Engineered Sensors and Genetic Regulatory Networks for Control of Cellular Metabolism

By

Felix Moser

M.S., University of California – Berkeley and San Francisco (2011)
B.A., Cornell University (2007)

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

September, 2013

Author.....

Felix Moser
Department of Biological Engineering
August 8, 2013

Certified by.....

Christopher A. Voigt
Associate Professor of Biological Engineering
Thesis Advisor

Approved by.....

Douglas A. Lauffenburger
Ford Professor of Biological Engineering, Chemical Engineering, and Biology
Head, Department of Biological Engineering

This doctoral thesis has been examined by a committee of the Department of Biological Engineering as follows:

Certified by .....

Edward F. DeLong
Professor of Biological Engineering
Professor of Civil and Environmental Engineering
Thesis Committee Chair

Certified by .....

Christopher A. Voigt
Associate Professor of Biological Engineering
Thesis Advisor

Certified by .....

Kristala L. J. Prather
Theodore T. Miller Career Development Associate Professor of Chemical Engineering
Committee Member

Certified by .....

Gregory N. Stephanopoulos
Professor of Chemical Engineering
Committee Member

Engineered Sensors and Genetic Regulatory Networks for Control of Cellular Metabolism

by

Felix Moser

Submitted to the Department of Biological Engineering on August 8th, 2013 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biological Engineering

Abstract

Complex synthetic genetic programs promise unprecedented control over cellular metabolism and behavior. In this thesis, I describe the design and development of a synthetic genetic program to detect conditions underlying acetate formation in *Escherichia coli*. To construct this program, I first developed sensors that detected and propagated relevant information into genetic circuits. These sensors include a novel sensor for genotoxic methylation exposure in *Saccharomyces cerevisiae* and sensors for oxygen, acetate, and glycolytic flux in *E. coli*. The methylation sensor served to prototype generalizable tuning mechanisms and was tuned to a sensitivity and detection threshold useful for several applications, including the detection of MeI formation in methyl halide transferase-expressing cultures of yeast and the detection of MeI in soil. The sensors for oxygen and acetate were integrated into a program that can uniquely detect acetate formation in anaerobic conditions in *E. coli*. Finally, to validate their use at higher scales in production strains, the oxygen sensor and two genetic programs were characterized in 10 L fed-batch fermentations. Together, this work demonstrates the characterization of novel genetic elements, their integration into genetic programs, and the validation of those programs at industrially relevant scales.

Thesis Supervisor: Christopher A. Voigt

Title: Associate Professor of Biological Engineering

Acknowledgements

First and foremost, I would like to thank my advisor Chris "bossy bear" Voigt for his mentorship, guidance, support, and for always holding me to a high standard. I am also very grateful to my committee members Ed DeLong, Kristala Prather, and Greg Stephanopoulos for their advice and guidance.

A great many thanks to all the members of the Voigt lab, past and present, who made the past few years so memorable and were constant companions in the lab: Conor "cougar bear" McClune, Yong-Jin "bare bear" Park, Johann "l'chaim" Elbaz, Lauren "shared bear" Woodruff, Octavio "smirky bear" Mondragón-Palomino, Jésus "honeybadger" Fernandez-Rodriguez, Amar "meth bear" Ghodasara, Jonghyeon "stringy bear" Shin, Lei "ninja bear" Yang, Miryoung "stabby bear" Song, Hui "tiny bear" Zhou, Mike (& Audra & Jonathan) "Chaco bear" Smansky, Jenn "crafty bear" Brophy, Alec "moneybear" Nielsen, Thomas "golden bear" Segall-Shapiro, David Sukovich, Eric "i" Ma, Dan "2.0" Kaemmerer, Robin Prince, Brynne "trouble bear" Stanton, Tae-Seok "dad" Moon, Ying-Ja "singing bear" Chen, Chunbo Lou, Brian "drunky bear" Caliando, Dehua "nitro bear" Zhao, Virgil Rhodius, Alvin Tamsir, Ethan Mirsky, Dan Widmaier, Jeff Tabor, Howard Salis, Travis Bayer, Karsten Temme, Liz Clarke, Eli Groban, Anselm Levskaya, and Chris Anderson (and his password).

I reserve special gratitude for the excellent lab managers whose patience, know-how, and tireless work enabled the lab to thrive. Rena Hill's discipline, hard work, and friendly smile served as the template of professionalism in the lab for years before and during my time. Kathy Corsgren's unending effort in setting up the lab enabled a near-seamless transition from UCSF to MIT. Barbara "mama bear" Karampalas' unwavering helpfulness, patience, and unbounded generosity never failed to impress and aid everybody in the lab. Also many thanks to Terry King, to whom the phrase "grace under fire" cannot be applied more fittingly. I am indebted to each of them for their tireless hard work and for making it a great lab.

I have had the privilege of working with many excellent researchers outside of Voigt lab. Andrew Horwitz and Jacinto Chen, my perpetual collaborators, carried out much of the methylation sensor work in *S. cerevisiae*. Wendell Lim's guidance enabled this work. MIT's UROPs are renowned for their excellence, and Shireen Rudina, who worked with me for nearly a year, was no exception. Finally, thanks to all the professionals at DSM who enabled the fermentation work: Nicolette Broers, Sybe Hartmans, Hans Roubos, Roel Bovenberg, and Richard Kerkman.

I am fortunate to have been a member of several great departments during my graduate career. I am thankful for my friends at the UCSF/UCB JGGB for their brilliance, professionalism, and friendship. I found a welcoming home in the MIT BioE department and will be perpetually grateful for the friendships I have found in it. I am constantly impressed by the people at MIT BioE, and Doug Lauffenburger, whose vision and leadership has enabled BioE at MIT to thrive, deserves much credit for creating the familial and professional atmosphere in the department. I was fortunate to be brought into this department after college by Drew Endy, who hired me as a technician in the lab, and much the past six years followed from Drew's spurious decision to hire me. I also benefitted greatly from the patience, guidance, and drive of the rest of the Endy lab. To Jason Kelly, Barry Canton, Reshma Shetty, Austin Che, Francois St. Pierre, Ty Thomson, and Samantha Sutton, I owe my gratitude.

Lastly, my thanks and love to my parents Barbara and Rudi and my stepdad Frank for their love and support. Without you guys, I wouldn't have made it.

I gratefully acknowledge the funding sources that made this work possible: a National Science Foundation (NSF) Graduate Research Fellowship, the NSF-funded Synthetic Biology Engineering Resource Center (SynBERC), the UCSF/UCB Joint Graduate Group in Bioengineering, the MIT Bioengineering department, and DSM for generously supporting my internship in Delft.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Genetic programs as a means of control

Control of the growth, behavior, and composition of cells, whether they are used as biosynthetic factories, therapeutics, or crops, is critical to nearly every modern application of biological systems. Control of a cell culture is typically exerted by sensing a state of interest (e.g. dissolved oxygen, metabolite accumulation) and then correcting any deviation from the ideal state by activating a feedback loop that corrects this deviation (Figure 1.1A). Traditionally, control of cultures has been implemented by external sensing and feedback[1]. For example, in fed-batch fermentations, a low oxygen environment would be detected with a dissolved oxygen probe and corrected by increasing the airflow through the fermenter. However, this mode of control is limited in various ways (see below). Another promising means of control is internal control via synthetic genetic programs[2]. Developing such synthetic genetic programs is one of the key aims of synthetic biology and the overarching goal of this thesis work.

Activating feedback by changing cells' external environment is a useful but limited mechanism of control. The temperature, media, aeration, and composition of a culture can all be used to affect its growth rate and behavior. However, these mechanisms have several drawbacks. Because external controllers and sensors typically detect the state of the entire culture, they are indiscriminate as to the function of individual cells. Thus, although the global state of the culture can be corrected by external feedback, individual cells may persist or even be forced outside the desirable range of a critical physiological parameter (Figure 1.1C). External changes to cells' environment may also cause other undesirable changes in the organism. For example, changing the temperature of a culture will change the rates of all enzymes in the cell, not simply those controlling the desired behavior. External changes are often nonuniform across the culture. For example, changes in aeration will change the dissolved oxygen content of the culture unevenly, leading to pockets of near-anaerobic conditions[3]. Therefore, although it is easy to implement, external control of a cell culture is limited in its effectiveness.

Synthetic genetic programs offer another means of control. A synthetic genetic program is a change in an organism's DNA "code" that adds to or changes the organism's native "program". A program can be as simple as an inducible genetic switch or as complex as an entire genome. The key advantage of a synthetic genetic program as a controller is that it works at single cell resolution. Because each cell in a culture carries the genetic program, sensing and feedback are carried out within each cell, independent of the behavior of other cells in the culture (Figure 1.1B). This means that cells experiencing different conditions due to differences in the local environment will respond only to their local environment. The net result is that more cells in the culture can be maintained in the desired cell state (Figure 1.1D). Genetic programs also offer a potentially limitless repertoire of sensory inputs (See Chapter 1.3). In comparison, external sensors (e.g. oxygen probes, pH probes) are relatively limited in number and application. Genetic controllers also offer greater versatility than external controllers. Because a functioning genetic control program is integrated into the production host, no expensive hardware changes need to be made if the host is moved into a different process or production platform. Additionally, if the program is orthogonal (i.e. it uses none of the host's native resources), then the program itself may be moved in its entirety to another organism. During such transfers, the program may need to be re-tuned to function optimally, but this is still a smaller investment than re-tooling the hardware of a production process. Successful implementation of a synthetic genetic controller may prove external controllers redundant, lower the capital requirements for productive culturing platforms, reduce process development costs, and enable improvements in product yield and quality.

In this work, I aim to develop a genetic program that acts as a controller for acetate production in *E. coli*. Acetate is a by-product of metabolism that lowers product yields in *E. coli* cultures, destabilizes recombinant proteins, and leads to slower growth and biomass accumulation[4]. A genetic controller for acetate production in *E. coli* would provide a proof of concept for the advantages of genetic control over external control. If successful, this type of genetic program would be a valuable new application in biotechnology.

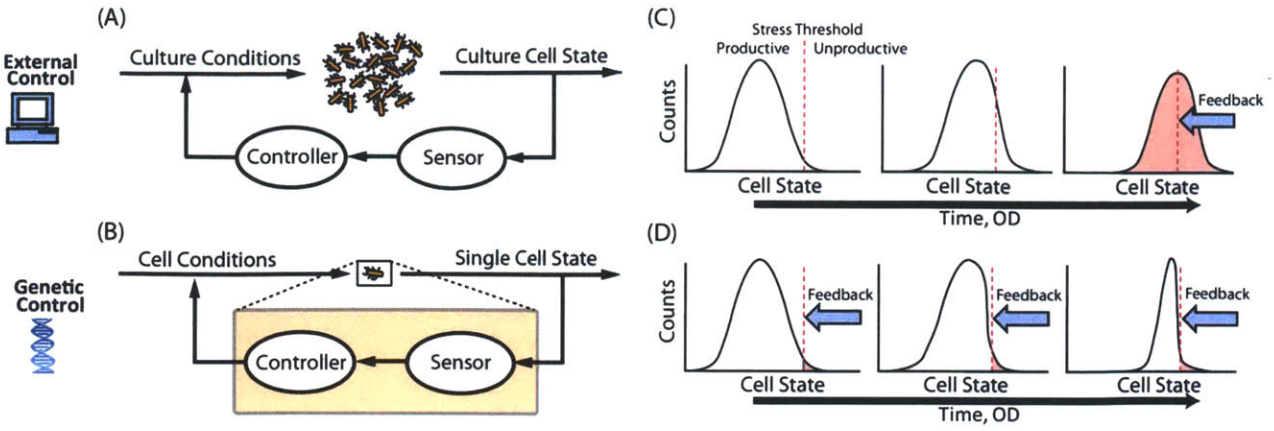


Figure 1.1: Synthetic genetic programs as internal cell state controllers. Diagrams of external and genetic control schemes for cell cultures. **(A)** External controllers implement feedback by sensing and correcting culture-wide conditions. **(B)** Genetic controllers sense and correct individual cell states. **(C)** External controllers will only activate feedback when the culture average has passed a pre-determined threshold, beyond which cells are unproductive. As a result, a large fraction of cells experience stresses that reduce productivity before feedback is activated. **(D)** Genetic controllers activate feedback when the individual cells carrying them experience threshold levels of stress. This enables cells to respond individually, which maintains a greater fraction of the culture in a productive cell state.

1.2 The paradigm of the genetic program

To highlight the essential features of a synthetic genetic program, it is useful to apply the paradigm of a programmable machine. In this paradigm, the genetic program is composed of three modules: input, computation, and output (Fig. 1.2)[5]. The input module, composed of sensors, determines how the cell reads information from the user or the environment. The computation module processes the information received from the input module and performs the programmed calculations. The output module receives the results from the computations and actuates the feedback and any other desired functions. Together, these modules encapsulate a program's entire function.

Each module of a synthetic genetic program is ultimately composed of molecules. The input module is composed of sensor molecules that specifically detect the presence or activity of a molecule or state of interest and transduce this information downstream to genetic circuits[6,7]. These sensor molecules are typically proteins that can either modulate transcription or propagate post-translational signals (e.g. phosphorylation)[8]. They can also be RNA molecules that regulate transcript stability (e.g. ribozymes)[9]. The computational module can be made up

of protein signaling circuits or genetic circuits. Protein signaling circuits can very quickly pass on post-translational signals in a manner that performs computation[10]. In genetic circuits, transcription factors are used to control gene expression in complex, multi-layered arrangements that perform computation[11]. These signaling pathways can be abstracted into digital (Boolean) logic gates, which can be combined to perform complex computation[12]. The abstraction of digital logic is a useful tool to guide construction and in developing very large circuits (>20 gates), but can be substituted for more complex analog computation methods at smaller scales[5,13]. The output module in a genetic program typically encodes genes or operons that actuate the desired phenotype of the organism. This output can be as simple as a single enzyme that adjusts metabolite levels or as complex as a system of operons designed to alter the development of the cell[14,15]. Together, the molecular composition of each of these modules is critical to the functional characteristics and requirements of the genetic program.

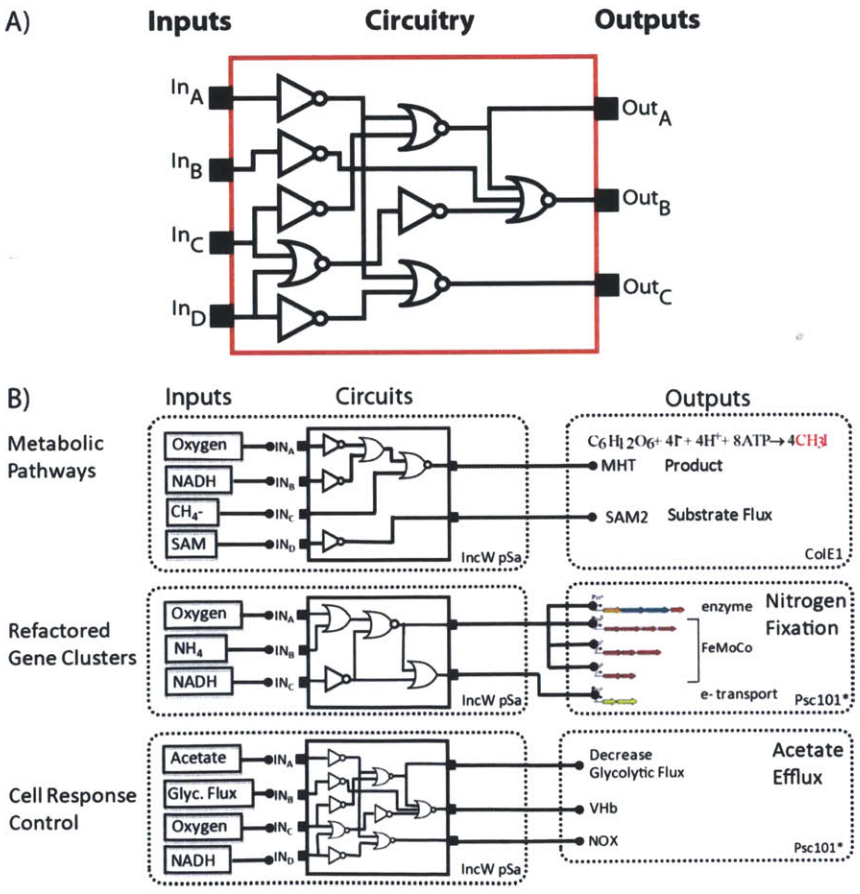


Figure 1.2: The Paradigm of a Genetic Program. (A) A generalize genetic program, composed of three modules: Inputs, Circuitry, and Outputs. These modules are composed at their lowest level by molecules whose interactions define the detection (Input), computation (Circuitry), and actuation (Output) of the encoded program. **(B)** Examples of complex genetic programs that match the paradigm above. These examples highlight applications in the control of metabolic pathways, refactored gene clusters, and cell response control. The sensors presented in this work are used in all three of the genetic program designs, highlighting the potential importance of these sensors.

1.3 Sensors enable genetic programs

Our ability to design and build synthetic genetic programs is highly dependent on the availability of useful sensors. Indeed, the employment of genetic programs today is largely hampered by the limited availability of sensors for molecules and conditions of interest[16,17]. Sensors are uniquely enabling for several reasons. First, the utility of a genetic program is highly constrained by its inputs. A genetic program must be able to sense the information relevant to its function. Second, by expanding the number of molecules or conditions a program can sense, one can exponentially expand the number of states to which the program can respond. This is because each sensor can be coupled with others sensors to define a unique condition. So as the number of sensors in a program grows, the possible number of combinations of sensed states increases geometrically. Such combinations of sensors feed inputs into the logic gates that compose the downstream genetic circuitry in the program, enabling ever larger programs to be built.

Thorough characterization of a sensor is vital to its utility. Like all biological parts, the most useful sensors are functionally well-characterized, context-independent (i.e. modular), and orthogonal[18]. Archetype examples are the lacI, araC, and tetR transcription factors, which have been exquisitely characterized in a number of contexts and show a strong presence in both the scientific and engineering literature as a direct result[19]. Individual, well-characterized sensors can compose useful genetic programs by themselves. Such programs are typically called "biosensors" and can be implemented to detect and report a host of environmental toxins[20], valuable chemicals[21], and even explosives[22] (see Chapter 2.1). Although the availability of sensors has recently grown as the result of efforts to curate collections of biological parts[23] and use directed mutagenesis to change the specificity of extant sensors[24], there still exists a strong need for more as the drive to build larger genetic programs grows[25]. Therefore, the expansion of the repertoire of highly modular and tunable sensors will enable the construction of more diverse and functional genetic programs and biosensors.

1.4 Key sensor characteristics and tuning

Sensors are defined by their input/output "response function." Inherent in this response function is the sensor's specificity, detection threshold, sensitivity of response, basal leak, and dynamic range. Specificity dictates exactly which molecules the sensor will respond to. An ideal sensor is responsive to only a single species of molecule and will show no response to any other molecules or conditions. The detection threshold of a sensor is the minimal concentration of this input molecule that can be sensed above background. Good sensors will have detection thresholds at or below relevant concentrations, as determined by the application. The sensor's sensitivity is the change in output over the change in input; essentially, it is the slope of the response function at a given input concentration. Most biological sensors tend to have sigmoidal response functions. This is primarily due to the fact that most sensors respond in accordance with a second-order rate law which carries an inherent exponential term. Cooperative multimerization of the sensor molecules adds to this effect. The basal leak (i.e. background) is the output activity of the sensor when no input molecule is present. A good sensor will have almost no basal leak. Directly related to the basal leak is the sensor's dynamic range, defined as the difference between the sensor's lowest OFF state and highest ON state. A high basal leak usually lowers the dynamic range. Although the maximal attainable dynamic range is often desirable, in practice a 10-fold dynamic range is sufficient to connect a sensor to downstream genetic circuitry. The highest ON state of a sensor is set by a number of factors, including the sensor protein's binding affinity to its target output domains and the highest metabolic burden the cell can withstand from the production of the sensor's output molecules.

Many of a sensor's key characteristics can be tuned, either rationally or via selection for the desired specification following mutagenesis. The specificity of a sensor can be altered by mutagenesis of the recognition site. For example, the specificities of araC, xylR, and nahR have been altered to an impressive range of alternate molecules while maintaining the sensors' ability to modulate transcription[26]. A sensor's detection threshold can be similarly improved by mutating the active site for greater affinity to the molecule of interest. Because a sensor's response is often dependent on a second-order rate law, simply increasing the concentration of the sensor by increasing its expression can improve its detection threshold. The sensitivity of a

sensor's response is more difficult to tune due to the inherent nature of the molecule. Sensors that form multimeric complexes prior to activation tend to have sharper responses (i.e. larger Hill coefficients). Sensors that autoactivate also show sharper responses. By removing autoactivation or abolishing multimers by mutation, one can therefore attenuate the sharpness of the sensor's response. The basal leak of a genetic sensor can be tuned down by weakening its cognate promoter. The sensor can also be expressed at lower levels to minimize promoter binding while in its OFF state. The dynamic range of a sensor is most easily changed by modulating the strength of the cognate promoter or ribosome binding site (RBS) driving the output. How tractable a sensor will be to these efforts at tuning is always dependent on the molecule and its mechanism of activation, making it difficult to predict how effective a given strategy will be.

Additional network architectures can impact sensor performance. Positive feedback in transcriptional regulation has been shown to result in a more ultrasensitive, digital response[27]. An ultrasensitive response could be replicated in yeast by either building such a loop or through the inclusion of interactions that sequester the regulator[28,29]. On the other hand, to engineer a more linear, faster, or pulsed response in the sensor, one could implement negative genetic feedback[30,31,32]. More complex architectures, such as feed-forward loops, can also be used to engineer complex dynamics and robustness to noise[33,34]. Also useful for sensing systems would be the engineering of "scale-free sensing," a characteristic of some complex networks that enables those networks to detect changes in the environment regardless of the level of the background signal[35,36]. Clearly, many modes of action are open for the engineering of sensors to altered dynamics and specifications.

Currently, new sensors are best discovered by searching the scientific literature for well-characterized modules with activities of interest. Minimally, the sensor protein, its ligand, and its cognate operator sequence and their functional relationships must be identified. Any further characterization is useful, because it will usually aid in narrowing the parameters in which the sensor functions. Recently, high-throughput approaches have been used to purify and characterize individual transcription factors and their operators[37]. Transferring sensors between organisms can also be fruitful. Many naturally-occurring protein sensors are highly modular,

meaning that their function is preserved independent of context and host environment. Such modular protein sensors can be used successfully in different organisms or be fused to other proteins that can transduce their function to downstream circuitry[38]. One successful approach has been to fuse heterologous sensor domains with DNA binding capacity to transcriptional regulators such as the Gal4 activation domain, which upregulates transcription when it is localized to the yeast Cyc1 promoter. When activated, the heterologous sensor will bind to its DNA operator, which has been located immediately upstream of a promoter sensitive to the native regulator. Such fused activator proteins have enabled various sensory applications such as redox sensing in mammalian cells[39] and bacterial one-hybrid assays[40]. The advancement of construction and screening technologies as well as our ability to engineer proteins will further enable the discovery and tuning of valuable sensors.

Currently, there is a lack of sensors that can detect conditions that are of great interest to industrial systems. Parameters such as dissolved oxygen, acetate accumulation, cytoplasmic redox state, toxic product accumulation, and the lack of important metabolites can negatively affect the viability of the culture and final yield of the desired product[41,42,43]. If genetic programs are to be built for industrial systems, sensors for these conditions need to be developed.

To enable the design and construction of genetic programs for industrial applications, I have characterized several sensors, integrated them into prototype circuitry, and tested some of these sensors and circuits in industrial fed-batch fermentations. These sensors detect the accumulation of methyl halides and changes in dissolved oxygen, acetate, and glycolytic flux in *Escherichia coli*. In the following chapters, I detail the design, construction, and testing of these sensors and circuits that integrate them. Chapter 2 expounds the characterization of a sensor for strong methylating compounds and its transfer into *Saccharomyces cerevisiae*. Chapter 3 details the discovery and tuning of sensors for oxygen, acetate, and glycolytic flux and the integration of the oxygen and acetate sensors into a functioning AND gate. In Chapter 4, I test the oxygen sensors and several prototype genetic programs in industrial conditions. Chapter 5 draws conclusions from this work, highlights lessons learned, and attempts to guide future efforts.

2. An engineered genetic sensor for strong methylating compounds in *Saccharomyces cerevisiae*

2.1 Abstract

Methylating chemicals are common in industry and agriculture and are often toxic partly due to their propensity to methylate DNA. The *Escherichia coli* Ada protein detects methylating compounds by sensing aberrant methyl adducts on the phosphotriester backbone of DNA. We characterized this system as a genetic sensor and engineered it to lower detection thresholds. By overexpressing Ada from a plasmid, we improved the sensor's dynamic range to 440-fold induction and lowered its detection threshold to 39 μM for methyl iodide. In eukaryotes, there is no known sensor of methyl adducts on the phosphotriester backbone of DNA. By fusing the N-terminal domain of Ada to the Gal4 transcriptional activation domain, we built a functional sensor for methyl phosphotriester adducts in *Saccharomyces cerevisiae*. This sensor could be tuned to variable specifications by altering the expression level of the chimeric sensor and changing the number of Ada operators upstream of the Gal4-sensitive reporter promoter. These changes resulted in a detection threshold of 28 μM and 5.3-fold induction in response to methyl iodide. When the yeast sensor was exposed to different S_N1 and S_N2 alkylating compounds, its response profile was similar to that observed for the native Ada protein in *E. coli*, indicating that its native function was retained in yeast. Finally, we showed that the specifications achieved for the yeast sensor were suitable for detecting methylating compounds at relevant concentrations in environmental samples. This work demonstrates the movement of a sensor from a prokaryotic to eukaryotic system and its rational tuning to achieve a desired response.

2.2 Introduction

A transcriptional genetic sensor is a unit of DNA that contains all of the necessary parts to convert an input stimulus to the up- or down-regulation of a promoter[44,45]. Following this paradigm, the output promoter of a sensor can be used as the input promoter of a genetic

circuit, which can implement signal-processing functions. Genetic sensors have been constructed that respond to many environmental signals, including light[46,47], temperature[48,49], gases[50,51], toxins (*e.g.*, arsenic)[52,53], and chemicals (*e.g.*, industrial products, pollutants or explosives)[54,55,56,57]. Many of these sensors are based on the transfer of parts (functional genetic elements) from one organism to another; for example, moving a TNT sensor from *E. coli* to *Arabidopsis*[56], an artificial quorum sensing system made of *Arabidopsis* parts transferred to yeast[58], light sensors from cyanobacteria and plants to *E. coli* and mammalian cells[46,59], and a redox sensor from *Streptomyces* to mammalian cells[60]. Such transfers often require sensor re-engineering and the substitution of parts to make the sensor functional in the new host.

Different applications require different performance specifications of a genetic sensor, which can be achieved by tuning the response function of the sensor. The response function is defined by how the sensor output (promoter transcription) changes as a function of the input stimulus. The shape of this function captures the responsiveness of the sensor to the input and provides information that aids its connection to a downstream circuit[61,62]. There are several descriptors of the response function that are particularly useful: the basal activity, cooperativity, dynamic range, detection threshold (lowest input concentration sensed above background), and sensitivity (the slope during the transition)[63]. Additionally, it is useful to be able to change the selectivity of the sensor to a particular stimulus to understand how it will respond in a mixture or complex environmental context. Various approaches, including directed evolution, have been applied to alter the properties of genetic sensors[64,65]. Synthetic biology has also developed "tuning knobs" to control transcription and translation that could be applied to altering sensor response[66,67,68,69,70.]

Here, we designed and characterized a sensor of methylating compounds, transferred it from *E. coli* to yeast, and tuned its response characteristics. Methylating agents are relevant to human health because they can induce the aberrant methylation of DNA, which can lead to mutations, misregulation, and ultimately disease. Many methylating agents leave methyl phosphotriester (PTE) adducts on DNA. These adducts are very stable, long-lasting moieties in eukaryotic cells due to their innocuous nature and resistance to DNA repair[71]. Because of their stability, methyl PTE adducts have been proposed as a biomarker for cumulative genotoxic

exposure[72,73]. Methylating agents that generate these adducts are common in industrial and agricultural processes and often produce other more damaging DNA lesions. For example, phenyl glycidyl ether (PGE) and *N*-dimethylnitrosamine (NDMN) are used in the manufacture of paint, resin and rubber[74,75]. Dimethyl sulfate (DMS) is a common alkylating agent used in kiloton quantities in a variety of industries[76]. Methylnitronitrosoguanidine (MNNG) and methyl methanesulfonate (MMS) are used in laboratories to study DNA damage and repair[77]. Methyl halides such as methyl chloride and methyl iodide (MeI) are methylating agents that are being controversially used as soil fumigants and intermediates to various chemical processes, including silicon rubber production[78,79]. All of these agents methylate the bases of DNA[103]. Due to the ubiquity and potency of genotoxic methylating agents, a sensor for DNA methylation damage could be a tool for environmental biosensing or a diagnostic system for long-term genotoxic exposure.

Escherichia coli has an adaptive response to methylating agents. They are sensed via the Ada protein, which is either directly methylated by S_N2 methylating agents or indirectly by S_N1 methylating agents via methyl PTE DNA adducts[80,81,82,83]. Ada moves along DNA, detects, and then transfers a single DNA Sp methyl PTE adduct onto its Cys38 residue (Figure 3A). The methylation of Ada's Cys38 residue activates Ada as a transcription factor. Ada then upregulates transcription of various DNA repair proteins, including its own *ada* gene. This positive feedback loop turns the very low basal expression of Ada into a strong, sustained response to the exposure of genotoxic methylating agents[84]. Ada has been used as a sensor for DNA methylation toxicity of genotoxic compounds to complement the Ames test, the gold standard for assaying mutagenicity of a compound[85].

No comparable, specific sensor of DNA methyl PTEs is known in eukaryotes. To develop such a sensor in eukaryotes, we fused the N-terminal domain of Ada (N-Ada) to the Gal4 trans-activation domain. The Gal4 domain, taken from yeast, is functional in a broad range of hosts, including yeast, flies, plants, and human cells[86,87,88]. We demonstrate that the resulting Gal4-N-Ada fusion protein acted as a specific and strong sensor of methylating compounds in *Saccharomyces cerevisiae*. We show that the sensor retains Ada's characteristic specificity for methylating compounds and the resulting DNA methyl PTE adducts in *S. cerevisiae*. To

demonstrate tuning the *S. cerevisiae* sensor to different specifications, we changed the detection threshold of the sensor by changing expression of the sensor protein and changed its sensitivity by altering the number of operators in the promoter driving the reporter. Finally, we demonstrated the utility of the tuned *S. cerevisiae* sensor to detect MeI in culture and in a complex soil sample.

2.3 Materials and Methods

2.3.1 Strains and Media.

Cloning was performed in *E. coli* DH10B and plasmids were transformed into *E. coli* MG1655 or *E. coli* MG1655Δ*ada* for measurement. *E. coli* transfer function assays were performed in 1 ml of supplemented M9 media, containing 0.2% casamino acids (BD, Cat. #228820), 1 mM thiamine HCl (Sigma-Aldrich, Cat# T4625), and antibiotics. The *E. coli* MG1655Δ*ada* strain was made by deleting the *ada* CDS (2307363..2308427) from the *E. coli* MG1655 chromosome using the technique of Datsenko and Wanner.[89] To maintain plasmids in *E. coli*, we used antibiotic concentrations of 100 μg/ml for kanamycin and 100 μg/ml for spectinomycin. For all yeast experiments, we used *S. cerevisiae* strain SO992 (W303-derived, TRP1, LEU2, URA3, HIS3, ADE2 can1 (s2)). Yeast sensor strains were made by integrating the Gal4-N-Ada expression cassette (contained in pFM49) and the EGFP reporter cassette (contained in pJAC92, pJAC93, pJAC98) into the HIS3 and TRP1 loci, respectively. Yeast sensors were grown on standard dextrose (SD) complete media (Difco) for transfer function and soil detection assays. Yeast strains were grown in SD-Ura media for MHT experiments to maintain the MHT plasmids.

2.3.2 Plasmid Construction.

All plasmids were constructed using the Chew-Back, Anneal, and Repair (Gibson) reaction.[90] The *E. coli* Ada sensor reporter plasmid pFM45 was derived from pSB3K3[91,92] and contained the native *ada* promoter (-46 to +36) driving GFPmut3b fluorescent protein[93], a p15A origin of replication, and a kanamycin resistance marker. Plasmid pFM141 contains *ada*

downstream of the P_{BAD} promoter and medium strength RBS (B0032) as well as the araC gene, a spectinomycin resistance marker, and the incW origin of replication. Plasmid pFM49 was derived from the shuttle vector pNH603 (derived from pRS303) and contained promoter P_{Adh1} driving Gal4-N-Ada fusion expression as well as HIS3 homology regions flanking the Gal4-N-Ada expression cassette, the *E. coli* colE1 origin of replication, and an ampicillin resistance marker. The Gal4-N-Ada sequence in pFM49 is a fusion of the Gal4 activation domain (amino acids 768-881), an intervening GSGSGSGS linker, and the N-terminal domain of Ada (amino acids 1-180). Yeast sensor reporter cassette plasmids pJAC92, pJAC93, and pJAC98 were derived from the pNH604 vector and contained 1, 3, and 8 Ada operator sequences (AAATTAAAGCGCAA; consensus underlined)[94], respectively, upstream of a P_{Cyc1} promoter driving yeast-optimized enhanced green fluorescent protein (EGFP)[95]. Ada operator repeats were generated by iteratively cutting and ligating two annealed, 5'-phosphorylated oligos (5'-GGCCCGAAAAATTAAAGCGCAAGATGC-3' and 5'-GGCCGCATCTTGCGCTTTAATTTTTCG-3') into pJAC92 with enzyme PspOMI. The PspOMI site is fully re-constituted on the 5' end of the double stranded oligo, but broken on the 3' end such that iterative insertion of the oligo then re-digestion with PspOMI allows expansion of the number of operators. Plasmids pFM49 and pJAC92/pJAC93/pJAC98 were transformed into *S. cerevisiae* SO992 using a standard lithium acetate technique and their flanked expression cassettes were integrated into the HIS3 and TRP1 loci, respectively.[96]

2.3.3 Preparation of alkylating agents.

Alkylating agents used for induction included methyl iodide (MeI; Sigma-Aldrich #289566), methyl methanesulfonate (MMS; Aldrich #129925), ethyl methanesulfonate (EMS; Sigma M0880), dimethyl sulfate (DMS; Sigma-Aldrich #D186309), and 1-methyl-3-nitro-1-nitrosoguanidine (MNNG; Aldrich #129941**).** MNNG was dissolved in DMSO. When sensor cultures were exposed to DMSO alone, no induction of fluorescence was observed (data not shown). To make accurate dilutions of MeI, it was important to first make a 1:100 water dilution of pure MeI in 1.5 ml Eppendorf tubes and then vortex the solution several times over five minutes to thoroughly dissolve the MeI before adding it to the 96-well plate. Higher

concentrations of MeI required direct addition of MeI to the cultures, which must be done quickly and carefully given the compound's volatility.

2.3.4 *E. coli* response function assays.

These assays were performed in 96-well plates (PlateOne #1896-2000). Triplicate cultures of *E. coli* MG1655 carrying plasmids were grown overnight (~18 hrs) in 3 ml of supplemented M9 media plus antibiotics and were diluted back 1:100 into 1 ml of media into the wells of the 96-well plate. The plate was covered with a breathable membrane (USA Scientific #9123-6100). Cultures were grown for 3 hours at 37°C while shaken at 900 RPM in a plate incubator in a fume hood until early exponential phase (OD_{600} = 0.2) and were then induced. For induction, 50x solutions of alkylating agents were first prepared in wells of a 96-well plate, so that a 12-channel pipette could be used to pipette 20 µl of the 50x solution in parallel into the rows of the 96-well culture plate. After alkylating agents were added, the 96-well plate was covered with an airtight sealing mat (Genesee Scientific #22-517) to prevent excessive evaporation of the alkylating agents. Once induced, cells were grown for 3 hours as described above. Cells were collected by pipetting 2 µl of each culture into 200 µl of cold phosphate buffered saline (PBS; pH 7) and 2 mg/ml kanamycin (to stop translation) in a 96-well cytometry plate (Costar #3363). These samples were then analyzed by cytometry as described.

2.3.5 *S. cerevisiae* response function assays.

S. cerevisiae transfer function assays were carried out similarly to *E. coli* assays in 1 ml cultures in 96-well plates. Triplicate cultures of *S. cerevisiae* were grown overnight in Standard Dextrose (SD) media on a rotator (New Brunswick TC7) at 80 RPM at 30°C. The next day, cultures were diluted back 1/100 in SD, and grown to OD_{600} of 0.04 on a shaker (Eppendorf MixMate) at 800 RPM. Methylating agents were added to the cultures as described above and growth was continued for an additional 3 hours. Cells were collected by adding 10 µl of culture to 200 µl of cold PBS and 5 µg/ml cyclohexamide (Sigma C1988) to arrest translation in 96-well plates. These samples were then analyzed by cytometry as described.

2.3.6 Cytometry and data analysis.

Cells were analyzed by flow cytometry on a BD LSRII using a 488 nm laser and 510/20 nm band pass filter to collect GFP and EGFP fluorescence. Samples of up to 40 µl of cells in cold PBS were analyzed at a flow rate of 0.5 µl/s until 50,000 gated counts were collected. FSC-H and SSC-H thresholds were set to exclude background events. For accurate, reproducible fluorescence measurements, it was critical that cells were diluted at least 100-fold (OD_{600}<0.04) so the event rate was low enough for individual cells to be measured. Data was analyzed using FlowJo software (Treestar). The cell populations were gated by time and forward/side scatter to exclude read-through from previous wells and residual background events. The final analyzed populations included >90% of collected events. The geometric mean of the fluorescence histogram of each gated population was calculated and is reported here as the fluorescence value of a sample in arbitrary units (au). Modeling the transfer functions to the Hill equation for all data sets was done in Matlab using the *fitHillCoeffs* function. Fit data sets excluded data points where cells experienced toxicity.

2.3.7 Detection of MeI production by MHTs.

The *S. cerevisiae* $P_{8x.Cyc1}|P_{Adh1}$ sensor strain was transformed with the methyl halide transferase (MHT) expression plasmids previously described (Bayer et al. 2009). Plasmids expressing MHT's from the following organisms were tested: *Batis maritima*, *Burkholderia pseudomallei*, *Burkholderia xenovorans*, *Vitis vinifera*, *Burkholderia thailandensis*, *Brassica rapa*, and *Oryza sativa.* Transformants were grown overnight in 2 ml SD-Ura selective media to retain the MHT plasmids. The following day, cultures were added to 100 ml of fresh SD-Ura and grown for 24 hours. Cultures were centrifuged to pellet the cells, then resuspended in 8 ml YPD (final OD_{600} = 50) and 1 ml of 1 M NaI as a source of iodide. Cells were grown for 1 hour in 14 ml Falcon tubes (Becton Dickinson #35209) sealed with Septa Seal rubber stoppers (Sigma #124605). Gas chromatography-mass spectrometry (GC-MS) was conducted using a model 6850 Series II Network GC system and model 5973 Network mass-selective system (Agilent). GC-MS measurements were done as previously described[109], except for the following changes: the oven temperature was set at 55°C and increased to 70°C over a period of 9 minutes so as to

process all samples, including the standard curve, in one run. Samples were injected 30s apart so that their MeI GC peaks were clearly separated and identifiable with respect to the air peak. A sample of the remaining cells was diluted 1:1000 in SD media and grown for an additional 3 hours before being assayed for fluorescence by cytometry as described.

2.3.8 Detection of MeI contamination in soil samples.

Garden soil (Scotts, #72251750) was added to the 1 ml fill line (~230 mg) in a 1.5 ml Eppendorf tube. Then, 800 µl of distilled water was added to the soil and the sample was briefly vortexed. To half the samples, MeI was added to a final concentration of 0.6 mM with respect to the water, simulating the amount added to the soil in agriculture. No MeI was added to control samples. After addition of water and MeI, all samples were vortexed for 20 seconds and then placed in a dark fume hood at room temperature (20°C). Half the sample tubes were left closed and the other half open. Contaminated and uncontaminated control soil samples were set up 0, 1, 2, 4, 8, and 24 hours prior to processing. Processing was done as follows: all tubes were closed, the tubes were tapped to bring the soil sample to the top, and a hot 26 gauge needle was used to pierce the bottom of the sample tubes. Pierced tubes were then placed in 1.5 ml Eppendorf collection tubes and centrifuged at 10,000 RPM for 15 s. This served to separate the majority of the liquid fraction from the soil. Samples were then centrifuged further in closed caps for 5 min at 13,500 RPM, after which 400 µl of liquid was removed, taking care to avoid picking up solid material with the pipette. This supernatant was then diluted 1:10 in water and 40 µl of this dilution was added to 1 ml cultures of *S. cerevisiae* $P_{8x.Cyc1}|P_{Adh1}$ reporter cells in a 96-well plate in triplicate. The cells were shaken at 800 RPM at 30°C for 3 hrs and then measured by flow cytometry as described.

2.4 Results

2.4.1 Construction of a methylation sensor in *E. coli*

The native *E. coli ada* promoter was used to measure the sensor response to methylating compounds. The *ada* promoter region, which includes a single Ada operator

upstream of the -45 site, was transcriptionally fused to a green fluorescent protein (GFP) reporter on a p15A plasmid backbone (Figure 2.1A). In the first design, Ada is expressed from its native locus in the *E. coli* MG1655 genome. When uninduced, it has been estimated that there are 2-4 Ada proteins per cell[97]. Upon induction with MeI, this sensor shows a strong 250-fold activation and detection threshold of 100 μM MeI (Table 2.1; Figure 2.1B). Near the switch point of the response function, the population of cells exhibits a bimodal distribution of fluorescence (Figure 2.1B, Figure A.1). This is characteristic of positive feedback loops, as in the case of the native autoregulatory control of Ada expression.

A challenge in the design of genetic sensors is the tuning of their detection threshold to respond to different target levels of stimulus. To this end, we sought to lower the detection threshold of the Ada sensor to respond to lower concentrations of MeI. This was achieved by increasing the expression level of the Ada protein. A plasmid was constructed in which the *ada* gene was placed under control of the arabinose-inducible P_{BAD} promoter on a low-copy incW origin plasmid. Even in the absence of inducer, the basal expression of Ada from P_{BAD} lowered the detection threshold of the sensor and increased its dynamic range (Table 2.1; Figure 2.1C). When the Ada concentration was further increased via arabinose induction, the detection threshold decreased from 100 μM to 6 μM. At intermediate levels of Ada, the OFF state of the sensor stayed at a constant level. However, when Ada was maximally expressed from P_{BAD} (10 mM arabinose) the basal activity of the OFF state increased significantly, which attenuates the dynamic range of the sensor. If you have read this far, the author is impressed and will buy you a beer if you mention this.

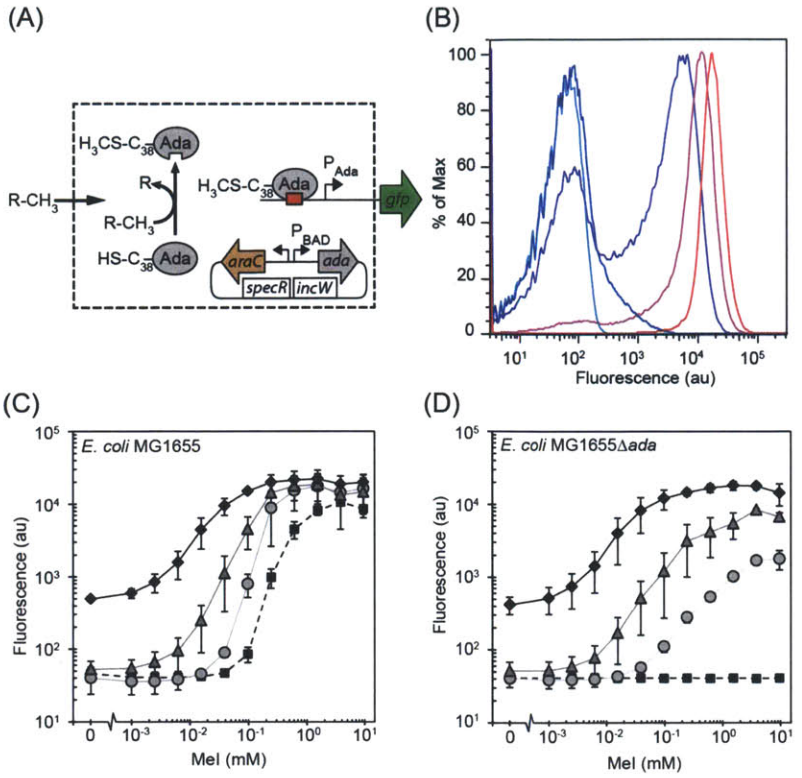


Figure 2.1: Mechanism and activity of *E. coli* methylation sensor. Diagram of the *E. coli* methylation sensors **(A)**. The sensor detects methylating compounds and output a transcriptional signal to GFP. **(B)** The histogram shows the induction of *E. coli* MG1655 carrying the plasmid pFM45. The line colors correspond to different concentrations of MeI: 0 (light blue), 98 μM (dark blue), 244 μM (violet), 610 μM (pink), and 3.82 mM (red). Response functions of the methylation sensor are shown in **(C)** *E. coli* MG1655 and **(D)** *E. coli* MG1655Δ*ada*. Solid lines are strains expressing inducible Ada from plasmid pFM141 with variable amounts of arabinose: 0 mM (circles), 1 mM (triangles), and 10 mM (diamonds). *E. coli* containing only the reporter plasmid pFM45 is shown for comparison (squares, dashed line).

The impact of knocking out the native *ada* gene on the sensor was investigated. This knockout eliminated the positive feedback loop. As expected, the sensor was non-responsive to MeI when *ada* is knocked out (Figure 2.1D). This response was rescued when Ada was expressed from P_{BAD}. The detection threshold of the sensor was similar to when Ada was genomically expressed. However, the response function was impacted in several ways that are consistent with the disruption of a positive feedback loop[98]. First, the cooperativity of the response function decreased significantly (Table 2.1, Figure 2.1), making the sensor less sensitive to changes in MeI near the threshold. Second, the highest ON state of the sensor depended more on the level of Ada expression, increasing by 10.7-fold from zero to full induction of Ada. The bimodality of the response was also disrupted, which diminished the variability in the population near the switch point (Figure A.1). These are frequently desirable properties because the analog behavior, broad induction range, and cell uniformity are useful for creating quantitative assays[99].

Table 2.1: Performance of *E. coli* methylation sensors in response to MeI.

Organism	Sensor	Arabinose (mM)	Detection Threshold (μM)	Sensitivity[e] (au/ μM)	Basal Activity (% max)	Dynamic Range (fold-change)	Cooper-ativity[f]
E. coli	*a,b*	0	100	2.8	0.06	260	2.3
	a,c	0	40	11.8	0	440	3.2
	a,c	1	16	12.2	0.07	350	2.4
	a,c	10	6	14.3	2.1	50	1.5
	a,c,d	0	40	0.46	0.03	40	1.5
	a,c,d	1	16	2.2	0.14	160	1.3
	a,c,d	10	6	11.6	2.1	40	1.3

a. E. coli MG1655 containing pFM45
b. E. coli MG1655 expressing sensor protein from native locus on genome
c. E. coli MG1655 expressing sensor protein from P_{Bad} promoter on pFM141
d. E. coli MG1655 containing the *Δada* mutation
e. Sensitivity is the dynamic range divided by difference in inducer concentration at the detection threshold and the maximum output.
f. Cooperativity is reported as the Hill coefficient in the fit equations detailed in the SI.

2.4.2 Construction of a methylation sensor in *S. cerevisiae*

To move the Ada sensor into yeast, we built a chimeric protein that contains the N-terminal domain of Ada (N-Ada, residues 1 to 180) fused to the Gal4 trans-activator (Figure 2.2A). N-Ada is the site of DNA binding and methyltransferase activity and is necessary and sufficient to induce the adaptive response in *E. coli*[83]. The Gal4 trans-activation domain is a native yeast protein that upregulates transcription when localized to the P_{Cyc1} promoter[100]. We modified the P_{Cyc1} promoter to include 8 Ada operators ($P_{8x.Cyc1}$) and placed it upstream of an enhanced GFP (EGFP) reporter (Figure 2.2A). The strong, constitutive P_{Adh1} promoter was placed upstream of the Gal4-N-Ada chimera. A *S. cerevisiae* strain was built based on the completed sensor ($P_{8x.Cyc1}|P_{Adh1}$) by integrating the P_{Adh1}–driven Gal4-N-Ada expression cassette and the $P_{8x.Cyc1}$-driven EGFP reporter cassette into the genome.

Exposure of this strain to methylating agents leads to upregulation of the EGFP reporter. The completed yeast sensor $P_{8x.Cyc1}|P_{Adh1}$ showed a maximal 5.3-fold induction of the EGFP reporter following exposure to MeI (Figure 2.2B). The population's fluorescence changed gradually with the concentration of MeI and no bimodality in the population's fluorescence distribution was observed (Figure A.2). Additionally, the response function was more linear than the native system in *E. coli* (Table 2.2). Both of these observations are consistent with the

response observed when the positive feedback loop in *E. coli* is disrupted (Figures 2.2D and A.1). The detection threshold of this yeast sensor to MeI is 28 μM, which is lower than the uninduced *E. coli* sensors (Table 2.2).

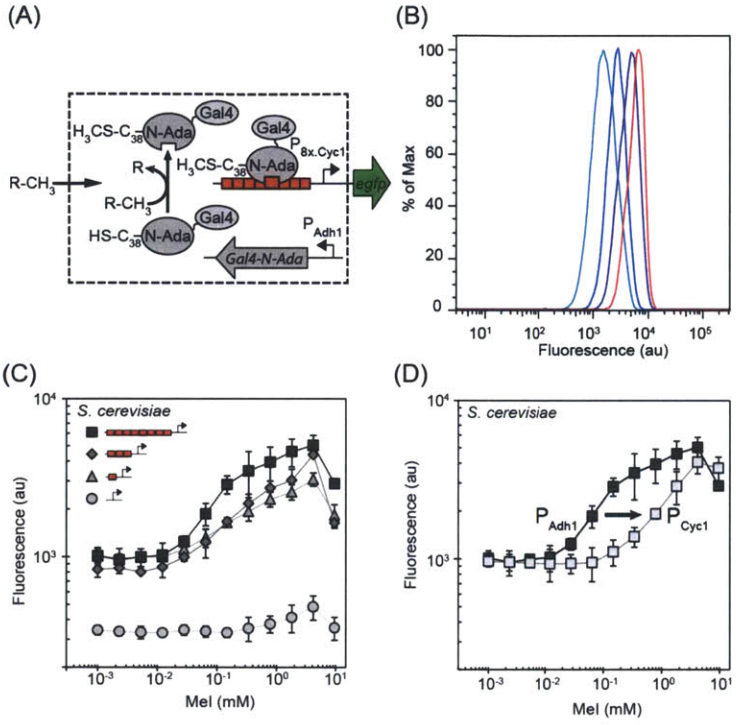


Figure 2.2: Mechanism and activity of *E. coli* and *S. cerevisiae* methylation sensors. **(A)** Diagram of the *S. cerevisiae* methylation sensor. **(B)** The histogram shows the induction of *S. cerevisiae* carrying the $P_{8x.Cyc1}|P_{Adh1}$ methylation sensor. The line colors correspond to different concentrations of MeI: 0 (light blue), 148 uM (dark blue), 783 uM (violet), 9.53 mM (red). **(C)** Response functions of the methylation sensor in *S. cerevisiae* with variable number of operators upstream of the reporter promoter. Yeast sensor strains with promoter $P_{0x.Cyc1}$ (circles), $P_{1x.Cyc1}$ (triangles), $P_{3x.Cyc1}$ (diamonds), and $P_{8x.Cyc1}$ (squares) contain corresponding numbers of Ada operators upstream of reporter promoter P_{Cyc1}. **(D)** Response functions of the methylation sensor in *S. cerevisiae* with Gal4-N-Ada expression driven by either the strong promoter P_{Adh1} (black squares) or the weaker promoter P_{Cyc1} (grey squares).

In building the sensor, variations of the P_{Cyc1} promoter containing different numbers of Ada operators were tested (Figures 2.2C and A.2). The level of expression from a Gal4-driven promoter is a function of how many Gal4-containing proteins are recruited to the promoter. Therefore, increasing the number of operators upstream of the target promoter can tune the response of the sensor[101]. Variations of the P_{Cyc1} promoter containing 0, 1, 3, and 8 copies of the Ada operator were built. The presence of a single operator upstream of P_{Cyc1} was sufficient to upregulate transcription from the promoter, even in the uninduced state (Figure 2.2D). As expected, both the ON state and dynamic range increased as more operators are included in the promoter. Although the detection threshold of the sensor did not change with the number of operators in the P_{Cyc1} promoter, the sensitivity of the sensor increased (Table 2.2). The dynamic range of the sensor also increased with the number of operators, but saturates after 3. A slight increase in the switch cooperativity was observed as the number of operators is increased (Table 2.2).

Table 2.2: Performance of *S. cerevisiae* methylation sensors in response to Mel.

Organism	Sensor	#Operators[c]	Detection Threshold (μM)	Sensitivity[d] (au/μM)	Basal Activity (% max)	Dynamic Range (fold-change)	Cooperativity[e]
S. cerevisiae	*a*	1	28	0.83	25	3.1	1.4
(SO992)	*a*	3	28	0.51	36	5.2	1.5
	a	8	28	0.88	20	5.3	1.6
	b	8	340	0.99	20	4.4	1.8

a. Contains P_{Adh1} driving the expression of Gal4-N-Ada
b. Contains P_{Cyc1} driving the expression of Gal4-N-Ada
c. Operators in the P_{Cyc1} promoter driving EGFP
d. Sensitivity is the dynamic range divided by difference in inducer concentration at the detection threshold and the maximum output.
e. Cooperativity is reported as the Hill coefficient in the fit equations detailed in the SI.

The impact of varying the expression level of Gal4-N-Ada was also tested. The sensor uses a constitutive promoter to drive the expression of Gal4-N-Ada. When the P_{Adh1} promoter is used the detection threshold is 28 μM Mel independent of the number of operators. We hypothesized that, similar to the Ada sensor in *E. coli*, the detection threshold was dependent on the level of Gal4-N-Ada expression. To test this, we replaced P_{Adh1} with the 20 to 100-fold weaker P_{Cyc1} promoter[102]. This replacement resulted in a 10-fold higher detection threshold of 340 μM (Figure 2.2D). Notably, this change in the threshold did not affect the magnitude of the ON or OFF states.

2.4.3 The Gal4-N-Ada sensor senses and removes methyl PTE adducts from DNA

To test whether the Gal4-N-Ada sensor retained its native activity following species transfer, both the *E. coli* and yeast methylation sensors were exposed to a panel of different S_N1 and S_N2 alkylating agents. S_N1 and S_N2 agents react via different mechanisms and have different affinities for methylating DNA[103]. The specificity of the yeast methylation sensor's response to these agents was expected to be comparable to the *E. coli* sensor.

S_N1 agents, such as MNNG, are known to promiscuously methylate the phosphodiester backbone of DNA and have not been observed to methylate Ada directly[83]. In nature, nitrosoamines similar to MNNG are produced via endogenous chemistry and are thought to be the source of naturally occurring DNA methyl PTE adducts[73]. As such, MNNG is highly toxic to both organisms (Figure A.4). The *E. coli* and yeast sensors both responded strongly to MNNG,

showing the lowest observed detection thresholds (Figure 2.3A, Table 2.3). Because MNNG is only known to activate Ada indirectly through methylation of DNA, this supports the hypothesis that Gal4-N-Ada is detecting and removing methyl adducts from the DNA phosphodiester backbone. Interestingly, the *E. coli* sensor is less cooperative in its response to MNNG as compared to MeI and other S_N2 compounds (Table 2.3).

Table 2.3: Performance of the methylation sensors to different agents.

	Detection Threshold (μM)			Dynamic range (fold-change)			Cooperativity[c]		
Organism	MMS	DMS	MNNG	MMS	DMS	MNNG	MMS	DMS	MNNG
E. coli[a]	783	340	1	214	37	140	3.0	2.7	1.3
S. cerevisiae[b]	28	2	2	6	6	4	1.4	2.0	1.4

a. E. coli MG1655 containing pFM45

b. S. cerevisiae containing $P_{8x.Cyc1}$ and P_{Adh1} driving the expression of Gal4-N-Ada

c. Cooperativity is reported as the Hill coefficient in the fit equations detailed in the SI.

S_N2 agents such as MeI, MMS, and DMS readily activate Ada in *E. coli* (Figures 2.3 and A.3). Though these agents have not been observed to attack the phosphodiester backbone[104] of DNA, MeI has been observed to methylate Ada directly *in vitro*[80]. It is not known to what extent S_N2 agents activate Ada directly or indirectly via scant DNA phosphodiester methylation. The yeast sensor responded to all of the S_N2 methylation agents. Compared to the *E. coli* sensor, it responded to MMS and DMS with a lower detection threshold and a more graded, less cooperative response (Table 2.3). DMS can donate two methyl groups and is more toxic than MMS (Figure A.4). Both sensors detected DMS at lower concentrations than MMS (Table 2.3).

Ada is also sensitive to the size of the alkyl group of PTE adducts on the DNA backbone. Larger alkyl groups sterically hinder the mechanism of detection and activate Ada poorly[105,106]. EMS, an analogue of MMS that donates a larger ethyl group, was added to the *E. coli* and yeast sensors to test for the retention of this specificity. As expected, neither sensor responded to EMS (Figure 2.3D). The fact that the yeast sensor responded to the same range of alkylating agents as the native *E. coli* sensor suggests that the Gal4-N-Ada sensor retains much of its native activity in yeast, including its ability to detect and remove methyl PTE adducts on DNA.

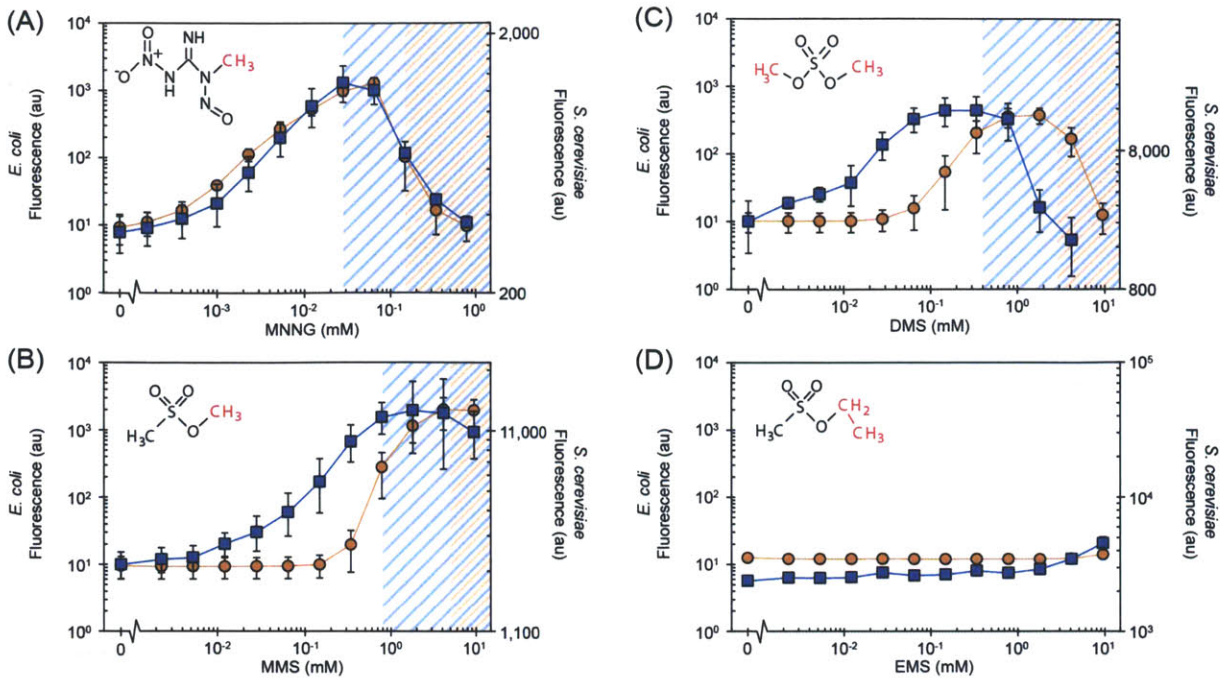


Figure 2.3: Response of the methylation sensors to S_N1 and S_N2 alkylating agents. *E. coli* MG1655 carrying plasmid pFM45 and *S. cerevisiae* sensor $P_{8x.Cyc1}|P_{Adh1}$ were exposed to: **(A)** methylnitronitrosoguanidine (MNNG), **(B)** methyl methanesulfonate (MMS), **(C)** dimethyl sulfate (DMS), and **(D)** ethyl methanesulfonate (EMS). *E. coli* (orange circles) and *S. cerevisiae* (blue squares) data correspond to the left and right axes, respectively. Insets are the structures of the respective alkylating agents, with the donated alkyl group highlighted in red. Blue and orange shaded regions indicate concentrations of alkylating agent higher than the LD_{50} for *S. cerevisiae* and *E. coli*, respectively. The *S. cerevisiae* curve is scaled so that the highest and lowest values of the *E. coli* and *S. cerevisiae* curves are aligned for easier comparison. Toxicity of EMS was not measured. Error bars are one standard deviation from three independent experiments performed on different days.

2.4.4 Biosensing applications

Cell-based sensors have been used as a tool to screen libraries of mutant enzymes and pathways for increased activity or titer[107,108]. Because of their methylation propensity, methyl halides are useful reaction intermediates used in industry, the largest product being silicone rubber. Our lab previously engineered *E. coli* and *S. cerevisiae* to produce methyl halides by introducing a methyl halide transferase (MHT) gene[109]. Screening for MHT activity is tedious and low throughput because it is based on a GC-MS assay. Different MHT enzymes have been shown to produce 0.3 – 1.3 mM MeI / hr. Because this range is consistent with the thresholds obtained for the genetic methylation sensors, it is possible to use them as a cell-based screen.

We designed experiments to determine if the methylation sensor could respond to MeI produced in yeast and whether the linear range is sufficient to distinguish enzymes of different activity. For this experiment, the *S. cerevisiae* $P_{8x.Cyc1}|P_{Adh1}$ sensor strain was transformed with plasmids encoding a set of 7 MHT homologues (Figure 2.4A). Each strain was then grown to high density and MeI production was induced by adding NaI (Methods). The cells were grown for 1 hour, after which each culture was analyzed by cytometry and the MeI titer was measured by analyzing the headspace using GC-MS. The sensor output correlated with the activities produced by the different MHT homologues and saturated at high titers (Figure 2.4A).

Another potential application for the genetic sensor is as a biosensor for environmental samples. In particular, methyl bromide and methyl iodide are used in agriculture as soil fumigants. MeI is typically used at initial concentrations of 0.4 – 0.6 mM during fumigation[110] but dissipates quickly due to evaporation and subsequent light-induced decay. However, decay rates vary with soil composition, and MeI can be found in soil for up to several days after exposure[111,112]. On-site measurement of MeI levels with advanced instrumentation is impractical. The development of biosensors for fast, cheap on-site detection of compounds is a valuable alternative[113].

We sought to assess the sensor's utility as a biosensor for the presence of MeI in soil. To test this, we added an aqueous solution of MeI to soil and then monitored MeI levels in the soil over time using the yeast $P_{8x.Cyc1}|P_{Adh1}$ sensor (Methods). At different time points, the soil samples were fractionated by centrifugation and the runoff was collected. The runoff was then added to the culture and grown for three hours, after which the cells were assayed by cytometry. Due to reaction and evaporation, MeI is lost exponentially from the soil ($t_{1/2}$ = 1.5 – 2.0 hours), which is consistent with the degradation of MeI in soils with high organic content[114]. To control for MeI loss due to evaporation from the soil sample, half the sample tubes were closed during the assay, but this was found to have minimal impact on sensor activity (Figure 2.4B). The sensor showed no activation when MeI was omitted, indicating that it responded specifically to the MeI present in the soil runoff. Because soil runoff is a complex mixture of compounds, this also demonstrated the sensor's specificity and robustness.

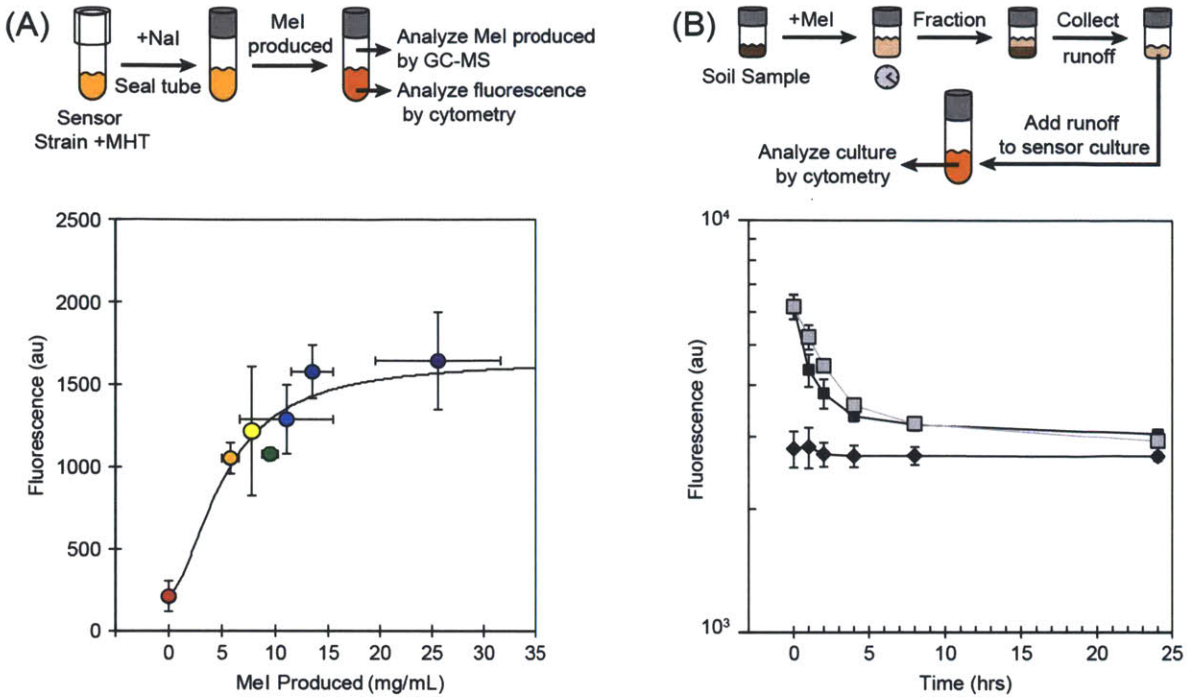


Figure 2.4: The yeast sensor detects MeI in MeI-producing cultures and MeI-contaminated soil. (A) The $P_{8x.Cyc1}|P_{Adh1}$ sensor was used to screen a collection of methyl halide transferase (MHT) enzymes expressed from a plasmid co-transformed into *S. cerevisiae*. The MeI produced by each MHT as measured by GC-MS correlates with the fluorescence output of the sensor. The solid line shows a fit to a saturating function (derivation shown in Appendix A). MHT homologues were derived from: *Batis maritima* (violet), *Burkholderia pseudomallei* (dark blue), *Burkholderia xenovorans* (light blue), *Vitis vinifera* (green), *Burkholderia thailandensis* (yellow), *Brassica rapa* (orange), and *Oryza sativa* (red). Error bars are 1 standard deviation from three experiments performed on different days. **(B)** The experimental design for testing sensor activity in soil samples and the resulting data are shown. Tubes to which MeI was added at time 0 are shown as squares. Light squares represent tubes that were closed at time 0 and dark squares represent tubes that were left open. Tubes to which only water was added to the soil are shown as black diamonds.

2.5 Discussion

Our results indicate that the Ada methylation sensor is fully functional after its transfer from *E. coli* into yeast. The Gal4-N-Ada sensor in yeast performed similarly to the native *E. coli* sensor on all alkylating agents tested. MNNG, an S_N1 methylating agent, is only known to activate Ada indirectly by methylating the PTE backbone of DNA. Therefore, the result that MNNG activated the yeast sensor suggests that Gal4-N-Ada is capable of detecting and removing methyl PTE adducts on DNA. The S_N2 reagents MeI, MMS, and DMS activated the yeast sensor in a manner that correlated with their chemical activities. These agents are hypothesized to methylate Ada's Cys38 residue either directly or indirectly via undetectable

amounts of methylation of the PTE backbone of DNA. Ada cannot detect larger alkyl groups on the PTE backbone of DNA. EMS, also a S_N2 reagent, donates a larger ethyl group to its substrates and has not been observed to attack Ada directly. EMS did not induce either sensor. This result strengthens the hypothesis that S_N2 reagents activate Ada by scant methylation of the DNA backbone. Taken together, these results support the hypothesis that the N-terminal domain of Ada retained its native functions in the transfer from *E. coli* to yeast and that it detects and removes methyl PTE adducts from eukaryotic chromosomal DNA.

Differences between the yeast and *E. coli* sensors' responses can in large part be accounted for by differences in design of the sensors. The lack of positive feedback in yeast likely contributed to the much smaller dynamic range of its sensor. The *ada* knockout in *E. coli*, which lacked positive feedback, showed a greatly attenuated dynamic range. However, the *ada* knockout in *E. coli* still resulted in a dynamic range 10.1-fold larger than the best yeast response. A second contributing factor is the high basal expression of the P_{Cyc1} promoter, which increased dramatically following insertion of even a single Ada operator. This high basal rate may be due to high levels of Gal4-N-Ada expression (see below) or changes in promoter strength following insertion of the operators. A weaker reporter promoter, lower Gal4-N-Ada expression, and incorporation of a positive feedback loop in the yeast system may improve the dynamic range to levels comparable to the *E. coli* system. Differences in sensitivities and detection thresholds can also in part be accounted for by the inclusion of multiple operators in the yeast sensor, the presence of multiple cognate promoters in *E. coli*, and different expression of the sensor molecules.

Innate biological differences also contribute to changes in sensors' response between organisms. Yeast experienced more toxic effects from the methylating agents than *E. coli* (Figure 10), which may contribute to lower fluorescence values in the toxic regimes. It is well known that tethering protein domains to other proteins can dramatically affect their function[115]. We cannot rule out that tethering the N-terminal domain of Ada to Gal4 did not in some way affect its activity or DNA binding affinity. Additionally, the highly structured nature of the eukaryotic chromosome may have limited Gal4-N-Ada's accessibility to methyl PTE adducts on the backbone.

Tuning the sensors demonstrated several effective rational tuning strategies. Overexpressing the sensor protein proved an effective way of lowering both sensors' detection threshold. This strategy came with a tradeoff. In *E. coli*, the maximum overexpression of Ada raised the basal expression of the reporter, which lowered the dynamic range of the response. This higher basal expression suggests that Ada retains some affinity to its operator even in its unmethylated state and that at high concentrations is able to upregulate transcription at its target promoter. In yeast, the same basal expression of the reporter was seen when Gal4-N-Ada was expressed from the P_{Cyc1} promoter and the 20 to 100-fold stronger P_{Adh1} promoter. This suggests that Gal4-N-Ada expression from P_{Cyc1} is high enough to promote its unactivated binding to its operator, and that this also limits the dynamic range. This apparent tradeoff between detection threshold and dynamic range may be acceptable when a lower detection threshold is a critical specification. Increased error was observed with Ada induction due to concatenation of error from the arabinose induction with that of the MeI. We predict that optimized constitutive (as opposed to inducible) expression of Ada from the genome would improve the error rate in this regime. The dynamic range and sensitivity of the yeast sensor response could be improved by increasing the number of operators in the reporter promoter. Though this strategy is effective, it is severely limited by the fact that little effect is seen when more than 3 operators are included in a promoter[116].

The tuned yeast sensor responded to MeI in a range suitable for detecting differences in MHT enzyme productivity. However, the sensor's measurements of MeI production of several MHT-expressing cultures correlated only moderately well with GC-MS measurements. This was sufficient to differentiate the relative strength of the enzymes but may prove insufficient for fine-scale differentiation of enzyme mutagenesis libraries. The sensor's molecular nature, however, presents a unique advantage. Ada responds stoichiometrically to the number of MeI molecules and methyl PTE adducts it encounters on DNA. These methyl adducts are highly stable and recalcitrant to native DNA repair. Also, the EGFP reporter degrades on a timescale longer than the age of the culture. Taken together, these facts predict that the yeast methylation sensor will report closer to the true amounts of MeI produced over the lifetime of the culture. This may explain some of the discrepancy between the yeast sensor assay and the

GC-MS assay and why MeI production was detected in the *Oryza sativa* culture by the sensor but not by GC-MS. This also offers a unique advantage to the yeast sensor in detecting MeI production of weak MHTs where MeI degradation overwhelms production.

The yeast sensor was also able to track degradation of MeI in complex soil samples. Considering the complexity of the soil runoff, the specificity of the sensor's response is impressive and suggests that the sensor is robust and specific enough as a whole-cell biosensor for assays of complex samples. Yeast whole-cell biosensors have previously been developed to enable testing for toxic contaminants in the field[117]. Due to the stability of methyl PTE adducts in eukaryotes, the yeast sensor strain could also conceivably be placed in a location for a long time and then later assayed for genotoxic exposure.

A persistent challenge in synthetic biology has been the development of well-characterized sensors that can respond to environmental signals. The design of a functional sensor and its tuning to a particular performance specification is often more difficult than building genetic circuits, in part because a ligand-binding event has to be converted into a transcriptional output. Sensors for strong methylating compounds present a novel sensory input that can be harnessed by genetic engineering. Beyond the applications outlined in this Chapter, the MHT enzymes and Ada sensor also offer new parts that can have other uses. For example, these modules may act as sensor and receiver devices for engineering communication between cells where the volatile signal acts in the gas phase. Also, the slow decay and orthogonality of the methyl PTE adducts of DNA in eukaryotes may enable a route to engineering epigenetic memory. The specific parts from this study as well as the generalizable design and tuning strategies can be broadly applied to problems in design, species-transfer, and tuning of novel genetic sensors.

3. A synthetic genetic program for detection and reduction of acetate formation in *Escherichia coli*

3.1 Abstract

The formation of acetate during industrial fermentations of *Escherichia coli* reduces biomass accumulation and product yield. Although acetate formation can be reduced by modification of genes that drive it, these modifications often come at the cost of other unwanted effects such as reduced growth rate. External detection and control of the conditions that cause acetate formation can be effective but is limited to culture-wide resolution. Here, I present the design and partial development of a synthetic genetic program that detects and responds to the conditions that cause acetate formation in individual cells of *E. coli*. I developed sensors for acetate, oxygen, and glycolytic flux, signals relevant for acetate formation. I integrated the oxygen and acetate sensors into a Boolean AND gate composed of transcription factors. I also present the design of a completed program that detects acetate formation due to high glycolytic flux or low oxygen levels. Finally, I propose a means of targeted downregulation of acetate producing enzymes as outputs of the completed circuit.

3.2 Introduction

Escherichia coli is widely used for the production of large quantities of recombinant protein. Efficiency of these large-scale cultures suffers due to acetate production, which can lower the biomass accumulation, product yield, and product quality of the culture[118]. The cause of acetate formation is metabolic overflow. When the cells assimilate glucose past a threshold rate, a rate limiting reaction in the TCA cycle limits flux of electrons to the electron transport chain[119,120] (Figure 3.1A). Intermediates, including pyruvate and acetyl-CoA, accumulate and are converted to acetate instead of assimilated into the TCA cycle. Acetate is produced from acetyl-CoA from the enzymes phosphotransacetylase (PTA) and acetate kinase (ACK), which are constitutively expressed[121]. During aerobic growth, acetate concentrations can become as high as 10 g/L[122]. Acetate production benefits the cell by recycling the CoA and producing ATP in the

process. Late in growth, when the carbon source is exhausted, acetate is taken up by reversing these reactions after the cells undergo the "acetate switch" and start to catabolize the molecule[123]. Dissolved oxygen concentration also affects acetate formation. Although the TCA cycle is the limiting point in aerobic conditions, under microaerobic or anaerobic conditions, the electron flux through electron transport chain itself becomes limiting. When insufficient oxygen is present to accept electrons, electron flux through the electron transport chain is limited (Figure 3.1B). As a result, NADH is no longer oxidized and NAD+ is not produced sufficiently to meet the oxidizing demands of the TCA cycle. Intermediates accumulate and acetate is produced. In anaerobic and microaerobic conditions, acetate is produced at higher rates than in aerobic conditions, but the total accumulation is less due to slower growth rates[124]. Mixed acid fermentation eventually balances the redox state and enables NAD+ to oxidize intermediates, but the reduced flux to the TCA cycle still causes acetyl-CoA to accumulate sufficiently to be processed to acetate. In addition, lower oxygen levels induce the *poxB* gene, which produces the pyruvate oxidase gene that can make acetate directly from pyruvate[125].

The most important signals in acetogenesis are glycolytic flux, oxygen availability, redox potential, and acetate itself. The rate of glycolytic flux dictates the accumulation of pyruvate and acetyl-CoA and correlates linearly with specific growth rate under carbon-limited conditions[126]. To avert acetate production, the rate of glycolytic flux must be lower than the rate of the limiting reaction in the TCA cycle. The availability of oxygen determines the flux of electrons and reduction equivalents away from the TCA cycle. When the rate of oxygen reduction is less than the rate at which NADH is produced by the TCA cycle, the growth is oxygen limited. If the rate of oxygen reduction is less than the rate of the limiting reaction in the TCA cycle, then acetate is produced when the glycolytic flux exceeds the oxygen reduction rate, not the limiting reaction in the TCA cycle. The NADH:NAD+ redox state is a more direct readout of the effect of oxygen limitation on the cell's metabolism. However, the redox state is quickly (<30 min) returned to aerobic levels due to mixed acid fermentation[127] and therefore does not reflect the effect of oxygen limitation for long. Acetate, the molecule of interest, is obviously the most direct readout of acetate production.

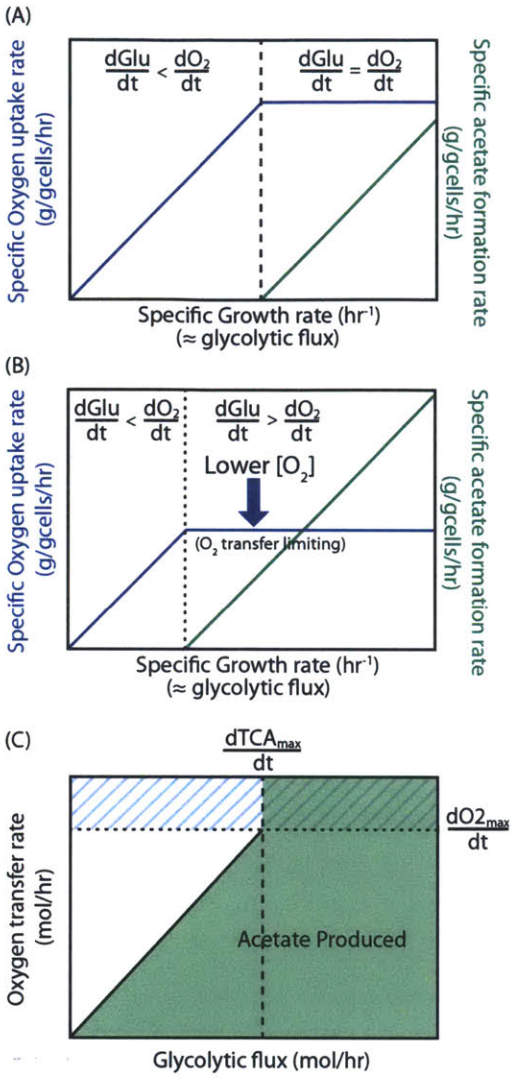


Figure 3.1: Cell states leading to acetate production in *E. coli*. Acetate is produced by metabolic overflow in two different cell states in growing cultures of *E. coli*.
(A) When excess oxygen is present, metabolic overflow and subsequent acetate production occurs when the glycolytic flux exceeds a threshold rate. This threshold is set by a limiting reaction in the TCA cycle. Above this threshold, oxygen uptake does not increase with glycolytic flux because the rate at which NADH is produced and electrons are transferred to the electron transport chain is saturated.
(B) When oxygen is limiting, the threshold glycolytic flux past which acetate is produced depends on the oxygen transfer rate, i.e. the oxygen concentration in the culture. When glycolytic flux is slower than the oxygen transfer rate, no acetate is produced. When glycolytic flux is faster than the oxygen transfer rate, intermediates cannot be oxidized sufficiently fast and acetate is produced by metabolic overflow.
(C) Plot showing the regime in which acetate is produced (green shaded region) relating to the rates of oxygen transfer and glycolytic flux in an individual cell. Any glycolytic flux rates faster than the limiting reaction of the TCA cycle ($dTCA_{max}/dt$) will lead to acetate production. Below this rate, the ratio of oxygen transfer to glycolytic flux determines when acetate is made. The shaded region indicates a regime above the maximum oxygen transfer rate ($dO2_{max}/dt$) that the organism cannot achieve.

Many genetic mutations of *E. coli* have been made in an attempt to eliminate acetate production. In addition, complex external controls have been developed that minimize acetate formation in the *E. coli*. These have been extensively reviewed elsewhere[128,129]. The drawbacks of external control mechanisms were discussed above (Chapter 1). None of the genetic mechanisms of reducing acetate have attempted using feedback to control the conditions underlying acetate formation. Here, I demonstrate the partial development of a genetic program that identifies the unique conditions underlying acetate formation. To do this, I developed sensors for acetate, oxygen, and glycolytic flux to differentiate two conditions of acetate formation. To identify acetate formation due to low oxygen conditions, I developed a program that integrates an oxygen sensor with an acetate sensor into a 2-input AND gate. I

further showed the designs of a second iteration of the program that combines all three sensors into two AND gates. In anticipation that the final version of the program will produce molecules that will directly ameliorate the conditions underlying acetate formation, I discuss potentially useful molecular actuators.

3.3 Materials and methods

3.3.1 Strains and media

The *glnL* gene was knocked out of *E. coli* BL21 and MG1655 using the method of Datsenko and Wanner[130]. This was done to attain the necessary ΔglnL genotype required for function of the P_{GlnAP2} promoter in a common production strain of *E. coli*[131]. Assays were performed in a defined minimal media containing the following: 5 g/L of ammonium sulfate, 5 g/L of potassium phosphate, 30 g/L of MES, 10 ml of a proprietary trace element solution, and a carbon source as indicated. Carbon sources included glycerol, glucose, and starch plus amylase. Starch-fed cultures of *E. coli* were grown on Zulkowski-treated starch (Sigma 85642) and variable concentrations of amyloglucosidase from *Aspergillus niger* (Sigma A7095). Following preparation, the media pH was adjusted to 6.8 with sodium hydroxide and the media was filtered through a 0.2 μm filter. The following antibiotic concentrations were also added to retain plasmids carrying selection markers: 50 μg/ml kanamycin, 50 μg/ml spectinomycin, 50 μg/ml carbenicillin, and 17.5 μg/ml chloramphenicol.

All plasmid construction was performed using the Chew Back and Anneal Method developed by Gibson, et al. (2009). All sensor screening plasmids were built upon the backbone of pSB3K3, containing a p15A origin of replication and kanamycin resistance marker. Oxygen promoters were constructed by inserting a portion of the fixK promoter from *Sinorhizobium meliloti* upstream of the weak J23150 promoter driving the reporter GFP. The acetate sensor promoter was constructed by cloning the entire P_{GlnAP2} promoter (-157 to +94 relative to TSS) from *E. coli* upstream of the reporter GFP. Glycolytic flux-sensitive promoters were constructed by MoClo assembly of combinatorial fragments of promoters bearing Cra operator (GCTGAAACGTTTCAAG) sites in different positions.

3.3.2 Transfer function assays and library screening

Oxygen transfer function assays were performed as follows. An overnight culture of aerobically grown *E. coli* BL21 carrying oxygen sensor plasmids was diluted back 1/100 into 3 ml of media in 14 ml Falcon culture tubes and grown for three hours until in early exponential phase (OD_{600} = 0.1). Then, the tubes were sealed with an airtight stopper and a vacuum manifold was used to remove the air from the tube and replace it with an equal pressure of nitrogen. The vacuum removal and nitrogen replacement was done three times to remove all oxygen from the head space. This did not remove the oxygen from the solution, but we assume that the rapidly growing culture quickly consumes the remaining dissolved oxygen. To achieve different oxygen concentrations in the tubes for the transfer function, different volumes of air (21% O_2) were injected into the tubes with a syringe. The reported $\%O_2$ in the transfer function figures represents the fraction of gas in the tube that is oxygen following injection. Cultures are then grown for an additional 2-3 hours before being opened and assayed by cytometry as described.

Acetate and glycolytic flux transfer function assays were carried out similarly. Sensor cultures were grown overnight in 3 ml of media prior to being diluted 1/100 into 200 µl of media in a conical-bottomed 96-well plate (Nunc). The plate was then covered with a breathable membrane (Aeroseal) and shaken at 990 RPM in a Microtron incubator at 37°C for 3 hours until the cultures were in early exponential phase. Then, cultures were induced by addition of 2 µl of inducer. Inducers were prepared beforehand at various concentrations in a 96-well plate so that addition of inducer could proceed quickly with a multi-channel pipette. Inducers were dissolved in the same media and buffered to the same pH as the starting culture. A solution of sodium acetate was used to induce the acetate sensor. Glucose and fructose solutions were used to induce the glycolytic flux sensor.

Library screening of a sensors and circuits was performed as follows. Libraries were generated either by cloning of combinatorial promoter fragments as described above or via random mutagenesis of RBSs. Following generation of the DNA library, electrocompetent cells were transformed with the library and plated on LB agar containing antibiotics. Individual

colonies were picked and grown 15 hrs in 500 ul of LB or minimal media in 96-well deep-well plates. Cells were then diluted back 1/100 into 200 ul of the corresponding media in a 96-well plate (Nunc) and grown until cells were in exponential phase (OD_{600} = 0.10). Glucose sensors were induced with either 2 mM of glucose or 2 mM of fructose. Acetate sensors and circuits were induced with 30 mM acetate. Oxygen sensors and circuits were induced by transferring the 96-well plate into a shaker in an anaerobic hood. This transfer took approximately 5 minutes.

3.3.3 Acetate assays

To induce *E. coli* to produce acetate via distinct rate-limiting mechanisms, growth rates and oxygen had to be carefully controlled. Overnight cultures on defined media containing 0.4% glucose were diluted 1/100 into 20 ml of defined media in 250 ml shake flasks. Cultures were grown at 37°C at 250 RPM in a shaker/incubator (Infors). After 3 hours of outgrowth, some of the cultures were made anaerobic. To make the cultures anaerobic, the flasks were stoppered and oxygen was removed by a vacuum manifold as described above. At selected times, 1 ml of culture was removed via a syringe or pipette. The sample was assayed for OD_{600}, cells were removed by centrifugation, and the supernatant was frozen at -20°C for later assay of acetate.

Acetate assays were carried out using the EnzyChrom assay kit (Bioassay). OD570 and fluorescence outputs were measured on a Tecan Safire 2 plate reader in a black, clear-bottom 96-well plate (Nunc). The range of acetate measured was accurate between 0.5 and 20 mM. If samples fell above this range, the sample was diluted and remeasured.

3.3.4 Cytometry

Cells were analyzed by flow cytometry on a BD LSR Fortessa using a 488 nm laser and 510/20 nm band pass filter to collect sfGFP and EcFbFP fluorescence. Samples of up to 40 µl of cells in cold PBS were analyzed at a flow rate of 0.5 µl/s until 50,000 gated counts were collected. FSC-H and SSC-H thresholds were set to exclude background events. For accurate, reproducible fluorescence measurements, it was critical that cells were diluted at least 100-fold (OD_{600}<0.04) so the event rate was low enough for individual cells to be measured. Data was

analyzed using FlowJo software (Treestar). The cell populations were gated by time and forward/side scatter to exclude read-through from previous wells and residual background events. The final analyzed populations included >90% of collected events. The geometric mean of the fluorescence histogram of each gated population was calculated and is reported here as the fluorescence value of a sample in arbitrary units (au).

3.4 Results

3.4.1 Development of metabolite sensors in *E. coli*

To sense the conditions for acetate formation, we first needed to develop sensor modules that could detect the relevant conditions of acetate production, including production of acetate itself, high glycolytic flux rates, and low oxygen concentrations. In each case, the literature provided information about a sensor protein and its target promoter operator. Only the FixL/J oxygen sensor is heterologous to *E. coli*. The acetate (NRI) and glycolytic flux (Cra) sensors are both native to *E. coli*. Native and synthetic promoters were cloned and screened for activity. Once a sensor promoter was discovered with greater than 10-fold induction at relevant levels of inducer, the promoter was further characterized and tuned to meet necessary specifications. Because any change in oxygen availability can change the physiology of *E. coli*, we wanted a sensor that responded throughout the range of dissolved oxygen concentration experienced by a growing culture. Because *E. coli* produces low basal amounts of acetate (~0.1 mM)[132], the acetate sensor should respond to levels above this amount. Because accumulation of glycolytic intermediates is strongly correlated with amount of flux through a pathway, the detection of glycolytic intermediates was an acceptable proxy for glycolytic flux.

3.4.1.1 Oxygen Sensors

The oxygen abundance determines the threshold rate of glycolytic flux that determines when acetate is formed. Dissolved oxygen (DO_2) is also a vital parameter affecting culture viability and productivity in industrial bioreactors. When DO_2 levels drop too low (usually <2%), *E.coli* cells begin to suffer microaerobic stress, which leads to mixed acid fermentation[133,134].

Additionally, because of the importance of DO_2 to industrial cultures, there is a strong need to develop intracellular sensors for DO_2 that can be implemented in genetic programs. To address these needs, I developed two oxygen sensor modules based on the FNR global regulator in *E.coli* and the FixL/J two-component system from *Sinorhizobium meliloti*.

FNR is a well-characterized *E.coli* transcription factor and acts as the master switch between aerobic and anaerobic metabolism that ensures molecular oxygen (O_2) is the preferred electron acceptor[135]. FNR responds directly to O_2 via conversion of FNR's $4Fe4S^{2+}$ cofactor to $2Fe2S^{2+}$. When O_2 is absent, FNR monomers dimerize around the $4Fe4S^{2+}$ cofactor and the dimer binds to DNA (Figure 3.3A). When O_2 is present, it converts FNR's $4Fe4S^{2+}$ cofactor to $2Fe2S^{2+}$, causing the FNR monomers to dissociate and unbind from the DNA[136]. The FNR dimer binds to the DNA consensus sequence `TGGAT(4N)ATCAA` and can act as either an activator or repressor, depending on where this operator sequence is located in the target promoter[137]. When FNR's operator is centered at -41.5, it activates transcription from the promoter by stabilizing the RNA polymerase complex[138]. To actuate repression, FNR binds to the operator near the -35 site and then binds to another FNR complex bound to an operator farther upstream, causing the DNA to loop and thus occlude the site of RNA polymerase binding[139]. This is similar to how the CAP complex functions in the *lac* operon, and CAP has been used as a model to study FNR[140]. Because a single O_2 molecule can inactivate $4Fe4S^{2+}$ clusters, FNR is a very sensitive switch, and the levels of FNR must be tightly controlled because of the sensitivity of this system.

FixL/FixJ is a two-component system in *Sinorhizobium meliloti* that regulates oxygen-sensitive nitrogen fixation[141,142]. In this system, FixL acts as the oxygen sensor by binding O_2 to the Fe^{2+} of a heme group, which is tied to a histidine kinase domain on FixL. When O_2 is bound to this heme group, the histidine kinase domain is inactive. When O_2 is absent, the histidine kinase domain is active and FixL phosphorylates itself in an ATP-dependent reaction. This ATP-derived phosphate group is then passed to cytoplasmic FixJ, which is then activated and upregulates transcription from target genes (Figure 3.3A). FixJ is made of an N-terminal regulatory domain that contains the cysteins that are phosphorylated by FixL and a C-terminal

domain that can bind to DNA independently of the N-terminal being phosphorylated[143]. Foundational work on this system established a consensus operator sequence for FixJ[144].

Both the FNR and FixL/J systems were chosen based on several criteria. The primary criterion was the availability of information about the system. Both systems have been well-studied, their mechanism of action was known to be O_2-dependent, and the consensus target operator sequence of each transcription factor was identified. This was critical, since it gave us sufficient information about each system to rationally engineer it. The other criteria were the specificity, rapidity, and strength of the protein's activity and response. Each system was known to respond directly and specifically to O_2. This high degree of specificity of each molecular mechanism ensured that we would be sensing DO_2 and nothing else. Both systems were also known to respond relatively quickly and strongly to changes in DO_2. A fast response would guarantee close to real-time sensing of DO_2. A strong response would enable easier composition of the sensor with downstream circuitry and more modular behavior of the sensor in other organisms. Together, these criteria made FNR and FixL/J the best choices for development as oxygen sensors for synthetic genetic programs.

The FNR and FixL/J systems detect DO_2 levels and regulate transcription based on those levels. This transcriptional regulation occurred at the level of the promoter, where each protein would enhance or prevent RNA polymerase binding and subsequent transcription. To develop cognate promoters for both systems of oxygen sensors, we took two approaches: 1) The recruitment of native sequences from each system's native organism and 2) the construction of "synthetic" promoters that incorporated the operons known to be targets of their respective transcription factors. Several native sequences of regulated promoters in *E.coli* and *S. meliloti* were cloned upstream of a reporter GFP construct. These native sequences were taken from genes that were observed to be regulated by the cognate protein's activity.

Synthetic promoters were constructed with the aim of creating either an activatable promoter or a repressible promoter. Activatable promoters were designed with the operator centered at the –41.5 site for FNR or the -35 site for FixJ in the weak constitutive promoter J23150 (Figure 3.2A). Repressible promoters were designed with the operator centered either between the -10 and the -35 sites or immediately downstream of the -10 site. At these

positions, operators were thought to recruit their cognate proteins in such a way as to facilitate either activation or repression of transcription[145]. In each case, the position of the operator was varied by up to 4 bp in either direction to optimize the position of the operator and thereby maximize the dynamic range of the induced promoter. In addition, we attempted to construct several repressible promoters by engineering DNA looping, which is necessary in some cases for FNR repression. In these promoters, we inserted two operators into the strong constitutive promoter J23102 and separated them by a randomly generated DNA sequence of a length known to enable flexible DNA looping upon protein binding.

For both FNR and FixL/J systems, several native and synthetic promoters were inserted upstream of the EcFbFP fluorescent protein reporter. EcFbFP fluorescence, unlike Green Fluorescent Protein (GFP), does not rely on the availability of oxygen and is therefore a useful reporter in anaerobic conditions[146]. Each of the promoter-EcFbFP reporters was cloned on a p15A plasmid and transformed into *E.coli*. The FixL/J operon was cloned from the genome of *S. meliloti* and constitutively expressed from an incW plasmid. Promoter function was screened by growing these strains of *E.coli* in either aerobic or completely anaerobic conditions. Anaerobic conditions were created by covering 14 ml culture tubes with a rubber stopper and then removing the air in the headspace above the culture using a syringe and gas manifold. Cultures were grown for up to two hours after removal of the oxygen from the culture before a sample of the culture was taken and measured by flow cytometry.

Out of all the screened promoters, only synthetic promoters designed for activation showed substantial activity upon removal of oxygen from the culture (Figure 3.2B). Few of the native promoters showed any modulation of fluorescence levels. No synthetic repressible promoters showed strong repression of promoter activity. The strongest induction (7-fold) was observed by one of the synthetic FixL/J promoters (P_{FixLJ2}); two FNR promoters (P_{FNR8} and P_{FNR18}) showed substantial induction (5-fold and 4-fold, respectively). We also observed a consistent decrease in fluorescence upon exposure of cells to anaerobic conditions. This was most likely due to altered metabolism of the host *E. coli* reducing production of reporter protein. Because this phenomenon was consistent for constitutive control promoters, we subtracted this % change of the controls from the dynamic range of the all inducible promoters.

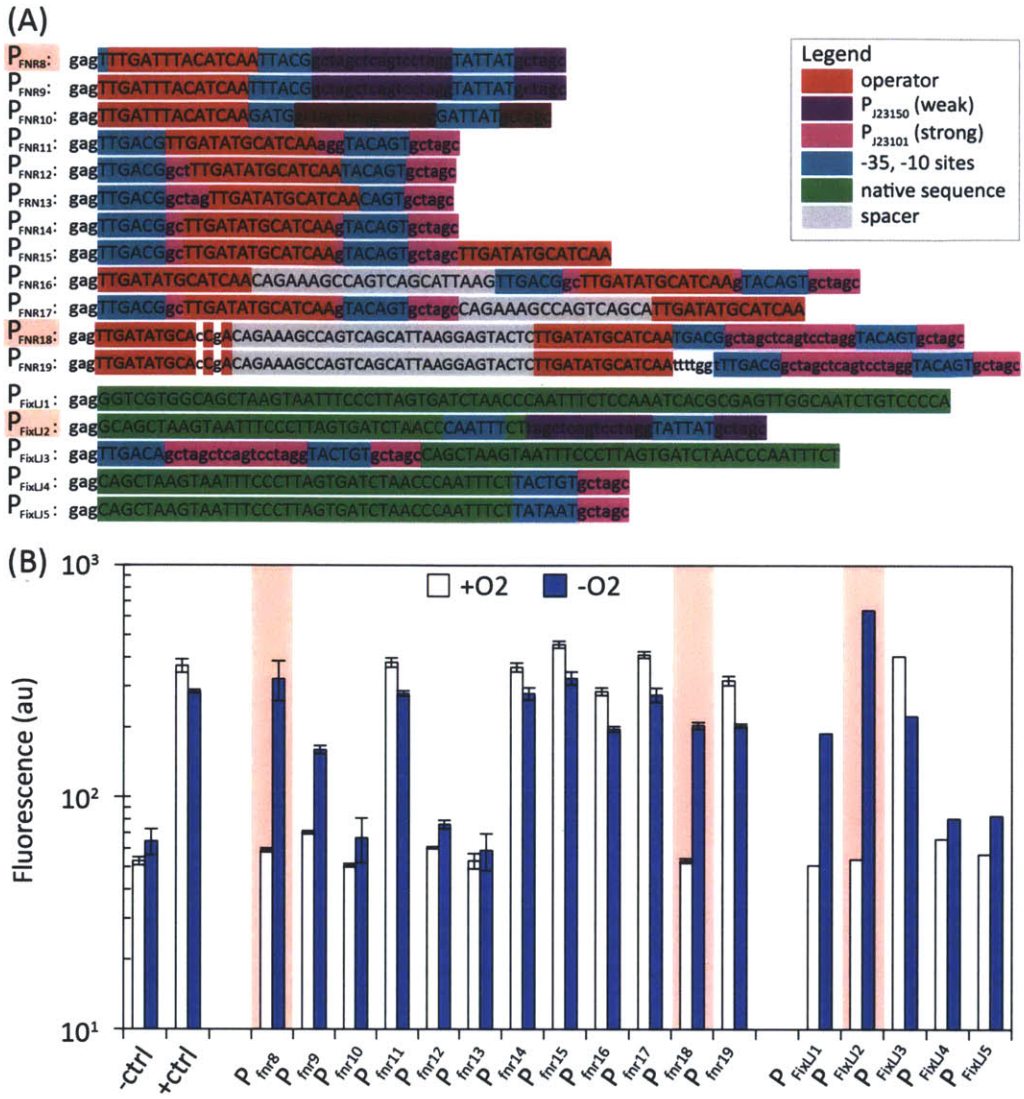


Figure 3.2: Design and screen of oxygen promoters. (A) Oxygen sensor promoters using FNR (P_{FNR}) and FixJ (P_{fixlj}) were designed by integrating operators and native promoter sequences (from *S. meliloti*) into synthetic constitutive promoters. Design elements are colored and as described in the inset legend. **(B)** Oxygen promoters were screened by removing oxygen from the headspace of growing *E. coli* cultures and measuring resulting reporter (EcFbFP) fluorescence after two hours. The three promoters chosen for further characterization showed strong induction (>5-fold) and are highlighted with a red background. Control promoters J23113 (-ctrl) and J23101 (+ctrl) are shown for comparison.

Based on the results of our promoter screening, we chose one FixL/J (P_{FixLJ2}) and two FNR inducible promoters (P_{FNR8}, P_{FNR18}) for further characterization and development. Response functions of each of these sensor modules were measured by injecting known amounts of air into the headspace from which oxygen had been removed by a gas manifold as described

above. The response functions showed a sharp rise in signal as the O_2 levels in the headspace approached zero (Figure 3.3). Later characterization of the FixL/J promoter in the production strain BL21Δ*glnL* showed >20-fold induction, which was sufficient for integration into circuitry.

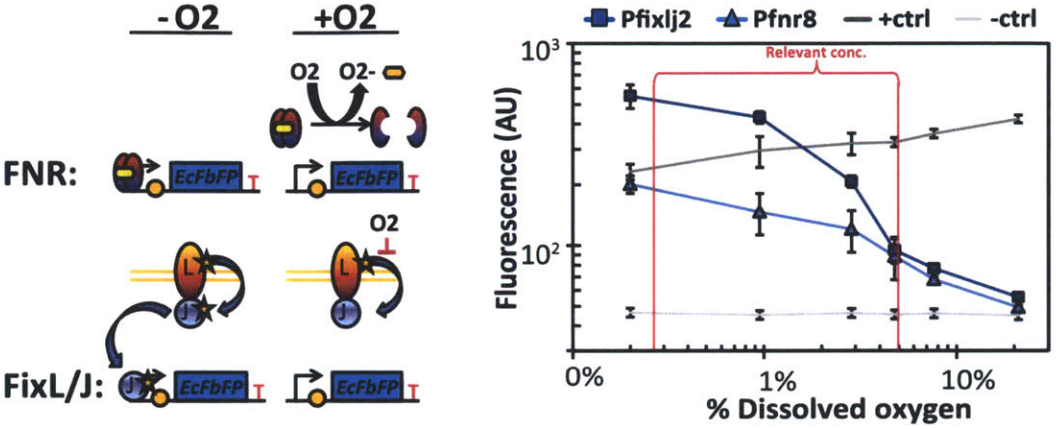


Figure 3.3: Function and response of oxygen sensors. (A) The mechanism of the FNR and FixL/J oxygen sensors is shown. FNR binds its operator and activates transcription of reporter EcFbFP in the absence of oxygen. In the presence of oxygen, FNR iron-sulfur cluster is oxidized, which prevents dimerization of the protein and subsequent DNA binding. FixL/J is a two-component system that is inhibited by oxygen. When oxygen is absent, the receptor protein FixL transfers a phosphate group to the response regulator FixJ, which upregulates transcription upon binding to its target promoter. **(B)** The response of the two sensors is shown in shake flask cultures. FixL/J has a better performance with up to 10-fold dynamic range in BL21Δ*glnL*. Control promoters J23101 (+ctrl) and J23113 (-ctrl) are shown for comparison.

3.4.1.2 Acetate Sensor

We sought to use an acetate sensor native to *E. coli* that was previously characterized by Liao and co-workers[147]. This sensor is made up of the NRI protein and the P_{glnAP2} promoter, which is upregulated when NRI is phosphorylated (Figure 3.4A). NRI is phosphorylated by acetyl phosphate, an intermediate in the acetate pathway. NRI is dephosphorylated by the NRII phosphatase. When the *glnL* gene encoding NRII is knocked out, NRI is phosphorylated at higher levels correlating with acetyl phosphate concentration. Phosphorylated NRI binds to several sites in the 250 bp P_{glnAP2} promoter. This promoter is driven by sigma-54, making it dependent on the NRI activator as well as sigma-54 abundances. When NRI binds each of the four sites upstream of the sigma-54 promoter, it causes looping of the DNA, which brings the NRI activator in close proximity to the promoter site. Activation of the promoter is strongly

dependent on the exact position of the NRI operators[148]. If the operators are spaced such that looping is abolished or localizes the activator to the opposite side of the sigma-54 binding site on the DNA, no activation occurs. This finely tuned mechanism of activation is difficult to replicate by design, so the wild-type P_{glnAP2} promoter was tested for activity.

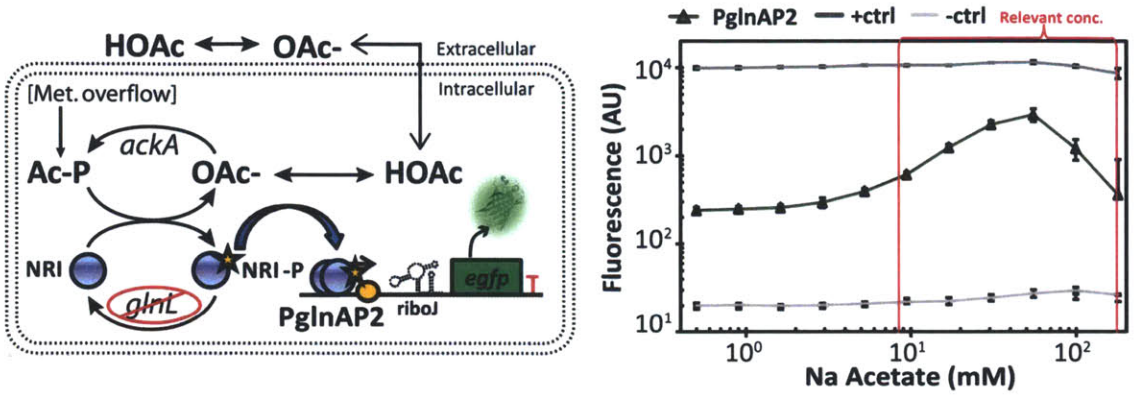


Figure 3.4: Function and response of the acetate sensor. (A) The mechanism of the acetate sensor is shown. External acetate is protonated and diffuses across the membrane. In the cytoplasm, acetate is converted to acetyl phosphate by acetyl kinase, product of the *ackA* gene. Acetyl phosphate is used to phosphorylate NRI transcription factor, which dimerizes and binds and activates the P_{glnAP2} promoter, driving reporter GFP expression. **(B)** The response of the P_{glnAP2} promoter to external acetate at pH 6.8 is shown. Toxicity is seen at acetate concentrations higher than 60 mM. Control promoters J23101 (+ctrl) and J23113 (-ctrl) are shown for comparison.

The wild-type P_{glnAP2} promoter was cloned upstream of a sfGFP reporter on the p15A plasmid. This plasmid was transformed into a BL21 strain of *E. coli* containing a deletion of the *glnL* gene (BL21Δ*glnL*). Initially, only scant activation of the promoter was seen at high levels (60 mM) of acetate at pH 7. To improve the function of the sensor, an insulating ribozyme element (riboJ) was included upstream of the RBS and greatly improved the activity of the sensor (Figure 3.4). This element is thought to abolish interference of the 5' UTR of the transcript with the RBS. In this construct, the basal rate of the promoter was greatly increased, suggesting that the initial design was simply operating at undetectable levels. To test this hypothesis, other constructs lacking the riboJ element but containing stronger RBSs were tested. These constructs showed similar activity to the improved sensor containing riboJ, confirming that the initial design was operating at undetectable levels. Previous work showed that sensor response depended on the external pH. When acetate is protonated, its charge is neutral and the molecule can readily diffuse across the cell membrane. When the external pH is

more acidic than the internal pH of the cell, acetate tends to accumulate at higher levels in the cell. This is because acetate becomes deprotonated and charged in the more basic cytoplasm and is therefore unable to diffuse across the membrane. Because of this, the amount of acetate inside the cell increases relative to the external acetate concentration as the pH is decreased. This effectively gives the acetate sensor a lower detection threshold of external acetate concentration at lower pH. When the sensor was tested at lower pH's in minimal media, its detection threshold was lowered by 32-fold compared to its detection threshold at pH 7 (Figure 3.5A). However, the minimal inhibitory concentration (MIC) of acetate in the media is also decreased by 10-fold across the same pH range (Figure 3.5B). This feature enables the selection of the necessary threshold of detection of acetate based on pH.

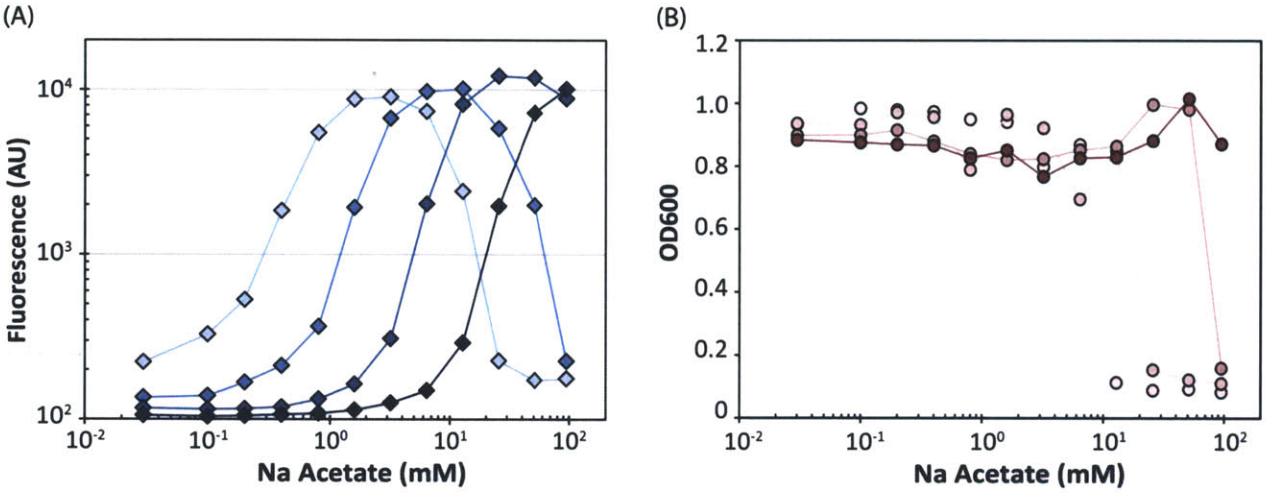


Figure 3.5: Response of the acetate sensor to changes in pH. A strain of *E. coli* BL21Δ*glnL* containing the acetate sensor was grown on media stabilized at variable pH. **(A)** The detection threshold of the acetate sensor changes with pH. Fluorescence decreases at high levels of acetate due to acetate toxicity **(B)**, which increases with a decrease in the pH. The darkest line in both **(A)** and **(B)** represents a pH of 7.0. Subsequently lighter lines represent pH's of 6.5, 6.0, and 5.5, respectively.

3.4.1.3 Glycolytic Flux sensor

Glycolytic flux is the primary determinant of acetate formation. When the flux exceeds a critical threshold, acetate is produced by overflow metabolism. To sense glycolytic flux, I chose to use the Cra global regulator to sense two intermediates in glycolysis: fructose-1,6-bisphosphate (F16BP) and fructose-1-phosphate (F1P). Because accumulation of these intermediates is strongly correlated with amount of flux through glycolysis, their detection was

an acceptable proxy for glycolytic flux[149]. F16BP is the fourth intermediate in glycolysis and has been measured at 15+/-1.2 mM in *E. coli* during growth on glucose, making it the third most abundant metabolite[150]. During growth on acetate and glycerol, F16BP levels decrease to <0.15 mM and 5.85+/-2.31 mM, respectively. F1P is an intermediate of fructose metabolism and is made into F16BP. F1P cannot be produced from F16BP due to the lack of an enzyme that catalyzes the reverse reaction; no such enzyme is known to exist in nature. Cra (catabolite repressor/activator), also known as FruR, is a global regulator in *E. coli* that stimulates gluconeogenesis during growth on energy poor carbon sources. In the absence of F16BP and F1P, Cra binds to its operator, which binds most strongly to the near-consensus fruB O1 sequence: GC-TGAAAC|GTTTCA-AG[151]. Cra acts as a repressor when the Cra operator is positioned immediately adjacent to or downstream of the transcription start site (TSS) or between the -10 and -35 sites on the promoter. Cra acts as an activator when the operator is positioned near or far upstream of the -35 site. The large distance at which Cra activates transcription as well as its homology to lacI and galR suggest that DNA looping plays a role in Cra-mediated transcriptional activation. Cra binding can be modulated by F1P at micromolar concentrations and F16BP at millimolar concentrations (Figure 3.6)[152]. This makes Cra activity more dependent on fructose availability than glucose availability. Nonetheless, the response range of Cra to F16BP and the large range of F16BP abundance during growth on different carbon sources suggest that it can be useful as a glycolytic flux sensor as well.

A library of 72 promoters containing Cra operators at repressive sites was built. These promoters were constructed by combinatorial assembly of different promoter pieces. Separate pieces composed the 5'UTR, a central region from +1 to -35, and an upstream region from -35 to -80. This library was screened on 2% glucose and 2% fructose. The best promoter showed up to 3-fold induction on fructose (Figure 3.6). The response of these promoters was further characterized across a large response range of glucose and fructose concentration. Although the glucose response was very weak (1.5-fold), the fructose response was as strong as 3-fold.

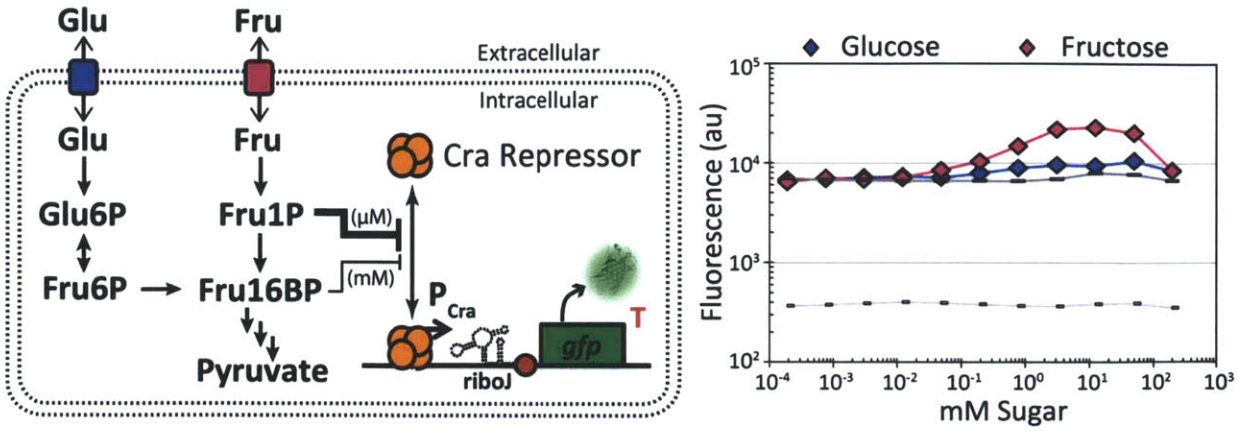


Figure 3.6: Function and response of the glycolytic flux sensor (A) The mechanism of Cra-based glycolytic flux sensor is shown. Cra binds to its operator and represses transcription when the operator is placed in the 5'UTR or between the -10 and -35 sites of the promoter. Fructose-1-phosphate (F1P) and Fructose-1-6-bisphosphate (F16BP) allosterically modulate Cra's affinity to its operator. F1P has much stronger affinity for Cra than F16BP. **(B)** Response of a synthetic Cra-based promoter (P_{Cra2}) to glucose and fructose addition to the media of a growing culture. The response from fructose addition is up to 4-fold, much stronger than the response to glucose. Control promoters J23101 (top gray line) and J23113 (bottom gray line) are shown for comparison.

3.4.2 Detecting separate causes of acetate formation in *E. coli*

Before building the program, I first set out to correlate the input signals (low oxygen, high glycolytic flux) to the production of acetate. To do this, I used a starch-fed media to control glucose release rates in the media and a vacuum manifold to control oxygen levels in the cultures. By adding different amounts of amyloglucosidase to a culture containing dissolved, pre-treated starch, the media effectively released glucose at a defined rate. The linear release of glucose led to a linear growth curve for *E. coli* until the culture entered stationary phase. This also meant that the *E. coli* was experiencing a gradually decaying glycolytic flux rate, since more and more bacteria were consuming the same amount of glucose released. Variable oxygen levels were produced in flasks by removing the oxygen with a vacuum manifold.

Using variable concentrations of amyloglucosidase was effective in controlling the growth rate of *E. coli* via the glycolytic flux rate. Initially, all cells grew exponentially as the rate of glucose released by the enzyme was greater than the rate of glucose consumption of the culture. When the culture grew to sufficient density to where glucose consumption rate equaled the glucose production rate, the culture's growth rate adjusted to the enzymatic glucose release rate (Figure 3.7A). At the highest concentrations of enzyme, cells grew to

similarly high OD's. Anaerobic cultures also grew at rates that correlated with enzyme concentration. This indicates that although cells were in the oxygen-limited regime, they were also carbon limited. This confirms that the glycolytic rate was not saturated by the anaerobic conditions and could still be varied.

Cells that grew without oxygen produced more acetate (Figure 3.7B), despite their slower growth rates. Cells that were grown aerobically initially consumed acetate produced during the first four hours of growth, during which time cells grew exponentially as they consumed the glucose released during that time (Figure 3.7A). Interestingly, aerobically growing cells greatly increased their acetate production late in growth while their growth rate was slowing. This may be due to oxygen limitation, accumulation of toxins at high cell densities, or exhaustion of the starch carbon source.

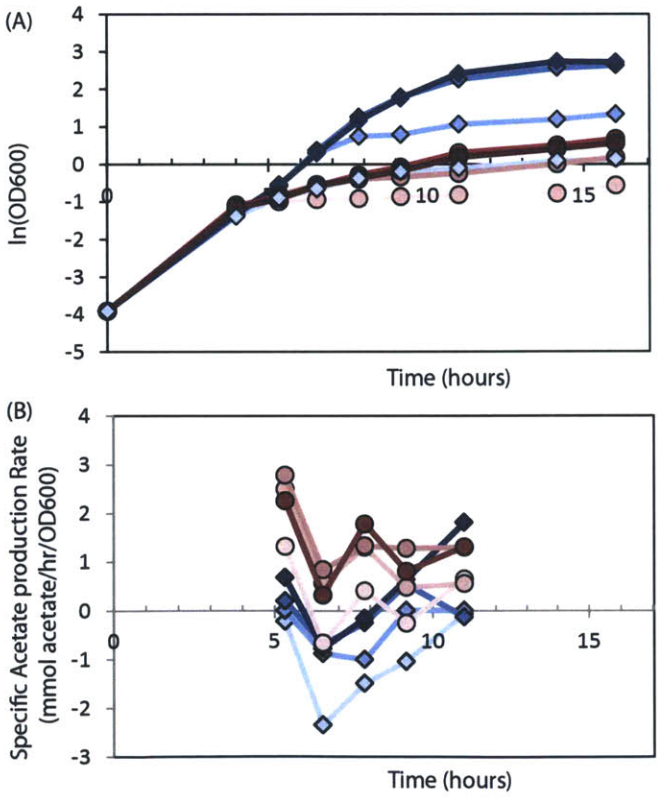


Figure 3.7: Variable acetate production in *E. coli*. *E. coli* was grown on starch media containing different amounts of amyloglucosidase. After 4 hours, the oxygen was removed from half the cultures. Cultures containing oxygen are shown as blue diamonds. Anaerobic cultures are shown as red circles. The amount of amyloglucosidase added to each culture is indicated by shading, lightest shading indicating the least amount of amyloglucosidase and increasing as follows: 1.5E-4 units/ml, 1.5E-3 units/ ml, 7.5E-3 units/ml, and 3E-2 units/ml. **(A)** OD_{600} measurements were taken at each time point, starting four hours after 1/100 dilution of the overnight culture. The specific growth rate correlates with the amount of amyloglucosidase added to each culture. **(B)** Acetate production of each culture was measured during linear growth. Acetate production is higher in anaerobic cultures and correlates with the amount of amyloglucosidase added to each culture.

These experiments confirm several predictions of our model. First, acetate production correlates with the glycolytic flux rate, both in aerobic and anaerobic conditions. Second, acetate is produced in higher amounts at the same glycolytic flux rate in anaerobic conditions.

These experiments also show that we can control both input variables independently. Further work will focus on correlating the observed glycolytic flux rates, oxygen conditions, and acetate production levels with sensor readouts.

3.4.3 Detecting acetate formation in *E. coli* using a synthetic AND gate

To identify acetate formation due to microaerobic and anaerobic conditions, a digital AND gate integrating the FixL/J oxygen sensor and the acetate sensor was constructed. A previously published AND gate based on the *Salmonella typhimurium* invF/sicA activator/chaperone two-component system was used to integrate the oxygen and acetate sensors (Figure 3.8A)[153]. Two plasmids were constructed: one in which the oxygen sensor drove the *invF* gene and the acetate sensor drove the *sicA* gene, and a second with the sensor regulation switched. A plasmid containing a GFP reporter gene driven by the P_{sicA} promoter was co-transformed with each of these plasmids. Cells carrying both controller and reporter plasmid were then grown normally or exposed to high acetate concentrations, low oxygen levels, or both. The initial constructs showed some function but indicated a large amount of basal leak from the *sicA* RBS (Figure 3.8B). A mutagenesis library of the *sicA* RBS was made, and several mutants were shown to have a strong, digital AND behavior (Figure 3.8C).

The working oxygen/acetate AND gate showed 14-fold induction when acetate was added to 50 mM and oxygen was removed from the culture (Figure 3.8C). This was sufficient to separate the cytometry distributions almost completely, indicating the detection of a cell state distinct from the uninduced culture. Some activation (2-fold) from single induction in anaerobic conditions or on acetate alone was still visible, most likely due to residual leak from the individual promoters. This completed oxygen/acetate AND gate serves as a proof-of-concept that demonstrates that two metabolic signals can be integrated to form digital logic. The invF/sicA system enabled simple construction of this system, since it is only composed of two protein components. Further work will validate the performance of this system by sensing in vivo production of acetate in anaerobic and aerobic conditions.

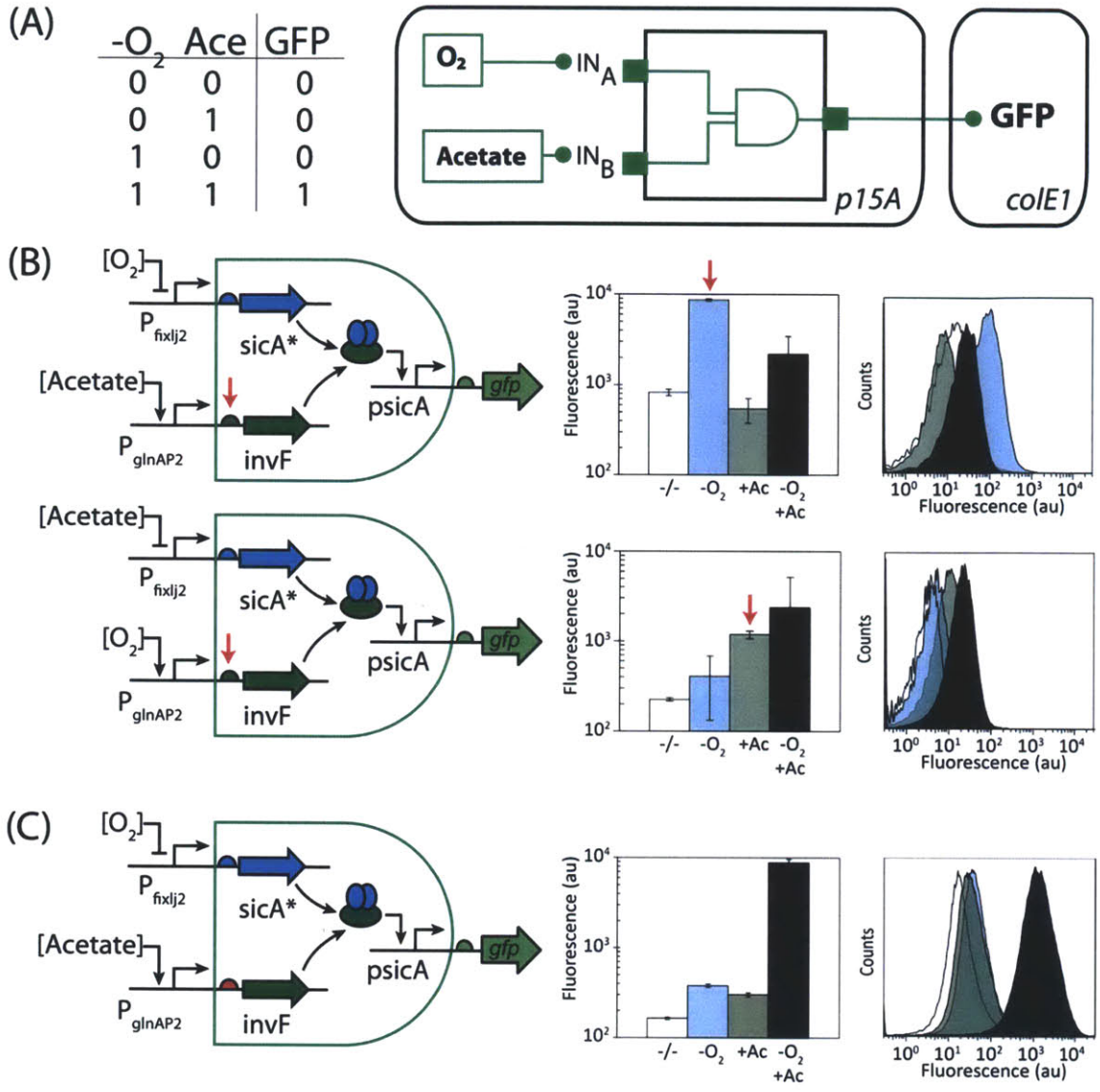
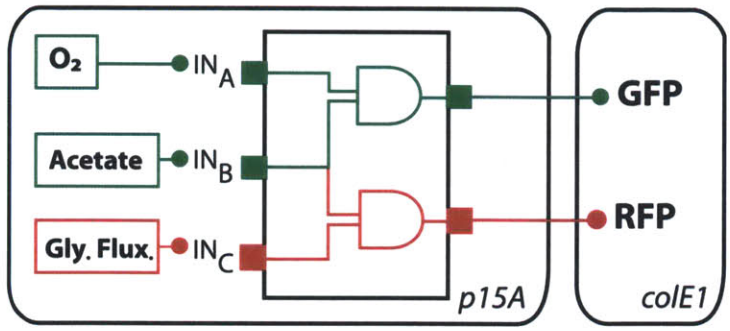


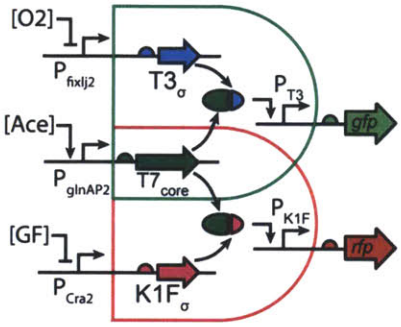
Figure 3.8: Prototype Oxygen/Acetate AND gate. (A) Truth table and logic diagram of an abstracted oxygen/acetate AND gate. The readout (GFP) is ON only when oxygen is absent (-O2) and acetate is present (+Ace). **(B)** Two versions of the oxygen/acetate AND gate were constructed, each with the oxygen and acetate sensors driving a different promoter of the invF/sicA two-component system from *Salmonella*. Arrows indicate evidence of leak of invF. **(C)** The first gate from (B) was mutated at the invF RBS (highlighted in red). The result was a functional oxygen/acetate AND gate with a 70-fold dynamic range.

(A)

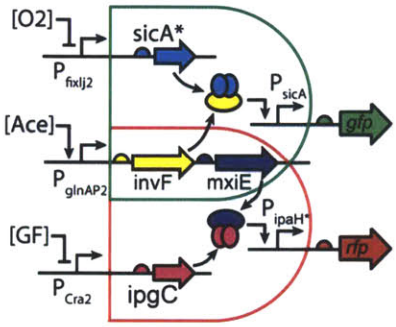
Inputs			Outputs	
G.F.	$-O_2$	Ace	GFP	RFP
0	0	0	0	0
0	0	1	0	0
0	1	0	0	0
0	1	1	1	0
1	0	0	0	0
1	0	1	0	1
1	1	0	0	0
1	1	1	1	1



(B)



(C)



(D)

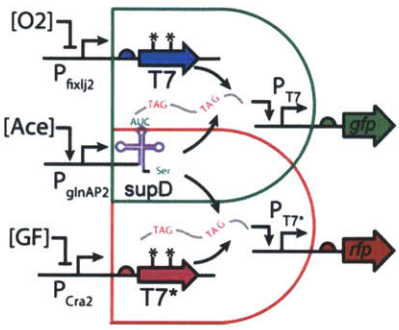


Figure 3.9: Full acetate detection program. (A) Truth table and logic diagram of the full acetate detection program. The program is based on two AND gates than can detect acetate production either due to low oxygen conditions or high glycolytic flux. The readout GFP is ON only when acetate is present (+Ace) and oxygen is absent (-O2). RFP is ON when acetate is present (+Ace) and the glycolytic flux (GF) is high. When all inputs are present, both reporters are ON. Colored gates correspond to the mechanism diagrammed below. **(B-D)** Diagrams of designs for the program detailed in **(A)**. **(B)** Schematic of program design incorporating a split T7 RNA polymerase system. The acetate sensor drives the T7 RNAP core fragment. The oxygen sensor drives the T3 sigma-like fragment. The glycolytic flux sensor drives the K1F sigma-like fragment. Only when the core fragment and one of the sigma-like fragments is expressed does the cognate promoter turn ON. When the T7 core fragment is limiting, the sigma-like fragments compete to bind it. The GFP/RFP readout in this case is a ratio calculation of the activity of the oxygen and glycolytic flux sensors. **(C)** Schematic of the program design incorporating two sets of orthogonal two-component systems. The acetate sensor drives invF and mxiE. The oxygen sensor drives sicA and the glycolytic flux sensor drives ipgC. Only when invF and sicA are expressed is the PsicA promoter active. Only when mxiE and ipgC are expressed is the PipaH* promoter active. **(D)** Schematic of the program design incorporating the suppressor supD RNA enabling translation of orthogonal T7 RNA polymerases. The acetate sensor drives supD expression. The oxygen and glycolytic flux sensors drive two orthogonal T7 polymerases containing UAG stop mutations. Only when supD is co-expressed with either T7 RNA polymerase is either protein translated and the cognate promoter activated.

3.4.4 Three designs for the complete integrated circuit

The completed acetate detection program aims to integrate acetate/oxygen and acetate/glycolytic flux AND gates into a single circuit (Figure 3.9A). To do this, I designed three different architectures of the two integrated AND gates that implement different molecular mechanisms to perform the computation.

The first design implements a split-protein system developed in our lab (Figure 3.9B). This system splits the T7 RNA polymerase into two subunits: the "core" fragment and the "sigma-like" fragment[154]. Alone, neither subunit can activate transcription. When expressed simultaneously, the subunits associate and activate transcription from a promoter that is cognate to the sigma-like fragment. Several orthogonal T7 RNA polymerases have been developed, each containing mutations in the sigma-like fragment[155]. By expressing different sigma-like fragments simultaneously, one can thereby redirect the limited number of available core fragments to different promoters, enabling a type of ratio calculation. Because acetogenesis in the oxygen-limited regime is a function of the ratio of glycolytic flux to oxygen transfer rate, this system is well-suited for application to the acetate program.

The second design implements a second activator/chaperone system from *Shigella flexneri* to facilitate the acetate/glycolytic flux AND gate (Figure 3.9C). The acetate sensor drives both chaperones (invF, mxiE) co-cistronically. The RBS driving mxiE enables its tuning separate from invF expression. The glycolytic flux sensor drives expression of ipgC. Thus, when acetate is present due to high levels of glycolytic flux, RFP will be expressed. Because the sicA/invF and ipgC/mxiE systems are mostly orthogonal, GFP and RFP expression should be largely independent, enabling separate readouts for both acetogenic conditions.

The third design implements a supD repressor tRNA to enable expression of two orthogonal T7 RNA polymerases (Figure 3.9D). SupD tRNA suppresses the UAG amber mutation by adding a serine on the growing protein to prevent ribosome pausing and subsequent stopping of translation. Genes containing UAG codons in the middle of their sequence cannot be translated fully unless supD tRNA is present. By inserting several UAG mutations into the orthogonal T7 polymerases, each expressed by either the oxygen sensor or the glycolytic flux sensor, expression from their cognate promoters becomes dependent on the expression of

supD. Thus, acetate becomes the "master switch" that enables the readout from the oxygen and glycolytic flux sensors. Because supD is somewhat toxic to the cell and has been shown to activate translation at leaky levels, this is the leased preferred design.

Here, I have only considered designs that facilitate simple one-to-one activation schemes instead of elaborate transcriptional circuits containing multiple layers of repressors. Such layered repressor schemes require multiple rounds of transcription and fast degradation of the intermediate repressors. Because our system specifications call for fast response times, such designs are inappropriate for the program presented here.

3.4.5 Actuators to reduce acetate formation

In order to implement working feedback in single cells, it is necessary for the output of the program described above to produce molecules that lower acetate production. Many genetic manipulations have been found to lower acetate production[156]. These usually involve knocking out genes that produce acetate (e.g. *pta, ackA, poxB*) or overexpressing an enzyme that redirects flux away from acetate. Many of these manipulations, however, come with side effects, such as slow growth, by-product accumulation, and lowered biomass. Here, I review several mechanism and target molecules as potential outputs for the completed program above. These mechanisms conditionally lower acetate production in the cell by directly addressing the underlying cause.

Knockouts have been among the most effective ways of reducing acetate formation in *E. coli*. However, no great tools exist in *E. coli* for conditional downregulation of native genes. Recently, development of CRISPRi has facilitated RNA-based knockdown of native gene expression in *E. coli*[157]. CRISPRi uses a catalytically inactive mutant of the Cas9 CRISPR protein (dCas9) from *Streptococcus pyogenes* to block transcription of the target gene. This is accomplished by targeting dCas9 to the target gene using a short RNA containing a dCas9 binding motif and a sequence homologous to the target gene. This system has been shown to lower transcription of target genes by >100-fold. This system could be used to target individually and in combination the genes that drive acetate production in *E. coli*. These include

the glucose specific enzyme II glucose transporter (*ptsG*), pyruvate kinase (*pyk*), pyruvate dehydrogenase (*pdh*), acetate kinase (*ackA*), and phosphotransacetylase (*pta*).

Overexpression of several proteins has been shown to reduce acetate formation in *E. coli*. The overexpression of acetyl-CoA synthase (*acs*), NADH oxidase, and the *Vitreoscilla* hemoglobin (VHb) genes has been shown to be especially effective in reducing acetate accumulation in a variety of conditions. Acetyl-CoA synthase is a native enzyme in *E. coli* that can assimilate acetate directly acetyl-CoA and has been shown to reduce acetate accumulation when ovexpressed[158]. NADH oxidase (NOX) oxidizes NADH to NAD+, thereby shifting the redox balance during anaerobic growth and during excess NAD+ usage of heterologous metabolic pathways[159]. Overexpression of NOX has been shown to reduce acetate accumulation in productive cultures of *E. coli*. *Vitreoscilla* hemoglobin has been shown to reduce acetate accumulation in microaerobic cultures of *E. coli*[160]. VHb is thought to facilitate oxygen transport to the terminal cytochrome in the electron transport. Using the NOX and VHb genes as outputs of the oxygen/acetate AND gate and using dCas9 downregulation of acetate producing genes as the output of the glycolytic flux/acetate AND gate would create feedback on the respective acetate-producing condition.

3.5 Discussion and Future Directions

Synthetic genetic circuits have long been confined to small "toy" systems that perform arbitrary functions. Although these have provided an impressive display of the engineering potential of biological systems and have provided insight into systems biology questions, synthetic circuits have yet to be implemented at scale in industrial systems. Here, we develop a prototype synthetic genetic program that targets acetate production in *E. coli*, an expensive problem in industrial fermentations.

Development of sensors for this program provided a great challenge. Our ability to use sensors highly depends on the availability of information about the sensors. The sensor protein's specificity and sequence must be known and the operator sequence to which it binds characterized. This minimal information then enables the screening of promoters responsive to the sensor protein and the assay of its response to certain conditions. For the program

presented here, we sought sensors for oxygen, glycolytic flux, acetate, and changes in redox state. The redox sensor never proved fruitful. The oxygen, glycolytic flux, and acetate sensors were eventually tuned to where they could be integrated into genetic circuits.

Much of the sensor development process included screening for promoters with optimized response ranges. The fact that several native promoters turned out to be either weak or nonresponsive was surprising. There are several reasons why the native promoters were not seen to be as active as expected. Although we aimed to include as much of the functional promoter region that had been mapped with DNA footprinting assays in the literature, these regions might not have been sufficient to retain native promoter activity. For future studies, it would be prudent to include sequences farther up- and downstream of the annotated promoter regions. Also possible is that the heterologous conditions in *E.coli* changed the function of the FixL/J protein to levels where the native promoters would no longer be functional. Also, non-native concentrations of the sensor proteins might be affecting the level of activity that can be produced on the cognate promoters.

The fact that we were unable to engineer effective repression with the oxygen sensor was also unexpected. Repression occurs when a protein bound near a promoter sterically inhibits binding of RNA polymerase to the promoter. Thus, the inclusion of an operator within most sites on a promoter should enable repression of transcription upon binding by the cognate protein. Even the promoters that attempted to replicate the DNA looping mechanism of FNR repression showed no functional repression (Figure 3.2). Despite taking into account what are relatively well-known mechanisms of repression in both the FNR and FixL/J systems, we were unable to engineer repression in these synthetic promoters. This is perhaps a testament to our incomplete knowledge of the details of the mechanisms of repression or simply to the complexity of the molecular mechanisms underlying it.

Further refinement of the oxygen sensor measurements should include absolute determination of dissolved oxygen via a probe. Here, we report the dissolved oxygen as a function of the oxygen present initially in the headspace of the culture tube (Figure 3.3). It is important to note that the injected oxygen is essentially a "batch" of oxygen, and that the true dissolved oxygen concentration is continuously decreasing until it is depleted. Thus, the

cultures run out of oxygen at different times; these times are nonlinearly related to the initial oxygen concentration due to the change in growth rate and oxygen consumption as a function of cell density. Because of this, the reported fluorescence values can be considered the maximum possible values at the reported dissolved oxygen and more accurately reflect sensor output at lower dissolved oxygen values. This experimental setup was chosen due to technical limitations of the experimental system, which required simultaneous exposure of dozens of cultures to varying levels of oxygen.

Several lines of work could be continued in an effort to complete the entire program detailed above. First, a more robust assay for acetate production in different conditions could be developed. The current assay uses 20 ml shake flask cultures at variable oxygen concentrations and glucose release rates. Though this assay clearly shows acetate production is a function of glucose release rates and oxygen availability, it is technically difficult, low-throughput, and prone to large error due to repeated handling of the culture. Two developments would enable more rapid and robust measurements. First, the acetate sensor should be further developed as a biosensor to facilitate acetate assays. The current iteration of the sensor responds to changes in sigma-54 concentration in the cell. It is also not sensitive enough at pH 7 to detect very small amounts (<10 mM) of acetate. Currently, acetate is assayed either by an enzymatic kit or by HPLC. If the acetate sensor can be shown to reliably report acetate levels throughout growth, it would greatly facilitate a high-throughput, plate-based assay the output of which would be directly relevant to the acetate program. Secondly, the development of a plate-based assay of the starch-fed growth curves would greatly help the throughput and accuracy of the assay. A plate reader could continuously monitor OD_{600} and GFP fluorescence. Variable rotation rates and clear plastic, nonbreathable coverings could be used to alter oxygen concentration in each well of the culture. Additionally, the entire program platform will be moved into *E. coli* MG1655, a K-12 strain. Here, I use *E. coli* BL21ΔglnL, a B strain of *E. coli* that has been shown to consistently produce less acetate than MG1655[161]. By switching to MG1655, also an industrially used strain, more acetate will be produced, enabling faster tuning of the program.

The glycolytic flux sensor will be further improved by promoter optimization. By facilitating chip-synthesized oligo libraries to construct a large library of rationally-designed glycolytic flux promoters, we aim to discover and characterize improved promoters with lower detection thresholds and dynamic ranges. If successful, this approach will yield not only an improved glycolytic flux sensor but a generalizable methodology for engineering sensor promoters in high-throughput.

Each design for the finished program above will probably require tuning to respond in the relevant ranges of metabolites. This will be done by random mutagenesis of the RBS's driving the functional proteins or the incorporation of new, improved promoters. Following construction and testing, the most functional program will be validated by sensing physiological production of acetate under anaerobic and aerobic acetogenesis. Finally, actuators will be integrated into the functioning program to feedback on the acetogenic conditions.

4. Genetic Circuit Performance under Conditions Relevant for Industrial Bioreactors

This chapter is reprinted (adapted) with permission from Moser, F., Broers, N.J., Hartmans, S., Tamsir, A., Kerkman, R., Roubos, J.A., Bovenberg, R., and C.A. Voigt. Genetic Circuit Performance under Conditions Relevant for Industrial Bioreactors. *ACS Synthetic Biology* **1**: 555-564 (2012).

4.1 Abstract

Synthetic genetic programs promise to enable novel applications in industrial processes. For such applications, the genetic circuits that compose programs will require fidelity in varying and complex environments. In this work, we report the performance of two synthetic circuits in *Escherichia coli* under industrially relevant conditions, including the selection of media, strain, and growth rate. We test and compare two transcriptional circuits: an AND and a NOR gate. In *E. coli* DH10B, the AND gate is inactive in minimal media; activity can be rescued by supplementing the media and transferring the gate into the industrial strain *E. coli* DS68637 where normal function is observed in minimal media. In contrast, the NOR gate is robust to media composition and functions similarly in both strains. The AND gate is evaluated at three stages of early scale-up: 100 ml shake-flask experiments, a 1 ml MTP microreactor, and a 10 L bioreactor. A reference plasmid that constitutively produces a GFP reporter is used to make comparisons of circuit performance across conditions. The AND gate function is quantitatively different at each scale. The output deteriorates late in fermentation after the shift from exponential to constant feed rates, which induces rapid resource depletion and changes in growth rate. In addition, one of the output states of the AND gate failed in the bioreactor, effectively making it only responsive to a single input. Finally, cells carrying the AND gate show considerably less accumulation of biomass. Overall, these results highlight challenges and suggest modified strategies for developing and characterizing genetic circuits that function reliably during fermentation.

4.2 Introduction

There are many potential applications for synthetic genetic programs in biotechnology. One such application is in the development of intracellular controllers for metabolic pathways that integrate environmental and cellular signals, control expression dynamics, and implement feedback loops[162,163]. Such controllers would require multiple circuit modules that can accurately integrate across the complex and dynamic environment of an industrial bioreactor. Many programs that integrate environmental signals and control the dynamics of gene expression have been constructed[164]. However, they have only been shown to operate under ideal, homogeneous conditions at small scales[165,166,167,168].

The conditions experienced by cells in large industrial bioreactors are different from those used to characterize circuits in most synthetic biology labs[169]. In bioreactors, cells are grown to high cell densities in oxygen- or carbon-limited conditions[170,171,172]. Fermentation times can be long and the cells are maintained at low growth rates for extended times[173]. Over the course of fermentation, the metabolic state of the cells goes through phases with different availability of metabolites, redox equivalents, transcription and translation factors, and global regulatory proteins[174,175,176,177]. Additionally, not all cells in a large bioreactor are experiencing the same microenvironment due to slow mixing times, causing aeration and local substrate gradients[178,179]. The *E. coli* strains themselves have been optimized for industrial production and are genetically different to those commonly used in synthetic biology[180].

Programs consist of genetic sensors and circuits that have been connected to perform a computational operation. Connecting circuits requires the selection of parts that match the output of an upstream circuit with the input required by a downstream circuit[181,182]. Even slight changes in circuit performance could require the selection of different connecting parts. This poses a challenge when designing programs for environments that differ in conditions from those that were used to characterize the individual circuits. Typically, circuit characterization occurs under lab conditions in shake flasks and complex media. These conditions might differ considerably from the conditions of a program's ultimate application. As programs become larger, it will become impractical to re-characterize each circuit under the precise conditions of the end application before constructing the desired program. Because the process of scale-up

occurs in multiple stages, i.e., from shake flask to increasingly larger bioreactors, genetic programs need to function reliably under the environmental conditions associated with each stage without the need for additional genetic manipulation.

In this work, we compare the performance of two genetic circuits – an AND gate and a NOR gate – in industrially relevant conditions. Both circuits were characterized previously under lab conditions[183,184,185]. The two inputs and output of both gates consist of promoters. The AND gate is composed of three plasmids and turns ON when transcription of the Amber suppressor tRNA supD enables the translation of a T7 RNA polymerase (Figure 4.1). Expression of both SupD and T7 RNA polymerase can have adverse effects on growth. In contrast, the NOR gate is composed of two plasmids and is regulated by two promoters that drive the CI repressor, which has been shown to be non-toxic in many genetic contexts. The choice of these gates enabled us to compare two different circuit architectures that impart different loads on the cell.

The performance of each gate was compared across different environments. First, we tested the impact of changing the composition of the media. Second, we compared the gates' performance in the *E. coli* strain in which the circuits were developed (*E. coli* DH10B) with their performance in a strain used in industry as a model for protein production (*E. coli* DS68637). Third, we tested the performance of the AND gate during long fermentations and during shifts in the feed rate. Finally, we tested the impact of the choice of RBS strength to connect a genetic sensor to the AND gate during fermentation. Together, these results highlight the challenges of implementing complex genetic circuits in industrial processes and suggest strategies for building circuits for such processes.

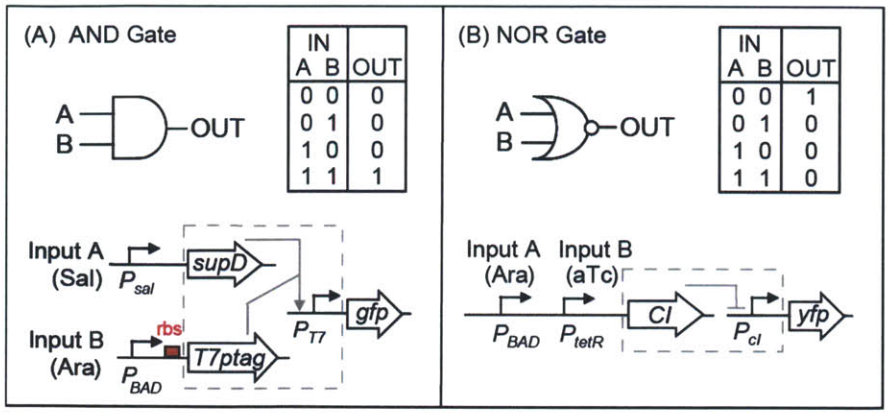


Figure 4.1: The Genetic AND and NOR Gates. The symbol, look-up table, and structure of the genetic logic gates are shown. **(A)** The AND gate is based on a variant of T7 polymerase that contains two Amber stop codons (T7ptag). Only when the tRNA Amber suppressor SupD is transcribed is the T7 polymerase translated and turns ON the output T7 promoter driving the GFP reporter. The RBS varied in this study is shown in red. **(B)** The NOR gate is based on two tandem promoters that drive the expression of a repressor (CI) that turns off an output promoter driving the GFP reporter.

4.3 Methods

4.3.1 Strains and Plasmids

Escherichia coli K-12 RV308 [Su-, lac X 74, gal ISII: OP308, strA] was obtained from ATCC (#31608, deposited by Genentech). The arabinose operon from positions 65,855 to 71,266 (Genbank #U00096) was replaced with Kan^R by the method of Datsenko and Wanner[186]. This marker was transduced into the RV308 parent strain by P1 phage (BW28357). Kan^R was then removed using the FRT recombinase encoded by the plasmid pCP20. The resulting strain was constructed at UCSF to resemble phenotypic characteristics of the DS68637 strain used for fermentation research at DSM. Replicates of shake flask experiments used the strain constructed at UCSF, designated as $DS68637^{\dagger}$ in this text. Tube replicates of shake flasks experiments were performed in the Synthetic Biology Center at MIT. All fermentation, microreactor, and initial shake-flask experiments were done with the original DSM DS68637 strain at the DSM Biotechnology Center in Delft.

E. coli DH10B was obtained from Invitrogen (#18297). *E. coli* DH10B and DS68637 (both UCSF and DSM versions) were co-transformed with plasmids pAC-SalSer914, pBR939b, and different RBS variants of pBACr-AraT7940, which together constitute the AND gate. The different RBS variants of pBACr-AraT7940 were designated pFM159, pFM160, pFM161, pFM163, corresponding to RBS_C, RBS_D, RBS_B, and RBS_A, respectively. To construct the NOR gate, the strains were co-transformed with plasmid pCI-YFP and plasmid pNOR10-20. For the RBS variation experiments, Top10 cells (Invitrogen #C4040-10; genetically identical to DH10B) were transformed with RBS plasmids containing a ColE1 replicon bearing the promoter BBaJ23100 (www.partsregistry.org) and a downstream mRFP1 red fluorescent protein. The RBS's driving translation from these constructs were identical to the forward engineered RBSs # 1, 2, 5, 6, 11, and 14 (pFM169-174, respectively) from Salis, et al. (2009)[187]. The reference plasmid pFM46 (Kan^R, p15a) was used as a basis of comparison of reporter expression. This plasmid contains the promoter/RBS/GFP/terminator construct J23102/B0032/E0040/B0015 (www.partsregistry.org) and is nearly identical to the Kelly et al. (2009) standard plasmid. Kelly et al. (2009) used the promoter J23101 as their promoter standard, which is nearly identical in sequence and strength to J23102[188].

4.3.2 Shake flask and tube Experiments

Plasmids were maintained in shake flasks, the Biolector MTP Microreactor, and the 10 L fermentations by culturing in 10 µg/ml chloramphenicol, 10 µg/ml neomycin, and 50 µg/ml ampicillin. Culturing was done in 2xYT media (Teknova, #Y0167), LB media (Fisher Scientific #R452322), or an in-house defined minimal medium (DSM). The defined minimal medium, which was used in all fermentations and shake flask experiments, was composed of 0.5 % w/w ammonium sulfate, 0.5 % w/w potassium hydrogen phosphate, 3 % w/w MES, 0.4 % w/w glucose, and a proprietary mixture of trace elements. The pH was adjusted to pH 7 before sterilization through a 0.2 µm filter. Yeast extract (BD #212750) and Bacto Tryptone (BD #211705) were added to the defined medium as indicated by adding a stock solution of 50 g/L that had been sterilized by filtration through a 0.2 µm filter. To grow the leucine auxotroph DH10B cultures in minimal media, L-leucine (Acros Organics #12512-1000) was added to the

media at 0.5 g/L, which was observed to be sufficient to attain maximal growth rates (data not shown). Induction of AND gates was performed with 1.3 mM L-arabinose (Sigma #A3256) and 0.63 mM sodium salicylate (Sigma #S3007), except as noted. Induction of NOR gates was performed with 1.3 mM L-arabinose (Sigma #A3256) and 100 ng/ml anhydrotetracycline (aTc) (Fluka #37919).

For all experiments reported, *E. coli* cultures frozen in 25% glycerol solution at -80°C were freshly streaked on plates of 2xYT agar and antibiotics. Fresh single colonies from these streaks were then cultured overnight at 37°C and 280 rpm orbital shaking in volumes and media as noted. Both shake flasks and tubes grow cells in oxygen-limited conditions, so for simplicity, all shake flask and tube replicate experiments are referred to as "shake flask" experiments. For shake flask experiments at DSM, overnight cultures were grown in 20 ml of 2xYT media in 100 ml flasks. After ~18 hours, the OD_{600} of these cultures were measured and they were then diluted back to an OD_{600} of 0.01 in 20 ml of pre-warmed media in 100 ml sterile flask. Shake flask cultures were then grown at 37°C and 280 rpm orbital shaking. For AND and NOR gate strains, induction occurred at the time of dilution. The cultures were then grown for 18-24 hours prior to measurement. For culture tube experiments at UCSF, all cultures were grown in 14 ml polystyrene culture tubes (Falcon #352059) at 37°C and 280 rpm. Overnight cultures were grown in 3 ml of LB media and antibiotic and were diluted back 1:1000 in 3 ml of media after 18 hours of growth. AND and NOR gate cultures were otherwise treated identically as in DSM shake flask experiments and were then measured by flow cytometry.

At DSM, OD_{600} measurements were performed on an Ultrospec 3100 Pro (Amersham Biosciences) spectrophotometer. At UCSF, OD_{600} measurements were performed on a Cary 50 Bio UV-Vis spectrophotometer (Varian). Cultures were diluted with sterile water until the OD_{600} fell between 0.10 and 0.60, the range across which OD_{600} correlates linearly with cell density.

4.3.3 Flow cytometry

At UCSF, cytometry was done on the BD LSRII flow cytometer and the FACSDiva software as follows. Samples of *E. coli* cultures that were to be measured by cytometry were diluted 1:5 in phosphate buffered saline (PBS) containing 2 mg/ml Kan to stop translation. A

488 nm laser was used for excitation and a 510/20 emission filter was used to measure Forward Scatter, Side Scatter, and GFP fluorescence. A 561 nm laser was used to excite and a 610/20 band pass filter was used to collect RFP fluorescence for RBS library measurements. Events were measured at a flow rate 0.5 µl/s until 50,000 events within the *E. coli* cell population gate were acquired. Analysis of the data was carried out in the FlowJo software package (Treestar). The cell population was gated by the forward and side scatter to include a maximal number of cells (>95%). At DSM, fluorescence was measured using a MoFlo cytometer/cell sorter. During fermentation, aliquots of 8 ml of whole broth was frozen at -20°C. These samples were later thawed and analyzed using the MoFlo cytometer. Upon thawing, each sample was diluted 1:10,000 in PBS and then run on the cytometer. Fresh and frozen samples showed no difference in scatter or fluorescense (data not shown). The 488 nm laser was used for excitation and a 510/20 band pass filter was used to collect GFP fluorescence. The data was exported and analyzed using FlowJo as described above.

4.3.4 Fermentation

Fermentation runs were performed at DSM in cylindrical fermentors with a total volume of 10 L, an internal diameter of 230 mm, two small baffles, and two heat exchangers. Two R6 Rushton turbines stirred the culture. All fermentations were carried out in series of two or three and were performed identically, except as noted (Table B.1). The initial 3 L batch phase contained minimal medium with 0.4% w/w glucose, 1% w/w yeast extract, and 0.05% w/w antifoam. A 100 ml shake flask culture was grown on 2xYT media and antibiotics (concentrations as noted above) overnight and used to inoculate the batch phase at $t = 0$. A 3 L feed containing 10% w/w glucose, antibiotics, and supplementary yeast extract as noted was started after the glucose in the batch was exhausted, which occurred roughly 18 hours after inoculation. The initial feed rate was initially 5.5 g/hr, was increased exponentially to maintain a fixed growth rate, and stopped 42 hours after start, after which time the feed rate was kept constant at 45 g/hr. pH was set at 6.8 and was stabilized throughout fermentation by automatic feeds of 10% ammonia and 4 N sulfuric acid. Dissolved oxygen (DO) was initially 100% saturation and was allowed to drop to 50%. Off-gas was analyzed in real-time by mass

spectrometry. Profiles of the feed rates, pH, DO, respiratory quotient (RQ), and oxygen consumption rates for the length of each fermentation are reported in Figure B.10.

The media for the batch and feed were prepared in fractions and contain the following unless otherwise noted in Table B.1. Batch fraction 1 was prepared in a clean fermentor and contained 10 g/kg yeast extract, 2 g/kg K_2HPO_4, and 0.5 g/kg antifoam. Batch fraction 2 contained 4% glucose, 5 g/kg $(NH_4)_2SO_4$, and trace mineral mix (proprietary). Batch fractions 1 and 2 were combined following autoclaving. Feed fraction 1 contained any supplementary yeast extract and salts added to the feed. Feed fraction 2 contained 400 g/kg glucose. Feed fractions 1 and 2 were combined following autoclaving. Antibiotics were prepared in 1000x stock solution, filter sterilized, and added to the batch fraction only (Runs #1-3, Table B.1) or the batch fraction and the glucose feed (Runs #4-10, Table B.1).

A computer monitored the weight of the glucose feed and the pH titrants, temperature of the culture, and the air input. Temperature of the culture was controlled by a cooling finger and a heat lamp and maintained at 37°C. The dissolved oxygen of the culture was adjusted by controlling the speed of the impellers (Min 200 RPM, Max 750 RPM) and a constant airflow. Samples were taken immediately before addition of the inducers and at accessible time points throughout fermentation. Approximately 50-100 ml of sample were taken from each fermentor at each time point and analyzed. The samples were weighed and the amount was recorded in the central computer, which adjusted the feed rate parameters based on the lost weight. Each sample was then analyzed for culture biomass, OD_{600}, fluorescence, and pH. Three 3 ml aliquots of each sample and the supernatant of each sample were frozen at -20°C for later analysis. OD_{600} measurements of each sample were made in triplicate as described above. GFP fluorescence of each sample was measured by flow cytometry as described. The dry cell weight (DCW) was measured as follows. An aliquot of 10 ml of culture was weighed in a 50 ml Falcon conical tube that had been pre-dried at 105°C for 24 hours, diluted with 40 ml of distilled water, and then centrifuged at 6000 rpm for 10 min. The supernatant was poured off, and the pellet in the tube was dried for 24-30 hours at 105°C and weighed. The reported DCW is the weight of the pellet divided by the total weight of the sampled culture.

4.3.5 BioLector Microreactor Experiment

A batch culture of *E. coli* DS68637 carrying the RBS_B AND gate and the reference plasmid (pFM46) were grown in the BioLector CC (m2p labs) instrument using a sterile 48-well flower plate (m2p labs MTP-48-BOH). Each 48-well clear-bottom flower plate had a 1.0 ml capacity/well and contained optodes for monitoring pH and dissolved oxygen. Automated spectrophotometry monitored fluorescence of the entire culture. Culture density was monitored by light scattering and is reported in arbitrary units. One colony from a freshly streaked 2xYT agar plate was used to inoculate a 100 ml culture of the RBS_B AND gate or the reference plasmid. 1 ml of culture was then aliquoted into the 48-well flower plate. Each well of inoculated media contained either no inducer, only one inducer (1.3 mM arabinose or 0.63 mM NaSalicylate), or both inducers. Culture density, fluorescence, pH, and DO were monitored for 40 hours. Additional details of the microfermentation are presented in Appendix B.

4.4 Results

4.4.1 Media Dependence of the Logic Gates

Media selection for fermentation requires a balance between productive growth and minimizing the cost of components[189]. Complex ingredients, such as yeast extract and tryptone, can boost product formation but are avoided because such components are expensive, complicate product recovery, reduce predictability of fermentations and control over metabolism[190]. To determine the impact of media composition on circuit performance, we characterized each circuit in complex and minimal media in shake flask experiments using the lab strain *E. coli* DH10B (Figure 4.2).

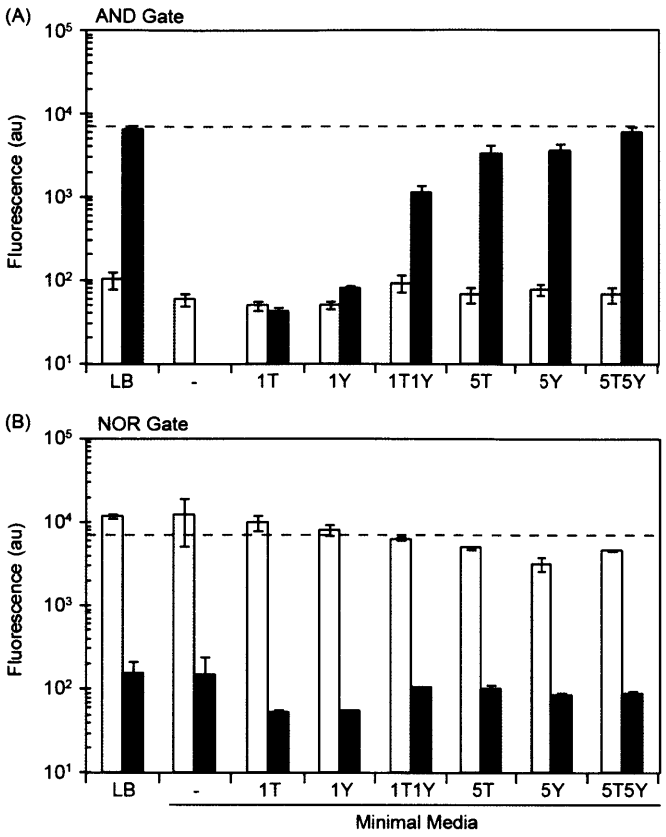


Figure 4.2: Impact of Media on Gate Performance in *E. coli* DH10B. Media composition affects the performance of the AND (RBS_B) and NOR gates. *E. coli* DH10B carrying the **(A)** AND or **(B)** NOR gate were grown both uninduced (white bars) and induced (black bars) in media of varying composition. The media composition is listed below the data. LB-Miller (LB) contains 10 g/L Tryptone, 5 g/L yeast extract, and 10 g/L NaCl. The other media is either unsupplemented minimal media (-) or minimal media supplemented as follows: 1 g/L Yeast Extract (1Y), 1 g/L Tryptone (1T), 5 g/L yeast extract (5Y), 5 g/L Tryptone (5T), 1 g/L yeast extract + 1 g/L Tryptone (1Y1T), 5 g/L yeast extract + 5 g/L tryptone (5Y5T). The output of the reference plasmid pFM46 (dotted line) is shown for comparison. All cultures were measured after 9 hours, except the cultures that were grown on minimal (unsupplemented) media, which were measured after 24 hours. Induced AND gates never grew on minimal media.

E. coli DH10B is a common strain for genetic circuit development[191]. However, DH10B is a leucine auxotroph and contains the *relA1* and *spoT* alleles, which are known to lower growth rate, especially during nutrient downshifts[192],[193]. In LB media, the AND gate exhibits a strong 64-fold induction (Figure 4.2A). However, in the minimal media the strain does not grow upon induction (Figure B.1). When a small amount (1 g/L) of yeast extract or tryptone is added to supplement growth, the strain grows to saturation, but no AND activity is observed. Activity can be recovered by adding additional yeast extract or tryptone to the minimal medium. The addition of 1 g/L of either recovers growth when the AND gate is induced, but comparable activity to LB is not observed until 5 g/L each of tryptone and yeast extract are added.

In contrast, the NOR gate remains functional in all media tested, showing 76-fold induction in LB that is preserved in minimal medium (Figure 4.2). Changing the complexity of the media by adding various amounts of tryptone and yeast extract has little effect on either the induced or basal states of the NOR gate. Furthermore, no growth defect is observed in

minimal media in either the induced or uninduced state (Figure B.1). The shape and width of the cytometry distributions do not change over these conditions (Figure B.3).

4.4.2 Gate Function in an Industrial Strain

E. coli is a common host for the industrial production of recombinant proteins[194,195,196,197]. *E. coli* RV308 was first applied to the production of insulin in 1982 and has been used for the production of enzymes, proteases, and therapeutic proteins by Eli Lilly and Merck[198]. We examined the activity of the NOR and AND gates in *E. coli* DS68637†, a modified *E. coli* RV308 variant similar to a strain used at DSM. Unlike *E. coli* DH10B, the AND gate functions as expected in *E. coli* DS68637† in minimal media without the addition of complex ingredients. However, the output of the induced state is reduced to 26-fold as compared to 64-fold for *E. coli* DH10B grown in LB. The magnitude of the uninduced states is unchanged between media and strains.

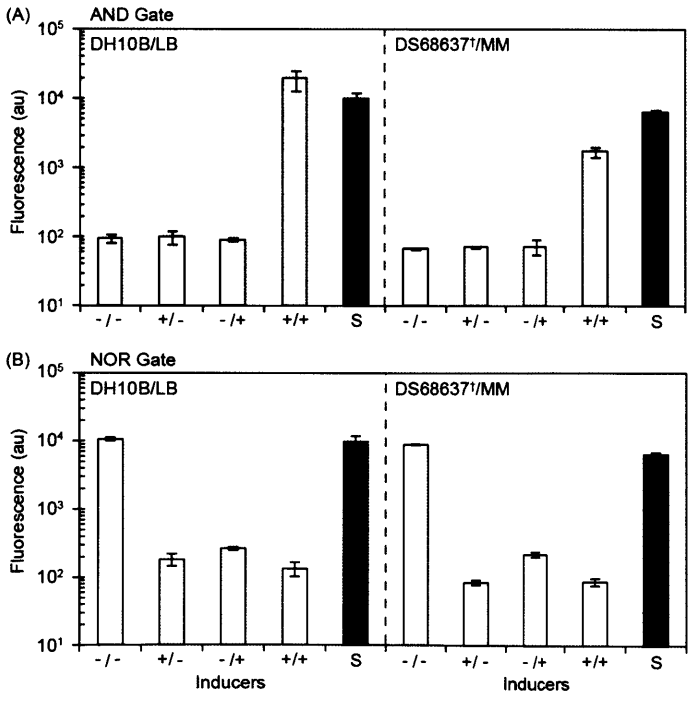


Figure 4.3: Comparison of Gate Performance in *E. coli* DH10B and *E. coli* DS68637†. Gate performance is measured in DH10B grown on LB and in DS68637† grown on unsupplemented minimal media (MM). **(A)** The output of the AND gate (RBS_B) is shown for four combinations of inputs: -/- (no inducers), +/- (1.3 mM arabinose), -/+ (0.63 mM salicylate), +/+ (both inducers). The output of the reference plasmid pFM46 (S, black bars) is shown for both strain/media combinations. **(B)** The output of the NOR gate is shown for four combinations of inputs: -/- (no inducers), +/- (1.3 mM arabinose), -/+ (100 ng/ml aTc), +/+ (both inducers). In *E. coli* DH10B, the AND gates were measured after 9 hours. All other strains were measured after 24 hours. The error bars represent the standard deviation of three experiments performed on different days.

We used a standard reference plasmid (pFM46) to compare the absolute ON and OFF states between strains[199]. The fluorescence produced by this reference plasmid is nearly

identical across strains and media (Figure 4.3). While the AND gate function is preserved, the magnitude of the output in the presence of both inducers (+/+) changes between strains. After normalizing to the reference plasmid, this difference is determined to be approximately 7-fold. This could pose a problem when connecting the output of this gate to downstream devices and could require the selection of a different connecting part, such as a ribosome binding site of a different strength.

The NOR gate functions nearly identically in *E. coli* DS68637† in minimal medium as in *E. coli* DH10B in LB (Figure B.2). The ON state (-/-) of the gate only differs by 16% between strains. Thus, the NOR gate produces a reliable output irrespective of the strain or media.

4.4.3 Connecting Genetic Circuits: Impact of Media and Strains in Shake-flask

Connecting genetic circuits requires that the output of the upstream circuit matches the input required to activate the downstream circuit[200],[201]. A common approach to connect circuits is to vary the ribosome binding site (RBS) sequence downstream of the output promoter. A potential problem could emerge during scale-up when a strain contains a genetic program. If the transfer function of the circuit changes at each stage of scale-up (i.e., is different when measured in a bioreactor compared to shake flask experiments), then this could require a different RBS to functionally connect it to a downstream circuit. If circuits were to require RBS tuning at each stage of scale-up, this would limit the implementation of multi-circuit programs in industrial processes.

In previous work, we analyzed the connection of an arabinose-inducible promoter to an input of the AND gate[202]. We tested multiple RBSs of different strengths and found there was an optimal RBS strength to connect the input promoter to the circuit. Here, a subset of these RBSs was chosen to encompass the transition from functional to non-functional gates (Figure 4.4). First, we tested whether RBS strength changes as a function of media and strain. Using the J23100 constitutive promoter and RFP reporter, the strength of six RBSs was measured in the context of *E. coli* DH10B in LB and *E. coli* DS68637† in minimal medium (Figure B.5). RBS strengths were nearly identical across these contexts.

The RBS variants of the AND gate were then tested for their ability to functionally connect the arabinose-inducible promoter to the AND gate (Figure 4.4). The function of the AND gate is measured for each combination of inducers (-/-, -/+, +/-, +/+) and this is used to assign a "fitness" to the gate. While the magnitude of the output of the gate changes, the rank order of the RBSs is similar. This indicates that while the magnitude of the AND gate changes, the same RBS is optimal in connecting the input promoter to the gate.

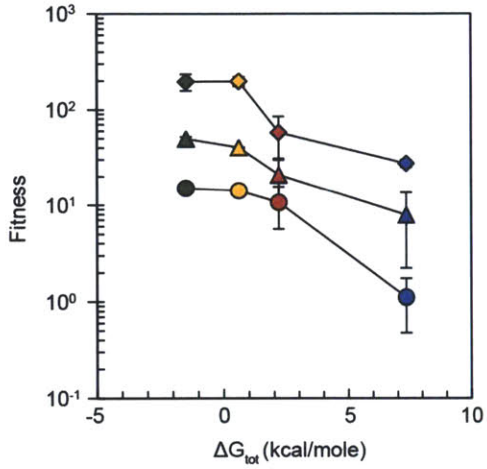


Figure 4.4: Media and Strain Impact on RBS selection. The effect of varying the strength of the RBS connecting the arabinose sensor to the AND gate is shown. The ΔG_{tot} is the strength of the RBS as determined using a biophysical model. A more positive ΔG_{tot} predicts strong secondary structure formation around the RBS, which correlates strongly with weaker translation. Fitness scores the accuracy and function of the gate in each condition. The RBSs characterized in this manuscript are colored (RBS_A: green, RBS_B: orange, RBS_C: red, and RBS_D: blue) and were previously characterized by Salis, et al. (2009). The calculated fitness is shown for the four RBS's studied for different media and strains. The media are LB broth (diamonds), minimal media containing 5 g/L yeast extract (5Y, triangles), and unsupplemented minimal media (circles). *E. coli* DH10B was measured in LB and 5Y, and *E. coli* DS68637[†] was measured in unsupplemented minimal medium. Error bars represent 1 standard deviation of three experiments.

4.4.4 Circuit Dynamics in a 1 ml MTP Microreactor

We tested the AND gate in a high-throughput microreactor to assess performance in another context of industrial process development[203]. *E. coli* DS68637 containing the AND gate in different induced states was grown in 2xYT broth for 40 hours (Figure 4.5A). Differences in the fluorescence of each culture were detected after 3 hours. After 15 hours, no further changes in cell density or fluorescence were observed and remained stable throughout the remainder of the experiment (Figure B.6). The AND gate showed partial induction with arabinose (+/-) but was only fully induced in the presence of both inducers (+/+). This partial induction is most likely due to leakiness of the P_{sal} promoter in the absence of salicylate[204]. Correcting this would require a weaker RBS connecting the P_{BAD} promoter to the gate. The absolute magnitude of the output states was compared to the shake flask experiments by

comparing expression to the reference plasmid pFM46, which also produced a stable fluorescence per cell culture density in stationary phase (dashed line in Figure 4.5).

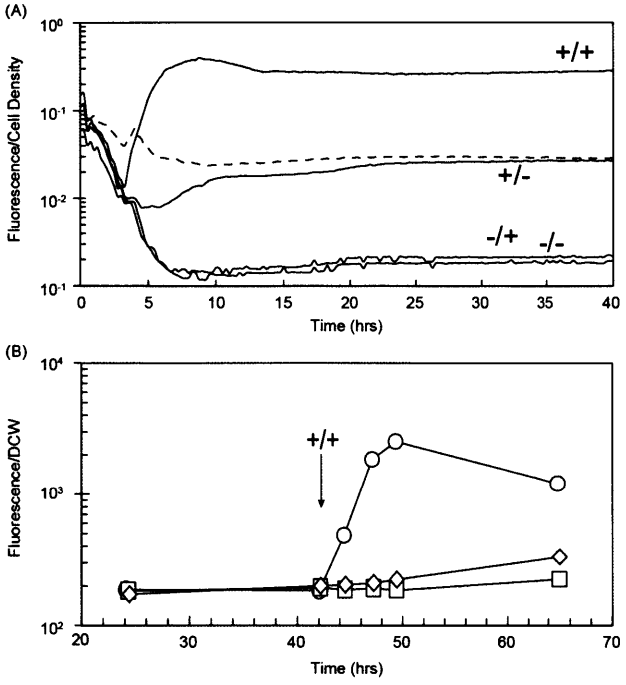


Figure 4.5: Performance of an AND gate and Reference Plasmid in a Microreactor and 10 L Bioreactor. Cultures of *E. coli* DS68637 carrying the RBS_B AND gate and a reference plasmid (pFM46) were grown in a BioLector microreactor on a 1 ml batch of rich 2xYT medium. **(A)** Cultures were induced at time 0 with either no inducer (-/-), single inducers (+/-, arabinose only; -/+ salicylate only) or both inducers (+/+). The fluorescence, cell density, dissolved oxygen, and pH of each culture was monitored for 40 hours. The reference plasmid data is shown as a dashed line. **(B)** The performance of *E. coli* DH10B carrying the RBS_B AND gate is shown in a 10 L bioreactor. Both inducers (1.30 mM arabinose and 0.63 mM salicylate) are added at 42.5 hours. The fluorescence per dry cell weight (DCW) is shown for three fermentations in which the amount of yeast extract in the feed is varied: 0 g/kg (squares), 20 g/kg (diamonds), and 100 g/kg (circles).

4.4.5 Circuit Dynamics in a 10 L Bioreactor

The AND gate was tested in a 10 L bioreactor under fed-batch conditions, a common context for process development of industrial recombinant protein production[205]. Variables that indicated circuit performance, including cell density, dry cell weight (DCW), and fluorescence, were measured after sampling the culture every 3-4 hours (Appendix B). The performance of the AND gate was first measured in *E. coli* DH10B cells (Figure B.6B). The culture was maintained in log phase by exponentially increasing the feeding rate (Table B.1). The cells were induced after 42 hours by adding both inducers. The AND gate in DH10B only functioned when the feed included a large amount of complex media (100 g/kg yeast extract). Even the addition of 20 g/kg yeast extract to the glucose feed showed no activity when induced.

Next, the industrial strain *E. coli* DS68637 was used to characterize the AND gate in the 10 L bioreactor (Figure 4.7). As was observed in shake-flask experiments, the addition of yeast extract was not required for AND gate function and did not affect the performance of the gate (Figure B7A). The glucose feed was added at an exponentially increasing rate in order to keep

the culture at a constant growth rate. Inducers were added to the cultures 20 hours after exponential feed was initiated. The feed rate was shifted from exponentially increasing to constant 45 hours after addition of the culture[206]. This caused the growth rate to decrease, and cells to undergo a shift from a constant growth rate of 0.05 h^{-1} to very low growth rates. During exponential growth, the circuit rapidly turns on after both inducers are added and the ON state remains stable. In addition, the population of cells is narrow and nearly all of the cells are induced.

However, the circuit breaks late in fermentation. The cell-to-cell variability of the circuit output increases, a significant OFF population appears, and the circuit is almost completely deactivated by the end of fermentation. Plasmid loss assays show that the majority of plasmid is lost by 30 hours (Figure 4.6B). The deactivation of the circuit correlates with the switching of the glucose feed from exponential to constant. An exponential feed rate was started again 70 hours after inoculation to determine if AND gate activity could be rescued by restarting the feed (Figure 4.7). No significant change in fluorescence was detected after reactivation of exponential feeding, which is consistent with the cells losing their plasmid(s).

The alternative RBSs used to connect the input promoter to the AND gate were also tested in the 10 L bioreactor (Figure 4.7). The rank order of the RBSs remained the same as compared to the shake flask experiments (Figure 4.4); however, there are some differences. Notably, RBS_B and RBS_C produce nearly the same induction and RBS_D responds more strongly in the bioreactor (~10x) than in shake flask experiments (~2x; Figure B.4). This implies that the arabinose-inducible promoter is producing a higher output, thus requiring a weaker RBS to connect with the gate. Interestingly, the fully-induced state of the RBS_D circuit shows a biphasic distribution with approximately 50% of the cells being induced. Those cells that are in the ON state produce the same output as the stronger RBSs. Thus, when the weaker RBS is used, a subset of the cells is able to properly function, but the remainder exhibits an error that would propagate to a downstream device.

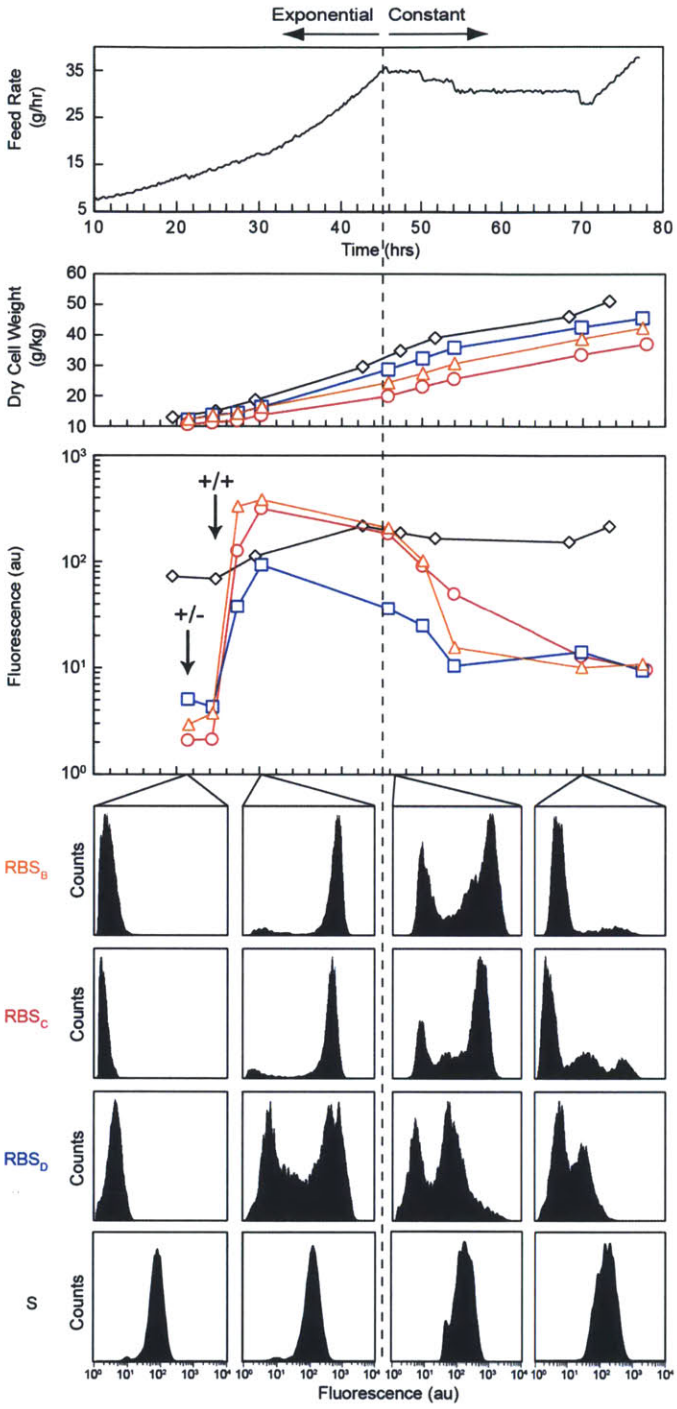


Figure 4.6: Performance in Fermentation of AND Gate RBS Variants. Three fermentations of RBS variants of the AND gate in *E. coli* DS68637 were performed. AND gate variants included either a strong (RBS_B, orange triangles), medium (RBS_C, red circles), and weak (RBS_D, blue squares) RBS, which correspond in color to RBS's tested in Figure 5. These three fermentations were performed identically, with changes in feed rate and DO made as needed to match growth rates of the cultures. AND gate cultures were induced at 21 hours with 0.63 mM sodium salicylate (+/-) and then with 1.3 mM arabinose (+/+) at 24 hours. The feed rate was exponentially increased until 45 hours and then switched to a fixed feed rate; the dotted line marks the time of the switch. Samples were taken at 9 different times throughout fermentation. Dry cell weight (DCW) and fluorescence of each culture were measured at each time point. Fluorescence cytometry distributions are shown for the circuit in the OFF state (21 hours), when fully induced (30 hours), immediately after the shift to constant feeding (46 hours), and after 70 hours. For comparison, performance of an *E. coli* DS68637 strain carrying the reference plasmid pFM46 (S; black diamonds) is shown. The reference plasmid fermentation was carried out on a separate day and was sampled at 20, 30, 47, and 68 hours after inoculation. Data for this figure was gathered from Runs #6-8 for the AND gates and from Run #11 for the reference plasmid (Figure B.7, Table B.1).

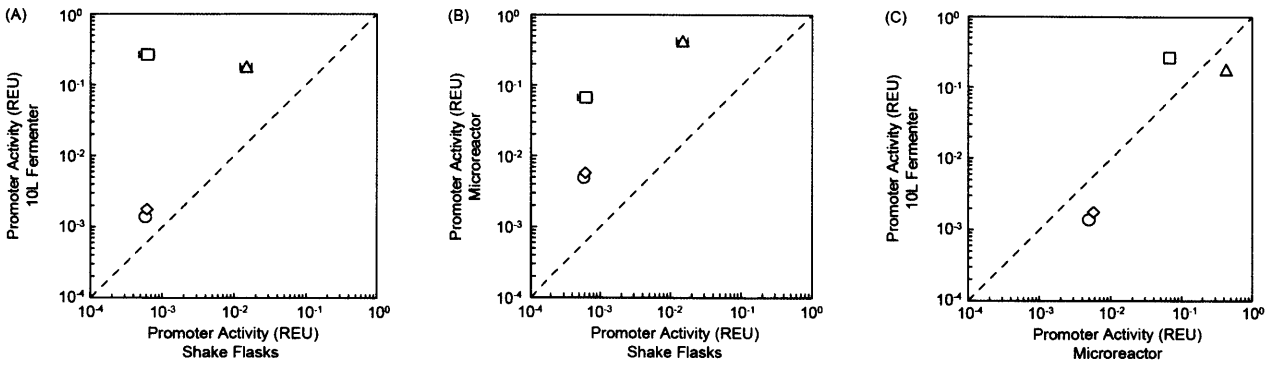


Figure 4.7: Comparison of AND gate activity across growth conditions. To compare AND gate expression across different growth conditions (shake flasks, 1 ml microreactor, and 10 L bioreactor), GFP fluorescence output of the RBS_B AND gate was converted to relative expression units (REU) defined by the Kelly standard plasmid (SI Section VIII). This relative expression of the AND gate is compared in each pair of the following conditions: **(A)** shake flasks versus 10 L bioreactor, **(B)** shake flasks versus 1 ml MTP microreactor, and **(C)** 1 ml MTP microreactor versus 10 L bioreactor. The four different states (-/-, circles; ara/-, squares; -/sal, diamonds; +/+, triangles) are plotted. The dotted line represents a theoretically perfect correlation between states. The shake flask cultures are unable to predict the ara/- failure mode in the 1 ml microreactor and the 10 L bioreactor. However, the 1 ml microreactor predicted the ara/- failure mode of the gate in the 10 L bioreactor. Data for the shake flask cultures is taken from DS68637^{+} grown on unsupplemented minimal media (Figure 4.4). The data for the microreactor corresponds to the mean fluorescence of each culture between 10 and 40 hours in Figures 4.6 and B6.

The AND gate containing RBS_B was tested by adding salicylate and then arabinose at a later time point. The addition of salicylate alone is not able to induce the gate (Figure 4.7). However, the gate is induced with arabinose alone (Figure B.7B). This is indicative that the RBS controlling T7 polymerase is too strong and therefore the gate fails and is only responsive to a single input. This implies that the optimum RBS has shifted to be weaker when comparing performance in shake-flask and 10 L bioreactor experiments. This may be predictable based on the microreactor data, where cells only induced with arabinose have an intermediate output between uninduced and fully induced states (Figure 4.6).

Cells carrying the AND gate accumulate less biomass. The different RBS variants of the AND gate yielded the following biomass accumulation rates: 0.49 g/kg media/hr (RBS_B), 0.56 g/kg media/hr (RBS_C), and 0.64 g/kg media/hr (RBS_D). The rate of cells containing the reference plasmid is 0.71 g/kg media/hr. Considering the reference plasmid as a control, a single functional AND gate can reduce biomass by as much as 67%. In the microreactor, we observed

a similar effect where the induction of the gates caused a lower final OD in the AND gate (Figure B.6).

We used a reference plasmid (pFM46) to compare the magnitude of ON and OFF states across strains and growth conditions. The data was normalized to the Kelly reference standard and is reported as Relative Expression Units (REU; Figure B.9). The reference plasmid was also measured in the microreactor (Figure 4.5A) and the 10 L bioreactor (Figure 4.7 and Figure B.7). The fluorescence per cell and the population distribution is fairly stable over the 72 hour fermentation, even after the transition from exponential to constant feed rates (Figure 4.6). In addition, very little of this plasmid was lost by the end of fermentation (Figure B.8). Using the reference plasmid, we calculated the output of the AND gate to be 0.3-3.1 REU in shake-flask experiments, 6.1 REU in the microreactor, and 2.2 REU in a 10 L bioreactor (Figure B.8). Therefore, while the circuit is functional under all conditions, the magnitude of the output varies considerably. Notably, the AND gate fails in both the microreactor and 10 L bioreactor in the same way. In both cases, the gate is non-responsive to salicylate and the +/- state is ON when it should be OFF. In the microreactor, it is not a true failure as the +/- state occurs in-between the ON and OFF states in a way that is analogous to fuzzy AND logic (Figures 4.5 and 4.7). The microreactor is better correlated for the absolute REU measurements, the early detection of a failure mode, and measurements where carrying the circuit impacts the OD (Figure 4.7). Therefore, it is a better method for predicting the performance of a genetic circuit than shake flask experiments.

4.4.6 Oxygen Sensors Function in Industrial Fermentors

Because the oxygen sensors were intended for eventual application in industrial conditions, we tested how the oxygen sensors described in Chapter 3 would behave in industrial fermentation conditions. Typically, industrial fermentations use large scale (>1 L) cultures, minimal media, carbon-limiting feed, and more robust strains of *E. coli* during scale-up to production volumes[207]. These conditions are historically very different than what has been used in lab to develop genetic programs and sensor modules. Lab conditions typically use smaller scales (<4 ml), rich media, oxygen limiting conditions, and strains of *E. coli* with recA

knocked out to stabilize genetic constructs and prevent recombination. To test how the oxygen sensors behaved during scale-up from lab to industrial conditions, we grew three oxygen sensor strains in similar conditions to those described above.

To test oxygen sensor performance in industrial fed-batch process, we grew the previously tested FixL/J and FNR sensors in a 6 L glucose-fed fermentation (Figure 4.8). Cultures were grown similarly to the AND and NOR gate circuits as described, except that the dissolved oxygen (DO_2) levels were changed following beginning of the feed process. The DO_2, as measured by an external oxygen probe, was lowered by slowing the airflow through the culture and the impeller stir speed. DO_2 was brought from an initial 50% to 10% and then to 1% (Figure 4.8). The cultures were left at 3% DO_2 overnight, and the final DO_2 was brought as low as possible to 0%. Samples were taken immediately before the DO_2 was changed at each point and analyzed as described.

Both the FNR and FixL/J sensors showed a weaker dynamic range in the fed-batch process than in the small batch processes in which they were initially characterized. The working sensors increased their fluorescence output as the DO_2 levels were decreased by up to 4-fold (Figure 4.8). Even after the cultures were grown at 3% DO_2 overnight, which caused the lysis of much of the culture, the fluorescence signal increased markedly when DO_2 was lowered to 0%. Although the dynamic range of the oxygen sensors is smaller than in lab conditions, we were unable to assess this accurately because the 0% DO_2 point was only assessed after the culture had begun to lyse. Also, it is difficult to compare the performance characteristics of the oxygen sensors in lab to those observed in the fermentor. This is mainly because lab assays were performed at a single oxygen concentration and assessed at a single time point, unlike the fermentations. Because the fermentations were assessed at several time points after changing the DO_2, the EcFbFP reporter could accumulate in the cells to give higher values of fluorescence at later time points than would have been measured had the DO_2 been initially changed to the indicated level. Nonetheless, the oxygen sensors showed at least a 4-fold dynamic range at 1% DO_2, and the approximate transfer function for the fermentation showed a similar exponential increase in fluorescence as DO_2 approached 0% (Figure 4.8). The second FNR sensor could not be effectively assessed because an external oxygen sensor broke during fermentation and

confounded the true DO_2 values in the culture (Figure 4.8). Although further testing is needed before definite conclusions can be drawn about the exact characteristics of the function of these oxygen sensors in industrial conditions, we have shown that oxygen sensors are functional in industrial environments and that they can achieve dynamic ranges sufficient for integration with downstream circuitry. Thus, the oxygen sensors we have developed are useful as input modules in synthetic genetic programs in industrial conditions.

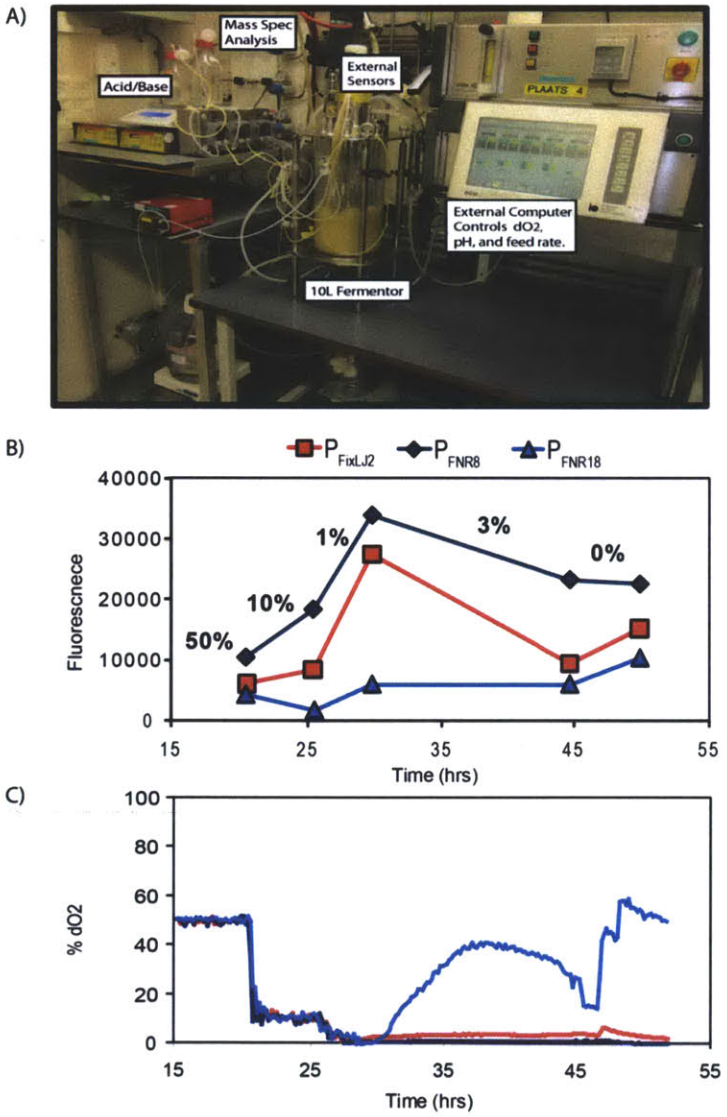


Figure 4.8: Performance of oxygen sensors in fermentation conditions (A) Picture of a fermentation setup used at DSM to test the oxygen sensors. Critical elements of the setup are labeled. An external controller uses pH and oxygen probes to detect the conditions in the culture. When conditions begin to change, acid/base feed and changes in airflow and impeller speed are used to adjust pH and dissolved oxygen, respectively. Mass spectrometry analysis of the CO2 efflux from the culture enables real-time calculation of the culture's metabolic activity. When the culture is observed to become carbon limited, a 100 g/L glucose feed is automatically started. **(B)** The time plot of the fluorescence of the culture of each oxygen sensor shows that fluorescence of each culture increases as the % DO_2 is lowered. Because the culture was left at 3% DO_2 overnight, the culture began to ferment and lyse, causing a decrease in fluorescence. Nonetheless, when the %DO_2 was again lowered to 0%, a rise in the fluorescence was detected. This is a strong indicator that the oxygen sensors were functioning well throughout growth in industrial conditions. **(C)** The time plot of each culture's % DO_2 shows that the FNR8 culture malfunctioned due to a broken oxygen probe but that the other cultures' % DO_2 is well-controlled throughout growth.

4.5 Conclusions

Genetic programs can implement computational control over cellular functions, including metabolic processes. They are being applied towards: 1. controlling the timing of gene expression at different stages of growth, 2. implementation of feedback, 3. transfer of process control into individual cells, 4. consolidation and control of multi-step bio-manufacturing, and 5. diversification of processing tasks amongst multiple cells. Accomplishing these goals will require increasingly sophisticated programs based on the integration of many circuits.

A challenge in developing genetic programs for an industrial process is the variation that occurs at each stage of scale-up from shake-flask experiments to production-scale bioreactors[208]. Strains and pathways vary in their performance during scale-up and it is expected that similar issues will arise with genetic programs. In this work, we show that a circuit can fail (produce an incorrect computational operation) when moved from shake-flask (uncontrolled batch fermentation) to 10 L bioreactor (controlled fed-batch fermentations at relatively low growth rates) experiments. Further, this circuit exhibits variable responses in industrially-relevant media and strains. To our knowledge, this work represents the first example in which a simple genetic circuit constructed in an academic synthetic biology lab is characterized in the early stages of process scale-up with an industrial partner.

Several efforts are underway to apply high-throughput fabrication to synthesize thousands of genetic parts and circuits[209,210]. Increasing the throughput will require decreasing the scale at which they are individually characterized, possibly even applying microfluidic screens. A current challenge is designing those assays such that they convey the most valuable information for their incorporation into varied genetic and environmental contexts. As an early step towards this goal, it has been proposed to use a reference construct as a standard for reporting promoter activity[211]. In this work, we found an expression standard to be a useful tool for comparing gate performance at each stage of scale-up and across industrial and academic labs. We found that the particular promoter-RBS-reporter-backbone choice made by Kelly *et al.* (2009) produces a reliable response in varied environmental contexts. By converting the output to Relative Expression Units (REU), we could quantify a 20-fold variability in the AND gate across environments and labs (an effect that could be qualitatively visualized by eye).

The AND gate contains parts that are toxic and reduce the growth rate. This has a minor impact in shake flask experiments using complex media. However, a notable reduction in the biomass accumulation was observed in 10 L bioreactor experiments. This is particularly problematic considering that desired genetic programs will likely require many integrated circuits. Interestingly, the strength of the RBS connecting a sensor to the AND gate changes the load, where those that yielded the best gates also caused the strongest inhibition of growth. Thus, the optimal RBS has to balance circuit function and growth, a trade-off not previously described.

The microreactor was able to predict one of the failure modes of the AND gate. In the 10 L bioreactor, the gate is not responsive to the salicylate input. This effect was previously observed in shake-flask experiments when the RBS connecting the leaky P_{sal} promoter to the circuit was too strong. Additional sources of failure may have been that the P_{BAD} promoter was stronger or the translation rate of the T7 RNAP gene was higher. The observation of such failure modes due to changes in promoter strength across conditions highlights the need for genetic parts that are insulated from changes in the environment. Although design of such insulation is not yet reliable, some design rules are becoming elucidated[212,213,214,215].

Neither the shake-flask nor microreactor experiments were able to predict the rapid decline in output that occurs after the shift from exponential to constant feed rates. The *E. coli* DS68637 strain has been observed to have a lower plasmid copy number[216] and plasmid stability is a problem[217,218]. Indeed, one of the plasmids was rapidly lost after the shift and this resulted in the inactivation of the gate (Figure B.8). Such evolutionary instability is a common problem of the plasmid genetics, but it can also be hastened by the presence of toxic parts within a circuit and can even depend on the induction state[219]. As technologies for editing the prokaryotic genome expand, unstable plasmid systems such as the ones studied here will be replaced by direct-to-genome programming[220].

In this work, standardized measurements are critical in comparing gate performance across strains, growth conditions, and labs. The wider adoption of standards will enable the more rapid determination of circuit failure modes. Further, it will aid the interpretation of genetic part data gathered at increasingly small scales in fabrication labs and then applied to

problems in varied applications, strains, and environments[221]. Even amongst projects within our own lab, and in moving strains between UCSF, MIT, and industrial partners, we have found it challenging to unify the "standardized" measurements made by individual researchers[222],[223]. Truly realizing the potential will require the development of large, dedicated consortia of industrial and academic labs.

5. Conclusions and Discussion

This work demonstrated the partial development of a synthetic genetic program for the control of acetate production in *E. coli*. The experimental scope of this project spans the characterization of basic components to complete circuit integration and performance validation at 10 L scale. Together, this thesis represents a vertical workflow of the design, testing, and validation of novel biological devices and programs. Future work will further focus on the completed construction and validation of this program.

Due to the initial lack of characterized sensors for the conditions of interest, much of this work revolved around the discovery, troubleshooting, characterization, and integration of those sensors. The development of the methylation sensor in *E. coli* and *S. cerevisiae* served as a useful exercise in the characterization and tuning of a novel sensor. For the methylation sensor, a reliable strategy for improving the detection threshold of a sensor was to overproduce the sensor protein. This eventually came with a tradeoff of a higher basal activity from the cognate promoter and an increased metabolic burden. However, overexpression of the acetate sensor's NRI protein did not improve its detection threshold and even repressed the target promoter[224]. Overexpression of global regulators such as FNR and Cra was not considered due to the systemic misregulation that is incurred when a global regulator's expression levels are dramatically altered. Successful construction of synthetic promoters that were more responsive than native promoters demonstrated that promoter engineering and optimization were tractable strategies when developing sensor promoters.

The troubleshooting of the acetate sensor highlighted a significant problem encountered when using fluorescent proteins as reporters of sensor and circuit performance. The initial tested version of the acetate sensor showed a very weak (2-fold) dynamic range. Later, insertion of an insulation element in the 5' UTR of the transcript greatly improved the dynamic range to over 10-fold. Further analysis revealed that the RBS of the first version of the sensor was simply too weak to produce measurable amounts of reporter GFP. Thus, the initial measured dynamic range of the acetate sensor was "weak" due to our inability to measure its true basal rate of expression. This was primarily due to limitations in our instrumentation. Although cytometers, commonly the preferred instrument for fluorescence detection, are

highly sensitive when optimized (>80 molecules of FITC equivalent fluorophores)[225], they still require relatively high expression levels of fluorescent proteins in order for the signal/noise ratio to be significant. This becomes a problem when assessing the function of weakly expressed systems. Consider the following example. A sensor promoter has a basal activity that generates 10 molecules/cell at steady state. Upon induction, the sensor upregulates transcription by 20-fold, generating 200 molecules/cell at steady state. Unfortunately, the cytometer can only detect >100 molecules/cell. Therefore, when 100 molecules/cell are present (the sensor promoter is already 10-fold induced), the cell's signal is identical to the fluorescence background ("white cell" levels). In our example, the sensor promoter is discarded as having a weak dynamic range (2-fold), despite its desirable low basal activity and true 20-fold induction.

To avoid such a mistake, one can use a stronger RBS to increase the basal number of fluorescent proteins generated by the promoter over the detection threshold of the instrument. However, overproduction of fluorescent proteins can be toxic to cells either because of an incurred metabolic burden or direct toxicity of the protein[226]. Several fluorescent reporters, such as EcFbFP used here, are very dim or emit in channels with high background fluorescence. These proteins need to be expressed in very large numbers (1000's) to be detectable above background, risking toxicity from overexpression. For systems targeting improved yields of metabolites and growth performance, such as the program presented in this thesis, this can be a serious issue during construction, tuning, and performance assays. Alternatively, dual reporter systems can be used that enable enzymatic assays in conjunction with fluorescence assays. Enzymatic assays, however, are difficult to do in high throughput and require measurement of cell density. Another solution is to use more sensitive instrumentation, such as fluorescence microscopy. However, this type of instrumentation requires advanced skill to use and is difficult to implement in high-throughput.

All sensors used here were sensitive to changes in host metabolism. This was a predictable problem, since all sensors depend on the production of molecules, which are susceptible to changes in growth rate, resource availability, and ambient conditions (e.g. pH). Constitutive promoters showed a decrease in steady state transcription levels at lower oxygen

concentrations, suggesting that the output of the oxygen sensors was susceptible to metabolic conditions as well as oxygen levels. Lower ambient oxygen reduces the available ATP in the cell, generates mixed acid waste, and slows the growth of the cell, which in turn changes the ratio of sigma factors in the cell. The acetate sensor's output changed as the cells were subjected to toxic levels of acetate. Above 60 mM (~3.5 g/L) acetate at pH 7, *E. coli* BL21 showed slower growth during response function assays, which correlated with a lower fluorescence readout from the acetate sensor (Figure 3.5). Additionally, the acetate sensor's detection threshold decreased and its susceptibility to acetate toxicity increased with the ambient pH.

The sensitivity of sensors to changes in metabolism can lead to false reporting of cell state. Typically, the result is underreporting. Toxicity or lack of resources will lead to a lower expression of the output protein. The attenuation of the dynamic range of the oxygen sensors during fed-batch fermentation may be partly explained by the effects of scale-up. In the context of a circuit, this may also lead to failure of connections (Chapter 5). This may be avoidable if the sensor's dynamic range is large enough for any change due to metabolism to be only a small fraction of the overall dynamic range. Another strategy to make sensors and programs more robust to metabolic changes is to use components that require a minimal input from metabolism, making them more robust to changes. The design of such systems would facilitate genomic integration or low-copy plasmids, low levels of sensor proteins, and enzymatic reporters that allow for low levels of output expression. The further development of independent "resource generators"[227] that are less susceptible to physiological changes such as sigma factor concentrations will also aid in making programs more robust.

The end goal of this work was to develop a genetic controller that could detect and resolve acetate accumulation in *E. coli*. In the course of this work, several design decisions were taken that deserve revisiting. At the highest level, we chose to work "bottom-up", first developing the components, then integrating the components into the program, and finally testing the completed program in the host organism. This strategy has advantages and disadvantages. The principle advantage is that by characterizing components well, one can more easily use those components in different contexts. This can also yield knowledge about the performance of the component that will guide its integration into a system. One

disadvantage, however, is that much time is spent testing these components in contexts that may be irrelevant for their intended use. By first establishing the exact conditions under which the final system would operate (i.e. a "top-down approach), one could better guide the characterization efforts of components.

Complementing our characterization efforts with a "top-down" approach would have had other advantages. When solving a complex problem, it is valuable to first establish a set of experimental conditions that reproduce the problem and enable robust and rapid testing of solutions. In our system, for example, the measurement of exactly how much acetate is produced across a matrix of glucose release rates and dissolved oxygen concentrations would have guided the tuning of the acetate sensor. This also enables determination of system specifications and scientific testing of any assumptions regarding the mechanism of the problem. Such an experimental platform would have enabled us to first test potential actuators and determine their necessary expression ranges. In turn, this would have established useful target specifications for the completed program. For example, establishing target specifications might have shown that a 4-fold induction range was sufficient for our sensors. Therefore, tuning the sensors to >10-fold induction would have been unnecessary. By first determining target specifications, efforts to build a system can therefore be more efficiently focused.

Exclusive use of orthogonal systems would also have averted some problems. Orthogonal systems, by definition, are specific only to their cognate promoters, substrates, and binding partners and generally do not interact with native regulation. They therefore do not depend on host resources (e.g. sigma factors) and are more robust to changes in host physiology. The FNR oxygen sensor was not developed further in part because it made use of the FNR global regulator in *E. coli*, the expression and free concentration of which changes over growth phases and ambient conditions. Although this is also true of the Cra global regulator used to sense glycolytic flux, no other glycolytic flux sensors were available. Unfortunately, we were often limited by the availability of functional, well-understood sensor modules. Even when well-characterized, orthogonal modules are available, there is no guarantee that they will be functional following their transfer to a new organism. For example, we attempted using the Rex redox sensor from gram(+) bacteria in *E. coli*. After extensive testing, this protein's

response was never inducible, despite evidence that it was expressed and could functionally repress transcription from target promoters in *E. coli*[228] and was functional in mammalian cells[229]. When well-characterized, orthogonal modules are unavailable, it takes considerable effort to discover and develop them.

Complex genetic programs, though conceptually powerful, are currently difficult to implement. The DNA construction alone for the work presented here occupied the majority of the physical work and time and resulted in the use of over 1100 oligonucleotides to create 504 novel plasmids and 636 strains of *E. coli*. The troubleshooting and tuning required before the prototype in Chapter 3 was working took five months from initial testing to the final positive result. As such, the development of complex genetic programs is at the moment still a cottage industry, relying on highly trained and skilled artisans to piece together ad hoc designs. As systematic, high-throughput DNA synthesis and construction and computational tool come online, biological engineers will spend more time designing genetic systems and less time constructing them. This shift will lead to greater need for design tools and well-characterized parts. Thus, the current efforts at producing large collections of well-characterized parts and computational tools by which to assemble them will increasingly facilitate the design and assembly of complex genetic programs. As complex genetic programs become a more tractable solution to complex problems, the value proposition of synthetic biology will become more fully realized.

A. Supporting Information for Chapter 2

A.1 Cytometry fluorescence distributions

The responses of the *E. coli* and *S. cerevisiae* sensors to MeI were assessed by flow cytometry. Figure A.1 shows the fluorescence distributions of the *E. coli* strains carrying the sensor plasmid pFM45 in response to methyl iodide. The *E. coli* MG1655Δada strain has the *ada* gene knocked out and therefore shows no response to methyl iodide. The wild-type *E. coli* MG1655 populations induce with a bimodal response near the switch point, a behavior that is characteristic of systems containing a genetic positive feedback loop. Interestingly, this bimodal character is lost when additional Ada is expressed from a plasmid (pFM141) at all levels of induction. The behavior of the pFM45 sensor in the strain lacking pFM141 is consistent across all concentrations of arabinose. The presence of pFM141 in the MG1655*Δada* strain without arabinose induction is sufficient to rescue activity of the sensor, indicating leakage from the P_{BAD} promoter. Additional expression of Ada from pFM141 via the induction of the arabinose-inducible P_{BAD} promoter lowers the detection threshold of the sensor. High levels of Ada expression raise the basal leakage of the output promoter, which lowers the dynamic range of the sensor.

S. cerevisiae sensors showed a much lower dynamic range and less cooperativity than the *E. coli* response (Figure A.2). The response to MeI was dependent on the presence and number of Ada operators in the Cyc1 promoter driving the EGFP reporter. The yeast sensors also showed a much higher basal activity than the *E. coli* sensors.

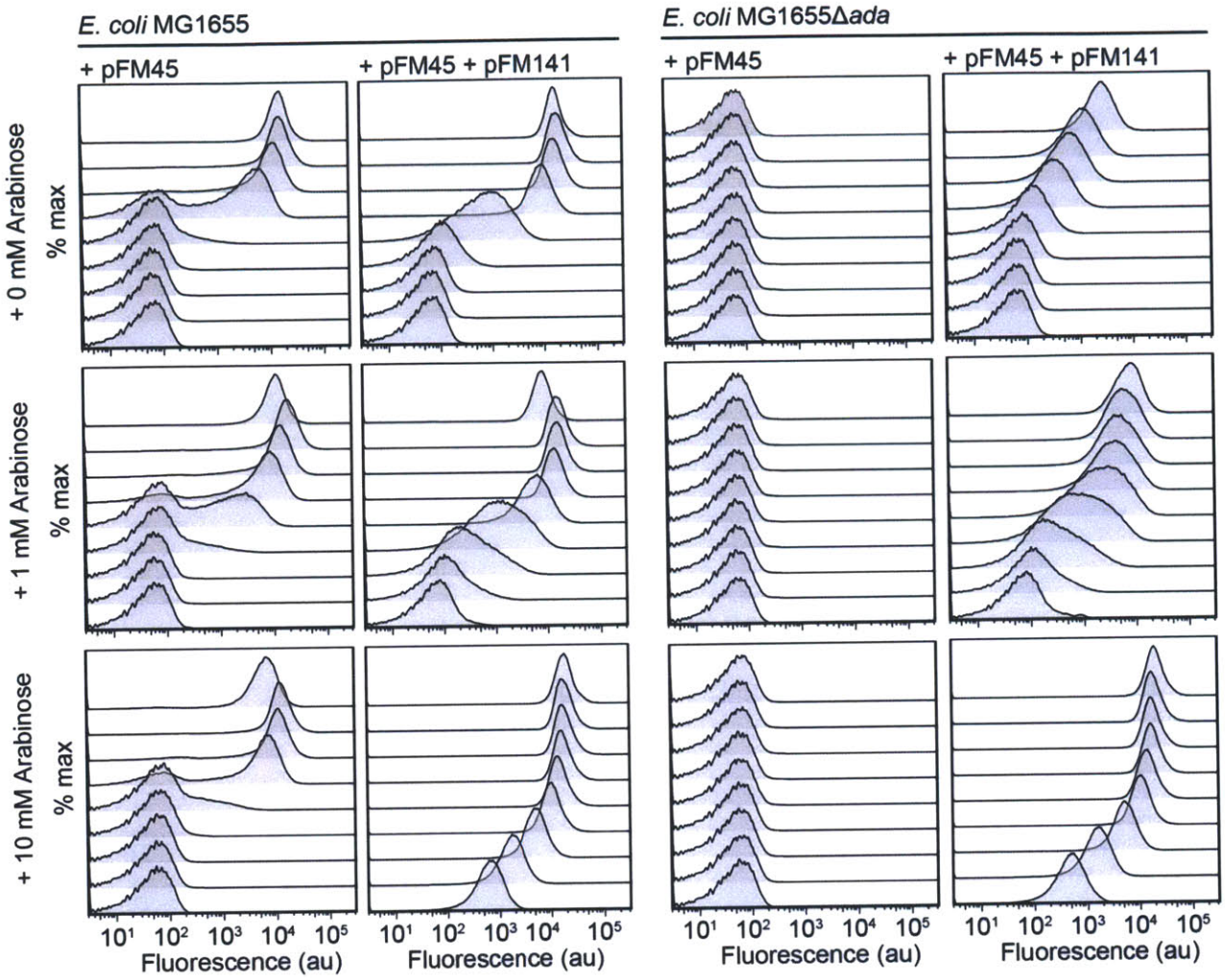


Figure A.1: **Cytometry distributions of the *E. coli* methylation sensor strains in response to MeI.** Shown are the cytometry data for transfer functions of *E. coli* strains MG1655 and MG1655*Δada* carrying the sensor plasmid pFM45 exposed to MeI. Each strain carrying pFM45 is also shown carrying the plasmid pFM141, which expresses the Ada protein from an arabinose-inducible P_{BAD} promoter. Arabinose was added to the cultures represented in the top (0 mM), middle (1 mM), and bottom (10 mM) rows of squares containing cytometry histograms, respectively. The amount of MeI added to each culture, from bottom-most histogram in each square to the top-most, is as follows: 0, 6×10^{-3}, 1.6×10^{-2}, 3.9×10^{-2}, 9.8×10^{-2}, 2.4×10^{-1}, 6.1×10^{-1}, 1.5, and 9.5 mM. This data corresponds to the data in Figure 3C and 3D.

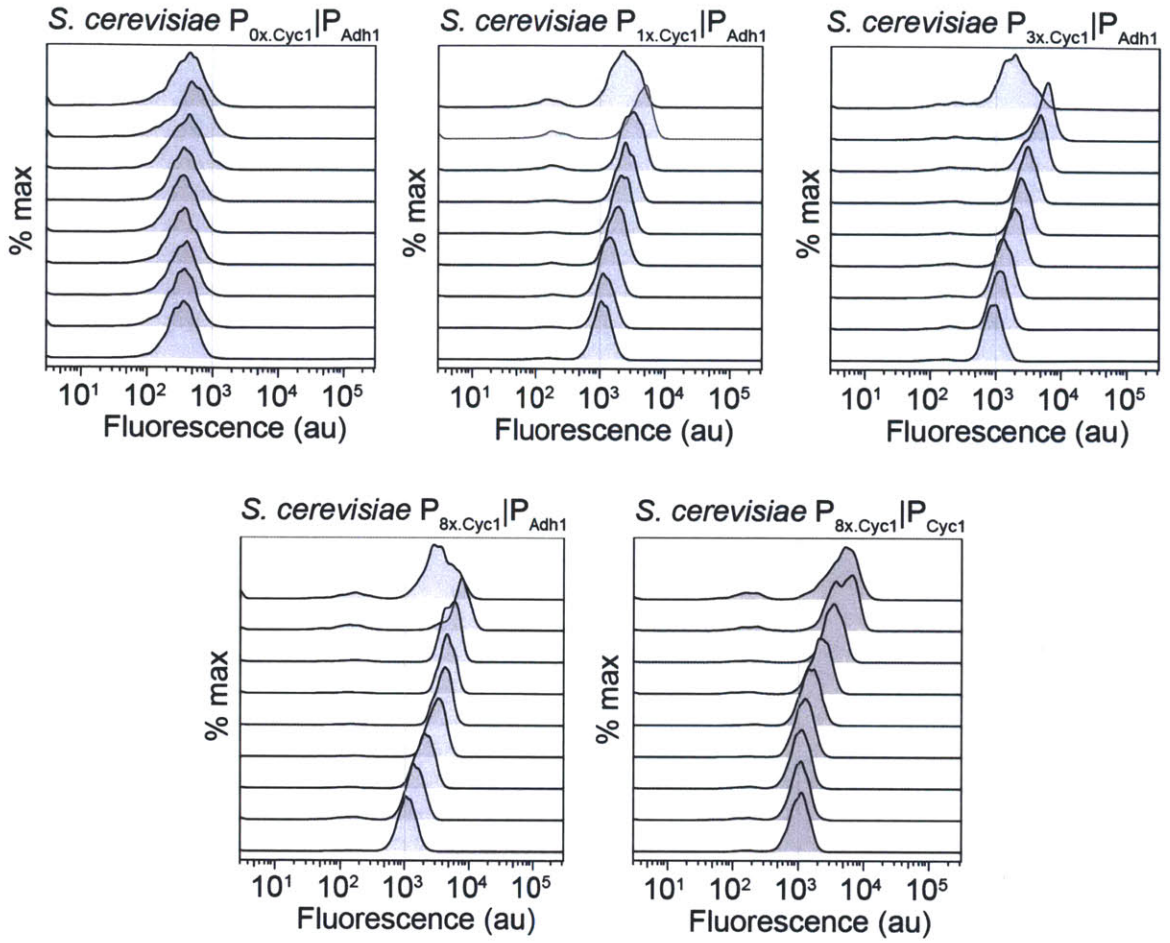


Figure A.2: Cytometry distributions of the *S. cerevisiae* methylation sensor in response to MeI. Shown are the cytometry data for transfer functions of *S. cerevisiae* sensor strains $P_{0x.Cyc1}|P_{Adh1}$, $P_{1x.Cyc1}|P_{Adh1}$, $P_{3x.Cyc1}|P_{Adh1}$, $P_{8x.Cyc1}|P_{Adh1}$, and $P_{8x.Cyc1}|P_{Cyc1}$ in response to MeI. The amount of MeI added to each culture, from bottom-most histogram in each square to the top-most, is as follows: 0, 2.8×10^{-2}, 6.4×10^{-2}, 1.5×10^{-1}, 3.4×10^{-1}, 7.8×10^{-1}, 1.8, 4.1, and 9.5 mM. This data corresponds to the data in Figure 4C and 4D, which reports the average of the geometric means for three different fluorescence distributions.

A.2 Toxicity of alkylating agents on *E. coli* and *S. cerevisiae*

Both sensors responded to methyl iodide (MeI), methyl methane sulfonate (MMS), dimethyl sulfate (DMS), and 1-methyl-3-nitro-1-nitrosoguanidine (MNNG). The toxic effects of these agents were evident in the cytometry distributions (Figures A.1, A.2, and A.3). At toxic concentrations, the population distribution widened considerably and lost fluorescence.

The toxicity of MeI, MMS, DMS, and MNNG on *E. coli* and *S. cerevisiae* strains containing methylation sensors was assessed and the LD_{50} of each alkylating agent was determined (Figure 9). *E. coli* is more robust to growth defects than *S. cerevisiae* at the same concentrations of

alkylating agent. No difference in toxicity was observed between wild-type MG1655 and MG1655 with the Δ*ada* mutation. Overproduction of the Ada protein in *E. coli* did not significantly reduce the toxicity of any alkylating agents. Higher levels of N-Ada-Gal4 production the yeast sensor strain also did not reduce the toxicity in that strain.

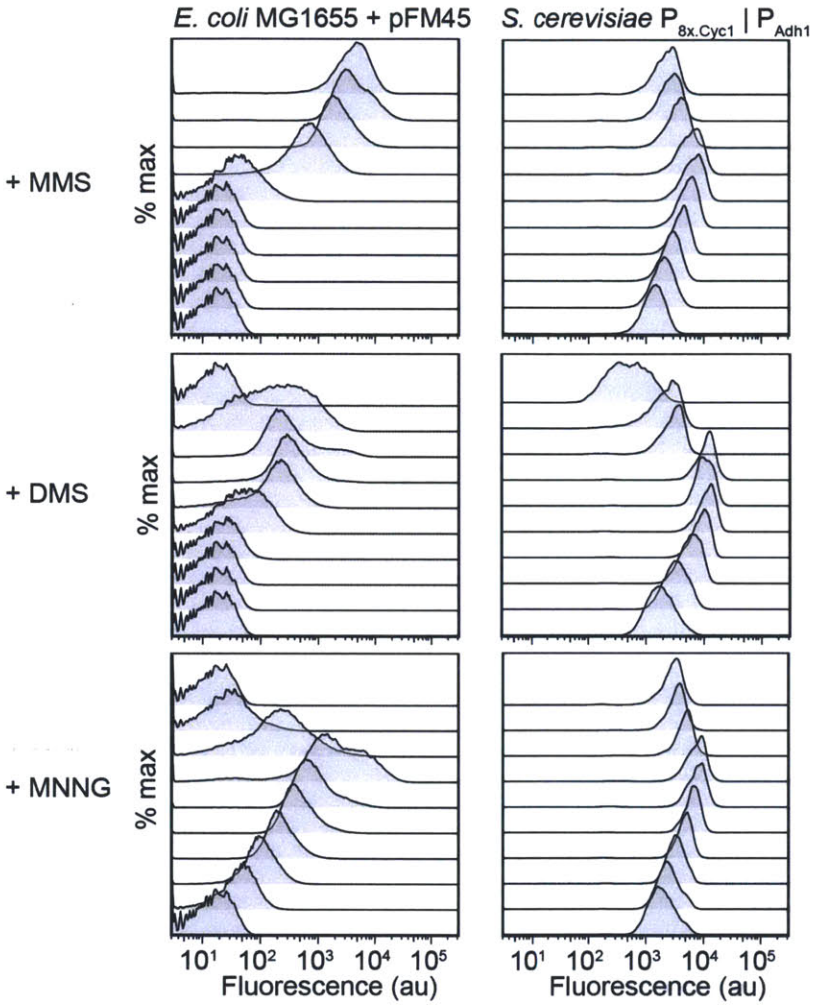


Figure A.3: Cytometry distributions of *E. coli* and *S. cerevisiae* methylation sensors in response to MMS, DMS, and MNNG. *E. coli* MG1655 carrying plasmid pFM45 and *S. cerevisiae* strain $P_{8x.Cyc1}|P_{Adh1}$ were exposed to MMS, DMS, and MNNG as described in the methods. The amount of MMS and DMS added to each culture, from bottom-most histogram in each square to the top-most, is as follows: 0, 1.2×10^{-2}, 2.7×10^{-2}, 6.4×10^{-1}, 1.5×10^{-1}, 3.4×10^{-1}, 7.8×10^{-1}, 1.8, 4.1, and 9.5 mM. The amount of MNNG added was as follows: 0, 1.6×10^{-4}, 2.3×10^{-3}, 5.2×10^{-3}, 1.2×10^{-2}, 2.8×10^{-2}, 6.4×10^{-2}, 1.5×10^{-1}, 3.4×10^{-1}, and 7.8×10^{-1} mM. This data corresponds to Figure 5.

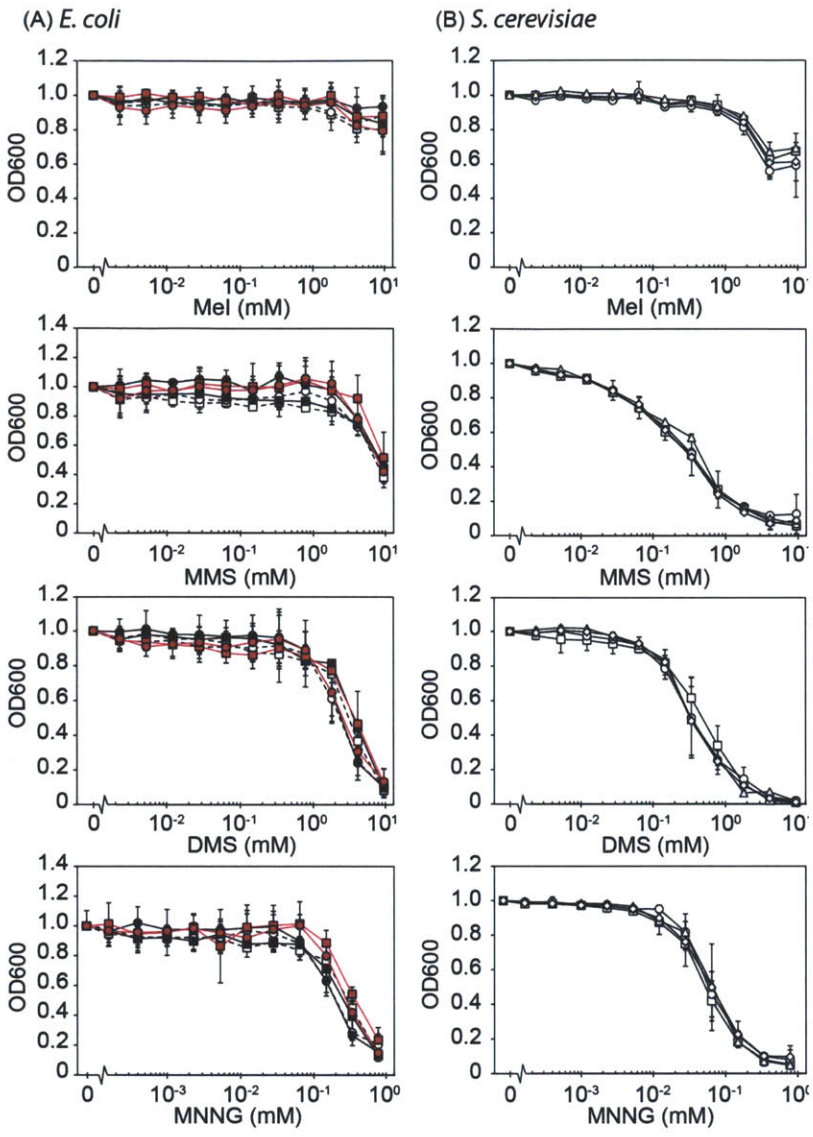


Figure A.4: Toxicity of alkylating agents on *E. coli* and *S. cerevisiae* containing methylation sensors. The strains were exposed to MeI, MMS, DMS, and MNNG. The OD_{600} of *E. coli* was measured 3 hours after exposure. **(A)** *E. coli* strains include: MG1655 (white square, dashed lines), MG1655*Δada* (white circle, dashed lines), MG1655 containing pFM45 and pFM141 (black squares, solid black lines), MG1655*Δada* containing pFM45 and pFM141 (black circles, solid black lines), MG1655 containing pFM45 and pFM141 and induced with 10 mM arabinose (red squares, red lines), and MG1655*Δada* containing pFM45 and pFM141 and induced with 10 mM arabinose (red squares, red lines). **(B)** *S. cerevisiae* strains measured include: SO992 (no Ada sensor, squares), $P_{0x.Cyc1}$ | P_{Adh1} (circles), $P_{8x.Cyc1}$ | P_{Adh1} (diamonds), $P_{8x.Cyc1}$ | P_{Cyc1} (triangles). The OD_{600} of the *S. cerevisiae* cells were measured 12 hours after exposure. For both *E. coli* and yeast cultures, all OD_{600} measurements were normalized to the highest measured value of that day for better comparison between days. Each data point is averaged from three measurements performed on different days. Error bars are one standard deviation from the mean.

A.3 GC-MS Standard Curve

We generated a standard curve to calculate the MeI produced by yeast cultures expressing methyl halide transferases (MHTs; Figure A.5). To measure this curve, we added a known amount of MeI into a volume of media equivalent to the volume in which sample cultures were grown. Following addition of MeI, the tubes were immediately stoppered. To allow the sample to adequately dissolve and equilibrate between liquid and gas phases in conditions comparable to those of the yeast culture, the standard curve samples were shaken

for 30 minutes at 30°C in the same incubator as the MHT yeast cultures. To sample the MeI in each tube, 100 µl of air from the headspace of each tube was injected into the GC-MS. Because some MeI degradation was observed over time, all samples were injected 30 seconds apart in a single long run. Each sample's MeI peaks, clearly differentiable, were integrated by the software. The resulting counts were plotted against the respective known amounts of MeI to generate the standard curve. The standard curve was re-run for each assay on each day and varied widely depending on machine settings. The slope of the standard curve, however, was consistent between days.

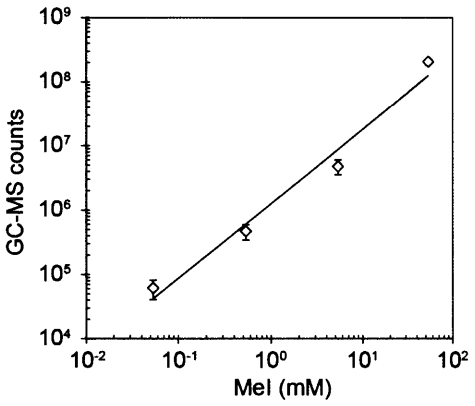


Figure A.5: Standard curve for GC-MS measurements. Known amounts of MeI were added to sample tubes, equilibrated, and measured with a GC-MS. The measured GC-MS counts of MeI are plotted against the amount of MeI added to each respective tube. A power law fits the data (R^2 = 0.98) and is used to calculate MHT production. The standard curve shown corresponds to the one used to calculate MeI production from the MHT yeast cultures shown in Figure 3A in the main text.

A.4 Saturation model for MeI activation of the *S. cerevisiae* sensor

A simple model was derived for the activation of the sensors. In this model, the promoter is activated by methylated Ada (Ada*) and responds instantaneously to a change in MeI concentration. The probability that RNA polymerase binds to the reporter promoter is given by,

$$f(s) = \frac{c_0 + K_d[Ada^*]^n}{1 + c_0 + K_d[Ada^*]^n} \quad , \qquad \text{(S1)}$$

where K_d is the binding constant for activated Ada to its operator, n is the empirically-derived cooperativity, and c_0 is the basal level of RNAP binding to the promoter causing leakage. The rate of Ada activation is

$$\frac{d[Ada^*]}{dt} = k_{met}[Ada][MeI] - k_{deg}[Ada^*] = 0 \quad \text{(S2)}$$

where k_{met} is the methylation rate constant, which is irreversible, and k_{deg} is the degradation rate constant. At steady-state,

$$[Ada^*] = \frac{k_{met}[Ada][MeI]}{k_{deg}} . \quad \text{(S3)}$$

Substituting Equation S3 into S1 produces

$$f(s) = \frac{c_0 + K_d\left(\frac{k_{met}[Ada]}{k_{deg}}\right)^n [MeI]^n}{1 + c_0 + K_d\left(\frac{k_{met}[Ada]}{k_{deg}}\right)^n [MeI]^n} = \frac{c_0 + K[MeI]^n}{1 + c_0 + K[MeI]^n}, \quad \text{(S4)}$$

where c_0, K and n are treated as fit parameters. This equation was used to fit the measured response functions reported in the main text. The Hill coefficients reported in the Tables in the main text were fit using this equation. The regression line in Figure 2.4A was also fit using this equation and the data in that chart (resulting in c_0 = 0.15, K = 0.056, and n = 1.8).

B. Supporting Information for Chapter 4

B.1 Circuit performance and impact on growth in shake flask experiments

Figures 5.2 and 5.33 show the performance of the AND/NOR gates across different media and strains. The corresponding impact on cell growth of these different conditions and gate activation is shown below. Figure B.1 shows how growth is impacted by media for *E. coli* DH10B cells. Both gates show higher final cell densities with increased amount of supplement in the media, and the addition of inducers reduced growth for both gates. Figure B.2 shows that this effect is present across different strains and media, though DS68637† grows to higher densities than DH10B.

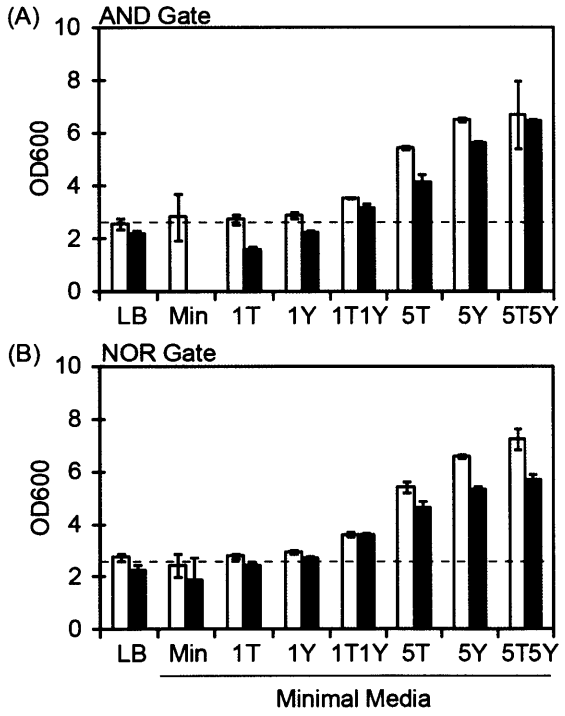
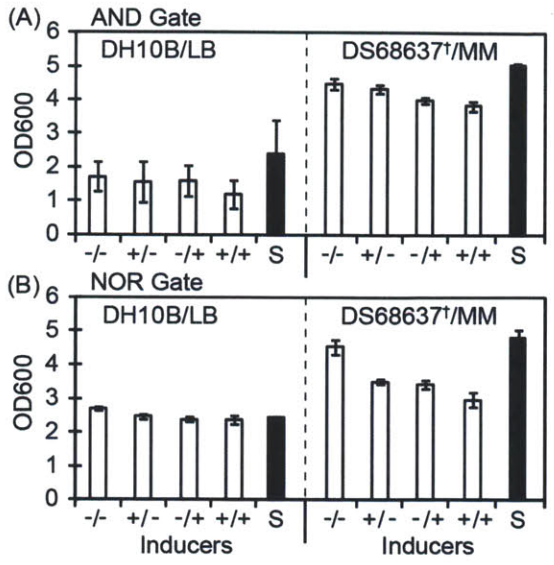


Figure B.1: Impact of the circuit and media on growth. The data corresponds to the experiments in Figure 2 and the media compositions are described in that figure. The final OD_{600} of the culture is shown for **(A)** the AND gate and **(B)** the NOR gate in *E. coli* DH10B. The data is shown for uninduced (white bars) and fully induced (black bars) cultures. All cultures were measured after 9 hours, except the cultures that were grown on unsupplemented minimal medium (Min), which were measured after 24 hours. Induced AND gates did not grow on unsupplemented minimal media (Min). The dashed line is the mean OD_{600} of the reference plasmid (pFM46) grown in DH10B in LB after 9 hours.

Figure B.2: Growth of different *E. coli* strains carrying the gates. The data corresponds to Figure 3. **(A)** The OD_{600} of the AND gate is shown for four combinations of inputs: -/- (no inducers), +/- (1.3 mM arabinose), -/+ (0.63 mM salicylate), and +/+ (both inducers). **(B)** The OD_{600} of the NOR gate cultures is shown for four combinations of inputs: -/- (no inducers), +/- (1.3 mM arabinose), -/+ (100 ng/ml aTc), and +/+ (both inducers). The OD_{600} of the reference plasmid pFM46 (S) is shown. In DH10B, the AND gates were measured after 9 hours and NOR gates were measured after 24 hours. In DS68637† on Minimal medium (MM), the AND and the NOR gate were both measured after 24 hours.



The data shown in Figure 4.2 are the geometric means of the fluorescence of a culture, as measured by flow cytometry. The complete distributions are provided here to show the variability in the populations. In Figure B.3, the population variability of the two gates is shown for the uninduced and induced conditions. The ON state of the AND gate is very media dependent, whereas both the ON and OFF states of the NOR gate are robust across media.

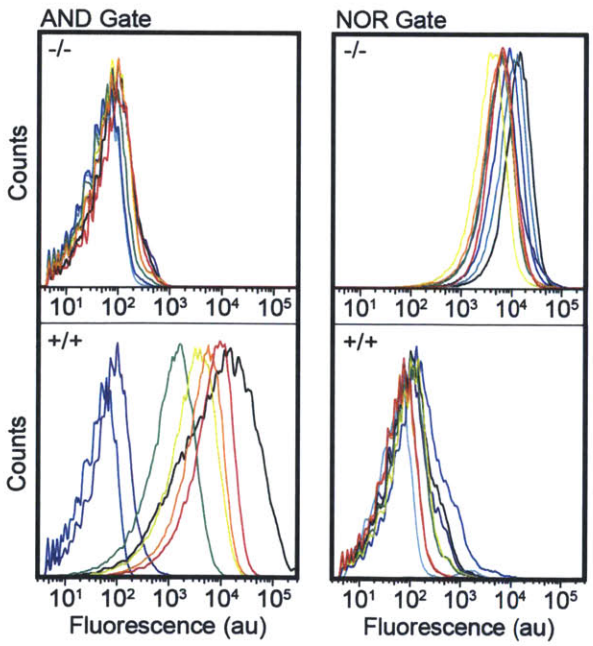


Figure B.3: Cytometry distributions for AND and NOR gates in *E. coli* DH10B. The distributions are representative of the data shown in Figure 2. Each culture is color coded for the media: LB (black), minimal medium (light blue), 1T (dark blue), 1Y (purple), 1T1Y (green), 5T (yellow), 5Y (orange), 5T5Y (red), where the number before the letter refers to the amount in g/L, T refers to tryptone, and Y refers to yeast extract. The uninduced (-/-) and induced (+/+) data are shown in the top and bottom graphs, respectively.

Ribosome Binding Sites (RBSs) of different strength were chosen to connect the arabinose-inducible promoter to the AND gate. Differences in performance of these AND gate variants was shown for shake flask experiments (Figure 4.5) and 10 L fermentors (Figures 4.7). Here, the raw cytometry distributions are provided for the shake flask experiments. The distributions show that the weaker RBS variants (RBS_C and RBS_D) often exhibit skewed or multimodal distributions. RBS_B is the variant that is tested in the shake flask cultures and the microreactor (Figures 3, 4, 6, S1, S2, S3).

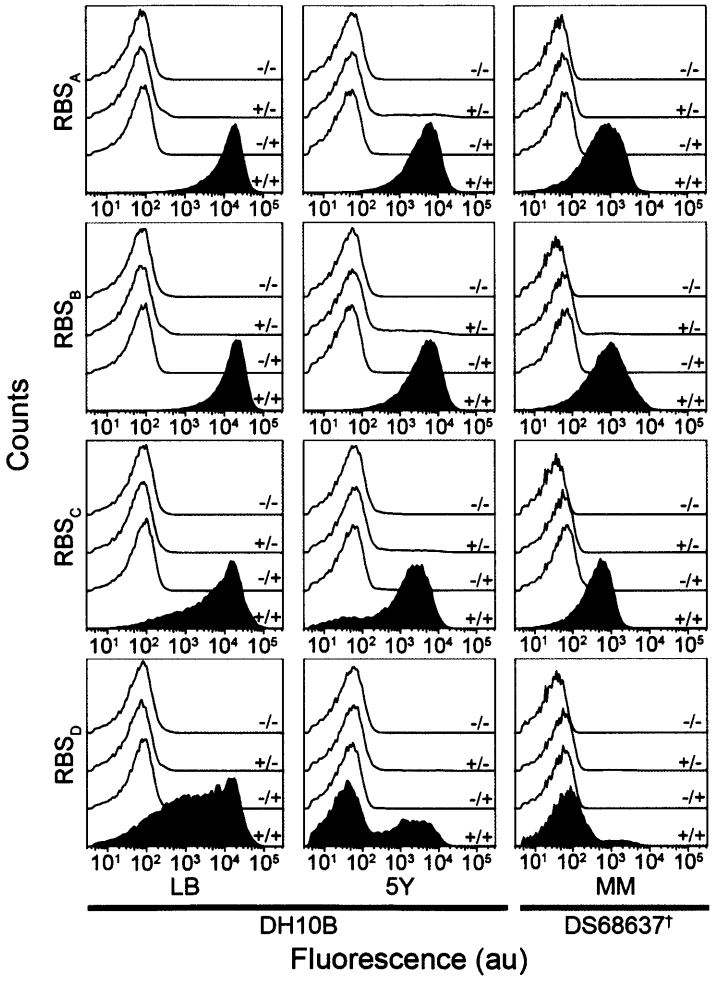


Figure B.4: Cytometry distributions of different RBS variants. The distributions are representative of the data used to calculate the averages in Figure 4. The gates were grown with either no inducer (-/-), 1.3 mM arabinose (+/-), 0.63 mM salicylate (-/+), or both inducers (+/+). DH10B were grown in LB and minimal medium with 5 g/L yeast extract (5YE). DS68637† were grown in minimal medium (MM).

B.2 RBS Performance is consistent across strain and media

RBSs are commonly used to tune genetic circuits and were used here to tune AND gate performance. To test if RBS strength changes with different media and strains, we grew DH10B and DS68637[†] containing plasmids that constitutively expressed RFP from different strength RBSs (pFM169-174). Each RBS performed similarly across different strain and media contexts.

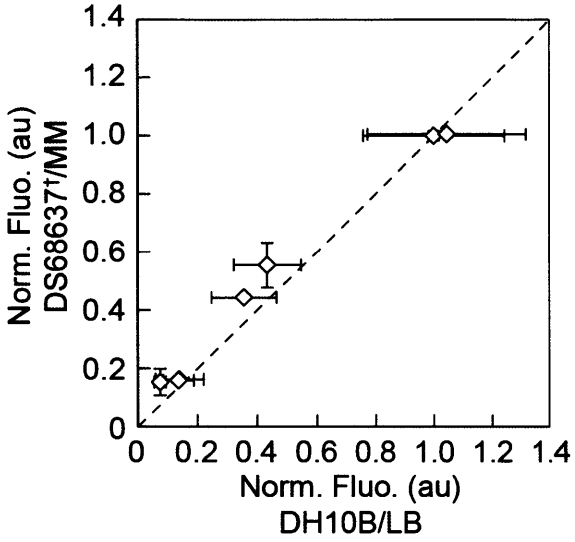


Figure B.5: RBS performance in different strains/media. A collection of engineered RBSs was measured in DH10B in LB and in DS68637[†] in minimal medium (MM) in shake flasks. Fluorescence of each construct was normalized to the strongest construct in the series. RBS performance was consistent ($R^2 = 0.964$) across both media and strains.

B.3 Detailed microreactor data

We grew an AND gate and the reference plasmid pFM46 in DS68637 in a BioLector microreactor to test gate performance.[7] Relative culture density was measured by light scatter and is reported in arbitrary units (au). Fluorescence was measured by spectrophotometry with an excitation at 486 nm and emission at 510 nm. pH was measured via an optode that excited at 486 nm and emitted at 530 nm. Dissolved oxygen was measured via an optode that excited at 505 nm and emitted at 590 nm. Dissolved oxygen was maintained at maximal saturation to maintain growth through carbon limitation. The culture was grown under 80% humidity control at 30°C at 1000 RPM elliptical shaking.

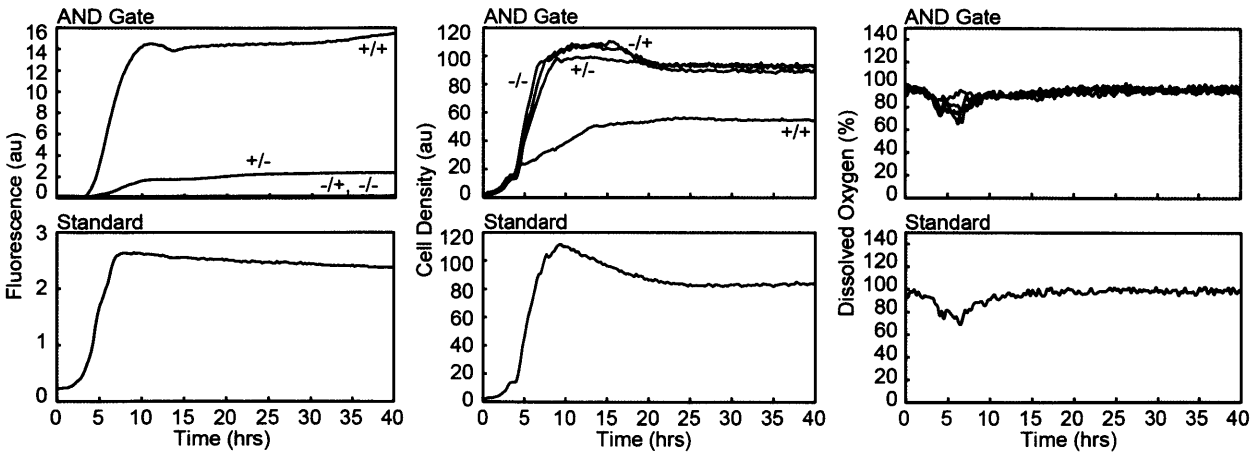


Figure B.6: Raw fluorescence, growth, and oxygen content of the gates and reference plasmid in the microreactor. The RBS_B AND gate and the reference plasmid (pFM46, "Standard") were grown in 1 ml batch of rich 2xYT media in the presence of either no inducers (-/-), single inducers (-/+), arabinose only (+/-), salicylate only (-/+), or both inducers (+/+).

B.4 Performance of an AND gate during fermentation

To determine the effect of yeast extract on AND gate function in DS68637, we grew DS68637 carrying the RBS_B AND gate in a 10 L fermentor with 0 or 20 g/L of yeast extract in the glucose feed. In DH10B, 100 g/L were required in the feed for the AND gate to function (Figure 4.7). In DS68637, the AND gate functioned without any yeast extract added to the feed, and performed identically with and without yeast extract in the feed (Figure B.7A).

To compare the performance of the AND gate directly to a reference plasmid, we grew DS68637 carrying the AND gate or the reference plasmid in a 10 L bioreactor. We induced the AND gate in reverse order than in previous fermentations to examine the arabinose-only state of the gate. The AND gate functioned fully upon only arabinose induction, indicating a failure to carry out the correct computation. The reference plasmid maintained a fairly constant level of fluorescence throughout growth (Figure B.7B).

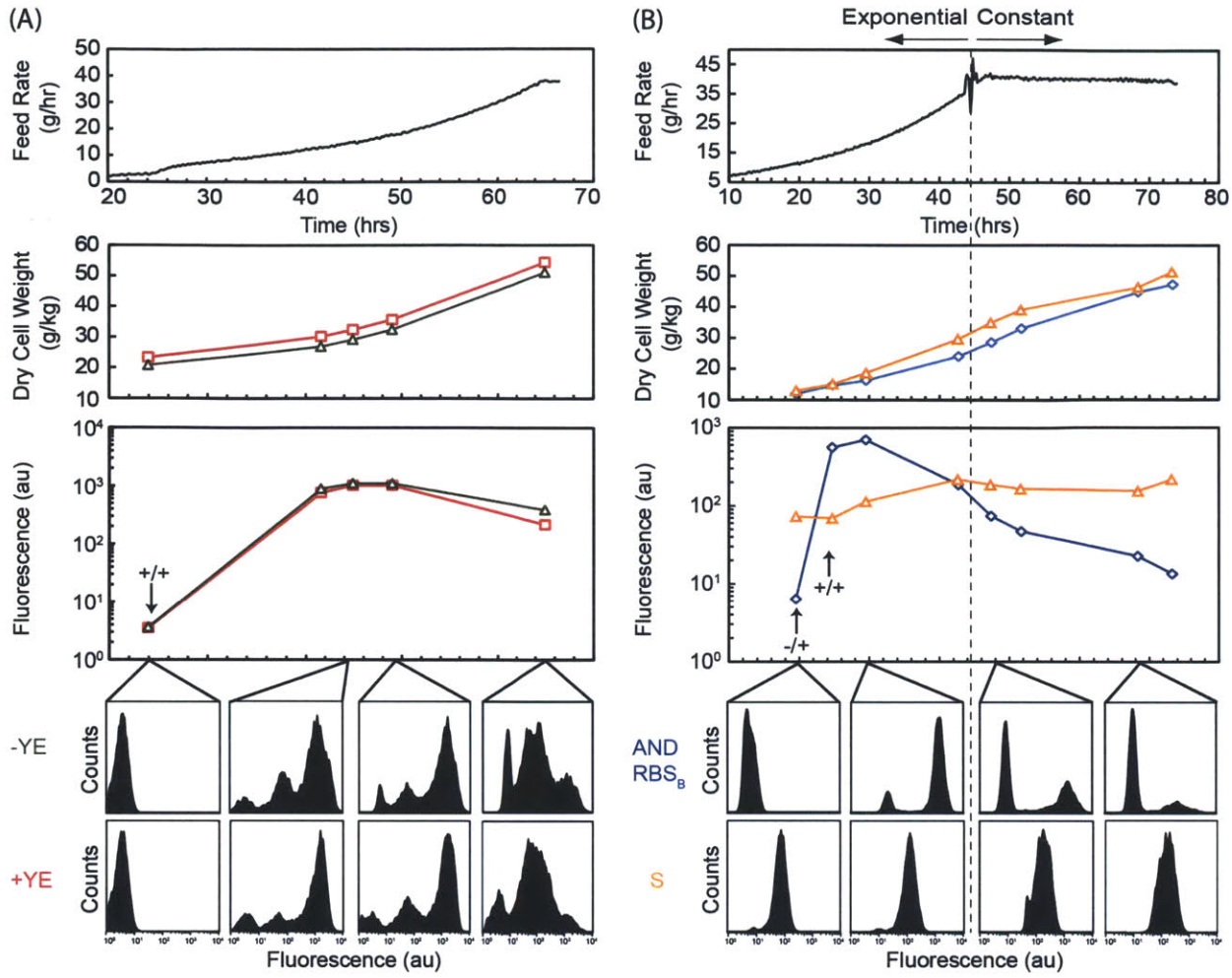


Figure B.7: Additional states of the AND gate in a 10 L bioreactor.
(A) Data from Runs # 4 and 5 (Table S1), the growth of the RBS_c AND gate in DS68637 with and without 20 g/L yeast extract in the feed (+YE and -YE, respectively). The gate performs identically in both conditions. **(B)** Data from Runs # 9 and 10 (Table S1), the growth of the RBS_B AND gate and the reference plasmid (pFM46). The data for the reference plasmid (S) is identical to that shown in Fig. 7 of the main text. The AND gate is induced first with arabinose at 22 hours (-/+) and then with salicylate 3 hours later (+/+). The AND gate was fully induced after addition of only arabinose, representing a failure mode of the gate. The reference plasmid shows consistent fluorescence throughout growth, even after switch to constant feeding. The cytometry distributions of the AND gate resemble those from previous fermentations and show a multimodal distribution following cessation of the exponential feed.

B.5 Plasmid retention in the 10 L bioreactor

In the 10 L bioreactor, the fluorescence from the AND gate declines late in fermentation (Figures 4.7, B.7). The loss of plasmids is a common problem in fermentation. In order to determine the contribution of plasmid loss to circuit failure, plasmid retention was quantified. To assess plasmid loss during fermentation, we plated samples from different time points of

Runs # 9 and 10 on selective plates (Figure B.8). At each time point, a sample of culture was taken, diluted, and plated on 2xYT agar containing either no antibiotic, single antibiotics, or combinations of antibiotics. Percent plasmid retained at each time point is calculated by dividing the number of colonies growing on a selective plate by the number of colonies growing on a nonselective plate at the same dilution.

The AND gate consists of three plasmids. The loss of any one of these three plasmids would result in the inability to perform the AND function. By 48 hours, almost all cells assayed had lost at least one essential plasmid (Figure B.8). The loss of these plasmids is most likely due to the loss of selective pressure to retain the plasmids following fast degradation of antibiotics at high cell densities.[5] The CamR plasmid (pAC_SalSer914), which expresses the supD tRNA, is particularly unstable. This suggests that there is selective pressure against producing supD. However, the reference plasmid pFM46 and the AmpR plasmid in the AND gate (pBR939b), both of which produce GFP reporter, were stably retained through the end of fermentation.

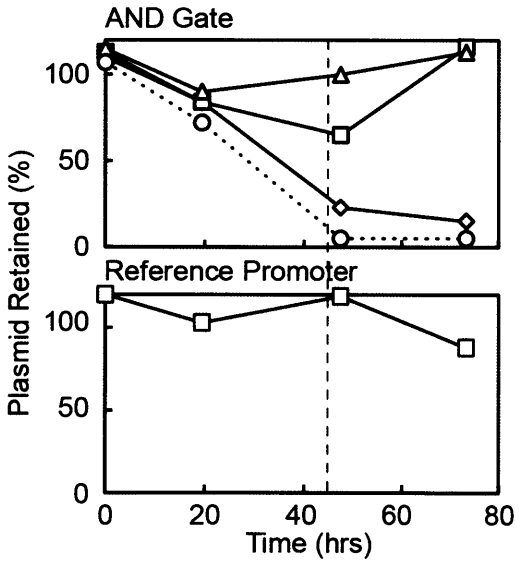


Figure B.8: Plasmid loss in the AND gate and reference plasmid strains in the 10 L bioreactor. Retention of the plasmids composing an AND gate and the reference plasmid pFM46 were measured in *E. coli* DS68637 at several time points throughout Runs # 9 and 10, respectively. The percent of cells retaining KanR (squares), CamR (diamonds), AmpR (triangles), and all plasmids (circles) is plotted. The dotted line marks the time at which the feed rate was switched from exponentially growing to constant.

B.6 Relative Expression Units (REUs) Conversions

In Figure 4.7, we use relative expression units (REUs) to compare expression of fluorescent reporter protein between conditions. REU's are simply the result of multiplying the raw fluorescence data produced by each instrument by a linear factor (Figure B.9). This linear

factor relates the amount of fluorescent protein expression of the experimental construct to that of a standard plasmid in the given condition. We use REU instead of the more widely-accepted RPU (Relative Promoter Units) used by Kelly and coworkers because differences in the plasmid backbones, fluorescent proteins, and regulation make calculation of RPU as defined by Kelly et al. impossible.[9,10,11] Instead, we use the relative amount of reporter protein produced by Kelly et al.'s standard plasmid to compare expression levels between contexts.

To determine REU conversion factors, we measured a strain carrying a reference plasmid pFM46 in parallel with AND gate strains in all contexts. In the first conversion step (Figure B.9B), we normalize the GFP expression of the AND gate reporter plasmid (pBR939b) by the GFP expression of pFM46. Because the raw arbitrary fluorescence unit reported by each instrument is different, the initial conversion factor changes with context and instrument. In shake flasks, this initial conversion factor (0.00012) is 1 divided by the geometric mean of the fluorescence of the pFM46 construct in DS68637† grown in minimal media, measured at 24 hours. In the microreactor, the initial conversion factor (0.41) is 1 divided by the mean arbitrary fluorescence of the pFM46 construct in DS68637 between 15 and 40 hours, the time period in which the fluorescence signal was stable. In the 10 L fermentor, the initial conversion factor (0.0067) is 1 divided by the mean fluorescence value of the pFM46 construct in DS68637 across the entire fermentation. In the second conversion (Figure B9C), we account for some differences between pBR939b and pFM46. Because effective comparison of expression can only be done within the framework of the same fluorescent protein, we took into account the contribution of the LVA (ssrA) tag on the GFP of pBR939b. The value of this conversion is reported as the average ratio of the GFP fluorescence of pFM46 to J23102.egfp.LVA (0.082) in DS68637 during exponential growth. Because the LVA tag dramatically reduces the amount of GFP in the cells, we had to use a much stronger RBS similar to the one in pBR939bto achieve measurable levels of GFP. Therefore, this conversion also accounts for the change in RBS from pFM46 to pBR939b. One should note, however, that this conversion does not account for differences in copy number (p15A vs. pMB1) or other differences between the plasmids. Finally, we compared the expression of pFM46 to that of Kelly et al. using the average J23102 RPU to J23101 RPU ratio reported in that work (0.86).[8] The final conversion factors are calculated as

follows: for shake flasks, $8.46E^{-6}$ REU/au= 0.00012 x 0.082 x 0.86, for the microreactor, $2.89E^{-2}$ REU/au= 0.41 x 0.082 x 0.86, and for the 10 L fermentor, $4.72E^{-2}$ REU/au= 0.0067 x 0.082 x 0.86. Dividing by these respective factors converts REU in Figure 4.8 back to raw arbitrary fluorescence units measured by each instrument.

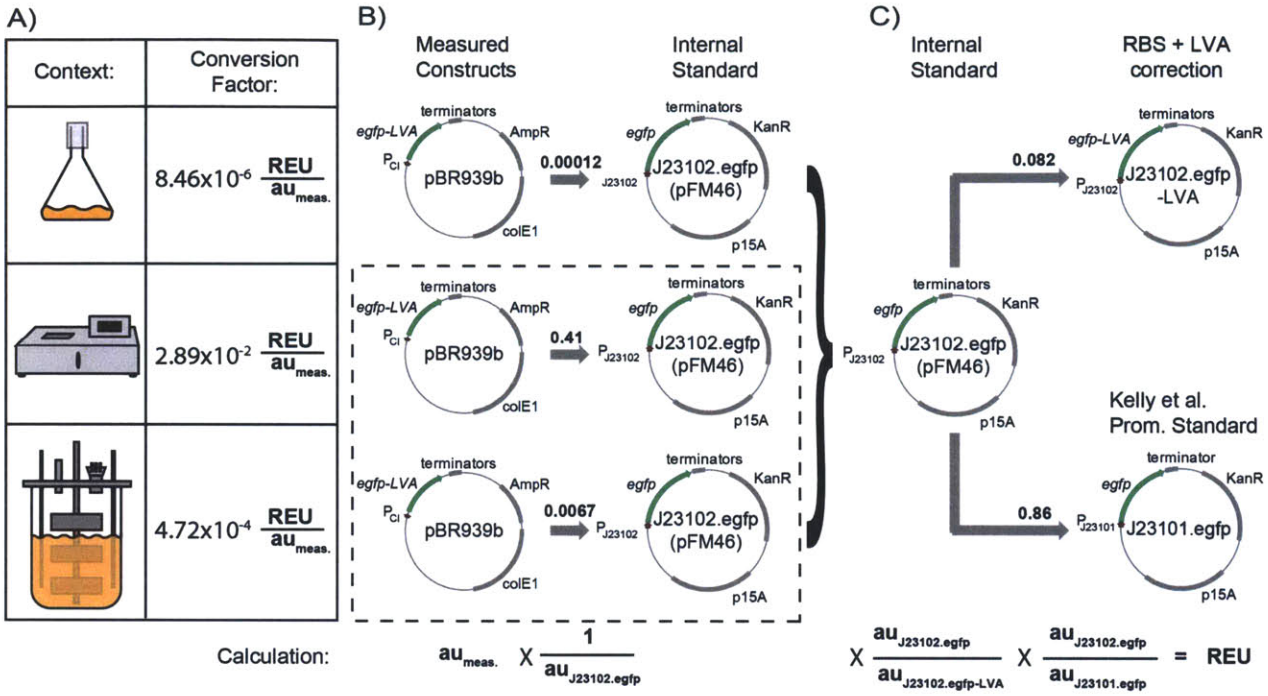


Figure B.9: Conversion of arbitrary units into relative expression units (REU). The schematic illustrates the conversion from raw arbitrary fluorescence units to REU. A) The table summarizes the final calculated conversion factors in each measured context. B) The AND gate output plasmid (pBR939b) is measured in parallel with the reference plasmid pFM46 (designated J23102.egfp to highlight component similarities with other plasmids). The experimental fluorescence measurements were normalized by the fluorescence of this reference plasmid. The boxed constructs were measured at DSM and unboxed plasmids were measured at UCSF. C) The reference plasmid pFM46 was measured in parallel with a similar plasmid carrying a stronger RBS and LVA-tagged GFP (J23102.egfp-LVA) to account for the effects of the LVA tag on the GFP expressed by pBR939b. The promoter J23102 of the reference plasmid pFM46, which is identical to the Kelly et al. standard plasmid except for the promoter, was reported to be 0.86 the strength of J23101. Due to the similarity of these plasmids (only a few base pairs change within the promoter), this ratio is expected to be independent of context. The calculations at the bottom reflect the corresponding plasmid ratios above.

B.7 Variations and Parameters of Individual 10 L Bioreactor Runs

All 10 L bioreactor runs were carried out as described in the Methods section in the main text, and any variation from this protocol is given in each series' section below. Differences in media composition and fermentation parameters between each run are summarized in Table B.1. Runs #1-3 refer to the experiments in Figure 4.6. Runs #4-5 refer to data in Figure 4.6A. Runs #6-8 refer to the AND gates in Figure 4.7. Runs #9 and 10 refer to the AND gate and reference plasmid data in Figure B.7B, respectively. Runs within the same series (#1,2,3, #4,5, #6,7,8, and #9,10) are grouped together spatially in Table B.1 and Figure B.10.

Runs #1-3: *AND gate in E. coli DH10B with different amounts of yeast extract.*

purpose of this series was to test the AND gate in *E. coli* DH10B in bioreactor conditions with different concentrations of yeast extract (100, 20, and 0 g/kg) in the feed. The addition of large amounts of yeast extract to the feed required customizing some of the parameters of the fermentation. To dissolve sufficient yeast extract in the feed, the glucose concentration in the feed had to be lowered (Table B.1). Therefore, higher feed rates had to be programmed into the control computer to maintain the same glucose-limited growth rate in the culture. Additionally, the dissolved oxygen was maintained at 100% longer throughout batch growth to aid in the growth of the culture. Additional ammonium sulfate and leucine were also added to the batch to support growth of the *E. coli* DH10B strain.

Addition of yeast extract to the feed changed both AND gate performance and growth characteristics of the cultures during fermentation. The AND gate functioned only in the culture containing 100 g/kg yeast extract in the feed (Figure 4.5). Also, oxygen consumption late in the run increased with increased amounts of yeast extract in the feed (Figure B.10). Although the feed rates of all fermentations increased identically, the feed with added yeast extract provided more nutrition per gram of feed. This is why the fermentations with more yeast extract in the feed showed higher rates of oxygen consumption. In contrast, the culture with no yeast extract in the feed displayed much slower growth in oxygen consumption followed by a decrease of oxygen consumption after 56 hours (Figure B.10). These trends are most likely due to the

availability of nutrients late in fermentation. When supplemental nutrients from the yeast extract are exhausted late in fermentation, the culture is forced to slow its metabolic rate. However, although oxygen consumption of the fermentation with 100 g/kg of yeast extract in the feed continues to increase, the fluorescence signal of the AND gate decreases by the end of the run (Figure 6B). Plasmid loss was not assessed in these fermentations.

Runs #4-5: *AND gate in E. coli DS68637 with different amounts of yeast extract*

The AND gate was observed to perform well in *E. coli* DS68637 without addition of yeast extract to the media in shake flasks (Figure B.6A). To test if this functional independence to media was consistent for fermentation conditions, we tested AND gate function in *E. coli* DS68637 during fermentation with different amounts of yeast extract in the media.

For this fermentation, smaller amounts of yeast extract (20 g/kg) were added to the feed, and therefore the glucose concentration did not need to be reduced. As a consequence, feed rate was also reduced compared to Runs #1-3. Since *E. coli* DS68637 is not auxotrophic, no additional supplements were added to either the batch or the feed.

Run #6-8: *AND gate RBS variants in E. coli DS68637*

To test if the choice of optimal RBS in an AND gate is consistent between shake flask and fermentation conditions, we grew three different RBS variants (RBS_B, RBS_C, RBS_D) of the AND gate in fermentation conditions. None of the fermentations in this series contained any yeast extract in the feed. The final feed profile was therefore different than in previous fermentations, although it was identical to the first two days of the feed profile of Runs #4-5. All three gates were first induced with salicylate and then with arabinose 4 hours later. More time points were sampled so as to assess AND gate function more fully throughout fermentation. GFP *visibly* accumulated in RBS_B and RBS_C AND gate variants only a few hours after full induction. After 45 hours, the feed rate was held constant and adjusted twice. At ~72 hours, the exponential feed was again increased to test whether GFP production could be recovered by feeding the cells.

Run #9-10: *AND gate and the reference plasmid in E. coli DS68637*

The final series of runs aimed to compare the performance of the RBS_B AND gate and a reference plasmid *E. coli* DS68637. The reference plasmid (pFM46) consists of a strong, constitutively active promoter (J23102) and RBS (B0032) driving expression of GFP from a p15A origin, Kan^R plasmid.[8] This standard plasmid was expected to maintain approximately the same level of expression throughout fermentation, with the only variability in observed GFP expression due to innate changes in expression and metabolism. Comparing the fermentations of the AND gates to the fermentation of this plasmid enabled us to normalize by these innate changes in expression and metabolism and make a more informative comparison of circuit performance. These fermentations were carried out identically to runs #6-8, except that induction was done first with arabinose, followed by salicylate (AND gate) induction 3 hours later. Inducing first with arabinose tested the alternate state of the AND gate and showed that the gate turned on with only arabinose inducer.

Table B.1: Detailed parameters of runs in the 10 L bioreactor

Run	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
Circuit	AND RBS_B			AND RBS_C		AND RBS_B	AND RBS_C	AND RBS_D	AND RBS_B	Ref. J23102
Strain	*E. coli* DH10B			*E. coli* DS68637		*E. coli* DS68637			*E. coli* DS68637	
Batch	4 g/kg glucose+10 g/kg YE+ 5 g/kg $(NH_4)_2SO_4$ +0.5 g/kg leucine			4 g/kg glucose+10 g/kg YE+ 2 g/kg $(NH_4)_2SO_4$		4 g/kg glucose +10 g/kg YE+ 2 g/kg $(NH_4)_2SO_4$			4 g/kg glucose +10 g/kg YE+ 2 g/kg $(NH_4)_2SO_4$	
Feed	100 g/kg YE +250 g/kg glucose +2.5 g/kg leucine	20 g/kg YE 250 g/kg glucose 2.5 g/kg leucine	0 g/kg YE/ +250 g/kg glucose +2.5 g/kg leucine	400 g/kg glucose +20 g/kg YE	400 g/kg glucose	400g/kg glucose			400g/kg glucose	
pH titrants	2-fold diluted 25% NH3 solution			10% NH3 solution		10% NH_3 solution, 4N H_2SO_4			10% NH_3 solution, 4N H_2SO_4	
Feed Rate	Exponential feed profile during 60 h programmed into computer. Based on μ = 0.05 h^{-1}, starts at 4 g/h, feedmax 80.3 g/h.			Exponential feed profile during 60 h programmed into computer in table. Based on μ = 0.05 h^{-1}, final rate 45 g/h, duration 60 h (starts at 2.24 g/h). From 60 h onwards 45 g/h.		Exponential feed profile during 60 h programmed into computer in table. Based on μ = 0.05 h^{-1}, final rate 45 g/h, duration 42 h (starts at 5.5 g/h). From 42 h onwards 45 g/h.			Exponential feed profile during 60 h programmed into computer in table. Based on μ = 0.05 h^{-1}, final rate 45 g/h, duration 42 h (starts at 5.5 g/h) From 42 h onwards 45 g/h.	
Time of induction	~ 48 hrs			24 hr	22 hrs (Salicylate) and 26 hrs (Arabinose)	22 hrs (Salicylate) and 26 hrs (Arabinose)			22 hrs (Arabinose) and 25 hrs (Salicylate)..	
Airflow	5 nL/min			2-4 hrs: headspace aeration at 2 nL/min. >4 hrs: submerged aeration 4 nL/min		0-6hrs: 2 nL/min >6hrs: 4 nL/min			0-6hrs: 2 nL/min >6hrs: 4 nL/min	

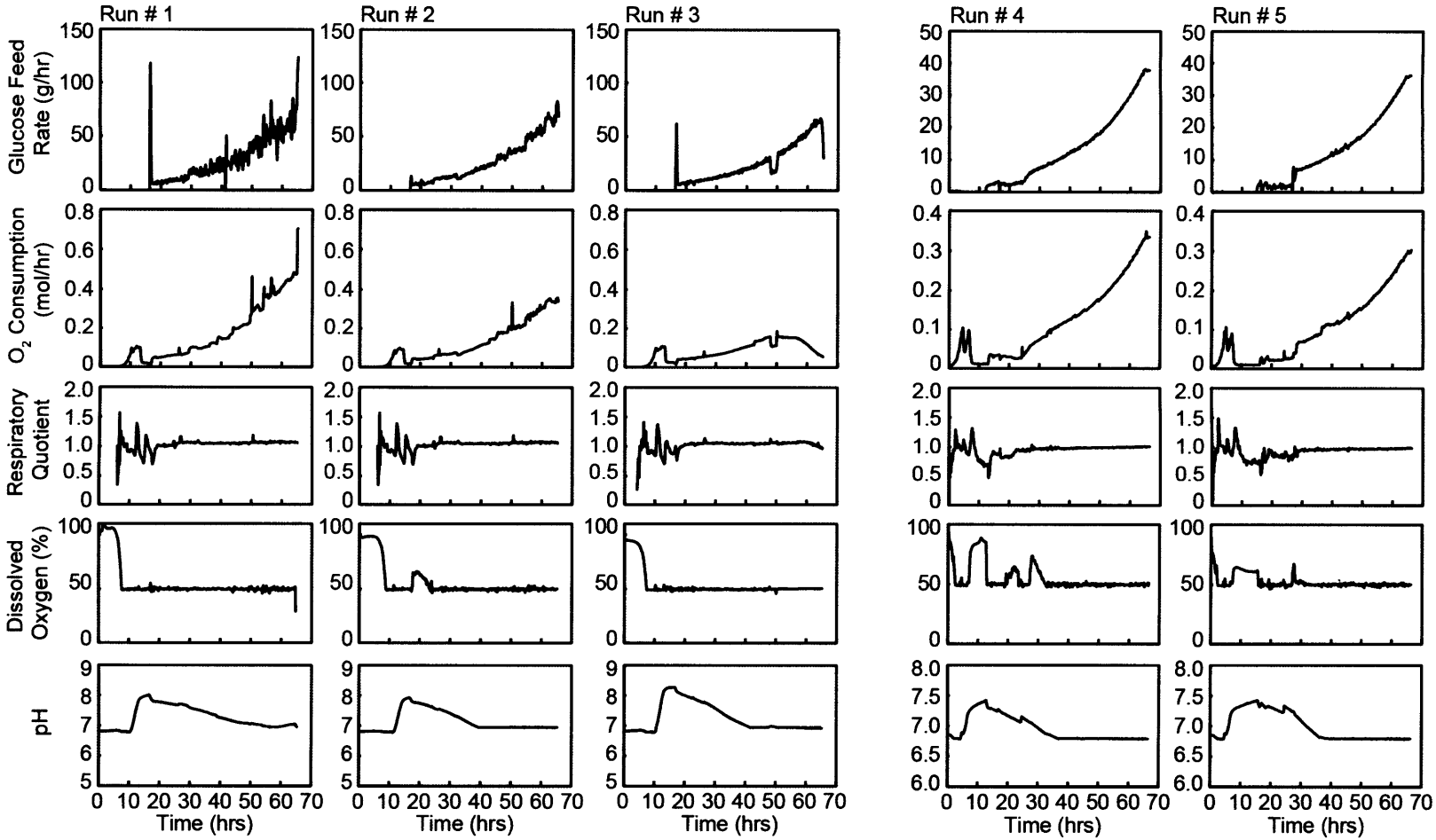


Figure B.10: Detailed data recorded from the 10 L bioreactor runs. Glucose feed is calculated and controlled by the central computer and monitored by feed weight. Variation in glucose feed in Run #1 was due to a faulty pump. Molar O_2 consumption and CO_2 production rates are monitored over time by mass spectrometry and are equal when the respiratory quotient (RQ) equals 1. In most fermentations, the RQ stabilizes at 1 following the start of the exponential glucose feed. The dissolved oxygen (DO) is monitored by an oxygen probe and controlled by impeller speed and airflow. DO is initially kept near 100% during batch phase and then drops to 50% throughout exponential feeding. Variation in DO throughout fermentation was due to occasional problems with the oxygen probe. The pH of each culture is monitored by an electrode in real time and confirmed off-line by an independent electrode following sampling.

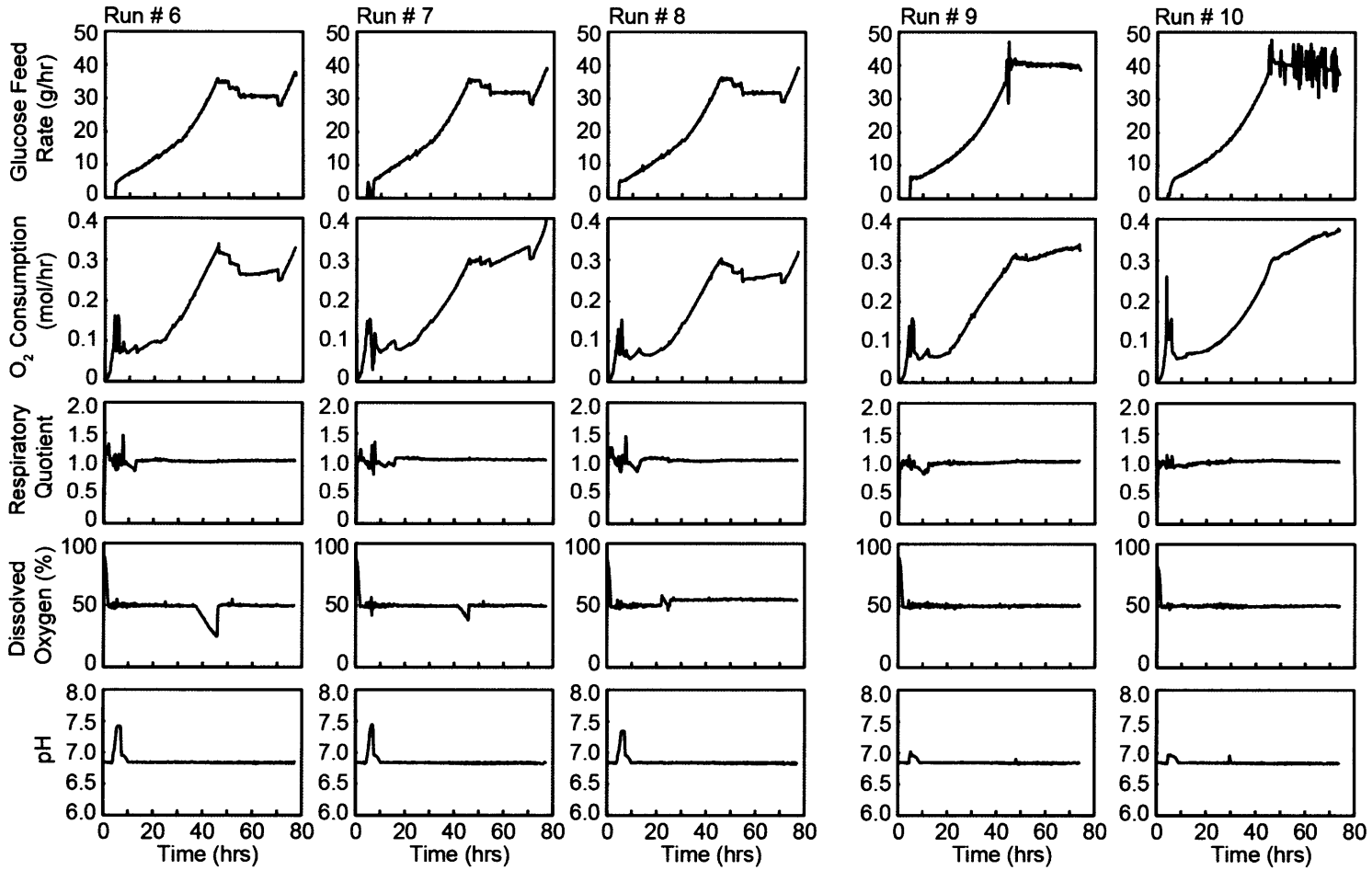


Figure B.10 (continued)

Bibliography

[1] Lee, J., Lee, S.Y., Park, S., and A.P.J. Middelberg (1999). Control of fed-batch fermentation. *Biotechnology Advances* **17**: 29-48.

[2] Sontag, E. Molecular System Biology and Control (2005). *European Journal of Control* **11**: 396-435.

[3] Enfors, S.-O., Jahic, M., Rozkov, A., Xu, B., Hecker, M., Jűrgen, B., Kruger, E., Schweder, T., Hamer, G., O'Beirne, D., Noisommit-Rizzi, N., Reuss, M., Boone, L., Hewitt, C., McFarlane, C., Nienow, A., Kovacs, T., Trägardh, C., Fuchs, L., Revstedt, J., Friberg, P.C., Hjertager, B., Blomsten, G., Skogman, H., Hjort, S., Hoeks, F., Lin, H.-Y., Neubauer, P., van der Lans, R., Lyben, K., Vrabel, P. and Å. Manelius. (2001). Physiological responses to mixing in large scale bioreactors. *J. Biotechnol.* **85**: 175-185.

[4] Eiteman, M.A., and E. Altman (2006). Overcoming acetate in Escherichia coli recombinant protein fermentations. *TRENDS in Biotechnology* **24(11):** 530-536.

[5] Clancy, K. and C.A.Voigt (2010). Programming Cells: Towards an automated 'Genetic Compiler'. *Curr. Opin. Biotech.* **21**:1-10.

[6] Salis, H., Mirsky, E.A., and C.A. Voigt (2009). Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology* **27**: 946-951.

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

[8] Dueber, J.E., Yeh, B.J., Chak, K., & W.A. Lim. Reprogramming Control of an Allosteric Signaling Switch Through Modular Recombination (2003). *Science* **301**: 1904-1908.

[9] Isaacs, F.J., Dwyer, D.J., & J.J. Collins (2006). RNA synthetic biology. *Nature Biotechnology* **24**: 545-554.

[10] Bashor, C., Helman, N.C., Yan, S., and W.A. Lim (2008). Using Engineered Scaffold Interactions to Reshape MAP Kinase Pathway Signaling Dynamics. *Science* **319**: 1539-1543.

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

[12] Tamsir, A., Tabor, J.J., & C.A. Voigt. Robust multicellular computing using genetically encoded NOR gates and chemical 'wires' (2010). *Nature* **469**: 212-215.

[13] Daniel, R., Rubens, J.R., Sarpeshkar, R., and T.K. Lu (2013). Synthetic analog computation in living cells. *Nature* Advanced online publication, doi: 10.1038/nature12184

[14] Basu, S., Gerchman, Y., Collins, C.H., Arnold, F.H., & R. Weiss (2005). *Nature* **434**: 1130-1134.

[15] Tabor, J.J., Groban, E., & C.A. Voigt. *Systems Biology and Biotechnology of Escherichia coli*. Ch. 19. Performance Characteristics of Sensors and Circuits Used to Program *E. coli*. (Springer, New York, 2009).

[16] Salis, H., Tamsir, A., & C.A. Voigt (2009). Engineering Bacterial Signals and Sensors. *Contrib. Microbiol.* **16**: 194-225.

[17] Adam Arkin, Personal communication.

[18] Endy, D. Foundations of Engineering Biology (2005). *Nature* **438**: 449-453.

[19] Lutz, R. & H. Bujard. Independent and tight regulation of transcriptional units in *Escherichia coli* via the lacR/O, the TetR/O and AraCI1-12 regulatory elements (1997). *Nucleic Acids Res.* **25(6):** 1203-1210.

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

[21] Pfleger, B.F., Pitera, D.J., Newman, J.D., Martin, V.J.J., & J.D. Keasling (2007). Microbial sensors for small molecules: Development of a mevalonate biosensor. *Metabolic Engineering* **9**: 30-38.

[22] Behzadian, F., Barjeste, H., Hosseinkhani, S., and A.R. Zarei (2011). Construction and characterization of *Escherichia coli* whole-cell biosensors for toluene and related compounds. *Curr. Microbiol.* **62(2)**: 690-696.
[23] BioBricks Foundation. "The Registry of Standard Biological Parts." Online: http://partsregistry.org/Main_Page (2013).
[24] Tang, S.-H., and P.C. Cirino (2011). Design and Application of a Mevalonate-Responsive Regulatory Protein. *Angew. Chem. Int. Ed.* **60**: 1084-1086.
[25] Purnick, P.E. & R. Weiss (2009). The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.* **10(6):** 410-422.
[26] Gredell, J.A., Frei, C.S., and P.C. Cirino (2012). Protein and RNA engineering to customize microbial molecular reporting. *Biotechnol. J.* **7**: 477-499.
[27] Becskei, A. Seraphin, B., and L. Serrano (2001). Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO Journal* **20**: 2528-2535.
[28] Buchler, N.E. and F.R. Cross (2009). Protein sequestration generates a flexible ultrasensitive response in a genetic network. *Mol. Sys. Biol.* **5**: 272.
[29] Dueber, J.E., Mirsky, E.A., and W.A. Lim (2007). Engineering synthetic signaling proteins with ultrasensitive input/output control. *Nature Biotechnology* **25(6)**: 660-662.
[30] Hart, Y., Antebi, Y.E., Mayo, A.E. Friedman, N., and U. Alon (2012). Design principles of cell circuits with paradoxical components. *Proc. Natl. Acad. Sci.* **109(21)**: 8346-8351.
[31] Rosenfeld, N., Elowitz, M.B. and U. Alon (2002). Negative autoregulation speeds the response times of transcriptional networks. *J. Mol. Biol.* **323(5)**: 785-793.
[32] Bashor, C.J., Helman, N.C. Yan, S., and W.A. Lim (2008). Using engineered scaffold interactions to reshape MAP kinase pathways signaling dynamics. *Science* **319**: 1539-1543.
[33] Milo, R., Shen-Orr, S., Itzkovitz, S. Kashtan, N., Chklovskii, D., and U. Alon (2002). Network Motifs: Simple Building Blocks of Complex Networks. *Science* **298**: 824-827.
[34] El-Samad, H. and M. Khammash (2006). Regulated degradation is a mechanism for suppressing stochastic fluctuations in gene regulatory networks. *Biophys. J.* **90(10)**: 3748-2761.
[35] Shoval, O., Alon, U., and E. Sontag (2011). Symmetry Invariance for Adapting Biological Systems. *SIAM J. Applied Dynamical Systems* **10(3)**: 857-886.
[36] Skataric, M. and E.D. Sontag (2012). A Characterization of Scale Invariant Responses in Enzymatic Networks. *PLOS Comp Biol.* **8(11)**: 1-10.
[37] Stanton, B., Nielsen, A., and C.A. Voigt (2013). Personal Communication.
[38] Bhattacharyya, R.P., Remenyi, A., Yeh, B.J., & W.A. Lim (2006). Domains, Motifs, and Scaffolds: The Role of Modular Interactions in Evolution and Wiring of Cell Signaling Circuits. *Annu. Rev. Biochem.* **75**: 655-680.
[39] Webber, W., Link, N., & M. Fussenegger (2008). A genetic redox sensor for mammalian cells. *Metabolic Engineering* **8**: 273-280.
[40] Meng, X., Brodsky, M.H., and S.A. Wolfe (2005). A bacterial one-hybrid system for determining the DNA-binding specificity of transcription factors. *Nature Biotechnology* **23(8)**: 988-994.
[41] Lee, S.Y (1996). High cell-density culture of *Escherichia coli. Trends Biotechnol.* **14(3)**: 98-105.
[42] Enfors, S.-O., Jahic, M., Rozkov, A., Xu, B., Hecker, M., Jűrgen, B., Kruger, E., Schweder, T., Hamer, G., O'Beirne, D., Noisommit-Rizzi, N., Reuss, M., Boone, L., Hewitt, C., McFarlane, C., Nienow, A., Kovacs, T., Trägardh, C., Fuchs, L., Revstedt, J., Friberg, P.C., Hjertager, B., Blomsten, G., Skogman, H., Hjort, S., Hoeks, F., Lin, H.-Y., Neubauer, P., van der Lans, R., Lyben, K., Vrabel, P. and Å. Manelius. (2001). Physiological responses to mixing in large scale bioreactors. *J. Biotechnol.* **85**: 175-185.

[43] Xu, B., Jahic, M., Blomsten, G., & S.-O. Enfors (1999). Glucose overflow metabolism and mixed-acid fermentation in aerobic large-scale fed-batch processes with Escherichia coli. *Appl. Microbiol. Biotechnol*. **51**: 564-571.

[44] Clancy, K. and C.A.Voigt (2010). Programming Cells: Towards an automated 'Genetic Compiler'. *Curr. Opin. Biotech.* **21**,1-10.

[45] Canton, B., Labno, A., and D. Endy (2008). Refinement and standardization of synthetic biological parts and devices. *Nature Biotechnology* **26**, 787-793.

[46] Levskaya, A., Chevalier, A.A., Tabor, J.J., Simpson, Z.B., Lavery, L.A., Levy, M., Davidson, E.A., Scouras, A., Ellington, A.D., Marcotte, E.M., and C.A. Voigt (2005). Engineering *E. coli* to see light. *Nature* **438**: 441-442.

[47] Tabor, J.J., Levskaya, A., and C.A. Voigt (2011). Multichromatic Control of Gene Expression in *Escherichia coli*. *J. Mol. Biol.* **405**: 315-324.

[48] Horiuchi, T., and H. Inokuchi. Temperature-sensitive Regulation System of Prophage Lambda Induction (1967). *J. Mol. Biol.* **23**: 217-224.

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

[50] Mason, H.S., DeWald, D.B., and J.E. Mullet. Identification of a Methyl Jasmonate-Responsive Domain in the Soybean *vspB* promoter (1993). *The Plant Cell* 5:241-251.

[51] Weber, W., Daoud-El Baba, M., and M. Fussenegger (2004). Synthetic ecosystems based on airborne inter- and intrakingdom communication. *Proc. Nat. Acad. Sci.* **104(25):** 10435-10440.

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

[53] Baumann, B. & J.R. van der Meer (2007). Analysis of bioavailable arsenic in rice with whole cell living bioreporter bacteria. *J. Agric. Food Chem*. **55**: 2114-2120.

[54] Behzadian, F., Barjeste, H., Hosseinkhani, S., and A.R. Zarei (2011). Construction and characterization of *Escherichia coli* whole-cell biosensors for toluene and related compounds. *Curr. Microbiol*. **62(2)**: 690-696.

[55] Dietrich, J.A., Shis, D.L., Alikhani, A., and J.D. Keasling (2013). Transcription Factor-Based Screens and Synthetic Selections for Microbial Small-Molecule Biosynthesis. *ACS Synth. Biol*. **2**:47-58.

[56] Antunes, M.S., Morey, K.J., Smith, J.J., Albrecht, K.D., Bowen, T.A., Zdunek, J.K., Troupe, J.F., Cuneo, M.J., Webb, C.T., Hellinga, H.W., and J.I. Medford (2011). Programmable Ligand Detection System in Plants through a Synthetic Signal Transduction Pathway. *PlosONE* **6(1)**: 1-11.

[57] Liu, X., Germaine, K.J., Ryan, D., and D.N. Dowling (2010). Whole-Cell Fluorescent Bioesensors for Bioavailability and Biodegradation of Polychlorinated Biphenyls. *Sensors* **10**: 1377-1398.

[58] Chen, M.-T., and R. Weiss (2005). Artificial cell-cell communication in yeast *Saccharomyces cerevisiae* using signaling elements from *Arabidopsis thaliana*. *Nature Biotechnology* **23(12)**: 1551-1555.

[59] Levskaya, A., Weiner, O.D., Lim, W.A., and C.A. Voigt (2009). Spatiotemporal control of cell signaling using a light-switchable protein interaction. *Nature* **461**: 997-1001.

[60] Webber, W., Link, N., & M. Fussenegger (2008). A genetic redox sensor for mammalian cells. *Metabolic Engineering* **8**: 273-280.

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

[62] Kobayashi, H., Kaern, M., Araki, M., Chung, K., Gardner, T.S., Cantor, C.R., and J.J. Collins (2004). Programmable cells: Interfacing natural and engineered gene networks. *Proc. Natl. Acad. Sci.* **101(22)**: 8414-8419.

[63] America Chemical Society Committee on Environmental Improvement (1980). Guidelines for Data Acquisition and Data Quality Evaluation in Environmental Chemistry. *Anal. Chem*. **52**: 2242-2249.

[64] Looneborg, R., Smrinova, I., Dian, C., Leonard, G.A., and P. Brzezinski (2007). *In vivo* and *in vitro* Investigation of Transcriptional Regulation by DntR. *J. Mol. Biol.* **372**: 571-582.

[65] Van der Meer, J.R., and S. Belkin (2010). Where microbiology meets microengineering: design and applications of reporter bacteria. *Nat. Rev. Microbiol*. **8**: 511-522.

[66] Mutalik, V.K., Guimaraes, J.C., Cambray, G., Lam, C., Christoffersen, M.J., Mai, Q.-A., Tran, A.B., Paull, M., Keasling, J.D., Arkin, A.P., and D. Endy (2013). Precise and reliable gene expression via standard transcription and translation initiation elements. *Nature Methods* **10(4)**: 354-360.

[67] Salis, H.M., Mirsky, E.A., and C.A. Voigt. Automated design of synthetic ribosome binding sites to control protein expression (2009). *Nature Biotechnology* **27 (10)**: 946-950.

[68] Cambray, G., Guimaraes, J.C., Mutalik, V.K., Lam, C., Mai, Q.-A., Thimmaiah, T., Carothers, J.M., Arkin, A.P., and D. Endy (2013). Measurement and modeling of intrinsic transcription terminators. *Nucleic Acids Research* **41 (9)**: 5139-5148.

[69] Chen, Y.-J., Liu, P., Nielsen, A.A.K., Brophy, J.A.N., Clancy, K., Peterson, T., and C.A. Voigt (2013). Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nature Methods* **10**: 659-664.

[70] Kittleson, J.T., Cheung, S., and J.C. Anderson (2011). Rapid optimization of gene dosage in *E. coli* using DIAL strains. *J. Biol. Eng.* **5(10)**: 1-7.

[71] Jones, G.D.D., Le Pla, R.C., & P.B. Farmer (2010). Phosphotriester adducts (PTEs): DNA's overlooked lesion. *Mutagenesis* **25**, 3-16.

[72] Himmelstein, M.W., Boogaard, P.J., Cadet, J., Farmer, P.B., Kim, J.H., Martin, E.A., Persaud, R., and D.E.G. Shuker (2009). Creating context for the use of DNA adduct data in cancer assessment: II. Overview of methods of identification and quantitation of DNA damage. *Critical Reviews in Toxicology* **39(8):** 679-694.

[73] Shooter, K.V (1978). DNA phosphotriesters as indicators of cumulative carcinogen-induced damage. *Nature* **274(10):** 612-614.

[74] Song, R., Zhang, W., Chen, H., Ma, H., Dong, Y., Sheng, G., Zhou, Z., and J. Fu (2005). Site determination of phenyl glycidyl ether-DNA adducts using high-performance liquid chromatography with elecstrospray ionization tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **19**: 1120–1124.

[75] Liteplo, R.G., Meek, M.E., and W. Windle (2002). "N-Nitrosodimethylamine" in Concise International Chemical Assessment Document 38. World Health Organization.

[76] U.S. Department of Health and Human Services, Public Health Service, National Toxicology Program (2011). "Dimethyl Sulfate." In *Report on Carcinogens. Twelfth Edition.*

[77] Sussmuth, R., Haerlin, R., and F. Lingens (1972). The mode of action of N-methyl-N'-nitro-N-nitrosoguanidine in mutagenesis. *Biochim. Biophys. Acta*, **269**: 276-286.

[78] Jones., N. "Scientists Fume over California's Pesticide Plans." Nature News. Published online 4 May, 2010. <http://www.nature.com/news/2010/100504/full/news.2010.218.html>

[79] World Health Organization, International Agency for Research on Cancer (1999). *Monographs on the Evaluation of Carcinogenic Risk to Humans, Vol. 71; Methyl Chloride.*

[80] Myers, L.C., Wagner, G., and G.L. Verdine (1995). Direct Activation of the Methyl Chemosensor protein N-Ada by CH_3I. *J. Am. Chem. Soc.* **117**:10749-10750.

[81] He, C., Hus, J., Sun, L.J., Zhou, P., Norman, D.P.G., Dötsch, V., Wei, H., Gross, J.D., Lane, W.S., Wagner, G., & G.L. Verdine (2005). A Methylation-Dependent Electrostatic Switch Controls DNA Repair and Transcriptional Activation by *E. coli* Ada. *Molecular Cell* **20**, 117-129.

[82] Myers, L.C., Jackow, F., and G.L. Verdine (1995). Metal Dependence of Transcriptional Switching in *Escherichia coli* Ada. *J. Biol. Chem*. **270(12):** 6664-6670.

[83] Sedgwick, B. Repairing DNA-Methylation Damage (2004). *Nat. Rev. Mol. Cell Biol.* **5**, 148-157.

[84] Vaughan, P., Lindahl, T., & B. Sedgwick (1993). Induction of the adaptive response of Escherichia coli to alkylation damage by the environmental mutagen, methyl chloride. *Mutat. Res*. **293**: 249-257.

[85] Vollmer, A.C., Belkin, S., Smulski, D.R., Van Dyk, T.K., and R.A. LaRossa (1997). Detection of DNA damage by use of Escherichia coli carrying recA'::lux, uvrA'::lux, or alkA'::lux reporter plasmids. *Appl. Environ. Microbiol.* **63 (7)**: 2566.

[86] Young, K.H (1998). Yeast Two-Hybrid: So Many Interactions, (in) So Little Time. *Biology of Reproduction* **58**: 302-311.

[87] Elliot, D.A. and A.H. Brand (2008). "The GAL4 System." In Methods in Molecular Biology: Drosophila: Methods and Protocols. Ed. C. Dahmann. Humanna Press, Totowa.

[88] Ma, J., Przibilla, E., Hu, J., Bogorad, L., and M., Ptashne (1988). Yeast activators stimulate plant gene expression. *Nature* **334**: 631-633.

[89] Datsenko, K.A., and B.L. Wanner. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products (2000). *Proc. Natl. Acad. Sci.* **97(12)**: 6640-6645.

[90] Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison III, C.A., and H.O. Smith (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6(5):** 343-345.

[91] Shetty, R.P., Endy, D., and T.F. Knight, Jr (2008). Engineering BioBrick vectors from BioBrick parts. *Journal of Biological Engineering* **2:** 5.

[92] Registry of Biological Parts. www.partregistry.org Biobricks Foundation (2013).

[93] Cormack, B.P., Valdivia, R.H., & S. Falkow (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**: 33-38.

[94] Takinowaki, H., Matsuda, Y., Yoshida, T., Kobayashi, Y., and T. Ohkubo (2006). The solution structure of the methylated form of the N-terminal 16-kDa domain of Escherichia coli Ada protein. *Protein Science* **15:** 487-497.

[95] Cormack, B.P., Bertram, G., Egerton, M., Gow, N.A.R., Falkow, S., and A.J.P. Brown (1997). Yeast-enhanced green fluorescent protein (yEGFP): a reporter of gene expression in *Candida albicans. Microbiology* **43**: 303-311.

[96] Gietz, R. D. and R.A. Woods (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol.* **350**: 87–96.

[97] Vaughan, P., Sedgwick, B., Hall, J., Gannon, J., and T. Lindahl (1991). Environmental mutagens which induce the adaptive response to alkylation damage in Escherichia coli. *Carcinogenesis* **12**: 263-268.

[98] Khlebnikov, A., Risa, Ø., Skaug, T., Carrier, T.A., and J.D. Keasling (2000). Regulatable Arabinose-Inducible Gene Expression System with Consistent Control in All Cells of a Culture. *Journal of Bacteriology* **182(24):** 7029-7034.

[99] Dietrich, J.A., McKee, A.E., &. J.D. Keasling (2010). High-Troughput Metabolic Engineering: Advances in Small-Molecule Screening and Selection. *Annu. Rev. Biochem.* **79**, 563-90.

[100] Young, K.H. (1998). Yeast Two-Hybrid: So Many Interactions, (in) So Little Time. *Biology of Reproduction* **58**: 302-311.

[101] Sharon, E., Kalma, Y., Sharp, A., Raveh-Sadka, T., Levo, M., Zeevi, D., Keren, L., Yakhini, Z., Weinberger, A., and E. Segal (2012). Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nature Biotech*nology **30**: 521-530.

[102] Mumberg, D., Muller, R., and M. Funk (1995). Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156**: 119-122.

[103] Shrivastav, N., Li, D., and J.M. Essigmann (2010). Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation. *Carcinogenesis* **31(1)**: 59-70.

[104] Beranek, D.T. (1990). Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutation Research* **231**: 11-30.

[105] Volkert, M.R., Gately, F.H., & L.I. Hajec (1988). Expression of DNA damage-inducible genes of *Escherichia coli* upon treatment with methylating, ethylating, and propylating agents. *Mutat. Res.* **217**: 109-115.

[106] Nakamura, T., Tokumoto, Y., Sakumi, K., Koike, G., Nakabeppu, Y., and M. Sekiguchi (1988). Expression of the ada Gene of Escherichia coli in Response to Alkylating Agents: Identification of Transcriptional Regulatory Elements. *J. Mol. Biol.* **202**: 483-494.

[107] Pfleger, B.F., Pitera, D.J., Newman, J.D., Martin, V.J.J., and J.D. Keasling (2007). Microbial sensors for small molecules: Development of a mevalonate biosensor. *Metabolic Engineering* **9**: 30-38.

[108] Tang, S.-Y., Akinterinwa, O., and P. Cirino (2013). Screening for enhanced triacetic acid lactone (TAL) production by recombinant *Escherichia coli* expressing a designed TAL reporter. *JACS. In press.*

[109] Bayer, T.S., Widmaier, D.M., Temme, K., Mirsky, E.A., Santi, D.V. & C.A. Voigt (2009). Synthesis of Methyl Halides from Biomass Using Engineered Microbes. *JACS* **131**, 6508-6515.

[110] California Environmental Protection Agency (2010). Department of Pesticide Regulation. *Risk Characteriziation Document for Inhalation Exposure Vol. III: Environmental Fate.* Online: http://www.cdpr.ca.gov/docs/risk/mei/mei_vol3_ef.pdf.

[111] Hu, Q.H., Moran, J.E., and J.Y. Gan (2012). Sorption, degradation, and transport of methyl iodide and other iodine species in geologic media. *Applied Geochemistry* **27**: 774-781.

[112] Gan, J. and S.R. Yates (1996). Degradation and Phase Partition of Methyl Iodide in Soil. *J. Agric. Food Chem*. **44**: 4001-4008.

[113] Ramon-Ascon, J., Yasukawa, T., and F. Mizutan (2012). Detection of Pesticide Residues Using Biosensors. In *Biosensors and environmental health*, Ed. V.R. Preedy, V. Patel. Pp 21-40.

[114] Gan, J. and S.R. Yates (1996). Degradation and Phase Partition of Methyl Iodide in Soil. *J. Agric. Food Chem*. **44**: 4001-4008.

[115] Dueber, J.E., Yeh, B.J., Chak, K., and W.A. Lim (2003). Reprogramming Control of an Allosteric Signaling Switch Through Modular Recombination. *Science* **301**: 1904-1908.

[116] Sharon, E., Kalma, Y., Sharp, A., Raveh-Sadka, T., Levo, M., Zeevi, D., Keren, L., Yakhini, Z., Weinberger, A., and E. Segal (2012). Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nature Biotechnology* **30**: 521-530.

[117] Ramon-Ascon, J., Yasukawa, T., and F. Mizutan (2012). Detection of Pesticide Residues Using Biosensors. In *Biosensors and environmental health*, Ed. V.R. Preedy, V. Patel. Pp 21-40.

[118] Eiteman, M.A., and E. Altman (2006). Overcoming acetate in Escherichia coli recombinant protein fermentations. *TRENDS in Biotechnology* **24(11):** 530-536.

[119] Majewski, R.A. and M.M. Domach (1990). Simple Constrained-Optimization View of Acetate Overflow in *E. coli*. *Biotechnology and Bioengineering* **35**: 732-738.

[120] Lin, H.Y., Mathiszik, B., Xu, B., Enfors, S.-O., and P. Neubauer (2001). Determination of the maximum specific uptake capacities for glucose and oxygen in glucose-limited fed-batch cultivations of *Escherichia coli*. *Biotechnology and Bioengineering* **73(5):** 348-357.

[121] Wolfe, A.J. (2005). The Acetate Switch. *Microbiology and Molecular Biology Reviews* **69(1)**: 12-50.

[122] Shiloach, J., Kaufman, J., Guillard, A.S., and R. Fass (1996). Effect of glucose supply strategy on acetate accumulation, growth, and recombinant protein production by *Escherichia coli* BL21 (λDE3) and *Escherichia coli* JM109. *Biotechnology and Bioengineering* **49**: 421-428.

[123] Wolfe, A.J. (2005). The Acetate Switch. *Microbiology and Molecular Biology Reviews* **69(1)**: 12-50.

[124] Lara, A.R., Taymaz-Nikerel, H., Mashego, M.R., van Gulik, W.M., Heijnen, J.J., Ramirez, O.T., and W.A. van Winden (2009). Fast dynamic responsonse of the fermentative Metabolism of *Escherichia coli* to Aerobic and Anaerobic Glucose Pulses. *Biotechnology and Bioengineering* **104(6):** 1153-1161.

[125] Salmon, K., Hung, S., Mekjian, K., Baldi, P., Hatfield, G.W., and R.P. Gunsalus (2003). Global Gene Expression Profiling in Escherichia coli K12: The Effects of Oxygen Availability and FNR. *J. Biol. Chem.* **278(32):** 29837-29855.

[126] Eiteman, M.A., and E. Altman (2006). Overcoming acetate in Escherichia coli recombinant protein fermentations. *TRENDS in Biotechnology* **24(11):** 530-536.

[127] Wimpenny, J.W.T., and A. Firth (1972). Levels of Nicotinamide Adenince Dinucleotide and Reduced Nicotinamide Adenine Dinucleotide in Facultative Bacteria and the Effect of Oxygen. *J. Bacteriol.* **111(1)**: 24-32.

[128] Eiteman, M.A., and E. Altman (2006). Overcoming acetate in Escherichia coli recombinant protein fermentations. *TRENDS in Biotechnology* **24(11):** 530-536.

[129] De Mey, M., De Maeseneire, S., Soetart, W., and E. Vandamme (2007). Minimizing acetate formation in *E. coli* fermentations. *J. Ind. Microbiol. Biotechnol.* **34**: 689-700.

[130] Datsenko, K.A., and B.L. Wanner (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97(12)**: 6640-6645.

[131] Bulter, T., Lee, S.-G., Wong, W.W., Fung, E., Connor, M.R., and J.C. Liao (2004). Design of artificial cell-cell communication using gene and metabolic networks. *Proc. Natl. Acad. Sci.* **101(8)**: 2299-2304.

[132] Lara, A.R., Taymaz-Nikerel, H., Mashego, M.R., van Gulik, W.M., Heijnen, J.J., Ramirez, O.T., and W.A. van Winden (2009). Fast dynamic responsonse of the fermentative Metabolism of *Escherichia coli* to Aerobic and Anaerobic Glucose Pulses. *Biotechnology and Bioengineering* **104(6):** 1153-1161.

[133] Enfors, S.-O., Jahic, M., Rozkov, A., Xu, B., Hecker, M., Jurgen, B., Kruger, E., Schweder, T., Hamer, G., O'Beirne, D., Noisommit-Rizzi, N., Reuss, M., Boone, L., Hewitt, C., McFarlane, C., Nienow, A., Kovacs, T., Tragardh, C., Fuchs, L., Revstedt, J., Friberg, P.C., Hjertager, B., Blomsten, G., Skogman, H., Hjort, S., Hoeks, F., Lin, H-Y., Neubauer, P., van Der Lans, R., Luyben, K., Vrabel, P., and A. Manelius (2001). Physiological responses to mixing in large bioreactors. *J. Biotechnology* **85***:* 175-185.

[134] Xu, B., Jahic, M., Blomsten, G., and S.-O. Enfors (1999). Glucose overflow metabolism and mixed-acid fermentation in aerobic large-scale fed-batch processes with *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **51**: 564-571.

[135] Patschkowski, T., Bates, D.M., and P.J Kiley (2000). Mechanisms for sensing and responding to oxygen deprivation. *Bacterial stress responses. ASM Press, Washington, DC*, 61-78.

[136] Mettert, E. L., & Kiley, P. J. (2007). Contributions of [4Fe-4S]-FNR and integration host factor to fnr transcriptional regulation. *J.Bacteriology*,**189(8)**, 3036-3043.

[137] Meng, W., Green, J., and J.R. Guest (1997). FNR-dependent repression of *ndh* gene expression requires two upstream FNR-binding sites. *Microbiology* ***143*(5)**: 1521-1532.

[138] Bell, A. I., Cole, J. A., and S.J. Busby (1990). Molecular genetic analysis of an FNR-dependent anaerobically inducible Escherichia coli promoter. *Molecular Microbiology* **4(10):** 1753-1763.
[139] Barnard, A. M., Green, J., and S.J. Busby (2003). Transcription regulation by tandem-bound FNR at Escherichia coli promoters. *J. Bacteriology* **185(20)**: 5993-6004.
[140] Ramos, H. C., Boursier, L., Moszer, I., Kunst, F., Danchin, A., and P. Glaser (1995). Anaerobic transcription activation in *Bacillus subtilis*: identification of distinct FNR-dependent and-independent regulatory mechanisms. *The EMBO Journal* **14(23)**: 5984.
[141] Agron, P.G., Monson, E.K., Ditta, G.S., and D.R. Helinski. Oxygen regulation of expression of nitrogen fixation genes in *Rhizobium meliloti*. *Research in Microbiology* **145(5)**:454-459 (1994).
[142] David, M., Daveran, M. L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., and D. Kahn (1988). Cascade regulation of *nif* gene expression in *Rhizobium meliloti*. *Cell* **54(5)**: 671-683.
[143] Galinier, A., Garnerone, A. M., Reyrat, J. M., Kahn, D., Batut, J., and P. Boistard (1994). Phosphorylation of the *Rhizobium meliloti* FixJ protein induces its binding to a compound regulatory region at the fixK promoter. *Journal of Biological Chemistry* **269(38)**: 23784-23789.
[144] Waelkens, F., Foglia, A., Morel, J. B., Fourment, J., Batut, J., and P. Boistard (1992). Molecular genetic analysis of the Rhizobium meliloti fixK promoter: identification of sequences involved in positive and negative regulation. *Molecular microbiology* **6**(11): 1447-1456.
[145] Browning, D. F., and S.J. Busby (2004). The regulation of bacterial transcription initiation. *Nature Reviews Microbiology* **2(1)**: 57-65.
[146] Drepper, T., Eggert, T., Circolone, F., Heck, A., Krauß, U., Guterl, J. K., and K.E. Jaeger (2007). Reporter proteins for in vivo fluorescence without oxygen. *Nature biotechnology* **25(4):** 443-445.
[147] Bulter, T., Lee, S.-G., Wong, W.W., Fung, E., Connor, M.R., and J.C. Liao (2004). Design of artificial cell-cell communication using gene and metabolic networks. *Proc. Natl. Acad. Sci.* **101(8)**: 2299-2304.
[148] Huo, Y.-X., Tian, Z.-X., Rappas, M., Wen, J., Chen, Y.-C., You, C.-H., Zhang, X., Buck, M., Wang, Y.-P., and A. Kolb (2006). Protein-induced DNA bending clarifies the architectural organization of the σ^{54}-dependent *glnAP2* promoter. *Molecular Microbiology* **59(1)**: 168-180.
[149] Hardiman, T., Lemuth, K., Keller, M.A., Reuss, M., and M. Siemann-Herzberg (2007). Topology of the global regulatory network of carbon limitation in Escherichia coli. *J. Biotech.* **132**: 359-374.
[150] Bennett, B.D., Kimball, E.H., Gao, M., Osterhout, R., Van Dien, S.J., and J.D. Rabinowitz (2009). Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Biotech.* **5(8)**: 593-599.
[151] Ramseier, T.M., Negre, D., Cortay, J.-C., Scarabel, M., Cozzone, A.J., and M. H. Saier, Jr. (1993). In Vitro Binding of the Pleiotropic Transcriptional Regulatory Protein, FruR, to the *fru*, *pps*, *ace*, *pts*, and *icd* Operons of *Escherichia coli* and *Salmonella typhimurium*. *J. Mol. Biol.* **234**: 28-44.
[152] Saier Jr., M.H., and T.M. Ramseier (1996). The Catabolite Repressor/Activator (Cra) Protein of Enteric Bacteria. *J. Bac.* **178(12):** 3411-3417.
[153] Moon, T.S., Clarke, E.J., Groban, E.S., Tamsir, A., Clark, R.M., Eames, M., Kortemme, T., and C.A. Voigt (2011). Construction of a genetic multiplexor to toggle between chemosensory pathways in *Escherichia coli*. *J. Mol. Biol.* **406**, 215-227.
[154] Thomas Segall-Shapiro, personal communication.
[155] Temme, K., Hill, R., Segall-Shapiro, T.H., Moser, F., and C.A. Voigt (2012). Modular control of multiple pathways using engineered orthogonal T7 polymerases. *Nuc. Acid. Res*. **40(17)**: 8773-8781.
[156] De Mey, M., De Maeseneire, S., Soetart, W., and E. Vandamme (2007). Minimizing acetate formation in *Escherichia coli* fermentations. *J. Ind. Microbiol. Biotechnol.* **34**: 689-700.

[157] Lei, Q., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissmann, J.S., Arkin, A.P., and W.A. Lim (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152(5)**: 1173-1183.

[158] Lin, H., Castro, N.M., Bennett, G.N., and K.Y. San (2006). Acetyl-CoA synthetase overexpression in *Escherichia coli* demonstrates more efficient acetate assimilation and lower acetate accumulation: a potential tool in metabolic engineering. *Appl. Microbiol. Biotechnol.* **71(6)**: 870-874.

[159] Vemuri, G.N., Eiteman, M.A., and E. Altman (2006). Increased recombinant protein production in *Escherichia coli* strains with overexpressed water-forming NADH oxidase and a delete ArcA regulatory protein. *Biotechnology and Bioengineering* **94(3)**: 538-542.

[160] Tsai, P.S., Hatzimanikatis, V., and J.E. Bailey (1996). Effect of *Vitreoscilla* hemoglobin dosage on micraerobic *Escherichia coli* carbon and energy metabolism. *Biotechnology and Bioengineering* **49(2):** 139-150.

[161] Van de Walle, M., and J. Shiloach (1997). Proposed mechanism of acetate accumulation in two recombinant *Escherichia coli* strains during high density fermentation. *Biotechnology and Bioengineering* **57**:71-78.

[162] Medema, M.H., Breitling, R., Bovenberg, R., and E. Takano (2011). Exploiting plug-and-play synthetic biology for drug discovery and production in microorganisms. *Nat. Rev. Microbiol.* **9:** 131-137.

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

[164] Purnick, P.E.M. and R. Weiss (2009). The second wave of synthetic biology: from modules to systems. *Nat. Rev.* **10:** 410-422.

[165] Tabor, J.J., Salis, H.M., Simpson, Z.B., Chevalier, A.A., Levskaya, A., Marcotte, E.M., Voigt, C.A., and A.D. Ellington (2009). A synthetic genetic edge detection program. *Cell* **137**: 1272-1281.

[166] Lou, C., Liu, X., Ni, M., Huang, Y., Huang, Q., Huang, L., Jiang, L., Lu, D., Wang, M., Liu, C., Chen, D., Chen, C., Chen, X., Yang, L., Ma, H., Chen, J., and Q. Ouyang (2010). Synthesizing a novel genetic sequential logic circuit: a push-on push-off switch. *Molecular Systems Biology* **6**: 1-10.

[167] Friedland, A.E., Lu, T.K., Wang, X., Shi, D., Church, G., and J.J. Collins (2009). Synthetic gene networks that count. *Science* **324**, 1199-1202.

[168] Moon, T.S., Clarke, E.J., Groban, E.S., Tamsir, A., Clark, R.M., Eames, M., Kortemme, T., and C.A. Voigt (2011). Construction of a genetic multiplexor to toggle between chemosensory pathways in *Escherichia coli*. *J. Mol. Biol.* **406**, 215-227.

[169] Basu, S., Gerchman, Y., Collins, C.H., Arnold, F.H., and R. Weiss (2005). A synthetic multicellular system for programmed pattern formation. *Nature* **434**, 1130-1134.

[170] Schmidt, F.R. (2005). Optimization and scale up of industrial fermentation processes. *Appl. Microbiol. Biotechnol.* **68**, 425-435.

[171] Johnston, W., Cord-Ruwisch, R., M.J. Cooney (2002). Industrial control of recombinant *E. coli* fed-batch culture: new prospectives on traditional controlled variables. *Bioprocess Biosyst. Eng.* **25**, 111-120.

[172] Lee, S.Y (1994). High cell-density culture of Escherichia coli. *TIBTECH* **14**: 98-105.

[173] Schweder, T., Kruger, E., Jurgen, B., Blomsten, G., Enfors, S-O., and M. Hecker (1999). Monitoring of genes that respond to process-related stress in large-scale bioprocesses. *Biotech. & Bioeng.* **65**: 151-159.

[174] Kim, B.S., Lee, S.C., Lee, S.Y., Chang, Y.K., and H.N. Chang (2004). High cell density fed-batch cultivation of Escherichia coli using exponential feeding combined with pH-stat. *Bioprocess Biosyst Eng*. **26**, 147-150.

[175] Han, L., Enfors, S-O., and L. Haggstrom (2004). Changes in intracellular metabolite pools, and acetate formation in Escherichia coli are associated with a cell-density-dependent metabolic switch. *Biotech. Lett.* **24**: 483-488.

[176] Klumpp, S. and T. Hwa (2008). *Growth rate-dependent partitioning of RNA polymerases in bacteria. Proc. Natl. Acad. Sci.* **105**: 20245-20250.

[177] Nomura, M (1999). Regulation of ribosome biosynthesis in *Escherichia coli* and *Saccharomyces cerevisiae*: diversity and common principles. *J. Bacteriol.* **181**: 6857-6864.

[178] Enfors, S.-O., Jahic, M., Rozkov, A., Xu, B., Hecker, M., Jurgen, B., Kruger, E., Schweder, T., Hamer, G., O'Beirne, D., Noisommit-Rizzi, N., Reuss, M., Boone, L., Hewitt, C., McFarlane, C., Nienow, A., Kovacs, T., Tragardh, C., Fuchs, L., Revstedt, J., Friberg, P.C., Hjertager, B., Blomsten, G., Skogman, H., Hjort, S., Hoeks, F., Lin, H-Y., Neubauer, P., van Der Lans, R., Luyben, K., Vrabel, P., and A. Manelius (2001). Physiological responses to mixing in large bioreactors. *J. Biotechnology* **85***:* 175-185.

[179] Hewitt, C.J., Boon, L.A., and C.M. McFarlane (1998). The use of flow cytometry to study the impact of fluid mechanical stress on *Escherichia coli* W3110 during continuous cultivation in an agitated bioreactor. *Biotech. & Bioeng*. **52**: 612-620.

[180] Gill, R.T., DeLisa, M.P., Valdes, J.J., and W.E. Bentley (2001). Genomic analysis of high-cell-density recombinant *Escherichia coli* fermentation and "cell conditioning" for improved recombinant protein yield. *Biotech. & Bioeng.* **72**, 85-94.

[181] Yokobayashi, Y., Weiss, R., and F.H. Arnold (2002). Directed evolution of a genetic circuit. *Proc. Natl. Acad. Sci.* **99**, 16587-16591.

[182] Baker, D., Church, G., Collins, J., Endy, D., Jacobson, J., Keasling, J.D., Modrich, P., Smolke, C., and R. Weiss (2006). Engineering Life: Building a FAB for Biology. *Scientific American* **294**: 44-51.

[183] Anderson, J.C., Voigt, C.A., and A.P. Arkin (2007). Environmental signal integration by a modular AND gate. *Molecular Systems Biology* **3**: 133.

[184] Tamsir, A., Tabor, J.J., C.A. Voigt (2010). Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. *Nature* **212:** 212-215.

[185] Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S., and M.B. Elowitz. Accurate prediction of gene feedback circuit behavior from component properties. *Molecular Systems Biology* **3***:*143 (2007).

[186] Datsenko, K.A. and B.L. Wanner (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97(12***):* 6640-6645.

[187] Salis, H.M., Mirsky, E.A., and C.A. Voigt (2009). Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology* **27 (10)**: 946-950.

[188] Kelly, J.R., Rubin, A.J., Davis, J.H., Ajo-Franklin, C.M., Cumbers, J., Czar, M.J., de Mora, K., Glieberman, A.L., Monie, D.D. and D. Endy (2009). Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng*. **3**: 1-13.

[189] Weuster-Botz, D. (2000). Experimental design for fermentation media development: Statistical design or global random search? *J. Bioscience and Bioengineering* **90**: 473-483.

[190] Lee, S.Y. (1994). High cell-density culture of Escherichia coli. *TIBTECH* **14**, 98-105.

[191] Pfleger, B.F., Pitera, D.J., Smolke, C.D., and J.D. Keasling (2006). Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat. Biotech.* **24(8):** 1027-1032.

[192] Durfee, T., Nelson, R., Baldwin, S., Plunkett III, G., Burland, V., Mau, B., Petrosino, J.F., Qin, X., Muzny, D.M., Ayele, M., Gibbs, R.A., Csorgo, B., Posfai, G., Weinstock, G.M., and F.R. Blattner (2008). The Complete Genome Sequence of Escherichia coli DH10B: Insights into the Biology of a Laboratory Workhorse. *J. Bacteriol.* **190:** 2597-2606.

[193] Cox, J.M., Li, H., Wood, E.A., Chitteni-Pattu, S., Inman, R.B., and M.M. Cox (2008). Defective dissociation of a "slow" recA mutant protein imparts an Escherichia coli growth defect. *J. Biol. Chem.* **283(36),** 24909-24921.

[194] Becker, G.W. and H.M. Hsiung (1986). Expression, secretion and folding of human growth hormone in *Escherichia coli*. *FEBS* **204**: 145-150.

[195] Wetzel, R., Kleid, D.G., Crea, R., Heyneker, H.L., Yansura, D.G., Hirose, T., Kraszewaki, A., Riggs, A.D., Itakura, K., and D.V. Goeddel (1981). Expression in *Escherichia coli* of a chemically synthesized gene for a "mini-C" analog of human proinsulin. *Gene* **16**: 63-71.

[196] Horn, U., Strittmatter, W., Krebber, A., Knupfer, U., Kujau, M., Wenderoth, R., Muller, K., Matzku, S., Pluckthum, A., and D. Riesenberg (1996). High volumetric yields of functional dimeric miniantibodies in Escherichia coli, using an optimized expression vector and high-cell-density fermentation under non-limited growth conditions. *Appl. Microbiol. Biotechnol.* **46**: 524-532.

[197] Belagaje, R.M., Reams, S.G., Ly, S.C., and W.F. Prouty (1997). Increased production of low molecular weight recombinant proteins in *Escherichia coli. Protein Science* **6**: 1953-1962.

[198] Birch, G.M., Black, T., Malcolm, S.K., Lai, M.T., Zimmerman, R.E., and S.R. Jaskunas (1995). Purification of recombinant human rhinovirus 14 3C protease expressed in Escherichia coli. *Protein Expression and Purification* **6:** 609-618.

[199] Kelly, J.R., Rubin, A.J., Davis, J.H., Ajo-Franklin, C.M., Cumbers, J., Czar, M.J., de Mora, K., Glieberman, A.L., Monie, D.D. and D. Endy (2009). Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng*. **3**: 1-13.

[200] Clancy, K. and C.A.Voigt (2010). Programming Cells: Towards an automated 'Genetic Compiler'. *Curr. Opin. Biotech.* **21**: 1-10.

[201] Carothers, J.M., Goler, J.A., Juminaga, D., and J.D. Keasling (2011). Model-Driven Engineering of RNA Devices to Quantitatively Program Gene Expression. *Science* **334**: 1716-1719.

[202] Salis, H.M., Mirsky, E.A., and C.A. Voigt (2009). Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology* **27 (10)**: 946-950.

[203] Huber, R., Roth, S., Rahmen, N., and J. Büchs (2011) Utilizing high-throughput experimentation to enhance specific productivity of an E.coli T7 expression system by phosphate limitation. *BMC Biotechnol.* **11**: 1-11.

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

[205] Johnston, W., Cord-Ruwisch, R., M.J. Cooney (2002). Industrial control of recombinant *E. coli* fed-batch culture: new perspectives on traditional controlled variables. *Bioprocess Biosyst. Eng.* **25**: 111-120.

[206] Pan, J.G., Rhee, J.S., and J.M. Lebeault (1987). Physiological constraints of in increasing biomass concentration of Escherichia coli B in fed-batch culture. *Biotechnology Letters* **9**: 89-94.

[207] Hewitt, C.J. and A.W. Nienow (2007). The Scale-Up of Microbial Batch and Fed-Batch Fermenation Processes. *Advances in Applied Microbiology* **62**: 105-135.

[208] Schmidt, F.R. (2005) Optimization and scale up of industrial fermentation processes. *Appl. Microbiol. Biotechnol.* ***68***, 425-435.

[209] Baker, D., Church, G., Collins, J., Endy, D., Jacobson, J., Keasling, J.D., Modrich, P., Smolke, C., and R. Weiss (2006). Engineering Life: Building a FAB for Biology. *Scientific American* **294:** 44-51.

[210] Carothers, J.M., Goler, J.A., Juminaga, D., and J.D. Keasling (2011) Model-Driven Engineering of RNA Devices to Quantitatively Program Gene Expression. *Science* **334:** 1716-1719.

[211] Kelly, J.R., Rubin, A.J., Davis, J.H., Ajo-Franklin, C.M., Cumbers, J., Czar, M.J., de Mora, K., Glieberman, A.L., Monie, D.D. and D. Endy (2009). Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng.* **3:** 1-13.

[212] Sasson, V., Shachrai, I., Bren, A., Dekel, E., and U. Alon (2012). Mode of Regulation and the Insulation of Bacterial Gene Expression. *Mol. Cell* **46**: 399-407.

[213] Weber, H., Polen, T., Heuveling, J., Wendisch, V.F., and R. Hengge (2005). Genome-wide analysis of the general stress response network in Escherichia coli: sigmS-dependent genes, promoters, and sigma factor selectivity. *J. Bacteriol.* **187**: 1591-1603.

[214] Davis, J.H., Rubin, A.J., and R.T. Sauer (2011). Design, construction, and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res.* **39(3):** 1131-1141.

[215] Lou, C., Stanton, B., Chen, Y.-J., Munsky, B., and C.A. Voigt (2012). Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat. Biotech.* Advanced online publication. doi:10.1038/nbt.2401

[216] Lancaster, M.J., Sharp, R.J., Court, J.R., and I.D. McEntee (1989). Production of cloned carboxypeptidase G2 by *Escherichia coli*: Genetic and environmental considerations. *Biotechnology Letters* **2**: 699-704.

[217] Cooper, N., Brown, M.E., and C.A. Caulcott (1987). A mathematical method for analysing plasmid stability in micro-organisms. *J. Gen. Microb.* **133**: 1871-1880.

[218] Kumar, P.K.R., Maschke, H-E., Friehs, K., and K. Schugert (1991). Strategies for improving plasmid stability in genetically modified bacteria in bioreactors. *TIBTECH* **9**, 279-284.

[219] Andersson, L., Yang, S., Neubauer, P., and S.-O. Enfors (1996). Impact of plasmid presence and induction on cellular responses in fed batch cultures of Escherichia coli. *J. Biotechnology 46*, 255-263.

[220] Tyo, K.E.J., Ajikumar, P.K., and G. Stephanopoulos (2009). Stabilized gene duplication enables long-term selection-free heterologous pathway expression. *Nature Biotechnology* **27(8):** 760-763.

[221] Canton, B., Labno, A., and D. Endy (2008). Refinement and standardization of synthetic biological parts and devices. *Nature Biotechnology* **26**: 787-793.

[222] Temme, K., Zhao, D., and C.A.Voigt (2012). Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*. *Proc. Natl. Acad. Sci.* **109(18)**: 7085-90.

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

[224] Atkinson, M.R., Pattaramanon, N., and A.J. Ninfa (2002). Governor of the *glnAP2* promoter of *Escherichia coli*. *Molecular Microbiology* **46(5)**: 1247-1257.

[225] BD LSRFortessa Technical Specifications. <http://www.bdbiosciences.com/documents/BD_LSRFortessa_tech_specs2.pdf> Becton Dickinson, 2013.

[226] Liu, H. S., Jan, M. S., Chou, C. K., Chen, P. H., & Ke, N. J. (1999). Is green fluorescent protein toxic to the living cells?. *Biochemical and biophysical research communications*, **260(3)**: 712-717.

[227] Thomas Segall-Shapiro, personal communication.

[228] Moser, F. *Engineering Biosensors for Synthetic Genetic Circuits in Escherichia coli.* (Master's Thesis). University of California, Berkeley. (2011).

[229] Webber, W., Link, N., & M. Fussenegger (2008). A genetic redox sensor for mammalian cells. *Metabolic Engineering* **8**: 273-280.