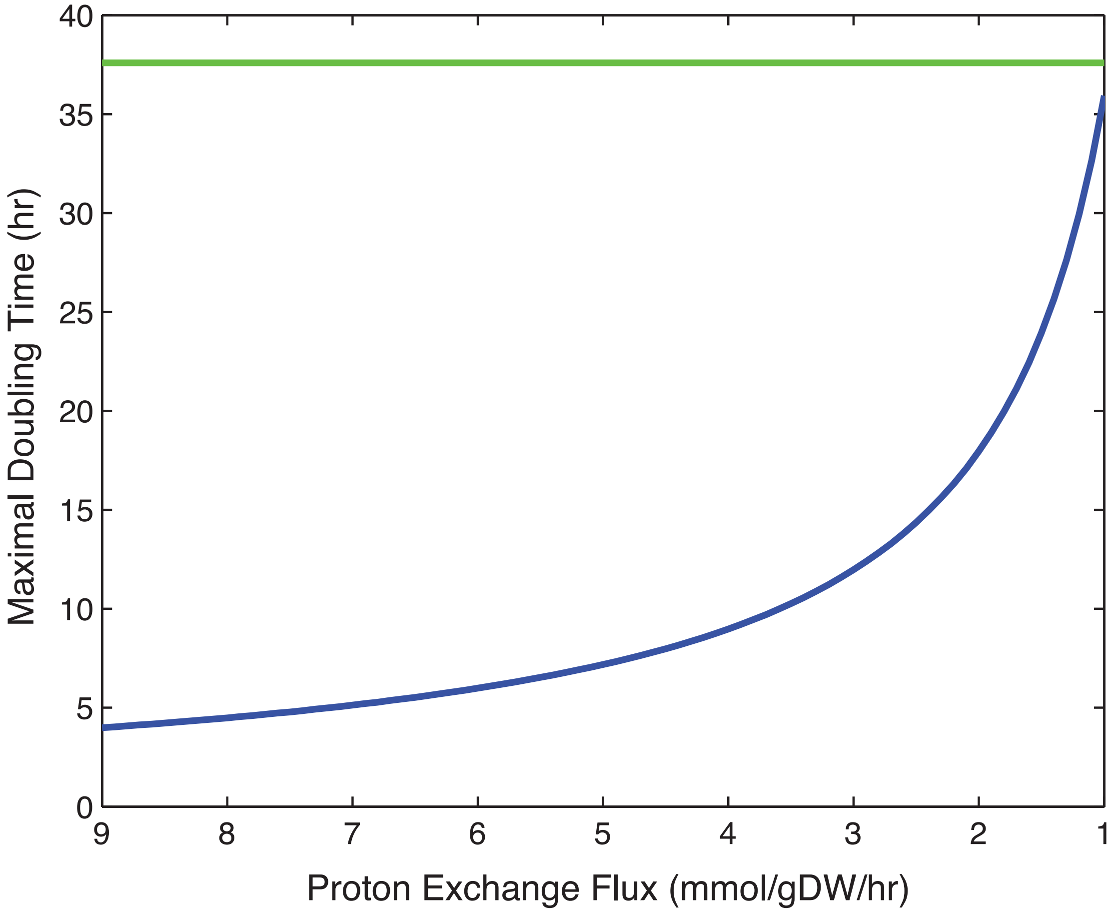


Figure 5.10: High Density Growth Capacity. Strain growth capacity depends on culture-cell density and media depth. Differences in strain maximal growth rate at low culture cell-densities (blue line) collapse to a single growth rate as the culture reaches a high cell-density (green line).

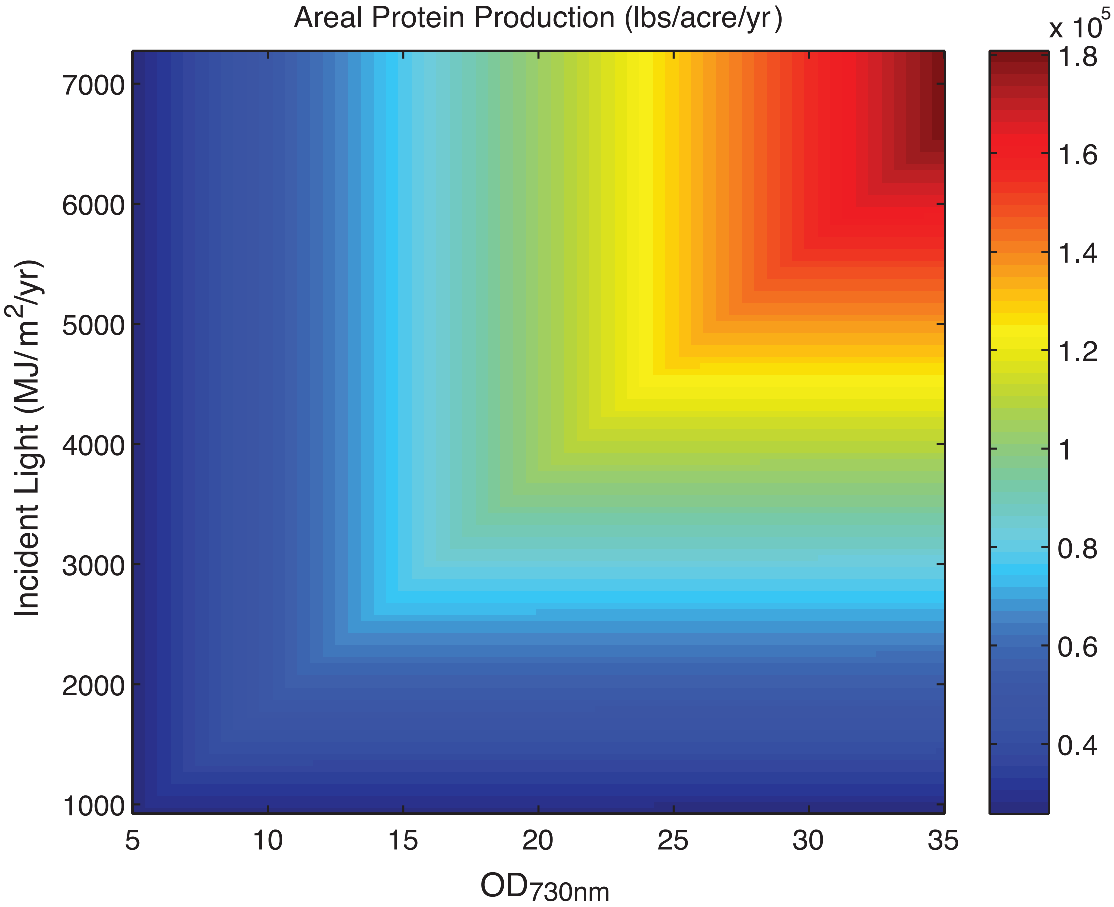


Figure 5.11: Areal Protein Production. Protein production (lbs/acre/yr) depends on culture-cell density and light intensity. At high cell-densities the strain translation capacity is no longer limiting, with areal yield scaling with incident light intensity.

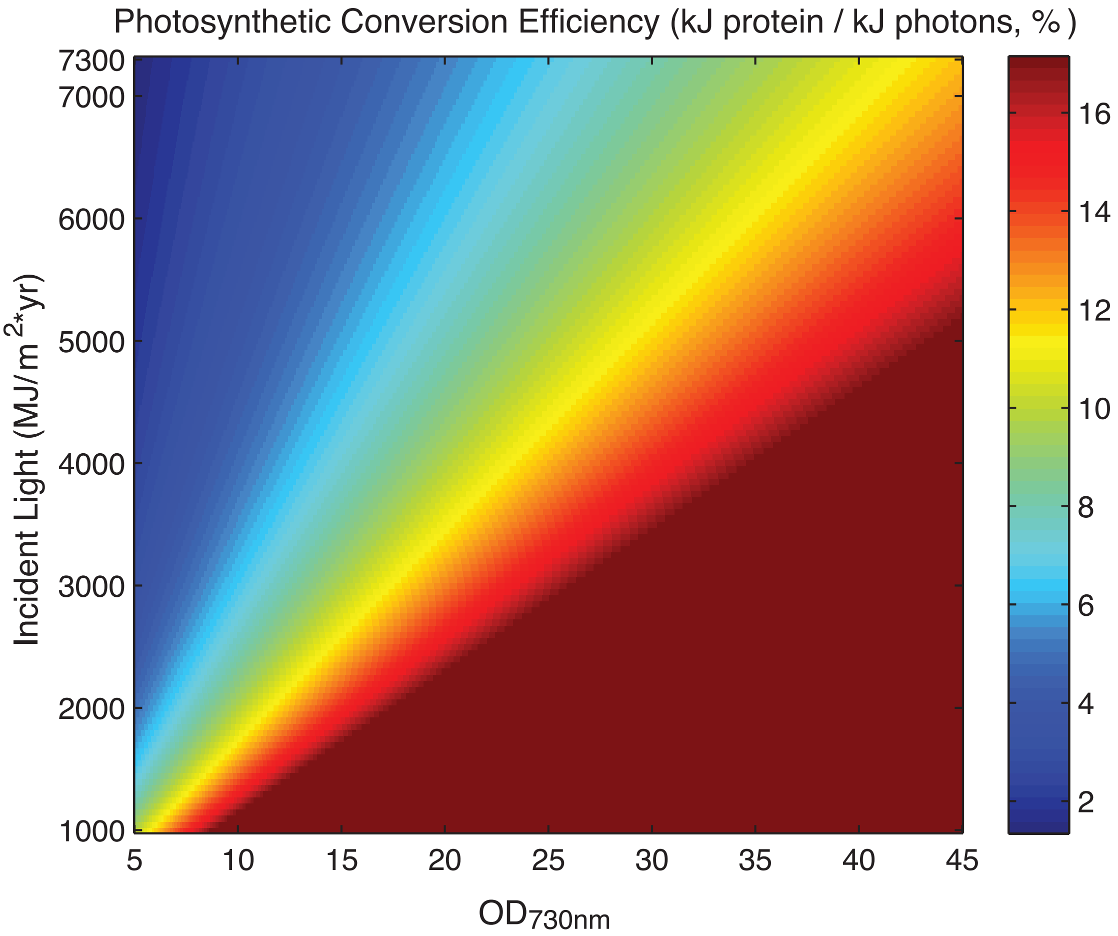


Figure 5.12: High-light Protein Production Efficiency. Model of protein photosynthetic efficiency (kJ protein / kJ incident PAR radiation) at high light intensities, approaching outdoor PFD conditions. The required cell-density for optimal photon utilization scales linearly with PAR flux.

Experimental Methods

High Throughput Cultivation

To develop a platform for high-throughput cultivation of cyanobacteria (Fig. 5.4a) for testing multiples of strains, expression constructs, culture conditions, or inducer levels, we retrofitted a Vertiga Micro-Expression Shaker (Glas-col; Jerra Haute, IN) to supply 2% CO_2 to the headspace of the shaker with the use of two mass flow controllers (Alicat Scientific; Tuscon, AZ) to support 4 standard L/min (3920 standard ml/min air and 80 standard ml/min pure CO_2). This gas stream was humidified with a water bottle in-line before the shaker to reduce evaporation from the culture. Mixing studies revealed 800rpm at an orbital offset of 0.070" worked best to generate vigorous mixing with a square-well, conical bottom, 2ml 96 Deep-well block (DWB) with 1ml of culture volume.

A manifold to hold a 96DWB and an LED panel used for camera systems with associated heat dissipation was designed internally using SolidWorks and manufactured by stereolithography (SLA) 3-D printing technology at RealizeInc (Noblesville, IN). The LED panels (Litepanels Micropro) had a 12×8 grid design in the same footprint as the 96DWB, allowing one LED to supply the light to every well. A UV absorbent film was overlaid above the LED panel and an optical diffusing film was overlaid over that. This allowed for $\sim$90% light transmittance but improved the light uniformity over the surface of the panel by $>$400%. The supplying AC power was tuned to achieve an intensity of 100 ± 10 $\mu E/m^2/s$ over the surface of the panel. Transmittance through the bottom of the 96DWB to the culture was $>$95%. PAR was measured using a LI-COR (Lincoln, NB) LI-250A light meter.

Strain Selection and Growth Rate Testing

Strains were purchased from The Pasteur Culture Collection of Cyanobacteria (PCC) at the Institut Pasteur (Paris, France) or the American Type Culture Collection

(ATCC; Manassas, VA). All strains purchased were isolated from saltwater or brackish water environments. The base medium for all strains was A+. This medium contained 18 g/l sodium chloride, 0.6 g/l potassium chloride, 1 g/l sodium nitrate, 5 g/l magnesium sulfate heptahydrate, 50 mg/l monobasic potassium phosphate, 266.4 mg/l calcium chloride, 30 mg/l tetrasodium EDTA dehydrate, 3.89 mg/l iron (III) chloride hexahydrate, 8.20 mM Tris pH 8.2, 4 μg/l vitamin B12, and 1 ml/l P1 metals solution (1000x P1 metals: 34.26 g/l H_3BO_3, 4.32 g/l MnC_{12}-$4H_2O$, 315 mg/l ZnCl, 30 mg/l MoO_3, 3mg/l $CuSO_4$-$5H_2O$, 12.15 mg/l CoC_{12}-$6H_2O$). Strains were cultivated on solid A+ medium (A+ medium with 15 g/l agar) at 30°C, ambient CO_2 concentration, and a light intensity of $\sim$15 μE/m^2/s with fluorescent lights. Single colonies were isolated and inoculated into 2 ml liquid A+ medium (in 22x175 mm^2 (50 ml) culture tubes with foam stoppers) and grown to $OD_{730nm} \approx$ 2-3 at 35°C, 150rpm, 2% CO_2 (v/v), and $\sim$80 μE/m^2/s with fluorescent lights. Cell-density was measured spectrophotometrically (using a BioTek MX II microplate reader) at 730nm. Samples were diluted in sterile, A+ medium to ensure the raw absorbance value was below 0.5. All strains were then inoculated to OD_{730nm} = 0.1 in 10 ml A+ medium (in a 125 ml glass, baffled shake flask) and grown for 120hr at 35°C, 150rpm, 2% CO_2 (v/v), and $\sim$80 μE/m^2/s with fluorescent lights. Samples were taken intermittently to measure cell-density. The OD_{730nm} vs. time data was then fit to a logistic population model,

$$X(t) = \frac{X_0 e^{\mu t}}{1 - \frac{X_0}{X_{sat}}(1 - e^{\mu t})} \tag{5.3}$$

This equation describes the cell-density, X, as a function of time, t, with the following constant parameters (X_0, the initial cell density, X_{sat}, the stationary phase cell-density, and μ, the maximal specific growth rate). A non-linear least-squares regression of the data with the logistic population model was undertaken in MATLAB (using the ”nlinfit” function) to determine the X_0, X_{sat}, and μ parameters. 95% confidence intervals were assessed for cell-density estimates, and the estimated parameters were examined for accuracy. The maximal specific growth rate, μ, an indication of the inherent photosynthetic capacity of the organism in question was then compared across strains. Strain doubling time was calculated from maximal specific growth rate as log(2)/μ.

Media Optimization

Media optimization was undertaken by varying the three major inorganic components contributing to cell-viability: Nitrogen, Phosphorus, and Iron. Sodium nitrate was varied between 1 and 20 g/l, monobasic potassium phosphate was varied between 50 and 250 mg/l, and iron (III) chloride hexahydrate was varied between 10 and 250 mg/l. The optimal concentrations of each component, as shown to optimize final cell-density (OD_{730nm}) at 168 hr (7 days), were 4 g/l sodium nitrate, 200 mg/l monobasic potassium phosphate, and 20 mg/l iron (III) chloride hexahydrate. These three component changes over A+ medium was named PB1.0. The effect of antifoam (BASF Industrol 204) concentration (10^{0}, 10^{-01}, 10^{-02}, 10^{-03}, 10^{-04}, 10^{-05} %) was then analyzed on bubble formation and cell-density over the course of a standard growth experiment. 0.001% v/v was determined to be the lowest concentration of antifoam shown to decrease bubble formation without effecting cell-density. This component change over PB1.0 medium was named PB1.1 and used for the protein production experiments described henceforth. Medium optimization enabled the long term cultivation of *Synechococcus sp. PCC 7002* without the need for medium replenishment (Fig. 5.13). This allowed the culture to reach higher cell-densities, and potentially improved production efficiencies (as described by the FBA above).

Salinity Tolerance

To determine the effect of varying salinities of the medium on the growth and robustness of PCC7002, we cultivated it in our 96DWB high-throughput platform with

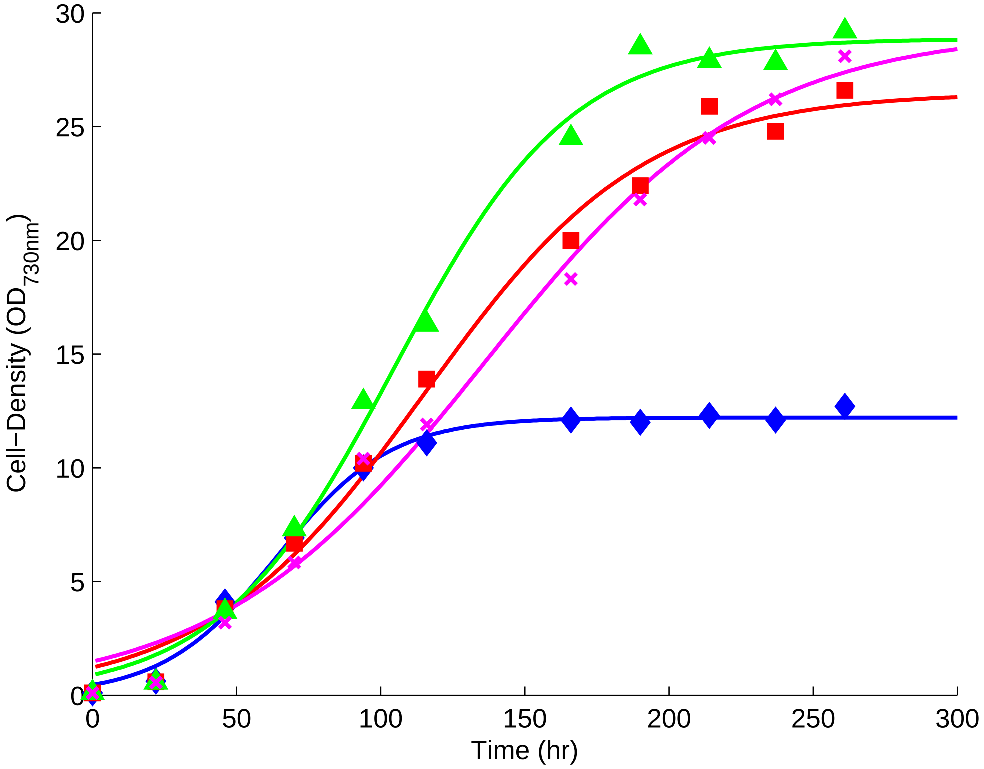


Figure 5.13: Long term PCC7002 cultivation. Growth Curves. PCC7002 cultured in A+ medium exhibits a plateau in growth (blue diamonds), which can be overcome by daily medium replenishment (red squares). Optimized PB1.1 medium permits extended growth profiles beyond 120hr both under conditions when the medium is replenished daily (purple x's) and without medium replenishment (green triangles).

PB1.1 medium across a concentration gradient of sodium chloride (Fig. 5.14). The robust growth under saline conditions confirms the strain as a viable photosynthetic nutrient catalyst without the need for freshwater. This feature might enable a protein production capacity in regions with typically poor agricultural productivities due to arid conditions.

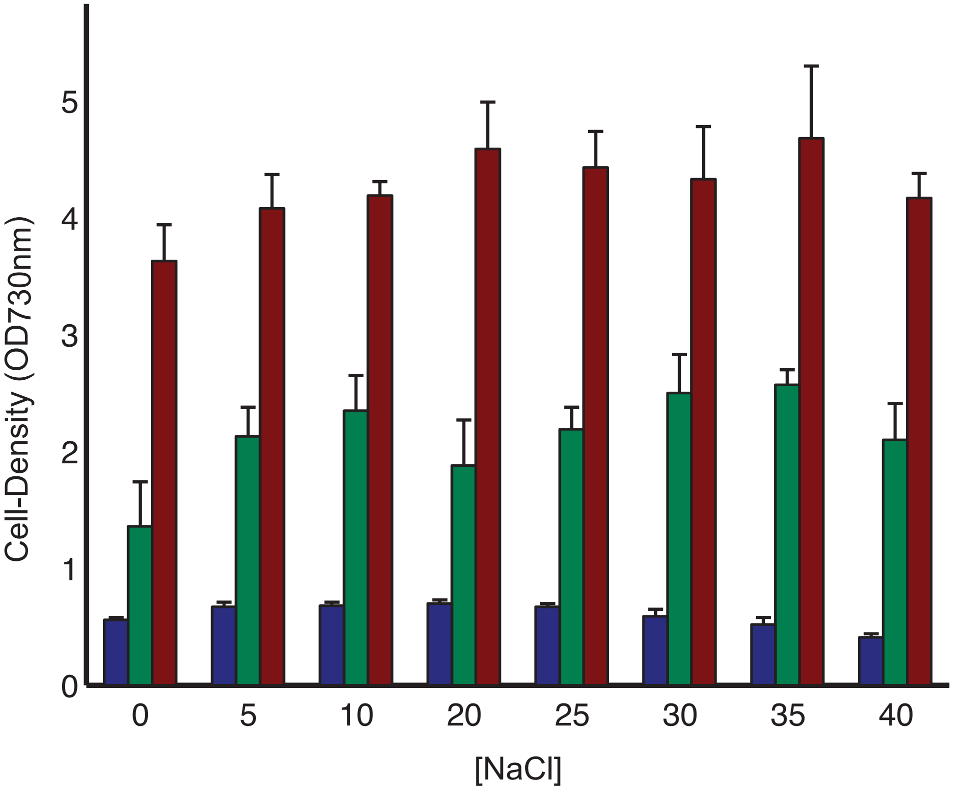


Figure 5.14: PCC7002 Salt Tolerance. Cell-density of PCC7002 vs. sodium chloride concentration. The cell-density at 24hr (blue bars), 65hr (green bars), and 90hr (red bars) demonstrate cell viability across a range of medium salinity. Error bars represent ± one standard deviation from four independent wells.

Promoter and Strain Construction

The expression constructs tested in Fig. 5.2c include the P*psbaII* promoter that drives expression of P*psbaII* gene in *Synechocystis elongatus PCC 7942*, the P*cpc** promoter that drives expression of the cpc operon in *Thermosynechococcus sp. BP-1*, the

P*cpc* promoter that drives expression of the cpc operon in *Synechocystis sp. PCC 6803*, the P*cro* promoter that drives expression of the cro gene in *Enterobacteria phage* λ, the P*trc* promoter that is hybrid promoter (ptrp-35+17bp+pLac-10), derived from *Escherichia coli*, the P*suf* promoter that drives expression of the sufB gene, and the P*rbc* promoter that drives expression of the rbcL gene in *Synechocystis elongatus PCC 7942*. All the promoters are amplified from their respective genome, inserting an EcoRI restriction site GAATTC at the 5' end and NcoI restriction site CCATGG at the 3' end. The PCR products, along with a pAQ1-*yfp* backbone plasmid for PCC7002, were then digested with EcoRI and NcoI restriction endonucleases. Each PCR product was ligated with the backbone plasmid by T4 ligase, inserting the promoter sequence upstream of the *yfp* gene (Fig. 5.15). Constructs were then transformed into Neb Turbo E.coli cells and sequence verified. All construct sequences in this work are included in Supplementary Dataset S2.

For cyanbacterial transformation of plasmids, 1 μg of each construct was mixed with 30 μl of *Synechococcus sp. PCC 7002* that was grown in A+ medium to OD_{730nm} $\approx$ 1.5. The mix was loaded on a grid agar plate with A+ medium supplemented with spectinomycin 100 μg/mL. The transformants were incubated at 35°C, 10 $\mu E/m^2/s$ and atmospheric CO_2 conditions. After 2 days, the growing bacteria were re-streaked on a similar grid agar plate (A+ spectinomycin 100 μg/ml). Plasmid segregation was monitored by colony PCR with primers flanking the integration regions.

Two highly expressing promoters on the pAQ1 plasmid system, (P*cpc** and P*cro*), were selected for construction of a cumate inducible promoter, implementing the *cymR* and *cymR* operator (*cymO*) regulatory elements from the *cym* operon of *Pseudomonas putida F1* (Choi et al., 2010). Sequences containing P*cro-cymO* with P*km-cymR* in cis on the reverse strand as well as P*cpc*-cymO* were synthesized. In *cymO* sequences, the *cumO* site was added at the +1 transcriptional site of the respective promoters. For construction of P*km-cymR*-P*cro-cymO-yfp*, a pAQ3-*yfp* backbone and the synthesized promoter sequence were digested using SacII and NcoI restriction endonucleases. The digested fragments were ligated togther by T4 ligase to construct the final

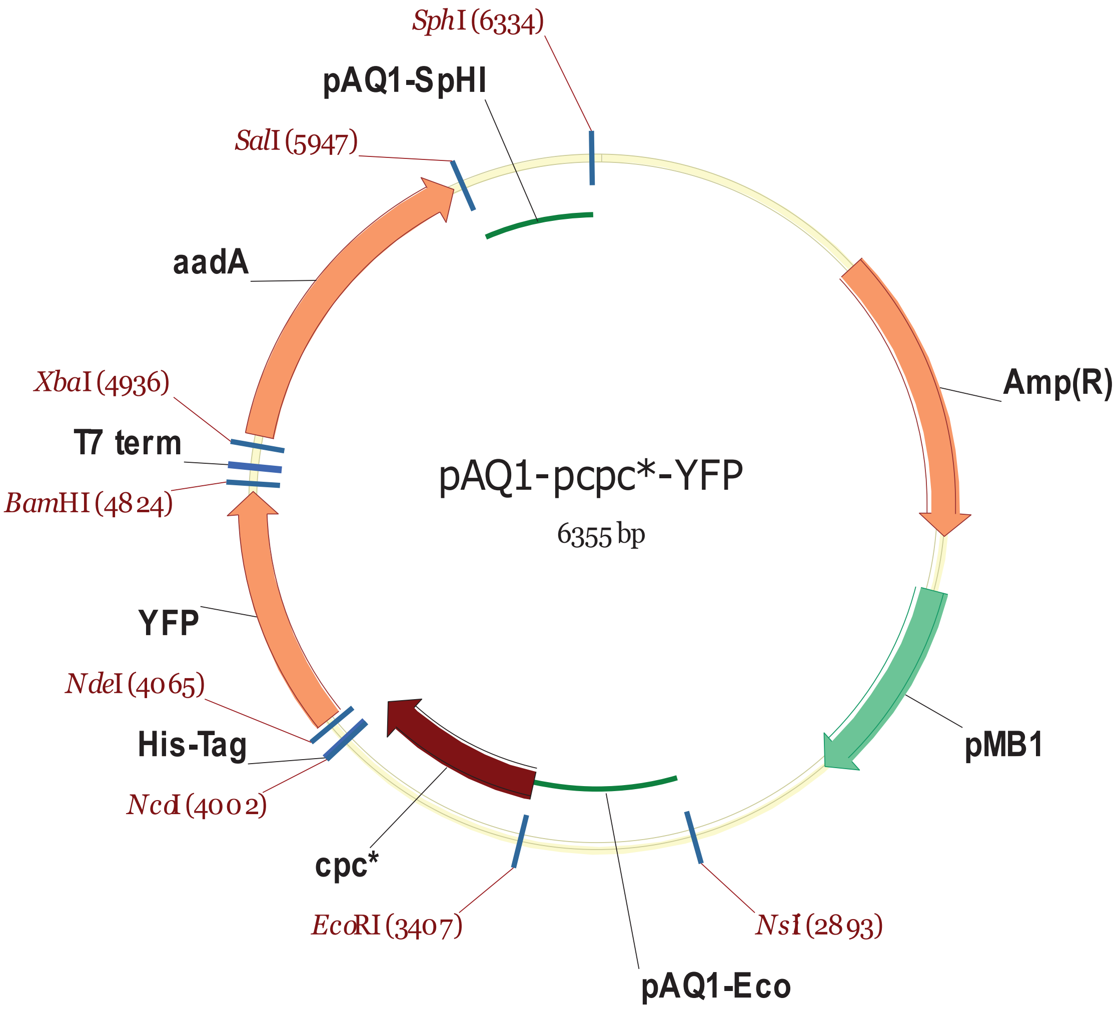


Figure 5.15: pAQ1 P*cpc-*yfp*.** pAQ1 plasmid map. The promoter upstream of *yfp* was substituted to screen expression strength.

plasmid. The P*km-cymR*-P*cpc*-cymO-yfp* was constructed by amplifying P*cpc*-cymO* with a 5' flanking MluI site and 3' NcoI site. The PCR product and P*km-cymR*-P*cro-cymO-yfp* plasmid were digested using MluI and NcoI restriction endonucleases. The digested fragments were ligated by T4 ligase, replacing P*cro-cymO* with P*cpc*-cymO*, to construct the final plasmid (Fig. 5.16a).

Of the constructs, only P*cro* exhibited inducible behavior. Testing whether P*cpc** could be induced more strongly with a lower concentration of repressor, P*cpc** was also constructed with the CymR repressor expressed in trans via chromosomal integration.

For construction of pAQ3-P*cpc*-cymO-yfp* with *cymR* in trans chromosomally, the P*cpc*-cumO* promoter sequence was amplified with a 5' flanking SacII and 3' NcoI site. The PCR product and pAQ3-*yfp* backbone were digested using SacII and NcoI restriction endonucleases. The digested fragments were ligated by T4 ligase to construct the final plasmid (Fig. 5.16b). The P*km-cymR* component for chromosomal integration was amplified from the P*km-cymR*-P*cro-cymO-yfp* with a 5' flanking EcoRI site and 3' flanking BamHI site. The PCR product and pA1133int backbone were digested using EcoRI and BamHI restriction endonucleases. The digested fragments were ligated by T4 ligase to construct the chromosomal construct. As this integrant would be co-selected with the pAQ3-P*cpc*-cymO-yfp* construct, the *aad1* spectinomycin resistance gene was replaced by the *aacC1* gentomycin resistance gene. This was done using *aacC1* amplification with 5' and 3' flanking AarI sites, for digestion and ligation into the final plasmid. Transformation was done as described above, with 100 μg/ml spectinomycin and 50 μg/ml gentomycin.

While this exhibited inducibility, the range of induction remained more narrow than P*cro/cymOP* (Fig. 3). To explore the variability in promoter strength being a product of the expression vector, we tested expression of constitutive P*cpc** and P*cpc* promoters on pAQ3 and chromosomal cassettes (Fig. 5.17). These were constructed as described above, amplifying the P*cpc** and P*cpc* promoters with restriction sites for insertion into the pAQ3, pA1133int, and pA1580 backbones. While P*cpc** was a stronger promoter than P*cpc* in pAQ1, their relative expression levels reversed in pAQ3

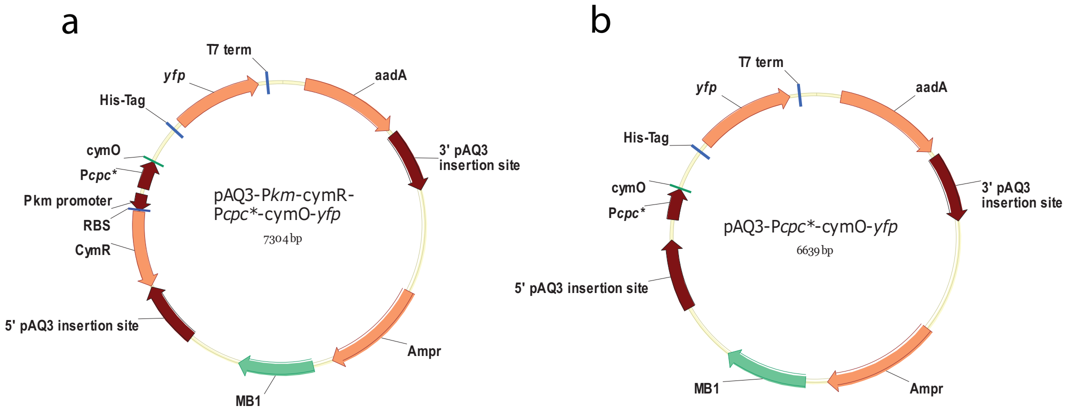


Figure 5.16: pAQ3 P*cpc*-cymO-yfp*. Cumate inducible plasmid map. (a) The P*cpc**-*cymO* plasmid with P*km-cymR* in cis on the reverse strand. (b) The P*cpc**-*cymO* plasmid with P*km-cymR* in trans.

(Fig. 5.18). This demonstrates the importance of characterizing promoters within the intended vector, as their activity exhibits specificity to the expression construct in *Synechococcus sp. PCC 7002*.

Promoter characterization was performed as described in Flask Experiments below, using 50μM cumate for expression from inducible promoters. Promoter activity was measured as fluorescence per OD$_{730nm}$, with samples in duplicate.

Flask Experiments

Shake-flask expression studies were undertaken in a similar manner as described above in the ”Strain Selection and Growth Rate Testing” section. The only difference being all medium contained 100 mg/l spectinomycin to maintain selection of the integrated plasmid, and the base medium used was PB1.1 and not A+. All strains were then inoculated to OD$_{730nm}$ = 0.2 in 10 ml PB1.1 medium (in a 125ml glass, baffled shake flask) and grown for 10 days at 35°C, 150rpm, 2% CO$_2$ (v/v), and $\sim$80 μE/m^2/s with fluorescent lights. Samples were taken intermittently to measure cell-density and fluorescence measurements to quantify yellow fluorescent protein (YFP).

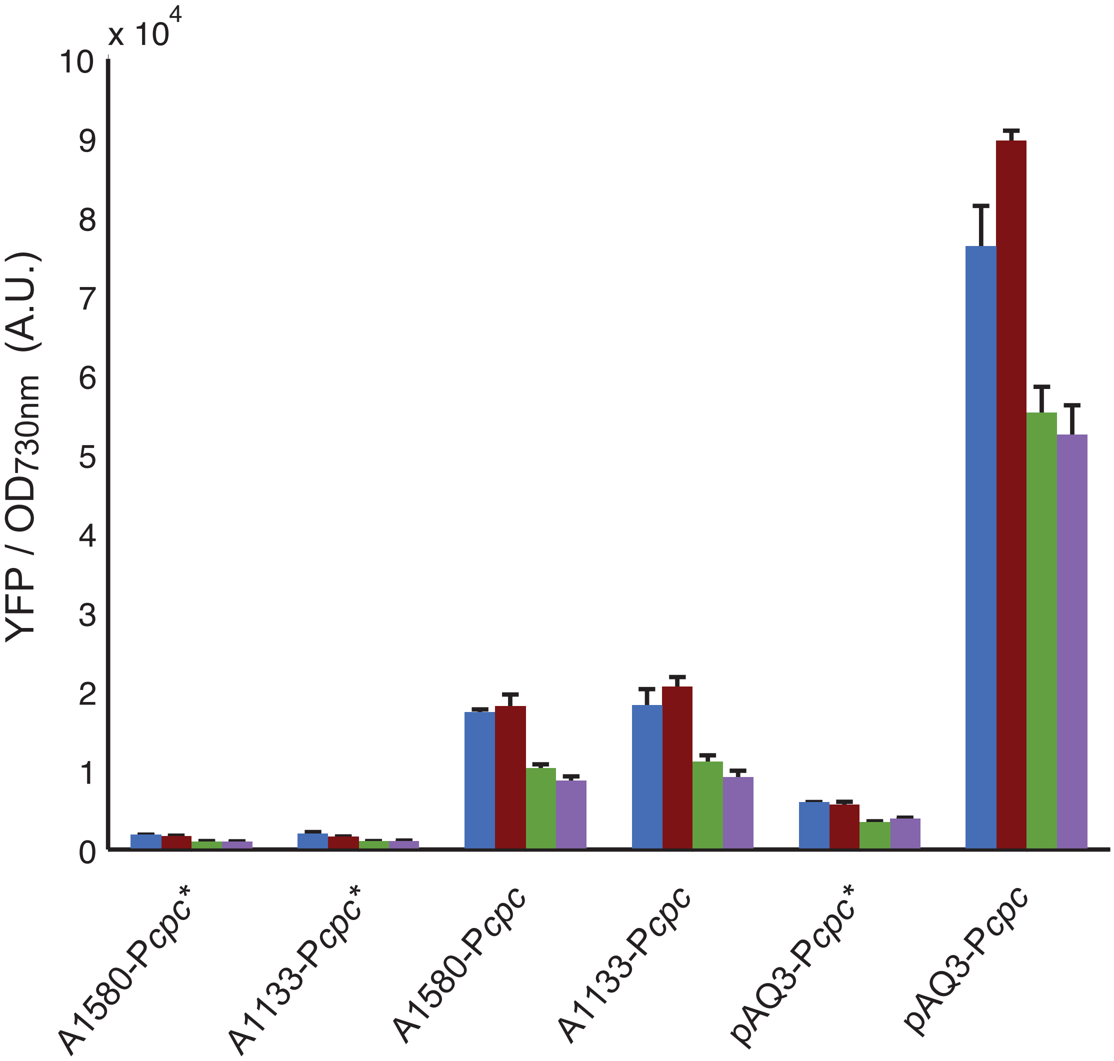


Figure 5.17: P*cpc*-cymO-yfp* and P*cpc-cymO-yfp* expression. Promoter characterization across expression vectors.

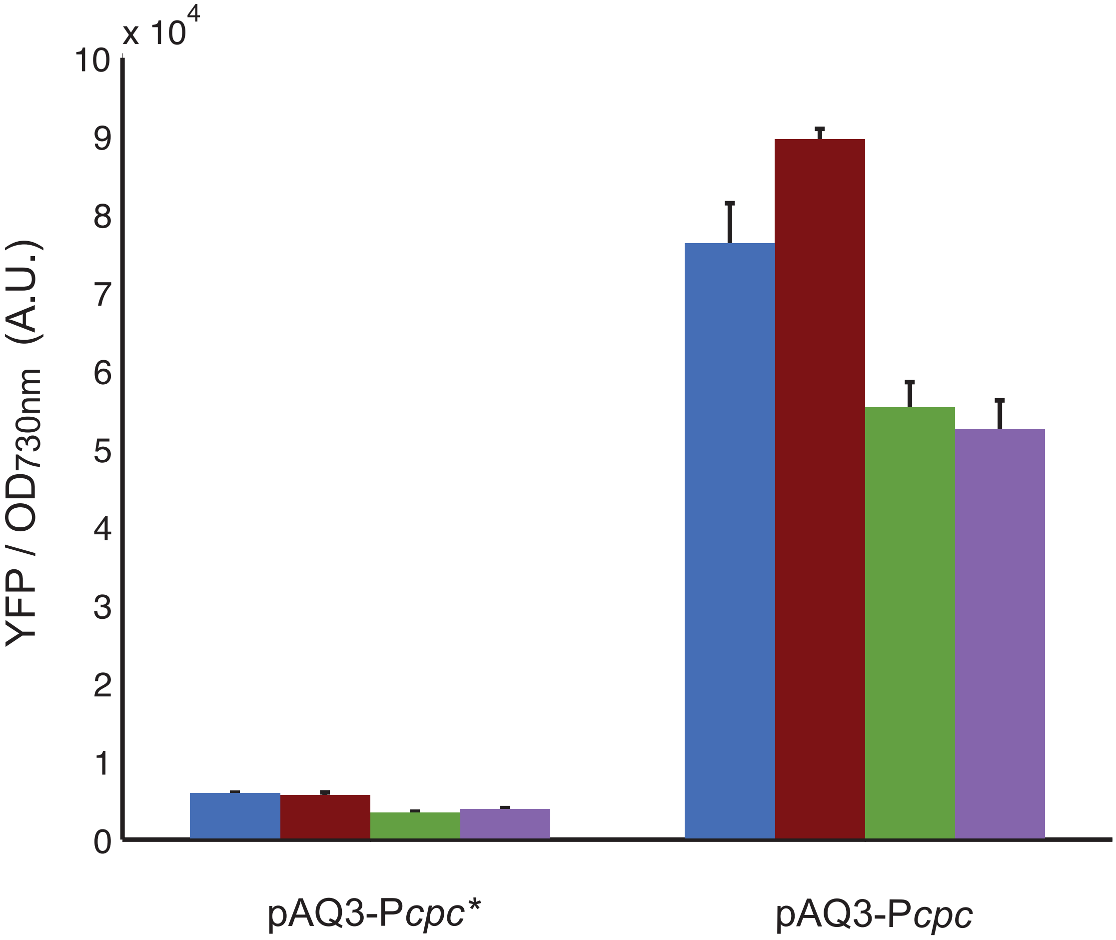


Figure 5.18: Comparison of P*cpc*-cymO-yfp* and P*cpc-cymO-yfp* expression strength in pAQ3 constructs.

YFP quantification was undertaken on a BioTek Synergy Mx plate reader. The optimal excitation and emission spectra were determined to be 505nm and 550nm respectively, and the detector sensitivity was set at 75. Recombinant YFP standard (MBL International Corporation) was diluted in PB1.1 medium at concentrations of 5.0, 2.5, 1.25, 1.0, 0.5, 0.25, 0.125, and 0.0625 mg/l. A fixed concentration of 0.5 mg/L YFP was then spiked into varying density cultures to determine its effect on quenching the fluorescence emission (Fig. 5.19). All samples for YFP quantification were diluted to OD$_{730nm}$ = 0.05 for measurement on the plate reader.

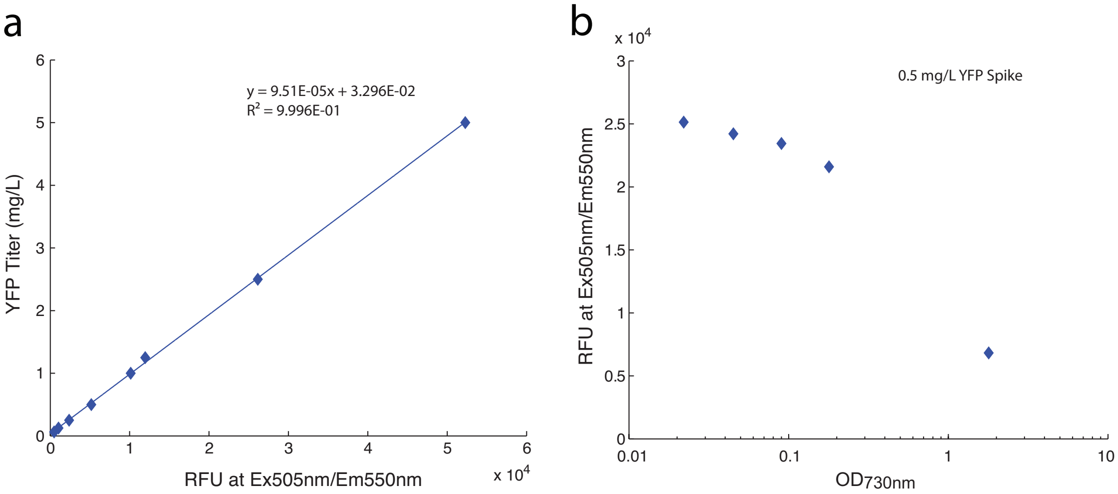


Figure 5.19: YFP Standard Curve. (a) RFU emission linear fit to a range of recombinant YFP concentrations. (b) A quenching experiment at 0.5 mg/L YFP to determine cell density dependence on fluorescence emission shows an optimal OD$_{730nm}$ range ¡ 0.1.

Measurements of photosynthetic efficiency were taken within the linear growth phase for flask experiments, each in triplicate. The rate of protein accumulation was averaged over 24 hours, normalized to the culture's OD$_{7}30nm$ for calculation of efficiency on a per-cell basis. Likewise, light intensity at each time point was normalized to the OD$_{7}30nm$, and averaged over 24 hours to calculate photon-flux per cell. Protein energy densities were taken as calculated for the model (23.32 kJ/g YFP and 23.645 kJ/g Ovalbumin). Finally, areal production per m^2 was calculated with 3184.7 mL/m^2 flask dimensions. A PAR photon energy of 2.26E+05 J/E was used as in the model Robertson et al. (2011). Calculation of biomass efficiency was done similarly, using the conversion from OD to dry cell mass of .28 g/L*OD$_{730nm}$ for PCC7002. A biomass energy density for PCC7002 of 26.5 kJ/g was used as in the model, in the upper range of the reported literature for algae 20-26.9 kJ/g (Weyer et al., 2010).

Promoters characterized in pAQ1 expression constructs are shown in Fig. 5.2. Cumate inducible promoters were tested in lower copy pAQ3 expression constructs. High protein production levels from pAQ1 slows the growth rate within the PCC7002 standard linear growth range (Fig. 5.20), as the cell is limited by its photosynthetic capacity for carbon metabolism (shown in Fig. 5.2). Promoters that more strongly express protein likewise reach a lower final OD$_{730nm}$ (Fig. 5.20b). To determine the direct effect of cell-density on protein photosynthetic efficiency, we instead built expression constructs on a lower-copy plasmid, pAQ3, to relieve the dependence between protein and biomass efficiency (Fig. 5.21). Any variability in growth rate between the ODs of induction did not correlate with the protein production efficiency (Fig. 5.21).

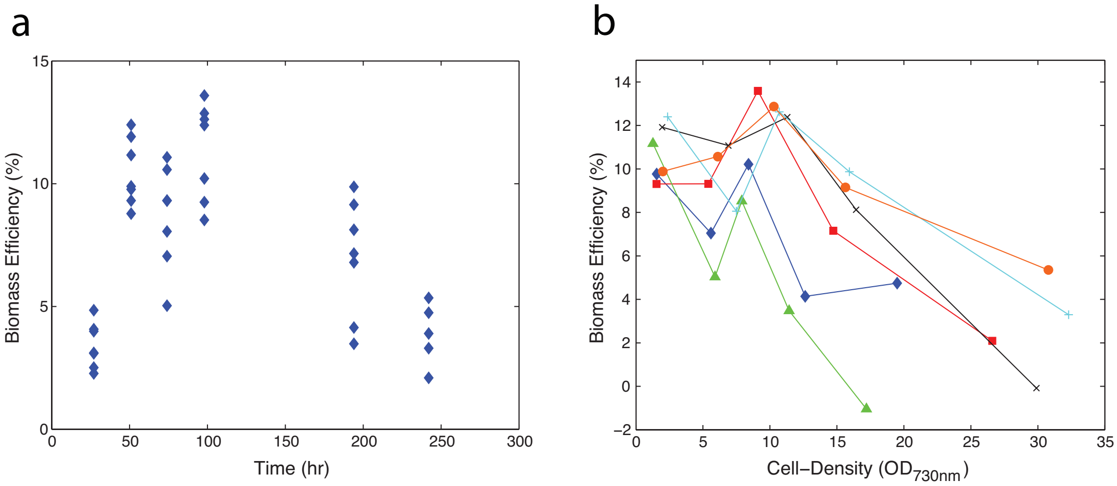


Figure 5.20: pAQ1 Biomass Efficiency. Photosynthetic efficiency of biomass production for PCC7002 producing protein from pAQ1. Strains expressing high-levels of protein exhibit decreasing efficiencies of biomass production over time. (a) Growth rate shifts over time as culture approaches carrying capacity. (b) Stronger promoters have lower growth rates (i.e. P*cpc**,green triangles vs. P*rbc*, red squares) and reach lower final ODs.

For cumate inducible constructs, inoculum from the P*cro/cymOP* PCC7002 strain was prepared as described previously in the ”Strain Selection and Growth Rate Testing” section. Here, PB1.1 was used as the production medium, and all flask cultures were inoculated to OD_{730nm} = 0.15. Cultures were induced with various concentrations of cumate (prepared at 50mM in molecular biology grade ethanol). Cultures were induced at OD_{730nm} of 1, 1.5, 2, and 6. Thereafter, samples were taken every 12 hours to monitor the cell-density, YFP titer, and samples for mRNA quantification were taken.

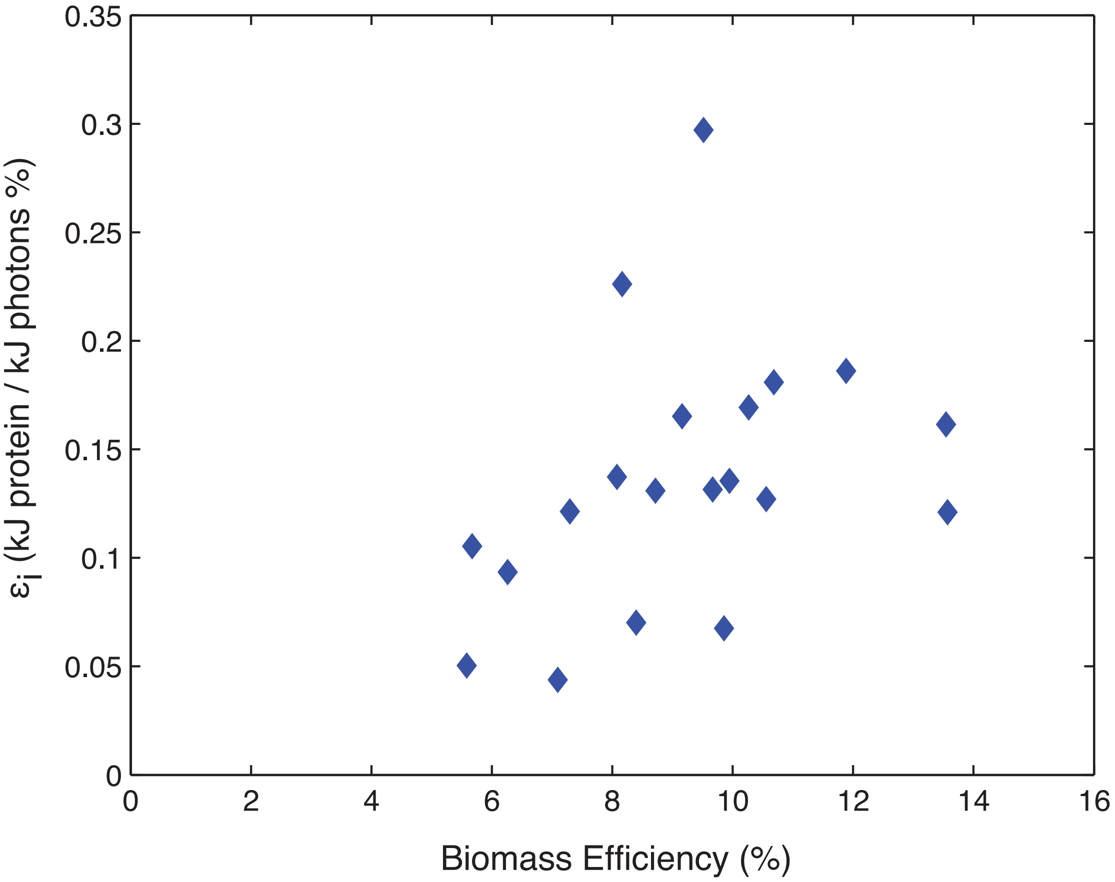


Figure 5.21: pAQ3 Carbon Metabolism. Photosynthetic efficiency of biomass and protein production for PCC7002. YFP expression levels from promoters integrated onto pAQ3 plasmids do not exhibit a growth dependence.

YFP mRNA was quantified across the experimental OD range to ensure consistent promoter activity (at induction levels of both 10 μM and 100 μM). This allows a comparison of the strain's photosynthetic efficiency upon induction at different cell-densities. Cell samples for mRNA extraction were taken upon the final YFP measurement for photosynthetic efficiency.

The equivalent of 1 ml of OD 2.5 of cells were precipitated by centrifuging at

1500g for 5 minutes and frozen at -20°C overnight. Cells were resuspended in 4 ml deionized water, thawed, lysed by adding 2 ml of lysis buffer (16 mM EDTA, 2% SDS), and heated to 95°C for 5 minutes. The total RNA was extracted by adding 4 ml phenol (Sigma) at 65°C and incubated at 65°C for 10 minutes while shaking every 5 minutes. The solution was centrifuged at 3000g for 10 minutes and the top aqueous layer was transferred to new tubes. 4 ml phenol was added and vortexed for a second round of phenol extraction followed by centrifugation in the same conditions. The top aqueous layer was transferred to new tubes and 4 ml phenol:chloroform:isoamyl alcohol = 25:24:1 (Sigma) was added and vortexed followed by centrifugation. This step was repeated once. The final aqueous layer was transferred to a new tube and the RNA was precipitated by adding 2 ml isopropanol and frozen at -20°C overnight. Following centrifugation at 20,000g for 20 minutes at 4°C, the precipitant was washed once with 70% EtOH and centrifuged with the same conditions. The supernatant was poured out and the precipitant air dried and resuspended in 100 μl RNase-free water. Residual DNA was removed from the RNA by DNase I (Promega) digestion following the manufacturer's protocol at 37°C for 1 hour. The reaction was stopped by adding 2 μl of stop solution and inactivated at 65°C for 10 minutes.

First-strand cDNA was synthesized using the New England BioLabs ProtoScript II First Strand cDNA Synthesis Kit. Standard protocols provided by the manufacturer were followed, using 2 μl of 60 μM random hexamer primers reacted with 200-600 ng of total RNA. The reaction was incubated at room temperature for 5 minutes before incubating at 42°C for 1 hour.

For real-time PCR quantification, cDNA was diluted by 2X, 20X, and 200X and measured in triplicate. qRT-PCR were performed with the SYBR Green master mix (Life Technologies SYBR Select Master Mix) by combining the cDNA templates with 200 nM primers that amplifies either yfp or a housekeeping gene, 16S rRNA or RuBisCo. Standard reaction conditions suggested by the manufacturer were followed. The sequences of the primers for each gene are listed in Table 5.2

Differences in light-use efficiency for protein production is not due to increased

Table 5.2: RT-PCR transcript quantification primers for YFP, 16S, and Rubisco.

Name	Target	Sequence
YFP-FW	YFP	AGAAGAACGGCATCAAGGTG
YFP-RV	YFP	GAACTCCAGCAGGACCATGT
16S rRNA-FW	16S rRNA	AGAGATGCGAGAGTGCCTTC
16S rRNA-RV	16S rRNA	AAGGGGCATGATGACTTGAC
RuBisCo-1-FW	RuBisCo	CACTTGCGAAGAAATGCTCA
RuBisCo-1-RV	RuBisCo	CGGAAGTGAATACCGTGGTT
RuBisCo-2-FW	RuBisCo	ACCCTCGGTTTCGTAGACCT
RuBisCo-2-RV	RuBisCo	AGGGTACCACCACCAAACTG

promoter activity at higher cell-densities (Fig. 5.22). Photosynthetic efficiency improves as cell-density increases, as light is better distributed between cells to match their translation capacity.

Nutrient Production

Screening for Ovalbumin Production

In screening for Ovalbumin expression, we tested a range of expression constructs. As we found above while characterizing the cumate inducible promoters, the copy number of the plasmid affects the performance of the promoter-protein pairing. We tested Ovalbumin production on high copy (~40) pAQ1, medium copy (~20-30) pAQ3, and low copy (~5) pAQ6 plasmids, as well as in multiple integration sites within the chromosome (Table 5.3). Construction of the plasmids was done as described above, using restriction digestion to replace yfp with the codon-oprimized gene for Ovalbumin, with N-terminal 8x His, in a number of expression constructs. 8xHis-ovalbumin with flanking restriction sites was amplified from the synthesized sequence.

Ovalbumin production studies were undertaken as described in the ”Flask Experiments” section, with samples in duplicate. Ovalbumin produced recombinantly was quantified in lysate by chip electrophoresis. Samples were prepared by adding 2 μl of

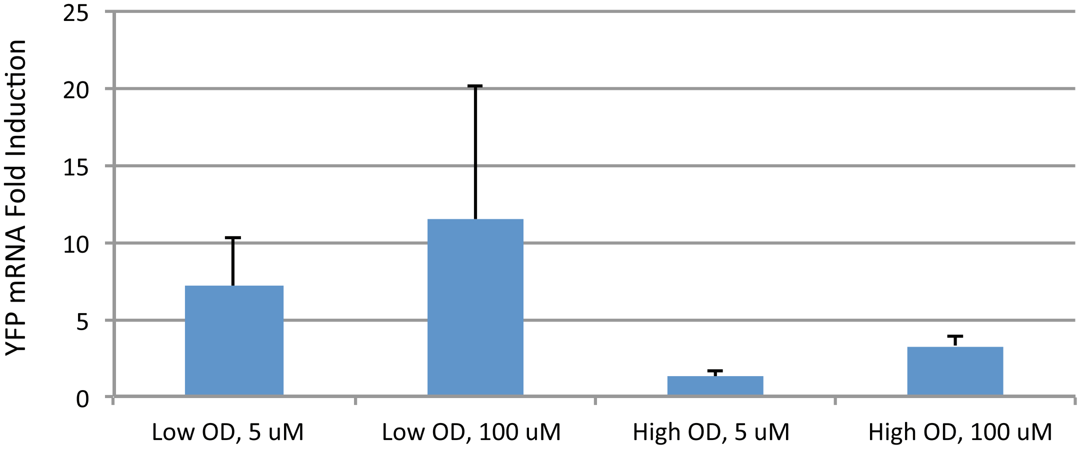


Figure 5.22: pAQ3 mRNA Levels. YFP mRNA is quantified relative to RuBisCo (primer set 2), averaged in triplicate. Absolute levels of RuBisCo remained consistent across the samples. YFP fold induction is the RuBisCo-normalized YFP mRNA levels over uninduced control.

lysate, which was normalized to OD_{730nm} 1.5, to 7 μl Protein Express sample buffer, heating at 95°C for 5 minutes, and then adding 35 μl of water. Analysis was completed using HT Low MW Protein Express LabChip® Kit following the manufacturer's protocol. A protein ladder ran every 12 samples for molecular weight determination (kDa) and quantification (ng/μl).

Ovalbumin produced recombinantly was additionally quantified in lysate using a dot blot method (Fig. 5.23). 110 μl of 0.2 μm filtered lysate was mixed with 110 μl 8.0M Guanidine Hydrochloride, 0.1M Sodium Phosphate (Denaturing Buffer) to allow for normalized protein binding and to ensure exposure of the binding epitope. A standard curve of native Ovalbumin (Sigma) was prepared in the same matrix as the samples, starting at 100 ng, diluting 2X serially to 1.56 ng. Invitrogen 0.45 μm nitrocellulose membrane was pre-wet in 1X PBS buffer for 5 minutes and then loaded onto Bio-Rad Dot Blot Apparatus. 300 μl of PBS was vacuumed through to further wet the membrane. Next, 200 μl the 1:1 Sample:Denaturing Buffer mixture was loaded into each well and allowed to drain through the dot blot apparatus by gravity for 30 minutes. Next, a

Table 5.3: Ovalbumin expression constructs.

Plasmid	Description	POI	Integration
pES665	pAQ1int-P*cpc**-[8xHis]ova-lox:Sp:lox	Ovalbumin	pAQ1
pES685	pAQ3int-P*cpc*-[8xHis]ova-lox:Sp:lox	Ovalbumin	pAQ3
pES690	pAQ6int-P*cpc**-[8xHis]ova-lox:Sp:lox	Ovalbumin	pAQ6
pES673	pA1103int-P*cpc**-[8xHis]ova-lox:Sp:lox	Ovalbumin	Chrom
pES678	pEX1580-P*cpc**-[8xHis]ova-lox:Sp:lox	Ovalbumin	Chrom
PCC7002	PCC7002 WT	-	-

300 μl PBS wash was performed on all wells by vacuum followed by loading 300 μl of Millipore Blok CH Noise Cancelling reagent and incubating for 60 minutes. After blocking, the membrane was washed with 300 μl of 1X PBS + 0.1% Tween 20. Next, antibody solution was prepared by adding 2.4 μl of Anti-Ovalbumin antibody (HRP) (Abcam) to 12 ml of Millipore Blok CH Noise Cancelling reagent (1:5000 dilution). 100 μl of the resulting antibody solution was added to each well and allowed to incubate for 30 minutes by gravity. After antibody incubation, three final washes are performed with 300 μl 1X PBS + 0.1% Tween 20 by vacuum. After washes, the nitrocellulose membrane was removed and placed into a reagent tray. 20 ml of Millipore Luminata Classico Western HRP substrate was added and allowed to incubate for 1 minute. After incubation, membrane was placed into imaging tray of Gel Doc™XR+ System (Bio-rad) and imaged using a chemiluminescent protocol.

Ovalbumin production was successful on constitutive P*cpc* with a pAQ3 integration vector (Fig. 5.24), achieving 6% of total protein with an average protein production efficiency of 0.41% kJ protein/kJ photon (Fig. 5.25). This surpasses the agricultural efficiency of egg protein production, which reaches 0.08% kJ protein/kJ photon assuming current yields of corn feed, averaging 160 bushels/acre/yr across the United States in 2013 (U.S. Department of Agriculture's National Agricultural Statistics Service, NASS).

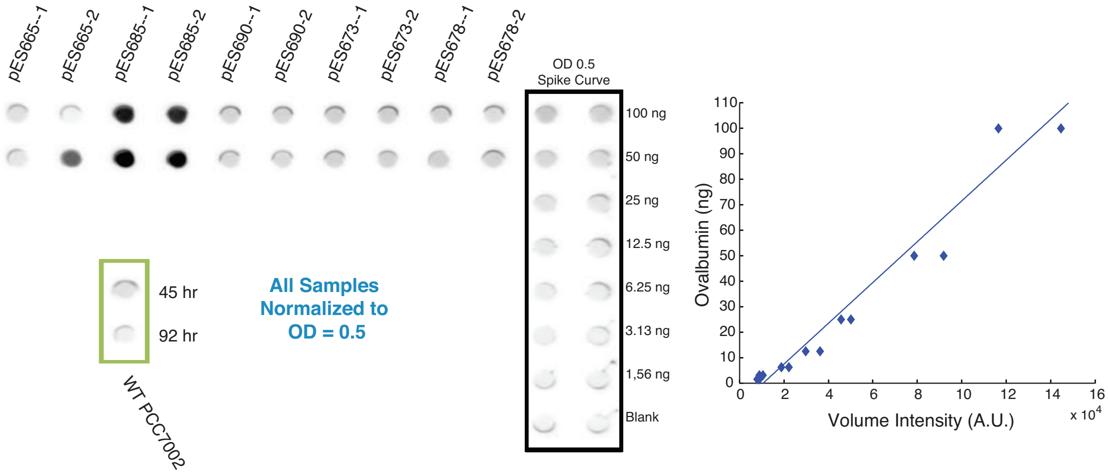


Figure 5.23: Ovalbumin Dot Plot Expression Screen. Detection of recombinant Ovalbumin in PCC7002 cell-lysate with standard curve for quantification.

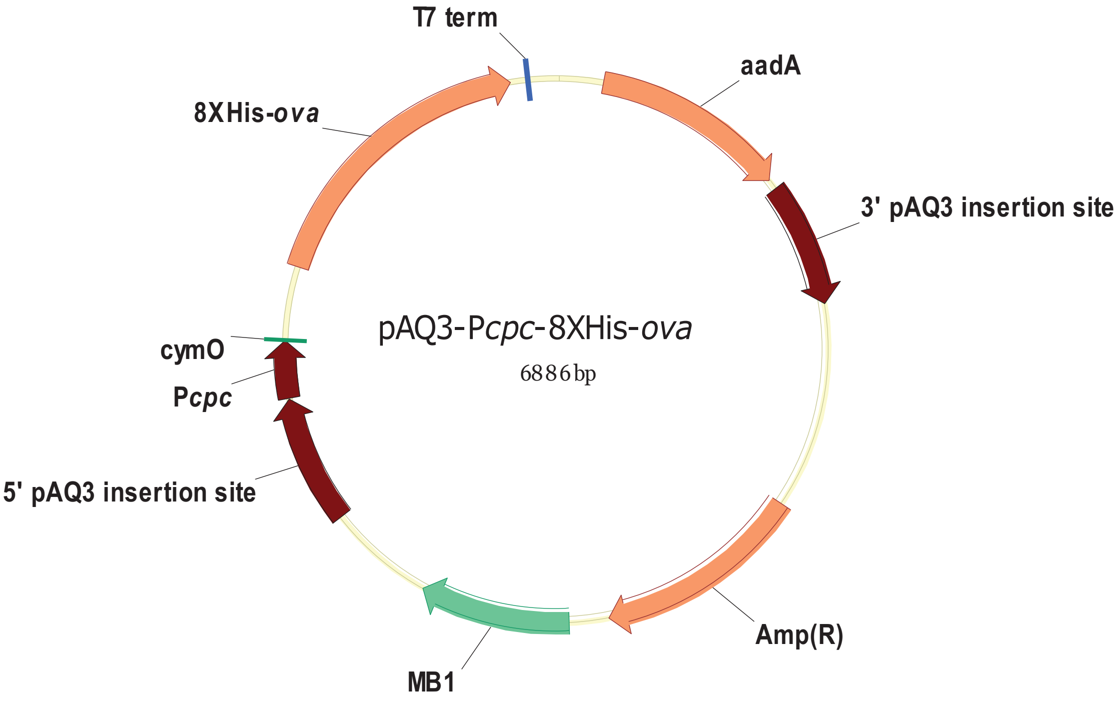


Figure 5.24: pAQ3 P*cpc-ova*. Plasmid map of successful Ovalbumin expression construct, P*cpc-ova* integrated on pAQ3.

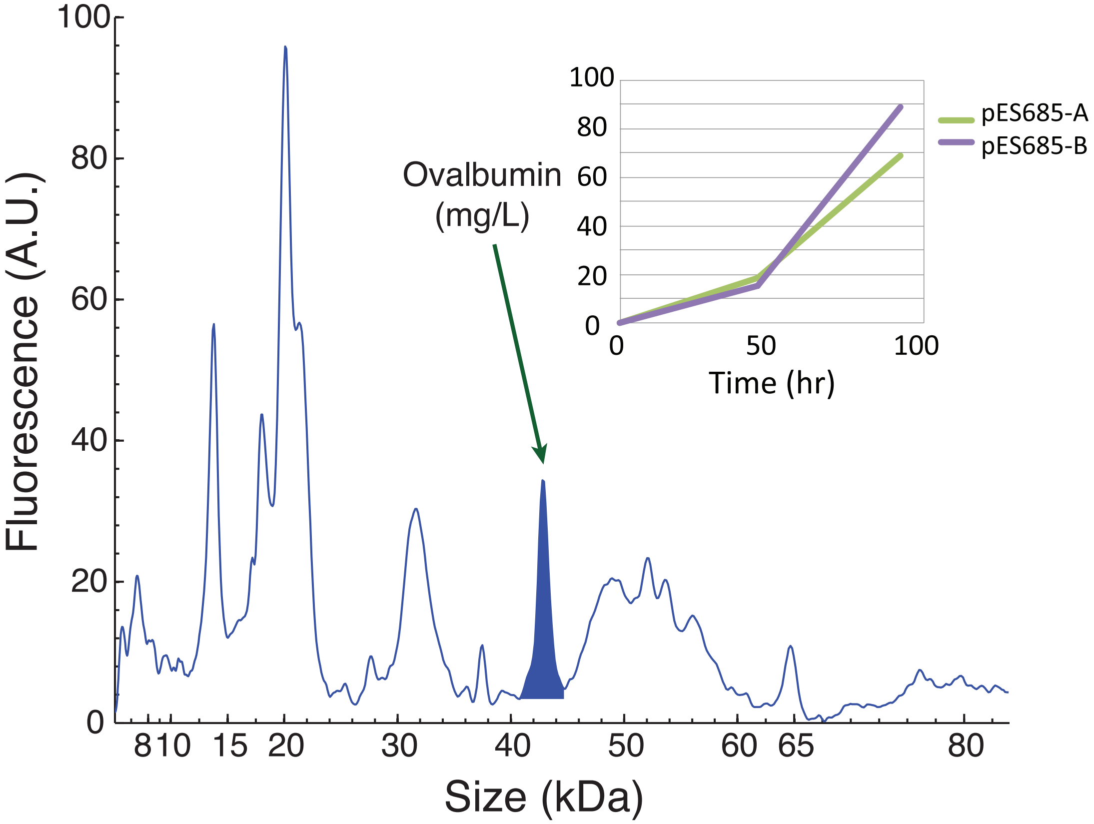


Figure 5.25: Ovalbumin Caliper Detection. Caliper detection of Ovalbumin protein produced in PCC7002. Successful expression is observed from pAQ3-P*cpc* at 6% of total protein content.

Lysis and Purification

Chemical lysis (with detergents) and mechanical lysis (by sonication) was determined to be inefficient at disrupting the cell-membranes of PCC7002 and its derivatives. Mechanical lysis by a microfluidizer has shown to be extremely efficient, but requires at least 50ml total volume, and is not high-throughput. Bead milling is a common method of cell disruption in which a large number of micron-scale beads are shaken vigorously with a sample of cell culture to produce a crushing action that lyses the cells. The QIAGEN TissueLyser II bead mill was chosen and various variables (bead size, bead type, duration of beat milling, frequency of oscillation, sample cell-density, and volume ratio of beads to sample) were examined to define optimal operating conditions to maximize lysis and minimize potential for proteolysis (by heat or protease activity). The optimal parameters were using 0.1mm zirconium beads with 5 min duration at 30Hz and an OD$_{730nm}$ between 5 and 15, and 1:1 volumetric ratio of beads to sample.

Cyanobacterial culture was centrifuged at 7,500g for 20 min at room temperature and the centrate was decanted and removed. Cell-pellet was re-suspended in IMAC equilibration buffer at a solids content of 20% v/v using a handheld mixer. Mechanical lysis was undertaken using a Microfluidics (Westwood, MA) M-110P Microfluidizer using 15,000 psi and three passes. This achieved $>$99.9% lysis as determined by plate counts. The lysate was then clarified by centrifugation at 12,000g for approximately 12 hr at 4°C. The centrate was then passed through a 0.22μm to further reduce the turbidity.

The digestibility of native and recombinant Ovalbumin was measured using an modified protocol from Martos et. al (Martos et al., 2010). Ovalbumin samples were first treated at a concentration of 2 g/L with simulated gastric fluid (0.03 M NaCl, titrated with HCl to pH 1.5 with a final pepsin:OVA ratio of 1:20 w/w) at 37°C. Time points were sampled from the reaction and quenched by addition of 0.2 M Na2CO3. After 120 min in simulated gastric fluid, the remaining reaction was mixed 50:50 with simulated intestinal fluid (15 mM sodium glycodeoxycholate, 15 mM taurocholic acid, 18.4 mM CaCl2, 50 mM MES pH 6.5 with a final trypsin:chymotrypsin:OVA ratio of 1:4:400

w/w) and neutralized with NaOH to pH 6.5. Time points were sampled from the reaction and quenched by addition of Trypsin/Chymotrypsin Inhibitor solution until 120 min. Sampled time points were then be analyzed by chip electrophoresis as described above.

In silico Prediction of Cyanobacteria Sec-type Leader Sequences

In-silico prediction model

The 48 Sec-type leaders examined in this study were selected using the SignalP prediction software (Nielsen et al., 1997; Bendtsen et al., 2004) as well as an in-house set of neural networks trained on a subset of the training dataset provided in the signalP 2.0 package. Using a combination of these two measures, the presence of an N-terminal signal sequence within all previously identified proteins of *Synechococcus sp. PCC 7002* was determined computationally.

The networks developed in-house used an identical prediction strategy to that described by Bendtsen et al. (Bendtsen et al., 2004), (i.e., they used two networks to predict whether a given amino acid was part of an N-terminal Sec-type signal sequence (S-score) and/or the cleavage site of the leader (C-score) and used the combined output of both networks to assess the presence of a N-terminal Sec-type signal sequence (D-score)), but differed in the training data used. Each network was trained using a 5 fold cross validation strategy where the input data was randomly split into five sets, three of these sets were used to actively train the network, the fourth was used during training as a validation dataset to prevent over fitting, and the fifth was used after training to assess model performance. Specifically, the networks each had two hidden layers, were trained using the gram negative bacteria training dataset provided in the signal 2.0 package, and implemented using the biopython v1.53 toolbox and python v2.6. Training was performed using a momentum of 0.05, a learning rate of 0.1, and training stopped when the prediction error, assessed using the validation dataset, started to increase, which indicated that the model was starting to over fit to the training data.

The S-score network was specifically trained using four pieces of data from each position in each sequence in the training dataset: the amino acid distribution of a window of 40 amino acids that included the 20 residues before and after each position, the amino acid distribution of the first 60 amino acids, the position index, and its identity as a signal sequence, cleavage, or normal residue. The C-score network was trained using similar data but used a 22 amino acid window around each cleavage site that included 20 amino acids N-terminal to the cleavage site and 1 amino acid C-terminal to the cleavage site. Given the disparity between the number of positions in the training set that were members of a signal sequence relative to those that were not, the negative examples were randomly sampled such than an equal number of positive and negative examples were selected for training.

The prediction statistics obtained from the 5 fold cross validation of the trained S and C networks used to predict the presence of an N-terminal signal sequence are shown in table 5.4. The Matthews correlation coefficient (MCC) was maximized using a D value cutoff of 0.32, yielding an MCC of 0.84, which indicates a high degree of correlation between the observed and predicted signal sequences. Similarly, the accuracy, sensitivity, and specificity are all approximately $\geqslant$0.9. Note that for signalP v4.0, the D value cutoff used was 0.57.

Table 5.4: Prediction statistics for trained S and C networks.

MCC	0.84
Accuracy	0.92
Sensitivity	0.95
Specificity	0.89

In an effort to increase the likelihood of selecting active, N-terminal Sec-type signal sequences from the synechococcal library, the maximal global homology of each candidate protein to secreted proteins contained within the signalP 2.0 training dataset was determined in order to identify those proteins with high homology to other proteins from gram-negative bacteria known to be secreted via the sec pathway. All alignments were performed using the FASTA algorithm with the BLOSUM50 substitution matrix, a gap open penalty of 10, and a gap extension penalty of 2 (Pearson and Lipman, 1988).

The tested set of protein sequences included all non-redundant sequences collected from searching NCBI protein database (07/10/2012) using the taxon ID 32049. Each sequence was processed using both signal 4.0 and the in-house model (Fig. 5.26), and the maximal sequence homology of each protein to the training set was assessed using a global-global optimal alignment. The results of these scans are shown in Fig. 5.27 and listed in Supplementary Dataset S3.

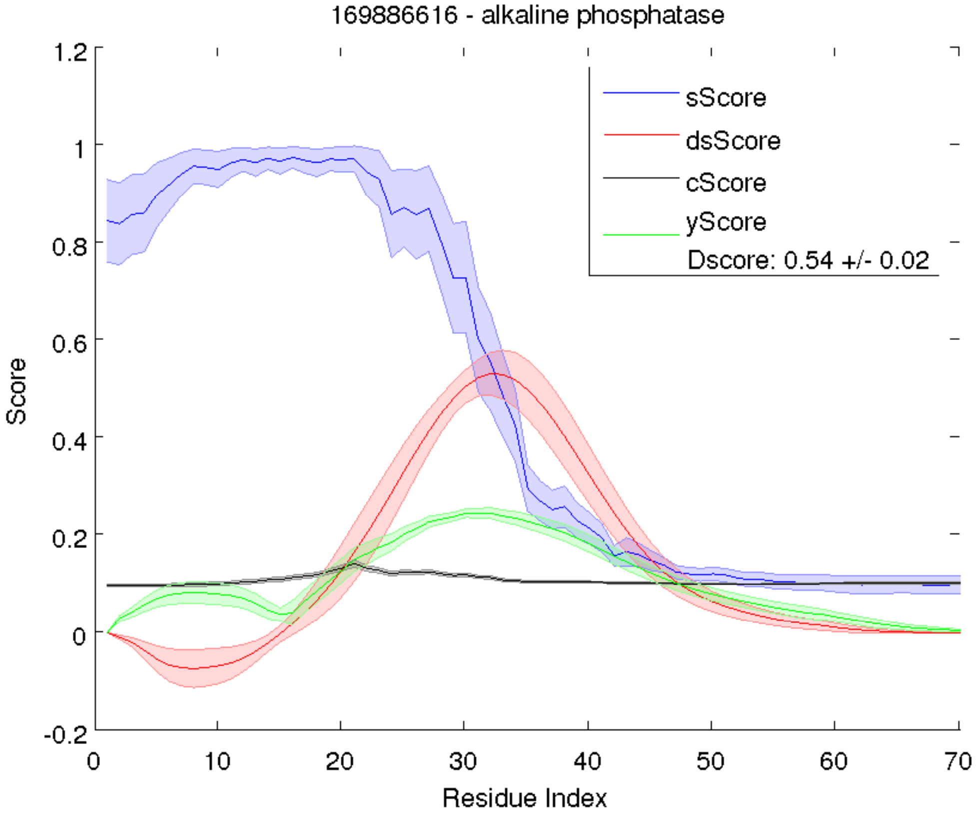


Figure 5.26: Signal Peptide Neural Network Scoring. An example network score for a *Synechococcus sp. PCC 7002* alkaline phosphatase.

The 48 leader sequences tested in this study (Supplementary Dataset S4) were

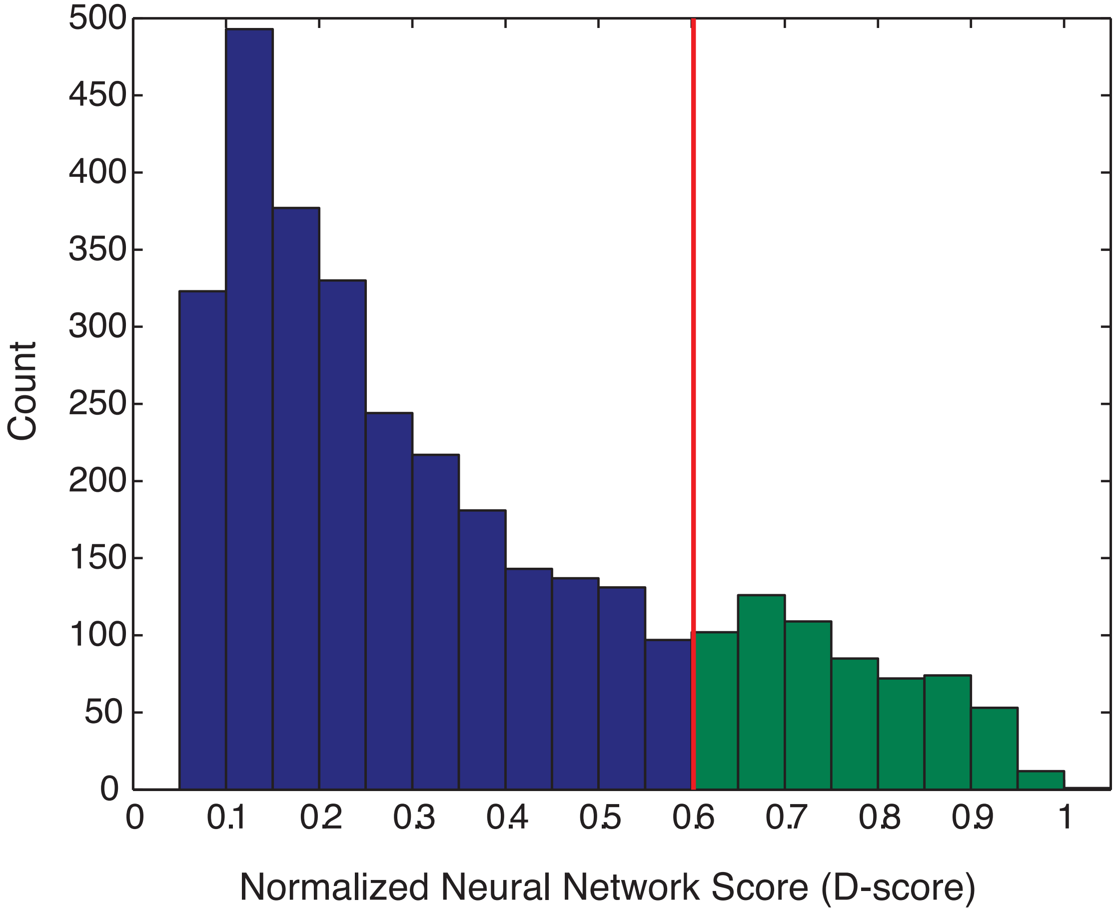


Figure 5.27: PCC7002 SignalP Network Score. The network D-score for the *Synechococcus sp. PCC 7002* proteome, highlighting the subset surpassing the SignalP v4.0 cutoff (red line).

selected by ranking all the putative secretion leaders based upon their computed network scores (i.e., the D-scores assigned to each sequence by both SignalP and the in-house trained network) and their calculated homology to the training set. The calculated values of each feature were each rescaled and mapped to the range [0,1] using the scaling function

$$x' = \frac{x - x_{min}}{x_{max} - x'_{min}} \tag{5.4}$$

where x is there original value, x_{min} is the minimum observed value, x_{max} is the maximum observed value, and x' is the scaled value, and an average of each scaled score was calculated. Note that for sequence homology, the negative logarithm of the expectation value was used. The specific secretion leader sequence for secretion screening was taken to be the SignalP v4.0 predicted leader sequence.

Screening predicted leader sequences

All 48 predicted secretion leaders were fused in N-terminal of lichenase and assembled downstream of the pAQ3-P*cro-CumO* promoter using an AarI site in the flanking region of vector primers. The leaders were constructed by hybridizing forward and reverse oligonucleotides that correspond to the gene leader sequence, with 4 bases of overhang compatible with the vector AarI sites. A Flag tag was added to C-terminal of lichenase for detection in Western Blot and DOT-BLOT. One of the leader sequences, leader 10 didn't transform in *Synechococcus sp. PCC 7002*.

1 μg of each construct was mixed with 30 μl of *Synechococcus sp. PCC 7002* that was grown in A+ medium to OD$_{730nm}$ $\approx$ 1.5. The mix was loaded on a grid agar plate with A+ medium supplemented with spectinomycin 100 μg/mL. The transformants were incubated at 35°C, 10 μE/m^2/s and atmospheric CO$_2$ conditions. After 2 days, the growing bacteria were re-streaked on a similar grid agar plate (A+ spectinomycin 100

μg/ml). Three colonies were picked and resuspended in 1 ml of A+ spectinomycin 100 μg/ml medium in 2 ml 96 deep well block and incubated in a 900 rpm plate shaker at 100 μE/m^2/s, 2% CO_2 (v/v), and 35°C. After 3 days, cells were diluted to 1 ml fresh A+ spectinomycin 100 μg/ml medium to final OD_{730nm} $\sim$0.2 in 2 ml 96 deep well block and incubated in similar conditions for 29 hrs when all the cultures reached OD_{730nm} 1. All cultures were then induced with 50 μM cumate and grown for another 18 hours. At this stage, the cells were harvested and centrifuged at 3000 rpm for 10 mins. The supernatant was filtered with 0.22 μM PVDF filter and the filtered supernatant was used for lichenase assay.

The filtered supernatant was incubated at 6°C for 30 minutes with lichenan as a substrate for release of glucose. The presence and quantity of the released glucose was measured with a dinitrosalicylic acid assay (Miller, 1959): 3-5 dinitrosalicylic acid reduces to 3-amino-5-nitrosalicylic acid in the presence of reducing sugars and has a linear range of approximately 0.1 to 1.00 mg D-glucose per mL at 100°C for 10 minutes. 3-amino-5-nitrosalicylic acid has an absorbance wavelength of 540 nm. For the purposes of this experiment the specific activity of lichenase was taken to be 192 μmol/min/mg for the native enzyme with substrate lichenan at pH 6.0 at 70°C (Scheer et al., 2011). The standard curve for dinitrosalicylic acid assay is shown in Fig. 5.28.

The results are summarized in Fig. 5.4e. Approximately 64% of the predicted leaders yielded strain activities of Lichenase greater than 0.5 μg/mL/OD_{730nm}, 52% of the leaders yielded activities greater than 0.75 μg/mL/OD_{730nm}, 37% of the leaders yielded activities greater than 1 μg/mL/OD_{730nm}, and 23% of the leaders yielded activities greater than 1.25 μg/mL/OD_{730nm}. The most effective leader sequences yielded activities $\approx$80-fold higher than that of lichenase expressed with no N-terminal fused signal peptide.

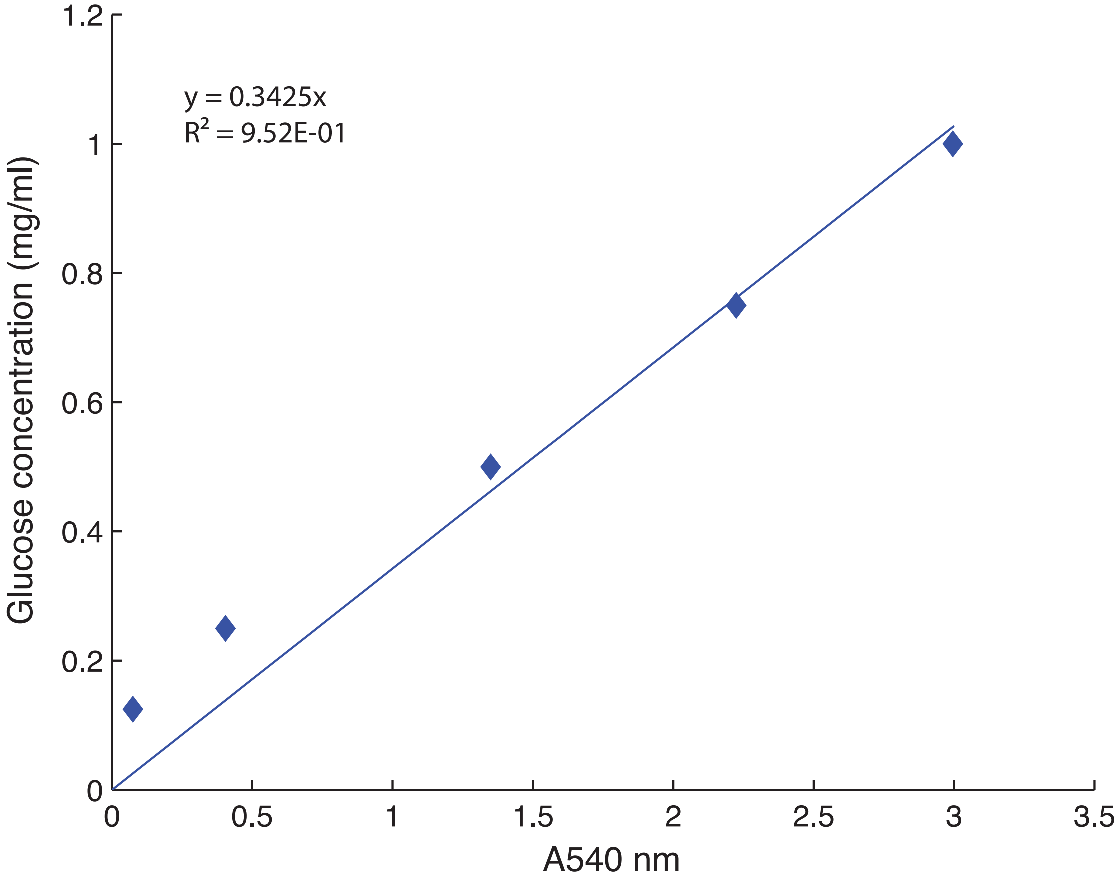


Figure 5.28: Glucose Standard Curve. A540nm linear fit of 3-5 dinitrosalicylic acid reduction across a range of glucose concentrations. Using the lichenase specific activity, the rate of glucose appearance determines lichenase concentration in the cell supernatant.

Acknowledgments

This work was supported by Flagship VentureLabs, the Gates Foundation, and NSF Graduate Research Fellowship Program under Grant No. DGE-1144086 (P.S.). We would like to thank the following colleagues for providing materials or experimental assistance: Gaozhong Shen (providing initial parent strains of interest), Roland Liu and Mini Zacharia (cloning expression constructs), Stephanie DeMarino (culturing techniques), Bryan Keen and Gavin McKeown (mechanical design work), Carine Robichon-Iyer and Luke Hamm (lichenase secretion leader library), Ryan Beck (lysis and purification), Vimal Vaidya (analytical characterization), Ying-Ja Chen (mRNA extraction), and Bonnie Jain (coding).

Chapter 5, in part, is currently being prepared for submission for publication of the material. Samayoa, P. Von Maltzahn, G., Boghigian, B., Basu, S., Young, D., Silver, N., Hamill, M. and Berry, D: Programming cyanobacteria to produce protein nutrients at supra-agricultural efficiencies.

Chapter 6

Summary and Conclusion

Synthetic biology offers the opportunity to interface with the natural functionality of microorganisms for therapeutic, industrial, and even computational pursuits. Central to fully realizing this potential is the rigorous design of engineered gene networks, from the mathematical basis for functional dynamics through analysis of the host-environment response at the point of application. In this thesis I have described the host selection, design, construction, and characterization of scalable synthetic circuits that robustly translate to commercially relevant bacteria. In coupling a synthetic oscillator to native regulatory control, we also revealed the functional scalability of synthetic networks that is unlocked by harnessing long-range bacterial communication. Finally, we demonstrated the potential to apply a rational, bottom-up construction of a synthetic circuit to enable autotrophic production of protein nutrients at superior efficiencies. These works reflect the capacity to integrate our rapid advancement in engineering biological circuits towards transformational applications.

Bibliography

2009: Issue brief: How to feed the world in 2050. In *Expert Meeting on How to feed the World in*, volume 2050.

Agemy, L., Friedmann-Morvinski, D., Kotamraju, V. R., Roth, L., Sugahara, K. N., Girard, O. M., Mattrey, R. F., Verma, I. M., and Ruoslahti, E., 2011: Targeted nanoparticle enhanced proapoptotic peptide as potential therapy for glioblastoma. *Proceedings of the National Academy of Sciences*, **108**(42), 17450–17455.

Allen, T. E., and Palsson, B. Ø., 2003: Sequence-based analysis of metabolic demands for protein synthesis in prokaryotes. *Journal of theoretical biology*, **220**(1), 1–18.

Alper, H., and Stephanopoulos, G., 2009: Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? *Nature Reviews Microbiology*, **7**(10), 715–723.

Anderson, C., 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. ISSN 00222836. doi:10.1016/j.jmb.2005.10.076.

Araus, J. L., Slafer, G. A., Buxo, R., and Romagosa, I., 2003: Productivity in prehistoric agriculture: physiological models for the quantification of cereal yields as an alternative to traditional approaches. *Journal of Archaeological Science*, **30**(6), 681–693.

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.

Balagadde, F., Song, H., Ozaki, J., Collins, C., Barnet, M., Arnold, F., Quake, S., and You, L., 2008: A synthetic Escherichia coli predator–prey ecosystem. *Molecular Systems Biology*, **4**.

Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H., and Weiss, R., 2005: A synthetic multicellular system for programmed pattern formation. *Nature*, **434**(7037), 1130–4.

Basu, S., Mehreja, R., Thiberge, S., Chen, M., and Weiss, R., 2004: Spatiotemporal control of gene expression with pulse-generating networks. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(17), 6355.

Bauer, H., Darji, A., Chakraborty, T., and Weiss, S., 2004: Salmonella-mediated oral dna vaccination using stabilized eukaryotic expression plasmids. *Gene therapy*, **12**(4), 364–372.

Beadle, C., and Long, S., 1985: Photosynthesis - is it limiting to biomass production? *Biomass*, **8**(2), 119–168.

Bendtsen, J., Nielsen, H., von Heijne, G., and Brunak, S., 2004: Improved prediction of signal peptides: Signalp 3.0. *Journal of Molecular Biology*, **340**(4), 783–795.

Berg, J., Tymoczko, J. L., and Stryer, L., 2006: *Biochemistry*. W.H. Freeman.

Berla, B. M., Saha, R., Immethun, C. M., Maranas, C. D., Moon, T. S., and Pakrasi, H. B., 2013: Synthetic biology of cyanobacteria: unique challenges and opportunities. *Frontiers in microbiology*, **4**.

Blake, W., Balázsi, G., Kohanski, M., Isaacs, F., Murphy, K., Kuang, Y., Cantor, C., Walt, D., and Collins, J., 2006: Phenotypic consequences of promoter-mediated transcriptional noise. *Molecular cell*, **24**(6), 853–865.

Blankenship, R. E., Tiede, D. M., Barber, J., Brudvig, G. W., Fleming, G., Ghirardi, M., Gunner, M., Junge, W., Kramer, D. M., Melis, A., et al., 2011: Comparing photosynthetic and photovoltaic efficiencies and recognizing the potential for improvement. *Science*, **332**(6031), 805–809.

Bose, J. L., Kim, U., Bartkowski, W., P., G. R., Overley, A. M., Lyell, N. L., Visick, K. L., and Stabb, E. V., 2007: Bioluminescence in vibrio fischeri is controlled by the redox-responsive regulator arca. *Molecular Microbiology*, **65**(2), 538–553.

Butzer, K. W., 1976: Early hydraulic civilization in egypt: a study in cultural ecology. *Prehistoric archeology and ecology (USA)*.

Chatterjee, A., Xiao, H., and Schultz, P. G., 2012: Evolution of multiple, mutually orthogonal prolyl-trna synthetase/trna pairs for unnatural amino acid mutagenesis in escherichia coli. *Proceedings of the National Academy of Sciences*, **109**(37), 14841–14846.

Chin, J. W., Cropp, T. A., Anderson, J. C., Mukherji, M., Zhang, Z., and Schultz, P. G., 2003: An expanded eukaryotic genetic code. *Science*, **301**(5635), 964–967.

Choi, Y. J., Morel, L., Le François, T., Bourque, D., Bourget, L., Groleau, D., Massie, B., and Míguez, C. B., 2010: Novel, versatile, and tightly regulated expression system for escherichia coli strains. *Applied and environmental microbiology*, **76**(15), 5058–5066.

Cookson, N. A., Mather, W. H., Danino, T., Mondragon-Palomino, O., Williams, R. J., Tsimring, L. S., and Hasty, J., 2011: Queueing up for enzymatic processing: Correlated signaling through coupled degradation. *Molecular Systems Biology*.

Cookson, S., Ostroff, N., Pang, W. L., Volfson, D., and Hasty, J., 2005: Monitoring dynamics of single-cell gene expression over multiple cell cycles. *Mol. Syst. Biol.*, **1**, msb4100032–E1–6.

Danino, T., Lo, J., Prindle, A., Hasty, J., and Bhatia, S. N., 2012: In vivo gene expression dynamics of tumor-targeted bacteria. *ACS synthetic biology*, **1**(10), 465–470.

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

Daunert, S., Barrett, G., Feliciano, J., Shetty, R., Shrestha, S., and Smith-Spencer, W., 2000: Genetically engineered whole-cell sensing systems: coupling biological recognition with reporter genes. *Chemical Reviews*, **100**(7), 2705–2738.

Duan, F., Curtis, K., and March, J., 2008: Secretion of insulinotropic proteins by commensal bacteria: Rewiring the gut to treat diabetes. *Applied and environmental microbiology*, **74**(23), 7437–7438.

Duan, F., and March, J., 2010: Engineered bacterial communication prevents vibrio cholerae virulence in an infant mouse model. *Science's STKE*, **107**(25), 11260.

Ellis, T., Wang, X., and Collins, J., 2009: Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nature biotechnology*, **27**(5), 465.

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

Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S., 2002: Stochastic gene expression in a single cell. *Science*, **297**(5584), 1183–6.

Endy, D., 2005: Foundations for engineering biology. *Nature*, **438**(7067), 449–53.

Ferrer-Lorente, R., Fernández-López, J. A., and Alemany, M., 2007: Estimation of the metabolizable energy equivalence of dietary proteins. *European journal of nutrition*, **46**(1), 1–11.

Ferry, M., Razinkov, I., and Hasty, J., 2011: Microfluidics for synthetic biology from design to execution. *Methods in enzymology*, **497**, 295.

Fischer, R., Byerlee, D., Edmeades, G. O., et al., 2009: Can technology deliver on the yield challenge to 2050. In *Expert Meeting on How to feed the World in*, volume 2050.

Fischer, R., and Edmeades, G. O., 2010: Breeding and cereal yield progress. *Crop Science*, **50**(Supplement_1), S–85.

Food and Agriculture Organization of the United Nations (FAO), 2003: Agriculture, food and water.

Forbes, N., 2010: Engineering the perfect (bacterial) cancer therapy. *Nature Reviews Cancer*, **10**(11), 785–794.

Forbes, N., Munn, L., Fukumura, D., and Jain, R., 2003: Sparse initial entrapment of systemically injected salmonella typhimurium leads to heterogeneous accumulation within tumors. *Cancer research*, **63**(17), 5188.

Fridovich, I., 1978: The biology of oxygen radicals. *Science*, **201**(4359), 875.

Friedland, A. E., Lu, T. K., Wang, X., Shi, D., Church, G., and Collins, J. J., 2009: Synthetic gene networks that count. *science*, **324**(5931), 1199–1202.

Fu, P., 2009: Genome-scale modeling of synechocystis sp. pcc 6803 and prediction of pathway insertion. *Journal of chemical technology and biotechnology*, **84**(4), 473–483.

Fuqua, W. C., Winans, S. C., and Greenberg, E. P., 1994: Quorum sensing in bacteria: the luxr-luxi family of cell density-responsive transcriptional regulators. *Journal of bacteriology*, **176**(2), 269.

Gahan, M., Webster, D., Wesselingh, S., and Strugnell, R., 2007: Impact of plasmid stability on oral dna delivery by¡ i¿ salmonella enterica¡/i¿ serovar typhimurium. *Vaccine*, **25**(8), 1476–1483.

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

Gast, T., 1985: Sensors with oscillating elements. *Journal of Physics E: Scientific Instruments*, **18**, 783.

Georgellis, D., Kwon, O., and Lin, E., 2001: Quinones as the redox signal for the arc two-component system of bacteria. *Science*, **292**(5525), 2314.

Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R.-Y., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., et al., 2010: Creation of a bacterial cell controlled by a chemically synthesized genome. *science*, **329**(5987), 52–56.

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., and Toulmin, C., 2010: Food security: the challenge of feeding 9 billion people. *Science*, **327**(5967), 812–818.

Golding, I., Paulsson, J., Zawilski, S., and Cox, E., 2005: Real-time kinetics of gene activity in individual bacteria. *Cell*, **123**(6), 1025–1036.

Goodland, R., Anhang, J., et al., 2009: Livestock and climate change: what if the key actors in climate change are... cows, pigs, and chickens?

Grilly, C., Stricker, J., Pang, W., Bennett, M., and Hasty, J., 2007: A synthetic gene network for tuning protein degradation in saccharomyces cerevisiae. *Molecular systems biology*, **3**(1).

Grobbelaar, J. U., 1994: Turbulence in mass algal cultures and the role of light/dark fluctuations. *Journal of Applied Phycology*, **6**(3), 331–335.

Grobbelaar, J. U., Nedbal, L., and Tichỳ, V., 1996: Influence of high frequency light/dark fluctuations on photosynthetic characteristics of microalgae photoacclimated to different light intensities and implications for mass algal cultivation. *Journal of applied phycology*, **8**(4-5), 335–343.

Gronenberg, L. S., Marcheschi, R. J., and Liao, J. C., 2013: Next generation biofuel engineering in prokaryotes. *Current opinion in chemical biology*.

Guido, N. J., Wang, X., Adalsteinsson, D., McMillen, D., Hasty, J., Cantor, C. R., Elston, T. C., and Collins, J. J., 2006: A bottom-up approach to gene regulation. *Nature*, **439**(7078), 856–60.

Guo, H., Zhang, J., Inal, C., Nguyen, T., Fruehauf, J. H., Keates, A. C., and Li, C. J., 2011: Targeting tumor gene by shrna-expressing salmonella-mediated rnai. *Gene Ther*, **18**(1), 95–105. doi:10.1038/gt.2010.112.

Ham, T., Lee, S., Keasling, J., and Arkin, A., 2008: Design and construction of a double inversion recombination switch for heritable sequential genetic memory. *PLoS One*, **3**(7), e2815.

Hamilton, J. J., and Reed, J. L., 2012: Identification of functional differences in metabolic networks using comparative genomics and constraint-based models. *PloS one*, **7**(4), e34670.

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

Heimann, D., and Rosenberg, S., 2003: Continuous intravenous administration of live genetically modified salmonella typhimurium in patients with metastatic melanoma. *Journal of immunotherapy (Hagerstown, Md.: 1997)*, **26**(2), 179.

Hoffman, R., 2011: Tumor-seeking salmonella amino acid auxotrophs. *Current Opinion in Biotechnology*.

Hohmann, E., Oletta, C., and Miller, S., 1996: Evaluation of a phop/phoq-deleted, aroa-deleted live oral salmonella typhi vaccine strain in human volunteers. *Vaccine*, **14**(1), 19–24.

Hong, S.-J., and Lee, C.-G., 2007: Evaluation of central metabolism based on a genomic database ofsynechocystis pcc6803. *Biotechnology and Bioprocess Engineering*, **12**(2), 165–173.

Hooshangi, S., Thiberge, S., and Weiss, R., 2005: Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(10), 3581.

Human Microbiome Project Consortium and others, 2012: Structure, function and diversity of the healthy human microbiome. *Nature*, **486**(7402), 207–214.

Janssen, M., Tramper, J., Mur, L. R., and Wijffels, R. H., 2003: Enclosed outdoor photobioreactors: light regime, photosynthetic efficiency, scale-up, and future prospects. *Biotechnology and Bioengineering*, **81**(2), 193–210.

Jasny, B., and Zahn, L., 2011: A celebration of the genome, part i. *Science*, **331**(6017), 546.

Jia, L. J., Wei, D. P., Sun, Q. M., Huang, Y., Wu, Q., and Hua, Z. C., 2007: Oral delivery of tumor-targeting salmonella for cancer therapy in murine tumor models. *Cancer Sci*, **98**(7), 1107–1112. doi:10.1111/j.1349-7006.2007.00503.x.

Kasinskas, R. W., and Forbes, N. S., 2006: Salmonella typhimurium specifically chemotax and proliferate in heterogeneous tumor tissue in vitro. *Biotechnol Bioeng*, **94**(4), 710–721. doi:10.1002/bit.20883.

Kasinskas, R. W., and Forbes, N. S., 2007: Salmonella typhimurium lacking ribose chemoreceptors localize in tumor quiescence and induce apoptosis. *Cancer Res*, **67**(7), 3201–3209. doi:10.1158/0008-5472.CAN-06-2618.

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.

Kelner, M., Bagnell, R., and Welch, K., 1990: Thioureas react with superoxide radicals to yield a sulfhydryl compound. explanation for protective effect against paraquat. *Journal of Biological Chemistry*, **265**(3), 1306.

Khalil, A., and Collins, J., 2010: Synthetic biology: applications come of age. *Nature Reviews Genetics*, **11**(5), 367–379.

Kiers, E. T., Rousseau, R. A., West, S. A., and Denison, R. F., 2003: Host sanctions and the legume–rhizobium mutualism. *Nature*, **425**(6953), 78–81.

Knoop, H., Zilliges, Y., Lockau, W., and Steuer, R., 2010: The metabolic network of synechocystis sp. pcc 6803: systemic properties of autotrophic growth. *Plant Physiology*, **154**(1), 410–422.

Kobayashi, H., Kærn, M., Araki, M., Chung, K., Gardner, T. S., Cantor, C. R., and Collins, J. J., 2004: Programmable cells: interfacing natural and engineered gene networks. *Proc. Natl. Acad. Sci. U. S. A.*, **101**(22), 8414–9.

Kohanski, M. A., DePristo, M. A., and Collins, J. J., 2010: Sublethal antibiotic treatment leads to multidrug resistance via radical-induced mutagenesis. *Molecular Cell*, **37**, 311–320.

Kotula, J., Kerns, S., Shaket, L., Siraj, L., Collins, J., Way, J., and PA, S., 2014: Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proceedings of the National Academy of Sciences*, **in press**.

Kramer, B., and Fussenegger, M., 2005: Hysteresis in a synthetic mammalian gene network. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(27), 9517.

Lee, H. H., Molla, M. N., Cantor, C. R., and Collins, J. J., 2010: Bacterial charity work leads to population-wide resistance. *Nature*, **467**(7311), 82–85.

Leveau, J., and Lindow, S., 2002: Bioreporters in microbial ecology. *Current opinion in microbiology*, **5**(3), 259–265.

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

Lm, Z., Luo, X., Feng, M., Li, Z., Ittensohn, M., Trailsmith, M., Bermudes, D., Lin, S., King, I., et al., 2001: Tumor amplified protein expression therapy: Salmonella as a tumor-selective protein delivery vector. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics*, **12**(3), 127–135.

Lobell, D. B., Schlenker, W., and Costa-Roberts, J., 2011: Climate trends and global crop production since 1980. *Science*, **333**(6042), 616–620.

Low, K., Ittensohn, M., Le, T., Platt, J., Sodi, S., Amoss, M., Ash, O., Carmichael, E., Chakraborty, A., Fischer, J., et al., 1999: Lipid a mutant salmonella with suppressed virulence and tnfα induction retain tumor-targeting in vivo. *Nature biotechnology*, **17**(1), 37–41.

Lu, T., and Collins, J., 2007: Dispersing biofilms with engineered enzymatic bacteriophage. *Proceedings of the National Academy of Sciences*, **104**(27), 11197.

Lu, T., and Collins, J., 2009: Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy. *Proceedings of the National Academy of Sciences*, **106**(12), 4629.

Lundberg, D. S., Lebeis, S. L., Paredes, S. H., Yourstone, S., Gehring, J., Malfatti, S., Tremblay, J., Engelbrektson, A., Kunin, V., del Rio, T. G., et al., 2012: Defining the core arabidopsis thaliana root microbiome. *Nature*, **488**(7409), 86–90.

Lutz, R., and Bujard, H., 1997: 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.

Martos, G., Contreras, P., Molina, E., and López-FandiŃo, R., 2010: Egg white ovalbumin digestion mimicking physiological conditions. *Journal of agricultural and food chemistry*, **58**(9), 5640–5648.

Mather, W., Bennett, M., Hasty, J., and Tsimring, L., 2009: Delay-Induced Degrade-and-Fire Oscillations in Small Genetic Circuits. *Biophys. J Phys Rev Lett*, **102**, 068105.

Matzas, M., Stähler, P., Kefer, N., Siebelt, N., Boisguérin, V., Leonard, J., Keller, A., Stähler, C., Häberle, P., Gharizadeh, B., et al., 2010: High-fidelity gene synthesis by retrieval of sequence-verified dna identified using high-throughput pyrosequencing. *Nature biotechnology*, **28**(12), 1291–1294.

McClelland, M., Sanderson, K., Spieth, J., Clifton, S., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R., and Wilson, R., 2001: Complete genome sequence of salmonella enterica serovar typhimurium lt2. *Nature*, **413**(6858), 852–856. ISSN 0028-0836. doi:10.1038/35101614.

McCord, J., and Fridovich, I., 1969: Superoxide dismutase. *Journal of Biological Chemistry*, **244**(22), 6049.

Merril, C., Scholl, D., and Adhya, S., 2003: The prospect for bacteriophage therapy in western medicine. *Nature Reviews Drug Discovery*, **2**(6), 489–497.

Messner, K., and Imlay, J., 1999: The identification of primary sites of superoxide and hydrogen peroxide formation in the aerobic respiratory chain and sulfite reductase complex of escherichia coli. *Journal of Biological Chemistry*, **274**(15), 10119.

Miller, G. L., 1959: Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Analytical chemistry*, **31**(3), 426–428.

Mondragón-Palomino, O., Danino, T., Selimkhanov, J., Tsimring, L., and Hasty, J., 2011: Entrainment of a population of synthetic genetic oscillators. *Science*, **333**(6047), 1315–1319.

Montagud, A., Navarro, E., de Córdoba, P. F., Urchueguía, J., and Patil, K., 2010: Reconstruction and analysis of genome-scale metabolic model of a photosynthetic bacterium. *BMC systems biology*, **4**(1), 156.

Montagud, A., Zelezniak, A., Navarro, E., de Córdoba, P. F., Urchueguía, J. F., and Patil, K. R., 2011: Flux coupling and transcriptional regulation within the metabolic network of the photosynthetic bacterium synechocystis sp. pcc6803. *Biotechnology journal*, **6**(3), 330–342.

Moon, T. S., Lou, C., Tamsir, A., Stanton, B. C., and Voigt, C. A., 2012: Genetic programs constructed from layered logic gates in single cells. *Nature*, **491**(7423), 249–253.

Moser, F., Broers, N. J., Hartmans, S., Tamsir, A., Kerkman, R., Roubos, J. A., Bovenberg, R., and Voigt, C. A., 2012: Genetic circuit performance under conditions relevant for industrial bioreactors. *ACS synthetic biology*, **1**(11), 555–564.

Mukherji, S., and Van Oudenaarden, A., 2009: Synthetic biology: understanding biological design from synthetic circuits. *Nature Reviews Genetics*, **10**(12), 859–871.

Mutalik, V. K., Qi, L., Guimaraes, J. C., Lucks, J. B., and Arkin, A. P., 2012: Rationally designed families of orthogonal rna regulators of translation. *Nature chemical biology*, **8**(5), 447–454.

Nagakura, C., Hayashi, K., Zhao, M., Yamauchi, K., Yamamoto, N., Tsuchiya, H., Tomita, K., Bouvet, M., and Hoffman, R. M., 2009: Efficacy of a genetically-modified salmonella typhimurium in an orthotopic human pancreatic cancer in nude mice. *Anticancer Res*, **29**(6), 1873–1878.

Nandagopal, N., and Elowitz, M. B., 2011: Synthetic biology: Integrated gene circuits. *Science*, **333**(6047), 1244–1248.

Navarro, E., Montagud, A., Fernández de Córdoba, P., and Urchueguía, J., 2009: Metabolic flux analysis of the hydrogen production potential in¡ i¿ synechocystis¡/i¿ sp. pcc6803. *international journal of hydrogen energy*, **34**(21), 8828–8838.

Nemunaitis, J., Cunningham, C., Senzer, N., Kuhn, J., Cramm, J., Litz, C., Cavagnolo, R., Cahill, A., Clairmont, C., and Sznol, M., 2003: Pilot trial of genetically modified, attenuated salmonella expressing the e. coli cytosine deaminase gene in refractory cancer patients. *Cancer gene therapy*, **10**(10), 737–744.

Nguyen, V. H., Kim, H. S., Ha, J. M., Hong, Y., Choy, H. E., and Min, J. J., 2010: Genetically engineered salmonella typhimurium as an imageable therapeutic probe for cancer. *Cancer Res*, **70**(1), 18–23. doi:10.1158/0008-5472.CAN-09-3453.

Nielsen, H., Engelbrecht, J., Brunak, S., and Von Heijne, G., 1997: Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein engineering*, **10**(1), 1.

Nogales, J., Gudmundsson, S., Knight, E. M., Palsson, B. O., and Thiele, I., 2012: Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. *Proceedings of the National Academy of Sciences*, **109**(7), 2678–2683.

Nordstrom, D., 2002: Worldwide occurrences of arsenic in ground water. *Science*, **296**(5576), 2143.

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

Pawelek, J., Low, K., and Bermudes, D., 1997: Tumor-targeted salmonella as a novel anticancer vector. *Cancer research*, **57**(20), 4537.

Pearson, W. R., and Lipman, D. J., 1988: Improved tools for biological sequence comparison. *Proceedings of the National Academy of Sciences*, **85**(8), 2444–2448.

Pedelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S., 2006: Engineering and characterization of a superfolder green fluorescent protein. *Nature Biotechnology*, **24**(1), 79–88.

Prindle, A., Samayoa, P., Razinkov, I., Danino, T., Tsimring, L. S., and Hasty, J., 2011: A sensing array of radically coupled genetic 'biopixels'. *Nature*, **481**(7379), 39–44.

Projan, S., 2004: Phage-inspired antibiotics? *Nature biotechnology*, **22**(2), 167–168.

Punt, P. J., van Biezen, N., Conesa, A., Albers, A., Mangnus, J., and van den Hondel, C., 2002: Filamentous fungi as cell factories for heterologous protein production. *Trends in biotechnology*, **20**(5), 200–206.

Qiang, H., Zarmi, Y., and Richmond, A., 1998: Combined effects of light intensity, light-path and culture density on output rate of spirulina platensis (cyanobacteria). *European Journal of Phycology*, **33**(2), 165–171.

Quan, J., and Tian, J., 2009: Circular polymerase extension cloning of complex gene libraries and pathways. *PloS one*, **4**(7), e6441.

Rackham, O., and Chin, J. W., 2005: A network of orthogonal ribosome· mrna pairs. *Nature chemical biology*, **1**(3), 159–166.

Rao, S., Hu, S., McHugh, L., Lueders, K., Henry, K., Zhao, Q., Fekete, R., Kar, S., Adhya, S., and Hamer, D., 2005: Toward a live microbial microbicide for hiv: commensal bacteria secreting an hiv fusion inhibitor peptide. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(34), 11993.

Regot, S., Macia, J., Conde, N., Furukawa, K., Kjellén, J., Peeters, T., Hohmann, S., De Nadal, E., Posas, F., and Solé, R., 2010: Distributed biological computation with multicellular engineered networks. *Nature*, **469**(7329), 207–211.

Remington, S., 2006: Fluorescent proteins: maturation, photochemistry and photophysics. *Current opinion in structural biology*, **16**(6), 714–721.

Richmond, A., Cheng-Wu, Z., and Zarmi, Y., 2003: Efficient use of strong light for high photosynthetic productivity: interrelationships between the optical path, the optimal population density and cell-growth inhibition. *Biomolecular Engineering*, **20**(4), 229–236.

Robertson, D. E., Jacobson, S. A., Morgan, F., Berry, D., Church, G. M., and Afeyan, N. B., 2011: A new dawn for industrial photosynthesis. *Photosynthesis research*, **107**(3), 269–277.

Rott, R., Zipor, G., Portnoy, V., Liveanu, V., and Schuster, G., 2003: Rna polyadenylation and degradation in cyanobacteria are similar to the chloroplast but different from escherichia coli. *Journal of Biological Chemistry*, **278**(18), 15771–15777.

Ruder, W. C., Lu, T., and Collins, J. J., 2011: Synthetic biology moving into the clinic. *Science*, **333**(6047), 1248–1252.

Saha, R., Verseput, A. T., Berla, B. M., Mueller, T. J., Pakrasi, H. B., and Maranas, C. D., 2012: Reconstruction and comparison of the metabolic potential of cyanobacteria cyanothece sp. atcc 51142 and synechocystis sp. pcc 6803. *PloS one*, **7**(10), e48285.

Schaafsma, G., 2000: The protein digestibility–corrected amino acid score. *The Journal of nutrition*, **130**(7), 1865S–1867S.

Scheer, M., Grote, A., Chang, A., Schomburg, I., Munaretto, C., Rother, M., Söhngen, C., Stelzer, M., Thiele, J., and Schomburg, D., 2011: Brenda, the enzyme information system in 2011. *Nucleic acids research*, **39**(suppl 1), D670–D676.

Schellenberger, J., Que, R., Fleming, R. M., Thiele, I., Orth, J. D., Feist, A. M., Zielinski, D. C., Bordbar, A., Lewis, N. E., Rahmanian, S., et al., 2011: Quantitative prediction of cellular metabolism with constraint-based models: the cobra toolbox v2. 0. *Nature protocols*, **6**(9), 1290–1307.

Seaver, L., and Imlay, J., 2001: Hydrogen peroxide fluxes and compartmentalization inside growing escherichia coli. *Journal of bacteriology*, **183**(24), 7182.

Sergeyenko, T. V., and Los, D. A., 2003: Cyanobacterial leader peptides for protein secretion. *FEMS microbiology letters*, **218**(2), 351–357.

Shastri, A. A., and Morgan, J. A., 2005: Flux balance analysis of photoautotrophic metabolism. *Biotechnology progress*, **21**(6), 1617–1626.

Sohka, T., Heins, R., Phelan, R., Greisler, J., Townsend, C., and Ostermeier, M., 2009: An externally tunable bacterial band-pass filter. *Proceedings of the National Academy of Sciences*, **106**(25), 10135.

Sprinzak, D., and Elowitz, M. B., 2005: Reconstruction of genetic circuits. *Nature*, **438**(7067), 443–8.

Stocker, J., Balluch, D., Gsell, M., Harms, H., Feliciano, J., Daunert, S., Malik, K., and van der Meer, J., 2003: Development of a set of simple bacterial biosensors for quantitative and rapid measurements of arsenite and arsenate in potable water. *Environ. Sci. Technol*, **37**(20), 4743–4750.

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

Tabor, J., Salis, H., Simpson, Z., Chevalier, A., Levskaya, A., Marcotte, E., Voigt, C., and Ellington, A., 2009: A synthetic genetic edge detection program. *Cell*, **137**(7), 1272–1281. ISSN 0092-8674.

Tamsir, A., Tabor, J. J., and Voigt, C. A., 2010: Robust multicellular computing using genetically encoded nor gates and chemical 'wires'. *Nature*, **469**(7329), 212–215.

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

Toso, J., Gill, V., Hwu, P., Marincola, F., Restifo, N., Schwartzentruber, D., Sherry, R., Topalian, S., Yang, J., Stock, F., et al., 2002: Phase i study of the intravenous administration of attenuated salmonella typhimurium to patients with metastatic melanoma. *Journal of clinical oncology*, **20**(1), 142–152.

Touati, D., Jacques, M., Tardat, B., Bouchard, L., and Despied, S., 1995: Lethal oxidative damage and mutagenesis are generated by iron in delta fur mutants of escherichia coli: protective role of superoxide dismutase. *Journal of bacteriology*, **177**(9), 2305.

van der Meer, J., and Belkin, S., 2010: Where microbiology meets microengineering: design and applications of reporter bacteria. *Nature Reviews Microbiology*, **8**(7), 511–522.

van Dijl, J. M., Hecker, M., et al., 2013: Bacillus subtilis: from soil bacterium to super-secreting cell factory. *Microbial cell factories*, **12**(1), 1–6.

Vermeulen, S. J., Campbell, B. M., and Ingram, J. S., 2012: Climate change and food systems. *Annual Review of Environment and Resources*, **37**(1), 195.

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

Wang, B., Wang, J., Zhang, W., and Meldrum, D. R., 2012: Application of synthetic biology in cyanobacteria and algae. *Frontiers in microbiology*, **3**.

Wang, Y.-H., Wei, K. Y., and Smolke, C. D., 2013: Synthetic biology: advancing the design of diverse genetic systems. *Annual review of chemical and biomolecular engineering*, **4**, 69.

Waters, C., and Bassler, B., 2005: Quorum sensing: cell-to-cell communication in bacteria. *Annu Rev Cell Dev Biol.*

Weber, W., and Fussenegger, M., 2011: Emerging biomedical applications of synthetic biology. *Nature Reviews Genetics*.

Westinghouse, G., 1887: System of electrical distribution. U.S. Patent No. 373,035.

Weyer, K. M., Bush, D. R., Darzins, A., and Willson, B. D., 2010: Theoretical maximum algal oil production. *BioEnergy Research*, **3**(2), 204–213.

Wijffels, R. H., Barbosa, M. J., et al., 2010: An outlook on microalgal biofuels. *Science*, **329**(5993), 796–799.

Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N., Kunin, V., Goodwin, L., Wu, M., Tindall, B., et al., 2009: A phylogeny-driven genomic encyclopaedia of bacteria and archaea. *Nature*, **462**(7276), 1056–1060.

Xiang, S., Fruehauf, J., and Li, C. J., 2006: Short hairpin rna-expressing bacteria elicit rna interference in mammals. *Nat Biotechnol*, **24**(6), 697–702. doi:10.1038/nbt1211.

Xu, Y., Alvey, R. M., Byrne, P. O., Graham, J. E., Shen, G., and Bryant, D. A., 2011: Expression of genes in cyanobacteria: adaptation of endogenous plasmids as platforms for high-level gene expression in synechococcus sp. pcc 7002. In *Photosynthesis research protocols*, 273–293. Springer.

Ye, H., Daoud-El Baba, M., Peng, R. W., and Fussenegger, M., 2011: A synthetic optogenetic transcription device enhances blood-glucose homeostasis in mice. *Science*, **332**(6037), 1565–1568. doi:10.1126/science.1203535.

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

Zhang, S., and Bryant, D. A., 2011: The tricarboxylic acid cycle in cyanobacteria. *Science*, **334**(6062), 1551–1553.

Zhang, X., and Cai, X., 2011: Climate change impacts on global agricultural land availability. *Environmental Research Letters*, **6**(1), 014014.

Zhu, X.-G., Long, S. P., and Ort, D. R., 2008: What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? *Current opinion in biotechnology*, **19**(2), 153–159.

Zhu, X.-G., Long, S. P., and Ort, D. R., 2010: Improving photosynthetic efficiency for greater yield. *Annual review of plant biology*, **61**, 235–261.