Engineering the interface between cellular chassis and synthetic biological systems

by

Bartholomew Canton

Submitted to the Department of Biological Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Biological Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2008

Author . . . .
Department of Biological Engineering
May 23, 2008

Certified by. . . . .
Drew Endy
Cabot Assistant Professor of Biological Engineering
Thesis Supervisor

Accepted by . . . .
Alan J. Grodzinsky
Professor of Electrical, Mechanical, and Biological Engineering
Chairman, Course XX Graduate Program Committee

Thesis Committee

Accepted by . . . .

Edward F. DeLong

Professor, Department of Biological Engineering & Department of

Civil and Environmental Engineering

Chairman of Thesis Committee

Accepted by . . . .

Drew Endy

Cabot Assistant Professor of Biological Engineering

Thesis Supervisor

Accepted by . . . .

Thomas F. Knight, Jr.

Senior Research Scientist, Computer Science & Artificial Intelligence

Laboratory

Thesis Committee Member

Accepted by . . . .

Uttam L. RajBhandary

Lester Wolfe Professor of Molecular Biology

Thesis Committee Member

Accepted by . . . .

F. William Studier

Biology Department, Brookhaven National Laboratory

Thesis Committee Member

3

Engineering the interface between cellular chassis and synthetic biological systems

by

Bartholomew Canton

Submitted to the Department of Biological Engineering
on May 23, 2008, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Biological Engineering

Abstract

The aim of my thesis is to help enable the engineering of biological systems that behave in a predictable manner. Well-established techniques exist to engineer systems that behave as expected. Here, I apply such techniques to two aspects of the engineering of biological systems. First, I address the design and construction of standard biological devices in a manner that facilitates reuse in higher-order systems. I describe the design and construction of an exemplar device, an engineered cell-cell communication receiver using standard biological parts (refined genetic objects designed to support physical and functional composition). I adopt a conventional framework for describing the behavior of engineered devices and use the adopted framework to design and interpret experiments that describe the behavior of the receiver. The output of the device is the activity of a promoter reported in units of Polymerases Per Second (PoPS), a common signal carrier. Second, I begin to address the coupling that exists between engineered biological systems and the host cell, or chassis. I propose that the coupling between engineered biological systems and the cellular chassis might be reduced if fewer resources were shared between the system and the chassis. I describe the construction of cellular chassis expressing both T7 RNA polymerases (RNAP) and orthogonal ribosomes that are unused by the chassis but are available for use by an engineered system. I implement a network in which the orthogonal ribosomal RNA and the gene encoding T7 RNAP are transcribed by T7 RNAP. In turn, the orthogonal ribosomes translate the T7 RNAP message. In addition, the T7 RNAP and orthogonal ribosomes express a repressor that inhibits transcription of both the T7 RNAP and orthogonal ribosomes. As a result, the orthogonal RNAP and ribosomes are auto-generating and self-regulating. The provision of resources unused by the cellular chassis and dedicated to an engineered biological system forms the beginnings of a biological virtual machine.

Thesis Supervisor: Drew Endy
Title: Cabot Assistant Professor of Biological Engineering

Acknowledgments

This period of my life, more than any other, has made me realize how completely indispensable the people around me are. Below are the people without whom this thesis would not have happened.

Drew Endy was an inspirational advisor for the last five years. I knew I wanted to do biological engineering and Drew helped me understand what that statement actually meant. He provided the resources to do anything I wanted, as long as it was vaguely sensible. His advice on how to think, talk, and write about my research has given me the freedom to pursue that which is most important.

The Endy Lab past and present, have been the finest fellow travelers I could have hoped to have on this occasionally bumpy road. Sitting in group meeting, I'm reminded how smart and helpful this group of colleagues are. Sitting in a kayak in Vermont, I'm reminded how fun this group of friends are. Deserving special mention is Jason, co-conspirator, baymate, and irrepressible source of good humor. I've also had the invaluable assistance of two very talented undergraduates, Ania Labno and Matt Gethers, who are both on the path to great achievements.

My thesis committee have been a wonderful source of advice and support as I tried to craft a worthwhile body of work. Their knowledge and expertize and body of work have been inspirational. They are, themselves, the best examples of scientists and engineers I could have hoped to have and their advice will make those footsteps that bit easier to follow.

The people of MIT have been very special. From my colleagues and friends in BE, to the Synthetic Biology Working Group, to the fifth floor bio folk. All have been a font of fun, advice, and conversations on life and work.

I found a group of friends roughly five hours after moving to America and five years hasn't done much to change that. I've shared a lot of very special memories with my friends in BE, they have been like family to me and I'll remember our shared

time in Boston as being very special. My time here has also been made more fun by my Irish friends who have helped me not become a complete yank and my friends in CRC with whom I have shared many a mile and many a smile. My housemates have been a particular rock of friendship and support. Amy, Hillary, Jason, and Victor have been a constant source of completely insane ideas (from Aardvarks to Zippo memory sticks), target practice, and most importantly warm friendship.

Over the past five years, Reshma has become the most special colleague, friend, and girlfriend I could have hoped for. I wouldn't be graduating without her stabilizing influence, and I'd be a lot less happy.

My Mom and Dad have had the longest and deepest influence on me and it is to them that I dedicate this thesis. Their love, support and sacrifice over my 24 year education is deeply appreciated and I hope I can do something sufficiently useful for them and society to repay their efforts.

Contents

List of Figures

14

15

List of Tables

Chapter 1

Introduction

1.1 Overview

We can imagine engineering biology for the good of society. For example, engineered microbes could be deployed to cure disease [3], to produce renewable energy [112], or to more evenly distribute global food production [70]. Our limited ability to engineer biology and societal concerns about negative consequences restricts the deployment of engineered biological systems outside the lab.

Both the technological and societal obstacles standing in the way of widespread biological engineering are due in part to the difficulty of constructing systems that behave in a quantitatively predictable manner. There are several reasons why constructing predictable systems is challenging. One significant problem is the paucity of quantitative data on the behavior of simple biological objects, alone or in combination. As a result, the behavior of systems comprising even small numbers of components is difficult to predict. Additionally, the methods for assembling and testing biological systems are typically ad hoc and resource intensive. With high associated costs and low probabilities for success, the engineering of biological systems to accomplish a human-defined function is a daunting prospect [44].

The complexity and capabilities of natural self-replicating machines are inspirational to engineers used to working with non-replicating systems that have limited abilities to self-repair and adapt. However, therein lies a second problem faced by

would-be biological engineers; the complexity of a cell and the ability to change and adapt to the environment makes cells a challenging environment in which to operate engineered biological systems. Our incomplete knowledge of the behavior of natural living systems means that even a well-engineered system hosted in a natural organism is likely to exhibit complex, unpredictable behavior.

In the remainder of this chapter I describe in more detail our current abilities to engineer biology and possible strategies to improve those abilities. In the remaining chapters, I describe my own work to make routine the engineering of biological systems that behave reliably and in the final chapter I propose future steps that may get us closer to the goal of rapidly deploying engineered biological systems in service of the global community.

1.2   Background

1.2.1   The function and characteristics of today's engineered biological systems

In recent years, biological engineers have made progress in constructing synthetic gene regulatory networks and simple metabolic networks. Today's gene regulatory networks take a small number of inputs and produce an observable phenotype. For example, researchers have constructed regulatory networks that produce oscillations in the level of a reporter protein [42], switch between expression of two proteins [50], form 2-d patterns on a plate [76, 13], or regulate the cell density of a continuous culture [139]. As another example, a general information processing system has been described that is active in human kidney cells [102]. Metabolic networks use heterologous enzymes expressed in cells to synthesize a useful small molecule. For example, researchers have engineered an *E. coli* strain to produce the useful chemical, 1, 3 propanediol at an industrial scale (US patent number 6514733). Examples also exist in organisms other than *E. coli*; production of a precursor to an antimalarial treatment has been demonstrated in yeast [103]. Significant resources are being directed

towards the construction of metabolic pathways in a range of microbes with the goal of synthesizing hydrocarbons at an industrial scale [112].

What characteristics do the systems described above share? (i) The number of components (for example, proteins, genetic regulatory elements etc.) in each system is small, typically less than ten. (ii) The components of the system are ad hoc objects harvested from nature that do not conform to standards for construction or performance. (iii) In all cases, the systems are hosted by a living system of which we have an incomplete understanding. (iv) The choice of host cell is also invariably a compromise. For example, *E. coli* is often the easiest host cell to culture but lacks many essential functions for useful applications (for example, an application employing photosynthesis). As a result, biological engineers must either work with a less engineerable organism or must port essential functionality to a cell like *E. coli*.

Systems from mature engineering disciplines are qualitatively different from existing engineered biological systems. Many-component systems are common, for example integrated circuits with billions of features. The components are highly refined and standardized (see Section 1.2.2). Finally, in constrast to biological engineering, which is characterized by the use of non-optimal natural living systems due to a lack of better alternatives, all aspects of an engineered system are usually refined and tailored to the application at hand. For example, every component of an internal combustion engine has been rationally designed and optimized to serve a role in the engine as a whole. Engineers have developed the ability to construct complex systems of multiple different classes (including chemical, mechanical, or electrical). The application of strategies developed for those classes of engineered systems to biology might allow the construction of equally complex and useful biological systems.

1.2.2 Learning from experience

Most mature fields of engineering depend on catalogs of synthetic and standardized parts (see [65] and `http://www.mcmaster.com` for examples). These collections of engineered parts are produced via purposeful processes of refinement and standardization. Raw materials taken from nature are purified, modified, and used to produce

21

a constrained set of synthetic objects that meet predefined design requirements – for example, silicon is purified and processed to form wafers for integrated circuits, and iron is purified, processed, and machined to form steel nuts and bolts.

Each newly sequenced genome provides biotechnology researchers with an increasing abundance of natural genetic "parts" to consider. These natural "parts lists" include protein coding sequences, regulatory elements for gene expression and signaling, and other functional genetic elements. However, such natural parts do not behave as a would-be biological engineer might naively expect. For example, the natural parts cannot be reliably reused in combination with each other.

It is possible that a collection of well-characterized parts designed to be readily combined into many-component systems would make the construction of useful engineered biological systems easier [72]. An early proposed standard for genetically-encoded parts gained little acceptance [101]. In recent years the BioBrick standard for genetically encoded parts has been described [71] and is beginning to be widely used [119, 76] (also see `http://igem.org` and `http://partsregistry.org`). Bio-Brick standard biological parts conform to simple rules for physical composition [71] and guidelines to aid functional composition [45].

Physical composition

Physical composition is the process by which two or more objects are materially connected. One example of a process for physical composition is BioBrick Standard Assembly [71], which allows genetic parts to be assembled in an idempotent manner since the product of two assembled parts is in the same format as the component parts (see Figure 1-1). BioBrick Standard Assembly makes use of standard prefix and suffix sequences that contain specific restriction endonuclease sites (EcoRI, XbaI, SpeI, PstI, and NotI). Consequently, to be compatible with BioBrick Standard Assembly, standard biological parts must be refined to remove all BioBrick restriction endonuclease recognition sites from the part coding sequence.

Functional composition

Functional composition is the process and means of connecting the functional inputs and outputs of parts together such that the behavior of the composite object (termed
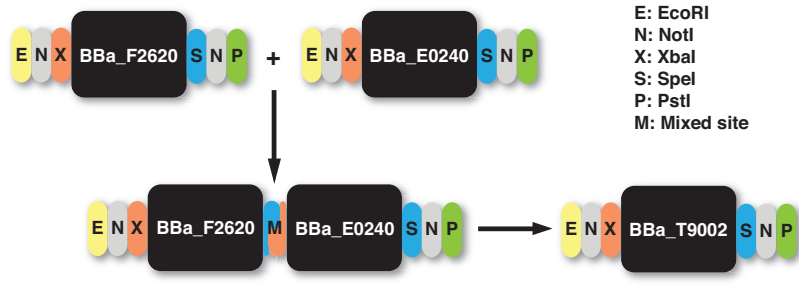
Figure 1-1: Idempotent assembly of BioBrick standard biological parts. Two parts, both bracketed by BioBrick prefix and suffix sequences, are digested with the appropriate enzymes. Following ligation, the restriction sites between the parts are destroyed and the assembled part has the same prefix and suffix sequences as the component parts. As a result, the assembled part can be recursively assembled with other BioBrick standard biological parts.

a device) is as expected, and not an emergent property of the connected objects or their interaction with the environment. To support reliable functional composition, standard biological parts must be designed to exhibit a number of properties, only some of which we currently understand. As a simple example, a standard signal carrier for device inputs and outputs supports the connection of engineered devices (see Figure 1-2). One example of a standard signal-carrier for transcription based devices is the flow of RNA polymerases along DNA, measured in PoPS (Polymerases Per Second) [45].
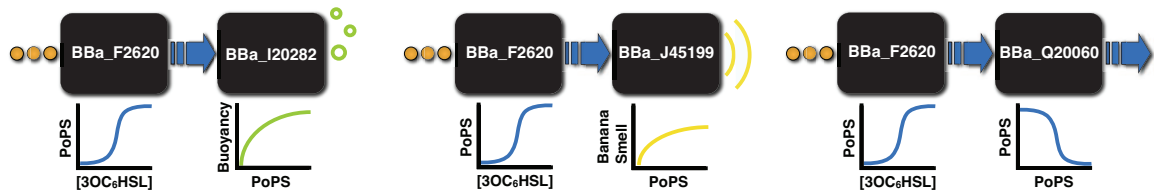
Figure 1-2: An illustration of functional composition using the PoPS common signal carrier. A device producing a PoPS output (BBa_F2620) can be connected to any device that accepts a PoPS input, making part and device reuse routine.

Building on this initial work, we next need to produce quantitative descriptions of our first-generation parts and devices. While there are many examples of quantitative descriptions of certain aspects of the behavior of synthetic biological networks [107], it is equally important that devices be designed and described to facilitate reuse. For example, a collection of bacterial promoters has recently been described in which

the strength of the promoters was measured by three different methods, including two different downstream coding sequences, with the intention of minimizing any contextual sequence or experimental dependencies in the reported promoter strengths [1]. Producing quantitative descriptions of devices that have also been designed to facilitate composition with other devices will enable the use and refinement of existing parts and devices. Furthermore, gathering data on such parts and devices will allow us to evaluate whether the principles of composition and abstraction can be usefully applied to the engineering of biological systems.

1.2.3 The cell as a chassis

In the previous section I described recent efforts to develop libraries of parts that can be combined to construct useful systems that behave in a reliable manner. However, those engineered systems do not act in isolation; rather, they depend on the presence of a host cell. For example, engineered systems depend on the host cell for spatial containment, insulation from a changing environment, and for cellular resources. Useful resources include the enzymes, substrates and cofactors needed for replication, transcription, translation, degradation, or other catabolic or anabolic processes. As such, the cell can be considered as a chassis and power supply, providing structural support, insulation, and resources to an engineered biological system (Knight, T.F. Jr., personal communication). In this thesis, I will use the term cellular chassis to refer to a host cell supporting an engineered biological system.

Many ad hoc modifications have been made to *E. coli* and other microbes to construct improved cellular chassis. *E. coli* strains with improved transformability [39] or increased genetic stability [85] relative to wild-type *E. coli* are commonly used. Modifications have been made to provide improved resources for transcription to an engineered system. The best example may be the T7 expression system [125]. The RNA polymerase (RNAP) from bacteriophage T7 is a highly active RNAP that recognizes an orthogonal set of promoters to *E. coli* RNAP. BL21(DE3) is a widely used *E. coli* strain in which T7 RNAP can be induced from a chromosomal copy of T7 gene 1 [125]. High activity and regulatable expression make T7 RNAP a useful alternative

to *E. coli* RNAP for the expression of genes in an engineered biological system. There is no widely used equivalent to the T7 expression system at the level of translation; however, a number of groups have described mutant ribosomes that recognize orthogonal ribosome binding site (RBS) sequences in *E. coli* [99, 20]. Improvements have also been made at the level of the materials used for gene expression. The availability of transfer RNA (tRNA) that recognize particular codons varies from organism to organism. *E. coli* strains (Rosetta series, EMD Biosciences; BL21-CodonPlus-RIL series, Stratagene) have been engineered to express tRNA that are rarely used in *E. coli*, allowing for increased expression of genes from organisms that have different frequencies of codon usage.

The changes described above mostly involve making incremental changes to cellular chassis. Recently, efforts have begun to make larger scale changes to bacterial genomes with the goal of producing cellular chassis that are optimized for a particular purpose. For example, an *E. coli* strain with all insertion elements removed is reported to exhibit greatly increased genetic stability and transformability [97]. Gibson and coworkers recently describe the chemical synthesis of a complete bacterial genome [51]. Knight and coworkers are developing a scientific and engineering understanding of the very simple organism, *M. florum* (Genbank record, AE017263). Other groups have proposed a minimal set of genes to encode a self-replicating cell [52]. The ability to synthesize genomes together with the specification of refined bacterial genomes suggests a future where cellular chassis are radically redesigned and optimized for the support of engineered biological systems rather than for survival and replication.

What might an ideal interface between a cellular chassis and an engineered biological system look like? (i) The interactions should be well-specified. Only with a full understanding of the interactions between the chassis and the system can we expect to engineer biological systems that always behave as expected. Questions of basic biology remain to be addressed before this goal can be achieved. (ii) The interactions should be standardized and few in number. A small number of standard interactions between the chassis and the system should make prediction of the behavior of both

the chassis and the system easier. Additionally, with a small number of standard interactions, the system is less likely to be specific to a particular cellular chassis but instead be portable between different cellular chassis. (iii) The supply of resources to the system should be stable with respect to time and variations in the environment. If an engineered system receives a constant supply of required resources, the behavior of the system is more likely to be predictable and reliable. Figure 1-3 shows a representation of an optimized, modular interface between the cellular chassis and an engineered biological system.
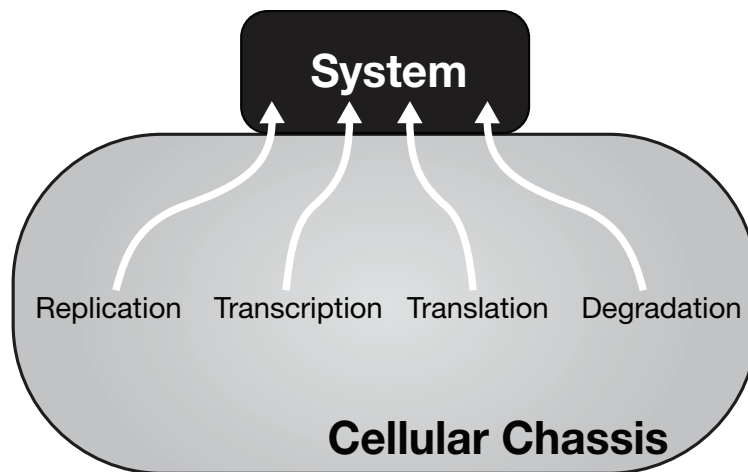
Figure 1-3: A representation of a modular chassis/system interface. The only interactions between the system and the chassis should be standardized and well-understood. The availability of resources to the system should be as invariant as possible to improve the predictability of system behavior.

An interesting example of a sophisticated interface between a system and a "chassis" is that of a virtual machine. The virtual machine is a commonly used concept in software engineering. Since software typically requires a particular platform (hardware and operating system), running software on multiple platforms is not easy. A solution is to provide a common interface on all platforms and design applications to interact with the interface rather than the underlying platform. The interface (or virtual machine) appears to the application as a complete platform, whereas in reality the virtual machine is itself an application running on the underlying platform. Applications designed to use the virtual machine can be ported to any platform for which a virtual machine exists. The analogous ability to routinely move engineered biological

systems between different cellular chassis would enable the reuse of systems and as a result spur the development of standard high-quality parts and systems rather than custom, ad hoc biological objects that cannot be easily reused.

In this thesis, I apply engineering techniques to two aspects of the engineering of biological systems that behave reliably. First, I address the design, construction, and characterization of standard biological devices in a manner that facilitates device reuse in higher-order engineered biological systems. I present a description of an exemplar biological device, a cell-cell communication receiver. Second, I address the engineering of improved cellular chassis, motivating my work with lessons learned from the description of the receiver. I describe the construction of cellular chassis expressing orthogonal RNAP and ribosomes that are unused by the cellular chassis but are available for use by an engineered biological system. The provision of resources unused by the cellular chassis and dedicated to an engineered biological system, forms the beginnings of a biological virtual machine. As a whole, my research can be viewed as the refinement of natural biological objects to produce a suite of components that will enable the routine construction of engineered biological systems that behave in a predictable manner.

Chapter 2

BBa_F2620, an engineered cell-cell communication receiver device

This chapter is based on a manuscript I wrote with an undergraduate, Ania Labno, and Drew Endy. All experiments were performed by Ania Labno and me.

2.1 Summary

The ability to quickly and reliably engineer many-component systems from libraries of standard interchangeable parts is one hallmark of modern technologies. Whether the complexity currently encountered in working with living systems will permit biological engineers to develop similar capabilities is a pressing research question. To help address this question I adapted a conventional framework for describing engineered devices in order to (i) facilitate the use and refinement of existing biological parts and devices, (ii) support research on enabling reliable composition of standard biological parts, and (iii) facilitate the development of abstraction hierarchies that simplify biological engineering. I used the adopted framework to describe one engineered biological device, BBa_F2620. BBa_F2620 is a genetically-encoded cell-cell communication receiver. Receivers convert environmental signals into intracellular signals. The resulting detailed description of the receiver is provided as a "datasheet" that facilitates reuse and serves as a template for describing many other genetically-encoded devices.

29

2.2 Introduction

While many biotechnology applications have been developed [5], the scope and scale of imaginable applications remain beyond current abilities [40, 109]. In part this is because the design and construction of engineered biological systems remains an ad hoc process for which costs, times to completion, and probabilities of success are difficult to estimate accurately [44]. Ideally, biological engineers might develop a design and construction framework that makes routine the incorporation of basic biological functions into many-component integrated genetic systems that always behave as expected. Mature engineering disciplines have developed similar frameworks by using the concept of abstraction to define sets of standardized, functional objects that can be used in combination, together with compositional rules [116] that specify how such objects should be assembled.

As described in Chapter 1, we are just starting to develop and apply composition rules and abstraction to the engineering of biology (see Figure 1-1 and 1-2. The challenge we now face is to produce quantitative descriptions of devices that have been designed to facilitate composition with other devices, thereby enabling the use and refinement of existing parts and devices. Quantitative descriptions of devices in the form of standardized, comprehensive datasheets are widely used in the electrical [65], mechanical, structural, and other engineering disciplines (see `http://www.mcmaster.com` for examples). A datasheet is intended to allow an engineer to quickly determine if the behavior of a device will meet the requirements of a system in which the device might be used. Such a determination is based on a set of standard characteristics of device behavior, which are the product of engineering theory and experience [65, 131, 19, 32]. The characteristics typically reported on datasheets are common across a wide range of device types, from sensors, to logic gates, to actuators: first, a definition of the function and interfaces of the device is provided (inputs and outputs); second, the operating context of the device is stated; third, measured characteristics describing the quantitative behavior of the device are given.

A crucial measured characteristic is the transfer function, detailing the static

relationship between device input(s) and output(s), and allowing prediction of the equilibrium behavior of composed devices. The dynamic behavior of the device is often reported so that the response time of the device can be compared to the expected timing of the overall system. Compatibility of device function with other devices or different operating conditions is important to report whenever the context in which the device operates is expected to vary. The reliability, or expected time to failure of a device, is also relevant whenever correct device performance over longer timescales is required. Finally, a description of the power and material resources consumed by a device informs the suitable choice of a power supply and resource pools for the system.

I propose adopting a similar framework for describing engineered biological devices. Despite the differences in materials and mechanisms, biological devices may often be defined with functions that are identical to the functions of electrical, mechanical, and other types of existing engineered devices. For example, biological equivalents of sensors [16, 47, 135], logic gates [58, 83, 4], and output devices [117] have all been demonstrated. Consequently, many of the characteristics found on existing device datasheets might also be useful for biological device datasheets. For example, the transfer function and dynamic behavior characteristics are directly applicable to any biological device with well-defined inputs and outputs. Compatibility of a biological device with genetic backgrounds, growth conditions, or other devices would also be useful information to biological engineers. Describing the reliability of a biological device is likely to be important but may require the invention of novel metrics due to the self-replicating and evolving nature of biological systems. For example, device failure across many generations might be measured by the number of culture doublings before a non-functional mutant becomes fixed in the population. Resource consumption, in the form of a demand for nucleotides, aminoacylated transfer RNAs, polymerases, ribosomes, and so on, is rarely reported for biological devices, yet such data would help biological engineers decide if a cellular chassis is suitable to support a particular device or combination of devices.

I applied the generic framework outlined above to develop a test datasheet for a ge-

netically encoded receiver (BBa_F2620, `http://parts.mit.edu/registry/index.php/Part:BBa\_F2620`). As outlined in Figure 2-1, the receiver builds on work by biologists [87, 88, 46] and early device engineers [16, 135, 114, 133, 2, 79, 29]; ad hoc engineered constructs, similar in function to BBa_F2620, have been used to control programmed pattern formation, cell culture density, and gene expression [13, 139]. BBa_F2620 is a composite device constructed via standard assembly [71] from five BioBrick standard biological parts: a promoter (BBa_R0040), a ribosome binding site (BBa_B0034), the luxR coding region (BBa_C0062), a transcription terminator (BBa_B0015), and the right lux promoter (BBa_R0062). Detailed descriptions for each part are freely available online via the Registry of Standard Biological Parts (`http://parts.mit.edu`). I defined the input to the receiver to be the extra-cellular level of a chemical (3-oxohexanoyl-N-homoserine lactone, $3OC_6HSL$) and the output to be a common gene expression signal, the flow of RNA polymerases along DNA (polymerase per second, or PoPS8). Hence, BBa_F2620 is a $3OC_6HSL$ to PoPS receiver. I choose to use a PoPS output for the receiver because PoPS possesses many characteristics likely to be necessary in a common signal carrier. First, it is a generic signal that can be used as the input to many other devices. Second, PoPS is a spatially-directed signal that can only pass via the DNA molecule connecting the output of an upstream device to the input of a downstream device.

Figure 2-1: (See next page) From discovery to a device. (1) Scientists identify a bioluminescent bacteria (*Vibrio fischeri*) that colonizes the light organ of several species including a squid (*Euprymna scolopes*) [88]. Bioluminescence is regulated via quorum-sensing (cell-cell density dependent communication) between individual *V. fischeri* bacteria [87]. (2) Biologists elucidate the minimal set of genetic elements encoding quorum-sensing regulated bioluminescence (the lux genes of *V. fischeri*) [46]. (3) The mechanisms and genetic sequences necessary for bacterial quorum-sensing are shared via peer-reviewed publications. Such publications are currently the major channel of communication between biologists and device engineers [88, 87, 46]. (4) Engineers construct a proof-of-principle device using a subset of the natural quorum-sensing regulatory elements [16, 135, 114, 133, 2, 79, 29]. (5) Engineers reimplement the receiver using BioBrick standard biological parts thereby
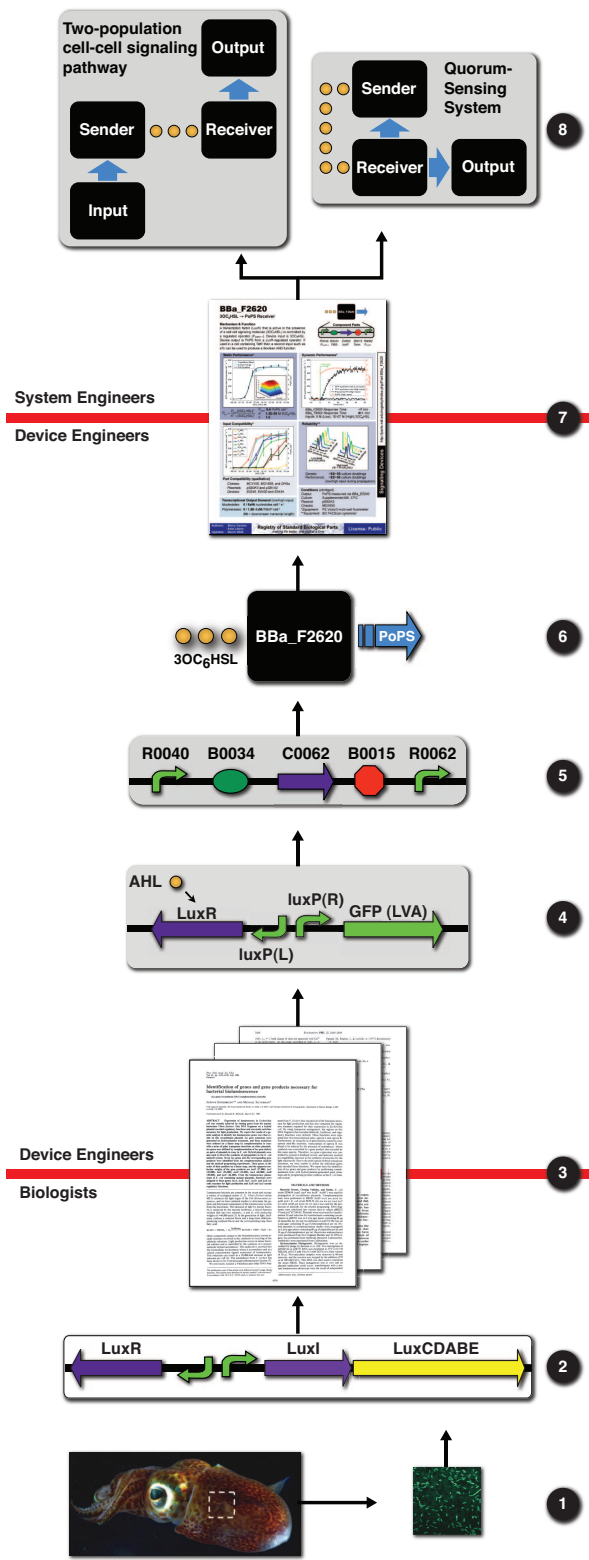
Figure 2-1: See previous page for caption.

enabling ready reuse of the device (This work). (6) Inputs and outputs to the device are defined and the component parts are no longer explicitly considered (This work). (7) The behavior of the receiver is characterized to produce a device datasheet. The datasheet forms the interface between device and system engineers, the latter of whom may choose to use the device, eliminating the need for extensive interaction between the two groups (This work). (8) The receiver is used in systems in which the device characteristics fulfill the system specification (`http://igem.org` and `http://parts.mit.edu`).

Note: If a datasheet for a suitable device does not exist, system engineers seeking a particular device may need to interact with designers of useful existing parts, or scientists who are expert in the relevant biological details, in order to develop a suitable device. While this process is slower than when a suitable characterized device is already available, it is crucial to the development of new devices and the refinement of existing ones.

2.3   Results

I used widely accessible technology to measure five characteristics that describe the behavior of the receiver under a particular set of operating conditions (described in Appendix A). In all experiments, I measured the behavior of the receiver indirectly by measuring green fluorescent protein (GFP) expression from a downstream reporter device (BBa_E0240). The combination of the receiver device and the reporter device is a composite system (BBa_T9002). I used independent experiments to parameterize a model of the behavior of the reporter device. This quantitative model allowed us to calculate the specific molecular output of the receiver from our observations of the dynamic behavior of the composite system (BBa_T9002). The detailed quantitative description of the receiver and its behavior are summarized on a device datasheet (Figure 2-2).

I determined the transfer function of the receiver across a range of $3OC_6HSL$ input concentrations (described in Appendix A). A Hill equation model with three parameters described the data well (Appendix A). The maximum, saturated output of the reporter was 490 ± 10 GFP molecules/cell/sec (uncertainties represent the 95%

BBa_F2620

3OC$_6$HSL → PoPS Receiver

BBa_F2620

Component Parts

R0040 B0034 C0062 B0015 R0062
P$_{LtetO-1}$ RBS luxR Term. P$_{lux,R}$

Mechanism & Function
A transcription factor (LuxR) that is active in the presence of a cell-cell signaling molecule (3OC$_6$HSL) is controlled by a regulated operator (P$_{LtetO-1}$). Device input is 3OC$_6$HSL. Device output is PoPS from a LuxR-regulated operator. If used in a cell containing TetR then a second input such as aTc can be used to produce a Boolean AND function.

Static Performance*

GFP synthesis rate (molecules cell^{-1} s^{-1})
PoPS cell^{-1}
[3OC$_6$HSL] (M)

- Population Mean
- Colony Range
- Hill Equation
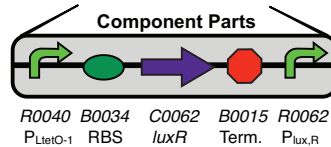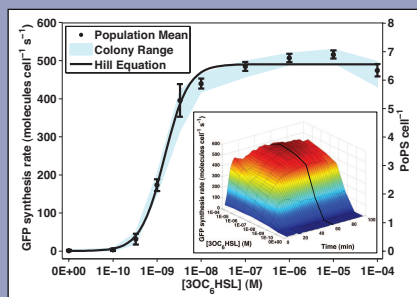
$$P_{out} = \frac{P_{max} [3OC_6HSL]^n}{K^n + [3OC_6HSL]^n}$$

P$_{max}$: **6.6** PoPS cell^{-1}
K: **1.5E-09** M 3OC$_6$HSL
n: **1.6**

Dynamic Performance*

GFP synthesis rate (molecules cell^{-1} s^{-1})
PoPS cell^{-1}
Time (min)
+ 3OC$_6$HSL

- GFP synthesis rate (Low Input)
- GFP synthesis rate (High Input)
- Polynomial Fit (High Input)
- PoPS (High Input)
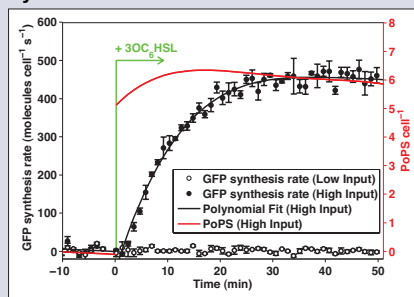
BBa_F2620 Response Time: **<1** min
BBa_T9002 Response Time: **6±1** min
Inputs: 0 M (Low), 1E-07 M (High) 3OC$_6$HSL

Input Compatibility*

GFP synthesis rate (molecules cell^{-1} s^{-1})
PoPS cell^{-1}
[AHL] (M)

- C$_4$HSL
- C$_6$HSL
- 3OC$_6$HSL
- C$_7$HSL
- C$_8$HSL
- 3OC$_8$HSL
- C$_{10}$HSL
- C$_{12}$HSL
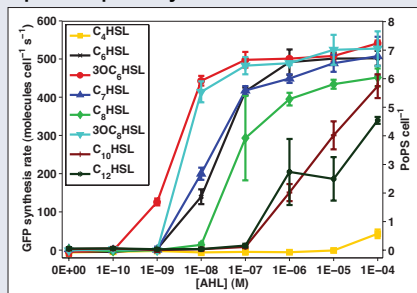
Part Compatibility (qualitative)

Chassis: MC4100, MG1655, and DH5α
Plasmids: pSB3K3 and pSB1A2
Devices: E0240, E0430 and E0434

Transcriptional Output Demand (low/high input)
Nucleotides: **0 / 6xNt** nucleotides cell^{-1} s^{-1}
Polymerases: **0 / 1.5E-1xNt** RNAP cell^{-1}
(**Nt** = downstream transcript length)

Reliability**

GFP (arbitrary units)
Doublings
Low Input (0 M 3OC$_6$HSL)
High Input (1E -7 M 3OC$_6$HSL)
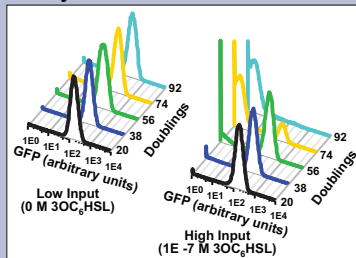
Genetic: **>92/>56** culture doublings
Performance: **>92/>56** culture doublings (low/high input during propagation)

Conditions (abridged)
Output: PoPS measured via *BBa_E0240*
Culture: Supplemented M9, 37ºC
Plasmid: pSB3K3
Chassis: MG1655
*Equipment: PE Victor3 multi-well fluirimeter
**Equipment: BD FACScan cytometer

http://parts.mit.edu/registry/index.php/Part:BBa_F2620

Signaling Devices

Authors: **Barry Canton**
Ania Labno
Updated: **March 2008**

Registry of Standard Biological Parts
making life better, one part at a time

License: Public

Figure 2-2: An exemplar data sheet that summarizes current knowledge of the behavior of the receiver (BBa_F2620). The structured description is designed to support rapid reuse of the receiver without the need to contact the scientists upon whose discoveries the receiver is based, or the engineers of the receiver itself. A general description of the device is included along with a condensed summary of relevant performance characteristics. The description of the receiver is also available in electronic format (BBa_F2620, `http://parts.mit.edu/registry/index.php/Part:BBa_F2620`). A glossary for the datasheet is provided in Appendix B.

confidence interval for the parameter). From the measured GFP synthesis rate, I estimated a maximum output from the receiver of 6.6 ± 0.3 PoPS/cell. The minimum observed output was determined to lie between 0 and 3 GFP molecules/cell/sec corresponding to device output of ~ 0 PoPS/cell. Given such a low minimum observed output, I did not include a basal PoPS output in the model describing the output of the receiver. The receiver switch point, the input required for half-maximal output, is $1.5E-9 \pm 3E-10$ M $3OC_6HSL$. The Hill coefficient describing the steepness of the transition from low to high output was determined to be 1.6 ± 0.4. The population distribution was monovariate at all input levels (not shown). Given that the output of the receiver varies over two logs of input concentration, the receiver might be used either as an analog device with a graded output or as a digital device (with the high and low output levels still to be defined).

I determined the dynamic response of the receiver by quantifying the time-dependent increase in the rate of fluorescent protein synthesis following a step increase in input level from 0 to 1E-7M $3OC_6HSL$ (Appendix A). Assuming a first-order linear response with time delay, I calculated a response time constant of 6 ± 1 min and a delay of 1.5 ± 0.5 min. Independent experiments demonstrated that the observed dynamic response is largely due to the maturation rate of GFP (described in Appendix A). The model of the reporter device was used to calculate the time-dependent response of the receiver given the observed response of the reporter. Using this method, I calculated a response time for the receiver of less than 1 min.

I measured receiver input specificity, which is the ability of the receiver to distinguish between its cognate input signal and similar chemical signals that might also be used in composite systems containing the receiver. Thus, input specificity describes the compatibility of the receiver within a particular set of related devices. I measured the response of the receiver to input signals carried by different acyl-homoserine lactones, both lacking the 3-oxo moiety and varying in side-chain length (described in Appendix A). The receiver responds to $3OC_6HSL$ and acyl-homoserine lactones with similar length side chains. Any device that produces one of this subset of like acyl-homoserine lactones may be used to send a signal to the receiver. The compounds

with the shortest and longest side chains produce very weak device responses, suggesting that the receiver could be used independently in parallel with other devices that respond to these compounds. The datasheet (Figure 2-2) also lists the compatibility of the receiver within a range of genetic backgrounds, output devices, and plasmids.

I measured the evolutionary reliability of the receiver coupled to the reporter device by following performance of the receiver as a function of culture doubling at low input levels (Appendix A). Since evolutionary reliability is known to be dependent on levels of recombinant protein expression [53], I measured the reliability of the receiver at low input levels so that GFP expression from the reporter device would be negligible. Receiver performance remained constant over 92 culture doublings. For comparison, I also measured the reliability of the composite system (BBa_T9002) at high input levels. Consistently, at high input levels, more than half the cells in the population were non-performing within 74 culture doublings. Sequence analysis of non-performing mutants indicated that system failure results from a deletion between DNA sequences that are repeated in both the receiver and the reporter devices. Additional experiments suggested that individual cells in our long-term stock contained two plasmid populations, a majority population encoding the intact composite system and a spontaneously arising minority population encoding the mutant system (Appendix A). The failure observed here is an emergent behavior specific to the combination of the receiver and reporter devices. Emergent behavior might be avoided by the development of appropriate design rules. For example, when system operation across many culture doublings is required, repeat sequences sufficient in length and proximity to promote deletion events should be avoided.

I computed the output demand of the receiver using the observed rates of downstream protein synthesis (Appendix A). The transcriptional output demand depends both on the output of the receiver and on the length of the transcript encoded by the downstream device (Appendix A). At low inputs, the output of the receiver is approximately zero and so places a negligible demand on the host cell. At high inputs, the output of the receiver requires 6.7 x Nt nucleotides/cell/sec and 0.15 x Nt polymerases/cell, where Nt is the number of nucleotides in the transcript being

produced from the output of the receiver. I did not measure the cellular resources required to produce the LuxR protein (BBa_C0062), an essential component of the receiver whose expression places an additional basal demand on the cell.

The natural biological system on which the design of the receiver is based has been used to produce other, functionally similar devices [16, 135, 114, 133, 2, 79, 29]. I compared the behavior of our receiver to these earlier systems (none of which were constructed from BioBrick standard biological parts) in order to begin to evaluate whether or not the performance of the receiver might depend on external factors such as host cell genetic background, culture conditions, or laboratory environment (Appendix C). None of the prior studies reported all the characteristics by which the receiver has been described here. What comparisons could be made suggested that the receiver switch point and response time are insensitive to host cell genotype or growth conditions but that the input compatibility is sensitive to host cell genotype or other variables. Importantly, two studies reported device switch points that are 100-fold or more different from all other studies. This variation is likely explained by the use of different isolates as sources of genetic materials (Appendix C); the amino acid sequences of the LuxR proteins used in these two studies differ by 25% from that used in the remaining studies.

2.4 Discussion

Here, I developed a generic framework for defining and describing standard biological devices in order to support the reuse and refinement of many devices. To test the utility of our framework, I used relatively well-understood biological mechanisms to design a device that converts the extra-cellular level of $3OC_6HSL$ to PoPS, a common intracellular signal carrier that can be accepted as input by many standard biological devices. I constructed the receiver from five standard biological parts. I used a reporter device also encoded by standard biological parts to measure the quantitative and dynamic behavior of the receiver. Three aspects of our work enable easy reuse of the receiver: (i) our use of standards that support the reliable physical composition

of genetic parts, (ii) a device design that produces an output signal that is a common signal carrier, and (iii) our extensive and quantitative device description. As evidence, while this manuscript was in preparation, I made freely available the DNA encoding BBa_F2620 and its accompanying datasheet via the Registry of Standard Biological Parts (`http://parts.mit.edu/`). Already, eighteen higher-order systems incorporating the receiver have been successfully assembled and contributed back to the Registry by teams in the International Genetically Engineered Machines Competition (`http://igem.org`).

The component parts of the receiver can be adapted to serve functions other than that chosen here. For example, the behavior of the receiver could be modified in a predictable manner by choosing, as input, one of the acyl-homoserine lactones similar to $3OC_6HSL$ to which I have demonstrated that the receiver responds. As a second example, in a host cell that constitutively expresses Tet repressor, the receiver can perform a logical AND operation, producing a high output only in the presence of $3OC_6HSL$ and anhydrotetracycline (aTc). As a final example, removing the promoter regulating the transcription of the LuxR coding region would produce a device that has both a PoPS input and a $3OC_6HSL$ input. The resulting three-terminal device could be used to perform an AND operation, or as a $3OC_6HSL$-dependent PoPS amplifier/attenuator. These examples highlight that there is value in considering the internal components, inputs, and outputs of the receiver in detail to design novel devices. However, such value is gained at the expense of the convenience afforded by choosing a well-described "black-box" device, such as the BBa_F2620 receiver.

Looking forward, much additional work is needed to make routine the engineering of many-component biological systems that behave as expected [108]. For example, the framework for describing device behavior introduced here, or an improved framework, should be applied to describe many devices and device combinations. When characterizing combinations of devices, special attention should be paid to combinations that fail to produce the behavior predicted given descriptions of the individual devices. Careful characterization and analysis of such emergent behaviors is needed to support the development of design rules that prevent interactions between devices

other than via the defined device inputs and outputs (such as the spontaneous selection for a deletion within the composite system, BBa_T9002). As a second example, standard input and output signal levels might be defined so that any two devices, when connected, would be well-matched. Understanding whether desired device behaviors (such as standard signal levels) can be best engineered via directed evolution, rational engineering, or a combined approach [29, 30, 138, 59] will help researchers to produce well-behaved devices more quickly.

Finally, since the receiver can be used in many systems and because I hope to promote the collaborative development and unfettered use of open libraries of standard biological parts and devices, all of the information describing the receiver is freely available via the Registry of Standard Biological Parts, as mentioned above. I encourage researchers to contribute improvements to the design and description based on experiences with the operation of the receiver (or other parts and devices) directly to the Registry. Ultimately, device descriptions such as that presented here should be available online in a machine-readable format that will enable the computer-aided design of many component engineered biological systems.

Chapter 3

Orthogonal transcription and translation in combination (VM1.0)

3.1 Introduction

In Chapter 1, I described past efforts to make natural living systems better chassis to support engineered biological systems. Future efforts to engineer cellular chassis should be responsive to the requirements of the systems being constructed by biological engineers. The receiver, BBa_F2620 (Chapter 2), is typical of the type of devices that biological engineers might use in the future to construct higher-order systems. Observations of the interactions between the receiver and the cellular chassis might suggest future directions for the engineering of cellular chassis.

A high output from the receiver causes a reduction in the growth rate of the cellular chassis (Section A.6). As a result, a careful prediction of how a system using the receiver will behave as receiver output changes must include consideration of changes in the growth rate of the cellular chassis due to the receiver. Engineering systems with predictable behavior would be easier if changes in system behavior did not affect the behavior of the cellular chassis. For example, the engineering of a

complex system such as a car would be more challenging if the torque provided by the motor varied with the volume of the radio. Also note that the change in growth rate determines, in part, the genetic reliability of the receiver. Were there no reduction in growth rate with increasing receiver output, the functional copy of the receiver would likely persist in the population longer (A.6).

What causes the observed coupling between the receiver and the cellular chassis? The input molecule, $3OC_6HSL$, is reported to have little if any interaction with the cell [128]. The maturation of GFP is also not expected to have a phenotypic effect on the chassis. As a result, the coupling between receiver and chassis is likely due to consumption of gene expression resources by the receiver and reporter. The resources necessary for gene expression (RNAP, ribosomes, nucleoside triphosphates (NTPs), aminoacylated tRNA etc.) are also used by the cell to supply systems necessary for cell growth. Hence, a consumption of resources leads to a reduction in chassis growth rate [53].

The sharing of resources between the receiver and the chassis was not made for reasons of sound engineering design, but rather due to convenience and a lack of alternative resource supplies. We can readily imagine several scenarios where extensive sharing of resources between engineered biological systems and cellular chassis might be undesirable. (i) When engineering a many-component system, where by virtue of the number of components, the system must consume a large fraction of chassis resources. Such a scenario is very plausible; the receiver and reporter described in Chapter 2, comprising approximately 40 copies of a three-gene network, places a noticeable demand on the cellular chassis (inferred from the observed change in growth rate). However, biological engineers might wish to construct systems with orders of magnitude more components than the receiver-reporter system. (ii) When the growth rate-dependent levels of cellular resources are insufficient for the system. For example, a system might need an increasing availability of resources as cell density increases whereas natural microbes tend to reduce the production of resources as cell density increases [18]. (iii) When it would be desirable to move a system between different cellular chassis. For example, the behavior of a system is likely to change significantly

when moved to a cellular chassis with a different availability of the resources needed by the system [10].

The arguments presented above motivate reducing the coupling between the cellular chassis and the system. Decoupling the behavior of different systems is widespread in biology. Eukaryotic cells make use of organelles to separate the components of different systems. For example, the nucleus contains the components necessary for the majority of cellular DNA replication and transcription, thereby spatially separating those processes from the rest of the cell. Since many biological processes occur in a shared space (e.g. the cytosol), biochemical specificity or orthogonality, is an alternative strategy for decoupling systems. For example, *E. coli* uses σ-factors that confer specificity to the general purpose RNAP [57]. As a related example, bacteriophage T7 makes uses of an RNAP and promoter sequences that are orthogonal to the host equivalents [24, 38].

As an example of engineered biological orthogonality, researchers have developed mutant *E. coli* ribosomes that recognize RBS sequences orthogonal to those recognized by wild-type (wt) *E. coli* ribosomes [62, 20, 99]. In the earlier work, de Boer and coworkers rationally designed a mutation to the anti Shine-Dalgarno (aSD) sequence of the 16s rRNA responsible for ribosome binding site (RBS) recognition so that functional ribosomes could be assembled that recognized an RBS sequence orthogonal to those recognized by *E. coli* ribosomes. In more recent work, Rackham and Chin used a positive and negative selection scheme to select ribosome-RBS pairs that were highly active but did not interact with wt *E. coli* ribosomes or RBS [99].

As described in Section 1.2.3, there are many interactions between a system and a cellular chassis. In this chapter and the next, I choose to begin to decouple usage of transcription and translation resources by a cellular chassis and an engineered biological system with the goal of making the behavior of the system and the cellular chassis more predictable. I focus on transcription and translation for several reasons. First, transcription and translation are critical processes for almost any biological system. Second, our understanding of transcription and translation is as complete as our understanding of any important molecular biological process. Third, there

are well documented examples of the emergent behavior that results from the shared use of transcriptional and translational resources by engineered systems and cellular chassis [53].

Ideally, we might duplicate all the resources associated with transcription and translation so that protein synthesis for the engineered system and chassis were not directly coupled. Note that an unavoidable coupling would still exist since central metabolism would be responsible for providing the energy, materials and precursors needed for both protein synthesis systems. It remains to be seen whether useful decoupling can be achieved despite a shared central metabolism. Here, I address the problem of producing a second population of RNAP and ribosomes that are not used by the chassis itself. If this second population of RNAP and ribosomes are unused by the chassis, they can be considered orthogonal to *E. coli* RNAP and wt *E. coli* ribosomes. Such orthogonality can be achieved by making use of an RNAP that does not recognize *E. coli* promoters and ribosomes that do not recognize *E. coli* RBS.

In this chapter, I build on past work to develop and characterize the behavior of genetic parts necessary to implement orthogonal RNAP and ribosomes in *E. coli*. Then, I demonstrate the use of orthogonal transcription and translation in combination to express a reporter protein. Finally, I compare the capacity of different combinations of transcription and translation to express a reporter protein. Figure 3-1 illustrates both the current sharing of RNAP and ribosomes and a potentially improved scenario where two, orthogonal populations of RNAP and ribosomes are provided by the cellular chassis.

3.2 Results

3.2.1 Orthogonal transcription

The T7 system for overexpression of target genes is widely used [56]. The combination of T7 RNAP and T7 promoters fits my criteria for an orthogonal RNAP since (i) T7 RNAP selectively recognizes T7 promoter sequences that are unrelated to *E.*
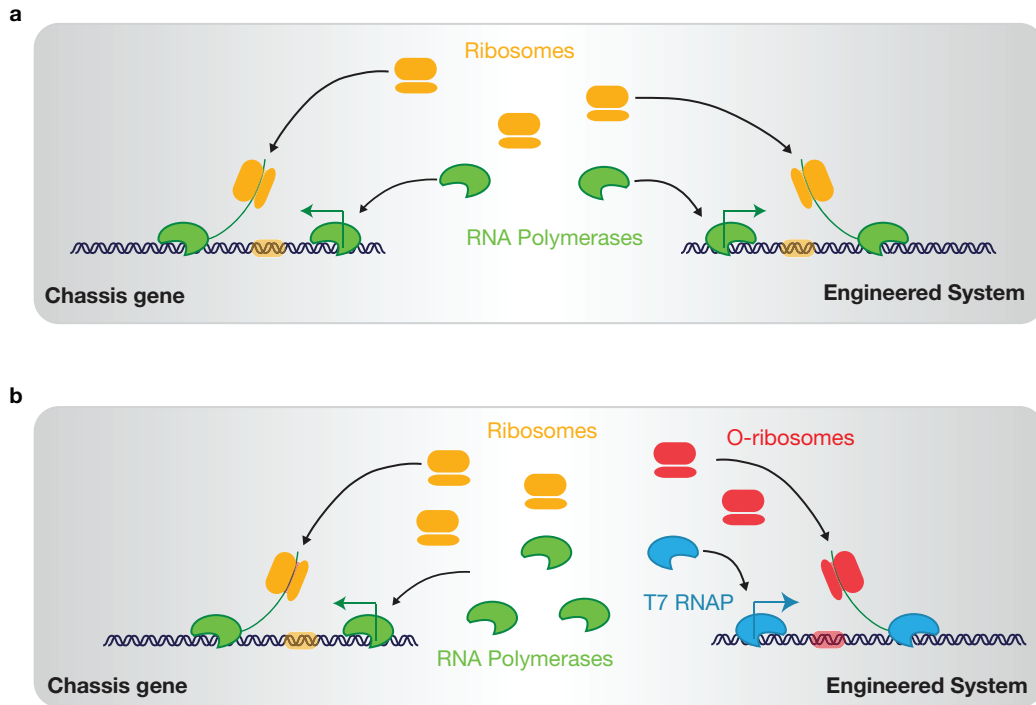
Figure 3-1: The provision of orthogonal RNAP and ribosomes can be used to reduce the number of shared components between an engineered system and the cellular chassis. (a) An engineered system designed to use *E. coli* RNAP and ribosomes draws from the same resource pool as the genes and systems that comprise the cell itself. Depending on the scale of the system, depletion of RNAP and ribosomes may occur, resulting in emergent behavior of the system and the cellular chassis. (b) If the cellular chassis contains a second, orthogonal, population of RNAP and ribosomes, systems can be engineered that do not draw from the RNAP and ribosome pool used by the cell. Expression of T7 RNAP [125] and O-ribosomes [99] in *E. coli* represent a potential implementation of a cellular chassis with two orthogonal populations of RNAP and ribosomes.

coli promoter sequences [24] and (ii) *E. coli* RNAP does not recognize T7 promoter sequences, as evidenced by the ability to maintain highly toxic genes in *E. coli* with T7 promoters upstream in the absence of T7 RNAP (F. William Studier, personal communication). I began by implementing the components of the T7 expression system (promoters and RNAP) in BioBrick format to facilitate the construction of networks involving those components.

The T7 RNAP is known to be highly active (180 nucleotides/s) [82] and the consensus T7 promoter initiates transcription at a high rate [63]. For many applications where maximizing protein expression is not the goal, lower rates of transcription initiation might be sufficient while also placing a less severe demand on the cellular chassis. One strategy for reducing the transcriptional activity of the T7 system is to use weaker promoters, although other alternatives exist [82, 141]. The strengths of all T7 promoters with a single base mutation from the consensus sequence have been measured relative to the conensus promoter sequence *in vitro* [64]. I chose a subset of the mutant library that should provide a range of T7 promoter strengths (see BBa_R0085, BBa_R0180-3 in Table H.1). The promoters were constructed in BioBrick format (see Figure 1-1). The construction process is described in Appendix D.

To confirm that the BioBrick T7 promoters were active and did produce a range of transcription initiation rates, I assembled the promoters with a GFP generator (a device that accepts a PoPS input and produces GFP). The configurations of the resulting reporters are described in Table H.2. The reporters can be used to measure the activity of the T7 promoters indirectly, by observing GFP expression.

I transformed the T7 promoter reporters into BL21(DE3) and used a multi-well fluorimeter to record the fluorescence of the cultures as a function of cell density after addition of 0.4 mM IPTG to the cultures. IPTG induces expression of T7 RNAP from the chromosome of BL21(DE3). The data is shown in Figure 3-2. Control cultures lacking IPTG in the culture media showed little increase in fluorescence as a function of cell density. In the presence of IPTG all cultures exhibited high levels of fluorescence. For the first five hours after addition of IPTG, there was a positive

correlation between the predicted strengths of the T7 promoters and the observed rate of increase in fluorescence. Interestingly, later in the growth of the cultures, the T7 promoters predicted to be weaker by Imburgio and coworkers continued to show increasing fluorescence while the stronger promoters exhibited a plateau in fluorescence accumulation. The observed result may be due to the high demand that a strong T7 promoter places on the cellular chassis. The increased demand due to the stronger promoters may have depleted some cellular resource thereby limiting the amount of GFP synthesis that could occur later in the experiment. Growth curves for the corresponding cultures are shown in Figure 3-3 and were measured as described in Appendix D.
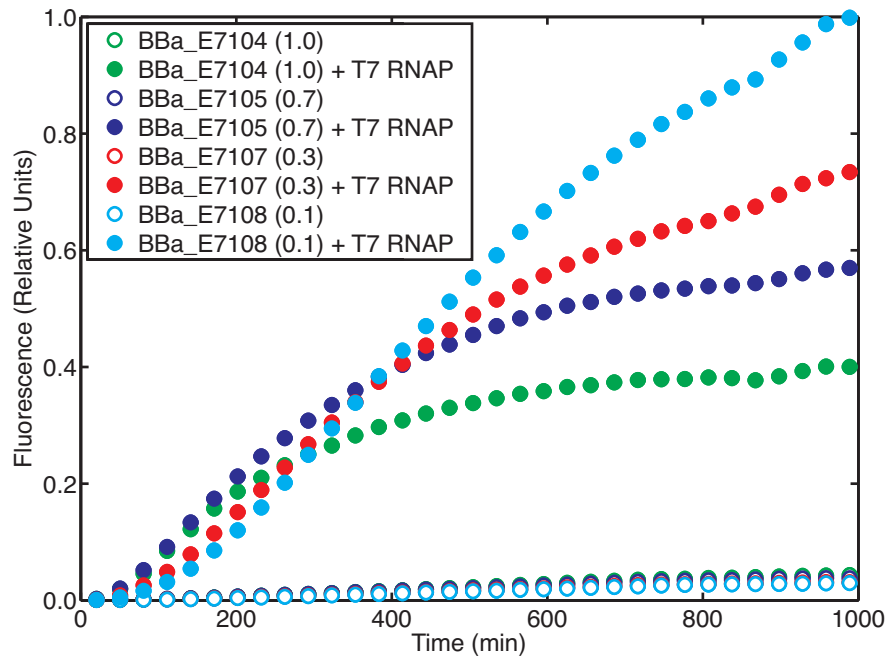
Figure 3-2: GFP expression regulated by BioBrick T7 promoters. Cultures each contained one of four synthetic T7 promoters in BioBricks format regulating a GFP generator. The four reporters are identified by BioBrick number (Table H.2. The expected relative strength of the T7 promoter in each reporter is given in parentheses. The relative strengths are predicted based on the measurements by Imburgio and coworkers [64]. Initially, increase in fluorescence is correlated with predicted promoter strength. However, the weaker T7 promoters maintain a high level of GFP expression for longer.

47
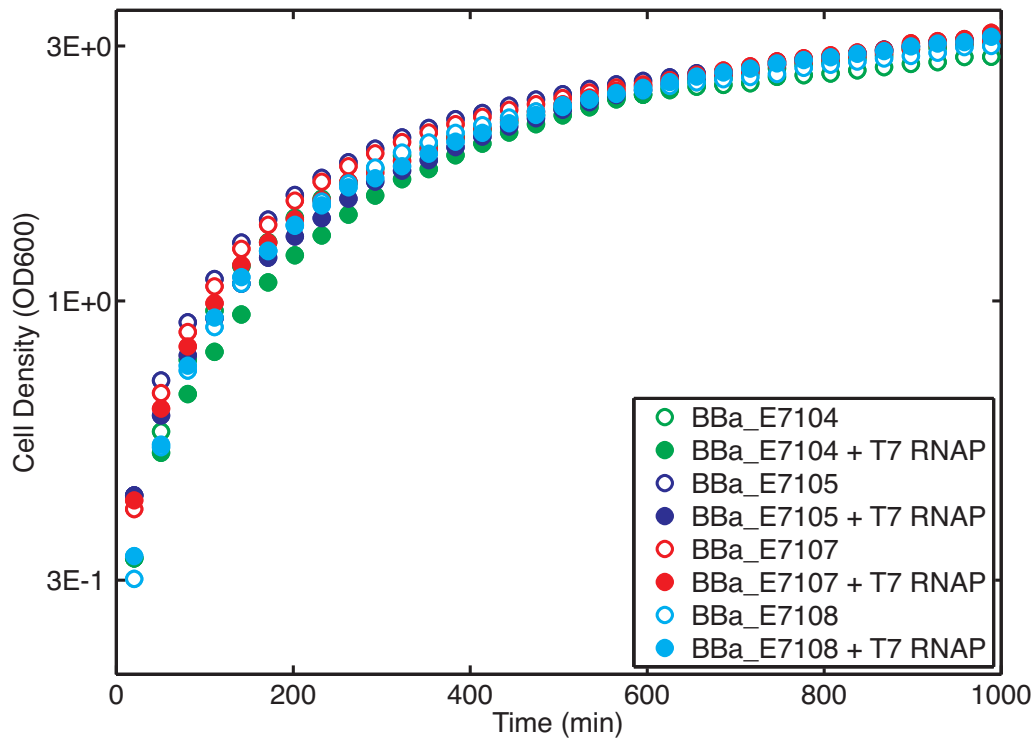
Figure 3-3: Growth curves for strains expressing T7 RNAP and carrying a T7 promoter-regulated GFP reporter. The growth curves shown correspond to the samples shown in Figure 3-2.

From these data I concluded that the T7 promoters were active in BioBrick format and that under the conditions tested the single basepair mutants described by Imburgio and coworkers did show a range of activities relative to the consensus promoter sequence (green data points in Figure 3-2). The data also suggested that the weakest T7 promoter (light blue data points in Figure 3-2), which was expected to be ~10% as strong as the consensus promoter, was still a highly active promoter that would be sufficiently strong for most applications while presumably placing less demand on the cellular chassis than the consensus promoter sequence.

3.2.2 Orthogonal translation

The orthogonal ribosome system described by Brink and coworkers appeared suitable for my purposes; a mutation in the aSD sequence at the 3' end of the 16s rRNA causes ribosomes assembled from that rRNA to recognize non-natural SD sequences in the RBS of messenger RNA (mRNA). The resulting ribosomes ("O-ribosomes" to use the terminology of Rackham and coworkers) are orthogonal in the sense that they should specifically recognize a set of RBS sequences not recognized by wt *E. coli* ribosomes and should not recognize *E. coli* RBS sequences (Figure 3-4). Note that only the 16s rRNA of the orthogonal ribosomes is mutant; the remainder of the ribosome is made up of wild-type components. Ideally, all ribosomal RNA and proteins would be orthogonal; however, there is reason to think that orthogonal rRNA may be sufficient for my purposes. The majority of regulation of ribosome synthesis happens at the level of transcription of the rRNA; ribosomal proteins are synthesized as needed when free, processed rRNA is present [18]. Hence, the levels of the orthogonal ribosomes should be determined for the most part by the transcriptional regulation of the rRNA rather than being coupled to the synthesis of wild-type components such as the ribosomal proteins.

I obtained a plasmid (pCH1497-ASD1, see Table H.5) carrying a copy of the *E. coli* rrnB ribosomal operon from Christopher Hayes. The operon carried a mutation at the 3' end of the *rrsB* gene (encoding the 16s rRNA) to carry the aSD sequence described by Brink and coworkers (see Table H). An arabinose-inducible P_{BAD} promoter
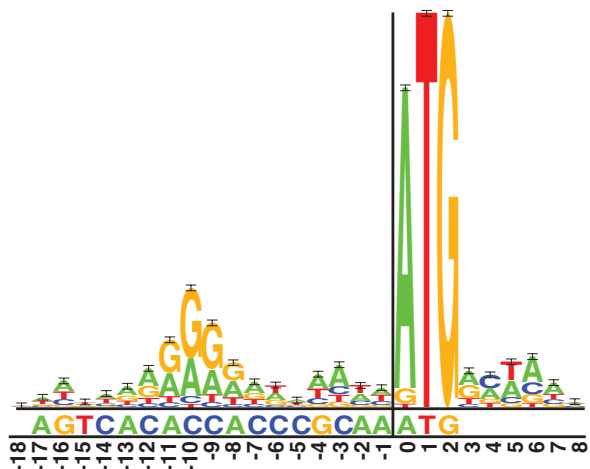
Figure 3-4: A comparison of *E. coli* and Rackham O-RBS sequences. The upper sequence logo shows the frequency of the base frequency at each location relative to the translation start site (labeled 0) for 149 *E. coli* RBS [115]. The lower sequence is the mRNA-A RBS sequence reported by Rackham & Chin [99]. For the sequence logo, the height of the letters corresponds to the degree of conservation at that location, measured in information (where the vertical black bars represent the maximum information, 2 bits). A perfectly conserved base will have a height of two, the height of the stack at a location that is completely random will be zero.

controls transcription of the mutant rrnB operon. I refer to the arabinose-inducible ribosomal rRNA cassette as an O-ribosome generator. The plasmid is derived from pACYC184 [26] (15-30 copies per cell, tetracycline resistance marker).

To test whether the O-ribosome generator was capable of expressing O-ribosomes, I constructed reporters that should express GFP only in the presence of O-ribosomes. The reporters (Table H.2) used some pre-existing parts: a synthetic TetR-repressible promoter (BBa_R0040, see Table H.1), the GFPmut3b gene (BBa_E0040, [31]) and a synthetic transcriptional terminator (BBa_B0015, see Table H.4). I designed RBS sequences (Table H.3) that included a mutant SD sequence complementary to the aSD sequence in the mutant rrnB operon carried by pCH1497-ASD1. Early RBS designs did not behave as expected (data not shown). Poor behavior was due to several problems, an early RBS (BBa_B0036) resulted in a cryptic *E. coli* RBS being formed across the junction between mixed connective site and the 5' end of the RBS sequence. A second RBS (BBa_B0037) formed a 7bp hairpin loop that appeared to

prevent any detectable GFP expression. A final RBS (BBa_B0070) was designed to not include any sequences that might be recognized by *E. coli* RBS and also to have little secondary structure when upstream of the GFP coding region. A reporter using that RBS (BBa_E70102, Table H.2) appeared to express GFP only when O-ribosomes were induced from the O-ribosome generator (See Figure 3-5). From the growth curves in Figure 3-6, it is clear that expression of the Brink orthogonal ribosomes was quite toxic to *E. coli* BL21(DE3).
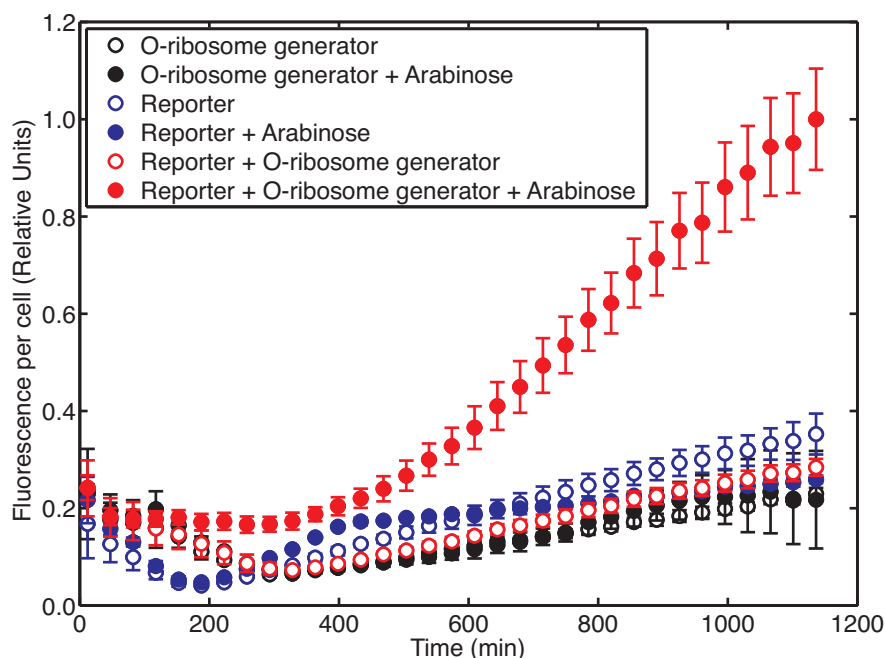
Figure 3-5: GFP expression regulated by a Brink O-RBS. Cultures contained either the O-ribosome generator, a reporter device (BBa_E70102) with an O-RBS or both. Data for cultures with 0.2% arabinose added to induce O-ribosome synthesis is shown with filled circles, empty circles represent measurements of cultures not containing arabinose. The expression of O-ribosomes and the presence of the reporter is necessary for fluorescence per cell to rise above background levels. Data points for the reporter and O-ribosome cultures with or without arabinose represent the average of three independent cultures, each measured in triplicate. All other data points represent the average of three replicates of single independent cultures. For all datapoints, the error bars represent the 95% confidence interval in the mean.

While I was debugging the Brink ribosomes and the reporter devices to show that the Brink ribosomes were functional, a new set of O-ribosomes was reported
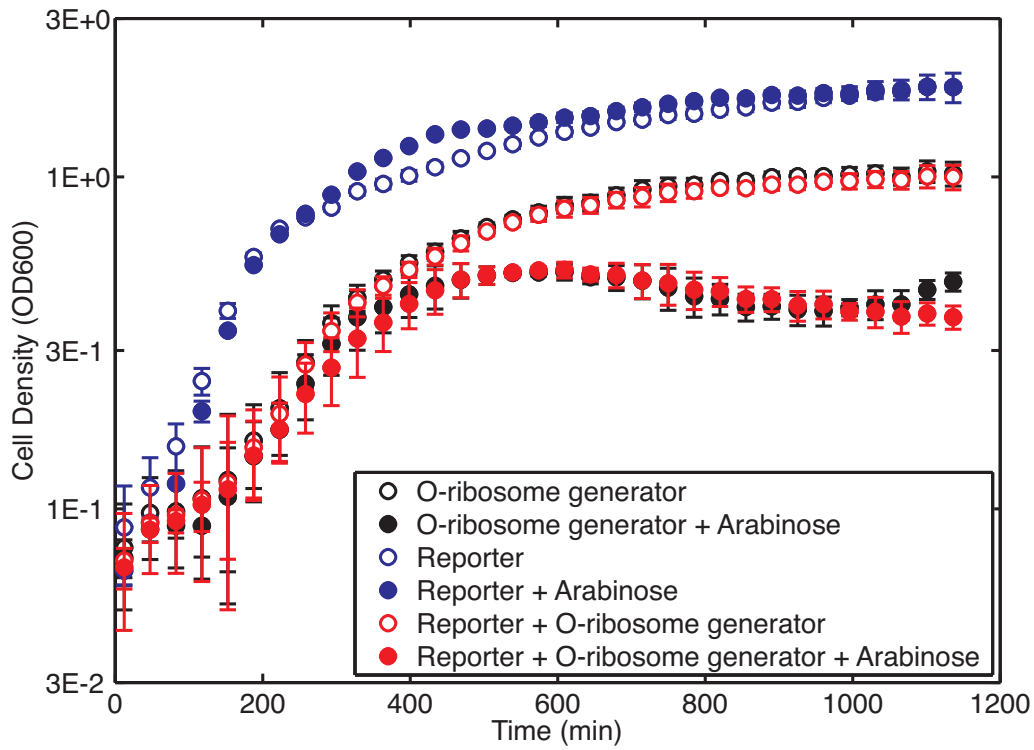
Figure 3-6: Synthesis of Brink O-ribosomes significantly affects culture growth. The samples correspond to those described in Figure 3-5. The data shows that carrying the O-ribosome plasmid causes a reduction in culture growth (red and black open circles). The observed decrease in cell density after 500 min for cultures expressing O-ribosomes may indicate cell lysis (red and black filled circles).

[99]. In addition to the improved methodology for finding O-ribosome/O-RBS pairs (described in Section 3.1), the new ribosomes were reported to exhibit much less toxicity to the cellular chassis than the earlier Brink O-ribosomes [99]. The Rackham and Chin O-ribosomes were randomized in two locations, the aSD sequence and also a distant sequence (at locations 721-722 in the 16s rRNA). For simplicity, I chose one of the O-ribosome sequences (rRNA-4 [99]) that required mutations only to the aSD, since that O-ribosome was reported to be as active as the others [99]. I modified the pCH1497-ASD1 plasmid (Appendix D) so that the resulting plasmid would have the 16s rRNA sequence matching rRNA-4 (see Table H). I named the resulting plasmid pCH1497-rRNA4.

Again, I designed GFP reporters that should express GFP only in the presence of O-ribosomes (BBa_E70201-3, see Table H.2). In light of my earlier difficulty in constructing a synthetic RBS that was efficiently translated by Brink O-ribosomes, I constructed three synthetic RBS sequences designed to be recognized by the Rackham O-ribosomes. The three RBS (BBa_B0072-4) comprised a six, nine, or eleven nucleotide sequence taken from mRNA-A (see Table 3.1) [99]. Note that BBa_B0072 and BBa_B0073 used the same spacing between RBS and start codon as used by Rackham & Chin whereas BBa_B0074 increased that spacing.

Table 3.1: The RBS-start codon region of the Rackham O-RBS and three BioBrick O-RBS.

Identifier[a]	Sequence[b]	Reporter[c]
mRNA-A	agtcaCACCACccgcaaatg	na
BBa_B0072	tactagagCACCACtactagatg	BBa_E70201
BBa_B0073	tactagagTCACACCACtactagatg	BBa_E70202
BBa_B0074	tactagagTCACACCACCCtactagatg	BBa_E70203

[a]The identifier is either that used by Rackham & Chin [99] or the BioBrick number denoted by a BBa prefix. [b]The locations of the capitalized letters in mRNA-A were randomized in the library. The capitalized letters in the BioBrick RBS sequences are taken from the Rackham RBS. The small letters represent the mixed connective site between assembled BioBrick parts. Note that the mixed connective site between an RBS and a coding region is truncated to allow optimal RBS - start codon spacing. [c]All reporters used in this work are described in greater details in Appendix H.

I assembled the BioBrick O-RBS sequences into GFP reporters as described in Appendix D and transformed the reporters into BL21(DE3) cells containing pCH1497-rRNA4. I grew the cultures in the presence or absence of arabinose as described in Appendix D and measured the fluorescence and culture density as a function of time after addition of arabinose (Figure 3-7). Each of the three RBS appeared capable of initiating translation of GFP although BBa_B0072 appeared to be significantly less active than the other RBS suggesting that at least the three nucleotides upstream of mRNA-A interacted with the O-ribosome as well as the sequence identified in the randomized library screen by Rackham & Chin. Growth curves (Figure 3-8) for the same cultures supported the claim that the Rackham O-ribosomes were less toxic to *E. coli* than the Brink O-ribosomes (compare to Figure 3-6).

3.2.3 Orthogonal transcription and translation in combination

Armed with functional T7 promoters and O-RBS, I next wished to demonstrate that both could be used in combination. I constructed a GFP reporter (BBa_E71205, Table H.2) that used the weakest T7 promoter (BBa_R0183) tested in Section 3.2.1 and the strongest O-RBS (BBa_B0073) tested in Section 3.2.2. This reporter should only express GFP in the presence of both T7 RNAP and O-ribosomes. As before, I transformed *E. coli* BL21(DE3) carrying pCH1497-rRNA4 (which I refer to as VM1.1, Appendix G) with the GFP reporter. 0.4 mM IPTG will induce expression of T7 RNAP and 0.2% arabinose will induce synthesis of O-ribosomes. I grew the cells in a multi-well fluorimeter and measured fluorescence as a function of culture density (Figure 3-9). Cultures contained neither, one, or both inducers. Detectable levels of GFP were produced only when both T7 RNAP and O-ribosomes were present. The growth curves for the same cultures show that expression of O-ribosomes reduced the final cell density of the cultures whereas expression of T7 RNAP appeared to have little affect on growth (Figure 3-10).

With a second population of RNAP and ribosomes in the cell, there are four
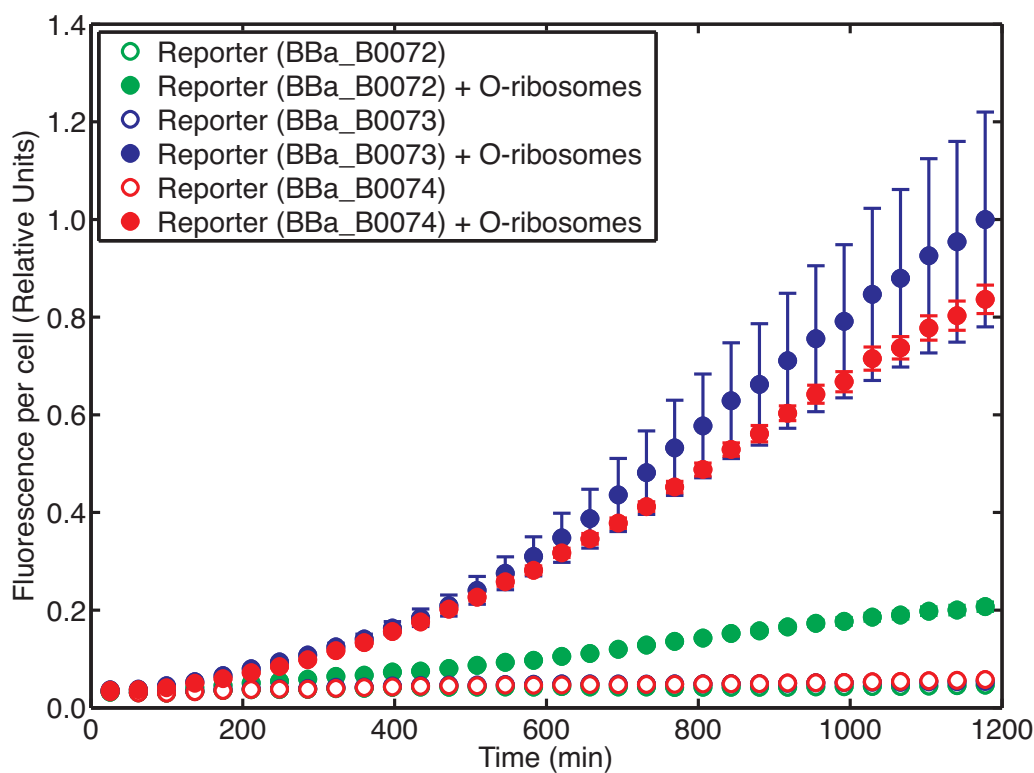
Figure 3-7: Rackham O-ribosomes can be used to express significant levels of GFP. Cultures consisted of cells carrying a reporter using one of three O-RBS (see Table 3.1). Cultures were grown either with or without 0.2% arabinose. Addition of arabinose induces expression of Rackham O-ribosomes from the plasmid-encoded O-ribosome generator. Each of the reporters led to an increase in fluorescence per cell in the presence of arabinose. The reporter using BBa_B0072 produced low levels of fluorescence relative to the other reporters. Each data point represents the mean of three independent cultures, each measured in triplicate. The error bars represent the 95% confidence interval in the mean.
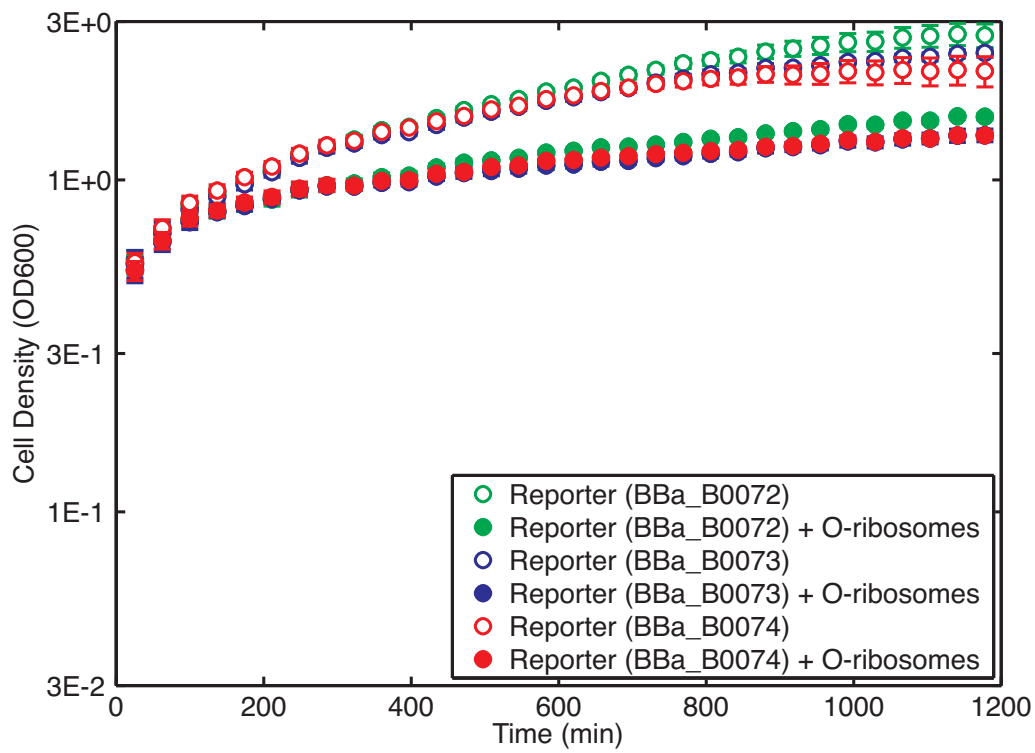
Figure 3-8: Synthesis of Rackham O-ribosomes has a moderate affect on culture growth. Each growth curve corresponds to a sample shown in Figure 3-7. Expression of O-ribosomes leads to a $\sim 50\%$ reduction in the final density of the culture. However, there is little evidence of a decrease in cell density at late times as observed when Brink O-ribosomes were expressed (3-6).
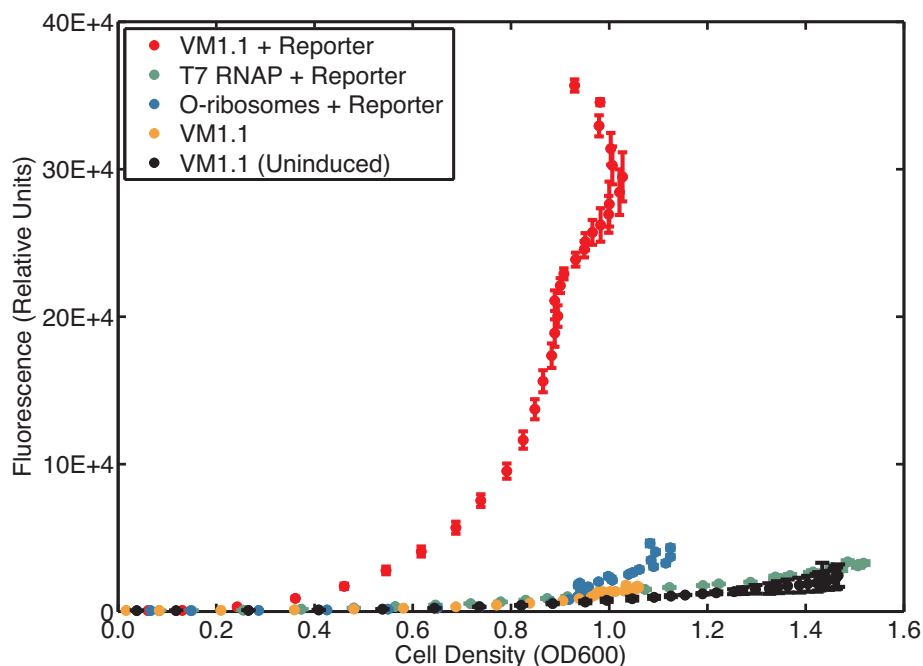
Figure 3-9: T7 RNAP and O-ribosomes can be used in combination to to express significant levels of GFP. VM1.1 cultures consisted of *E. coli* BL21(DE3) cells carrying the plasmid-encoded O-ribosome generator. In some cases the cells also contained a reporter. Cultures were grown with or without the inducers needed to synthesize T7 RNAP (0.4 mM IPTG) or O-ribosomes (0.2% arabinose). Data is shown for VM1.1 cultures with no reporter in the absence of IPTG and arabinose (black filled circles) and in the presence of both inducers (light blue filled circles). Induction of both T7 RNAP and O-ribosomes (red filled circles) was necessary to accumulate fluorescence detectable above background (black filled circles). The presence of O-ribosomes (dark blue filled circles) led to a low level of GFP expression late in the growth of the cultures suggesting there was leaky expression of T7 RNAP in stationary phase. The data suggests that transcription and translation of GFP from the reporter by *E. coli* RNAP and ribosomes is negligible. The data shown is for representative cultures where each data point is the mean of three samples each inocculated from an independent culture. The error bars represent the 95% confidence interval in the mean.
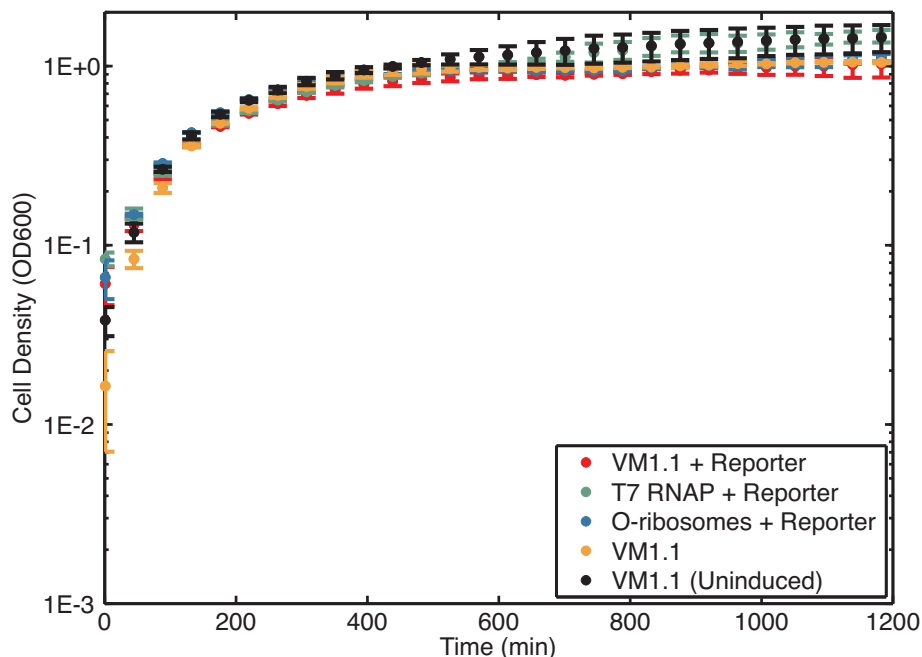
Figure 3-10: Growth curves for cultures expressing GFP using T7 RNAP and O-ribosomes. Each growth curve corresponds to a sample shown in Figure 3-9. The decrease in culture growth rate when O-ribosome synthesis is induced is minor.

possible combinations of RNAP and ribosomes that can be used to drive protein synthesis. In the preceding sections, I have described reporters that require either T7 RNAP or O-ribosomes or both in order to express GFP. I obtained a fourth reporter from the Registry (BBa_I7101, see Appendix H) that used the same constitutive *E. coli* promoter used in the O-ribosome reporters (BBa_E70201-3) and the same *E. coli* RBS used in the T7 RNAP reporters (BBa_E7104-5, BBa_E7107-8). I transformed BL21(DE3) carrying pCH1497-rRNA4 with each of these reporters. I measured the fluorescence as a function of cell density using the multi-well fluorimeter (Figure 3-11). Note these measurements were performed concurrently with those measurements shown in Figure 3-9. As expected, all reporter devices expressed GFP in the presence of both inducers but not in the absence of inducers. Of note, the reporter for *E. coli* RNAP and O-ribosomes (green circles) produced significantly less GFP than the other cultures. By comparing the green data series to the blue lines we can infer that the O-ribosome system with the current O-RBS has significantly less capacity for GFP expression than *E. coli* ribosomes and the *E. coli* RBS. Also of interest is the

fact that all three other reporters have roughly similar patterns of GFP expression as a function of cell density. BBa_I7101 produces a higher level of GFP earlier which is unsurprising for a constitutive reporter. By contrast, the other reporters must accumulate either or both T7 RNAP and O-ribosomes in order to express GFP.

Figure 3-11: GFP can be expressed using different combination of *E. coli* and orthogonal transcription and translation. Cultures were grown of VM1.1 cells containing a reporter that used either *E. coli* or orthogonal transcription and translation (see legend for color code). A control culture not containing a reporter device was included as a measure of fluorescence background. Each reporter led to an increase in fluorescence above background levels. The reporter using *E. coli* transcription and orthogonal translation (dark green) led to significantly less GFP expression than other combinations of transcription and translation. All other combinations of transcription and translation led to similar GFP expression patterns. The data shown is for representative cultures where each data point is the mean of three samples each inocculated from an independent culture. The error bars represent the 95% confidence interval in the mean.

In all experiments described above the source of T7 RNAP was from a chromosomal copy in BL21(DE3). To allow greater flexibility in use of T7 RNAP, I constructed a BioBrick version of the T7 RNAP coding sequence (BBa_I2032, see Appendix D). I cloned the coding sequence together with an *E. coli* RBS downstream of an arabinose-

inducible BioBrick promoter (BBa_I0500, Table H.1) based on the P_{BAD} promoter. I cloned the resulting T7 RNAP generator into a BioBrick vector pSB4C5 (a Bio-Brick plasmid with a pSC101 ori [60]). pSB4C5 is expected to be maintained at approximately five copies per cell and to have low cell-cell variability in copy number relative to other multi-copy plasmids (Johann Paulsson, personal communication). I transformed *E. coli* TOP10 (Invitrogen, Carlsbad, Ca.) with the plasmid-encoded T7 RNAP generator and pCH1497-rRNA4 (the resulting strain was termed VM1.2) and measured GFP expression from the reporter device that requires both T7 RNAP and O-ribosomes for expression in a manner similar to before (Figures 3-12 and 3-13). Cultures induced with arabinose produced high levels of GFP, whereas cultures carrying either the T7 RNAP generator or the O-ribosome generator alone with the reporter produced little GFP expression. From these data, I concluded that the plasmid-encoded T7 RNAP generator could produce high levels of active T7 RNAP.

Finally, I transformed VM1.2 cells (TOP10 strain carrying the T7 RNAP generator and the O-ribosome generator) with the four reporter devices using different combinations of *E. coli* and orthogonal transcription and translation enzymes. Again, I measured the fluorescence and cell density of these strains as a function of time after addition of inducer (Figure 3-14). The results were quite different from those for VM1.1 (BL21(DE3) with the plasmid-encoded O-ribosome generator). First, expression of GFP using *E. coli* RNAP and orthogonal ribosomes was even lower than for VM1.1 (see Figure 3-11). Second, the reporter using T7 RNAP and wt *E. coli* ribosomes led to an extremely rapid increase in fluorescence that saturated early. Note that those cells also reached a low final cell density suggesting the rapid expression of high levels of GFP hindered culture growth. Third, expression using *E. coli* RNAP and wt *E. coli* ribosomes appeared to lead to a higher level of GFP expression than T7 RNAP and O-ribosomes, whereas in VM1.1, both were similar.
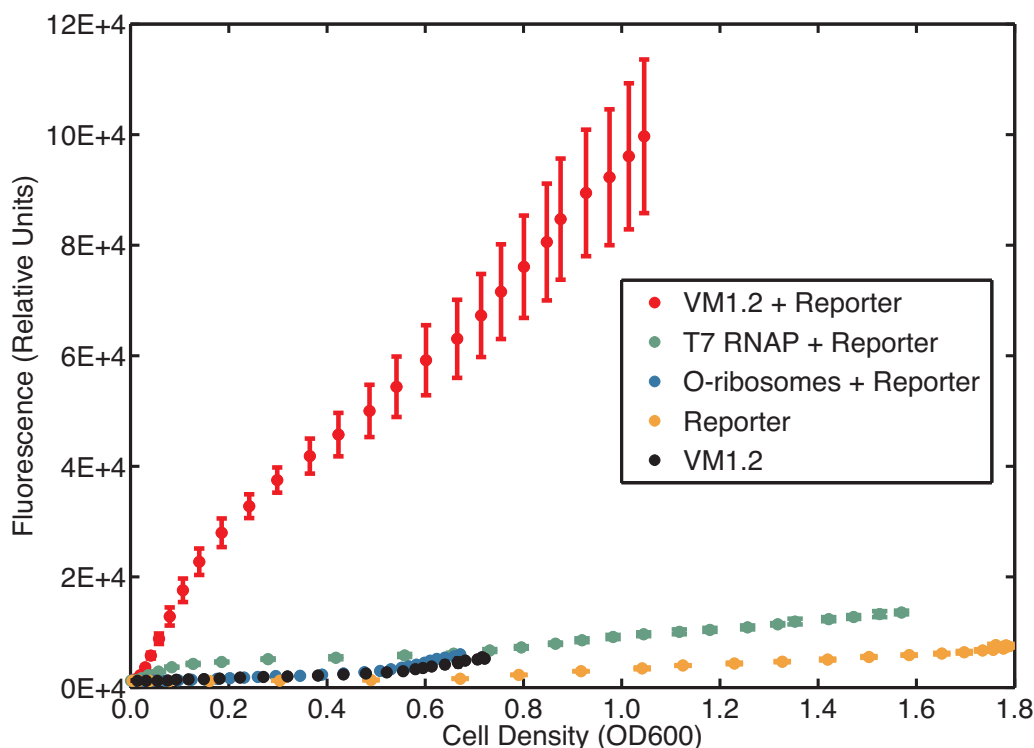
Figure 3-12: A plasmid-encoded T7 RNAP can be used together with O-ribosomes to express significant levels of GFP from a reporter. Cultures consisted of *E. coli* TOP10 carrying either the plasmid-encoded O-ribosome generator, or the plasmid-encoded T7 RNAP generator or both. In some cases the cells also contained a reporter. Data is shown for VM1.2 cultures with no reporter in the absence of IPTG and arabinose (black filled circles) and in the presence of both inducers (gold filled circles). Cultures were grown with or without 0.2% arabinose which induces expression of both O-ribosomes and T7 RNAP. Induction of both T7 RNAP and O-ribosomes (red filled circles) was necessary to produce fluorescence detectable above background (black filled circles). Each data point represents the mean of three independent cultures, each measured in triplicate. The error bars represent the 95% confidence interval in the mean.
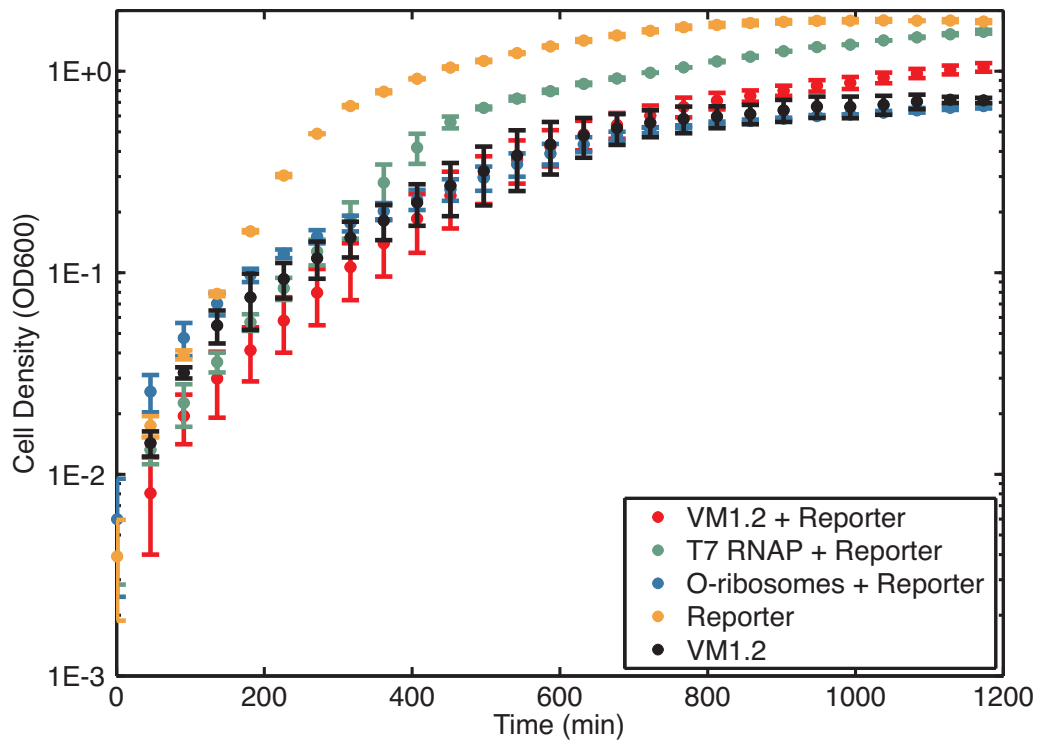
Figure 3-13: Growth curves for cultures expressing GFP using plasmid-encoded T7 RNAP and O-ribosome generators. Each growth curve corresponds to a sample shown in Figure 3-12. Expression of either T7 RNAP or O-ribosomes or both leads to a reduction in the final density in the culture.
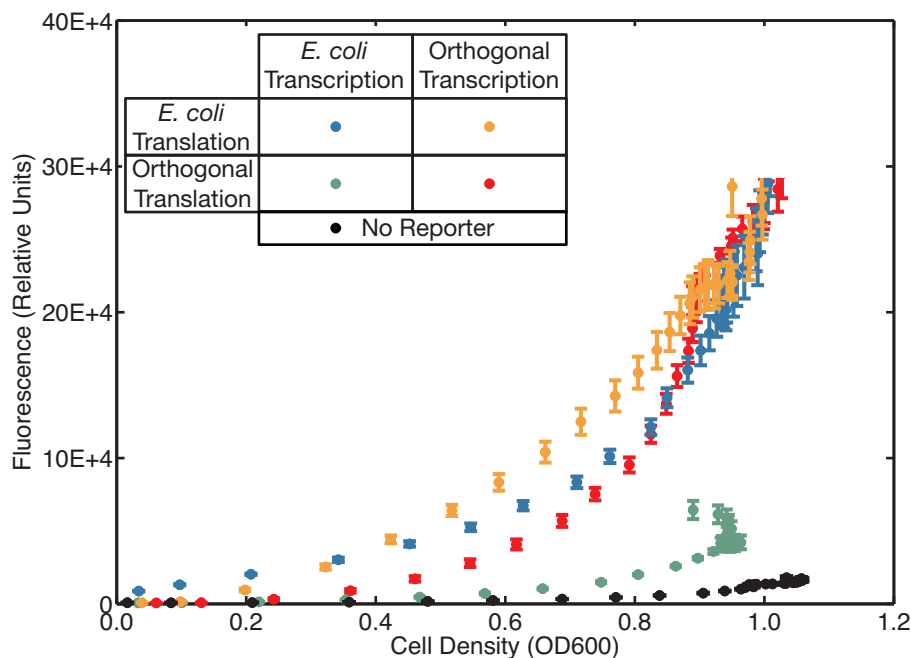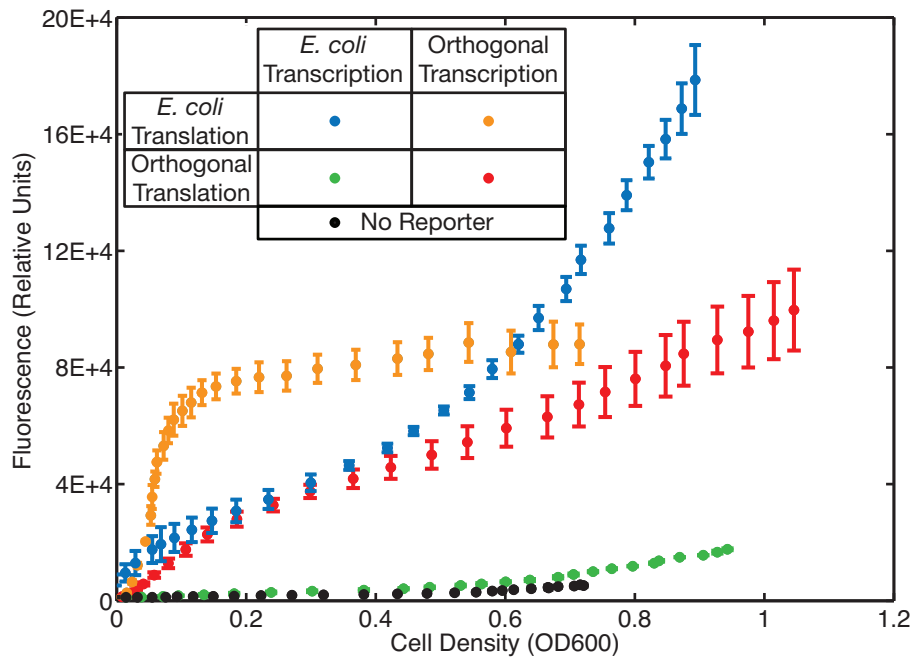
Figure 3-14: GFP can be expressed using different combination of *E. coli* and orthogonal transcription and translation. Cultures consisted of *E. coli* TOP10 cells carrying the plasmid-encoded O-ribosome generator. The cells also contained a reporter that used either *E. coli* or orthogonal transcription and translation (see legend for color code). A control culture not containing a reporter device was included as a measure of fluorescence background. Each reporter led to an increase in fluorescence above background levels. The reporter using *E. coli* transcription and orthogonal translation (green) led to significantly less GFP expression than other combinations of transcription and translation. Each data point represents the mean of three independent cultures, each measured in triplicate. The error bars represent the 95% confidence interval in the mean.

3.3   Discussion

Summary

In this Chapter, I argue that the behavior of engineered systems and cellular chassis would be more predictable if there were a modular relationship between systems and chassis with fewer shared resources. To that end, I describe the implementation of orthogonal transcription and translation working in coordination to transcribe and translate a reporter as a surrogate for an engineered biological system. Also, I develop a toolbox of standard biological parts to construct networks using orthogonal transcription and translation. Those components include T7 promoters of varying strength, O-RBS and the T7 RNAP.

My experiments using dedicated translation with or without dedicated transcription, show that transcription of plasmid-encoded mutant rRNA operons can have a dramatic affect on the growth rate of the cellular chassis. The Rackham O-ribosomes appear to have a less deleterious effect than the Brink O-ribosomes. This result suggests that the effect is not solely due to transcription of a plasmid-encoded rRNA operon and the associated demand on the cell, but is also due to a specific effect of the O-ribosomes, as has been reported for the Brink O-ribosomes [99].

Since both the O-ribosomes and the T7 RNAP have the potential to affect the behavior of the cellular chassis, care should be taken to use O-ribosomes and T7 RNAP under conditions where the effects on the cellular chassis are likely to be minimal. For example, *E. coli* BL21(DE3) appears to experience less of a reduction in growth rate in the presence of O-ribosomes than *E. coli* TOP10. Also, cultures grown in rich media such as LB or Neidhardt rich media [89] may experience less of a reduction in growth rate since there is greater availability of resources for transcription and translation, such as nucleotides and amino acids. Experiments comparing the behavior of the T7 RNAP- and O-ribosome-expressing strains in different media would be interesting to perform in the future.

The difficulties I described in implementing orthogonal translation relative to implementing dedicated transcription were in part because the O-ribosome technology

from which I was building was less developed than the T7 RNAP transcription system. However, another reason why orthogonal translation was more difficult to implement was the complex relationship between translational efficiency and the sequence at the 5' end of the message. Consider the issues I faced with secondary structure and cryptic *E. coli* RBS in the implementation of Brink O-ribosomes and reporter devices (Section 3.2.2). The assembly of translational units from standard biological parts (an RBS and a coding sequence) is likely to continue to require careful consideration of the parts and the DNA sequence. By constrast, at a transcriptional level, the composition of promoters and downstream elements seems to exhibit less context dependence.

One significant advantage of the T7 expression system is the low background; in the absence of T7 RNAP there is little expression of a target gene. However, there may remain some basal transcription by *E. coli* RNAP due to other promoters in the plasmid or nonspecific transcription initiation. In some commonly used vectors, some highly toxic genes cannot be stably maintained in *E. coli* when cloned with a T7 promoter and RBS even in the absence of T7 RNAP (F. William Studier, personal communication). The use of orthogonal transcription and translation in combination offers a potential method to reduce basal levels of protein synthesis lower than can be achieved with the T7 expression system since O-ribosome synthesis can be maintained at a low level when target gene expression is not wanted.

The data shown in Figures 3-11 and 3-14 allow estimates of the capacity of the different transcription and translation systems. As discussed in Section 3.2.3, the T7 RNAP and promoter appears capable of producing a high level of transcription relative to the level produced by *E. coli* RNAP and the *E. coli* promoter tested (Table H.1). In contrast, the O-ribosomes and O-RBS appear to cause significantly less translation than wt *E. coli* ribosomes and the *E. coli* RBS tested. This reduced capacity might be due either to a lower level or lower activity of O-ribosomes relative to wt *E. coli* ribosomes, or the O-RBS may initiate translation less efficiently than the *E. coli* RBS. An engineer using orthogonal RNAP and ribosomes to transcribe and translate an engineered biological system might have specific capacity requirements

which could be achieved by appropriate adjustment of either the levels of the orthogonal RNAP and ribosomes or by adjusting the strength of the orthogonal promoters and RBS. Depending on the application, an engineer would likely want the capacity of orthogonal transcription and translation to be matched, so as to use cellular resources most efficiently.

Is the use of orthogonal RNAP and ribosomes sufficient to partially decouple the behavior of a an engineered biological system from the cellular chassis? The answer to this question depends in part on which resources are responsible for the coupling (such as enzymes, materials, cofactors, or energy supply). With the orthogonal transcription and translation system developed here, we can begin to address this question, since expression of a reporter gene does not use *E. coli* RNAP nor the pool of wt *E. coli* ribosomes but does make use of *E. coli* resources such as nucleotides, amino acids, and energy.

It must be remembered that *E. coli* RNAP and wt *E. coli* ribosomes are required to express the T7 RNAP and O-ribosomes. Ideally, a network might be designed in which T7 RNAP and O-ribosomses could be expressed without the need for *E. coli* RNAP and ribosomes. Also, the levels of T7 RNAP and O-ribosomes are essentially unregulated after induction. Changes in culture growth rate, protein degradation, or synthesis rate will all alter the steady state level of T7 RNAP and O-ribosomes. In Chapter 4, I address both the sequestration of *E. coli* RNAP and wt *E. coli* ribosomes and the regulation of the orthogonal RNAP and ribosomes.

Chapter 4

An auto-regulated network of orthogonal RNAP and ribosomes (VM2.0)

4.1   Introduction

Control of the levels of the molecular species that comprise an engineered biological system is likely to be a key enabling ability for biological engineers. The receiver described in Chapter 2, serves as an illustrative example of the importance of controlling species levels. If the output of the receiver is to be determined solely by the level of the input signaling molecule ($3OC_6HSL$), the level of the constitutively expressed LuxR protein must not be so low as to limit promoter activation and must not be so high as to inhibit cell growth. Maintaining LuxR levels within the useful range depends on the availability of the cellular resources necessary to express LuxR, such as RNAP, ribosomes, molecular building blocks etc. I observed that the behavior of the receiver was dependent on the culture density in which the device was operating, with a decreasing output for a given input level as cell density increased. Such an effect is likely caused by a changing availability of cellular resources as growth rate changes [18]. A refined receiver design might use regulation to maintain LuxR levels

within the useful range in a manner that is robust to variations in the availability of cellular resources.

An alternative approach to enabling the construction of biological systems that behave reliably is to engineer the cellular chassis to minimize the variation in the availability of resources across culture conditions and growth phases. The idea of providing a constant environment for engineered systems was introduced in Chapter 1 and is a standard approach in mature engineering disciplines. For example, in computer science a virtual machine provides a constant environment for software, regardless of the underlying hardware platform. In a second example, electrical systems will commonly make use of surge protectors and voltage regulators to ensure the system receives a power supply that is constrained to useful levels. Cellular chassis that use analogous approaches to provide a constant supply of resources (or at least a supply of resources that always matches the demands of a system) such as molecular building blocks and enzymes are likely to make the behavior of engineered biological systems using those resources more predictable.

Natural biological systems also regulate the levels of key molecular species to provide a constant environment for cellular subsystems. For example, exponential or balanced growth of *E. coli* is defined by constrained relative levels of cellular species; the cell essentially grows in a homeostatic state [35]. As a more specific example, *E. coli* employs a complex network to regulate the level of ribosome synthesis, the general principles of which are understood although not all details are known [94, 35]. Broadly speaking, ribosome levels are maintained at a level sufficient to synthesize cellular proteins, increasing with increasing growth rate and decreasing as the cells enter stationary phase [67, 18]. It should be noted that ribosome synthesis is regulated in a fashion that presumably optimizes the growth and survival of *E. coli*. Levels of ribosomes (and other key cellular resources) may not be optimal for a biological engineer seeking to power an engineered biological system. Furthermore, the regulation network controlling ribosome synthesis is also unlikely to be optimized for human understanding, thereby making the engineering of cellular chassis tailored to a particular human-defined application difficult.

A standard engineering strategy for controlling the level of a variable is to make use of feedback. Feedback is the process by which the level of some variable is sensed and used as a signal to set the level of the variable itself. In contrast, if the level of a variable does not feedback to determine its own level, the control is termed "open loop". Examples of feedback networks abound in biology [54].

If the level of a feedback controlled species negatively affects its own level, the feedback is termed negative feedback. Negative feedback is used to hold a variable at a fixed level and is widely used in biological networks [113, 106]. Any decrease (or increase) in the level of a negative feedback-regulated species leads to an increase (or a decrease) in synthesis of the species until the original steady state is achieved once more. Negative feedback also offers the advantage of accelerating the return to a steady state following a perturbation to the level of the controlled species [106].

If the level of a feedback regulated species positively affects its own level, the feedback is termed positive feedback. Any decrease (or increase) in the level of a negative feedback-regulated species leads to a decrease (or an increase) in synthesis of the species until some external limit is reached. Consequently, with a positive feedback loop, the level of the controlled species can "run away" in the absence of other regulation. An engineered positive feedback network was demonstrated by Studier and coworkers [37]. In the described network, a T7 autogene was constructed in which the coding sequence for T7 RNAP was placed downstream of a consensus T7 promoter. Any basal transcription by *E. coli* RNAP leads to expression of T7 RNAP that in turn initiates transcription from the T7 promoter thereby positively regulating its own synthesis. I describe in Section 4.2 the strategies adopted by the authors to prevent run away synthesis of T7 RNAP. The T7 autogene represents an implementation of a positive feedback loop to construct a self-generating system (at least in terms of the polymerase); there is little need for *E. coli* RNAP to be present other than, perhaps, to provide a basal input to start the positive feedback loop. The behavior of such a network is likely to be largely decoupled from the availability of *E. coli* RNAP (assuming there is an excess of ribonucleotides in the cell [21, 35]).

In Chapter 3, I described the development of an orthogonal protein synthesis

system using open loop control (the levels of the orthogonal ribosomes and RNAP do not feedback to determine their own levels). In this chapter, I describe work to use positive and negative feedback to regulate the levels of O-ribosomes and T7 RNAP. My intention was to construct a network that is both self-sustaining and self-regulating. I present data demonstrating that the orthogonal ribosomes and RNAP have specific interactions and that I can modulate the species levels by modifying parameters of the network.

4.2   System design

When designing a regulatory network to control expression of the orthogonal ribosomes and RNAP described in Chapter 3, I had four design goals in mind. First, I wanted the network to produce a fixed level of ribosomes and polymerases that are robust to sources of variation, such as the sudden sequestration of either orthogonal ribosomes or RNAP by an engineered system. Second, I wanted to reduce the demand for *E. coli* RNAP and wt *E. coli* ribosomes so the orthogonal RNAP and ribosome network would place a reduced load on the cellular chassis. Third, I wanted the levels of the orthogonal ribosomes and RNAP to be relatively low so that they could be used for easily detectable expression of a reporter device but would not lead to a high level of expression that was toxic to the cell.

As described in Section 4.1, Dubendorff and coworkers constructed a T7 auto-gene that was stably maintained in an *E. coli* strain [37]. Stable maintenance of the autogene was dependent on constitutive expression of T7 lysozyme, which binds to and inhibits T7 RNAP. Additionally, for the autogene to be stably maintained, constitutive expression of LacI repressor was necessary. LacI bound to an operator site just downstream of the T7 promoter and further inhibited the positive feedback loop. The autogene exhibited some of the characteristics of the regulatory network I wished to construct. Namely, T7 RNAP expression was self-sustaining with little need for *E. coli* RNAP and could be stably maintained in *E. coli*. However, for the purposes of stably maintaining a low level of RNAP in the cellular chassis the auto-

gene network has some drawbacks. First, two additional proteins (T7 lysozyme and LacI) must be expressed by *E. coli* RNAP and wt *E. coli* ribosomes in order to limit the activity of the positive feedback loop. Second, the autogene is likely to exhibit the unstable equilibrium typical of positive feedback loops. Any perturbation in the level of T7 RNAP will lead to an increasing perturbation in the same direction raising the possibility of "run away" levels of T7 RNAP and saturation of the gene expression resources of the cellular chassis.

Taking inspiration from the Dubendorff autogene, I designed two T7 autogene networks incorporating several changes (see Figure 4-1). First, I constructed the network using BioBrick standard biological parts to allow ready construction of network variants quickly and reliably (see Chapter 1). Second, I constructed a mutant T7 promoter (BBa_R0184, see Table H.1) expected to be approximately 10-fold weaker than the consensus T7 promoter sequence used by Dubendorff and coworkers. The promoter made use of the same LacI binding site as the T7*lac* promoter [36]. Third, I placed a *lacI* coding sequence downstream of the LacI-repressible T7 promoter, thereby constructing a negative feedback loop. The negative feedback loop was intended to both reduce and stabilize the activity of the T7 promoter via LacI repression. Fourth, I placed the T7 RNAP coding sequence downstream of the *lacI* gene. Fifth, I placed a T7 promoter (based on Tϕ) downstream of the T7 RNAP coding sequence (BBa_B0016, see Table H.4). Sixth, I used a low copy plasmid vector for the system, pSB4C5, expected to maintain approximately five copies per cell [60]. Finally, I designed two variants of the T7 autogene device (T7 autogenerator), employing two different translation systems. In the first, I used *E. coli* RBS to regulate translation of the T7 RNAP and *lacI* mRNA. In the second, I used the orthogonal RBS (BBa_B0073 described in Chapter 3) to regulate translation of the T7 RNAP and *lacI* mRNA.
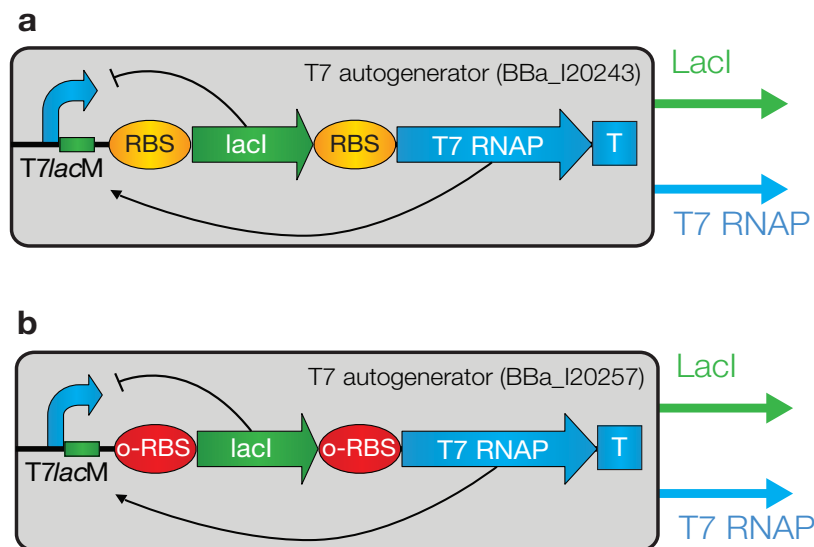
Figure 4-1: Two variants of the T7 autogenerator design. (a) A T7 autogenerator device using *E. coli* RBS. (b) A T7 autogenerator design using o-RBS. The positive and negative feedback loops due to the T7 RNAP and LacI respectively are indicated. T7*lac*M is a LacI-repressible T7 promoter (see Section 4.4.1 for details. The useful outputs from the device are T7 RNAP and LacI.

4.3 A model of the T7 autogenerator

4.3.1 A simple two-species model

The T7 autogenerators use a combination of positive and negative feedback to regulate expression of the T7 RNAP. Such combinations of positive and negative feedback can lead to unstable dynamic behavior of the network such as sustained oscillations. Rather than trusting intuition that the device would lead to stable low-level expression of T7 RNAP as designed, I chose to build a computational model of the network. The goal of the modeling work was (i) to gain insight into what species concentrations could be expected for plausible system parameters and (ii) to analyze the stability (the propensity of species levels to maintain a steady state) of the network. Since the values of many of the biochemical parameters in the network are uncertain, I started by building a simple model of the T7 autogenerator including just three species, T7 RNAP, LacI and a DNA operator.

In constructing the model, I made several assumptions:

1. The species concentrations can be represented as continuous dynamic variables. I ignored the discrete stochastic nature of the physical system.

2. Transcription and translation of the protein species can be represented as a single first-order processes with a single rate constant. I further assumed that basal synthesis rates of T7 RNAP and LacI can be represented as zeroth order processes with the same rate constant.

3. Induced synthesis of both proteins is proportional to the fraction of T7 RNAP-promoter complexes, which are formed via a reversible second order reaction between free T7 RNAP and free promoters.

4. LacI is active as a multimer. While the LacI protein forms a tetramer [86], the kinetic order for repressor-operator binding is approximately two since there is only one operator site at the promoter [61].

5. Repressor-operator complexes lead to no protein synthesis.

6. T7 RNAP and LacI are synthesized in a fixed ratio since they are cotranscribed but may have different translational efficiencies.

7. The same net protein degradation rates (degradation plus dilution due to growth) apply for each protein and degradation can be represented as a first-order reaction.

8. All binding and unbinding reactions are fast relative to protein synthesis and can be assumed to be in equilibrium.

Based on these assumptions, I represented the system by the following biochemical reactions, with species and parameters defined in Tables 4.1 and 4.2 -

$$
\begin{array}{lll}
nL + D & \overset{K_d^1}{\longleftrightarrow} & L_n.D \qquad \text{(Repressor-operator binding)} \\[4pt]
R + D & \overset{K_d^2}{\longleftrightarrow} & R.D \qquad \text{(T7 RNAP-promoter binding)} \\[4pt]
R.D & \overset{k_t}{\longrightarrow} & R \qquad\qquad \text{(T7 RNAP synthesis)} \\[4pt]
R.D & \overset{r \cdot k_t}{\longrightarrow} & L \qquad\qquad\;\; \text{(LacI synthesis)} \\[4pt]
\phi & \overset{k_l}{\longrightarrow} & R \qquad\quad \text{(Leaky synthesis of T7 RNAP)} \\[4pt]
\phi & \overset{r \cdot k_l}{\longrightarrow} & L \qquad\qquad \text{(Leaky synthesis of LacI)}
\end{array}
$$
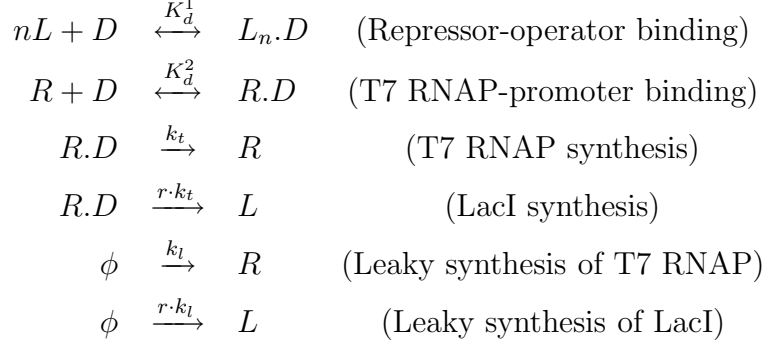
Table 4.1: T7 autogenerator model species

Variable	Definition	Units
$[R]$	T7 RNAP concentration	nM
$[L]$	LacI concentraion	nM
$[D]$	Operator concentraion	nM

From these reactions I derive mass action kinetics rate equations for both the T7 RNAP and LacI -

$$
[\dot{R}] \;=\; \left(\frac{k_l[D]_{tot} + k_t[D]_{tot}\left(\frac{1}{K_d^2}\right)[R]}{1 + \left(\frac{1}{K_d^2}\right)[R] + \left(\frac{1}{K_d^1}\right)[L]^n} \right) - \gamma[R] \tag{4.1}
$$

$$
[\dot{L}] \;=\; r \cdot \left(\frac{k_l[D]_{tot} + k_t[D]_{tot}\left(\frac{1}{K_d^2}\right)[R]}{1 + \left(\frac{1}{K_d^2}\right)[R] + \left(\frac{1}{K_d^1}\right)[L]^n} \right) - \gamma[L] \tag{4.2}
$$

Since the synthesis term for LacI differed from that of the T7 RNAP by a proportionality constant, the levels of LacI and T7 RNAP were also related by the constant of proportionality, r. As a result, I reduced the system to one equation where $[L]$ was replaced by $r[R]$:

$$
[\dot{R}] \;=\; \left(\frac{k_l[D]_{tot} + k_t[D]_{tot}\left(\frac{1}{K_d^2}\right)[R]}{1 + \left(\frac{1}{K_d^2}\right)[R] + \left(\frac{1}{K_d^1}\right)(r[R])^n} \right) - \gamma[R] \tag{4.3}
$$

At steady state $[\dot{R}]$ was zero and I solved for the steady state level of T7 RNAP,

Table 4.2: T7 autogenerator model parameters

Parameter	Definition	Value	Units	Reference
K_d^1	LacI-operator dissociation constant	1	$(nM)^n$	[91]
n	Kinetic order of LacI-operator binding	2	unitless	[61]
K_d^2	T7 RNAP-promoter dissociation constant	1	nM	[11, 126]
k_t	T7 RNAP synthesis rate	70	T7 RNAP/s/active promoter	Derived[a]
k_l	T7 RNAP basal synthesis rate	0.01	T7 RNAP/s/active promoter	Assumed
r	Ratio of LacI synthesis rate to T7 RNAP synthesis rate	1	unitless	Assumed
γ_{RNAP}	Combined degradation and dilution rate of T7 RNAP	1.9E-4	1/s	Derived[b]
γ_{LacI}	Combined degradation and dilution rate of LacI	1.9E-4	1/s	Derived[b]
D_{tot}	Total number of promoters per cell	5	copies/cell	[60]

[a]T7 RNAP synthesis rate per active promoter was assumed to equal the rate of protein synthesis per mRNA per second and the steady state concentration of mRNA per promoter copy. Protein synthesis per mRNA per second was estimated from published values [34]. mRNA steady state concentration was calculated as the ratio of the mRNA synthesis rate per copy divided by the degradation rate of mRNA. mRNA synthesis rate per DNA copy was estimated from kinetic measurements for T7 promoters [68, 63] and mRNA degradation rate was calculated from [14]. [b]Assuming a stable protein and a 60 min culture doubling time

$[R]_{ss}$, by setting the left hand side of Equation 4.3 equal to zero. I noted that for typical parameter values, the basal protein synthesis term is small relative to the activated protein synthesis term and can be neglected in calculating the steady state level of T7 RNAP. Next, I noted that for typical parameter values (Table 4.2), the third term in the denominator of the synthesis term was large relative to the first and second terms. Making these approximations yielded a simple solution for the steady state level of T7 RNAP in terms of certain parameter values:

$$[R]_{ss} = \frac{1}{r} \cdot \left(\frac{k_t D K_d^1}{\gamma K_d^2} \right)^{\frac{1}{n}} \qquad (4.4)$$

Solving Equation 4.4 using the parameter values given in Table 4.2 yielded a steady state level of T7 RNAP of 1346 proteins per cell (the same for LacI since r is assumed equal to one). Relative to expression of heterologous proteins in *E. coli*, which can reach 50% of total cell protein [53], the calculated levels of T7 RNAP and LacI seemed moderate.

Using this simple model one can begin to estimate how changes to parameter values affects device behavior. First, each variable affects R_{ss} in an intuitive fashion. For example, as r increases, more LacI is made relative to the amount of T7 RNAP and so the steady state level of T7 RNAP decreases (see Figure 4-2). As the synthesis rate of RNAP (k_t) increases, so too do steady state levels of T7 RNAP (see Figure 4-3).

Also note from Equation 4.4 that as the kinetic order of LacI repression (n) increases, sensitivity of the steady state level of T7 RNAP to changes in promoter copy number (D), degradation and growth rate (γ) and protein synthesis rate (k_t) decreases. For a T7 RNAP generator device that uses open loop control such as that presented in Chapter 3, the steady state level of T7 RNAP is linearly dependent on each of those variables (equations not shown). The increased stability of the steady state to changes in parameter values is due to the LacI-based negative feedback. Later in this chapter I will discuss how two parameter values, r and k_t, can be varied experimentally.

4.3.2 Dynamic behavior of the T7 autogenerator

The model presented above was sufficient to gain insight into the magnitudes of the steady state protein levels of the network. However, such a simple model is insufficient to describe all possible dynamic behaviors of the physical network represented by the model [61]. Additionally, the combination of negative and positive feedback makes the dynamics of the T7 autogenerator hard to predict based on human intuition alone. For particular values of rate constants and kinetic orders, the topology of the T7 autogenerator might exhibit stable, oscillatory or toggle switch (bistable) behavior [9]. The desired behavior of the T7 autogenerator is to produce a non-varying level of
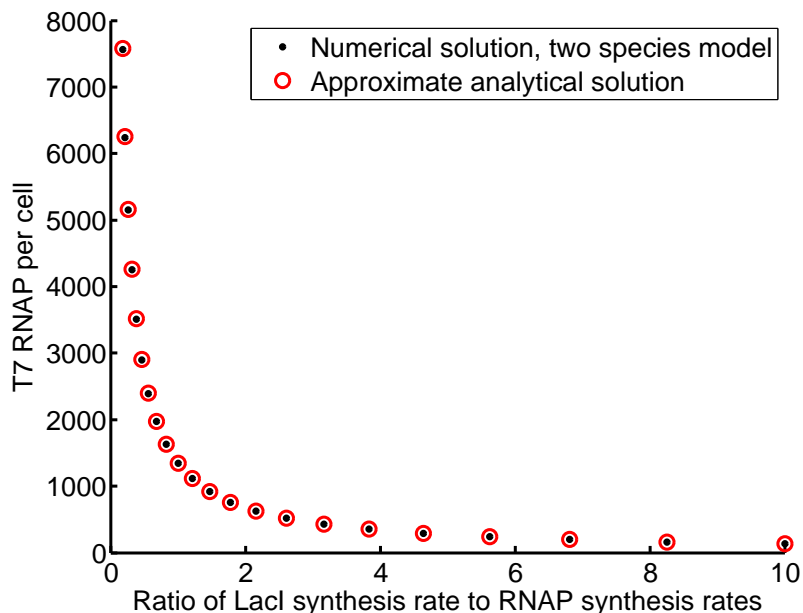
Figure 4-2: Dependence of steady state T7 RNAP levels on the ratio of LacI expression to that of T7 RNAP. Steady state levels of T7 RNAP were calculated over a tenfold increase and a tenfold decrease in the value of r given in Table 4.2 (from 0.1 to 10). All other parameter values were as given in Table 4.2. Steady state levels were calculated by two methods, first by numerical solution of Equations 4.1 (black dots) and second, via Equation 4.4 (red circles). Both methods are in close agreement. MATLAB (The Mathworks Inc.) was used for all computational analysis.

T7 RNAP, although a demonstration of oscillatory behavior might be of wide interest [42, 9, 49, 12]

Network topologies similar to that of the T7 autogenerator occur widely in nature, the lactose and arabinose systems found in *E. coli* being two well-known examples [113]. Consequently, the dynamic behavior of such networks as a function of rate constants and reaction orders have been studied in detail [113, 9, 12]. Such models are more biochemically detailed than that presented above; for example, the levels of mRNA are explicitly considered as well as protein levels. As a result they can better predict the full range of dynamic behaviors of the physical network. The criteria for the behavior of the T7 autogenerator network topology to be stable have been determined (see Fig. 1 of [9]). Assuming typical values for the half lives of the mRNA and proteins in the network, a stable steady state is dependent only on the kinetic orders of T7 RNAP activation of both itself and LacI and also the kinetic orders of
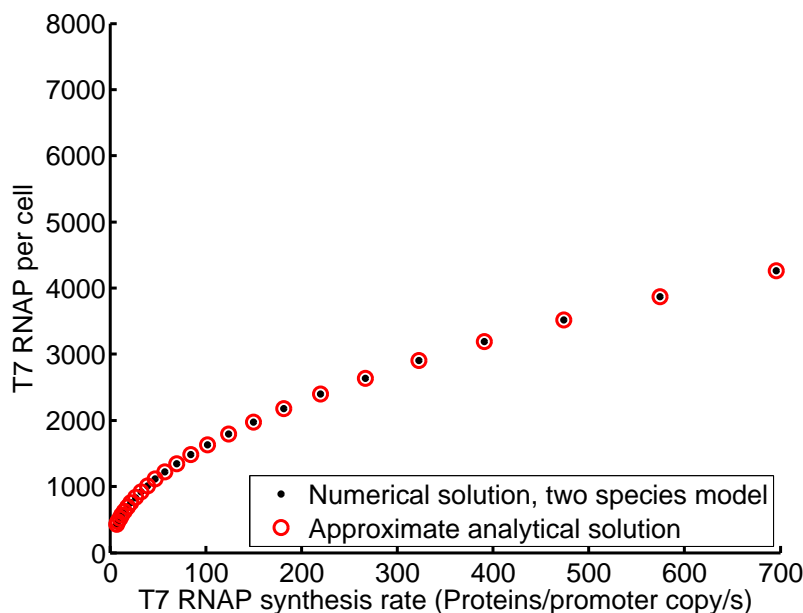
Figure 4-3: Dependence of steady state T7 RNAP levels on the protein synthesis rate of T7 RNAP. Steady state levels of T7 RNAP were calculated over a tenfold increase and a tenfold decrease in the value of k_t given in Table 4.2 (from 7 to 700 proteins/promoter copy/s). All other parameter values were as given in Table 4.2. Steady state levels were calculated by two methods, first by numerical solution of Equations 4.1 (black dots) and second, via Equation 4.4 (red circles). Both methods are in close agreement. MATLAB (The Mathworks Inc.) was used for all computational analysis.

LacI regulation of itself and T7 RNAP. The reaction order for T7 RNAP, which is active as a monomer and which has just one site of action per gene, is close to 1. The kinetic order for LacI regulation is approximately -2 since the repressor forms a tetramer capable of binding two operator sites but the engineered T7 promoter that I will describe in Section 4.2.1 has only one LacI binding site [113]. For these kinetic orders, the T7 autogenerator lies well within the region of parameter space expected to produce stable behavior.

From the analysis presented above, I expected that the T7 autogenerator was likely to be stable and the demand on a cellular chassis due to T7 RNAP and LacI synthesis might not overwhelm the capacity of the chassis. Of course, with the simplicity of the models considered, such predictions cannot be taken for granted. Consequently, further investment in computational analysis of the T7 autogenerator should await

78

quantitative experimental data on device behavior.

4.3.3 Transcription of orthogonal rRNA by T7 RNAP

The ability to transcribe orthogonal rRNA without sequestering *E. coli* RNAP would reduce demand for cellular resources, and decouple O-ribosome synthesis from the regulation of *E. coli* transcription. I considered transcription of orthogonal rRNA by T7 RNAP as a potential way to remove the need for *E. coli* RNAP. Two groups have previously attempted to use T7 RNAP to synthesize rRNA that can be assembled into functional ribosomal particles [121, 77]. Although there are differences in the materials and methods used by the two groups, there are some common findings from both studies. First, T7 RNAP can be used to transcribe a wild-type ribosomal operon. The resulting transcripts can be correctly processed by RNase III to yield the correct length rRNA precursors. Furthermore, under certain conditions, the rRNA precursors can be assembled into functional ribosomal subunits. In the experiments performed by Steen and coworkers, mature subunits were formed only when T7 expression was induced late in the cell growth phase. In the work of Lewicki *et al.*, T7 RNAP transcribed rRNA assembled into functional ribosomal subunits in large numbers at a reduced temperature (25°C). At 37°C, only a small fraction (∼10%) of T7 RNAP transcribed rRNA was found in functional ribosomal particles. The majority of T7 RNAP transcribed rRNA was found in non-funcional 50S and 30S subunits. This result was thought to be due to one of two factors: either the high level of transcription of the ribosomal operon by T7 RNAP oversaturates the ribosome-assembly machinery or the faster speed of T7 RNAP relative to *E. coli* RNAP may not be well-matched to the ribosome assembly process. The experiments reported by the authors do not appear sufficient to rule out either of these possibilities. One final difference between the studies involved the transcription start site used. Due to differing cloning strategies used by the two groups the immature rRNA had different sequences added to the 5' end (see Table H.7 for a comparison). Such additions may alter RNA stability [100] or the efficiency of rRNA processing [120].

From these two prior studies I presumed it likely that a set of conditions can be

found under which T7 RNAP can be used to transcribe rRNA that is assembled into functional ribosomes. However, making this process work efficiently might require the use of non-standard conditions, such as lower growth temperatures, induction of T7 RNAP only late in the growth cycle, the use of weak T7 RNAP promoter sequences, or the use of a slower-transcribing mutant of T7 RNAP [82].

The combination of the T7 autogenerator using O-ribosomes and a T7 RNAP-regulated O-ribosome generator opens the intriguing possibility of constructing a virtual machine network that requires negligible numbers of *E. coli* RNAP or ribosomes. The interaction between the two devices is shown schematically in Figure 4.4.5. I note that the dynamics of the VM2.0 network are likely to be different from those of the T7 autogenerator presented above and likely warrant further computational studies.
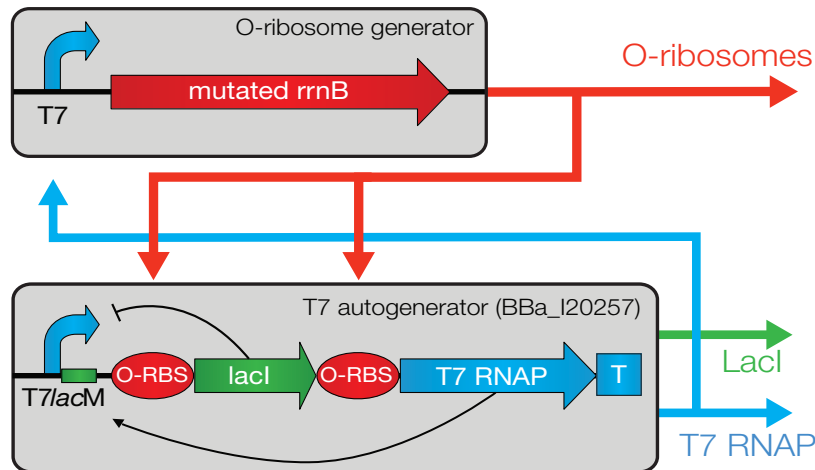
Figure 4-4: The VM2.0 network in schematic form. VM2.0 consists of two devices, a T7 autogenerator using orthogonal translation and a T7 promoter-regulated O-ribosome generator. The interactions between the two devices are shown by heavy lines. The heavy blue line represents activation by T7 RNAP and the heavy red line represents activation by O-ribosomes. Interactions internal to a device are shown as thin black lines. T7 RNAP and O-ribosomes generated by VM2.0 are also available for use by an engineered biological system.

4.4 Results

4.4.1 Testing the VM2.0 component parts

I constructed several new parts during assembly of the VM2.0 network. All construction methods are detailed in Appendix E. The first part I constructed was a T7 promoter with a LacI binding site, BBa_R0184 (Table H.1). This promoter would be used to transcribe the T7 RNAP and *lacI* genes, and the orthogonal rRNA (T7*Lac*M in Figure 4.4.5). The promoter was very similar to the previously described T7*lac* promoter [36]. I modified that promoter by introducing the same A to C mutation at the -16 base as described in Chapter 3 for the T7 promoter, BBa_R0183. Prior characterization of mutant T7 promoter strengths suggested that BBa_R0184 should be ∼10% of the strength of the T7*lac* promoter [64].

I constructed a *lacI* coding region (BBa_C0013) based on an existing BioBrick part, BBa_C0012 comprising a *lacI* gene fused to a 33bp ssrA degradation tag sequence. The sequence for BBa_C0012 was taken from a previously published sequence [42]. I constructed BBa_C0013 by PCR amplification of the *lacI* coding sequence from BBa_C0012 without the *ssrA* tag as described in Appendix E.

I performed a simple experiment to ensure that the LacI-repressible T7 promoter (BBa_R0184) could be repressed by the protein product of BBa_C0013. First, I constructed a reporter (BBa_I2035, see Table H.2) in which the LacI-repressible T7 promoter regulates transcription of a gene encoding GFP on a high copy plasmid (pSB1AT3, Table H.5). Second, I constructed a gene expression cassette (BBa_I2043) in which transcription of *lacI* (BBa_C0013) is regulated by the receiver (BBa_F2620) on a mid copy plasmid (pSB3k3, Table H.5). I transformed the two plasmids into *E. coli* BL21(AI) (Invitrogen). *E. coli* BL21(AI) has a chromosomally integrated copy of T7 gene 1 under the control of the ara P_{BAD} promoter. Addition of arabinose to the medium of a culture of *E. coli* BL21(AI) induces transcription of T7 gene 1. I grew the strain carrying both the LacI expression cassette and the GFP reporter with or without 0.2 % arabinose and 0.4 mM IPTG and measured the fluorescence and cell density of the cultures as described in Appendix E. I also measured the

fluorescence and cell density of cultures of cells carrying either the reporter or the LacI expression cassette alone. Fluorescence per unit OD600 was used as an indirect measure of the activity of the LacI-repressible T7 promoter. The resulting data is shown in Figure 4-5. As expected, when the T7 RNAP was not induced, little GFP accumulation was observed (see the first black and white bars in Figure 4-5). Induction of T7 RNAP led to high accumulation of GFP in the strain without the LacI expression cassette confirming that the T7 promoter was active (compare first and second black bars). The strain carrying the LacI expression cassette as well as the reporter expressed little GFP in the presence of T7 RNAP suggesting LacI was indeed capable of repressing the T7 promoter (compare second black and white bars). Adding IPTG (expected to inhibit LacI repression) as well as inducing T7 RNAP was sufficient to cause GFP expression from the strain containing the reporter and the LacI expression cassette (compare the second white bar to the fourth white bar). Note that although transcription of the *lacI* gene was regulated by the receiver, addition of $3OC_6HSL$ to induce the receiver appeared to be unnecessary to produce sufficient LacI to repress the T7 promoter. From these data I concluded that the T7 promoter was active, and could be repressed by the LacI repressor. There were two other observations that might be worth investigating further. First, addition of IPTG appears to reduce GFP levels when T7 RNAP expression is induced (compare the second and fourth black bars). Second, when T7 RNAP expression is induced, IPTG appeared to only partially recover GFP expression levels, suggesting either that inhibition of LacI was incomplete, or that the LacI cassette plasmid affected the level of GFP expression via some other mechanism (compare fourth black and white bars).

4.4.2 A T7 autogenerator using *E. coli* translation

Having constructed a functional T7 RNAP BioBrick (see chapter 3, part BBa_I2032), and now a LacI-repressible T7 promoter and an active LacI repressor, I assembled a T7 autogenerator (BBa_I20243). The final assembly stage involved assembling the T7 promoter with a part consisting of the T7 RNAP coding sequence, the *lacI* gene (both with *E. coli* RBS sequences upstream) and the T7 transcriptional terminator
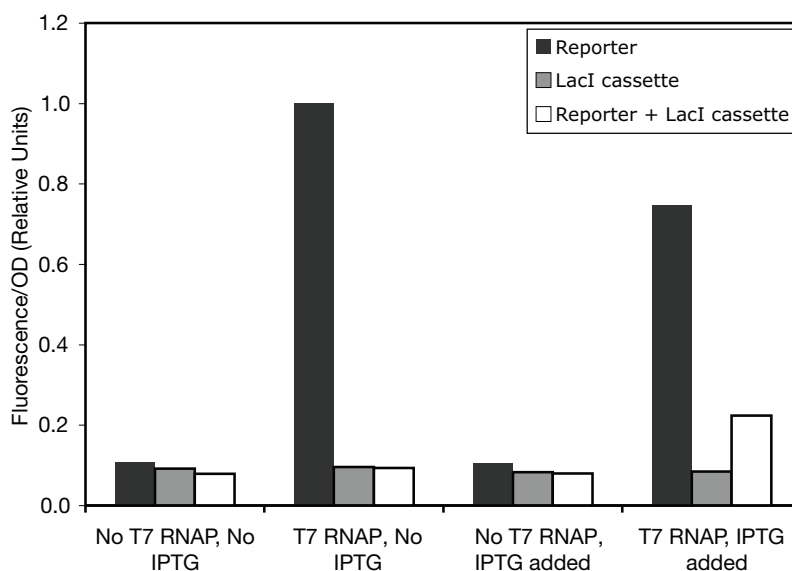
Figure 4-5: BBa_R0184 is an active and LacI-repressible T7 promoter. Cultures of BL21(AI) carrying a GFP reporter (black bars), a LacI expression cassette (grey bars), or both (white bars) were grown with or without induction of T7 RNAP expression (0.2% arabinose) and with or without inhibition of LacI (0.4 mM IPTG). Cells expressing T7 RNAP and containing the reporter alone expressed high levels of GFP (compare first and second black bars). LacI, in the absence of inhibitor was able to repress GFP expression (compare second black and white bars). Note that inhibition of LacI was insufficient to recover the same level of GFP expression as in the absence of LacI.

Tϕ (BBa_B0016) (Figure 4-1a). Several attempts to assemble this final stage of the T7 autogenerator into the common *E. coli* cloning strain TOP10 (Invitrogen) proved unsuccessful. Attempts to clone the autogenerator into a TOP10 strain carrying a T7 lysozme-expressing plasmid (pLysS [124]) also failed. However, I was able to clone the final construct into *E. coli* D1210 [110]. A possible reason why the T7 autogenerator should be stable in *E. coli* D1210 but not *E. coli* TOP10 is because the lacI^q genotype of D1210 leads to higher intracellular levels of LacI expression than are found in TOP10. The additional LacI may be capable of repressing the T7 promoter sufficiently to render the T7 autogenerator stable. If constitutive levels of LacI are critical to the stability of the T7 autogenerator, that would suggest that the T7 autogenerator itself does not express enough LacI for stability. Although the T7 autogenerator could be stably maintained in D1210, the culture had significantly

slower growth than the parent strain (compare the blue diamonds and red triangles of Figure 4-6). Nevertheless, the ability to stably maintain the T7 autogenerator at all, and particularly to be able to do so in the absence of T7 lysozyme, was an encouraging intermediate result on the way to constructing the VM2.0 network.
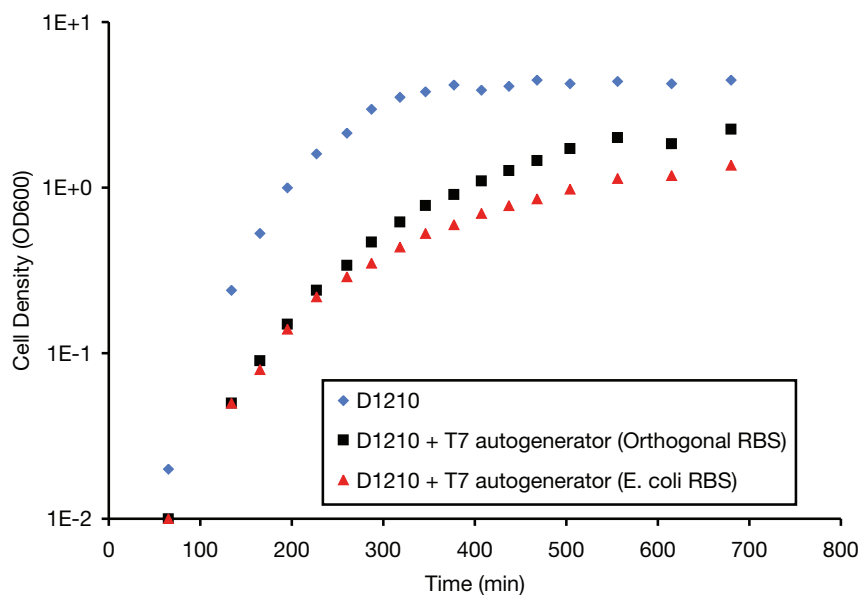
Figure 4-6: The presence of a plasmid encoding a T7 autogenerator significantly reduces the growth of *E. coli* D1210. Two T7 autogenerators were compared, one using *E. coli* ribosomes and one using orthogonal ribosomes (Figure 4-1). The T7 autogenerator using orthogonal RBS reaches an OD600 of approximately half that reached by the parent strain, *E. coli* D1210. The T7 RNAP autogenerator using with *E. coli* RBS reaches an OD600 of approximately 25% that reached by the parent strain. Data was measured and processed as described in Appendix E.

4.4.3 A T7 autogenerator using orthogonal translation

I assembled a new version of the T7 autogenerator (BBa_I20257), making use of orthogonal RBS sequences in place of the *E. coli* RBS sequences used in the T7 autogenerator described in Section 4.4.2. Otherwise, the autogenerator was unchanged. As an anecdotal observation, this version of the T7 autogenerator was readily cloned in *E. coli* TOP10 cells lacking orthogonal ribosomes. However, in *E. coli* D1210, there is still a significant reduction in growth rate relative to the parent strain (Figure 4-6).

To test whether the autogenerator produced active T7 RNAP, I cotransformed the autogenerator into *E. coli* TOP10 cells with or without the O-ribosome generator described in Chapter 3. Additionally, the cells included the reporter described in Chapter 3 and Table H.2 (BBa_E71205), which requires both T7 RNAP and O-ribosomes for significant production of GFP fluorescence. The cells were grown in LB medium with or without 0.2 % arabinose to induce production of O-ribosomes as described in Appendix E. The fluorescence and cell density of the cultures were measured following centrifugation and resuspension in Phosphate-Buffered Saline (PBS). The fluorescence per OD600 was calculated as an indirect measure of GFP per cell (see Figure 4-7). In the absence of arabinose there appears to be little accumulation of GFP for any combination of devices. In the presence of arabinose, which induces expression of the O-ribosomes, detectable levels of GFP accumulate only when both the T7 autogenerator and the O-ribosome generator are present. As an anecdotal observation, the fluorescence of these cultures is low relative to that of other cultures measured during this project. Nevertheless, the data suggests that the O-ribosome-regulated T7 autogenerator was capable of producing T7 RNAP in the presence of O-ribosomes.

4.4.4   Transcription of orthgonal rRNA by T7 RNAP

I next modified the O-ribosome generator described in Chapter 3 by replacing the P_{BAD} promoter with a range of T7 promoters. The assembly process is described in Appendix E. The T7 promoters were designed to produce low level activity relative to the consensus T7 promoter sequence. Some promoters included a LacI binding site to more tightly regulate promoter activity. The individual promoters are described in Table 4.3 and the nucleotide sequence junctions between the T7 promoters and the orthogonal rRNA are shown for two representative promoters in Table H.7.

I transformed the resulting constructs into BL21(DE3) together with a GFP reporter (BBa_E70202, Table 3.1). The GFP reporter is transcribed constitutively by an *E. coli* promoter and translation of the GFP mRNA is controlled by an orthogonal RBS. I grew the resulting strains with or without 0.8 mM IPTG to induce expression
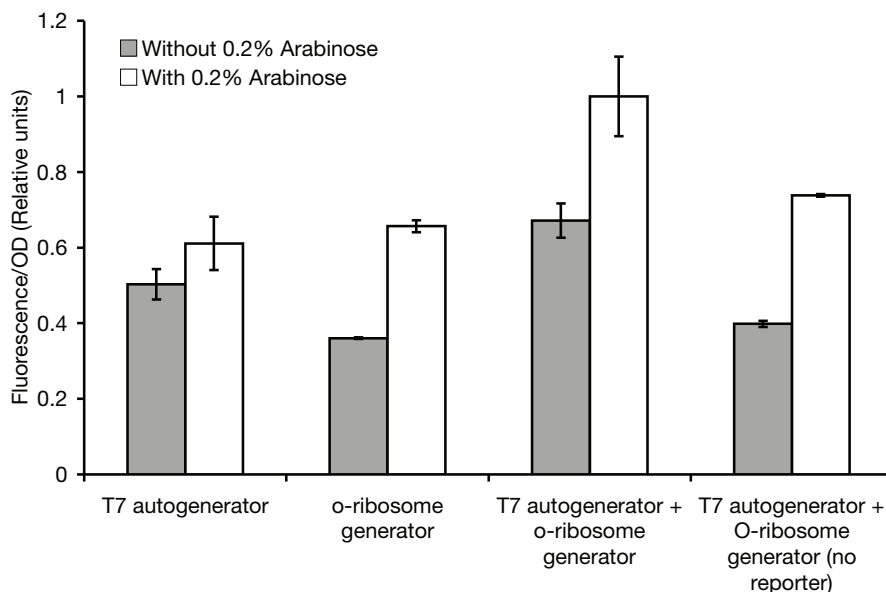
Figure 4-7: The O-ribosome-regulated T7 autogenerator can produce T7 RNAP in the presence of O-ribosomes. Relative GFP expression was measured for cells containing either the T7 autogenerator, the O-ribosome generator or both. Cells contained a GFP reporter requiring both T7 RNAP and O-ribosomes unless indicated. In the presence of both the T7 autogenerator and the O-ribosome generator, GFP expression was detectable but less than two fold greater than cellular autofluorescence. Data was measured and processed as described in Appendix E. The data points represent the mean of three independent cultures and the error bars represent the standard deviation for the three samples.

of T7 RNAP from the chromosome of BL21(DE3). I measured the fluoresence and OD600 of the cultures after 15 hours of growth (see Appendix E for details). After the subtraction of appropriate backgrounds, the ratio of fluorescence to OD600 is an indirect measure of the amount of GFP per cell. The amount of GFP per cell is taken as an indirect indicator of the amount of orthogonal rRNA transcribed by the T7 RNAP and assembled into functional O-ribosomes. The resulting data is shown in Figure 4-8. There is a clear increase in GFP expression when T7 RNAP expression is induced for all promoters (compare white bars to grey bars in Figure 4-8). The strengths of the promoters (as predicted from published experiments [64]) agree at least in rank order with the observed GFP expression. In a somewhat unexpected result, the addition of the LacI binding site increased GFP expression and, by in-

Table 4.3: T7 promoter variants used to transcribe the orthogonal rRNA operon

BioBrick #	Mutation (Strength[a])	LacI site	Sequence[b]
BBa_R0183	-16C (0.09)	No	tCatacgactcactata gggaga
BBa_R0184	-16C (0.09)	Yes	tCatacgactcactata ggggaattgtgagcggataacaattcc
BBa_R0185	-10G (0.04)	Yes	taatacgGctcactata ggggaattgtgagcggataacaattcc
BBa_R0186	-9A (< 0.03)	Yes	taatacgaAtcactata ggggaattgtgagcggataacaattcc
BBa_R0187	-9G (< 0.03)	Yes	taatacgaGtcactata ggggaattgtgagcggataacaattcc

[a]Promoter strength relative to the consensus T7 promoter sequence [64]. The promoters were measured without a LacI binding site downstream and included the +1 to +6 sequence. [b]The mutated base is capitalized. A space separates the transcribed sequence and the non-transcribed sequence. The LacI binding site is underlined.

ference, the expression of O-ribosomes. The addition of the 20bp hairpin encoded by the LacI binding site at the 5' end of the immature rRNA may either prevent degradation of the RNA or more efficiently promote rRNA processing. From these data I concluded that T7 RNAP could be used to direct transcription of orthogonal rRNA that could then be assembled into functional O-ribosomes. I further conclude that by modulating the strength of the T7 promoter, that the level of the orthogonal ribosomes can be controlled. In separate experiments (data not shown) I noted that transcription of orthogonal rRNA appeared to work equally well at relatively low ODs and also at both 37°C and 30°C.

4.4.5 VM2.0

Having demonstrated a stable T7 autogenerator device being translated by O-ribosomes and a set of O-ribosome generators transcribed by T7 RNAP, I attempted to put the two devices together in the same cell. I was unable to cotransform the T7 autogenerator (on pSB4C5) with any of the T7 RNAP-transcribed O-ribosome generators (on plasmids conferring tetracycline resistance and a p15a ori) into *E. coli* TOP10. No colonies were observed on plates containing the appropriate antibiotics after sev-
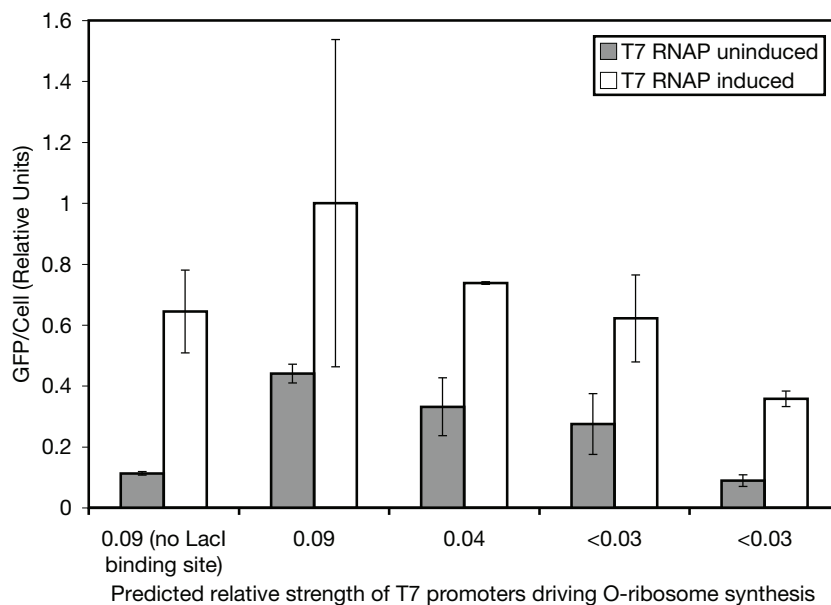
Figure 4-8: T7 RNAP can transcribe orthogonal rRNA that is assembled into functional ribosomes. Furthermore, O-ribosome levels can be modulated by changing the strength of the T7 promoter upstream of the rRNA. I measured the relative expression of GFP of *E. coli* BL21(DE3) cultures with or without of 0.8 mM IPTG. The cells contained an O-ribosome generator with transcription regulated by a range of T7 promoters. Cultures also contained a reporter that uses *E. coli* transcription and orthogonal translation. Data was measured and processed as described in the Appendix E. The data points represent the mean of three independent cultures and the error bars represent the standard deviation for the three samples.

eral days. In a control cotransformation, I obtained > 1000 transformants when transforming the T7 autogenerator with the P_{BAD}-regulated O-ribosome generator described in Chapter 3. This result suggested that a specific interaction between the T7 autogenerator and the T7 RNAP-transcribed O-ribosome generator rendered cells transformed with the two devices non-viable.

A likely explanation for my inability to cotransform TOP10 with the combined network of O-ribosomes and T7 RNAP was that there was more transcription and translation of the network species than could be tolerated by *E. coli* TOP10. I attempted the same transformation into *E. coli* D1210 which has a lacI^q genotype. The increased LacI levels might decrease activity of the T7 promoter to non-toxic levels. I obtained cotransformants in D1210 that stably maintained both plasmids.

To investigate whether the O-ribosomes and T7 RNAP were being expressed in D1210, I transformed a third plasmid into the cells, carrying the orthogonal GFP reporter described in Chapter 3, BBa_E71205. As a control, I also transformed cells with a similar plasmid to that carrying BBa_E71205 but with no reporter. As further controls, I cotransformed the GFP reporter with the T7 autogenerator and the T7 RNAP-transcribed O-ribosome generator separately. Relative levels of GFP per cell were estimated as described in Appendix E and the data is shown in Figure 4-9. Although measured GFP levels were low relative to those seen for VM1.0 in Chapter 3, it was clear that both the T7 autogenerator and the O-ribosome device were necessary to lead to expression of GFP. Since I showed in Chapter 3 that the GFP reporter device required the presence of both T7 RNAP and O-ribosomes to produce GFP, it appears likely that VM2.0 is expressing both T7 RNAP and O-ribosomes in D1210. Given the low level of measured GFP, it would appear that either there are few available T7 RNAP to transcribe the reporter, or few O-ribosomes available to translate the reporter. Such a result may have been due to low total levels of either T7 RNAP or O-ribosomes in the cells. Alternatively, either T7 RNAP or O-ribosomes may have been titrated out through competition among the various T7 promoters and orthogonal RBS sequences in the VM2.0 network and reporter.

The version of the VM2.0 network presented here was stable and could be used to

activate a reporter. Ideally, however, the VM2.0 network would be stable in a common cloning strain such as *E. coli* TOP10. I hypothesized that the reason the network was stable in D1210 but not TOP10 was because the network was not making enough LacI to reduce the activity of the T7 promoters and the wild-type level of LacI in TOP10 was also insufficient to repress the network. To further confirm that the ability of D1210 to stably maintain VM2.0 was related to the LacI^q phenotype, I added 0.8 mM IPTG to growing D1210 cells carrying the VM2.0 device. Negligible growth of the cultures was observed following the addition of IPTG (data not shown), lending support to the theory that the stability of VM2.0 in D1210 was LacI dependent.
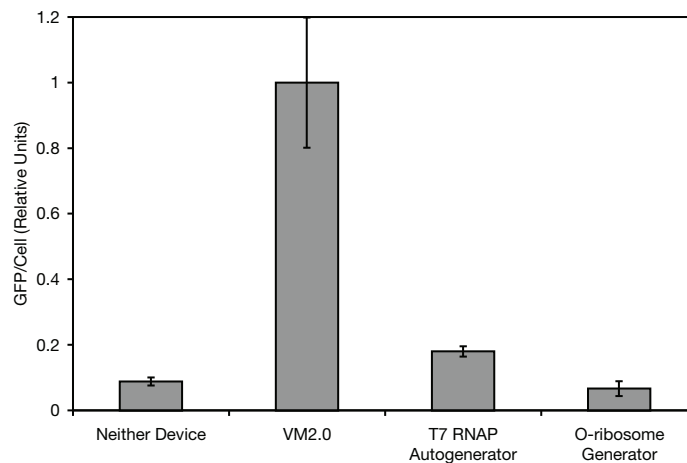
Figure 4-9: *E. coli* D1210 carrying the VM2.0 mutants can activate a reporter. The cultures contained neither, both, or one of the devices that comprise VM2.0. All cultures carried the GFP reporter. Both devices must be present to produce significant GFP expression. Data was measured and processed as described in the E section. The data represents the mean of three independent cultures and the error bars represent ± one standard deviation.

4.4.6   VM2.2, a redesigned version of VM2.0

I redesigned the VM2.0 networks with the goal of producing a network that was stable in *E. coli* TOP10. Given that the device was stable in *E. coli* strain D1210 but not TOP10, I focused my attention on increasing the level of LacI produced by the network. I identified two approaches to increasing LacI production. The first approach was to increase the rate at which the *lacI* message was translated by the

90

T7 autogenerator. The second approach was to focus on increasing the rate at which the *lacI* message was being transcribed.

The efficiency with which a message is translated is determined, in part, by the secondary structure of the 5' end of the transcript [100]. I used mfold [142] to analyze the secondary structure of the first 180 bases of the polycistronic mRNA carrying the *lacI* and T7 RNAP messages. The secondary structure with the minimum free energy as calculated by mfold is shown in Figure 4-10. Two hairpin structures of interest can be observed. The first is formed from the partly symmetric LacI binding site just downstream of the transcription start site of the T7 promoter. The second hairpin, including seven Watson-Crick base pairs, is formed from the mixed XbaI/SpeI site resulting from the BioBrick cloning scheme and the orthogonal RBS sequence. Such partial occlusion of the RBS is known to reduce the efficiency of translation of the message [33]. It was possible that modifying the sequence of the message to disrupt one or both of these hairpins might increase the efficiency with which the *lacI* message is translated. I identified single base pair substitutions, insertions, and deletions that would disrupt the hairpin formed by the RBS and the BioBrick mixed XbaI/SpeI site (`TACTAGAG`). One such mutation was a deletion of the last `G` in the mixed XbaI/SpeI site. A plasmid carrying that mutation was successfully cotransformed with the O-ribosome generator plasmid into *E. coli* TOP10 and both plasmids could be stably maintained. Experiments are underway currently to determine whether the network expresses both T7 RNAP and O-ribosomes.

In the second approach, I considered making changes at the levels of the component parts of the T7 autogenerator rather than at the level of the DNA sequence. I designed a new T7 autogenerator (termed VM2.2) from preexisting parts (Figure 4-11). The rationale for the new design was as follows -

1. Since the genes for LacI and T7 RNAP were cotranscribed in the original T7 autogenerator design, simply increasing the strength of the T7 promoter would likely increase expression of both LacI and T7 RNAP which might be counterproductive. As a result, the redesigned VM2.2 network separated the two genes into separate transcription units, each with a T7 promoter and terminator, Tϕ.
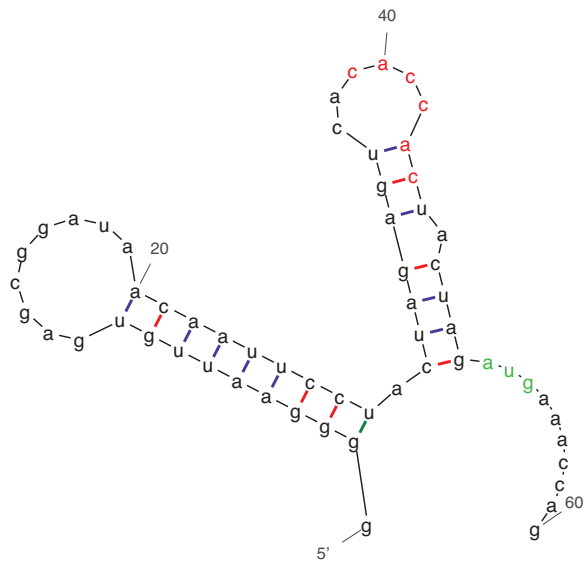
Figure 4-10: Secondary structure of the 5' end of the LacI mRNA transcribed from the LacI-repressible T7 promoter in VM2.0. mfold was used to find the minimum energy fold of the first 180 bases of the mRNA (only the first 60 bases are shown). Two hairpins are formed. One hairpin is formed due to the partially symmetric LacI binding site. The second hairpin is formed from the BioBrick scar site (between promoter and RBS) and the RBS itself. Bases expected to interact with the 16s rRNA of the orthogonal ribosome are highlighted in red and the start codon is highlighted in green.

I used the same T7 promoter (BBa_R0184) to transcribe T7 RNAP as used in the original T7 autogenerator.

2. LacI expression might be increased by using a stronger T7 promoter than that used in the original T7 autogenerator. I tested two different T7 promoters described in Chapter 3 (BBa_R0180 and BBa_R0182) with expected strengths of 0.72 and 0.3 relative to the T7 consensus promoter sequence [64].

3. The order of the transcription units was likely to be important since the downstream unit may experience read-through transcription from the upstream unit. I presumed this to be the case since the Tϕ terminator is known to have an efficiency of approximately 80%. Given that the goal of the redesign was to increase LacI expression rather than T7 RNAP expression, I placed the T7 RNAP transcription unit upstream of the LacI unit.
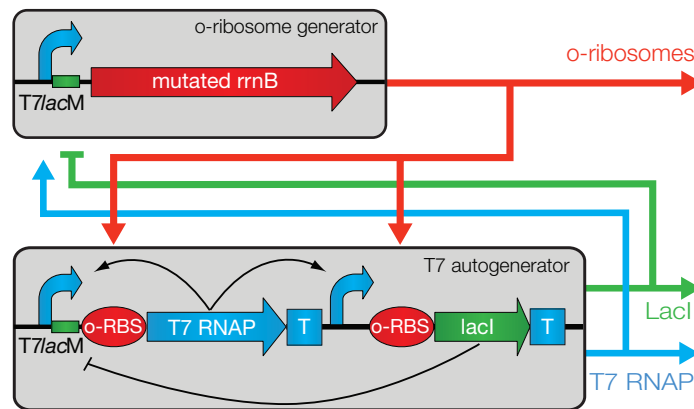
Figure 4-11: The VM2.2 network schematic. VM2.2 consists of two devices, the T7 autogenerator using orthogonal translation and a T7 promoter-regulated O-ribosome generator. The O-ribosome generator was identical to that of the VM2.0. The redesigned T7 autogenerator used a separate, stronger T7 promoter to transcribe the *lacI* gene.

To investigate whether either of the newly designed T7 autogenerator devices might be stably maintained in *E. coli* TOP10 in the presence of a T7 RNAP-transcribed O-ribosome generator, I performed cotransformations with different variants of both devices in a combinatorial fashion. I tested the P_{BAD}-regulated O-ribosome generator described in Chapter 3 and the four T7 promoter-regulated O-

ribosome generators described in Section 4.4.4. In addition to the two redesigned T7 autogenerator devices, I included the original T7 autogenerator device in the co-transformation experiment. The original T7 autogenerator did not produce colonies when transformed with any T7 promoter-regulated O-ribosome generator although the same T7 autogenerator was cotransformed successfully with the P_{BAD}-regulated O-ribosome generator, confirming the results described in Section 4.4.5. Both the redesigned T7 autogenerators were cotransformed successfully with all O-ribosome generators tested. Colonies from cotransformations with the redesigned T7 auto-generators were used to inoculate LB cultures containing the appropriate antibiotics. The cells appeared to be viable and the plasmids were stably maintained during batch culture growth.

Having obtained stable clones of VM2.2 in *E. coli* TOP10, I proceeded to test whether the T7 RNAP and O-ribosomes were expressed. Similar to before, I trans-formed different combinations of the devices with a GFP reporter device that required both T7 RNAP and O-ribosomes to express detectable levels of GFP. I grew the re-sulting strains in a multi-well fluorimeter as described in Appendix E, taking repeated measurements of fluorescence and absorbance. The measurements showing fluores-cence as a function of cell density (measured in units of OD600) are shown in Figure 4-12. The data suggests that significant accumulation of GFP occurs only in the presence of both the T7 autogenerator and the O-ribosome generator. It should be noted that the presence of the T7 autogenerator appears to affect the relationship between cellular autofluorescence and OD. As evidence, compare the black and green data points to the blue and cyan data series, all of which are expected to have little or no fluorescence due to GFP expression. Given how similar the green and black data points are to each other and how similar the blue and cyan data points are to each other, I infer that there is little GFP being made when only one of the devices is present.

Growth curves for the same cultures are shown in Figure 4-13. The growth curves show that there is a lag in growth for cells carrying VM2.2 but little decrease in maximum growth rate. Cells carrying both VM2.2 and a GFP reporter device exhibit

94
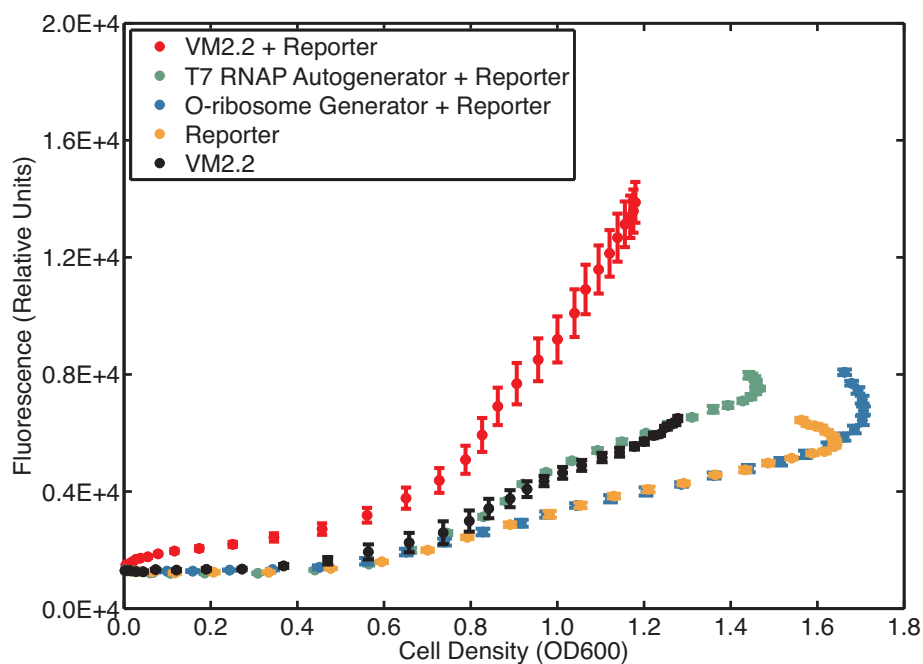
Figure 4-12: *E. coli* TOP10 cells carrying VM2.2 express can express GFP from a reporter. Relative expression of GFP as a function of cell density (measured in units of OD600) from *E. coli* TOP10 cultures was measured. The cultures contained neither, both, or one of the devices that comprise VM2.2, in addition to the orthogonal reporter. All devices must be present to produce significant GFP expression (red filled circles). Note that cells containing the T7 autogenerator device produce a different level of autofluorescence to cells not containing the T7 autogenerator as can be observed by comparing the black and green filled circles to the blue and gold filled circles that represent cultures with no T7 autogenerator. Data was measured and processed as described in Appendix E. The data points represent the mean of three independent cultures (each measured in triplicate) and the error bars represent the 95% confidence interval for the mean.

both a lag in growth and also a reduction in maximum growth rate.
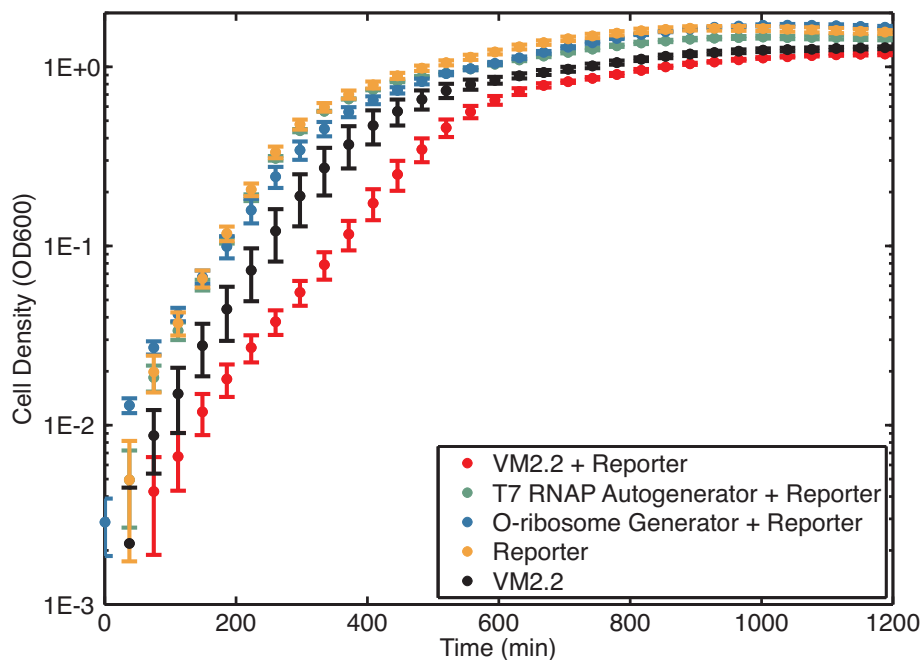
Figure 4-13: Growth curves for *E. coli* TOP10 cultures carrying the same devices as shown in Figure 4-12. VM2.2 led to a lag in cell growth and also a reduction in growth rate (black filled circles). VM2.0 and the reporter device led to a more marked decrease in growth rate (red filled circles). Data was measured and processed as described in the Appendix E. The data points represent the mean of three independent cultures (each measured in triplicate) and the error bars represent the 95% confidence interval for the mean.

While the fluorescence data presented above suggested that both T7 RNAP and O-ribosomes were being expressed, the data could be explained by a scenario where both T7 RNAP and O-ribosomes were not being expressed by each other but rather were being expressed at low levels due to basal transcription and possibly translation by *E. coli* RNAP and ribosomes. In that scenario, the presence of either device alone might not lead to GFP expression whereas the presence of both devices would lead to GFP expression. However, two experiments described earlier suggested that the two devices are likely to transcribe and translate each other. First, in Section 4.4.4, I demonstrated that T7 RNAP can transcribe orthogonal rRNA that is assembled into functional ribosomes. Second, in Section E.6, I demonstrated that orthogonal ribosomes can translate T7 RNAP.

To confirm more explicitly that the T7 RNAP and O-ribosomes were regulating each other in the VM2.2 system, I made use of reporter devices described in Chapter 3 that use either orthogonal transcription or orthogonal translation but not both. For example, BBa_E7108 is a GFP reporter with a T7 promoter and an *E. coli* RBS. I transformed this reporter device with the T7 autogenerator in strains either with or without a T7 promoter-regulated O-ribosome generator. If basal *E. coli* transcription was responsible for the expression of T7 RNAP, then only the T7 autogenerator should be necessary activate the reporter, BBa_E7108. As can be seen by comparing the first two bars in Figure 4-14, the T7 autogenerator is not sufficient to cause significant expression from the reporter device, while the addition of the O-ribosome generator leads to high level of GFP expression. From this result I conclude that the orthogonal ribosomes must be translating the T7 RNAP message.

I also performed the converse experiment, in which I make use of a reporter device (BBa_E70202) with an *E. coli* promoter and an orthogonal RBS and transform that device with the O-ribosome generator either in the presence or the absence of the T7 autogenerator. The experiment yielded a similar result: there is a small but significant increase in GFP fluorescence when the T7 autogenerator is added to cells containing the reporter and the T7 promoter-regulated O-ribosome generator (compare bars three and four in Figure 4-14).

Also shown in Figure 4-14 is the expression of the GFP reporter device requiring both O-ribosomes and T7 RNAP. The relative "capacity" of the different transcription and translation systems can be estimated by comparing bars two, four, and five in Figure 4-14. At the level of transcription, capacity refers to the ability of the T7 RNAP-promoter combination to transcribe the target gene. At the level of translation, capacity refers to the abilty of the O-ribosomes and the o-RBS to translate the target message. By comparing bars two and five, one can infer that the translational capacity of the O-ribosomes and the o-RBS is significantly less than the translational capacity of wt *E. coli* ribosomes and the *E. coli* RBS used in BBa_E7108. A comparison of bars four and five suggests that the capacity of the T7 transcription system (T7 RNAP and a weak T7 promoter) is significantly higher than that of the *E. coli*
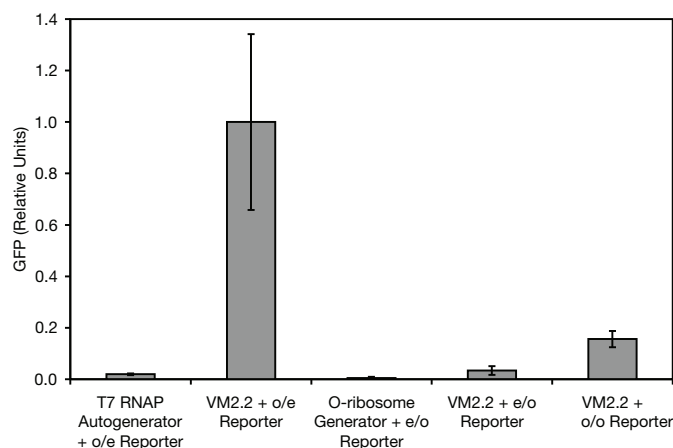
Figure 4-14: Expression of T7 RNAP by the VM2.2 network depends on the presence of O-ribosomes. Similarly, expression of O-ribosomes depends on the presence of T7 RNAP. Cultures contained a reporter device that uses orthogonal transcription and *E. coli* translation (o/e reporter) or *E. coli* transcription and orthogonal translation (e/o reporter), or orthogonal transcription and translation (o/o reporter). Additionally, the cultures contained combinations of the T7 autogenerator and the O-ribosome generator. The data shows that not only the T7 autogenerator but also the O-ribosome generator are necessary for high expression of the o/e reporter, indicating that T7 RNAP is in fact being produced almost entirely by O-ribosomes. Similarly, both devices are necessary for expression from the e/o reporter, although expression is significantly lower. For reference, GFP expression from the o/o reporter is also shown. Data was measured and processed as described in the Appendix E. The data points represent the mean of three independent cultures and the error bars represent ± one standard deviation.

transcription system (*E. coli* RNAP and the *E. coli* promoter used in E70202).

4.4.7 Variants on the VM2.2 network

As described in Section 4.4.6, clones were obtained for several combinations of the various T7 promoter-regulated O-ribosome generators and redesigned T7 autogenerators. In addition to the pair of devices characterized in Section 4.4.6, I transformed several other combinations of T7 autogenerators and O-ribosome generators with the reporter that requires both T7 RNAP and O-ribosomes for GFP expression (BBa_E71205) and the appropriate empty vector control. I grew cultures for all combinations of two T7 promoter-regulated O-ribosome generators and two redesigned T7 autogenerator variants in LB medium and measured the fluorescence to OD600 ratio as a surrogate for the level of GFP per cell. The data is shown in Figure 4-15.
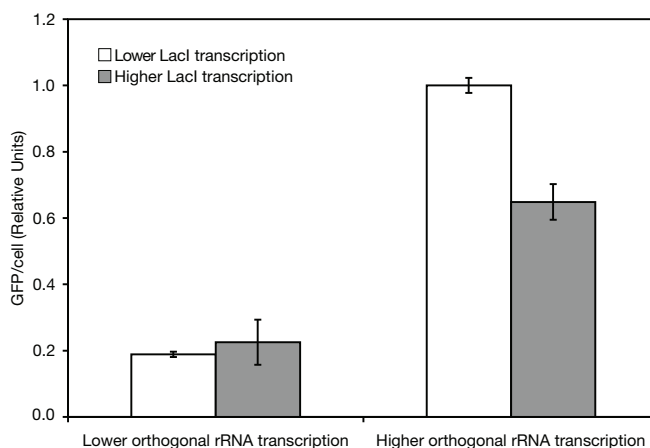
Figure 4-15: Changing the strength of the promoters regulating O-ribosome and LacI synthesis can modulate the protein synthesis capacity of VM2.2. Relative levels of GFP expression in cells bearing variants of the VM2.2 network and a GFP reporter was measured. Increasing the transcription of the orthogonal rRNA increased GFP expression. For the pair of bars on the right, an increase in *lacI* transcription reduced GFP expression. Data was processed as described in the Appendix E. The data points represent the mean of three independent cultures (each measured in triplicate) and the error bars represent ± one standard deviation.

All strains showed significant levels of GFP expression. Furthermore, there were clear differences in expression between the devices. It is of interest to attempt to

correlate the different observed behaviors to the differences among the devices. First, increasing the strength of the T7 promoter transcribing the orthogonal rRNA leads to a clear increase in GFP expression. We might expect that increasing the expression of O-ribosomes will increase the translation rate of T7 RNAP. Recalling the model presented in Section 4.2, increasing the translation rate of T7 RNAP is represented by increasing k_t and so the observed experimental data is in qualitative agreement with the model predictions (Figure 4-3). Comparing the second pair of grouped bars, reducing the *lacI* transcription lead to higher GFP expression. Increasing the transcription rate of *lacI* is equivalent to increasing the ratio of LacI expression to T7 RNAP expression (r) in the simple model presented in Section 4.2. Again, the observed experimental data is in qualitative agreement with the model predictions (Figure4-2) I was not able to distinguish a change in GFP expression as *lacI* transcription changes for the first pair of grouped bars. One possible explanation for this observation is that the capacity of O-ribosomes when being transcribed at the lower level may have limited the expression of GFP levels.

4.5   Discussion

In the preceding sections, I have demonstrated that the T7 RNAP is self-transcribing and also transcribes the orthogonal rRNA. I have demonstrated that the O-ribosomes translate the T7 RNAP. Further, I have shown that there is little transcription and translation of the engineered network by *E. coli* RNAP and wt *E. coli* ribosomes. The VM2.0 and VM2.2 networks are the first instances of an engineered network of self-regulating ribosomes and RNAP. Finally, I made deliberate changes to make the network stable in *E. coli* TOP10.

The ability of the VM2.2 network to function without the need for *E. coli* RNAP and wt *E. coli* ribosomes (except for basal transcription to start the network) has interesting engineering potential. For example, the behavior of the network is likely to be less sensitive to changes in RNAP and ribosome availability as a function of growth rate under conditions where RNAP and ribosome levels limit gene expression. This

characteristic of the VM network may be useful for operation of devices in stationary phase or during periods of varying nutrient availability. As a second example, by modifying the VM network, an engineer could tailor the levels or the regulation of T7 RNAP and O-ribosomes within certain limits yet to be determined to suit a particular application with the expectation that the behavior of the cell will not be significantly affected. By contrast, modifications to the levels or regulation of *E. coli* RNAP or ribosomes is likely to have a significant impact on the behavior of the cell. As a final example, the regulatory network governing T7 RNAP and O-ribosome synthesis is significantly simpler than that regulating *E. coli* RNAP and ribosome synthesis. As a result, the response of the VM network to consumption of RNAP and ribosomes by an engineered system should be easier to predict than the corresponding response of the *E. coli* network.

Where possible, I constructed all networks using standard biological parts. As demonstrated in Section 4.4.6, the use of standard biological parts made it routine to construct redesigned variants of the T7 autogenerator. That those changes had the desired result is a demonstration that the BioBrick framework of standard biological parts can be a powerful tool for rapidly protoyping networks that exhibit a desired behavior. Nevertheless, thoughtful choice of network design can make the construction of subsequent network variants more routine. For example, the first T7 autogenerator transcribed the T7 RNAP and the LacI repressor on a single transcript, thereby directly coupling the transcript levels of both genes. The subsequent redesign of the autogenerator in which both genes were transcribed separately, made it easier to independently modify the transcript levels of the two genes.

There are easily imagined ways in which the VM2.2 network might be improved upon. First, the capacities, both absolute and relative, of the transcription and translation might be modified. For example, data presented in Figure 4-14 suggests that the T7 transcription system has a high capacity for transcription, whereas the orthogonal translation system has a relative low capacity. As a result, unnecessary transcription is occurring that is not leading to protein synthesis. It might be worthwhile to better balance the relative capacities of the transcription and translation

101

systems. Many approaches could be pursued to better balance the capacities. For example, a T7 RNAP mutant that is less processive might reduce the capacity of the transcription system [82]. Alternatively, new RBS sequences or mutant 16s rRNA sequences could be selected for that initiate translation at a higher rate and so make better use of the capacity of the transcription system. Note, it is possible that the RBS sequence used here might be much more active in a sequence context different from that tested. Finally, it would be interesting to evaluate what fraction of orthogonal rRNA is being assembled into functional O-ribosomes. If this assembly process is inefficient, reducing the level of orthogonal rRNA transcription or other approaches to increasing assembly efficiency might be usefully pursued.

The "usability" of the network might also be improved. For example, VM2.2 currently exists on two separate plasmids. If the two devices were carried on a single plasmid, or if both devices were inserted into the chromosome, system engineers would have more freedom to make use of plasmid origins of replication and antibiotic markers in engineered biological systems. It would also be interesting to observe how the behavior of the network changed as the number of the devices was reduced. Another interesting development of the VM2.2 network would be a variant that included a user-controllable on/off switch since the current implementation of the network is constitutively on. An independently controllable (using, for example, a repressor other than LacI) T7 promoter regulating transcription of the orthogonal rRNA might allow the network to be turned on and off at will. Alternatively T7 RNAP activity could be regulated via a temperature sensitive mutant of T7 RNAP [78]. Finally, in this thesis I exclusively use GFP as a reporter of the activity of orthogonal RNAP and ribosomes. It would be helpful to demonstrate that the virtual machine networks can efficiently synthesize a range of other target genes.

Finally, I re-emphasize the fact that an engineered biological system using any of the virtual machine networks still shares many resources with the cellular chassis, including ribosomal proteins, nucleotides, and charged tRNA (to list only a few). Transcription of orthogonal ribosomal RNA presumably upregulates the expression of ribosomal proteins, which must be translated by wt *E. coli* ribosomes. It may be

possible to supplement the media with certain resources such as nucleotides and amino acids such that the intracellular levels of those resources are at a saturating level. However, an open question remains as to whether the degree of decoupling brought about by independent populations of RNAP and ribosomes alone is significant. In Chapter 5, I discuss ways to answer this question.

It is worth considering certain differences between the VM1.0 network and the VM2.2 network. For example, in the VM1.0 network, usage of T7 RNAP and O-ribosomes by an engineered systems does not affect the synthesis of the T7 RNAP and O-ribosomes. This decoupling between engineered system and VM makes the interactions between system and VM easier to predict; the disadvantage is that a demand for T7 RNAP and O-ribosome levels is not balanced by an upregulation in their synthesis. In the VM2.2 network, I have made the reverse compromise, T7 RNAP and O-ribosome levels should be robust to changes in demand at the cost of increased coupling between system and VM, since both compete for the same T7 RNAP and O-ribosomes. A more detailed analysis of an optimal compromise between coupling and robustness, making use of experiments on the different VM networks and a detailed computational study, is warranted.

We can also make some simple comparisons of the VM1.2 and VM2.2 networks. Both networks used *E. coli* TOP10 as the cellular chassis. The same plasmids were used in each network to carry the T7 RNAP generator and O-ribosome generator. Finally, the experimental protocols to measure GFP synthesis as a function of culture growth were performed identically in both cases (compare Figures 3-12, 3-13, and 4-12, 4-13). Differences in observed behavior should be due to differences in the network architecture (arabinose induction of T7 RNAP and O-ribosomes in VM1.2 and auto-regulation for VM2.2). A comparison suggests that the VM2.2 network leads to much lower GFP expression than VM1.2 under the conditions tested even though the VM2.2 network is always on while the VM1.2 network was induced immediately prior to addition of cultures to the multi-well fluorimeter. This result suggests that either levels of orthogonal RNAP and ribosomes are lower in VM2.2 or that competition between the various T7 promoters and O-RBS in the VM2.2 network reduce the

availability of T7 RNAP and O-ribosomes to the GFP reporter. However, I note that the VM2.2 network reaches a significantly higher final culture density possibly because the levels of GFP synthesis were lower.

Chapter 5

Future work & conclusions

In this thesis I adopted and adapted a framework for describing engineered devices and applied that framework to an early standard biological part, BBa_F2620. Device descriptions, such as that provided for BBa_F2620, will become both simpler and much more useful if the behavior of the device is independent of the behavior of the cellular chassis, since the description could be expected to hold across many chassis and culture conditions. I adopted the concept of a chassis to propose an ideal interaction between a host cell and an engineered biological system where the cellular chassis provides a constant pool of resources to the system, and the behavior of the system and chassis is otherwise largely decoupled. To that end, I engineered devices that provide a second population of RNAP and ribosomes in *E. coli* for exclusive use by simple gene-expression devices. Finally, I showed that the orthogonal RNAP and ribosomes can regulate their own synthesis. By describing a receiver in a quantitative manner that supports predictable reuse of the device, and designing cellular chassis that may interact with engineered biological systems in a more predictable way, my contributions should spur the development of engineered biological systems that behave predictably and reliably. In the remainder of this chapter I describe work underway and work that might be embarked upon to realize this goal.

5.1 Characterizing and making use of the virtual machine

In Chapters 3 & 4 I use networks of T7 RNAP and O-ribosomes to transcribe and translate a simple engineered system. Future experiments should examine the behavior of the orthogonal RNAP and ribosome networks in more detail. For example, do the levels of the orthogonal RNAP and ribosomes reach a steady state during culture growth? If so, how robust is that steady state to variations in growth rate, resource availability or other variations in culture conditions? Are the levels of free orthogonal RNAP and ribosomes held fixed despite a demand for orthogonal RNAP and ribosomes by an engineered biological system as expected for a feedback regulated network? Finally, can the behavior of the network and an associated engineered system be predicted by a computational model?

To evaluate the utility of the virtual machine, experiments should be designed and performed that would clarify whether an engineered biological system that uses the virtual machine exhibits less dependence on the state of the chassis than an engineered system that uses the same RNAP and ribosomes as the chassis. This question could be addressed by comparing the behavior of reporters that use orthogonal transcription and translation to those that use chassis transcription and translation as a function of growth rate and other culture conditions. At the same time, the behavior of the chassis should be observed to determine whether, for a given level of reporter gene expression, the cellular chassis experiences less of a reduction in growth when reporter expression is driven by orthogonal RNAP and ribosomes. Such experiments could be elegantly performed by placing both reporters on the one plasmid, so the behavior of the reporters could be compared independent of confounding sources of external noise [43].

The experiments just proposed would begin to suggest whether the major sources of coupling between the system and the chassis are at the level of machinery (RNAP and ribosomes) or materials (nucleotides, amino acids, etc.). Elucidation of the limiting components shared between the chassis and the system would inform and moti-

vate the design of new cellular chassis that decoupled the consumption of the limiting species in the transcription and translation processes. For example, under conditions where amino acid consumption by the system affects the behavior of the chassis or vice versa, the use of non-natural amino acids and orthogonal tRNA and tRNA synthetase pairs could be used to decouple the demand of the system and the chassis [73, 137, 27]. Alternately we could establish culture conditions such that resources that are shared by an engineered system and a cellular chassis do not limit the behavior of either system or chassis. For example, under some culture and nutrient conditions, ribosomes limit expression of a heterologous protein [129] so the sharing of other gene expression resources may not be critical. It is important to stress that regardless of how many decoupled pools of resources are provided to the system, the coupling between the system and the chassis due to the unavoidable sharing of central metabolic pathways is likely to remain for the foreseeable future. Whether that fundamental coupling will ever permit largely independent operation of the chassis and the system is an open research question.

In this thesis I have implemented a second population of RNAP and ribosomes into the cell that are orthogonal to the endogenous RNAP and ribosomes. In the future, still more orthogonal populations of RNAP and ribosomes could be added to the cellular chassis allowing for independent control of transcription and translation of multiple target genes. T7 mutants with orthogonal specificities to the wt T7 RNAP have been described [105] and multiple orthogonal pairs of O-ribosomes have also been described [99].

Since the cytoplasm of a microbe such as *E. coli* is a single space shared, I make use of biochemical specificity to decouple the requirements of the system and the chassis for RNAP and ribosomes. A more technically challenging approach would make use of spatial separation to decouple the resource requirements of a system and a chassis. Spatial separation of an engineered biological system would allow fine control of the interactions between the system and the rest of the cell. Also, independent populations of a component could be maintained in two spatial compartments allowing reuse of components. Eukaryotic cells make use of or-

ganelles that spatially separate particular processes from the rest of the cell via a membrane. For example, mitochondria are isolated compartments responsible for much of the ATP synthesis of a eukaryotic cell. Similarly, chloroplasts are isolated compartments dedicated to photosynthesis. Engineers might consider co-opting a natural organelle to insulate and confine an engineered biological system within a larger self-replicating machine. Alternatively, a synthetic organelle (a "Synthosome") might be constructed for the sole purpose of housing engineered biological systems (`http://parts.mit.edu/igem07/index.php/UCSF/Organelle\_Intro`).

A software virtual machine allows the same code to operate on different hardware platforms. A useful biological virtual machine should allow an engineered system to operate in different cellular chassis. For example, an engineer might want to move a particular engineered system from *E. coli* to *Bacillus subtilis*. Although *E. coli* and *B. subtilis* have an identical consensus promoter sequence for the exponential phase RNAP holoenzmye, *B. subtilis* promoters must be much closer to the consensus sequence than *E. coli* promoters due to differences in the RNAP holoenzyme [6, 134]. As a result, promoters that work in *E. coli* work differently or not at all in *B. subtilis*. The anti Shine-Dalgarno sequence of the 16s rRNA is identical in both species. However, differences between the two organisms exist at the level of translation; *B. subtilis* lacks the S1 r-protein that contributes to the ribosome-mRNA interactions [104]. By porting a virtual machine network such as VM2.2 from *E. coli* to *B. subtilis*, we might provide the beginnings of a common interface for engineered biological systems in both organisms. Since the behavior of the VM2.2 network is expected to be largely independent of the endogenous RNAP and ribosomes, the behavior of a T7 promoter might be more similar in both organisms than an *E. coli* promoter. I can be less confident that an O-RBS would behave more similarly in both organisms due to the lack of an S1 protein in *B. subtilis* but a suitably designed O-RBS that did not interact with S1 might be less sensitive to the differing ribosomal protein usage. Clearly, there are other factors that might make a system behave differently in *B. subtilis* relative to *E. coli* such as different propensities for anti-termination or RNA degradation but the porting of orthogonal RNAP and ribosomes from *E. coli* to

B. subtilis might represent the beginning of the construction of a standard interface to both species.

For VM2.2 or a similar network to be successfully ported to *B. subtilis*, the orthogonal rRNA must to be compatible with the *B. subtilis* rRNA processing enzymes and r-proteins. To my knowledge this compatibility has not been tested for *E. coli* rRNA and *B. subtilis* ribosomal proteins, but there are multiple reports of rRNA compatibility between organisms. For example, *B. stearothermophilus* 16s rRNA has been successfully assembled with *E. coli* ribosomal proteins to form 30s subunits [90]. Also Squires and coworkers have shown that *Salmonella typhimurium* and *Proteus vulgaris* rRNA operons can be transcribed in *E. coli* and the resulting rRNA assembled into functional ribosomes [8].

Chan and coworkers demonstrated a viable "refactored" bacteriophage T7 genome [25]. The refactoring made the genome more amenable to human understanding while retaining all genetic elements thought to be necessary for phage viability. We have a rapidly developing ability to manipulate DNA fragments on the scale of bacterial genomes [51, 66]. As described in Chapter 1, re-engineered or refactored cellular chassis with more predictable behavior are likely to make the behavior of engineered biological systems carried by the chassis more predictable also. Regulation of the synthesis of *E. coli* ribosomes and RNAP might be re-engineered to make the availability of those species more useful or more predictable to a biological engineer. A likely consequence of such a refactoring would be cellular chassis with a compromised ability to grow or adapt to a changing environment. Such a tradeoff might be worth making as a stepping stone to the construction of cellular chassis that are both highly predictable and also robust. By analogy, early steam-powered vehicles were initially more unreliable than horse-drawn vehicles but the greater ease of engineering engines caused fuel-powered vehicles to rapidly exceed the scale and range of horse-drawn vehicles.

A network of RNAP and ribosomes similar to VM2.2 might be a starting point for a refactored RNAP and ribosome synthesis network powering a cellular chassis. Researchers have constructed *E. coli* strains in which all seven chromosomal rRNA

operons have been either deleted or disrupted [7]. A plasmid-encoded rRNA operon with wild-type regulation (the P1 and P2 promoters) is responsible for the synthesis of all rRNA used by the cell. Modifications to the regulation or sequence of ribosomal rRNA can be made by swapping in different plasmids carrying modified rRNA operons. Such a strain represents an ideal platform for testing novel regulation networks for ribosome synthesis. Together with an undergraduate, Matt Gethers, I have obtained several variants of the rrn deletion strains (under the terms of an MTA from Cathy Squires) and have begun construction of an rRNA operon regulated by a T7 promoter in a plasmid that can be swapped into the rrn deletion strains. The rRNA operon encodes wild-type 16s rRNA that can translate any *E. coli* gene. This ribosome generator, together with the T7 autogenerator using *E. coli* RBS described in Section 4.4.2 should form a self-regulating network that could supply ribosomes to the rrn deletion strains. With some tuning of network parameters the rrn deletion strains might be propagated with a ribosome population regulated with an entirely synthetic network, decoupled from growth rate and stress regulation. Such strains are unlikely to be robust or genetically stable but might represent an important stepping stone towards a robust and highly predictable cellular chassis.

5.2 Optimizing device design and description for use with cellular chassis

In Section 2.4, I discussed several useful steps that engineers might take towards building families of engineered biological parts and devices that can be assembled into higher-order systems that carry out a useful function. In this section, I focus on how the design and description of engineered devices might be optimized to facilitate predicting the interactions of those devices with cellular chassis.

The input and output signal levels of the receiver were not set by design, but as an unpredictable outcome of the parts that comprise the receiver. I chose the copy number of the device based on empirical rules of thumb of how many copies of a

reporter are necessary to ensure expression of GFP that is readily detectable while not being overly toxic to the chassis. Both the plasmid copy number and the input and output signal levels of the receiver are parameters that might be optimized in order to produce a device that is useful but also places a minimal demand on the cellular chassis. Work is underway to begin to construct a model of device behavior that elucidates the tradeoff between useful signal levels and the demand placed on the cellular chassis [118].

In Section 2.3, I described measurements in absolute molecular units of the output of the receiver device. I used those measurements to estimate the transcriptional output demand of the device in terms of nucleotides consumed per second and RNAP sequestered when the receiver is steady state operation. These simple estimates might be refined and extended to other demands placed by the receiver on the cellular chassis. Ideally, an estimate of the total demand of an engineered system could be generated by a Computer Aided Design (CAD) package based on the parts comprising an engineered biological system. Such predictions would allow an engineer to make an informed estimate of the demand being placed on the cellular chassis and to compare the demand of different systems.

5.3   Demand modeling

An accurate model capable of predicting the total demand that an engineered biological system places on a cellular chassis would be useful, but that use would be limited by our lack of information on the capacity of a cellular chassis to supply the applied demand. A complementary effort to specifying the resource demand of an engineered system would be to develop a model of the chassis resource pools and the response of the chassis to a depletion of those resource pools. To begin to address this question, I have established a collaboration with Javier Carrera-Montesinos, Guillermo Rodrigo, and Alfonso Jaramilla (at the Ecole Polytechnique, Paris) to construct a model of the cellular chassis. We have begun by constructing a simple empirical model of the cellular chassis based on published data on chassis resources as a function of chassis

growth rate [18, 98]. The published data included estimates of the composition of cells in different media (and hence different growth rates) and calculated the intracellular concentration of species such as RNAP, ribosomes, amino acids, nucleotides etc. From this data we derive empirical relations relating a reduction in cellular resources to a reduction in growth rate of the cells. Next, we model an engineered system that consumes certain cellular resources as a function of plasmid copy number, promoter strength etc. The resource-growth rate relations can then be used to predict the reduction in growth rate of a cellular chassis supplying the engineered biological system. Surprisingly, using this simple model we can make predictions on the growth rate reduction of cellular chassis carrying the receiver device (BBa_F2620) or other published systems [95] that agree quite closely with the experimental measurements. Such simple models could be usefully extended by accounting for the various feedback loops that regulate the synthesis of cellular gene expression resources. These feedback loops presumably modulate the availability of cellular resources in response to resource consumption.

5.4   Outlook

To summarize, we should continue to develop engineered parts and systems that are ever more reliable and chassis that are both designed to be more reliable and for which better models exist. Going forward, both systems and chassis should be engineered to work better together. I envision a future in which libraries of standardized cellular chassis exist where the performance characteristics of each have been carefully described. Based on a quantitative understanding of the system and available chassis choices, engineers can choose a chassis that is well-matched to the application at hand, yielding a self-replicating machine that performs efficiently and predictably.

Appendix A

Methods and Supplementary Information for BBa_F2620

A.1 Genetic material, manipulations, strains, and reagents

I constructed BBa_F2620 via BioBricks standard assembly [71] from five BioBrick standard biological parts: BBa_R0040 (tetR regulated promoter), BBa_B0034 (synthetic RBS), BBa_C0062 (luxR gene), BBa_B0015 (composite transcriptional terminator), and BBa_R0062 (right lux promoter). The prefix, BBa, denotes a BioBrick part from the alpha release of BioBrick standard biological parts (http://parts.mit.edu/). The initial letter of the part number specifies the part type (e.g., F denotes a cell-cell signaling part). The remaining numbers identify the specific part. DNA encoding each part (10-50 pmoles) was digested using restriction enzymes according to the manufacturers directions (New England Biolabs, NEB). Parts and vectors were then purified by gel electrophoresis (1% TAE agarose gel, 9 V/cm) and extracted using a QiaQuick gel extraction kit (Qiagen). Ligation reactions using T4 DNA ligase were carried out in a 6:1 part:vector molar ratio according to the manufacturers directions (NEB). Ligation products were transformed by chemical transformation [28] into *E. coli* DH5α [85]. A further BioBricks assembly of BBa_F2620 and BBa_E0240

resulted in the composite system, BBa_T9002, which was used for all characterization experiments. BBa_T9002 was ligated into a standard BioBrick plasmid, pSB3K3 (http://parts.mit.edu/). All assemblies were sequenced at the MIT Biopolymers lab and the sequence data analyzed in VectorNTI v10.1 (Invitrogen). For characterization experiments, BBa_T9002 was transformed into *E. coli* MG16554 by chemical transformation [28], sequence verified as described above and stored in a 20% glycerol stock at -80°C.

I obtained all acyl-homoserine lactones (AHLs) from Sigma-Aldrich (See Table A.1). AHLs were dissolved in ethyl acetate acidified with glacial acetic acid (0.01% v/v) to a stock concentration of 1E-1 M. Prior to each experiment, I serially diluted these stock concentrations in deionized H_2O (pH 5) to obtain solutions ranging in concentration from 1E-2 M to 1E-8 M. Making fresh dilutions prior to each experiment ensured that there was no loss of activity of low concentration stocks as had been noted in early experiments with low concentration stocks that had been reused multiple times (data not shown).

Table A.1: AHL variants used in this study.

Name	Abbreviation	Species
Butanoyl-homoserine lactone	C_4HSL	*P. aeruginosa* [48]
3-oxohexanoyl-homoserine lactone	$3OC_6$HSL	*V. fischeri* [41]
Hexanoyl-homoserine lactone	C_6HSL	*C. violaceum* [84]
Heptanoyl-homoserine lactone	C_7HSL	*E. psidii* R. IBSBF 435T [96]
3-oxooctanoyl-homoserine lactone	$3OC_8$HSL	*A. tumefaciens* [140]
Octanoyl-homoserine lactone	C_8HSL	*B. cepacia, V. fischeri* [48]
Decanoyl-homoserine lactone	C_{10}HSL	*B. pseudomallei* [127]
Dodecanoyl-homoserine lactone	C_{12}HSL	Synthetic

A.2   Static performance characterization protocol

1. Three 5 ml cultures of M9 minimal medium [111] supplemented with 0.2% casamino acids and 1mM thiamine hydrochloride (supplemented M9 medium) and antibiotic (kanamycin, 20 μg/ml) were inoculated with single colonies of $\sim$ 2

mm diameter from a freshly streaked plate of MG1655 containing BBa_T9002. One 5 ml culture was inoculated with a single colony from a freshly streaked plate of MG1655 containing a BBa_T9002 mutant lacking a GFP expression device (T9002m) described in the stability section.

2. Cultures were grown in 17 mm test tubes for 15 hrs at 37°C with shaking at 70 rpm.

3. Cultures were diluted 1:1000 into 5.5 ml of pre-warmed fresh medium and grown to an OD600 of 0.15 under the same conditions as before (this growth took 4.5 hrs on average).

4. Twenty-four 200 μl aliquots of each of the cultures were transferred into a flat-bottomed 96 well plate (Cellstar Uclear bottom, cat. # T-3026-16, Greiner).

5. 2μl of the stock concentrations of the cognate AHL, 3-oxohexanoyl-homoserine lactone (3OC$_6$HSL), was added to each well to yield 8 different final concentrations (0, 1E-10, 1E-9, 1E-8, 1E-7, 1E-6, 1E-5 and 1E-4 M). Three replicate wells were measured for each concentration of 3OC$_6$HSL. Three wells were each filled with 200 μl of medium to measure the absorbance background. Three further wells were each filled with 200 μl of the BBa_T9002 mutant culture to measure fluorescent background.

6. The plate was incubated in a Wallac Victor3 multi-well fluorimeter (Perkin Elmer) at 37°C and assayed with an automatically repeating protocol of absorbance measurements (600 nm absorbance filter, 0.1 sec counting time through 5 mm of fluid), fluorescence measurements (488 nm excitation filter, 525 nm emission filter, 0.5 sec, CW lamp energy 12902 units), and shaking (1 mm, linear, normal speed, 5 sec). Time between repeated measurements was 2 min and 21 sec. Approximately 6 min elapsed between beginning addition of 3OC$_6$HSL to the wells and the first plate reader measurement. 3OC$_6$HSL was added in order of increasing concentration to minimize GFP synthesis during plate loading. Cells appear to grow exponentially for the duration of the plate reader

measurement protocol (see Figure A-1 for representative growth curves).

7. I repeated steps 1 through 6 on three separate days to obtain data for nine colonies from a single plate.

8. I processed the data to compute the PoPS output from BBa_F2620 as is described in the data analysis section (below). The data for each colony tested was averaged across the three replicate wells. The mean for all colonies was then averaged to obtain a population mean. The time and dose dependent input-output surface is shown in Figure A-2. Following an initial transient response, device output reached an approximate steady state.

9. The snapshot transfer function in Figure 2-2 is the 60 min time-slice from the surface shown in Figure A-2 (highlighted as a heavy black line). Error bars in the Static performance section of Figure 2-2 represent the 95% confidence interval in the population for the nine independent samples. The cyan shaded region represents the range of the nine independent samples.

10. To gain further information about the transition region of the transfer function, measurements were subsequently taken at two intermediate $3OC_6HSL$ concentrations (3.3E-09 M and 3.3E-08 M) using the same protocol described above. Measurements were simultaneously taken at a subset of the original concentrations to ensure the new data was consistent with the earlier data. The new data was processed simultaneously with the original data, with the exception that only six independent colonies were measured for the intermediate $3OC_6HSL$ concentrations.

11. To estimate parameters that characterize the measured transfer function, I used least squares estimation to fit a simple model (Equation A.1) to the data. A Hill equation derived from simple biochemical equations describes the data well ($R^2 = 0.99$) :

$$P_{out} = \frac{P_{max}[3OC_6HSL]^n}{K^n + [3OC_6HSL]^n} \tag{A.1}$$

where P_{out} is the PoPS per cell output of BBa_F2620, P_{max} is the maximum output level, K is the switch point, or input level needed to produce a half maximal output level, and n is the Hill coefficient describing the steepness of the transition from low output to high output. I estimated uncertainty in the fitted parameter values from the estimated coefficient covariance matrix using the nlparci function in MATLAB.
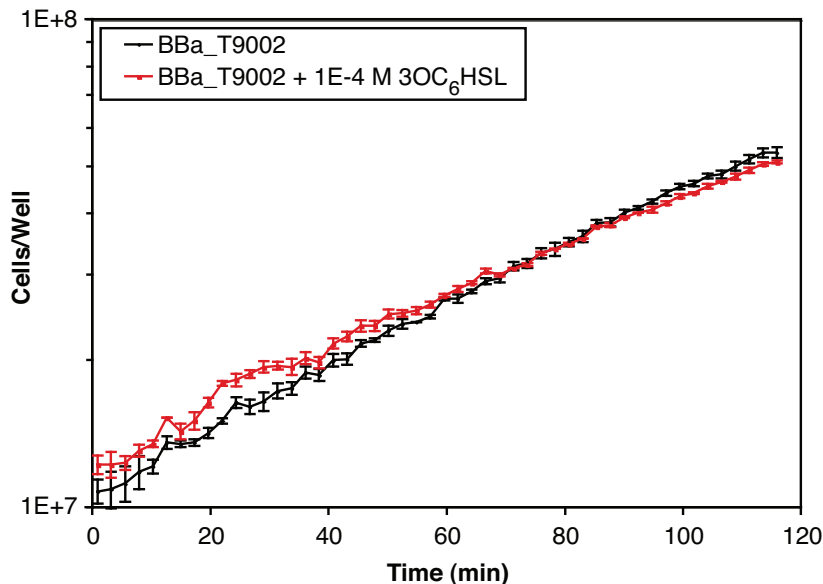
Figure A-1: Cell growth curves in the multi-well fluorimeter. Growth curves of cultures of MG1655 bearing BBa_T9002 from the transfer function experiment under zero input and an input of 1E-4 M 3OC$_6$HSL. The data points represent the mean of three identical samples; the error bars represent the 95% confidence interval in the estimate of the mean. The growth curves suggest that exponential growth continues for the course of the experiment.

A.3 Dynamic response characterization protocol

1. Four cultures, three of MG1655 bearing BBa_T9002 and one of MG1655 bearing T9002m were prepared as described in steps 1-3 of Section A.2 above.

2. Six 200 μl aliquots of each culture were transferred into a flat-bottom 96 well
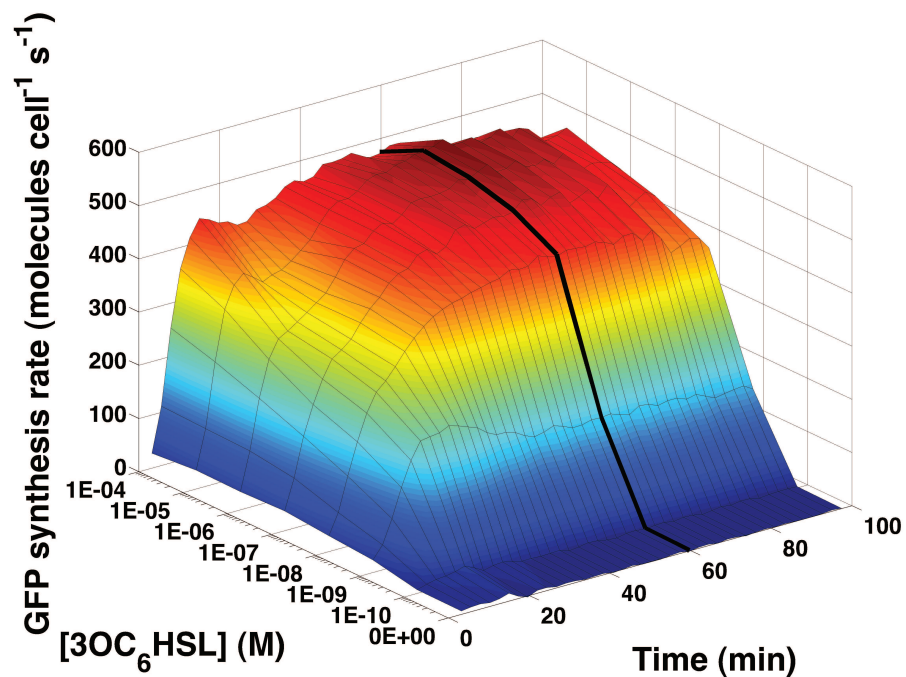
Figure A-2: Time and dose dependent input-output response of BBa_T9002 following addition of 3OC$_6$HSL. The surface shown is the weighted average of cultures inoculated from nine colonies. The heavy black line is the 60 min time slice that defines the transfer function shown in Figure 2-2.

plate.

3. Three wells were each filled with 200 μl of medium to measure the absorbance background.

4. The plate was incubated in a Wallac Victor3 multi-well fluorimeter at 37°C and assayed with an automatically repeating protocol of fluorescence measurements, absorbance measurements, and shaking (all as above). Time between repeated measurements was 69 sec.

5. After 10 min in the multi-well fluorimeter, $3OC_6HSL$ was added to three of the wells for each culture to a final concentration of 1E-7 M.

6. The plate was again incubated in the multi-well fluorimeter and assayed as before.

7. Before calculating the time dependent PoPS output from BBa_F2620 as described in the data analysis section (below) I first fit simple functions to the BBa_E0240 output data for those samples to which $3OC_6HSL$ was added. I used least-squares estimation to fit a straight line to the data prior to addition of $3OC_6HSL$ and a fourth order polynomial to the data after addition of $3OC_6HSL$. The fitting process was necessary to reduce the noise in the calculated PoPS values.

8. In the response time section of Figure 2-2 the data points represent the mean of the three cultures (each measured in triplicate) and the error bars represent the standard deviation of the three cultures. The least-squared fits are plotted as solid black lines.

A.4 Input compatibility

1. Two cultures, one of MG1655 bearing BBa_T9002 and one of MG1655 bearing BBa_T9002m were prepared as described in steps 1-3 of the transfer function

section above. However, in this case the overnight cultures were diluted into 20 ml of fresh medium in a 200 ml flask and shaken at 220 rpm during growth.

2. Three of the eight AHL variants (Table A.1) were preloaded into a flat-bottom 96 well plate to eight different final concentrations (0, 1E-10, 1E-9, 1E-8, 1E-7, 1E-6, 1E-5, and 1E-4 M). Three wells were each filled with 200 μl of media to measure the absorbance background. Three further wells were filled with 200 μl of the BBa_T9002m culture to measure the fluorescent background.

3. Seventy-two 200 μl aliquots of the BBa_T9002 culture were transferred to the plate. Three replicate wells were filled for each concentration of each AHL.

4. The plate was incubated in a Wallac Victor3 multi-well fluorimeter at 37°C and assayed with an automatically repeating protocol of absorbance measurements, fluorescence readings, and shaking (all as described for static performance). Time between repeated measurements was 2 min and 21 sec.

5. Steps 1 through 4 were repeated once with three more of the AHL variants and again with the final two AHL variants. The time between repeated measurements was kept fixed in each case.

6. Data processing is described below in the data analysis section. Under the Input Compatibility heading of Figure 2-2, snapshot transfer functions are plotted for each AHL variant at the 60 min time point similar to the transfer function experiment (above). The data points represent the mean of the three replicate wells of each measurement. The error bars represent the standard deviation of the three replicates.

A.5   Reliability

1. A 5 ml culture of supplemented M9 medium and antibiotic (kanamycin, 20 μg/ml) was inoculated with a single colony from a fresh plate.

2. The culture was grown in a 17 mm test tube for 15 hrs at 37°C with shaking at 70 rpm.

3. The culture was diluted 1:400 into 5 ml of fresh medium and grown under identical conditions for a further 10 hrs. The culture was diluted 1:4096 into two identical 5 ml cultures. $3OC_6HSL$ was added to one of the cultures to an input level of 1E-7 M.

4. These two cultures were propagated in a similar manner with 1:400 and 1:4096 dilutions in the morning and evening respectively every day for 5 days.

5. Each day, following the overnight incubation of cultures, a sample of the overnight cultures was stored in a 20% glycerol solution at -80°C to allow later sequencing. Samples to be sequenced were streaked on LB plates containing kanamycin (20 μg/ml). Plasmid DNA from five colonies on each plate was purified using a Qiagen Spin Miniprep Kit (Qiagen). DNA was sequenced and analyzed as described earlier using standard primers internal and external to the receiver and reporter device (VF2, VR, C0062VF from http://parts.mit.edu).

6. Each morning, a second culture was inoculated by a 1:400 dilution from the overnight culture for both the high and low input condition. These second cultures were grown in the absence of $3OC_6HSL$ for 8 hrs. This growth period diluted-out the accumulated GFP from the culture propagated in the presence of $3OC_6HSL$ before assaying performance.

7. Samples from both of the second copies were induced with a high input level of $3OC_6HSL$ (1E-7 M) at 37°C with shaking at 70 rpm for 45 min. Single-cell fluorescence measurements were carried out on a FACScan flow cytometer with a 488 nm Argon excitation laser and 525 nm emission filter (Becton-Dickinson). During each flow-cytometer measurement, data was collected from 50,000 cells. 2 μl of Sphero fluorescent beads (0.87 μm diameter), Spherotech) in 500 μl H2O was used as a control for experiment-to-experiment variation of cytometer

performance. FACScan data were analyzed using Cell Quest (Becton-Dickinson) and FlowJo (FlowJo).

8. The genetic reliability of a device is defined as the number of replication events until a mutant becomes fixed in the population, displacing the original device; the performance reliability of a device is defined as the number of replication events until the majority of the devices in the population have lost the ability to respond correctly to an input (Jason Kelly and Drew Endy, personal communication). Both the genetic and the performance stability were estimated from the FACS data shown in Figure 2-2 and genetic reliability was confirmed by sequence analysis.

A.6 Source of the reliability mutant

Repeated experiments following the stability protocol (Section A.5) consistently resulted in the accumulation of the same mutant, named T9002m, after a similar number of doublings. Sequence analysis of individual clones from the culture at the end of the stability experiment showed that T9002m had a deletion in BBa_T9002 between two 183 bp homologous regions (Figure A-3). The repeatability of the results suggested that there was pre-existing genetic variation within each clone at the start of each stability experiment. I measured the growth rate of cells carrying BBa_T9002 and T9002m when grown under the same conditions as in the stability experiment (Figure A-4). Exponential doubling times of 59 and 54 min were calculated for BBa_T9002 and T9002m respectively under high input conditions. A simple model (not shown) suggests that this doubling time difference is consistent with the observed population takeover by a pre-existing mutant within 74 doublings.

To test the hypothesis that pre-existing genetic variation is the source of the mutant device, plasmid DNA was purified from four colonies grown on a plate streaked from the master glycerol stock of MG1655 bearing BBa_T9002 using a Qiaprep spin miniprep kit (Qiagen). The purified DNA was used as template for PCR reactions. Several PCR products were detected by gel electrophoresis, with identical results for
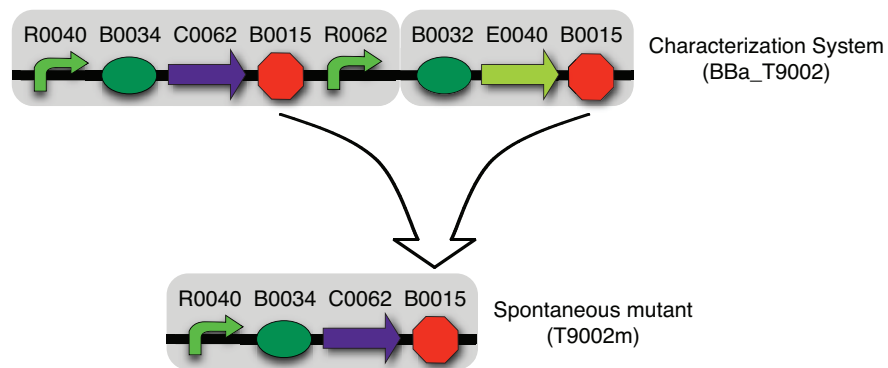
Figure A-3: Comparison of full length and mutant device. Part schematic of BBa_T9002 (top) and the mutant (bottom) that accumulated in the culture when propagated for > 74 generations in high $3OC_6HSL$ input conditions. BBa_T9002 contains two 183 bp direct repeats comprised of BBa_B0015 plus bracketing sequences.
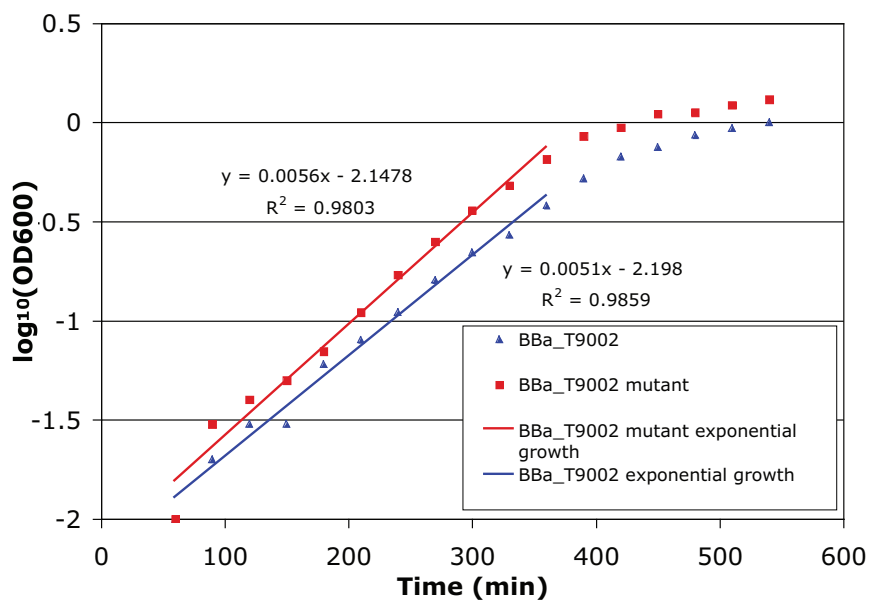
Figure A-4: Comparison of growth rates for full length and mutant devices in MG1655. Growth curves for MG655 bearing BBa_T9002 and T9002m. Growth conditions are described in the supplementary text. Based on the least squares fits to the exponential portion of the curves, the wild-type and mutant cells have doubling times of 59 and 54 min, respectively.

each template DNA sample. A minor band running at the speed expected for the mutant device was used as template for a further PCR reaction. The PCR product was sequenced and found to match the sequence of the mutant device.

To ensure the previous result was not due to an artifact of the PCR process, a further test for the presence of the mutant plasmid in the glycerol stock was performed. To enrich for the mutant plasmid, purified DNA from the glycerol stock was digested with restriction enzymes, BsrG1 and HincII (NEB), which recognize nucleotide sequences in BBa_T9002 but not T9002m. The digested DNA was purified using a QiaQuick PCR purification kit (Qiagen) and used to transform chemically competent *E. coli* TOP10 cells (Invitrogen). The transformed cells were plated on LB plates containing kanamycin (20 μg/ml) and 1E-7 M 3OC$_6$HSL. A fraction of the colonies were green, as expected for cells transformed with the full-length device, and the remainder of the colonies were not visibly green. Colony PCR reactions were performed on colonies that were not visibly green. The PCR products were analyzed by gel electrophoresis. The PCR product of all non-green colonies ran at the speed expected for the mutant device, while the PCR products of all green colonies ran at the speed expected for BBa_T9002. This result, together with the result of the PCR approach described above, suggests that either the mutant device was present on a minority plasmid population within individual cells in the glycerol stock, or that the mutant device occurs consistently with high frequency during growth on plates or liquid culture prior to DNA purification.

I then investigated whether occurrence of the mutant plasmid when transforming cells with BBa_T9002 was repeatable. I purified DNA from visibly green colonies used in the experiment described above. The parent cell of each colony should be transformed with a single plasmid based on the transformation protocol used and hence the colony should contain a homogenous plasmid population. The purified DNA was analyzed similarly to that described above and the results again showed that the transformed cells carried a small fraction of mutant plasmids. This data suggests that the mutant occurs at a high frequency in both recA+ (MG1655) and recA mutant cells (TOP10), consistent with a replication slippage mechanism for

repeat deletion [80].

A.7 Transcriptional demand

The output demand produced by BBa_T9002 was determined by estimating the number of RNA polymerases (RNAP) and nucleotides needed to produce the output transcript. For example, the estimated high output PoPS was 6.7 PoPS cell^{-1}. Assuming that in steady state, 6.7 mRNA are produced per sec, 6.7 x Nt nucleotides are incorporated into mRNA per sec, where Nt is the length of the mRNA produced by the output of BBa_F2620. Assuming an RNAP speed of 45 nucleotides per sec [130], approximately 0.15 x Nt RNAP must be actively engaged in transcribing the DNA downstream of BBa_F2620.

A.8 GFP maturation rate

1. I followed a similar method to that used to measure the maturation rate of GFP in yeast [55].

2. Four cultures, three of MG1655 bearing BBa_T9002 and one of MG1655 bearing T9002m were prepared as described in steps 1-3 of the transfer function section above.

3. Twelve 200 μl aliquots of each culture were transferred into a flat-bottom 96 well plate.

4. Three wells were each filled with 200 μl of medium to measure the absorbance background. $3OC_6HSL$ was added to six of the wells for each culture to a final concentration of 1E-6 M.

5. The plate was incubated in a Wallac Victor3 multi-well fluorimeter at 37°C and assayed with an automatically repeating protocol of fluorescence measurements, absorbance measurements, and shaking (all as above). Time between repeated measurements was 54 sec.

6. After 20 min, spectinomycin (Sigma-Aldrich) was added to a final concentration of 500μM to three wells containing $3OC_6HSL$ and to three wells not containing $3OC_6HSL$ for each culture.

7. The plate was again incubated in the multi-well fluorimeter and assayed as before.

8. Relative GFP accumulation was plotted as a function of time (Figure A-5).

9. I fit an exponential model [55] via least squares estimation to the GFP accumulation data. Least squares estimation was performed by the nlinfit algorithm in MATLAB.

10. I measured an average time constant for GFP maturation of 6.3 min with a 95% confidence interval over three replicates of ±0.4 min. The corresponding average GFP maturation rate is 1.8E-3 sec^{-1}.

A.9   Data processing

Data processing consisted of two stages. First, I calculated the output of BBa_E0240 using the relative measurements of the plate reader. Second, I used this data and our knowledge of BBa_E0240 to calculate the output of BBa_F2620 using an approach similar to previous models relating GFP expression to promoter output [75].

Raw data from the Wallac Victor3 multi-well fluorimeter was processed by first subtracting the appropriate backgrounds. The absorbance of wells containing supplemented M9 medium, A_{media}, was subtracted from the sample absorbance data, A_{raw}. The resulting data, $A_{corrected}$, was assumed to be directly proportional to the number of cells in the well.

$$A_{corrected} = A_{raw} - A_{media} \tag{A.2}$$

Similarly, the fluorescence data for the GFP-free BBa_T9002 mutant, G_{cells}, was subtracted from the sample fluorescence data, G_{raw}, and the resulting data $G_{corrected}$
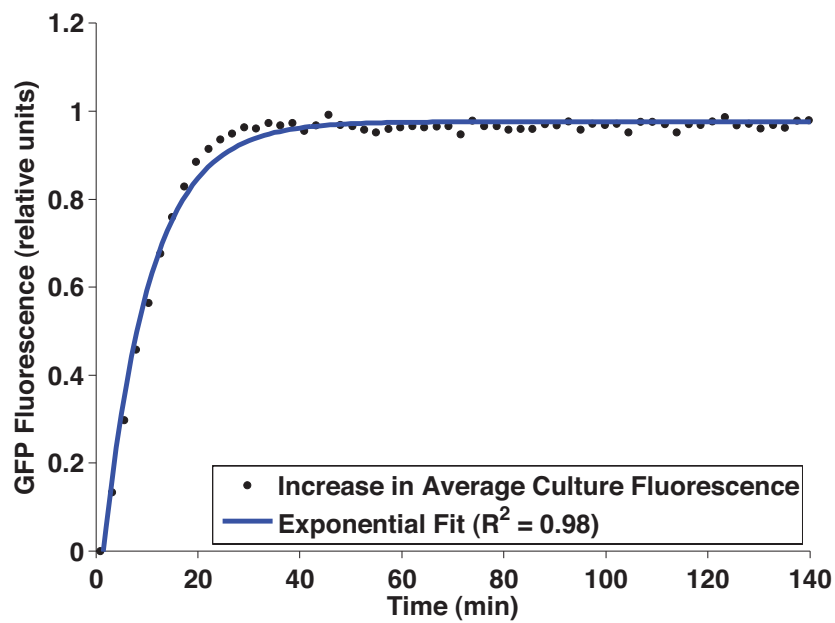
Figure A-5: Accumulation of GFP due to maturation. A time course of GFP fluorescence following addition of spectinomycin (500 μM) for a representative culture. The data has been normalized by subtracting the fluorescence level at the zero time point and via scaling by the maximum fluorescence level. An exponential function describing GFP maturation [55] has been fit to the data.

was assumed proportional to the total number of GFP molecules in the well.

$$G_{corrected} = G_{raw} - G_{cells} \tag{A.3}$$

The data was then converted to absolute units (CFU/well and GFP molecules/well) using the calibration method described below. The conversion equations used are shown in Equations A.4 and A.5.

$$OD = 3.11 A_{corrected} - 0.016 \tag{A.4}$$

$$GFP = 7.0E8 G_{corrected} + 6.0E11 \tag{A.5}$$

Mean synthesis rates of GFP per cell, S_{cell}, were calculated by assuming the total GFP synthesis rate, S_{total}, to be equal to the time differential of GFP. S_{cell} was then calculated via Equation A.7. Note that since I have measured the total amount of GFP in the well and since I have assumed that GFP is not degraded, I can calculate the total synthesis rate of GFP and hence the per cell synthesis rate of GFP without considering dilution due to cell growth.

$$S_{total} = \frac{dGFP}{dt} \tag{A.6}$$

$$S_{cell} = \frac{S_{total}}{OD} \tag{A.7}$$

To interpret the behavior of BBa_F2620 from our observations of BBa_E0240, I employed an ODE model relating the output of BBa_E0240 to its input (the output of BBa_F2620). I defined the input to BBa_E0240 to be the time dependent rate of mRNA synthesis, $PoPS(t)$ (mRNA per cell per sec). I defined the output of BBa_E0240 to be the synthesis rate of mature GFP, S_{cell} (GFP molecules per cell per sec). The model includes two species and four parameters. The differential equations governing the levels of the two species are:

$$[\dot{M}] = PoPS(t) - \gamma_M[M] \tag{A.8}$$

$$[\dot{I}] = \rho[M] - (a + \gamma_I[I] \tag{A.9}$$

$$S_{cell} = a[I] \tag{A.10}$$

where $[M]$ is the concentration of mRNA per cell, $[I]$ is the concentration of immature GFP, $PoPS(t)$ is the time dependent rate of mRNA synthesis, γ_M is the degradation rate of mRNA, ρ is the constant rate of protein synthesis per mRNA, a is the maturation rate of GFP, and γ_I is the degradation rate of immature GFP (incorporating degradation and dilution due to cell growth).

I parameterized the model using published and unpublished data and via experiments using BBa_T9002. I assumed a value of 4.8E-3 sec^{-1} for γ_M based on unpublished measurements performed in our lab on an almost identical mRNA, produced by BBa_I7107 (the transcript in the current study has two extra A nucleotides on the 5' end but an unchanged secondary structure). This value is also consistent with published data on mRNA decay in *E. coli* [14]. I assumed that dilution of mRNA due to cell growth was negligible relative to degradation.

I estimated a value for ρ of 0.4 proteins per sec per mRNA based on unpublished measurements from the Endy Lab on BBa_I7107. Again, this value is consistent with published data for translation rates per mRNA [34, 69].

I measured an average GFP maturation rate of 1.8E-3 sec^{-1} as described above. Finally, I assumed that immature GFP is stable so that degradation was negligible relative to dilution due to growth. An average dilution rate of 2E-4 sec^{-1} was calculated from the multi-well fluorimeter absorbance data (corresponding to a 55 min doubling time).

At the 60 min timepoint used in the snapshot transfer functions, the above model for BBa_E0240 behavior is in steady state. Hence, I used the steady state relationship of Equation A.11 to calculate the specific output of BBa_F2620 from the observed output of BBa_E0240.

$$PoPS^{ss} = \frac{\gamma_M(a + \gamma_I)S_{cell}^{ss}}{a\rho} \tag{A.11}$$

To calculate the transient output of BBa_F2620 in the response time experiments, I rearranged the model (Equations A.8, A.9, A.10 to relate the time dependent PoPS

output to measured values of S_{cell}.

$$PoPS(t) = \frac{\ddot{S}_{cell} + (\gamma_I + \gamma_M + a)\dot{S}_{cell} + \gamma_M(\gamma_I + a)S_{cell}}{a\rho} \qquad \text{(A.12)}$$

A.10 Data calibration

Data measured using the Wallac Victor3 multi-well fluorimeter was converted to absolute units to better allow repetition of the experiments and comparison of data from other researchers. Absorbance measurements were related to colony forming units (CFU) by growing a culture of MG1655 bearing BBa_T9002 to different cell densities and then plating an appropriate dilution on plates of Luria-Bertani medium [111] containing 20μg/ml kanamycin. For each culture density, three identical plates were made and the resulting colonies were counted. A straight line ($R^2 = 0.9912$) was used to fit the data relating the number of colony forming units (CFU) per ml of culture to the absorbance measurements. The equation of the straight line was used to convert all absorbance measurements to CFU/well.

Fluorescence measurements were converted to an absolute number of GFP molecules by relating the fluorescence of cultures of MG1655 bearing BBa_T9002 to the fluorescence of purified GFP using a two-step calibration. In the first step, purified GFP mut3b [31] (a generous gift from Chris Farrell) was serially diluted in cell lysate of wild-type MG1655. Lysis was performed using B-PER II, a non-denaturing bacterial lysis solution (Pierce) using a modified form of the manufacturers protocol to produce $> 90\%$ lysis as verified by a plating assay. Briefly, cells grown in supplemented M9 medium to an OD600 of 0.86 were centrifuged at 13000 rpm in a tabletop centrifuge for 4 min at 4°C and resuspended in 10mM Tris buffer containing 100mM NaCL adjusted to pH 8. Samples were centrifuged in the same manner as before, the supernatant removed and resuspended in 50 μl B-PER II. 1 unit of Benzonase was added to each sample and all samples were incubated at room temperature for 30 min. The fluorescence of the GFP dilutions was measured in the Wallac Victor3 multi-well fluorimeter and the data plotted against the mass of GFP in each dilution. A straight line

($R^2 = 0.99$) was fit to the data. In the second step, cultures of MG1655 containing BBa_T9002 were induced using 1E-7 M $3OC_6HSL$ for varying lengths of time leading to varying intracellular concentrations of GFP. Sample fluorescence was measured in the fluorimeter and the remainder of the cultures were lysed in an identical manner to that used for the purified GFP dilutions. The fluorescence of the lysed samples was measured and plotted against the fluorescence prior to lysis. A straight line (R^2 = 0.9711) was used to fit the data. The fit equations for the purified GFP dilutions and the lysed samples were combined to derive a straight-line relationship between the fluorescence of unlysed cultures and the number of GFP molecules/well. This relationship was used to convert all fluorescence measurements to GFP molecules per well.

Appendix B

Glossary of an exemplar datasheet

The following is a glossary of terminology and concepts employed on the datasheet of BBa_F2620 (Figure 2-2).

- *BBa_F2620*: The unique part number assigned to the device. The prefix, BBa, denotes a BioBrick part from the alpha release of BioBrick standard biological parts collection (`http://parts.mit.edu`). F denotes a cell-cell signaling device and the remaining numbers identify the specific device.

- *Static Performance*: This section contains data describing the steady-state relationship between the input and output of the device. The transfer function shows the input/output relationship 60 min after addition of input signal at which time the reporter device (BBa_E0240) is assumed to be at steady state. Hence, there is a linear relationship between the measured GFP synthesis rate and the PoPS output of the receiver. The inset shows the time and dose dependent response of the receiver, the 60 min time point is indicated by a solid black line.

- *Population Mean*: The mean output level for either six or nine independent cultures at a given input level. Error bars represent the 95% confidence interval of the mean of the independent cultures.

- *Colony Range*: A range bounded by the lowest and highest outputs among the

independent cultures at a given input level.

- *Hill Equation*: An equation relating the PoPS output per cell of the receiver (P_{out}) to the input concentration of $3OC_6HSL$. P_{max} represents the maximum output of the receiver, K is the device switch point and n is the Hill coefficient.

- *Dynamic Response*: This section describes the response of the receiver to a step increase in input level at 0 min. The mean GFP synthesis rates measured for three cultures of the composite part (BBa_T9002) are shown as filled (high input) or empty (low input) circles. Error bars represent standard deviations across the independent cultures. The solid black lines are a linear fit to the data (Appendix A). The time dependent PoPS output from the receiver (shown as a solid red line) was calculated using a model of the dynamic behavior of the reporter device (Appendix A).

- *Response time*: The time for the output of the receiver to reach 67% of its final value was estimated from the calculated PoPS output of the receiver. The response time of the composite part (BBa_T9002) was calculated by fitting an exponential function (not shown) to the GFP synthesis rate data after the addition of 1E-7 M $3OC_6HSL$.

- *Input compatibility*: The dose response of the receiver to a variety of signaling compounds similar to $3OC_6HSL$ is presented. The data points represent the mean of three independent cultures and the error bars represent the standard deviation of the data for the three independent cultures.

- *Part compatibility*: A list of other biological objects with which the receiver is known to be qualitatively functional.

- *Chassis*: An organism, or genetic background, that can be used to support and power a particular engineered biological device. Details of specific genetic backgrounds can be found online (`http://parts.mit.edu`).

- *Reliability*: The ability of the device to continue to function over many generations is reported. Here, FACS data shows the response of the device to a high input signal as a function of culture doublings. Two cases are shown, one in which the culture is propagated under low input conditions and one in which the culture is propagated under high input conditions.

- *Genetic Reliability*: The number of culture doublings before a mutant device comprises at least 50% of the population. The reported figures are derived from the FACS data and confirmed by DNA sequencing analysis.

- *Performance Reliability*: The number of culture doublings before 50% of the population is unable to correctly respond to an input. The reported figures are derived from the FACS data.

- *Transcriptional Output Demand*: The receiver requires resources from the cellular chassis in order to function. The demand for resources related to transcription are presented as a function of the length of the transcript produced by the output of the receiver.

- *Conditions*: The growth conditions and measurement methods used to characterize the receiver are summarized on the datasheet. Full details can be found in Appendix A.

- *License*: The ownership, sharing, and innovation terms by which the authors provide access to, and use of, the receiver together with the associated characterization data.

Appendix C

Comparison of previous genetic constructs similar to BBa_F2620

Study	Source of Key Elements[a]	E. coli Chassis[b]	Plasmid (ori, resistance)	Conditions (medium, temperature)	Switch Point[c] (nM)	Response Time[c] (min)	Input Compatibility[c,d] (summary)	Reliability[c]	Demand[c]
BBa_F2620	*V. fischeri* MJ1	MG1655	pSB3k3 (p15A, Kn^r)	Supplemented M9 medium, liquid culture, 37°C	1.5±0.3	< 1	Figure 1 (broad specificity)	> 92/ < 74	6.7xNt nuc cell^{-1} s^{-1} 0.15xNt RNAP cell^{-1}
[135]	*V. fischeri* MJ1	JM109	PSB401 (p15A, Tet^r)	LB, liquid culture, 30°C	~12	N/A	Figure 2 (A, B) [135] (broad specificity)	N/A	N/A
[79]	*V. fischeri* MJ1	WM54	pAL103 (p15A, Tet^r)	LB, liquid culture, 37°C	~ 5	N/A	N/A	N/A	N/A
[2]	*V. fischeri* MJ1	MC4100	pJBA132 (p15A, Kn^r)	LB, liquid culture, 30°C	4-8	Figure 4 [2] (broad specificity)	Figure 3 (C) [2]	N/A	N/A
[133]	*V. fischeri* MJ1	DH5α	pRCV-3 (pMB1, Ap^r)	LB, liquid culture, 30°C	N/A	Figure 6 [133]	N/A	N/A	N/A
[29]	*V. fischeri* MJ1	DH5α	pLuxR (p15A, Kn^r)	LB, liquid culture, 37°C	10	N/A	Figure 2 (A) [29] (unresponsive to C$_8$HSL)	N/A	N/A
[30]			pLuxGFPuv (ColE1, Cmr)				Figure 2 [30] (unresponsive to C$_8$HSL)		
[114]	*V. fischeri* ES114	VJS533	pHV200I$^-$ (ColE1, Ap^r)	LB, liquid culture, 30°C	> 100	N/A	Table 1 and text [114] (narrow specificity)	N/A	N/A
[16]	*V. fischeri* ES114	VJS533	pHV200I$^-$	Bioassay medium, liquid culture, 23-25°C	> 5000	Figure 1[16]	N/A	N/A	N/A

Genetic constructs based on the same DNA source can vary in genetic organization and regulatory elements. [a]Sequence differences exist between the two listed strains. For example, the MJ1 and ES114 LuxR proteins share 75% identity and 89% similarity at the amino acid level. [b]See Appendix F for genotype information. [c]See Chapter 2 for a description of each characteristic. [d]Input compatibility is defined using a response switch point cutoff two orders of magnitude from the cognate AHL switch point.

Table C.1: Comparison of previous genetic constructs similar to BBa_F2620.

Appendix D

Methods for VM1.0

D.1 BioBrick assemblies

Except where indicated, all genetic material was assembled via BioBricks standard assembly [71] or the three-antibiotic method [119]. Briefly, DNA encoding each part (10-50 pmoles) was digested using restriction enzymes according to the manufacturers directions (New England Biolabs, NEB). When necessary, parts and vectors were then purified by gel electrophoresis (1% TAE agarose gel, 9 V/cm) and extracted using a QiaQuick gel extraction kit (Qiagen). Ligation reactions using T4 DNA ligase were carried out according to the manufacturers directions (NEB). Ligation products were transformed by chemical transformation[28, 93]. All assemblies were sequenced at the MIT Biopolymers lab and the sequence data analyzed in VectorNTI v10.1 (Invitrogen).

D.2 Construction of BioBrick T7 promoters

I obtained a plasmid carrying the T7 consensus promoter sequence in BioBrick format from Sriram Kosuri that had not been tested. I constructed the BioBrick mutant T7 promoters by ordering complementary oligos that encoded the promoters and the appropriate BioBrick prefix and suffix. Complementary oligos were annealed. The annealed oligos were then cloned via BioBrick standard assembly.

D.3 Measurement protocol for T7 reporter experiment

1. 5 ml cultures of Neidhardt rich media [89](purchased from Teknova) supplemented with 0.4% glycerol and antibiotic (Ampicillin, 50 μg/ml) were inoculated with single colonies (2mm ϕ) from freshly streaked plates of BL21(DE3) containing the reporter devices.

2. Cultures were grown in 17 mm test tubes for 12 hrs at 37°C with shaking at 70 rpm.

3. Cultures were diluted 1:1000 into 5 ml of fresh medium and grown for 2 hours.

4. Two 200 μl aliquots of each of the cultures were transferred into a flat-bottomed 96 well plate (Cellstar Uclear bottom, cat. # T-3026-16, Greiner).

5. IPTG was added to a final concentration of 0.4 mM to half of the wells.

6. The plate was incubated in a Wallac Victor3 multi-well fluorimeter (Perkin Elmer) at 37°C and assayed with an automatically repeating protocol of absorbance measurements (600 nm absorbance filter, 0.1 sec counting time through 5 mm of fluid), fluorescence measurements (488 nm excitation filter, 525 nm emission filter, 0.5 sec, CW lamp energy 12902 units), and shaking (1 mm, linear, normal speed, 5 sec). Time between repeated measurements was 10 min.

7. Absorbance measurements were converted to OD600 by subtracting the absorbance of a well containing media only and then using an empirical calibration relating absorbance measured by the plate reader to OD600. Sterile water was automatically added to the wells at each repeat to counteract evaporation.

D.4 pCH1497-ASD1

An annotated Genbank file containing the nucleotide sequence of pCH1497-ASD1 is available at the following permanent url `http://dspace.mit.edu/handle/1721.1/`

41842.
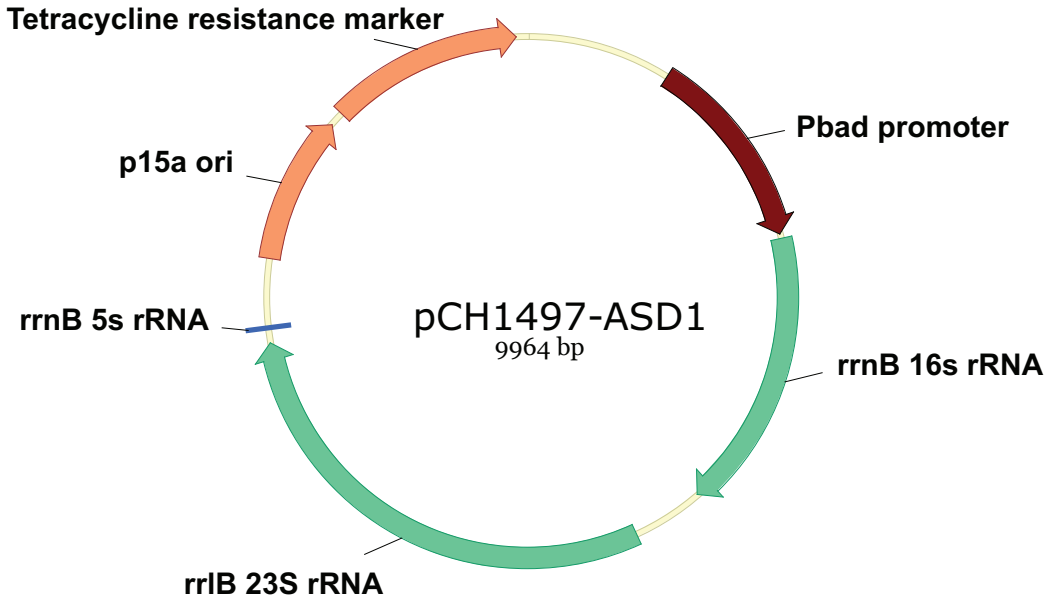
Figure D-1: Plasmid map of pCH1497-ASD1. pCH1497-ASD1 was obtained from Christopher Hayes (UCSB). 16s rRNA sequence includes a mutated aSD sequence matching that previously described[20].

D.5   Construction of BioBrick O-RBS

All orthogonal RBS were constructed via oligo annealing and ligation-based cloning (See Section D.2).

D.6   Measurement protocol for Brink O-ribosome experiment

1. 5 ml cultures of Neidhardt rich media [89](purchased from Teknova) supplemented with 0.4% glycerol and antibiotic (Ampicillin, 50 μg/ml; Tetracycline, 5 μg/ml) were inoculated with single colonies (2mm ϕ) from freshly streaked plates of BL21(DE3) containing the reporter devices. Three cultures were started from individual colonies for the strain carrying the reporter and the O-ribosome generator. One culture was started for each of the other strains.

2. Cultures were grown in 17 mm test tubes for 12.5 hrs at 37°C with shaking at 70 rpm.

3. Cultures were diluted between 1/500 and 1/250 (to yield equal final cell densities) into 5 ml of fresh medium.

4. Six 200 μl aliquots of each of the cultures were transferred into a flat-bottomed 96 well plate (Cellstar Uclear bottom, cat. # T-3026-16, Greiner).

5. Arabinose was added to a final concentration of 0.2% to half of the wells.

6. The plate was incubated in a Wallac Victor3 multi-well fluorimeter (Perkin Elmer) at 37°C and assayed with an automatically repeating protocol of absorbance measurements (600 nm absorbance filter, 0.1 sec counting time through 5 mm of fluid), fluorescence measurements (488 nm excitation filter, 525 nm emission filter, 0.5 sec, CW lamp energy 12902 units), and shaking (1 mm, linear, normal speed, 5 sec). Time between repeated measurements was approximately 12 min. Sterile water was automatically added to the wells at each repeat to counteract evaporation.

7. Absorbance measurements were converted to OD600 by subtracting the absorbance of a well containing media only and then using an empirical calibration relating absorbance measured by the plate reader to OD600.

D.7 Construction of pCH1497-rRNA4

pCH1497-rRNA4 was constructed via 'Round-the-horn PCR [92] using pCH1497-ASD1 as template. The forward primer was 4rRNA-f (`5'-TTG TGG TAc ctt aaa gaa gcg tac ttt gta gtg ctc aca cag-3'`). The reverse primer was 4/10rRNA-r (`5'-TGA TCC AAC CGC AGG TTC CCC TAC-3'`). The ribosomal operon and the P_{BAD} promoter were sequenced at the MIT Biopolymers lab and the sequence data analyzed in VectorNTI v10.1 (Invitrogen).

D.8 Measurement protocol for Rackham O-ribosome experiment

1. 5 ml cultures of Neidhardt rich media [89](purchased from Teknova) supplemented with 0.4% glycerol and antibiotic (Ampicillin, 50 μg/ml; Tetracycline, 5 μg/ml) were inoculated with single colonies (2mm ϕ) from freshly streaked plates of BL21(DE3) containing the described devices.

2. Cultures were grown in 17 mm test tubes for 12.5 hrs at 37°C with shaking at 70 rpm.

3. Cultures were diluted between 1/500 and 1/250 (to yield equal final cell densities) into 5 ml of fresh medium and grown for two hours.

4. Six 200 μl aliquots of each of the cultures were transferred into a flat-bottomed 96 well plate (Cellstar Uclear bottom, cat. # T-3026-16, Greiner).

5. Arabinose was added to a final concentration of 0.2% to half of the wells.

6. The plate was incubated in a Wallac Victor3 multi-well fluorimeter (Perkin Elmer) at 37°C and assayed with an automatically repeating protocol of absorbance measurements (600 nm absorbance filter, 0.1 sec counting time through 5 mm of fluid), fluorescence measurements (488 nm excitation filter, 525 nm emission filter, 0.5 sec, CW lamp energy 12902 units), and shaking (1 mm, linear, normal speed, 5 sec). Time between repeated measurements was approximately 12.5 min. Sterile water was automatically added to the wells at each repeat to counteract evaporation.

7. Absorbance measurements were converted to OD600 by subtracting the absorbance of a well containing media only and then using an empirical calibration relating absorbance measured by the plate reader to OD600. The fluorescence data was normalized at each timepoint by dividing by the absorbance of the sample.

D.9 Measurement protocol for the VM1.1 experiment

1. 1 ml cultures of Neidhardt rich media [89](purchased from Teknova) supplemented with 0.4% glycerol and antibiotic (Ampicillin, 50 μg/ml; Tetracycline, 5 μg/ml) were inoculated with single colonies (2mm ϕ) from freshly streaked plates of BL21(DE3) containing described devices.

2. Cultures were grown in 12 mm test tubes for 5 hrs at 37°C with rotating at 220 rpm.

3. Cultures were diluted 1/500 into 3 ml of fresh medium and grown for two hours.

4. Twelve 200 μl aliquots of each of the cultures were transferred into a flat-bottomed 96 well plate (Cellstar Uclear bottom, cat. # T-3026-16, Greiner).

5. Arabinose was added to a final concentration of 0.2% to one quarter of the wells, IPTG was added to a final concentration of 0.4 mM to another quarter of the wells, and both inducers were added to a third quarter.

6. The plate was incubated in a Wallac Victor3 multi-well fluorimeter (Perkin Elmer) at 37°C and assayed with an automatically repeating protocol of absorbance measurements (600 nm absorbance filter, 0.1 sec counting time through 5 mm of fluid), fluorescence measurements (488 nm excitation filter, 525 nm emission filter, 0.5 sec, CW lamp energy 12902 units), and shaking (1 mm, linear, normal speed, 5 sec). Time between repeated measurements was approximately 15 min. Sterile water was automatically added to the wells at each repeat to counteract evaporation.

7. Absorbance measurements were converted to OD600 by subtracting the absorbance of a well containing media only and then using an empirical calibration relating absorbance measured by the plate reader to OD600. The fluorescence data was normalized at each timepoint by dividing by the absorbance of the sample.

D.10 Construction of a T7 RNAP BioBrick, BBa_I2032

T7 gene 1 was PCR amplified from purified T7 DNA (obtained from Sriram Kosuri) with the BioBrick prefix and suffix added. The forward primer was `5'-CGC TTC TAG ATG AAC ACG ATT AAC ATC GCT AAG AAC GAC TTC-3'` and the reverse primer was `5'-CCA GCT GCA GCG GCC GCT ACT AGT ATT ACG CGA ACG CGA AGT CCG-3'`.

D.11 Measurement protocol for the VM1.2 experiment

1. 5 ml cultures of Neidhardt rich media [89](purchased from Teknova) supplemented with 0.4% glycerol and antibiotics as appropriate (Ampicillin, 50 μg/ml; Tetracycline, 5 μg/ml; Chloramphenicol, 34 μg/ml) were inoculated with single colonies (2mm ϕ) from freshly streaked plates of TOP10 containing described devices.

2. Cultures were grown in 17 mm test tubes for 15 hrs at 37°C with rotating at 70 rpm.

3. Cultures were diluted 1/500 (adjusted to reach the same final density for each culture) into 5 ml of fresh medium and grown for three hours.

4. Six 200 μl aliquots of each of the cultures were transferred into a flat-bottomed 96 well plate (Cellstar Uclear bottom, cat. # T-3026-16, Greiner).

5. Arabinose was added to a final concentration of 0.2% to half the wells.

6. The plate was incubated in a Wallac Victor3 multi-well fluorimeter (Perkin Elmer) at 37°C and assayed with an automatically repeating protocol of absorbance measurements (600 nm absorbance filter, 0.1 sec counting time through 5 mm of fluid), fluorescence measurements (488 nm excitation filter, 525 nm emission filter, 0.5 sec, CW lamp energy 12902 units), and shaking (1 mm,

145

linear, normal speed, 5 sec). Time between repeated measurements was approximately 15 min. Sterile water was automatically added to the wells at each repeat to counteract evaporation.

7. Absorbance measurements were converted to OD600 by subtracting the absorbance of a well containing media only and then using an empirical calibration relating absorbance measured by the plate reader to OD600. The fluorescence data was normalized at each timepoint by dividing by the absorbance of the sample.

Appendix E

Methods for VM2.0

E.1 BioBrick assemblies

Except where indicated, all genetic material was assembled via BioBricks standard assembly [71] or the three-antibiotic method [119]. Briefly, DNA encoding each part (10-50 pmoles) was digested using restriction enzymes according to the manufacturers directions (New England Biolabs, NEB). When necessary, parts and vectors were then purified by gel electrophoresis (1% TAE agarose gel, 9 V/cm) and extracted using a QiaQuick gel extraction kit (Qiagen). Ligation reactions using T4 DNA ligase were carried out according to the manufacturers directions (NEB). Ligation products were transformed by chemical transformation[28, 93]. All assemblies were sequenced at the MIT Biopolymers lab and the sequence data analyzed in VectorNTI v10.1 (Invitrogen).

E.2 Construction of BBa_R0184

BBa_R0184 was constructed via PCR amplification using purified pET-11d vector (Novagen) as template. Phusion DNA polymerase (Finnzymes) was used for the PCR. The forward primer, R0184F (5'-GTT TCT TCG AAT TCG CGG CCG CTT CTA GAG tca tac gac tca cta tag gg-3') included single base pair mismatch relative to the template sequence that encoded the -16 A to C mutation. The reverse primer,

R0184R was `5'-GTT TCT TCC TGC AGC GGC CGC TAC TAG TAG GAA TTG TTA TCC GCT CA-3'`. The forward and reverse primers added the BioBrick prefix and suffix onto the T7 promoter. The amplified product was digested and cloned via BioBricks standard assembly.

E.3 Construction of BBa_C0013

BBa_C0013 was constructed via PCR amplification using purified BBa_C0012 cloned in pSB1A2 as template. Phusion DNA polymerase (Finnzymes) was used for the PCR. The forward primer was the standard BioBrick sequencing primer, VF2 (tgccac-ctgacgtctaagaa, http://partsregistry.org). The reverse primer (C0013R, `5'-GTT TCT TCC TGC AGC GGC CGC TAC TAG TAT TAT TAC TGC CCG CTT TCC AGT C-3'`), included a region that was homologous to a sequence in BBa_C0012 immediately prior to the 33 bp ssrA sequence. The amplified product was digested and cloned via BioBricks standard assembly.

E.4 Measurement protocol for the LacI repressible T7 promoter experiment

1. 5 ml cultures of Neidhardt rich media [89](purchased from Teknova) supplemented with 0.4% glycerol and antibiotic as appropriate (Ampicillin, 50 μg/ml; Tetracycline, 5 μg/ml; Kanamycin, 10 μg/ml) were inoculated with single colonies (2mm ϕ) from freshly streaked plates of BL21(AI) (Invitrogen) containing the described devices.

2. Cultures were grown in 17 mm test tubes for 12 hrs at 37°C with rotation at 70 rpm.

3. Cultures were diluted 1:1000 into 5 ml of fresh medium and grown for four hours.

4. Appropriate inducers were added to the cultures. Arabinose was added to a final concentration of 0.2%, IPTG was added to a final concentration of 0.4 mM and grown for 12 hours.

5. 200 μl aliquots of each of the cultures were transferred into a flat-bottomed 96 well plate (Cellstar Uclear bottom, cat. # T-3026-16, Greiner).

6. The plate was assayed in a Wallac Victor3 multi-well fluorimeter (Perkin Elmer) with an absorbance measurement (600 nm absorbance filter, 0.1 sec counting time through 5 mm of fluid) and a fluorescence measurement (488 nm excitation filter, 525 nm emission filter, 0.5 sec, CW lamp energy 12902 units).

7. Absorbance measurements were converted to OD600 by subtracting the absorbance of a well containing media only and then using an empirical calibration relating absorbance measured by the plate reader to OD600.

E.5   T7 autogenerator growth measurements

1. 5 ml cultures of LB supplemented with antibiotic as appropriate (Chloramphenicol, 34 μg/ml) were inoculated with single colonies (2mm ϕ) from freshly streaked plates of *E. coli* D1210 carrying the listed devices.

2. Cultures were grown in 17 mm test tubes for 12 hrs at 37°C with rotation at 70 rpm.

3. Cultures were diluted 1:500 into 50 ml of fresh medium in 250 ml baffled flasks and grown at 37°C with shaking at 220rpm.

4. 1 ml samples were taken from the cultures at the indicated times after dilution and the OD600 of the cultures were measured via a CO8000 cell density meter (WPA, Cambridge, UK).

E.6 Measurement protocol for a T7 autogenerator using orthogonal translation

1. 5 ml cultures of LB supplemented with antibiotics as appropriate (Ampicillin, 50 μg/ml; Tetracycline, 5 μg/ml; Chloramphenicol, 34 μg/ml) were inoculated with single colonies (2mm ϕ) from freshly streaked plates of TOP10 containing the described devices. Three cultures started from single colonies were started for each strain.

2. Cultures were grown in 17 mm test tubes for 15 hrs at 37°C with rotation at 70 rpm.

3. Cultures were diluted 1:500 into 5 ml of fresh medium and grown for three hours.

4. Cultures were split in half and arabinose was added to a final concentration of 0.2% to half of the cultures.

5. Cultures were grown in 17 mm test tubes for 12 hrs at 37°C with rotation at 70 rpm.

6. 200 μl aliquots of each of the cultures were transferred into a flat-bottomed 96 well plate (Cellstar Uclear bottom, cat. # T-3026-16, Greiner).

7. The plate was assayed in a Wallac Victor3 multi-well fluorimeter (Perkin Elmer) with an absorbance measurement (600 nm absorbance filter, 0.1 sec counting time through 5 mm of fluid) and a fluorescence measurement (488 nm excitation filter, 525 nm emission filter, 0.5 sec, CW lamp energy 12902 units).

8. Absorbance measurements were converted to OD600 by subtracting the absorbance of a well containing media only and then using an empirical calibration relating absorbance measured by the plate reader to OD600.

E.7 Construction of T7 RNAP-regulated O-ribosome generators

Complementary oligos encoding the promoters (BBa_R0183-7) were ordered. The primers included an NsiI site upstream of the promoter and an XhoI site downstream as well. For example, the primers for R0183 were NsiIR0183XhoIF (`5'-GTT TCA TGC ATt cat acg act cac tat agg gag aCT CGA GGA AGA-3'`) and NsiIR0183XhoIR `5'-TCT TCC TCG AGT CTC CCT ATA GTG AGT CGT ATG AAT GCA TGA AAC-3'`). The primers were annealed and digested with NsiI and XhoI (NEB) reaction conditions recommended by the manufacturer. pCH1497-rRNA4 was digested with NsiI and XhoI and gel purified. Ligation, transformation, and sequencing were all performed as described in Section E.1. The resulting plasmids were named pCH1497-rRNA4(R018x), where the x was replaced by 3, 4, 5, 6, 7 depending on the promoter.

E.8 Measurement protocol for T7 RNAP-regulated O-ribosome generators

1. 5 ml cultures of LB supplemented with antibiotics as appropriate (Ampicillin, 50 μg/ml; Tetracycline, 5 μg/ml) were inoculated with single colonies (2mm ϕ) from freshly streaked plates of BL21(DE3) containing the described devices. Three cultures started from single colonies were started for each strain. A culture was also grown with no reporter device as a fluorescence background control.

2. Cultures were grown in 17 mm test tubes for 12 hrs at 37°C with rotation at 70 rpm.

3. Cultures were diluted 1:500 into 5 ml of fresh medium and grown for 3 hours.

4. Cultures were split in half and IPTG was added to a final concentration of 0.8 mM to half of the cultures.

5. Cultures were grown in 17 mm test tubes for 12 hrs at 37°C with rotation at 70 rpm.

6. 200 μl aliquots of each of the cultures were transferred into a flat-bottomed 96 well plate (Cellstar Uclear bottom, cat. # T-3026-16, Greiner).

7. The plate was assayed in a Wallac Victor3 multi-well fluorimeter (Perkin Elmer) with an absorbance measurement (600 nm absorbance filter, 0.1 sec counting time through 5 mm of fluid) and a fluorescence measurement (488 nm excitation filter, 525 nm emission filter, 0.5 sec, CW lamp energy 12902 units).

8. Absorbance measurements were converted to OD600 by subtracting the absorbance of a well containing media only and then using an empirical calibration relating absorbance measured by the plate reader to OD600. The fluorescence background of cells not containing a reporter device was subtracted from all cultures.

E.9 Measurement protocol for VM2.0

This protocol was performed identically to that described in Section E.6. However prior to assaying in the multi-well fluorimeter, the cultures were resuspended in an equal volume of Phosphate-buffered saline to reduce the fluorescence background of the cultures.

E.10 Measurement protocol for the VM2.2 experiment

Note this experiment was performed identically to that described in Section G-2 except no arabinose was added after aliquoting to the plate.

E.11 Measurement protocol for the VM2.2 experiment with various reporter devices

Note this experiment was performed identically to that described in Section G-2 except no arabinose was added after aliquoting to the plate.

E.12 Measurement protocol for the VM2.2 variants

1. 5 ml cultures of LB supplemented with antibiotics as appropriate (Ampicillin, 50 μg/ml; Tetracycline, 5 μg/ml; Chloramphenicol, 34 μg/ml) were inoculated with single colonies (2mm ϕ) from freshly streaked plates of BL21(DE3) containing the described devices. Three cultures started from single colonies were started for each strain. In addition to the strains plotted in Figure 4-9, three cultures were also started for each VM2.2 variant with an empty plasmid (the same plasmid carrying the reporter).

2. Cultures were grown in 17 mm test tubes for 12 hrs at 37°C with rotation at 70 rpm.

3. Cultures were diluted 1:500 into 5 ml of fresh medium and grown for 10 hours.

4. 200 μl aliquots of each of the cultures were transferred into a flat-bottomed 96 well plate (Cellstar Uclear bottom, cat. # T-3026-16, Greiner).

5. The plate was assayed in a Wallac Victor3 multi-well fluorimeter (Perkin Elmer) with an absorbance measurement (600 nm absorbance filter, 0.1 sec counting time through 5 mm of fluid) and a fluorescence measurement (488 nm excitation filter, 525 nm emission filter, 0.5 sec, CW lamp energy 12902 units).

6. Absorbance measurements were converted to OD600 by subtracting the absorbance of a well containing media only and then using an empirical calibration

relating absorbance measured by the plate reader to OD600. The fluorescence background of the cells not containing a reporter device was subtracted from all cultures.

Appendix F

Genotype information

Table F.1: *E. coli* strains used or cited in this thesis

Strain	Genotype	Reference
BL21(AI)	F$^-$ ompT gal dcm lon hsdS$_B$(r$_B^-$m$_B^-$) araB::T7 gene 1-tetA	Invitrogen
BL21(DE3)	F$^-$ ompT gal dcm lon hsdS$_B$(r$_B^-$m$_B^-$) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])	[125]
DH5α	F$^-$, endA1, glnV44, thi-1, recA1, relA1, gyrA96, deoR, nupG, φ80dlacZΔM15, Δ(lacZYA-argF)U169, hsdR17(r$_K^-$m$_K^+$), λ-	[85]
DH10B	F$^-$ endA1, recA1, galE15, galK16, nupG, rpsL, ΔlacX74, φ80lacZΔM15, araD139, Δ(ara,leu)7697, mcrA, Δ(mr-hsdRMS-mcrBC), λ-	[23, 39]
D1210	HB101, lacI^q, lacY$^+$	[110]
HB101	F$^-$, mcrB, mrr, hsdS20(r$_B^-$m$_B^-$), recA13, leuB6, ara-14, proA2, lacY1, galK2, xyl-5, mtl-1, rpsL20(Sm^R), glnV44, λ-	[17]
JM109	endA1, glnV44, thi-1, relA1, gyrA96, recA1, mcrB$^+$, Δ(lac-proAB), e14-, [F', traD36, proAB$^+$, lacI^q, lacZΔM15], hsdR17(r$_K^-$m$_K^+$)	[136]
MC4100	[araD139] δ(argF- lac)169, LAM-, e14-, flhD5301, Δ(fruK-yeiR)725(fruA25), relA1, rpsL150(strR), rbsR22, ΔfimB-fimE)632(::IS1), deoC1	[22]
MG1655	F$^-$, ilvG-, rfb-50, rph-1, λ-	[15]
TOP10	DH10B	Invitrogen
VJS533	ara, Δ(lac-proAB)X111, rpsL, φ80lacZΔM15, recA56	[123]
WM54	MG1655, ΔlacX74	[79]

Appendix G

Virtual machine configurations

G.1 VM1.1

Figure G-1: The VM1.1 network schematic. Transcription of T7 gene 1 and the orthogonal rRNA is regulated by inducible promoters. The P_{BAD} promoter regulates transcription of the orthogonal rRNA and the *lac*UV5 promoter regulates transcription of T7 gene 1.

Strain: *E. coli* BL21(DE3) (Appendix F)

T7 RNAP source: T7 gene 1 is carried on a lambda lysogen (DE3) in *E. coli* BL21(DE3) (Appendix F). Transcription of T7 gene 1 is regulated by the *lac*UV5 promoter and can be induced by addition of IPTG to the culture medium [125].

O-ribosome source: A plasmid-encoded copy of a mutated rrnB operon (Table H.6) is under the control of a P_{BAD} promoter. Transcription of the mutated rrnB

operon can be induced by addition of arabinose to the culture medium. The plasmid is pCH1497-rRNA4. Construction of pCH1497-rRNA4 is described in Appendix D.
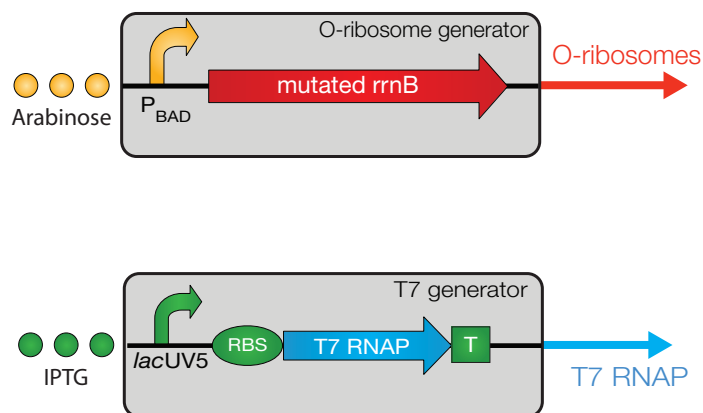
G.2   VM1.2
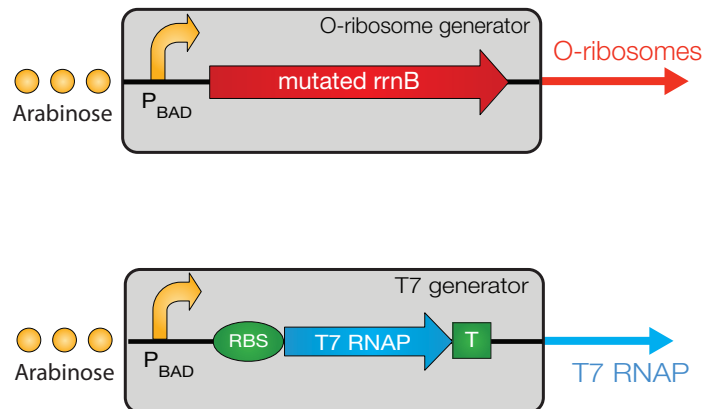
Figure G-2: The VM1.2 network schematic. Transcription of T7 gene 1 and the orthogonal rRNA is regulated by arabinose-inducible P_{BAD} promoters. Transcription downstream of T7 gene 1 is terminated by a synthetic *E. coli* terminator. A synthetic *E. coli* RBS regulates translation of the T7 gene 1 message.

Strain: *E. coli* TOP10 (Appendix F)

T7 RNAP source: T7 gene 1 is carried on pSB4C5 (see Table H.5). The transcription unit (BBa_I20279) consists of an arabinose-inducible P_{BAD} promoter (BBa_I0500), a synthetic *E. coli* RBS (BBa_B0032), T7 gene 1 (BBa_I2032), and a synthetic transcription terminator (BBa_B0015).

O-ribosome source: Identical to VM1.1.

G.3   VM2.0

Strain: *E. coli* D1210 (Appendix F)

T7 RNAP source: A T7 autogenerator is carried on pSB4C5 (Table H.5). The transcription unit (BBa_I20257) consists of a LacI-repressible T7 promoter (BBa_R0184),
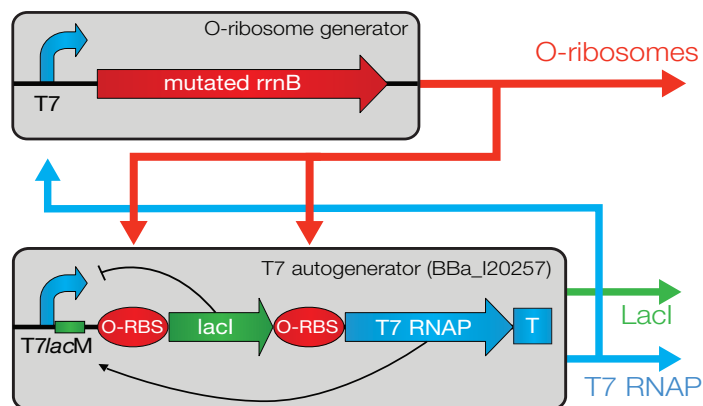
Figure G-3: The VM2.0 network schematic. A LacI-repressible T7 promoter transcribes both a *lacI* gene and T7 gene 1. Transcription of orthogonal rRNA is regulated by a T7 promoter. Translation of the *lacI* message and the T7 gene 1 message is regulated by synthetic O-RBS.

an O-RBS (BBa_B0073), a *lacI* gene (BBa_C0013), an O-RBS (BBa_B0073), T7 gene 1 (BBa_I2032), and the Tϕ transcription terminator (BBa_B0016).

O-ribosome source: A plasmid-encoded copy of a mutated rrnB operon is under the control of a T7 promoter (BBa_R0183). There is no T7 terminator downstream of the rrnB operon. The plasmid is pCH1497-rRNA4(R0183), which was constructed by replacing the P_{BAD} promoter in pCH1497-rRNA4 with the T7 promoter. Construction of pCH1497-rRNA4(R0183) is described in greater detail in Appendix D. The nucleotide sequence at the junction between the promoter and the mutated rrnB operon is shown in Table H.7.

Notes: This network was not stable in *E. coli* TOP10 but was stable in *E. coli* D1210. I was also unable to establish a variety of similar networks using LacI-repressible T7 promoters (BBa_R0184-7) in place of the T7 promoter regulating transcription of the orthogonal rRNA in *E. coli* TOP10.

G.4 VM2.2

Strain: *E. coli* TOP10 (Appendix F)

T7 RNAP source: A T7 autogenerator is carried on pSB4C5 (Table H.5). The tran-
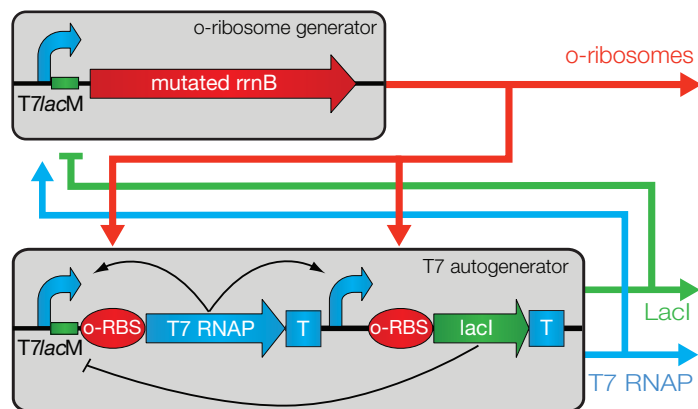
Figure G-4: The VM2.2 network schematic. Separate T7 promoters transcribe the *lacI* gene and T7 gene 1. The T7 promoter regulating T7 gene 1 transcription is LacI-repressible. Transcription of orthogonal rRNA is regulated by a LacI-repressible T7 promoter. Synthetic O-RBS regulate translation of both the *lacI* message and the T7 gene 1 message.

scription unit (BBa_I20281) consists of a LacI-repressible T7 promoter (BBa_R0184), an O-RBS (BBa_B0073), T7 gene 1 (BBa_I2032), the Tϕ transcription terminator (BBa_B0016), a T7 promoter (BBa_R0180), an O-RBS (BBa_B0073), a *lacI* gene (BBa_C0013), and the Tϕ transcription terminator (BBa_B0016).

O-ribosome source: A plasmid-encoded copy of a mutated rrnB operon is under the control of a LacI-repressible T7 promoter. There is no T7 terminator downstream of the rrnB operon. The plasmid is pCH1497-rRNA4(R0184). Construction of pCH1497-rRNA4(R0184) is described in Appendix D.

Notes: A variety of similar networks were also constructed and tested (see Section 4.4.7). Those network variants made use of an alternate O-ribosome generator with a putatively weaker LacI-repressible T7 promoter (BBa_R0185) replacing BBa_R0184. An alternate T7 autogenerator (BBa_I20280) was also tested that had a putatively weaker T7 promoter (BBa_R0182) regulating transcription of the *lacI* gene.

Appendix H

Genetic constructs

Information regarding the sequence, sub-parts, or source of the various genetic constructs used in my thesis is provided in this appendix. Further details for many of the constructs are available from the registry (`http://parts.mit.edu`).

Table H.1: Promoters used in this work.

BioBrick #	Sequence[a]	Reference
BBa_R0040	`tccctatcagtgatagagattgacatccctatcagtgatagagatactgagcac`	[81]
BBa_I0500	`See http://partsregistry.org`	[74]
BBa_R0085	`taatacgactcactatagggaga`	[38]
BBa_R0180	`-t--------------------`	[64]
BBa_R0181	`g---------------------`	[64]
BBa_R0182	`-------t--------------`	[64]
BBa_R0183	`-c--------------------`	[64]
BBa_R0184	`-c-----------------gaattgtgagcggataacaattcc`	[64, 36]
BBa_R0185	`-------g-----------gaattgtgagcggataacaattcc`	[64, 36]
BBa_R0186	`--------a----------gaattgtgagcggataacaattcc`	[64, 36]
BBa_R0187	`--------g----------gaattgtgagcggataacaattcc`	[64, 36]

BBa_R0040 and BBa_I0500 are *E. coli* promoters. All other promoters listed are T7 promoters.
[a]BBa_R0085 is the consensus T7 promoter sequence. For all other T7 promoters, only bases that differ from the consensus sequence are shown.

Table H.2: GFP reporters used in this work

Identifier[a]	Classification[b]	Promoter[c]	RBS[d]	Terminator[e]
BBa_I7101	e/e		BBa_B0032	
BBa_I7102			BBa_B0036	
BBa_E70101			BBa_B0037	
BBa_E70102	e/o	BBa_R0040	BBa_B0070	BBa_B0015
BBa_E70201			BBa_B0072	
BBa_E70202			BBa_B0073	
BBa_E70203			BBa_B0074	
BBa_E7104		BBa_R0085		
BBa_E7105		BBa_R0180		
BBa_E7107	o/e	BBa_R0182	BBa_B0032	BBa_B0016
BBa_E7108		BBa_R0183		
BBa_I2035		BBa_R0184		
BBa_E71205	o/o	BBa_R0183	BBa_B0073	

All reporters with the exception of BBa_I2035 were carried on pSB1A3, BBa_I2035 was carried on pSB1AT3 (see Table H.5). BBa_E0040, based on GFPmut3b [31] was the GFP coding region used in all reporters. [a]The BioBrick number of the reporter. [b]e/e denotes a reporter that uses *E. coli* transcription and translation. e/o denotes a reporter that uses *E. coli* transcription and orthogonal translation. o/e denotes a reporter that uses orthogonal transcription and *E. coli* translation. o/o denotes a reporter that uses orthogonal transcription and translation. [c]All promoter sequences are provided in Table H.1. [d]All RBS sequences are provided in Table H.3. [e]All terminator sequences are provided in Table H.4.

Table H.3: RBS used in this work.

BioBrick #	Sequence[a]	Reference
BBa_B0032	tactagagTCACACAGGAAAGtactagatg	[132]
BBa_B0036	tactagagGTGTGtactagatg	[20]
BBa_B0037	tactagagGTGTGTCTAGtactagatg	[20]
BBa_B0070	tactagagTCTACGTGTGTCAAGtactagatg	[20]
BBa_B0072	tactagagCACCACtactagatg	[99]
BBa_B0073	tactagagTCACACCACtactagatg	[99]
BBa_B0074	tactagagTCACACCACCCtactagatg	[99]

BBa_B0032 is an *E. coli* promoter. All other promoters listed are orthogonal RBS. [a]The sequence of the RBS is shown in capital letters, along with the bracketing sequences resulting from BioBrick assemblies and the start codon downstream. Additional bases may be added to the 5' end of the transcript depending on the promoter design.

Table H.4: Terminators used in this work.

BioBrick #	Sequence[a]	Reference
BBa_B0015	ccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctt	see partsregistry.org
	tcgttttatctgttgtttgtcggtgaacgctctctactagagtcac	
	actggctcaccttcgggtgggcctttctgcgtttata	
BBa_B0016	ctagcataaccccttggggcctctaaacgggtcttgaggggtttttttg	[38]

BBa_B0015 is a synthetic *E. coli* terminator composed of two natural terminators (TE from bacteriophage T7 and T1 from *E. coli* rrnB). BBa_B0016 is based on the T7-Tϕ terminator from bacteriophage T7.

Table H.5: Plasmids used in this work

Identifier	Replicon	Antibiotic resistance	Copy number [111]
pSB1A3	modified pMB1 derived from pUC19	Amp^r	500-700
pSB1AT3	modified pMB1 derived from pUC19	Amp^r & Tet^r	500-700
pSB3k3	p15A derived from pMR101	Kan^r	10-12
pSB4C5	rep101, repA derived from pSC101	Cam^r	~ 5
pCH1497	p15a derived from pACYC184	Tet^r	10-12

Table H.6: Mutations to rrnB leading to recognition of orthogonal RBS

Identifier	ASD[a]	Reference
wt rrnB 16s rRNA	^{1535}CCUCC^{1539}	[122]
pCH1497-ASD1	^{1535}CACAC^{1539}	[20]
pCH1497-rRNA4	^{1535}UUGUGG^{1540}	[99]

Table H.7: Nucleotide sequences of the promoter regions of T7 RNAP-transcribed ribosomal rRNA transcripts from this work and others

Identifier	Sequence[a]	Reference
wt *E. coli* rrnB (P2 promoter)	gcgccgctgagaaaaagcgaagcg	[15]
pT7-1	gggagaccacaacggttccctctagcgggatccggtaccggccgctgagaaaaagcgaagcg	[77]
pAR3056	gggagaccacaacgtttccctctagcgggatccggtaccgg	[121]
pCH1497-rRNA4(R0183)	gggagactcgagtattatgcacaccccgcgccgctagaaaaagcgaagcg	This work
pCH1497-rRNA4(R0184)	ggggaattgtgagcggataacaattccctcgagtattatgcacaccccgcgccgctagaaaaagcgaagcg	This work

[a]Sequences are shown from the transcriptional start site of each construct through to the first nucleotide of the ribosomal operon transcribed in common by all constructs.

Bibliography

[1] H Alper, C Fischer, E Nevoigt, and G Stephanopoulos. Tuning genetic control through promoter engineering. *Proc Natl Acad Sci U S A*, 102(36):12678–12683, Sep 2005.

[2] J B Andersen, A Heydorn, M Hentzer, L Eberl, O Geisenberger, B B Christensen, S Molin, and M Givskov. *gfp*-based *N*-acyl homoserine-lactone sensor systems for detection of bacterial communication. *Appl Environ Microbiol*, 67(0099-2240 (Print)):575–85, 2001.

[3] J C Anderson, E J Clarke, A P Arkin, and C A Voigt. Environmentally controlled invasion of cancer cells by engineered bacteria. *J Mol Biol*, 355(4):619–627, 2006.

[4] J C Anderson, C A Voigt, and A P Arkin. Environmental signal integration by a modular AND gate. *Mol Syst Biol*, 3:133, 2007.

[5] A P Arkin and D A Fletcher. Fast, cheap and somewhat in control. *Genome Biol*, 7(8):114, 2006.

[6] I Artsimovitch, V Svetlov, L Anthony, R R Burgess, and R Landick. RNA polymerases from *Bacillus subtilis* and *Escherichia coli* differ in recognition of regulatory signals *in vitro*. *J Bacteriol*, 182(0021-9193):6027–35, 2000.

[7] T Asai, C Condon, J Voulgaris, D Zaporojets, B Shen, M Al-Omar, C Squires, and C L Squires. Construction and initial characterization of *Escherichia coli* strains with few or no intact chromosomal rRNA operons. *J Bacteriol*, 181(12):3803–3809, 1999.

[8] T Asai, D Zaporojets, C Squires, and C L Squires. An *Escherichia coli* strain with all chromosomal rRNA operons inactivated: complete exchange of rRNA genes between bacteria. *Proc Natl Acad Sci U S A*, 96(5):1971–6, 1999.

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

[10] S Bagh, M Mazumder, T Velauthapillai, V Sardana, G Q Dong, A B Movva, L H Lim, and D R McMillen. Plasmid-borne prokaryotic gene expression: Sources

of variability and quantitative system characterization. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 77(2):021919, 2008.

[11] R P Bandwar, Y Jia, N M Stano, and S S Patel. Kinetic and thermodynamic basis of promoter strength: Multiple steps of transcription initiation by T7 RNA polymerase are modulated by the promoter sequence. *Biochemistry*, 41(11):3586–3595, 2002.

[12] N Barkai and S Leibler. Circadian clocks limited by noise. *Nature*, 403(6767):267–268, 2000.

[13] S Basu, Y Gerchman, C H Collins, F H Arnold, and R Weiss. A synthetic multicellular system for programmed pattern formation. *Nature*, 434(1476-4687):1130–4, 2005.

[14] J A Bernstein, A B Khodursky, P Lin, S Lin-Chao, and S N Cohen. Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc Natl Acad Sci U S A*, 99(15):9697–9702, 2002.

[15] F R Blattner, G Plunkett III, C A Bloch, N T Perna, V Burland, M Riley, J Collado-Vides, J D Glasner, C K Rode, G F Mayhew, J Gregor, N W Davis, H A Kirkpatrick, M A Goeden, D J Rose, B Mau, and Y Shao. The complete genome sequence of *Escherichia coli* K-12. *Science*, 277(5331):1453–1474, 1997.

[16] K J Boettcher and E G Ruby. Detection and quantification of *Vibrio fischeri* autoinducer from symbiotic squid light organs. *J Bacteriol*, 177(4):1053–8, 1995.

[17] H W Boyer and D Roulland-Dussoix. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J Mol Biol*, 41(3):459–472, 1969.

[18] H Bremer and P Dennis. *Escherichia coli and Salmonella: Cellular and Molecular Biology*, chapter Modulation of chemical composition and other parameters of the cell by growth rate. ASM Press, 1996.

[19] K Brindley. *Sensors and transducers*. Heinemann Professional Publishing, Portsmouth, New Hampshire, USA, 1988.

[20] M F Brink, M P Verbeet, and H A de Boer. Specialized ribosomes: highly specific translation *in vivo* of a single targetted mRNA species. *Gene*, 156(2):215–22, 1995.

[21] M H Buckstein, J He, and H Rubin. Characterization of nucleotide pools as a function of physiological state in *Escherichia coli*. *J Bacteriol*, 190(2):718–726, 2008.

[22] M J Casadaban. Transposition and fusion of the *lac* genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. *J Mol Biol*, 104(3):541–555, 1976.

[23] M J Casadaban and S N Cohen. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J Mol Biol*, 138(2):179–207, 1980.

[24] M Chamberlin, J McGrath, and L Waskell. New RNA polymerase from *Escherichia coli* infected with bacteriophage T7. *Nature*, 228(5268):227–231, 1970.

[25] L Y Chan, S Kosuri, and D Endy. Refactoring bacteriophage T7. *Mol Syst Biol*, 1:2005.0018, 2005.

[26] A C Chang and S N Cohen. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol*, 134(3):1141–1156, 1978.

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

[28] C T Chung, S L Niemela, and R H Miller. One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci U S A*, 86(7):2172–2175, 1989.

[29] C H Collins, F H Arnold, and J R Leadbetter. Directed evolution of *Vibrio fischeri* LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones. *Mol Microbiol*, 55(3):712–23, 2005.

[30] C H Collins, J R Leadbetter, and F H Arnold. Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR. *Nat Biotechnol*, 24(6):708–12, 2006.

[31] B P Cormack, R H Valdivia, and S Falkow. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*, 173(0378-1119 (Print)):33–8, 1996.

[32] C W de Silva. *Sensors and actuators: control system instrumentation*. CRC Press, Boca Raton, Florida, 2007.

[33] M H de Smit and J van Duin. Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proc Natl Acad Sci U S A*, 87(0027-8424):7668–72, 1990.

[34] M H de Smit and J van Duin. Translational standby sites: how ribosomes may deal with the rapid folding kinetics of mRNA. *J Mol Biol*, 331(0022-2836):737–43, 2003.

[35] P P Dennis, M Ehrenberg, and H Bremer. Control of rRNA synthesis in *Escherichia coli*: a systems biology approach. *Microbiol Mol Biol Rev*, 68(4):639–68, 2004.

[36] J W Dubendorff and F W Studier. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with *lac* repressor. *J Mol Biol*, 219(1):45–59, 1991.

[37] J W Dubendorff and F W Studier. Creation of a T7 autogene. cloning and expression of the gene for bacteriophage T7 RNA polymerase under control of its cognate promoter. *J Mol Biol*, 219(1):61–8, 1991.

[38] J J Dunn and F W Studier. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J Mol Biol*, 166(0022-2836):477–535, 1983.

[39] T Durfee, R Nelson, S Baldwin, G Plunkett III, V Burland, B Mau, J F Petrosino, X Qin, D M Muzny, M Ayele, R A Gibbs, B Csorgo, G Posfai, G M Weinstock, and F R Blattner. The complete genome sequence of *Escherichia coli* DH10B: insights into the biology of a laboratory workhorse. *J Bacteriol*, 190(7):2597–2606, 2008.

[40] F Dyson. The darwinian interlude, 2005.

[41] A Eberhard, A L Burlingame, C Eberhard, G L Kenyon, K H Nealson, and N J Oppenheimer. Structural identification of autoinducer of *Photobacterium fischeri* luciferase. *Biochemistry*, 20(9):2444–9, 1981.

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

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

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

[45] D Endy, I Deese, and C Wadey. Adventures in synthetic biology. *Nature [online]*, 2005.

[46] J Engebrecht and M Silverman. Identification of genes and gene products necessary for bacterial bioluminescence. *Proc Natl Acad Sci U S A*, 81(13):4154–8, 1984.

[47] C French, J Nicholson, F Bizzari, J Aleksic, Y Cai, S Seshasayee, S Ivakhno, B Davidson, J Wilson, K de Mora, H Ma, L Kozma-Bognar, and A Elfick. Arsenic biosenseor: a step further. *BMC Systems Biology*, 1(Suppl 1):S11, 2007.

[48] C Fuqua, S C Winans, and E P Greenberg. Census and consensus in bacterial ecosystems: the LuxR-LuxI family of quorum-sensing transcriptional regulators. *Annu Rev Microbiol*, 50:727–751, 1996.

[49] J Garcia-Ojalvo, M B Elowitz, and S H Strogatz. Modeling a synthetic multi-cellular clock: repressilators coupled by quorum sensing., 2004.

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

[51] D G Gibson, G A Benders, C Andrews-Pfannkoch, E A Denisova, H Baden-Tillson, J Zaveri, T B Stockwell, A Brownley, D W Thomas, M A Algire, C Merryman, L Young, V N Noskov, J I Glass, J C Venter, C A Hutchison III, and H O Smith. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science*, 319(5867):1215–1220, 2008.

[52] J I Glass, N Assad-Garcia, N Alperovich, S Yooseph, M R Lewis, M Maruf, C A Hutchison III, H O Smith, and J C Venter. Essential genes of a minimal bacterium. *Proc Natl Acad Sci U S A*, 103(2):425–430, 2006.

[53] B R Glick. Metabolic load and heterologous gene expression. *Biotechnol Adv*, 13(2):247–61, 1995.

[54] R F Goldberger. Autogenous regulation of gene expression. *Science*, 183(4127):810–816, 1974.

[55] A Gordon, A Colman-Lerner, T E Chin, K R Benjamin, R C Yu, and R Brent. Single-cell quantification of molecules and rates using open-source microscope-based cytometry. *Nat Methods*, 4(2):175–181, 2007.

[56] S Graslund, P Nordlund, J Weigelt, B M Hallberg, J Bray, O Gileadi, S Knapp, U Oppermann, C Arrowsmith, R Hui, J Ming, S dhe Paganon, H Park, A Savchenko, A Yee, A Edwards, R Vincentelli, C Cambillau, R Kim, S Kim, Z Rao, Y Shi, T C Terwilliger, C Kim, L Hung, G S Waldo, Y Peleg, S Albeck, T Unger, O Dym, J Prilusky, J L Sussman, R C Stevens, S A Lesley, I A Wilson, A Joachimiak, F Collart, I Dementieva, M I Donnelly, W H Eschenfeldt, Y Kim, L Stols, R Wu, M Zhou, S K Burley, J S Emtage, J M Sauder, D Thompson, K Bain, J Luz, T Gheyi, F Zhang, S Atwell, S C Almo, J B Bonanno, A Fiser, S Swaminathan, F W Studier, M R Chance, A Sali, T B Acton, R Xiao, L Zhao, L C Ma, J F Hunt, L Tong, K Cunningham, M Inouye, S Anderson, H Janjua, R Shastry, C K Ho, D Wang, H Wang, M Jiang, G T Montelione, D I Stuart, R J Owens, S Daenke, A Schutz, U Heinemann, S Yokoyama, K Bussow, and K C Gunsalus. Protein production and purification. *Nat Methods*, 5(2):135–146, 2008.

[57] T M Gruber and C A Gross. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol*, 57:441–466, 2003.

[58] C C Guet, M B Elowitz, W Hsing, and S Leibler. Combinatorial synthesis of genetic networks. *Science*, 296(5572):1466–70, 2002.

[59] E L Haseltine and F H Arnold. Synthetic gene circuits: design with directed evolution. *Annu Rev Biophys Biomol Struct*, 36:1–19, 2007.

[60] K Hasunuma and M Sekiguchi. Replication of plasmid pSC101 in *Escherichia coli* K12: requirement for dnaA function. *Mol Gen Genet*, 154(3):225–230, 1977.

[61] W S Hlavacek and M A Savageau. Rules for coupled expression of regulator and effector genes in inducible circuits. *J Mol Biol*, 255(1):121–139, 1996.

[62] A Hui and H A de Boer. Specialized ribosome system: preferential translation of a single mrna species by a subpopulation of mutated ribosomes in escherichia coli. *Proc Natl Acad Sci U S A*, 84(14):4762–6, 1987.

[63] R A Ikeda. The efficiency of promoter clearance distinguishes T7 class II and class III promoters. *J Biol Chem*, 267(0021-9258):11322–8, 1992.

[64] D Imburgio, M Rong, K Ma, and W T McAllister. Studies of promoter recognition and start site selection by T7 RNA polymerase using a comprehensive collection of promoter variants. *Biochemistry*, 39(34):10419–30, 2000.

[65] Texas Instruments. *TTL logic: standard TTL, Schottky, low-power Schottky*. Texas Instruments, Dallas, Texas, 1988.

[66] M Itaya, K Tsuge, M Koizumi, and K Fujita. Combining two genomes in one cell: Stable cloning of the *Synechocystis* PCC6803 genome in the *Bacillus subtilis* 168 genome. *Proc Natl Acad Sci U S A*, 102(44):15971–15976, 2005.

[67] A Jacobson and D Gillespie. Metabolic events occurring during recovery from prolonged glucose starvation in *Escherichia coli*. *J Bacteriol*, 95(3):1030–1039, 1968.

[68] Y Jia, A Kumar, and S S Patel. Equilibrium and stopped-flow kinetic studies of interaction between T7 RNA polymerase and its promoters measured by protein and 2-aminopurine fluorescence changes. *J Biol Chem*, 271(48):30451–30458, 1996.

[69] D Kennell and H Riezman. Transcription and translation initiation frequencies of the *Escherichia coli lac* operon. *J Mol Biol*, 114(1):1–21, 1977.

[70] J C King. Biotechnology: a solution for improving nutrient bioavailability. *Int J Vitam Nutr Res*, 72(1):7–12, 2002.

[71] T F Knight. Idempotent vector design for standard assembly of BioBricks. *Synthetic Biology Working Group technical reports*, 2003.

[72] T F Knight. Engineering novel life. *Mol Syst Biol*, 1:2005.0020, 2005.

[73] C Kohrer, E L Sullivan, and U L RajBhandary. Complete set of orthogonal 21st aminoacyl-tRNA synthetase-amber, ochre and opal suppressor tRNA pairs: concomitant suppression of three different termination codons in an mRNA in mammalian cells. *Nucleic Acids Research*, 32(21):6200–6211, 2004.

[74] N L Lee, W O Gielow, and R G Wallace. Mechanism of araC autoregulation and the domains of two overlapping promoters, Pc and PBAD, in the L-arabinose regulatory region of *Escherichia coli. Proc Natl Acad Sci U S A*, 78(2):752–756, 1981.

[75] J H Leveau and S E Lindow. Predictive and interpretive simulation of green fluorescent protein expression in reporter bacteria. *J Bacteriol*, 183(23):6752–6762, 2001.

[76] A Levskaya, A A Chevalier, J J Tabor, Z B Simpson, L A Lavery, M Levy, E A Davidson, A Scouras, A D Ellington, E M Marcotte, and C A Voigt. Synthetic biology: engineering *Escherichia coli* to see light. *Nature*, 438(7067):441–442, 2005.

[77] B T U Lewicki, T Margus, J Remme, and K H Nierhaus. Coupling of ribosomal-RNA transcription and ribosomal assembly *in vivo* - formation of active ribosomal-subunits in *Eschericia coli* requires transcription of ribosomal-RNA genes by host RNA-polymerase which cannot be replaced by bacteriophage-T7 RNA-polymerase. *J of Mol Biol*, 231(3):581–593, Jun 1993.

[78] R Liang, X Liu, J Liu, Q Ren, P Liang, Z Lin, and X Xie. A T7-expression system under temperature control could create temperature-sensitive phenotype of target gene in *Escherichia coli. J Microbiol Methods*, 68(3):497–506, 2007.

[79] A Lindsay and B M M Ahmer. Effect of sdiA on biosensors of N-acylhomoserine lactones. *J Bacteriol*, 187(14):5054–8, 2005.

[80] S T Lovett. Encoded errors: mutations and rearrangements mediated by mis-alignment at repetitive DNA sequences. *Mol Microbiol*, 52(5):1243–1253, 2004.

[81] R Lutz and H Bujard. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res*, 25(6):1203–10, 1997.

[82] O V Makarova, E M Makarov, R Sousa, and M Dreyfus. Transcribing of *Escherichia coli* genes with mutant T7 RNA polymerases: stability of *lacZ* mRNA inversely correlates with polymerase speed. *Proc Natl Acad Sci U S A*, 92(0027-8424):12250–4, 1995.

[83] A E Mayo, Y Setty, S Shavit, A Zaslaver, and R Alon. Plasticity of the *cis*-regulatory input function of a gene. *PLoS Biol*, 4(4):e45, 2006.

[84] K H McClean, M K Winson, L Fish, A Taylor, S R Chhabra, M Camara, M Daykin, J H Lamb, S Swift, B W Bycroft, G S Stewart, and P Williams. Quorum sensing and *Chromobacterium violaceum*: exploitation of violacein production and inhibition for the detection of *N*-acylhomoserine lactones. *Microbiology*, 143 (Pt 12)(1350-0872 (Print)):3703–11, 1997.

[85] M Meselson and R Yuan. DNA restriction enzyme from *E. coli*. *Nature*, 217(5134):1110–1114, 1968.

[86] B Müller-Hill. *The* lac *operon*. Walter de Gruyter, Berlin; New York, 1996.

[87] K H Nealson. Autoinduction of bacterial luciferase. occurrence, mechanism and significance. *Arch Microbiol*, 112(1):73–79, 1977.

[88] K H Nealson and J W Hastings. Bacterial bioluminescence: its control and ecological significance. *Microbiol Rev*, 43(4):496–518, 1979.

[89] F C Neidhardt, P L Bloch, and D F Smith. Culture medium for enterobacteria. *J Bacteriol*, 119(3):736–47, 1974.

[90] M Nomura. Engineering of bacterial ribosomes: replacement of all seven *Escherichia coli* rRNA operons by a single plasmid-encoded operon. *Proc Natl Acad Sci U S A*, 96(5):1820–2, 1999.

[91] S Oehler, E R Eismann, H Kramer, and B Müller-Hill. The three operators of the *lac* operon cooperate in repression. *EMBO J*, 9(4):973–979, 1990.

[92] OpenWetWare. 'Round-the-horn site-directed mutagenesis — openwetware, 2007. [Online; accessed 12-May-2008].

[93] OpenWetWare. TOP10 chemically competent cells — openwetware, 2008. [Online; accessed 11-May-2008].

[94] B J Paul, W Ross, T Gaal, and R L Gourse. rRNA transcription in *Escherichia coli*. *Annu Rev Genet*, 38:749–770, 2004.

[95] S W Peretti and J E Bailey. Simulations of host-plasmid interactions in *Escherichia coli*: Copy number, promoter strength, and ribosome binding site strength effects on metabolic activity and plasmid gene expression. *Biotechnol Bioeng*, 29(3):316–328, 1987.

[96] A M Pomini, G P Manfio, W L Araujo, and A J Marsaioli. Acyl-homoserine lactones from *Erwinia psidii* R. IBSBF 435^t, a guava phytopathogen (*Psidium guajava* L.). *J Agric Food Chem*, 53(16):6262–5, 2005.

[97] G Posfai, G Plunkett III, T Feher, D Frisch, G M Keil, K Umenhoffer, V Kolisnychenko, B Stahl, S S Sharma, M de Arruda, V Burland, S W Harcum, and F R Blattner. Emergent properties of reduced-genome *Escherichia coli*. *Science*, 312(5776):1044–6, 2006.

[98] J Pramanik and J D Keasling. Effect of *Escherichia coli* biomass composition on central metabolic fluxes predicted by a stoichiometric model. *Biotechnol Bioeng*, 60(2):230–238, 1998.

[99] O Rackham and J W Chin. A network of orthogonal ribosome·mRNA pairs. *Nat Chem Biol*, 1(3):159–166, 2005.

[100] R Rauhut and G Klug. mRNA degradation in bacteria. *FEMS Microbiol Rev*, 23(3):353–370, 1999.

[101] D Rebatchouk, N Daraselia, and J O Narita. NOMAD: a versatile strategy for *in vitro* DNA manipulation applied to promoter analysis and vector design. *Proc Natl Acad Sci U S A*, 93(20):10891–10896, 1996.

[102] K Rinaudo, L Bleris, R Maddamsetti, S Subramanian, R Weiss, and Y Benenson. A universal RNAi-based logic evaluator that operates in mammalian cells. *Nat Biotechnol*, 25(7):795–801, 2007.

[103] D Ro, E M Paradise, M Ouellet, K J Fisher, K L Newman, J M Ndungu, K A Ho, R A Eachus, T S Ham, J Kirby, M C Y Chang, S T Withers, Y Shiba, R Sarpong, and J D Keasling. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, 440(7086):940–943, 2006.

[104] MW Roberts and JC Rabinowitz. The effect of *Escherichia coli* ribosomal protein S1 on the translational specificity of bacterial ribosomes. *Journal of Biological Chemistry*, 264(4):2228–2235, 1989.

[105] M Rong, B He, W T McAllister, and R K Durbin. Promoter specificity determinants of T7 RNA polymerase. *Proc Natl Acad Sci U S A*, 95(2):515–519, 1998.

[106] N Rosenfeld, M B Elowitz, and U Alon. Negative autoregulation speeds the response times of transcription networks. *J Mol Biol*, 323(5):785–93, 2002.

[107] N Rosenfeld, J W Young, U Alon, P S Swain, and M B Elowitz. Gene regulation at the single-cell level. *Science*, 307(5717):1962–1965, 2005.

[108] N Rosenfeld, J W Young, U Alon, P S Swain, and M B Elowitz. Accurate prediction of gene feedback circuit behavior from component properties. *Mol Syst Biol*, 3:143, 2007.

[109] R Rucker. *Wetware*. Avon Books, New York, New York, USA, 1988.

[110] J R Sadler, M Tecklenburg, and J L Betz. Plasmids containing many tandem copies of a synthetic lactose operator. *Gene*, 8(3):279–300, 1980.

[111] J Sambrook and D.W. Russell. *Molecular Cloning*. Cold Spring Harbor Press, Cold Spring Harbor, New York, USA, 2nd edition, 2001.

[112] D F Savage, J Way, and P A Silver. Defossiling fuel: how synthetic biology can transform biofuel production. *ACS Chem Biol*, 3(1):13–16, 2008.

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

[114] A L Schaefer, B L Hanzelka, A Eberhard, and E P Greenberg. Quorum sensing in *Vibrio fischeri*: probing autoinducer-LuxR interactions with autoinducer analogs. *J Bacteriol*, 178(0021-9193 (Print)):2897–901, 1996.

[115] T. D. Schneider and R. M. Stephens. Sequence logos: A new way to display consensus sequences. *Nucleic Acids Res.*, 18:6097–6100, 1990.

[116] W Sellers. On a uniform system of screw threads. *Journal of the Franklin Institute*, 47:344, May 1864.

[117] N C Shaner, P A Steinbach, and R Y Tsien. A guide to choosing fluorescent proteins. *Nat Methods*, 2(12):905–9, 2005.

[118] R P Shetty. *Engineering transcription-based logic*. PhD thesis, Massachusetts Institute of Technology, 77 Massachussetts Avenue, Cambridge, MA, USA, 2008.

[119] R P Shetty, D Endy, and T F Knight Jr. Engineering BioBrick vectors from BioBrick parts. *J Biol Eng*, 2(1):5, 2008.

[120] A K Srivastava and D Schlessinger. Mechanism and regulation of bacterial ribosomal RNA processing. *Annu Rev Microbiol*, 44:105–129, 1990.

[121] R Steen, A E Dahlberg, B N Lade, F W Studier, and J J Dunn. T7 RNA polymerase directed expression of the *Escherichia coli* rrnB operon. *EMBO J*, 5(5):1099–103, 1986.

[122] J A Steitz and K Jakes. How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. *Proc Natl Acad Sci U S A*, 72(12):4734–4738, 1975.

[123] V Stewart and J Jr Parales. Identification and expression of genes *narL* and *narX* of the *nar* (nitrate reductase) locus in *Escherichia coli* K-12. *J Bacteriol*, 170(4):1589–97, 1988.

[124] F W Studier. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J Mol Biol*, 219(1):37–44, 1991.

[125] F W Studier and B A Moffatt. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol*, 189(1):113–30, 1986.

[126] A Ujvari and C T Martin. Thermodynamic and kinetic measurements of promoter binding by T7 RNA polymerase. *Biochemistry*, 35(46):14574–14582, 1996.

[127] R L Ulrich, D Deshazer, E E Brueggemann, H B Hines, P C Oyston, and J A Jeddeloh. Role of quorum sensing in the pathogenicity of *Burkholderia pseudomallei*. *J Med Microbiol*, 53(Pt 11):1053–1064, 2004.

[128] R Van Houdt, A Aertsen, P Moons, K Vanoirbeek, and C W Michiels. *N*-acyl-L-homoserine lactone signal interception by *Escherichia coli*. *FEMS Microbiol Lett*, 256(0378-1097 (Print)):83–9, 2006.

[129] J Vind, M A Sorensen, M D Rasmussen, and S Pedersen. Synthesis of proteins in *Escherichia coli* is limited by the concentration of free ribosomes. expression from reporter genes does not always reflect functional mRNA levels. *J Mol Biol*, 231(3):678–88, 1993.

[130] U Vogel and K F Jensen. The RNA chain elongation rate in *Escherichia coli* depends on the growth rate. *J Bacteriol*, 176(10):2807–2813, 1994.

[131] A Ward and Halstead R.H. *Computation structures*. MIT Press, Cambridge, Massachussetts, USA, 1990.

[132] R Weiss. *Cellular Computation and Communications using Engineered Genetic Regulatory Networks*. PhD thesis, Massachusetts Institute of Technology, 2001.

[133] R Weiss and T Knight. Engineered communications for microbial robotics. In *DNA6 Sixth International Meeting on DNA Based Computers*, Lecture Notes in Computer Science. Springer Berlin / Heidelberg, June 2000.

[134] F W Whipple and A L Sonenshein. Mechanism of initiation of transcription by *Bacillus subtilis* RNA polymerase at several promoters. *J Mol Biol*, 223(2):399–414, 1992.

[135] M K Winson, S Swift, L Fish, J P Throup, F Jorgensen, S R Chhabra, B W Bycroft, P Williams, and G S Stewart. Construction and analysis of *luxCDABE*-based plasmid sensors for investigating *N*-acyl homoserine lactone-mediated quorum sensing. *FEMS Microbiol Lett*, 163(2):185–92, 1998.

[136] C Yanisch-Perron, J Vieira, and J Messing. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, 33(1):103–119, 1985.

[137] S Ye, C Kohrer, T Huber, M Kazmi, P Sachdev, E C Y Yan, A Bhagat, U L RajBhandary, and T P Sakmar. Site-specific incorporation of keto amino acids into functional G protein-coupled receptors using unnatural amino acid mutagenesis. *Journal of Biological Chemistry*, 283(3):1525–1533, 2008.

[138] Y Yokobayashi, R Weiss, and F H Arnold. Directed evolution of a genetic circuit. *Proc Natl Acad Sci U S A*, 99(26):16587–16591, 2002.

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

[140] L Zhang, P J Murphy, A Kerr, and M E Tate. Agrobacterium conjugation and gene regulation by *N*-acyl-L-homoserine lactones. *Nature*, 362(6419):446–448, 1993.

[141] X Zhang and F W Studier. Mechanism of inhibition of bacteriophage T7 RNA polymerase by T7 lysozyme. *J Mol Biol*, 269(1):10–27, 1997.

[142] M Zuker. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*, 31(13):3406–3415, 2003.