Massachusetts Institute of Technology
Biological Engineering Division

Thesis Proposal
Doctor of Philosophy

Title:

Engineering the Interface Between Cellular Chassis and Integrated Biological Systems

Date of Submission:

June 23, 2005

SUBMITTED BY: Barry Canton
31 Ames Street, Room 68-558D
Cambridge, MA 02139

SIGNED: _____

SUPERVISOR: Drew Endy
Assistant Professor
Biological Engineering Division

SIGNED: _____

ACADEMIC OFFICE: Dalia Gabour
Academic Administrator
Biological Engineering Division

SIGNED: _____

Contents

Acknowledgments

All of the following individuals deserve my sincere thanks. Many of the ideas behind this project and synthetic biology in general were first developed by Tom Knight. In the context of this project, it was Tom who described the host cell as a power supply and chassis for engineered biological systems. My supervisor, Drew Endy provided invaluable advice, suggestions, and support. Kathleen McGinness and Chris Hayes have provided plasmids and advice on the construction of dedicated translation machinery and more. Ruben Gonzalez provided several ideas on how to measure ribosome and polymerase levels. Members of the Endy and Knight labs have always been available for long and stimulating discussions about engineering biology and this project in particular. Funding for this work has been generously provided by the National University of Ireland. Finally, I wish to thank the members of my thesis committee, listed below.

Edward F. DeLong
Biological Engineering Division & Department of Civil and Environmental Engineering
MIT

Drew Endy
Biological Engineering Division
MIT

Thomas Knight Jr.
Computer Science and Artificial Intelligence Laboratory
MIT

Uttam L. RajBhandary
Biology Department
MIT

F. William Studier
Biology Department
Brookhaven National Laboratory

Abstract

The engineering of biological systems with predictable behavior is a challenging problem. One reason for this difficulty is that engineered biological systems are embedded within complex and variable host cells. To help enable the future engineering of biological systems, I will study and optimize the interface between an engineered biological system and its host cell or "chassis". Other engineering disciplines use modularity to make interacting systems interchangeable and to insulate one system from another. Engineered biological systems are more likely to work as predicted if system function is decoupled from the state of the host cell. Also, specifying and standardizing the interfaces between a system and the chassis will allow systems to be engineered independent of chassis and allow systems to be interchanged between different chassis. To this end, I will build dedicated transcription and translation systems, independent from the equivalent host cell systems. In parallel, I will develop test systems and metrics to measure the interactions between an engineered system and its chassis. Lastly, I will explore methods to"port" a simple engineered system from a prokaryotic to a eukaryotic organism so that the system can function in both organisms.

1 Overall Objective & Specific Aims

My goal is to enable the rational engineering of biological systems inside cells. Here, I propose to study and optimize how engineered biological systems interface with their host cells. The host cell acts as a chassis that provides transcription, translation and other processes to the engineered system. Currently, engineered system performance is unpredictable because of the variation introduced by the chassis. Engineered system performance can be made more predictable by reducing the number of interactions between the system and the chassis. Improved predictability depends on the remaining interactions being specified. Prediction would be easier if those remaining interactions were unchanging. If these interactions are also standardized, then different engineered systems can be "plugged" into different chassis more reliably.

To achieve my overall objective I will:

1. **Integrate dedicated transcription and translation systems into a chassis**

   I will construct systems that synthesize dedicated transcription and translation machinery that is independent of the equivalent host cell machinery and thus less sensitive to host cell state. In addition, I will make the synthesis of the dedicated machinery more independent of the host cell.

2. **Develop test systems & measurements to evaluate chassis response to system demands**

   I will design and build a suite of test systems that can be used to place demands on specific host cell systems. In parallel, I will investigate the principle measurements that must be made to describe the chassis response to the demands of the test systems. The test systems and measurements will be suitable to characterize a range of chassis.

3. **Evaluate the ability of a wild-type chassis and a dedicated system containing chassis to supply an engineered system**

   I will use the test systems and response assays to measure the response of a chassis to applied demands. I will do this for a wild-type chassis and for a chassis containing the dedicated transcription and translation systems. Furthermore, I will compare the performance of an engineered system in both chassis.

4. **Investigate "porting" an engineered system between two chassis**

   To examine the portability of engineered systems, I will investigate three different approaches to porting an engineered system between *E. coli* and *S. cerevisiae* so that the system functions in both chassis.

2 Background & Significance

Engineering biological systems: past and future

It is becoming possible to engineer simple multi-component systems in living organisms based on transcriptional logic [Elowitz 00, Gardner 00, Hooshangi 05]. The ability to engineer biological systems offers clear benefits to society through advances in medicine and biotechnology. While today's engineered biological systems hint at a future ability to design and build complex systems with many components, the engineering of functional systems is still difficult and time consuming, more akin to art than engineering. Furthermore, current engineered systems are highly sensitive to host physiology and environmental conditions [Elowitz 02, Rosenfeld 05].

I believe the future engineering of biological systems will be greatly facilitated by adopting some of the concepts that have proved useful in other engineering disciplines. Central among these concepts are the ideas of standardization of components (e.g. http://parts.mit.edu) and abstraction, which lead to the concept of modularity. Currently, engineered biological systems are dependent on natural host cells. Constructing modular systems is made difficult by the complexity of the host cells and the numerous interactions between the host cell and the engineered system. Because of the complex interaction, the entire host cell must be considered when engineering a system. The development of engineered systems would be accelerated if system engineers did not have to consider all the details of the host cell. Modularization can be achieved by making the interactions between the engineered system and the host cell simpler and standardized.

The chassis/system interface

Engineered biological systems typically rely on the host cell for the processes of replication, transcription, translation and degradation and the requisite energy and materials to power those processes. In this way, the cell acts as a power supply and chassis that insulates and drives the system [Knight, T.F. Jr., personal communication]. Hereafter, I will use the word chassis to refer to the chassis and the power supply combined. In an ideal chassis/system relationship, perturbations in the external environment or the chassis should not be transmitted to the system. Conversely, changes in the function of the system should not affect the function of the chassis. However, today's engineered biological systems use the same cellular processes and resources as the chassis itself, leading to a coupling of the behavior of the chassis and the system. Consequently, the system experiences much of the same noise and variation that the chassis inherently experiences. In addition, by sharing chassis resources the system affects the function of the chassis. If engineered systems could by supplied from different resource pools than those used by the chassis, then system performance would be more independent of the state of the chassis. In addition, if the chassis supply to the engineered system could be standardized, it would be possible to "plug" a system

into a well chosen chassis and expect the system to work. Thus, with a specified and standardized interface, system design should become independent of chassis design. This idea is shown schematically below where the existing complex interface between the chassis and the engineered system is contrasted with a simplified and standardized interface.

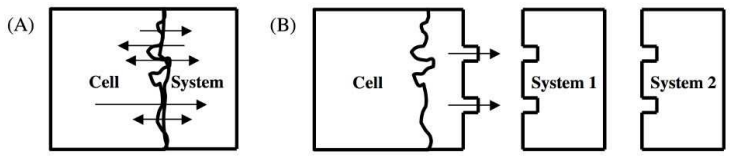


Figure 1: Designing a standard interface between the chassis and the system allows for improved predictability of performance and interchangeability of systems. (A) shows the current approach wherein the chassis/system interface is complex, making it difficult to reliably match system to chassis. (B) shows the improved case where a simple, standard interface exists, making it easy to power different systems with a general chassis.

Past work on the chassis/system interface

Transcription
Methods for recombinant gene expression have been reviewed recently [Sorensen 05,Baneyx 99]. *E. coli* is the most commonly used chassis for protein expression [Schmidt 04] and one of the most common expression systems is the T7 system (using the pET series of plasmids) developed by Studier and coworkers [Dubendorff 91a, Studier 90]. The system consists of the T7 consensus promoter that is not recognized by *E. coli* RNA polymerase (RNAP) but only by the dedicated, highly processive RNAP from bacteriophage T7. Because of the high specificity of the T7 promoter, highly toxic proteins can be cloned in *E. coli* in the absence of T7 RNAP while in the presence of T7 RNAP there is a high level of expression of the target gene. Expression levels can be varied over a wide range by modulating T7 RNAP activity via T7 lysozyme or by repressing the T7 promoter [Zhang 97, Dubendorff 91a]. The expression system requires a chassis that produces the T7 RNAP, such as *E. coli* BL21(DE3). The T7 system has been made flexible with a range of plasmids with varying promoter strengths, ribosome binding site (RBS) strengths and restriction sites to suit a range of applications.

The P_{BAD} expression system is also used for recombinant protein expression [Guzman 95]. The P_{BAD} system consists of plasmid borne elements of the araBAD operon and permits the graded induction of the P_{BAD} promoter. It is important to note that this graded induction is at a population level only - individual chassis have an all or none induction. The all or none induction is due to positive feedback in the regulation network of the araBAD operon. Linear induction of the P_{BAD} promoter in individual chassis was achieved by removing the positive regulation controlling the transporter gene that was responsible for the positive feedback [Khlebnikov 02]. The P_{BAD} system, like the Tac promoter system [pGex, Amersham

Biosciences] and the λPL system [pLEX, www.invitrogen.com], all differ from the T7 system in that they rely on *E. coli*'s own transcription system.

Translation

Most recombinant gene expression systems utilize the chassis' natural translation system. However, Hui and coworkers have developed a dedicated translation system [Brink 95,Hui 87]. Ribosomes recognize the RBS via a complementary RNA sequence on the 16s rRNA subunit termed the anti Shine-Dalgarno (ASD) sequence [Shine 74]. By using a mutated ASD (5'-CACAC-3'), Hui and coworkers produced a sub-population of ribosomes that did not significantly translate host cell mRNA. 80% of the protein expressed by this sub-population was encoded by mRNA containing an RBS complementary to the mutated ASD. The dedicated ribosomes accumulated in large numbers (∼60% of total ribosomes) and decreased the numbers of wild-type ribosomes by ∼18% but did not significantly affect chassis growth unless actively translating a reporter gene. Also, it has been shown that translation can initiate from a mutant initiation codon [Varshney 90]. A mutant initiator tRNA is required for translation of mRNA containing the mutant initiation codon. It is believed that overexpression of the mutant initiator tRNA will bind to a fraction of the chassis ribosomes, again creating a sub-population of dedicated ribosomes [RajBhandary, U.L., personal communication]. Using a mutant initiation codon rather than a mutant RBS is an alternative way to selectively translate mRNA. It should be possible to combine the mutant initiation codon with the mutated RBS of Brink and coworkers to allow highly specific translation of particular mRNAs.

Genetic engineering of chassis

Additional steps have also been taken to improve a chassis' ability to supply an engineered system. For example, the removal of proteases such as ompT and lon from BL21 decrease protein degradation rates. Deletion of the recombinase encoded by *recA* improves the stability of genetic information in *E. coli* strain DH5α. To assist the folding of highly expressed proteins, plasmids have been used to overexpress molecular chaperones [Ikura 02]. A range of different chassis exhibiting the above features have been engineered for use with the T7 expression system. Plasmids have been used to supply tRNAs necessary for recombinant protein production but that are rare in the chassis as a shortage of any charged tRNA can lead to the stringent response in *E. coli* [Dieci 00, Neidhardt 63]. Another approach to improving protein yield involves only producing protein once the chassis have reached a critical density so that continued growth is not required [Studier 05].

Chassis response to recombinant gene expression

Chassis responses to recombinant gene expression hamper the operation of the expression systems. Bacteria have a range of responses to an imposed stress, most notably a modulation of growth rate. Other responses include the heat shock response, the envelope stress response, the cold shock response, the SOS response, the stringent response and the general stress response [Gill 00, Wick 04]. However, these are not independent responses but rather different groupings of common individual responses. These responses are triggered to cope with stresses such as difficulty in folding proteins, resource starvation, unfavorable

extracellular pH, etc. The triggering of these responses by high levels of recombinant gene expression should not be surprising as bacteria are presumably not naturally optimized to supply large demands from a foreign gene expression system [Hoffmann 04]. These responses lead to a range of disadvantageous changes in chassis physiology including plasmid instability, increased cell lysis and increased genetic variation [Gill 00]. For example, a shortage of charged tRNAs invokes the stringent response, leading to a large reduction in the rate of ribosome synthesis and a change in maintenance of plasmids using the ColE1 origin of replication among other effects [Baracchini 88, Herman 95].

Currently, there are qualitative rather than quantitative guidelines on how to best express a recombinant protein in a chassis [Glick 95]. These qualitative guidelines are adequate to allow high expression of a recombinant protein once a successful combination of expression system and chassis has been found. However, an engineer who wishes to design a multi-component system with defined kinetic and steady state behavior would benefit from more quantitative information describing the response of a chassis to an engineered system [Basu 04,Elowitz 00]. For example, it would be useful for a system engineer to know that a system can take a specified number of ribosomes from the free ribosome population of the chassis before any chassis stress responses are triggered.

Defining an engineered system and a chassis

Engineered System Definition
We can think of an engineered system as a regulated set of genes, mRNAs, proteins, and other molecules. An engineered system has a set of demands that must be supplied by the chassis in order to function. While every engineered system is likely to demand replication, transcription, translation and degradation processes from its chassis, what will vary from system to system is the quantity of each demand. The demands of a system can be grouped into *materials*, *machinery*, and *energy* demands. For example, the translation demand of a system will be composed of a materials demand for amino acids per second, a machinery demand for numbers of ribosomes (and numerous other proteins), and an energy demand for GTP molecules per second. The different demands of an engineered system depend on two sets of parameters, chassis-specific parameters and system-specific parameters. System-specific parameters include promoter strength, transcript length, RBS strength etc. A system designer should specify the system-specific parameters and hence, the total demand of the engineered system as a function of chassis-specific parameters.

The specific sequence of the coding regions in an engineered system can affect gene expression [Bonomo 05]. The sequence of a coding region determines the frequency of amino acid usage, the stability of the coding region's mRNA, the presence of cryptic regulatory elements, and protein solubility. For the predictable design of engineered systems, such effects are undesirable. In this project I will attempt to ignore such sequence specific factors. It will be the responsibility of the system engineer to ensure that the individual parts in the system function correctly. For the purposes of this project, I will only use engineered systems

containing parts that are well tested and understood.

Systems should be defined by a set of performance characteristics and the demands placed on the chassis by the system. This information might take the form of a system data sheet, an example of which is shown in Section 6 for an Acyl-Homoserine Lactone detecting system [BBa_F2620, http://parts.mit.edu]. However, the behavior of BBa_F2620 is dependent on the specific chassis used and the precise experimental conditions under which it is tested. This is an example of how the definition of reusable, reliable systems is difficult.

Chassis Definition
The chassis is a host cell that is capable of supplying the demands of an engineered system. The chassis uses resources from its environment to supply the engineered system while simultaneously insulating the engineered system from the environment. The chassis preserves and replicates the genetic information of the system, produces the system's messages, proteins, and degrades all of the system's species. Therefore, it must be capable of supplying each of the individual demands of the system described above.

A chassis should be defined by a set of performance characteristics and its ability to supply system demands. These specifications could be listed in a chassis data sheet, an illustrative example of which is shown in Section 6 for *E. coli* strain MG1655. If the chassis can be engineered to stably supply the demands of a system without unwanted interactions then it will be easier for engineers to predict and specify the behavior of systems.

What I have described above is a two-way flow of information and materials across the interface between the chassis and the engineered system. It is beyond the scope of this project to consider every flow across the interface. Instead, I will focus on some of the most important flows that determine the behavior of the engineered system and the chassis. The relationship between a chassis and an engineered system is shown schematically in Figure 2.

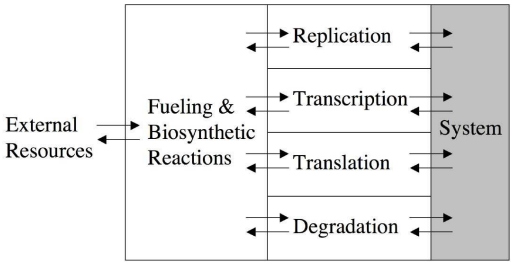


Figure 2: The relationship between a chassis and an engineered system. The chassis is unshaded.

3 Research Design & Methods

3.1 Integrate dedicated transcription and translation systems into a chassis

3.1.1 Design and build VM1.0

To attempt to simplify the interaction between the chassis and the system, I will build transcription and translation systems dedicated to serving engineered systems. Dedicated transcription and translation systems have already been developed individually by Studier, Brink and coworkers. The primary purpose of dedicated systems is to decouple the demands of the engineered system from the demands of the chassis. A secondary reason to build dedicated systems is to construct, by analogy to the JAVA programming language, a "virtual machine". Just as a JAVA virtual machine allows a JAVA program to run on any computer architecture, so a biological virtual machine might allow a system to work predictably in any biological chassis that carries a virtual machine. For example, an engineered system in two different bacterial species will behave differently due to the genetic differences between the bacteria. With dedicated transcription and translation systems forming the basis of a virtual machine in both bacteria, the same system could be expected to behave similarly in both chassis. Ideally, the engineered system should not compete for the components of systems used by the chassis itself. As a first step towards this end, the dedicated systems proposed here will separate the machinery demands of the system and the chassis but not the materials and energy demands.

I will begin with the simplest implementation of dedicated transcription and translation systems. For transcription, this means controlling transcription with a T7 promoter in a BL21(DE3) *E. coli* chassis. For translation, I will start with the Brink specialized rrnB operon encoded on a plasmid. In addition to having both systems separate, they will be combined to form the first version of a virtual machine, VM1.0. An engineered system that can be supplied by the VM1.0 will not compete with chassis systems for ribosomes or RNAP. The current status of this work is described in the Preliminary Results section of this document.

3.1.2 Design and build VM2.0

In parallel to testing this simple virtual machine, I will design and construct a second generation virtual machine, VM2.0. VM2.0 will incorporate more developed versions of both the dedicated transcription and translation systems. The objectives of VM2.0 will be to reduce and control the demand that the virtual machine places on the chassis and secondly, to further decouple system function from chassis function. A number of the features of VM2.0 are described below.

In VM1.0, chassis RNAP is required to produce the T7 RNAP, meaning the dedicated

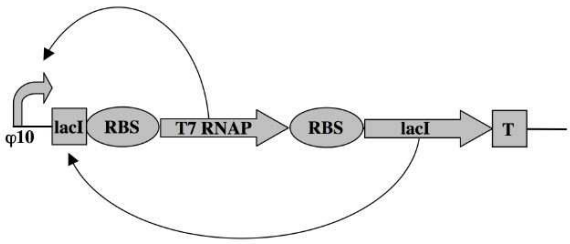


Figure 3: Arrangement of genetic elements to produce a T7 RNAP autogene. Transcription of its own gene by T7 RNAP leads to positive feedback while the production of LacI inhibits transcription from the T7 promoter. Constitutive expression of T7 lysozyme will inhibit T7 RNAP activity.

transcription system is subject to variation in the chassis transcription system. However, by placing a T7 RNAP gene under the control of a T7 promoter (creating a T7 autogene), the transcription of the dedicated T7 RNAP is only dependent on the availability of chassis transcription materials, not machinery [Dubendorff 91b]. Any non-specific transcription of the gene by *E. coli* RNAP will lead to positive feedback and high levels of transcription of the T7 RNAP gene. The high levels of T7 RNAP produced by this positive feedback can be lowered through the use of negative feedback and inhibition of transcription. Negative feedback can be achieved by placing a LacI DNA binding site downstream of the T7 promoter and placing the *lacI* gene under the control of a T7 promoter [Dubendorff 91a]. Inhibition of T7 RNAP activity can be achieved by constitutively producing the T7 lysozyme from a plasmid such as pLysE or pLysS [Dubendorff 91b]. This regulatory network is shown schematically in Figure 3. Should the levels of T7 RNAP prove to be too high despite the negative regulation, a weaker T7 promoter could be used instead, such as ϕ3.8.

The plasmid encoded rrnB operon of the first generation dedicated translation system places a heavy demand on the chassis as evidenced by the large reduction in growth rate described in the Preliminary Results Section. I will alleviate some of this demand by integrating the rrn operon encoding the dedicated rRNA into the chromosome. Efficient methods for integrating genetic information into the *E. coli* chromosome have been described previously [Yu 00].

I will place the dedicated rrn operon under the control of a T7 promoter. By doing so, the r-proteins will be the only components of the dedicated systems that require *E. coli* transcription machinery. I will use the RBS recognized by the dedicated translation machinery to control the translation of T7 RNAP. By doing so, the dedicated transcription system will be entirely synthesized by dedicated machinery. The regulatory network resulting from the proposed changes is shown in Figure 4.

I will investigate other means of producing dedicated ribosomes. RajBhandary and coworkers have shown that overexpression of a mutant initiator tRNA can produce a population of ribosomes that only initiate translation of transcripts containing a mutant initiation codon [Varshney 90]. I will investigate using this as an alternative or complementary means to produce dedicated ribosomes. These mutant tRNAs could also be used as a means to

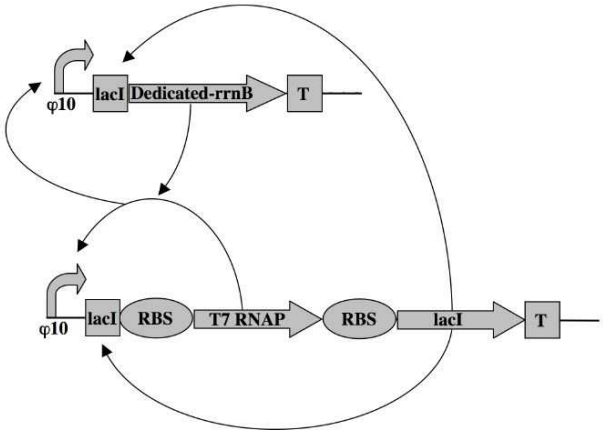


Figure 4: The T7 RNAP autogene arrangement shown previously but with added control of the rrnB operon. The dedicated transcription machinery is translated by the dedicated ribosomes.

selectively turn off the translation of an engineered system. This is a useful capability for examining system demands and for decoupling the translation of different systems.

3.2 Develop test systems & measurements to evaluate chassis response to system demands

I will design and implement ways to place a certain demand on the chassis and ways to measure the response to that demand. These methods will not be specific to a given chassis but can be used to characterize a range of potential chassis.

3.2.1 Design test systems with specific demands

Interactions between the chassis and an engineered system can be tested by placing clearly specified demands on the chassis. I will design and build a suite of test systems that will allow me to independently increment some of the demands placed on the chassis. Some of these test systems use the host cell's transcription and translation systems and some require the dedicated systems described in the previous subsection. The individual demands of an engineered system are coupled. For example, the number of actively transcribing RNAP determines the number of nucleotides polymerized per unit time, in other words, the transcription machinery demand determines the materials demand. Despite this coupling of demands, if single demands can be independently placed on the chassis, the effect of demands on the chassis can be probed in a more detailed way.

I will build the following systems to examine transcription and translation demand. It is not yet clear that this is an adequate list of systems to successfully probe the range of demands, I will continue to examine others. In Specific Aim 3, I will detail my proposal to test chassis

response to these system demands.

Total Transcription Demand

I will build a family of promoters of different strengths controlling untranslated transcripts of two different lengths. These test systems will place varying total transcription demands on the chassis. I will use a subset of a library of promoters constructed by Moyle and coworkers [Moyle 91]. The library is generated from 36 single-base pair mutations in the -35 and -10 hexamers of the *ant* promoter of P22. LacZ fusions were used to measure the relative activity of the promoter mutants in *S. typhimurium*. Should this method of varying transcription demand prove unsuccessful, transcription demand could be varied by using the P_{BAD} linear induction system [Khlebnikov 02]. I believe using a family of promoters is preferable to using linear induction as it removes the need to produce a repressor or consider non-specific effects of the inducer on the chassis. I will measure the rate of accumulation of the reporter RNA via Northern blotting. This data specifies the number of RNAP actively transcribing the reporter RNA and hence, the total transcription demand. To ensure that the test systems only place a transcription demand on the chassis, there must be negligible degradation of the reporter RNA, i.e. the test system must not place a significant demand on the chassis degradation systems. For this reason, a tRNA is an ideal choice for the reporter RNA [Lopez 94]. Total transcription demand can be varied either by increasing promoter strength, and hence loading more RNAP onto the reporter coding region, or by making the coding region longer. I will use both promoter strength and coding region length as independent ways to vary the total transcription demand of an engineered system. I expect that the total transcription demand should depend only on the number of RNAP engaged in transcription, not the way the demand is generated.

Total Translation Demand

This can be varied with a set of test systems that produce a single protein but that have RBSs of varying strength. Such a set of RBSs has been studied by Ringquist and coworkers [Ringquist 92]. I will determine the rate of accumulation of reporter protein via quantitative Western blotting and hence, specify the total translation demand. An alternative to using Western blotting might be measure the numbers of ribosomes on a single mRNA via an RNA gel shift assay [Plambeck 03]. Similar to the transcription case above, a highly stable protein (e.g GFP) should be used to ensure that the test system does not place a significant demand on the chassis degradation system. It will be necessary to take into account the fact that varying RBS strength will affect mRNA stability [Yarchuk 92].

The dedicated transcription and translation systems described above will allow me to vary the material and energy demands on the chassis independent of the machinery demand.

Material Demand of Transcription

The material demand of transcription can be varied independent of the machinery demand if a test system does not use the same transcription machinery as the chassis' own machinery. This can be achieved by using the T7 expression system of Studier and coworkers. I will build a family of T7 promoters of different strengths controlling an untranslated transcript. The relative activities of all single-base mutation variants of the T7 consensus promoter sequence

are known and I will use a subset of these promoters in the test systems [Imburgio 00]. The demand can be measured in a similar way to that outlined above to measure the total transcription demand.

Material and Energy Demand of Translation
Similar to above, the materials and energy demand of translation can be varied independent of the demand for translation machinery by using the dedicated translation system described in Specific Aim 1. I will use a set of test systems that produce a single stable protein but that have RBSs of varying strength. Such a set of RBSs compatible with the dedicated translation system does not exist. Hence, I will need to design and build a small set of RBSs of varying strength. Measurement of the demand these test systems place on the chassis will be similar to that described above for total translation demand.

3.2.2 Develop techniques to measure chassis response to system demands

Once I have ways to place specific demands on the chassis, I must also implement techniques to measure the chassis response to a given demand. The simplest way to measure response to an imposed demand is to measure the change in chassis growth rate. I can measure growth curves for large numbers of cultures in parallel with time resolution on the order of seconds using an automated multi-well fluorimeter with photometric capabilities [http://www.perkinelmer.com]. Hence, it is straightforward for me to measure growth rates for each engineered system and chassis of interest. I have confirmed that the growth characteristics of chassis in the plate reader are similar to those obtained by batch culture in an aerated flask.

The triggering of chassis stress responses as a result of a high engineered system demand leads to the upregulation of a wide range of genes. Some of the key genes upregulated during these responses are *lon*, *recA*, *rpoS*, *spoT*, *relA*, and *groEl*. I will choose a small subset of these genes and implement a practical strategy to measure the levels of those genes. The most common and reliable techniques for measuring the expression of numerous genes in the chassis are Northern-blotting, real time PCR, and RNA microarrays. These methods differ in cost, throughput and accuracy. The best measurement technique to choose will be partly dependent on which genes I decide to measure.

The use of fluorescent reporters fused to proteins and promoters is an easy way to report on the relative expression levels of key genes. By calibrating fluorimetry results against absolute measurements from Northern and Western blotting it should be possible to estimate numbers of molecules using fluorescent fusions to proteins of interest. Fluorescent proteins have been fused to the β' subunit of *E. coli* RNAP [Cabrera 03b, Cabrera 03a]. Similarly, YFP has been fused to a stably bound ribosomal protein (L25) in *S. cerevisiae* [Hurt 99]. However, care must be taken when designing protein fusions. For example, the RNAP fusion in *E. coli* leads to instability of the RNAP and growth defects at 37^{o}C. By contrast, the ribosomal fusion in *S. cerevisiae* did not lead to growth defects.

To fully monitor chassis response to demand it is also necessary to measure the levels of

certain small molecules. Perhaps the most important of these are the stringent response alarmones, ppGpp and pppGpp, together known as (p)ppGpp [Dennis 04]. These nucleotide concentrations have been measured via HPLC [Zhang 02]. Previous research has shown that the method of purification of (p)ppGpp and other nucleotides from host cell extract can significantly alter such measurements [Dennis 04]. Hence, the purification technique will warrant careful consideration.

Comprehensively measuring all physiological changes in a chassis in response to a demand would be a very significant undertaking. Here, I propose to implement existing techniques to enable me to measure certain key reporters of chassis physiology. One important question to be answered here is whether it is feasible to measure these reporters when testing a wide range of engineered systems in multiple chassis.

1. **RNAP and Ribosome Levels -** I will repeat the CFP fusion to the *E. coli* RNAP. In addition I will attempt to fuse a fluorescent reporter to an *E. coli* ribosomal protein. L23, the *E. coli* homolog of L25 in *S. cerevisiae*, appears to be a good candidate. Measuring the chassis fluorescence via flow cytometry will report the relative levels and variation of ribosomes and RNAPs. I will investigate the possibility of fusing a fluorescent protein to T7 RNAP or fusing a fluorescent reporter gene to the promoter controlling transcription of the T7 RNAP gene.

2. **(p)ppGpp -** I will investigate the feasibility of measuring (p)ppGpp levels as engineered system demand varies. I will also examine a technically simpler alternative; measuring the expression of the two (p)ppGpp synthetases. These are the protein products of the *relA* and *spoT* genes. The caveat to measuring *spoT* expression is that the enzyme encoded by this gene has varying (p)ppGpp 3'-pyrophosphohydrolase activity as well as (p)ppGpp synthetase activity [Xiao 91]. This means that the expression level of *spoT* does not necessarily correlate with (p)ppGpp levels.

3. **Stress Response Genes-** I will implement a combination of mRNA measurements and fluorescent fusions as a way to measure the upregulation of a small subset of chassis response genes including *lon*, *recA*, *rpoS* and *groEl* [Gill 00].

3.3 Evaluate the ability of a wild-type chassis and a dedicated system containing chassis to supply an engineered system

In this section of the project I hope to show that dedicated transcription and translation systems can supply engineered systems more reliably than wild type transcription and translation systems. If this is the case, dedicated transcription and translation systems will form the basis of a standard interface between chassis and engineered systems. If dedicated systems do not become a standard and practical way to supply engineered systems in the time frame of this project, I hope to at least have better elucidated how best to interface engineered systems with chassis.

3.3.1 Measure chassis response to varying engineered system demands

I will use the test systems and metrics developed in Specific Aim 2 to measure chassis response to a range of system demands. Two types of chassis will be considered, a wild-type *E. coli* chassis and an *E. coli* chassis containing a virtual machine. The goal of these experiments is to produce transfer curves that describe how the chassis response varies with increasing individual system demands. Illustrative examples of how these transfer curves may appear is shown in Figure 5. This data will characterize the ability of the chassis to supply a given level of demand and elucidate what characteristics should be measured to specify future potential chassis. These experiments should also give insight into whether one demand type is significantly more costly for the chassis to supply than others. For example, it has been hypothesized that a chassis is less able to supply a high translation demand than a high transcription demand due to the energetic costs of producing ribosomes, charging tRNAs, etc. [Vind 93]. Confirmation of this hypothesis would direct me to focus future work on optimizing the translation interface between chassis and system. The dedicated transcription and translation systems themselves place a demand on their chassis. I will measure the level of this demand and evaluate whether it significantly impacts chassis performance. These measurements can then be integrated into a specification of a chassis. This specification might be summarized in a chassis data sheet as discussed in the Background & Significance Section.

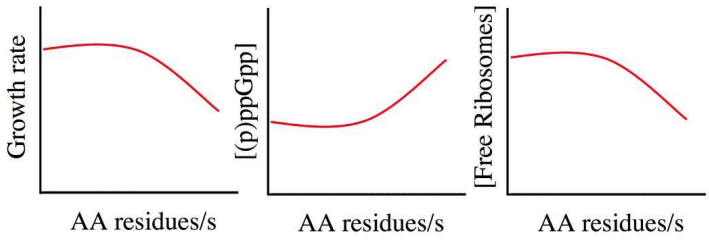


Figure 5: Illustrative transfer curves showing how key chassis parameters might vary with an increasing system demand, in this case that demand is for translation materials (AA = Amino Acids).

3.3.2 Measure the performance of a simple engineered system in different chassis

I will measure the ability of a wild-type *E. coli* chassis and an *E. coli* chassis containing a virtual machine to power a simple GFP expression system. This will help me to evaluate the utility of the dedicated systems in comparison to a wild-type *E. coli* chassis. Testing will consist of measuring GFP expression at both a population level using a multi-well fluorimeter and at a single chassis level using flow cytometry. I will examine the levels and variation of

engineered system species. I expect the levels of system mRNA and protein to be largely dependent on the numbers of dedicated machinery in the system. Since the dedicated systems will produce lower levels of free ribosomes and RNAPs to power the system, the extrinsic noise in the expression of the fluorescent protein may increase [Swain 02].

3.4 Investigate "porting" an engineered system between two chassis

As discussed in the Background & Significance Section, the engineering of biological systems would be made easier if systems could easily be moved, or "ported" from one chassis to another. I discuss three approaches to allow the porting of engineered systems between chassis below. I propose to investigate the practicality of the three approaches and pursue whichever offers the most utility. Specifically, I will investigate the practicality of porting engineered systems between *S. cerevisiae* and *E. coli*.

The first approach to porting a system between two chassis is to ensure that all the genetic elements that comprise the system function in both chassis. For example, every promoter and RBS in the engineered system should be recognized by transcription and translation machinery in each chassis. If the transcription and translation machinery that recognizes the system are dedicated to that system, then a standard interface, or virtual machine, has been constructed in each chassis. To allow the same genetic elements to be recognized in two different chassis, either identical transcription and translation machinery must exist in both chassis or the genetic elements must be recognizable by the different transcription and translation machinery in both chassis. This might be difficult to achieve if one chassis was prokaryotic and one chassis was eukaryotic, as the sequences of promoters, RBSs, and other genetic elements recognized by prokaryotic and eukaryotic chassis are significantly different. However, there is evidence to suggest that it is possible in some instances. For example, it has previously been shown that T3 and T7 phage RNAP can be used to transcribe genes in both prokaryotes and eukaryotes [Deuschle 89, Fuerst 86, He 95]. This shows that the dedicated transcription systems described in Specific Aim 1 can be ported between prokaryotic and eukaryotic chassis. In addition, there is some evidence that ribosomal components can be transferred between *S. cerevisiae* and *E. coli* as evidenced by the work of Asai and coworkers who transferred the GTPase domain of the *S. cerevisiae* ribosome into the *E. coli* ribosome [Asai 99].

The second approach to porting an engineered system between two chassis is to ensure that the same piece of DNA should function in both chassis. This relaxes the requirement in the previous approach that every genetic element should function in both chassis. In this approach it is not necessary to build a virtual machine in each chassis. As an example of this approach, a system engineer could use appropriate translation start signals for both chassis in the one engineered system, to form a composite translation start signal. This was the approach taken by He and coworkers who constructed a translation start signal incorporating an EMCV-IRES (internal ribosome entry site from encephalomyocarditis virus) and a

prokaryotic Shine-Dalgarno sequence that cause translation to initiate at the same codon in *E. coli* and *S. cerevisiae*. This approach does not require the same machinery to be present in both chassis, which should make it easier than the first approach. However, it is important to ensure that elements of the engineered system used in one chassis do not have unwanted function in the other chassis.

The third approach to porting systems between chassis dispenses with the definition of a system as a physical piece of DNA and instead defines a system in terms of its function. For example, in this approach an oscillator in two different chassis is the same system if its function is the same even if the DNA encoding it is very different in the two chassis. In this approach, a system engineer should design the function of the system and a computer program should then specify the DNA encoding that function that will work in the chosen chassis. For example, to implement repression of transcription, the software would need a database of regulatable promoters and the cognate repressors for those promoters. If a system engineer wished to design a system encoding repression of transcription the software would choose a repressor/promoter pair known to function correctly in the desired chassis.

The first approach is presumably the most difficult, however, it is the only approach that truly delivers a standardized and insulated interface between chassis and engineered system. As such, it is worthwhile to pursue this approach if possible. Once I have constructed a functioning virtual machine in *E. coli*, I will begin to investigate the equivalent virtual machine in *S. cerevisiae*. In addition, I will investigate the feasibility of implementing the second and third approaches.

4 Preliminary Results

Experimental Outline
To date, I have focused on the dedicated transcription and translation systems. In particular, I have acquired the necessary strains and plasmids to assemble the first generation virtual machine. I am constructing reporter systems to ensure the dedicated systems are functional before undertaking the more rigorous testing described in Specific Aim 3. The systems, strains, and experimental conditions used are described below along with the preliminary results that have been obtained.

Systems
pCH1497 encodes a plasmid borne copy of the rrnB operon containing a 16s rRNA gene with the mutated ASD used by Brink and coworkers [Hayes, C, personal communication]. The rrn operon is under the control of an inducible P_{BAD} promoter on a pACYC derived plasmid (p15A *ori*, 10-15 copies per chassis, tetracycline resistance marker).

I obtained a reporter system (pSD1-rbsK2) to test the function of the dedicated translation machinery [McGinness, K, personal communication]. It contains the *rbsK* gene under the control of an IPTG inducible promoter. pSD1-rbsK2 has previously been shown to produce mRNA that are successfully translated by ribosomes containing rRNA from pCH1497.

I constructed a GFP reporter system using an RBS (5'-GTGTG'3') complementary to the mutated ASD from standard biological parts [BBa_I7102, http://parts.mit.edu]. BBa_I7102 combines BBa_R0040, BBa_B0036, BBa_E0040, and BBa_B0015; a tet repressible promoter, RBS, GFP orf, and a transcriptional terminator respectively. BBa_I7102 is encoded on a standardized plasmid [pSB4A3, http://parts.mit.edu]. pSB4A3 contains a pSC101 *ori* and ampicillin resistance. BBa_I7102 is the simple GFP expression system that can be used to measure the ability of a chassis containing dedicated translation machinery to supply an engineered system as described in Specific Aim 3.

BBa_I7101 is identical to BBa_I7102 but contains a wild-type RBS. BBa_I7101 was also placed on pSB4A3. This system tests the ability of a wild-type chassis, lacking dedicated machinery, to supply an engineered system. Plasmid copy numbers, mRNA numbers and protein numbers have already been measured for this system in our lab [Braff, J.C., Conboy, C.M., personal communication].

I have built a third GFP reporter system (BBa_I7103) using the consensus T7 promoter, the mutant RBS, GFP coding region and a T7 transcriptional terminator (Tϕ) (parts BBa_R0085, BBa_B0036, BBa_E0040, and BBa_B0016). BBa_I7103 is also encoded on pSB4A3. This reporter system produces mRNA and GFP when transcribed by the dedicated transcription and translation machinery. As such, it will be used to test the ability of VM1.0 to supply an engineered system. It should be noted that Tϕ (BBa_B0016) has an efficiency of approximately 80%, meaning that the average transcript length produced by

this system may be significantly longer than expected. I have not yet tested this plasmid.

Strains
E. coli strain MG1655 was used for all experiments performed to date. Since MG1655 contains a full araBAD operon, induction with arabinose leads to an all or none induction as described in the Background & Significance section. When I begin testing the dedicated transcription system and VM1.0, I will use BL21(DE3). It may be necessary to use a pLysS or pLysE plasmid as described by Studier and coworkers to inhibit the high activity of T7 RNAP in BL21(DE3).

Experimental Conditions
I am using Neidhardt rich defined media for all cultures [Neidhardt 74]. Cultures are grown overnight at 37^{o}C, diluted 1/1000 and grown in a 96-well plate using a Victor3 fluorimeter [http://www.perkinelmer.com]. Optical Density (OD) at 600nm is recorded at regular intervals. Fluorescence emission at 535nm in response to excitation at 488nm is used as a relative measure of GFP levels. Arabinose is used to induce the P_{BAD} promoter on pCH1497 at concentrations up to 0.02% w/w.

Results
To confirm plasmid pCH1497 produced rRNA that was incorporated into functional, dedicated ribosomes in MG1655, I cotransformed pCH1497 and pSD1-rbsK2 into MG1655. Cultures of MG1655 and MG1655 containing both pCH1497 and pSD1-rbsK2 in the presence and absence of arabinose and IPTG were grown. Host cells were lysed and the protein was run out on a polyacrlyamide gel (Figure 6) and stained with Coomassie Blue. Only when transcription was induced from pSD1-rbsK2 and pCH1497 (lane 5) was a clear band seen where the protein product of *rbsK* should be. This confirmed that the mutant RBS was being specifically translated by the dedicated ribosomes. Results from the gel also suggest that there was some basal level of production of RbsK without induction of pCH1497 where trace levels of the reporter protein were seen (Lane 6). I attribute the trace levels of RbsK to non-specific translation by wild-type ribosomes from the mutant RBS rather than leaky expression from the P_{BAD} promoter controlling the rrn operon. This is because the P_{BAD} promoter is known to have a very tight "off" state. Repeating the experiment with MG1655 containing the reporter plasmid pSD1-rbsK2 but not pCH1497 should confirm that this is the case. Approximately equal numbers of lysed cells were used for each lane.

I next measured the growth characteristics of chassis containing the reporter systems (Figure 7). It is clear that the GFP reporter using *E. coli* translation machinery (BBa_I7101) grows at a very similar rate to untransformed *E. coli*. MG1655 containing BBa_I7102 but lacking pCH1497 also grows at a similar rate. The addition of pCH1497 to MG1655 with BBa_I7102 greatly slows chassis growth even without inducing the dedicated translation machinery. Once the OD of MG1655 rises above background, it has a minimum doubling time of 30mins. MG1655 containing pCH1497 and BBa_I7102 but with no induction of pCH1497 has a minimum doubling time of 75mins. I attribute the differences in chassis growth rates to the demand of replicating two plasmids and producing two antibiotic resistances. Tetra-

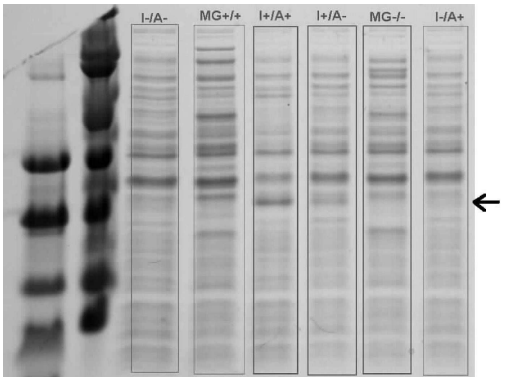


Figure 6: Polyacrylamide gel confirming dedicated ribosomes selectively translated the mutant RBS. I+/- indicates the presence/absence of IPTG, A+/- indicates the presence/absence of arabinose. The arrow marks the expected location of RbsK. A band is evident at this location in lane 5 where *rbsK* and the dedicated rrn operon are being transcribed. Some basal expression of RbsK can be seen in lane 6 where the *rbsK* mRNA is being produced but there is no induction of the rrn operon.

cycline resistance, in particular, is known to significantly affect chassis growth rate. When the P_{BAD} promoter is induced, OD barely rises above that of the background. I attribute the apparent lack of growth over the course of this experiment to the positive feedback in the araBAD operon and the relatively high arabinose concentration. This presumably leads to maximal induction of the P_{BAD} promoter and may lead to toxic levels of intracellular arabinose.

GFP measurements were also taken (Figure 7). This data is consistent with expectations based on the growth data shown in Figure 4. BBa_I7101 exhibits high levels of fluorescence and BBa_I7102 without pCH1497 exhibits fluorescence levels similar to those of untransformed MG1655. This suggests that there is negligible non-specific translation of this reporter system. The negligible fluorescence shown for BBa_I7102 with pCH1497 induced is due to the fact that the chassis grow slowly and does not necessarily indicate that GFP is not being produced. Indeed, it appears that the GFP level as a function of OD increases for BBa_I7102 in the presence of dedicated ribosomes at a rate similar to that of BBa_I7101 (data not shown). I am currently attempting to better confirm this observation.

These experiments showed the growth rate of chassis with dedicated translation machinery needs to be improved. I experimented with differing levels of arabinose induction on MG1655 transformed with pCH1497. Growth curves for different arabinose concentrations are shown in Figure 8. This data shows that induction with 0.002% arabinose reduces the doubling time to approximately 65mins, significantly faster than when 0.02% arabinose is used. I am currently testing whether using lower levels of arabinose will allow chassis driving BBa_I7102 to grow faster and produce detectable levels of GFP. Should this prove unsuccessful, an alternative is to use a chassis with linear induction of the P_{BAD} promoter. This should allow reduced induction of pCH1497 and hence an increase in growth rate of the chassis.

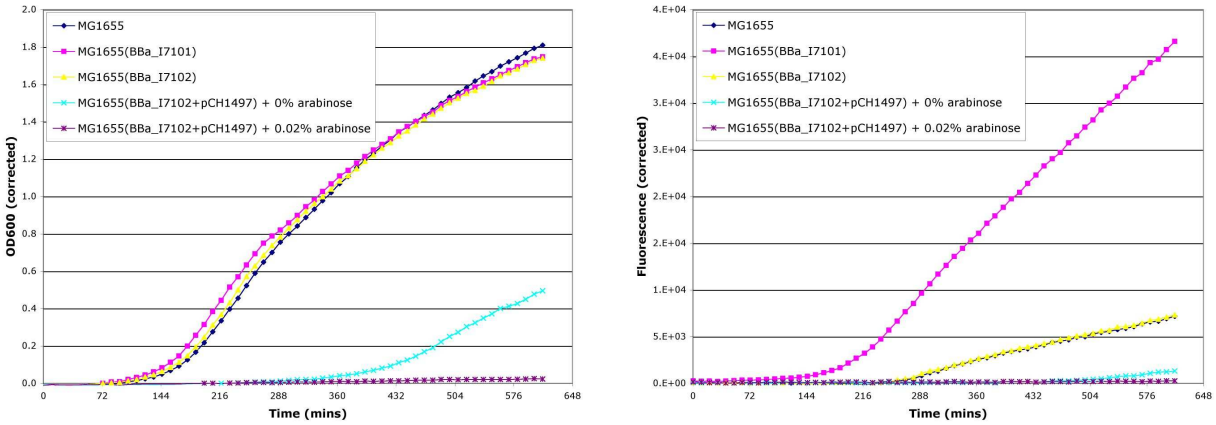


Figure 7: Growth curves for untransformed MG1655 and MG1655 containing I7101, I7102 and both I7102 and pCH1497 with and without arabinose induction are shown on the left. The corresponding GFP measurements are shown on the right. All curves are the mean of samples grown from three different cultures.

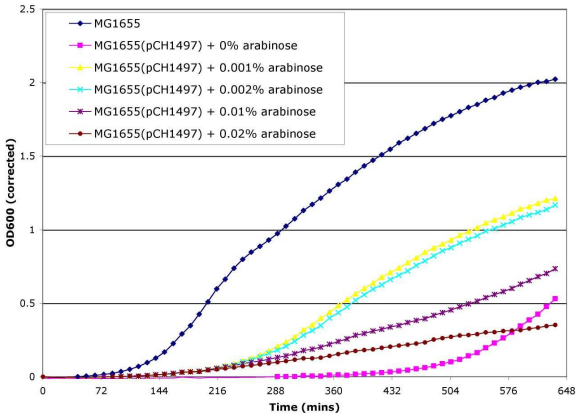


Figure 8: Growth curves for wild type MG1655 and MG1655 containing pCH1497 in the presence of different levels of arabinose. All curves are the mean of samples grown from three different cultures.

Research Timeline

These early experiments are part of the large body of work necessary to achieve my specific aims. A workflow diagram (Figure 9) shows the order in which I intend to carry out this work and indicates some of the intermediate project goals.

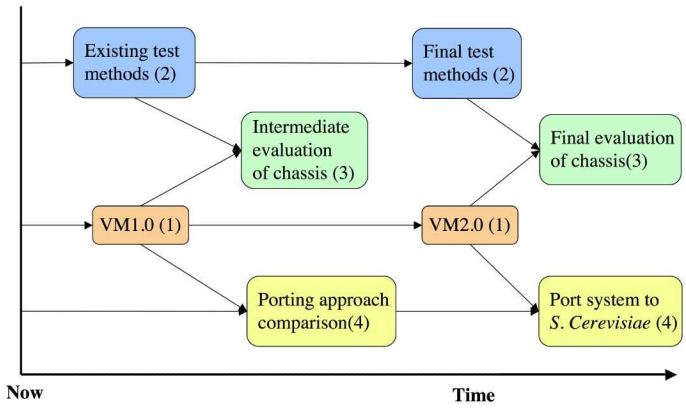


Figure 9: A timeline of the intermediate and final goals of the project. Each intermediate goal is associated with a single specific aim. The specific aims are indicated by numbers in brackets and a unique color.

5 References

[Asai 99] T Asai; D Zaporojets; C Squires & 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, vol. 96, no. 5, pages 1971–6, 1999.

[Baneyx 99] F Baneyx. *Recombinant protein expression in Escherichia coli.* Curr Opin Biotechnol, vol. 10, no. 5, pages 411–21, 1999.

[Baracchini 88] E Baracchini; R Glass & H Bremer. *Studies in vivo on Escherichia coli RNA polymerase mutants altered in the stringent response.* Mol Gen Genet, vol. 213, no. 2-3, pages 379–87, 1988.

[Basu 04] Subhayu Basu; Rishabh Mehreja; Stephan Thiberge; Ming-Tang Chen & Ron Weiss. *Spatiotemporal control of gene expression with pulse-generating networks.* Proc Natl Acad Sci U S A, vol. 101, no. 17, pages 6355–60, 2004.

[Bonomo 05] Jeanne Bonomo & Ryan T Gill. *Amino acid content of recombinant proteins influences the metabolic burden response.* Biotechnol Bioeng, vol. 90, no. 1, pages 116–26, 2005.

[Brink 95] M F Brink; M P Verbeet & H A de Boer. *Specialized ribosomes: highly specific translation in vivo of a single targetted mRNA species.* Gene, vol. 156, no. 2, pages 215–22, 1995.

[Cabrera 03a] Julio E Cabrera & Ding J Jin. *The distribution of RNA polymerase in Escherichia coli is dynamic and sensitive to environmental cues.* Mol Microbiol, vol. 50, no. 5, pages 1493–505, 2003.

[Cabrera 03b] Julio E Cabrera & Ding Jun Jin. *Construction, purification, and characterization of Escherichia coli RNA polymerases tagged with different fluorescent proteins.* Methods Enzymol, vol. 370, no. 0076-6879, pages 3–10, 2003.

[Dennis 04] Patrick P Dennis; Mans Ehrenberg & Hans Bremer. *Control of rRNA synthesis in Escherichia coli: a systems biology approach.* Microbiol Mol Biol Rev, vol. 68, no. 4, pages 639–68, 2004.

[Deuschle 89] U Deuschle; R Pepperkok; F B Wang; T J Giordano; W T McAllister; W Ansorge & H Bujard. *Regulated expression of foreign genes in mammalian cells under the control of coliphage T3 RNA polymerase and lac repressor.* Proc Natl Acad Sci U S A, vol. 86, no. 14, pages 5400–4, 1989.

[Dieci 00] G Dieci; L Bottarelli; A Ballabeni & S Ottonello. *tRNA-assisted overproduction of eukaryotic ribosomal proteins.* Protein Expr Purif, vol. 18, no. 3, pages 346–54, 2000.

[Dubendorff 91a] J W Dubendorff & F W Studier. *Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor.* J Mol Biol, vol. 219, no. 1, pages 45–59, 1991.

[Dubendorff 91b] J W Dubendorff & 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, vol. 219, no. 1, pages 61–8, 1991.

[Elowitz 00] M B Elowitz & S Leibler. *A synthetic oscillatory network of transcriptional regulators.* Nature, vol. 403, no. 6767, pages 335–8, 2000.

[Elowitz 02] Michael B Elowitz; Arnold J Levine; Eric D Siggia & Peter S Swain. *Stochastic gene expression in a single cell.* Science, vol. 297, no. 5584, pages 1183–6, 2002.

[Fuerst 86] T R Fuerst; E G Niles; F W Studier & B Moss. *Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase.* Proc Natl Acad Sci U S A, vol. 83, no. 21, pages 8122–6, 1986.

[Gardner 00] T S Gardner; C R Cantor & J J Collins. *Construction of a genetic toggle switch in Escherichia coli.* Nature, vol. 403, no. 6767, pages 339–42, 2000.

[Gill 00] R T Gill; J J Valdes & W E Bentley. *A comparative study of global stress gene regulation in response to overexpression of recombinant proteins in Escherichia coli.* Metab Eng, vol. 2, no. 3, pages 178–89, 2000.

[Glick 95] B R Glick. *Metabolic load and heterologous gene expression.* Biotechnol Adv, vol. 13, no. 2, pages 247–61, 1995.

[Guzman 95] L M Guzman; D Belin; M J Carson & J Beckwith. *Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter.* J Bacteriol, vol. 177, no. 14, pages 4121–30, 1995.

[He 95] B He; W T McAllister & R K Durbin. *Phage RNA polymerase vectors that allow efficient gene expression in both prokaryotic and eukaryotic cells.* Gene, vol. 164, no. 1, pages 75–9, 1995.

[Herman 95] A Herman & G Wegrzyn. *Effect of increased ppGpp concentration on DNA replication of different replicons in Escherichia coli.* J Basic Microbiol, vol. 35, no. 1, pages 33–9, 1995.

[Hoffmann 04] Frank Hoffmann. *Stress Induced by Recombinant Protein Production in Escherichia coli.* Advances in Biochemical Engineering/Biotechnology, vol. 89, pages 73–92, 2004.

[Hooshangi 05] Sara Hooshangi; Stephan Thiberge & Ron Weiss. *Ultrasensitivity and noise propagation in a synthetic transcriptional cascade.* PNAS, vol. 102, no. 10, pages 3581–3586, 2005.

[Hui 87] A Hui & 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, vol. 84, no. 14, pages 4762–6, 1987.

[Hurt 99] E Hurt; S Hannus; B Schmelzl; D Lau; D Tollervey & G Simos. *A novel in vivo assay reveals inhibition of ribosomal nuclear export in ran-cycle and nucleoporin mutants.* J Cell Biol, vol. 144, no. 3, pages 389–401, 1999.

[Ikura 02] Koji Ikura; Tsuyoshi Kokubu; Shunji Natsuka; Akira Ichikawa; Motoyasu Adachi; Kazuyo Nishihara; Hideki Yanagi & Shigeru Utsumi. *Co-overexpression of folding modulators improves the solubility of the recombinant guinea pig liver transglutaminase expressed in Escherichia coli.* Prep Biochem Biotechnol, vol. 32, no. 2, pages 189–205, 2002.

[Imburgio 00] D Imburgio; M Rong; K Ma & W T McAllister. *Studies of promoter recognition and start site selection by T7 RNA polymerase using a comprehensive collection of promoter variants.* Biochemistry, vol. 39, no. 34, pages 10419–30, 2000.

[Khlebnikov 02] A Khlebnikov; T Skaug & Jay D Keasling. *Modulation of gene expression from the arabinose-inducible araBAD promoter.* J Ind Microbiol Biotechnol, vol. 29, no. 1, pages 34–7, 2002.

[Lopez 94] P J Lopez; I Iost & M Dreyfus. *The use of a tRNA as a transcriptional reporter: the T7 late promoter is extremely efficient in Escherichia coli but its transcripts are poorly expressed.* Nucleic Acids Res, vol. 22, no. 7, pages 1186–93, 1994.

[Moyle 91] H Moyle; C Waldburger & M M Susskind. *Hierarchies of base pair preferences in the P22 ant promoter.* J Bacteriol, vol. 173, no. 6, pages 1944–50, 1991.

[Neidhardt 63] F C Neidhardt. *Properties of a bacterial mutant lacking amino acid control of RNA synthesis.* Biochim Biophys Acta, vol. 68, no. 0006-3002, pages 365–79, 1963.

[Neidhardt 74] F C Neidhardt; P L Bloch & D F Smith. *Culture medium for enterobacteria.* J Bacteriol, vol. 119, no. 3, pages 736–47, 1974.

[Plambeck 03] Craig A Plambeck; Ann H Y Kwan; David J Adams; Belinda J Westman; Louise van der Weyden; Robert L Medcalf; Brian J Morris & Joel P Mackay. *The structure of the zinc finger domain from human splicing factor ZNF265 fold.* J Biol Chem, vol. 278, no. 25, pages 22805–11, 2003.

[Ringquist 92] S Ringquist; S Shinedling; D Barrick; L Green; J Binkley; G D Stormo & L Gold. *Translation initiation in Escherichia coli: sequences within the ribosome-binding site.* Mol Microbiol, vol. 6, no. 9, pages 1219–29, 1992.

[Rosenfeld 05] Nitzan Rosenfeld; Jonathan W. Young; Uri Alon; Peter S. Swain & Michael B. Elowitz. *Gene Regulation at the Single-Cell Level.* Science, vol. 307, no. 5717, pages 1962–1965, 2005.

[Schmidt 04] F R Schmidt. *Recombinant expression systems in the pharmaceutical industry.* Appl Microbiol Biotechnol, vol. 65, no. 4, pages 363–72, 2004.

[Shine 74] J Shine & L Dalgarno. *The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites.* Proc Natl Acad Sci U S A, vol. 71, no. 4, pages 1342–6, 1974.

[Sorensen 05] Hans Peter Sorensen & Kim Kusk Mortensen. *Advanced genetic strategies for recombinant protein expression in Escherichia coli.* J Biotechnol, vol. 115, no. 2, pages 113–28, 2005.

[Studier 90] F W Studier; A H Rosenberg; J J Dunn & J W Dubendorff. *Use of T7 RNA polymerase to direct expression of cloned genes.* Methods Enzymol, vol. 185, no. 0076-6879, pages 60–89, 1990.

[Studier 05] F William Studier. *Protein production by auto-induction in high density shaking cultures.* Protein Expr Purif, vol. 41, no. 1, pages 207–34, 2005.

[Swain 02] Peter S Swain; Michael B Elowitz & Eric D Siggia. *Intrinsic and extrinsic contributions to stochasticity in gene expression.* Proc Natl Acad Sci U S A, vol. 99, no. 20, pages 12795–800, 2002.

[Varshney 90] U Varshney & U L RajBhandary. *Initiation of protein synthesis from a termination codon.* Proc Natl Acad Sci U S A, vol. 87, no. 4, pages 1586–90, 1990.

[Vind 93] J Vind; M A Sorensen; M D Rasmussen & 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, vol. 231, no. 3, pages 678–88, 1993.

[Wick 04] Lukas M. Wick & Thomas Egli. *Molecular Components of Physiological Stress Responses in Escherichia coli.* Advances in Biochemical Engineering/Biotechnology, vol. 89, pages 1–45, 2004.

[Xiao 91] H Xiao; M Kalman; K Ikehara; S Zemel; G Glaser & M Cashel. *Residual guanosine 3',5'-bispyrophosphate synthetic activity of relA null mutants can be eliminated by spoT null mutations.* J Biol Chem, vol. 266, no. 9, pages 5980–90, 1991.

[Yarchuk 92] O Yarchuk; N Jacques; J Guillerez & M Dreyfus. *Interdependence of translation, transcription and mRNA degradation in the lacZ gene.* J Mol Biol, vol. 226, no. 3, pages 581–96, 1992.

[Yu 00] D Yu; H M Ellis; E C Lee; N A Jenkins; N G Copeland & D L Court. *An efficient recombination system for chromosome engineering in Escherichia coli.* Proc Natl Acad Sci U S A, vol. 97, no. 11, pages 5978–83, 2000.

[Zhang 97] X Zhang & F W Studier. *Mechanism of inhibition of bacteriophage T7 RNA polymerase by T7 lysozyme.* J Mol Biol, vol. 269, no. 1, pages 10–27, 1997.

[Zhang 02] X Zhang; P Dennis; M Ehrenberg & H Bremer. *Kinetic properties of rrn promoters in Escherichia coli.* Biochimie, vol. 84, no. 10, pages 981–96, 2002.

6 Sample Data Sheets

BBa_F2620

$3OC_6HSL \rightarrow$ PoPS Receiver

Author(s): Barry Canton [bcanton@mit.edu] Last Update: May 10, 2005

PDL Devices

Description

A transcription factor [LuxR] that is active in the presence of cell-cell signaling molecule [$3OC_6HSL$] is controlled by an operator [TetR]. *Device input* is $3OC_6HSL$. *Device output* is PoPS produced at a LuxR-regulated operator.

Usage

Full PoPS output at high $3OC_6HSL$ levels and high plasmid copy [e.g., pSB1A2] results in a reduced cell growth rate (see *Load* section). If used in a cell containing TetR then a second input signal [aTc] can be used to produce a logical *AND* function.

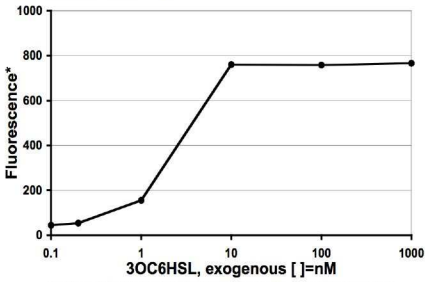
Characteristics

Input Swing: #-# nM $3OC_6HSL$, exogenous
Output Swing: #-# PoPS
Switch Point: 2 nM $3OC_6HSL$, exogenous
LH Latency: # seconds
HL Latency: # seconds

Key Components

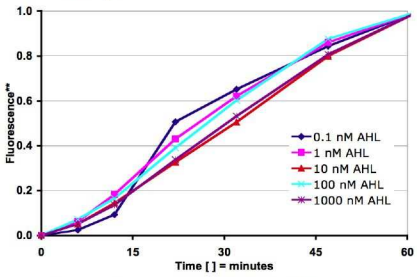
BBa_R0040: TetR-regulated operator
BBa_C0062: luxR ORF
BBa_R0062: LuxR-regulated operator

Transfer Function



*Device output measured indirectly via PoPS-driven fluorescence from *BBa_E0430*, [ ] = geometric mean, arbitrary units. Host cell *MC4100*, device carried on *pSB3K3*, 100ml batch flask, supplemented M9 media, FACScan cytometer [see *MIT SBWG FACS protocol*].

Latency



**Device output measured indirectly via PoPS-driven fluorescence from *BBa_E0430*, [ ] = geometric mean, arbitrary units. Host cell *MC4100*, device carried on *pSB3K3*, 200ul 96-well plate, VictorV plate reader [see *MIT SBWG plate reader protocol*].

Load

NTP/sec/copy: # NTP per second
AA/sec/copy: # AA per second

Stability

Genetic: > # replication events*
Operational: > # replication events*

Compatibility

Device has been shown to work in *MC4100*, *MG1655*, and *DH-5α*.
Device has been shown to with on *pSB3K3* and *pSB1A2*.
Device has been shown to work with *E0430* and *E0434*.
Crosstalk with input molecular similar to *$3OC_6HSL$*.
Crosstalk with systems containing *TetR*.

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

MG1655

E. coli K strain

Author(s): Barry Canton [bcanton@mit.edu]

Last Update: May 10, 2005

E. Coli Chassis

Description

MG1655 was the first sequenced K strain of *E. coli* and is closely related to wild type *E. coli*. It has relatively simple nutritional requirements and a minimum doubling time of 20mins. The physiology of MG1655 is well understood and a wide range of techniques exist to engineer it into more advanced chassis for engineered systems.

Usage

When growing in Neidhardt rich defined media, MG1655 has been shown to be able to supply a wide range of engineered systems. Its optimal growth temperature is 37°C but it grow at temperatures between 4°C and 42°C.

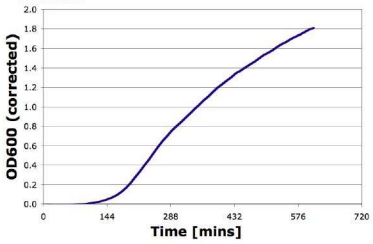
Supply Characteristics

Transcription capacity: # RNAP max.
Translation capacity: # Ribosomes max.
Protein degradation capacity: # AA/min max.
Replication capacity: # base pairs of plasmid DNA

Key Machinery

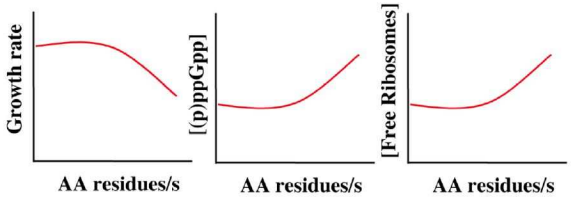
Replication: wt *E. coli* DNA polymerases
Transcription: wt *E. coli* RNAP
Translation: wt *E. coli* Ribosomes
Degradation: wt *E. coli* proteases

Growth



*Growth measured in Victor3 fluorimeter [Perkin Elmer] at 37°C in supplemented M9 minimal media.

Supply



**Translation Supply measured using SB standard translation demand test suite. Chassis were grown in Neidhardt rich defined media. Performance metrics measured with SB standard chassis evaluation protocols.

Growth Parameters

Min. Doulbing Time: # mins
Max. Density: # cfu/ml in LB

Stability

Genetic: > # mutations per doubling
Plasmid: > # doublings before 50% loss of pSB1A2

Compatibility

Chassis has been shown to grow in *LB, M9 minimal media, Neidhardt rich defined media*
Chassis is compatible with plasmids *pSB3K3, pSB4A3, pSB2k3,* and *pSB1A2*.
Chassis has been shown to grow in chemostat, batch culture, microscope slide.
Systems including Bba_R0052 have been shown to be toxic to MG1655