Improvement of D-Glucaric Acid Production in *Escherichia coli*

by

Eric Chun-Jen Shiue

M.S. Chemical Engineering Practice
Massachusetts Institute of Technology, 2011

B.S. Chemical Engineering
University of California, Berkeley, 2008

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

Doctor of Philosophy in Chemical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2014

Signature of Author .....
Eric Chun-Jen Shiue
Department of Chemical Engineering
December 17, 2013

Certified by .....
Kristala L. Jones Prather
Associate Professor of Chemical Engineering
Thesis Supervisor

Accepted by .....
Patrick S. Doyle
Professor of Chemical Engineering
Chairman, Committee for Graduate Students

Improvement of D-Glucaric Acid Production in *Escherichia coli*

by

Eric Chun-Jen Shiue

Submitted to the Department of Chemical Engineering on December 17, 2013 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Chemical Engineering

Abstract

D-glucaric acid is a naturally occurring compound which has been explored for a plethora of potential uses, including biopolymer production, cancer and diabetes treatment, cholesterol reduction, and as a replacement for polyphosphates in detergents. This molecule was identified in 2004 as a "Top Value-Added Chemical from Biomass" by the U.S. Department of Energy (Werpy and Petersen, 2004), implying that production of D-glucaric acid could be economically feasible if biomass were used as a feedstock. A biosynthetic route to D-glucaric acid from D-glucose has been constructed in *E. coli* by our group (Moon et al., 2009b), and the goal of this thesis has been to improve the economic viability of this biological production route through improvements to pathway productivity and yield.

One part of this thesis involved the application of metabolic engineering strategies towards improving D-glucaric acid productivity. These strategies targeted MIOX, which had been identified previously as the least active pathway enzyme. Directed evolution of MIOX led to the isolation of a 941 bp DNA fragment which increased D-glucaric acid production 65% from a *myo*-inositol feed. Fusion of MIOX to SUMO, a eukaryotic post-translational protein tag, significantly increased soluble expression and stability, resulting in a 75% increase in D-glucaric acid production from a *myo*-inositol feed.

A second part of this thesis attempted to apply synthetic biology strategies towards improving pathway productivity. Manual, delayed expression of MIOX via time-resolved addition of chemical inducers was shown to improve productivity approximately five-fold. However, inducers are generally too costly for use in industrial production processes, so we attempted to develop genetic circuits which could delay MIOX expression autonomously, eliminating the need for costly chemical inducers. Although the attempts to create robust, controllable genetic timers in this thesis were unsuccessful, these attempts provided significant insight into limitations currently preventing widespread application of synthetic biology devices to metabolic engineering problems.

A third part of this thesis explored strain engineering as a strategy for improving the yield of D-glucaric acid on D-glucose. Deletion of *pgi* and *zwf* was demonstrated to prevent *E. coli* from consuming D-glucose as well as eliminate catabolite repression effects in the presence of D-glucose. Finally, both D-glucaric acid productivity and yield were shown to be increased significantly in this Δ*pgi* Δ*zwf* strain.

Overall, this thesis reports significant strides towards commercially viable titers of D-glucaric acid as well as interesting avenues of research for further pathway improvements.

Thesis Supervisor: Kristala L. Jones Prather
Title: Asssociate Professor of Chemical Engineering

In Memory of Tzong-Yann Shiue

Acknowledgments

This thesis would not have been possible without the guidance, support, and encouragement of innumerable people during my time at the Institute. First and foremost, I would like to thank my thesis advisor, Kris Prather, for her constant encouragement and enthusiasm over the years as well as her scientific guidance. I am extremely grateful for the scientific latitude she allowed me, even if it resulted in occasional wandering down rabbit holes. I would also like to thank my thesis committee, Professors Narendra Maheshri, Ron Weiss, and Dane Wittrup for preventing me from becoming too lost in said rabbit holes and for providing helpful advice over the years.

I would like to acknowledge SynBERC not only for funding this work but for providing numerous opportunities for personal growth. I am grateful for the opportunity to serve in leadership roles through the Student and Postdoc Association and to mentor students through the Research Experience for Undergraduates program. I would also like to acknowledge the MIT School of Chemical Engineering Practice for providing me with an industrial perspective on research and development. My internship experiences through the Practice School Program were extremely memorable, and I am thankful to have had the opportunity to participate in such a great program.

To past members of the Prather Lab, thank you for welcoming me as a first year and for the many helpful suggestions you have given (and continue to give) regarding graduate school, research, and life in general. Your advice, support, and friendship has been an invaluable resource. To present members of the Prather Lab, thank you for continuing to make the Prather Lab an enjoyable place to work, for interesting scientific discussions, and for commiserating when things didn't quite work out. Thank you also for not breaking the HPLC too often! To Gwen Wilcox, thank you for keeping the lab running, for keeping all of us in line, and for answering all my questions over the years.

I would also like to thank the many undergraduate students that I have had the opportunity to mentor during my time at MIT: Matt Metlitz, Lauren Allen, Garielle Stiggers, Jane He, Alexis Courtney, Ha-Young Kang, Jessica Wang, and Hannah Dotson for their work on various aspects of this thesis. Thank you for helping me become a better mentor and educator.

To my friends: thank you for for camping, hiking, ice skating, playing cards, cooking, baking, watching movies, making music, and hanging out with me. I have enjoyed getting to know all of you and hope that our friendships will continue past our graduate school careers.

Finally, thank you to my parents, Tzong-Yann and May-Feng Shiue. I would not be the person I am today without the love, support, and enormous independence that you have given me.

ERIC C. SHIUE

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Abstract

D-glucaric acid is a naturally occurring compound which has been explored for a myriad of potential uses, including biopolymer production and cancer treatment. This molecule was identified in 2004 as a "Top Value-Added Chemical from Biomass" by the U.S. Department of Energy (Werpy and Petersen, 2004), implying that production of D-glucaric acid could be economically feasible if biomass were used as a feedstock. A biosynthetic route to D-glucaric acid from D-glucose has been constructed in *E. coli* by our group, and D-glucaric acid titers of 1.13 g/L were obtained in the initial pathway demonstration (Moon et al., 2009b). However, much higher productivity and yield is necessary if biological production of D-glucaric acid is to become economically feasible. This chapter begins with a brief overview of D-glucaric acid and previous work regarding its production in recombinant *E. coli*. This overview is followed by a summary of metabolic engineering and synthetic biology tools which could be applied to improve the productivity and yield of D-glucaric acid in recombinant *E. coli*; difficulties which may arise when using synthetic biology devices as tools for metabolic engineering are also discussed. Finally, the motivation for pursuing this thesis, specific thesis aims, and an outline of the thesis are presented.

Portions of this chapter have been published in:

Shiue, E and Prather, KLJ. Synthetic biology devices as tools for metabolic engineering. Biochemical Engineering Journal. 65:82-89 (2012).

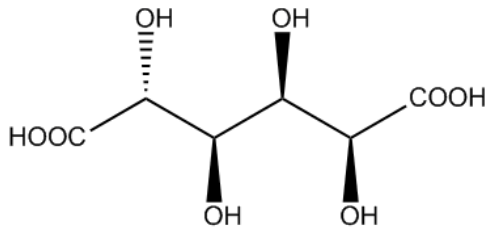


Figure 1.1 | Chemical structure of D-glucaric acid.

1.1 D-Glucaric Acid

D-glucaric acid (Fig. 1.1) is a naturally occurring compound found in fruits, vegetables, and mammals (Perez et al., 2008). This compound has been investigated for a wide variety of therapeutic and commercial uses: as an antitumor agent via inhibition of carcinogen-DNA binding (Gupta and Singh, 2004), as a cholesterol-reducing agent (Walaszek et al., 1996), and as a potential replacement for polyphosphates in detergents (Dijkgraaf et al., 1987). Additionally, in 2004 a U.S. Department of Energy report entitled "Top Value-Added Chemicals from Biomass" identified several potential uses for D-glucaric acid and its derivatives (Fig. 1.2): glucarolactones could function as novel solvents and glucaramides could serve as monomers for biodegradable polymers (Werpy and Petersen, 2004). In fact, production of a hydroxylated nylon from D-glucaramide monomers has already been demonstrated (Kiely et al., 1994). With a plethora of potential uses, D-glucaric acid continues to be a product of commercial interest.

Several engineering problems must be solved before these applications for D-glucaric acid can be realized, however. The availability of a cheap, clean source of D-glucaric acid would certainly facilitate the exploration of glucaric acid and its derivatives as a building block for complex molecules. Current methods for the production of D-glucaric acid involve the chemical oxidation of D-glucose, frequently with nitric acid as the solvent and oxidant. However, nitric oxidation of D-glucose suffers from low yields (approximately 40% of the theoretical value) and requires high temperatures, generating numerous oxidation products from which D-glucaric acid must be separated (Mehltretter and Rist, 1953). Catalysts such as vanadium pentoxide and 4-acetylamino-2,2,6,6-tetramethyl-1-piperidinyloxy have been shown to increase D-glucaric acid yields (Pamuk et al., 2001, Merbouh et al., 2002); however, such catalysts are generally quite expensive. By avoiding costly catalysts and harsh reaction conditions, biological production of D-glucaric acid offers the potential for a cheaper and more environmentally friendly process.

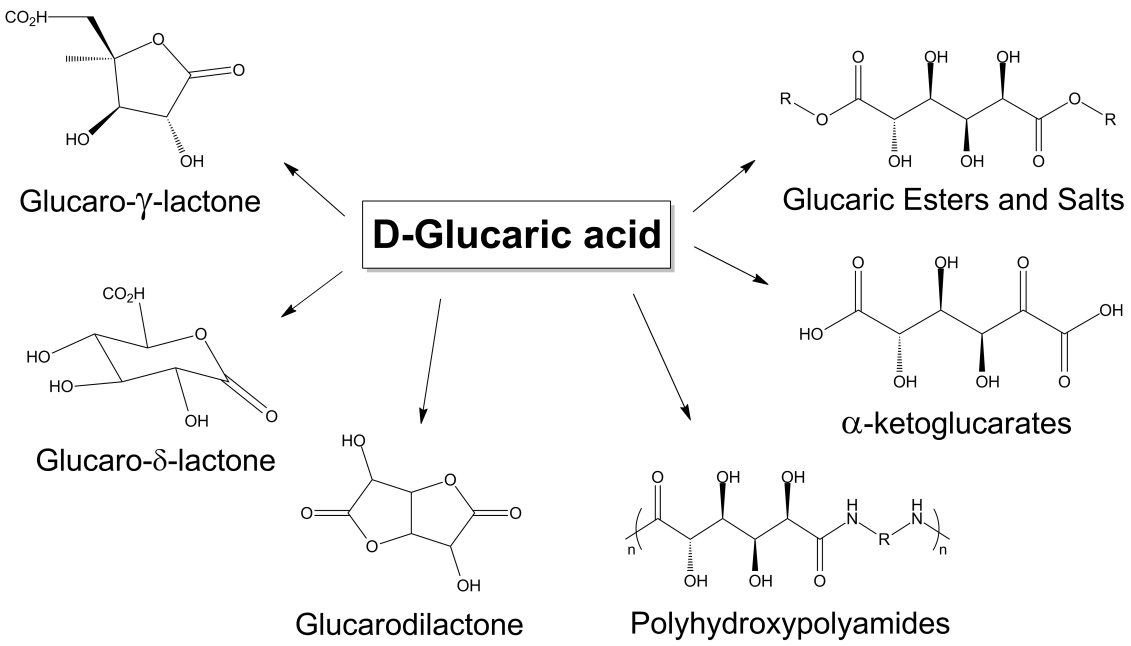


Figure 1.2 | D-glucaric acid derivatives of possible commercial interest (Werpy and Petersen, 2004).

1.2 D-Glucaric Acid Production in Recombinant *E. coli*

A pathway for the biosynthesis of D-glucaric acid from glucose exists in animals, but this pathway is complex, involving numerous steps. To construct a simpler pathway towards D-glucaric acid in *E. coli*, Moon combined the principles of retrosynthesis with databases of enzyme function (e.g., BRENDA, KEGG) to generate several novel routes to D-glucaric acid from glucose (Moon, 2006). One pathway, termed "Pathway 1," was successfully constructed in *E. coli* (Fig. 1.3) (Moon et al., 2009b). This pathway consists of three heterologous enzymes from three different organisms. Briefly, glucose is transported into the cell through the phosphotransferase system (PTS) of *E. coli*, generating glucose 6-phosphate. Glucose 6-phosphate is then isomerized to *myo*-inositol-1-phosphate by *myo*-inositol-1-phosphate synthase (INO1) from *Saccharomyces cerevisiae*. An endogenous phosphatase dephosphorylates *myo*-inositol-1-phosphate to produce *myo*-inositol, which is then oxidized to D-glucuronic acid by *myo*-inositol oxygenase (MIOX) from *Mus musculus* (mouse). D-glucuronic acid is then further oxidized by uronate dehydrogenase (Udh) from *Pseudomonas syringae* to produce D-glucaric acid. Using this pathway, titers of 1.13 g/L D-glucaric acid were obtained from 10 g/L glucose after optimization of culturing and induction conditions.

Analysis of the heterologous pathway enzymes revealed MIOX to be the least active enzyme, with an ac-

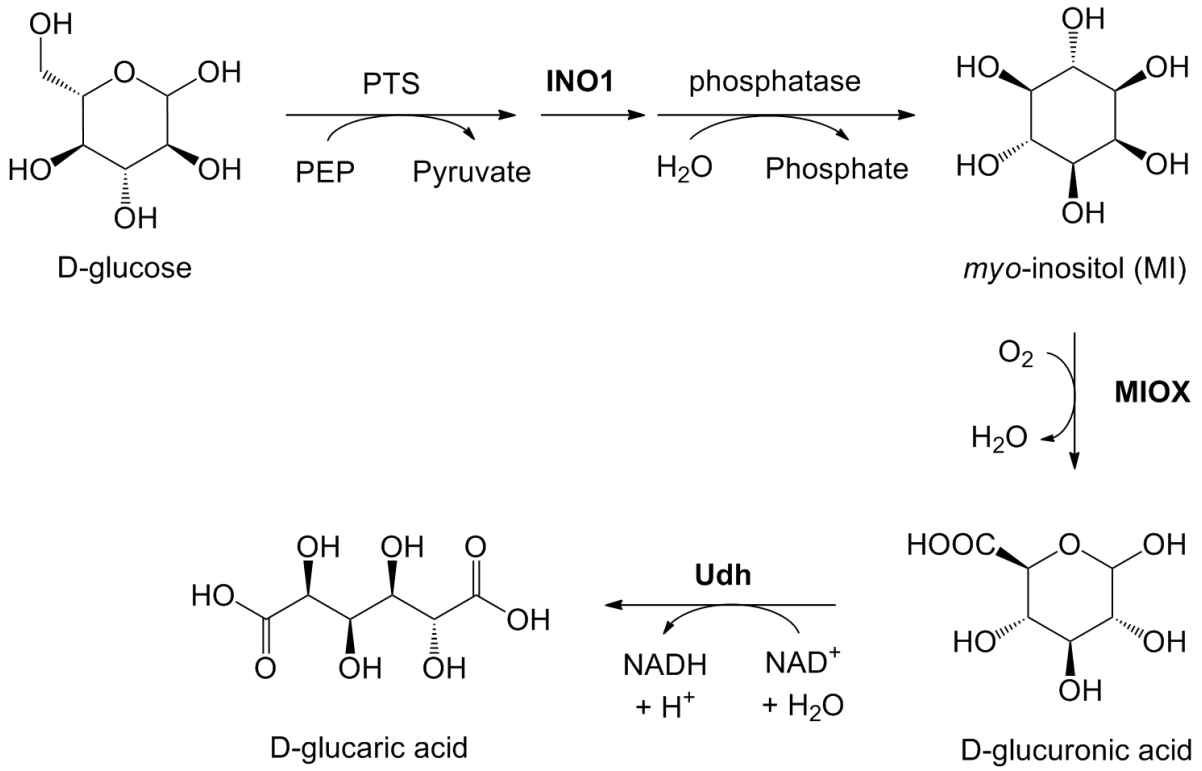


Figure 1.3 | Heterologous pathway for production of D-glucaric acid from D-glucose in *E. coli*.

tivity an order of magnitude lower than INO1 and several orders of magnitude lower than Udh (Moon et al., 2009b). It was also observed that addition of *myo*-inositol to the culture medium significantly increased MIOX activity in the exponential phase; however, this gain in activity was quickly lost in stationary phase, pointing to potential instabilities in MIOX. Previous works have attempted to increase D-glucaric acid production by using protein scaffolds to colocalize the three D-glucaric acid pathway enzymes (Dueber et al., 2009, Moon et al., 2010). Because INO1 actively produces *myo*-inositol from glucose 6-phosphate, the local concentration of *myo*-inositol around the enzyme should be higher than the bulk *myo*-inositol concentration. By colocalizing MIOX in close proximity to INO1 using protein scaffolds, it was hypothesized that the higher local concentration of *myo*-inositol would result in increased activation of MIOX. Indeed, D-glucaric acid titers of approximately 2.5 g/L were achieved with the protein scaffolds (Moon et al., 2009b). However, MIOX activity in this system was low relative to previously measured activities. Further work, including application of metabolic engineering and synthetic biology strategies, will be necessary to further improve D-glucaric acid production in *E. coli*.

1.3 Metabolic Engineering Tools and Synthetic Biology Devices

The advent of of recombinant DNA technology in 1972 (Jackson et al., 1972) opened a new realm of biology to exploration: by moving "parts" (e.g., genes) of one organism into another, scientists could now study individual biological aspects of an organism in isolation. Scientists soon realized that the ability to move genes between organisms offered them the opportunity to produce nonnative proteins in genetically tractable hosts, leading to the production of human insulin in *Escherichia coli* in 1978 (Johnson, 1983). Since this first demonstration of heterologous protein production, a large number of recombinant protein products have been produced, including human growth hormone, interferon (McCormick et al., 1984), factor VIII (Kaufman, 1991), and spider silk (Xia et al., 2010). In addition, heterologous genes have been combined into "pathways" which generate a myriad of nonnative biochemical products, including isoprenoids (Ajikumar et al., 2010, Wang et al., 2011b, Alper et al., 2005b, Yamano et al., 1994), hydroxyacids (Tseng et al., 2009, Tseng et al., 2010, Lee et al., 2008), biofuels (Bond-Watts et al., 2011, Atsumi and Liao, 2008), polyketides (Pfeifer et al., 2001, Piasecki et al., 2011), and biopolymers (Park et al., 2011, Xia et al., 2010). A particularly notable product is 1,3-propanediol, which is microbially synthesized in industry and is widely used as a polymer building block (Kurian, 2005).

1.3.1 Metabolic Engineering Tools

The field of metabolic engineering focuses on improving the productivity of heterologous or natural pathways (Raab et al., 2005, Stephanopoulos, 1999, Khosla and Keasling, 2003). A wide variety of strategies have been developed over the past thirty years to maximize flux through the pathway of interest while minimizing the burden imposed by the expression of heterologous genes. These methods include overexpression of pathway enzymes, deletion of off-pathway reactions which siphon flux away from the desired product, protein engineering for improved enzyme properties, and removal of regulatory mechanisms which may limit pathway flux. This section provides a brief overview of these well-established strategies.

Overexpression of Pathway Enzymes

The first strategy that is nearly always attempted by metabolic engineers is overexpression of pathway enzymes. High copy plasmids are frequently used to express heterologous pathway proteins at levels orders of magnitude higher than native proteins: by increasing the amount of catalyst available for product formation, it is hoped that flux towards the product of interest can be increased as well. This strategy has been successful

for a myriad of cases (Xu et al., 2011, Michnick et al., 1997, Wisselink and Mars, 2004, McKenna and Nielsen, 2011, Lütke-Eversloh and Stephanopoulos, 2008); however, strong overexpression from high copy plasmids often places an enormous metabolic burden on the cell, leading to reduced growth and cell viability. Additionally, overexpression of some proteins, particularly ones which interact with cell metabolism, can be toxic to the cell (Solomon et al., 2013). In many cases, slight overexpression from low copy plasmids is sufficient for metabolic engineering purposes (Jones et al., 2000).

Deletion of Off-Pathway Reactions

Another common strategy for increasing pathway productivity and yield is the deletion of off-pathway reactions which divert flux away from the product of interest. In the case of native product pathways, enzymes often exist within the host organism which convert pathway intermediates or the desired product into unwanted byproducts. By deleting the genes that encode for these enzymes, consumption of the product of interest can be abolished, leading to increased product yields. Additionally, the deletion of off-pathway reactions may also lead to increases in the concentration of pathway intermediates, which may in turn lead to increased pathway productivity. For example, the L-valine biosynthesis pathway competes with the L-isoleucine pathway for pyruvate. To increase productivity of L-valine, Park et al. deleted *ilvA* (L-threonine dehydratase), which produces 2-ketobutyrate from L-threonine, from *E. coli* (Park et al., 2007). Since L-isoleucine biosynthesis involves the condensation of pyruvate with 2-ketobutyrate, Park et al. hypothesized that the deletion of *ilvA* would lead to lower 2-ketobutyrate levels and a slower rate of pyruvate consumption for L-threonine formation, leaving a greater proportion of pyruvate for L-valine synthesis. *panB* and *leuA*, which siphon 2-ketoisovalerate away from the L-valine pathway, were also knocked out (Park et al., 2007). In the case of heterologous pathways, byproduct formation is often less of a concern, since the product of interest and its intermediates are less likely to be actively utilized by the host strain. However, the promiscuous nature of many native enzymes can still result in unwanted byproduct formation.

More recently, the availability of *in silico* genome-scale models of cellular metabolism (Feist and Palsson, 2008) and the development of flux balance analysis (Almaas et al., 2004, Raman and Chandra, 2009) has allowed for the identification of potential knockout targets which appear to be unrelated to the pathway of interest. Because these models capture global metabolism, they are able to predict global cellular responses to genomic perturbations. This method has been used successfully to identify potential knockout targets for overproduction of lycopene, resulting in a 8.5-fold increase in production over the recombinant wild type (Alper

et al., 2005b). Several computational algorithms, including OptKnock and OptStrain, have been developed to perform this type of genome optimization automatically (Burgard et al., 2003, Pharkya et al., 2004).

Protein Engineering

In many heterologous production pathways, one particular enzyme limits flux through the pathway. To improve pathway productivity, these bottleneck enzymes can be engineered for enhanced activity through random or rational means. Directed evolution is a powerful tool for protein engineering via random mutagenesis of the protein coding sequence, and this method has been used extensively and successfully to alter or improve various protein properties (Collins et al., 2005, Hibbert and Dalby, 2005, Hawkins et al., 2007, Atsumi and Liao, 2008). Methods such as error-prone PCR and whole-cell mutagenesis are capable of generating enormous sequence/structure diversity in a very short amount of time (Marcheschi et al., 2013); however, the identification of mutations which improve protein activity is a colossal task. The success of a random directed evolution approach therefore hinges on the availability of a selection or high-throughput screen for the desired protein property.

When a crystal structure of the protein of interest is available or when a homology model can be constructed, a more rational engineering approach is often fruitful. Knowledge about the location of a protein's active site allows specific residues to be selected for mutagenesis, significantly decreasing the sequence/structure space that must be explored. For example, Leonard et al. were able to increase levopimaradiene synthesis in *E. coli* approximately 2,600-fold by using an LPS (levopimaradiene synthase) homology model to identify fifteen key residues for targeted mutagenesis in addition to overexpression of pathway enzymes (Leonard et al., 2010).

Removal of Regulatory Mechanisms

Metabolic pathways in *E. coli* and other organisms are often heavily regulated to conserve precious cellular resources. Feedback inhibition, or inhibition of an upstream pathway enzyme by a downstream metabolite, is a common regulatory motif (Alon, 2006). This inhibition can severely limit pathway productivity, especially when native metabolic pathways are involved in product formation. L-valine, for example, inhibits acetohydroxy acid synthase isoenzyme III (*ilvH*), the first step in the biosynthesis of L-valine from pyruvate. Park et al. mutated *ilvH* to remove the feedback inhibition in order to overproduce L-valine in *E. coli* (Park et al., 2007). Regulatory mechanisms can also be removed by deletion of regulatory proteins or sigma factors: Choi et

al. demonstrated a twofold increase in 1-propanol production in a strain lacking *rpoS*, a stationary phase sigma factor which regulates several stress response genes in *E. coli*.

1.3.2 Synthetic Biology Devices

Synthetic biology aims to facilitate the engineering of biology through the characterization and standardization of a set of biological "parts" (Canton et al., 2008, Endy, 2005). Together with a set of composition and abstraction rules, such a collection of well-characterized parts would enable biological engineers to build "devices" capable of performing complex biological functions such as sensing cell state (Anderson et al., 2007, Kobayashi et al., 2004), counting of cellular events (Friedland et al., 2009, Burrill and Silver, 2011), and implementing computational logic (Anderson et al., 2007, Wang et al., 2011a, Lou et al., 2010, Callura and Dwyer, 2010). Although the metabolic engineering strategies discussed in Sec. 1.3.1 have been successful in improving product titers for a large number of metabolic pathways, more fine-tuned control over cellular processes is often required (Solomon and Prather, 2011). In addition, heterogeneities at the population level, which become magnified at industrial scales, often necessitate control of metabolic processes at the single-cell level. Synthetic biology devices have the potential to provide metabolic engineers with novel ways of exerting cellular-level control over heterologous production pathways. This section presents, in order of complexity, several synthetic biology devices which have potential relevance for the field of metabolic engineering.

Orthogonal Inducible Promoters

To date, the expression of heterologous pathways has relied on a surprisingly small number of inducible promoters, including arabinose-inducible P_{BAD}, lactose and IPTG-inducible P_{lac}, as well as tetracycline-inducible P_{tet}. A set of acyl homoserine lactone-responsive quorum sensing promoters is also available (Fuqua et al., 1994). In many cases, this limited diversity is sufficient for controlling metabolic pathways of interest. However, as metabolic pathways become more complex, individual control over specific elements in the pathway becomes limited by the number of orthogonal promoters available. This problem can be alleviated by using combinatorial promoter libraries to tune expression levels for a given inducer concentration (Cox et al., 2007) or by combining pathway genes into operons and varying the intergenic regions (Pfleger et al., 2006). However, the appropriate expression level for each pathway enzyme is rarely known *a priori*, so these solutions can be more time-consuming than solutions involving orthogonal promoters. In order to

cope with the increasing complexity of engineered metabolic pathways, new, orthogonal inducible promoters are needed. For metabolic engineering purposes, these promoters should respond to stimuli not normally found in cellular metabolism, as the effect of recombinant protein expression on intracellular metabolite levels is often difficult to predict, especially when the recombinant proteins actively siphon carbon flux away from cellular metabolism.

Light-sensitive promoters, which regulate gene expression in response to light of various wavelengths, are one example of orthogonal inducible promoters. These promoters rely on light-oxygen-voltage (LOV) domains (Möglich and Moffat, 2007, Fedorov et al., 2003) or light-sensor modules from phytochromes which are found in a diverse set of bacteria and plants (Yeh et al., 1997, Ni et al., 1999, Tabor et al., 2011). Recently, light-responsive promoters were used to implement *in silico* feedback regulation of a synthetic genetic circuit in *Saccharomyces cerevisiae* (Milias-Argeitis et al., 2011). An advantage of light-responsive promoters is that induction is inexpensive and can be applied transiently, as opposed to the small molecule-responsive promoters described above, where turning off expression requires the removal of inducer. Furthermore, induction can often be reversed using a different wavelength of light.

Temperature-sensitive promoters are another example of orthogonal inducible promoters. Qoronfleh et al. describe temperature-inducible promoters in *E. coli* which induce expression 36-fold during extended growth at 20°C (Qoronfleh et al., 1992). Temperature-inducible promoters have also been identified in mammalian systems: Thaisuchat et al. report a novel CHO promoter that is induced more than 2-fold upon a temperature shift from 37°C to 33°C. In the context of metabolic engineering, however, it is important to recognize that a temperature change will affect the physiological and metabolic state of a cell (Caspeta et al., 2009). The effects of such changes on the metabolic pathway of interest should be considered before employing a temperature-inducible promoter.

The growing number of orthogonal promoters gives metabolic engineers an expanding array of knobs with which to control cellular processes. The continued development of orthogonal promoters is therefore critical to ensuring the applicability of synthetic biology devices to metabolic engineering.

Promoter and Ribosome Binding Site (RBS) Libraries

Traditional metabolic engineering has focused on the overexpression of heterologous pathways in an attempt to outcompete endogenous metabolic pathways. However, protein production is highly resource-intensive, and protein overexpression places an enormous metabolic load on the cell, often resulting in low

growth rates and productivity (Glick, 1995, Neubauer et al., 2003). Physiological changes which may occur upon protein overexpression can lead to unpredictable, nonlinear relationships between expression level and productivity (Klumpp et al., 2009). In many cases, low-level expression is sufficient for productivity (Jones et al., 2000); however, the shape of the productivity landscape varies dramatically between different pathways.

Several synthetic promoter libraries have been generated which can be leveraged to explore the productivity landscape. By varying the spacer sequences between the -35 and -10 regions, Jensen et al. constructed a library of *L. lactis* promoters spanning a 400-fold range of promoter activity (Jensen and Hammer, 1998). Alper et al. created a set of constitutive promoters spanning a 196-fold range of activity via error-prone PCR of the bacteriophage P_L-λ promoter (Alper et al., 2005a). Schalabach et al. took a different approach, varying the sequence of the transcription factor binding site upstream of a CMV promoter (Schlabach et al., 2010). Other efforts have focused on developing predictive models to relate promoter sequence to promoter strength (De Mey et al., 2007). However, due to the large sequence space and relative lack of understanding regarding polymerase-promoter interactions, the development of such predictive models remains a daunting task.

In addition to promoter libraries, RBS libraries have been explored as another avenue for regulating gene expression levels. Salis et al. describe a statistical thermodynamic model which can be used to predict the rate of translation initiation of mRNAs and to design synthetic ribosome binding sites (Salis et al., 2009). It has been shown that mRNA secondary structure, which is determined by the DNA sequence surrounding the RBS, can affect the rate of translation initiation (de Smit and van Duin, 1990). The model developed by Salis et al. has been incorporated into a software program called the RBS Calculator which can be used to design synthetic RBS sequences with translation initiation rates varying over a 100,000-fold range (Salis, 2011).

A combination of these synthetic promoter and RBS libraries would yield an enormous, finely gradated library of expression "devices" for fine-tuning gene expression. In addition, the models that have been developed allow metabolic engineers the opportunity to precisely engineer protein expression levels in heterologous pathways.

State Sensors

The variability in industrial-scale fermentations is often due to incomplete mixing, which creates microenvironments within the fermentation vessel (Bylund et al., 1998). For example, characteristic mixing times in large-scale bioreactors have been shown to vary from 2 to over 3,000 seconds (Lara et al., 2006). While

strategies have been developed to take advantage of incomplete mixing, these strategies rely on controlling macroscopic fermentation parameters (Ganduri et al., 2005, Patnaik, 2002). "State sensors" are genetic devices which allow a cell to sense and respond to particular attributes about its microenvironment, such as temperature, dissolved oxygen, and metabolite concentrations. State sensors may be comprised of inducible promoters, aptamers, riboswitches, and other genetic elements necessary for sensing of cell state. These genetic devices would allow a cell to change its gene expression profile to maximize productivity in the given microenvironment. The state sensors that have been developed are capable of responding over a wide range of time scales from milliseconds (Desai and Gallivan, 2004, Radley et al., 2003) to tens of minutes (Cohen et al., 2001, Surette et al., 1999), making them useful tools for controlling ever-fluctuating industrial bioprocesses.

As oxygen is commonly a limiting nutrient in fermentations, mildly hypoxic conditions are frequently encountered. In *Saccharomyces cerevisiae*, the native *DAN1* promoter has been shown to respond to strictly anaerobic conditions (Cohen et al., 2001). Nevoigt et al. used directed evolution to engineer two *DAN1* promoter variants that respond to mildly hypoxic conditions (Nevoigt et al., 2007). Although this promoter would not be active in prokaryotic organisms, this study demonstrated that directed evolution could be used to modulate the regulatory properties of promoters.

A secondary consequence of oxygen limitation is a shift of redox state within the cell, which is manifested as a shift in the intracellular cofactor balance. Because many heterologous pathways rely on cofactors to push flux towards product, changes in cofactor balance can affect product titers dramatically (Cheol et al., 2010, Chemler et al., 2010). Weber et al. report a redox sensor for mammalian cells consisting of the NADH-responsive REX protein from *Streptomyces coelicolor* fused to the transactivating VP16 domain from *Herpes simplex* and a promoter PROP containing a ROP operator site (Weber et al., 2006). High intracellular levels of NADH prevent REX from binding to ROP, resulting in repression of gene expression (Brekasis and Paget, 2003). Weber and coworkers demonstrate that this redox sensor is able to detect shifts in redox balance due to changes in substrate and oxygen concentrations.

The sensing of intracellular metabolites can also be an advantageous feature, as genes can be upregulated to consume toxic intermediates or downregulated in the absence of precursors. This "just-in-time" expression, in which genes in a metabolic pathway are expressed only when necessary, can be found in a wide variety of bacterial species (Zaslaver et al., 2006, Zaslaver et al., 2004). Sensing of intracellular metabolites is often mediated by mRNA-metabolite interactions, where the binding of a metabolite acts to repress or activate protein expression (Nudler and Mironov, 2004). The development of *in vitro* selection methods in 1990

allowed for the creation of synthetic aptamers that recognize new metabolites (Tuerk and Gold, 1990,Ellington and Szostak, 1990), and the development of new screening methods allowed for the creation of aptamers that modulate gene expression, called riboswitches (Desai and Gallivan, 2004, Topp and Gallivan, 2010, Carothers et al., 2011). While synthetic aptamers are capable of recognizing several classes of molecules, the applicability of riboswitches to current compounds of interest, such as hydroxyacids and biofuels, remains to be seen.

Promoters which are induced in response to intracellular metabolites are also capable of acting as state sensors. The *glnAP2* promoter in *E. coli* was used by Farmer and Liao as a state sensor for intracellular concentrations of acetyl phosphate (Farmer and Liao, 2000). A high level of acetyl phosphate, which is indicative of excess glycolytic flux, induces expression from the *glnAP2* promoter. By using this sensor to control expression of two genes in the lycopene biosynthesis pathway, Farmer et al. achieved increased lycopene titers by balancing metabolic flux and minimizing the accumulation of toxic byproducts. More recently, Dahl et al. used whole-genome transcript arrays to identify native *E. coli* promoters which respond to the presence of farnesyl pyrophosphate, a toxic intermediate in the pathway for isoprenoid biosynthesis. These promoters were then used as state sensors to dynamically regulate pathway genes and minimize accumulation of farnesyl pyrophosphate inside the cell (Dahl et al., 2013).

Intracellular metabolite sensing can also be achieved via engineered allostery. Radley et al. designed a two-domain protein fusion in which the folding of one domain prevents folding of the other domain, a phenomenon termed "mutually exclusive folding" (Radley et al., 2003). By tying the folding of one domain to ligand binding, an allosteric switch can be created, destroying or enhancing the activity of an enzyme in response to the concentration of a particular metabolite. Because such events occur at the post-translational level, allosteric switches are capable of quickly responding to changes in metabolite concentrations.

In addition to creating substrate and oxygen gradients, incomplete mixing in large-scale fermentations can also lead to gradients in cell density. In the absence of sensors for metabolite and oxygen concentrations, local cell density can be used as an indicator of each cell's microenvironment. The ability to sense cell density, known as "quorum sensing," is present in several species of bacteria (Fuqua et al., 1994, Surette et al., 1999, Uroz et al., 2009, Dunlap, 1999, Pesci et al., 1997). Quorum sensing relies on the synthesis of a signaling molecule which freely diffuses out of the cell and into the extracellular medium. Buildup of the signaling molecule occurs as cell density increases until the concentration of the signaling molecule is sufficiently large to effect binding of a transcriptional activator to its cognate promoter. A positive feedback loop ensures

a population-wide response once the specified cell density is reached. Quorum sensing circuits have been shown to reliably induce gene expression at a given cell density (Kobayashi et al., 2004).

It has recently been suggested that enzymatic and transcriptional regulation of central metabolism in *E. coli* can lead to systems-level responses (Kotte et al., 2010). Coupled with an understanding of the intricate relationship between various metabolic fluxes, the variety of state sensors being developed affords metabolic engineers with the opportunity to implement carefully crafted control strategies which leverage many aspects of cellular metabolism to maximize product titers.

Spatiotemporal Controllers

Spatiotemporal control of cellular processes can be a powerful tool for metabolic engineering. Gene expression can be repressed via spatial sequestration of transcription factors from DNA, the frequency of protein-protein interactions can be enhanced through colocalization, and "just-in-time" gene expression can minimize the metabolic burden that is placed on a cell. Synthetic spatiotemporal controllers provide metabolic engineers with yet another set of tools for manipulating gene expression and flux through metabolic pathways.

Compartmentalization is frequently used in eukaryotic systems to isolate sensitive cellular processes from the rest of the cell environment. For example, disulfide bond formation, which requires an oxidizing environment, is localized to the endoplasmic reticulum in eukaryotes (Tu and Weissman, 2004). Levskaya et al. report a device which translocates proteins to the plasma membrane in response to light (Fig. 1.4A) (Levskaya et al., 2009). The device consists of a phytochrome (PhyB) and transcription factor (PIF), which bind under red light and dissociate under far-red light. By tethering PhyB to the plasma membrane, Levskaya and coworkers are able to recruit any protein fused to PIF to the membrane in a reversible, light-dependent manner.

Many natural systems use enzyme colocalization to enhance flux through metabolic pathways. For example, polyketide synthases consist of long chains of catalytic modules through which intermediates are channeled (Khosla et al., 2009). Such substrate channeling serves to minimize loss of intermediates through diffusion, sequester unstable and/or toxic intermediates, as well as enhance reaction rates by increasing local substrate concentrations. Inspired by natural systems, Dueber et al. incorporated three different metazoan protein interaction domains (SH3, PDZ, and GBD) into a synthetic scaffold (Dueber et al., 2009). By attaching a peptide ligand specific for a particular interaction domain to a protein, the protein can be directed to a partic-

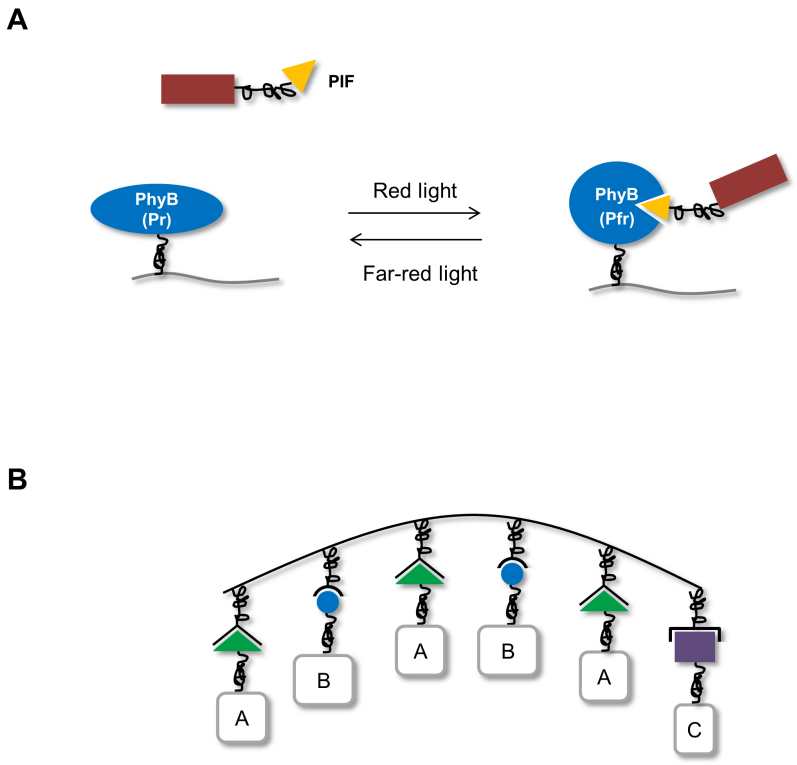


Figure 1.4 | Spatiotemporal controllers. **(A)** Light-responsive localization. The phytochrome PhyB undergoes a conformational change when exposed to red light, changing from the Pr form to the Pfr form. Phytochrome interaction factor (PIF) only binds to the Pfr form of PhyB. By fusing PhyB to the plasma membrane, proteins fused to PIF can be recruited in a light-dependent manner. **(B)** Colocalization using a protein scaffold. Three different protein interaction domains (triangle, circle, and rectangle) allow three different proteins (A, B, and C) to be colocalized, increasing pathway flux through substrate channeling.

ular location on the scaffold (Fig. 1.4B). The use of three orthogonal protein interaction domains allows three different proteins to be colocalized. By varying the arrangement and number of protein interaction domains on the scaffold, the stoichiometry and geometry of the system can be optimized for the metabolic pathway of interest. Using this scaffold device, Dueber et al. were able to achieve a 77-fold improvement in product titers. By applying the scaffold technology to the D-glucaric acid pathway, Moon et al. increased product titers 5-fold over a non-scaffolded control (Moon et al., 2010).

Temporal control of natural metabolic pathways has been observed and has been postulated to help the cell achieve a particular production goal while minimizing enzyme production (Zaslaver et al., 2004). In fact,

the genes in many metabolic pathways are arranged into operons in the order in which they are needed (Zaslaver et al., 2006). Synthetic genetic timers would allow metabolic engineers to mimic this regulatory scheme; however, there has been little work in this area. Riboswitches containing aptamers which bind metabolic intermediates could be used to delay expression of downstream genes until needed. If the development of metabolite-specific riboswitches proves difficult, quorum sensing circuits can be used to delay gene expression until a certain cell density is reached.

Logic Gates

The complexity and applicability of synthetic biology devices can be increased significantly by the introduction of computational logic. For example, genetic inverters (NOT gates) could find application in situations where a reduction of gene expression is desired. Traditional inducible systems exhibit a positive, monotonic relationship between expression level and inducer concentration: when more inducer is added, protein expression level increases until the system becomes saturated. To reduce gene expression, the inducer must be removed from the system, often a nontrivial task. In contrast, an inverter produces high gene expression when no inducer is present, and the addition of inducer results in a decrease in gene expression. Inverters typically consist of a repressor element expressed from an inducible promoter, with the gene of interest under the control of a repressible promoter (Fig. 1.5A). Inverters have been incorporated into a number of synthetic biology devices (Kobayashi et al., 2004, Wang et al., 2011a, Song et al., 2011).

AND gates form the cornerstone of digital computation and could find wide use in metabolic engineering, especially when multiple inputs (e.g., from multiple state sensors) need to be processed. Generally, biological AND gates rely on complementarity of function between two biological parts: for example, the AND gate constructed by Anderson et al. contains a T7 RNA polymerase with two amber stop codons (Anderson et al., 2007). While the gene is easily transcribed, it cannot be translated without the amber suppressor tRNA *supD* (Fig. 1.5B). Wang et al. employ a similar approach, expressing the co-activating genes *hrpR* and *hrpS* from separately inducible promoters (Wang et al., 2011a). Alternatively, operator sites for different regulatory proteins can be combined into a single promoter (Fig. 1.5C); binding or dissociation of both regulatory proteins is therefore required for gene expression. Cox et al. used a combinatorial approach to develop a library of AND gates from four different operator sites (Cox et al., 2007). Other groups have worked to develop models which facilitate the forward engineering of AND gates (Ramalingam et al., 2009).

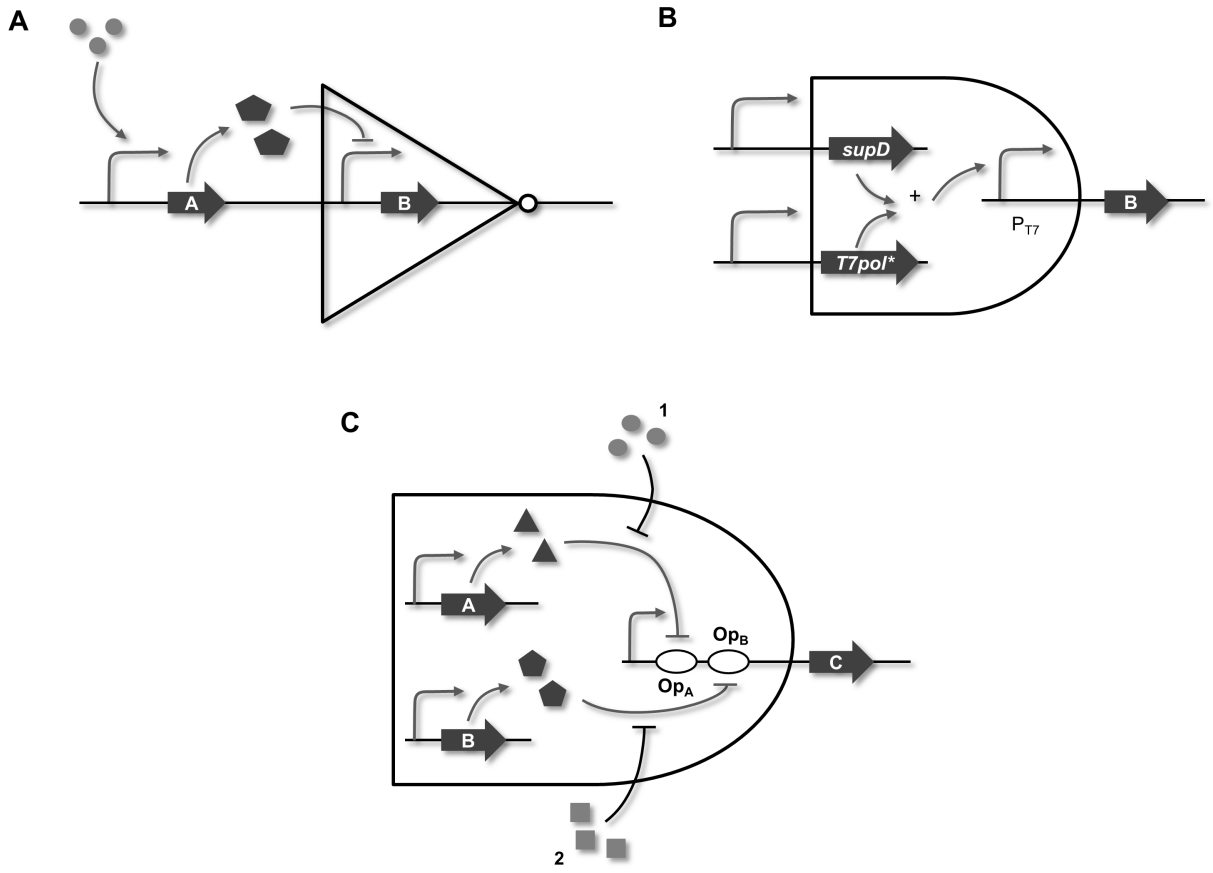


Figure 1.5 | Logic gates. **(A)** Inverter. Expression of repressor protein A is activated by addition of a small molecule inducer (circles), resulting in a decrease in expression of protein B. **(B)** AND gate constructed by Anderson et al. (Anderson et al., 2007). Because the T7 RNA polymerase gene *T7pol* contains two amber stop codons, the amber suppressor SupD must be present for translation of the RNA polymerase to occur. Successful translation of T7 RNA polymerase allows for expression of protein B, which is under the control of a T7 promoter (P_{T7}). **(C)** AND gate constructed by combining operator sequences (Cox et al., 2007). Addition of inducer 1 (circles) causes repressor A to dissociate from its operator site Op_A, while addition of inducer 2 (squares) causes repressor B to dissociate from its operator site Op_B. Both inducers must be added to the system in order for protein C to be expressed.

1.3.3 Discussion

While useful in isolation, the synthetic biology devices described above have the potential to become even more useful when combined to perform complex sensing and actuation tasks at the single-cell level. In industrial processes, where process parameters can vary simply due to incomplete mixing, the cellular-level control which synthetic biology devices are capable of can be a valuable tool for optimizing productivity. However, the integration of multiple devices into a single cell remains a difficult task. Several barriers to the integration of synthetic biology devices into industrial processes are discussed below.

Impedance Matching

In electrical engineering, it has long been known that the addition of a load to a circuit affects the electrical characteristics of the circuit. This phenomenon is due to the internal resistance of the electrical generator: as the power drawn by the load increases, the power lost in the generator increases as well. The concept of impedance matching was developed to address this problem: by matching the output impedance of the generator to the input impedance of the load, the amount of power that is transferred to the load is maximized.

As with electronic circuits, the addition of a biological device to an existing circuit can strongly influence circuit behavior. Therefore, to ensure functionality across all conditions, some amount of "impedance matching" is required: the output range of one device must be matched to the input range of its downstream device. Biological or metabolic fluxes are not perfectly analogous to energy fluxes; therefore, this matching is seldom a trivial task, usually requiring re-tuning of the expression level of each component within a device. For example, directed evolution was employed to modulate the input/output relationships of devices within a non-functional circuit, turning it into a functional one (Yokobayashi et al., 2002). While model-based device construction methods have been successful in producing functional, forward-engineered devices, these methods require fully characterized device components (Wang et al., 2011a, Salis et al., 2009, Wu et al., 2011, Ellis et al., 2009). These methods are also unable to predict secondary effects that arise when multiple devices are combined. The applicability of model-based methods to larger, more complex systems, where uncharacterized components may be involved, however, is unknown.

Context Dependence

The transfer of a production pathway to industry typically involves a change in host strain and medium composition, commonly called a change in context. Unfortunately, even small changes in host strain and/or medium composition can have drastic effects on circuit behavior. Wang et al. observed a twofold difference in maximal protein expression from an inducible promoter system when the device was moved from *E. coli* MC1061 to MC4100 (Wang et al., 2011a). Additionally, a switch from a glucose feed to a glycerol feed resulted in a dramatic shift in the system's induction curve. Due to the incredibly complex nature of biology, the differences that result from a change in context are difficult to predict *a priori*: modifications in host strain and/or medium composition can have far-reaching and often unpredictable effects on the cell's physiology and metabolome. To fully address this problem, a more fundamental understanding of the relationship between context and cell behavior is needed.

The optimal input/output relationship for a particular device may be strongly context-dependent, as circuit behavior may change unpredictably each time a new device is added to the system (Del Vecchio et al., 2008). If these changes in context affect any of the interactions between device components, a complete re-tuning of the system may be required. One solution to this problem is to build circuits that are robust to parameter changes. Another solution is to build and characterize large libraries of circuits that can be swapped with the existing circuit for re-optimization. With standardized assembly methods (Knight, 2003), swapping circuits is a trivial task; however, the screening of large circuit libraries remains a challenge. In many cases, high-throughput screens for increased metabolite production do not exist.

Orthogonality

Although a multitude of synthetic biology devices are available, the incorporation of multiple devices into a single system can be a daunting task due to the lack of insularity: in contrast to electronic circuits, insulated wires do not exist in biological systems to connect devices to one another. In addition, the relative similarity of biological components can lead to undesired crosstalk between devices and device components. One way to alleviate this problem is to spatially separate devices within a cell; however, the transmission of signals between devices remains a problem. Another solution is to increase the number of parts available: with more parts, the chance of crosstalk decreases. Efforts have been made to develop orthogonal gene circuits (Wang et al., 2011a, An and Chin, 2009) as well as devices capable of handling multiple inputs and outputs (Pasotti

et al., 2011). Nevertheless, the tolerable amount of crosstalk for an integrated set of devices remains to be seen.

Scalability

The synthetic biology devices described in this section have all been built, tested, and shown to function at the bench scale. One key question that remains, however, is whether these devices can function at the industrial scale. The enormous fermentation volumes used in industry necessitate large vessels, which presents a challenge for light-inducible devices: because light only penetrates a short distance into the vessel, uniform and synchronized induction of the entire population is impossible. Such large vessels also suffer from incomplete mixing, which presents a challenge for quorum sensing devices: because quorum sensing relies on the extracellular diffusion of signaling molecules to regulate gene expression, the local concentration gradients that form due to incomplete mixing may preclude synchronized behavior across the population.

On the other hand, many of these synthetic biology devices provide cells with the ability to sense their microenvironment and respond accordingly, an ability which is generally absent in traditional bioprocesses. These state sensors are capable of actuating responses over a range of time scales, from the transcriptional (tens of minutes) (Cohen et al., 2001, Surette et al., 1999) to the post-translational (milliseconds to seconds) levels (Desai and Gallivan, 2004, Radley et al., 2003). This capability promises to be a boon to industrial bioprocessing, allowing engineers to convert spatial and temporal inhomogeneities into tighter and more efficient control of metabolic pathways.

1.3.4 Conclusions

Metabolic engineering revolves around controlling cellular metabolism and flux through heterologous pathways to maximize production of a desired molecule. By providing elegant methods for gathering information about cells and their environment, processing that information, and modulating gene expression in response, devices from synthetic biology promise to be a useful addition to the metabolic engineer's toolbox. In fact, some devices have already been used to increase product titers (Dueber et al., 2009, Moon et al., 2010). However, synthetic biology devices remain largely untested in industrial settings, and the complexity of biology makes each application of a device a significant feat of engineering. Nevertheless, an awareness of the potential pitfalls and challenges that may occur upon application at large scale is critical for translation towards industrial use. With such a perspective, continued work with these devices may help elucidate design rules or

aid in the development of models that will facilitate the integration of these devices into industrial processes.

1.4 Thesis Objectives

Given the commercial potential for D-glucaric acid, the goal of this thesis is to improve the economic viability of biological production of D-glucaric acid by improving process productivity and yield. Because MIOX has been identified previously as the least active pathway enzyme, this thesis seeks to improve D-glucaric acid productivity via metabolic engineering and synthetic biology strategies targeted towards increases in MIOX activity. Additionally, this thesis attempts to improve the yield of D-glucaric acid on D-glucose via strain engineering to further improve the commercial prospects of D-glucaric acid.

More specifically, this thesis aims to:

- Improve D-glucaric acid pathway productivity via engineering of MIOX
- Explore synthetic biology devices for time-dependent control of gene expression and their application to D-glucaric acid production
- Engineer *E. coli* for improved yield of D-glucaric acid on D-glucose

1.5 Thesis Organization

This thesis is organized into five chapters. Chapter 1 provides an overview of metabolic engineering strategies and synthetic biology devices and discusses their application to industrial processes. Chapter 1 also provides background information on D-glucaric acid and its production in recombinant *E. coli*. Work regarding each of the specific thesis aims is then discussed in separate chapters: Chapter 2 describes protein engineering of MIOX for improved activity, Chapter 3 discusses efforts to develop synthetic biology devices for delayed gene expression, and Chapter 4 outlines efforts to engineer *E. coli* for improved D-glucaric acid yield on D-glucose. Finally, Chapter 5 provides conclusions and recommendations for future work.

Chapter 2

Application of Metabolic Engineering Strategies

Abstract

D-glucaric acid is a naturally occurring compound which has been explored for a myriad of potential uses, including biopolymer production and cancer treatment. Several additional uses for D-glucaric acid were identified in 2004, when the U.S. Department of Energy named the compound as a "top value-added chemical from biomass" (Werpy and Petersen, 2004). A biosynthetic route to produce D-glucaric acid from glucose in *E. coli* has been constructed previously in our group. This pathway consists of three heterologous enzymes from three different sources, and D-glucaric acid titers of 1.13 g/L from a 10 g/L glucose feed were obtained in the initial pathway demonstration (Moon et al., 2009b). Analysis of the pathway revealed the second heterologous enzyme in this pathway, *myo*-inositol oxygenase (MIOX), which catalyzes the conversion of *myo*-inositol to D-glucuronic acid, to be the least active enzyme. Soluble expression of MIOX was also observed to be much lower than that of the other two heterologous enzymes. To increase pathway productivity, we pursued two parallel strategies in this work to increase MIOX solubility and activity: protein fusion tags for increased solubility and directed evolution for increased activity. An N-terminal SUMO fusion to MIOX resulted in significantly increased soluble expression of MIOX and a corresponding 75% increase in D-glucaric acid production from *myo*-inositol. While our directed evolution efforts did not yield an improved MIOX variant, our screen isolated a 941 bp DNA fragment whose expression led to a 65% increase in D-glucaric acid production from *myo*-inositol. Overall, we report the production of up to 4.85 g/L (23 mM) of D-glucaric acid from 10.8 g/L (60 mM) *myo*-inositol in recombinant *E. coli*. In addition, we report the first evidence of *myo*-inositol transport by PtsG in *E. coli* as well as evidence of a pH-related limitation on D-glucaric acid productivity.

Portions of this chapter have been published in:

Shiue, E and Prather, KLJ. Improving D-glucaric acid production from *myo*-inositol in *E. coli* by increasing MIOX stability and *myo*-inositol transport. Metabolic Engineering. 2013. (doi: 10.1016/j.ymben.2013.12.002)

2.1 Introduction

D-glucaric acid, a compound which occurs naturally in fruits, vegetables, and mammals, has been investigated for a wide variety of therapeutic and commercial uses, including cholesterol reduction (Walaszek et al., 1996), diabetes treatment (Bhattacharya et al., 2013), and cancer therapy (Gupta and Singh, 2004). D-glucaric acid has also been explored as a replacement for polyphosphates commonly found in detergents, which can be damaging to the environment (Dijkgraaf et al., 1987). Additional uses for D-glucaric acid and its derivatives were identified in the U.S. Department of Energy's "Top Value-Added Chemicals from Biomass" list published in 2004: glucarolactones could function as novel solvents and glucaramides could serve as monomers for biodegradable polymers (Werpy and Petersen, 2004). In fact, production of a hydroxylated nylon from D-glucaramide monomers has already been demonstrated (Kiely et al., 1994). With a plethora of potential uses, D-glucaric acid continues to be a product of commercial interest.

Current methods for the production of D-glucaric acid involve the chemical oxidation of D-glucose, frequently with nitric acid as the solvent and oxidant. However, nitric acid oxidation of D-glucose suffers from low yields (approximately 40%) and requires high temperatures, generating numerous oxidation products from which D-glucaric acid must then be separated (Mehltretter and Rist, 1953). Catalysts such as vanadium pentoxide and 4-acetylamino-2,2,6,6-tetramethyl-1-piperidinyloxy have been shown to increase D-glucaric acid yields (Pamuk et al., 2001, Merbouh et al., 2002); however, such catalysts are generally quite expensive. By avoiding costly catalysts and harsh reaction conditions, biological production of D-glucaric acid offers the potential for a cheaper and more environmentally friendly process.

A biosynthetic route to D-glucaric acid from D-glucose consisting of three heterologous genes has been constructed in recombinant *E. coli* by our group (Fig. 1.3) (Moon et al., 2009b). Briefly, D-glucose is imported into E. coli through the native phosphotransferase system (PTS), generating glucose 6-phosphate. Glucose 6-phosphate is then isomerized to *myo*-inositol-1-phosphate by *myo*-inositol-1-phosphate synthase (INO1) from *Saccharomyces cerevisiae*. An endogenous phosphatase dephosphorylates *myo*-inositol-1-phosphate to produce *myo*-inositol, which is then oxidized to D-glucuronic acid by *myo*-inositol oxygenase (MIOX) from *Mus musculus* (mouse). D-glucuronic acid is then further oxidized by uronate dehydrogenase (Udh) from *Pseudomonas syringae* to produce D-glucaric acid. Using this pathway, D-glucaric acid titers of 1.13 g/L were achieved from 10 g/L glucose after optimization of induction and culture conditions. Analysis of the pathway revealed MIOX to be the least active enzyme, with a specific activity an order of magnitude lower than INO1

and several orders of magnitude lower than Udh (Moon et al., 2009b). Further experiments demonstrated a linear, positive correlation between final D-glucaric acid titers and MIOX specific activity, and strategies targeted towards increasing MIOX specific activity were able to improve titers up to 2.5 g/L (Moon et al., 2010).

Numerous metabolic engineering strategies are typically employed to increase pathway productivity, including overexpression of pathway enzymes (Lütke-Eversloh and Stephanopoulos, 2008), enzyme colocalization via synthetic protein scaffolds (Dueber et al., 2009), and deletion of off-pathway reactions (Alper et al., 2005b). Given the correlation between final D-glucaric acid titers and MIOX specific activity, we pursued two parallel, protein-oriented strategies to improve D-glucaric acid production: protein fusion tags and directed evolution. Protein fusion tags have been used extensively to increase soluble expression and to aid proper folding of recombinant proteins (Young et al., 2012a). The end goal of the vast majority of these applications, however, is a purified protein product which can then be studied *in vitro*. Additionally, because the effects of a fusion tag on a target protein's structure and function are rarely predictable, fusion tags are often cleaved from the protein of interest to reconstitute the protein's native sequence and structure. Furthermore, the fusion tag which results in optimal soluble expression for a particular protein is rarely known *a priori*. The unpredictable effects of protein fusion tags have therefore limited their usage in metabolic engineering contexts, where proteins must function even while fused: *in vivo* cleavage of the fusion tag would require expression of a potentially promiscuous protease, and freshly cleaved proteins may simply aggregate into insoluble and inactive inclusion bodies. Although increased product formation due to increased enzyme solubility has been reported, these gains were not due to protein fusion tags (Zhou et al., 2012). In contrast, directed evolution has been used successfully for a number of metabolic engineering problems (Marcheschi et al., 2013).

Previous works have sought to improve D-glucaric acid production from D-glucose (Moon et al., 2010, Dueber et al., 2009). In this work, we focused on improving D-glucaric acid production from *myo*-inositol rather than D-glucose, a decision based on the observation that addition of *myo*-inositol to the culture medium significantly increased MIOX specific activity during exponential phase (Moon et al., 2009b). Additionally, the production of large quantities (approximately 20 g/L) of *myo*-inositol from D-glucose in bacterial cultures expressing INO1 has been reported previously (Hansen et al., 1999). We hypothesized that by feeding *myo*-inositol rather than D-glucose, we could leverage the previously observed *myo*-inositol "activation" effect to drive flux towards D-glucaric acid more effectively and explore the ability to increase productivity beyond previously achieved levels.

2.2 Materials and Methods

2.2.1 *E. coli* Strains and Plasmids

E. coli strains and plasmids used in this study are listed in Table 2.1, and oligonucleotides used are listed in Table 2.2. All molecular biology manipulations were performed according to standard practices (Sambrook and Russell, 2001). *E. coli* DH10B was used for transformation of cloning reactions and propagation of all plasmids. MG1655 was obtained from the Coli Genetic Stock Center (CGSC #6300). Knockout of *recA* was achieved with λ-Red mediated recombination using pRedET (Gene Bridges, Heidelberg, Germany). PCR primers KOrecA-F and KOrecA-R (Table 2.2) were used to amplify the recombination cassette from pKD13 (Datsenko and Wanner, 2000), and MG1655 was transformed with this PCR product. The *kan* selection cassette was cured from successful deletion mutants using FLP recombinase expressed from pCP20. Knockout of *endA* was performed using the λ-Red recombination system contained in pKD46 (Datsenko and Wanner, 2000) and primers KOendA-F/KOendA-R. The λDE3 lysogen was integrated site-specifically into this double knockout strain using a λDE3 Lysogenization Kit (Novagen, Darmstadt, Germany), generating strain M2 (MG1655(DE3) Δ*endA* Δ*recA*). To eliminate the native ability of *E. coli* MG1655 to consume D-glucuronic and D-glucaric acids, knockouts of *gudD* and *uxaC* were performed using the λ-Red recombination system contained in pKD46 (Datsenko and Wanner, 2000) with primers KOgudD-F/KOgudD-R and KOuxaC-F/KOuxaC-R, respectively, generating strain M2-2 (MG1655(DE3) Δ*endA* Δ*recA* Δ*gudD* Δ*uxaC*).

Strain M2-2 or its derivatives were used for production experiments. Deletion of *ptsG* from M2-2 was achieved by P1 transduction with Keio collection strain JW1087-2 as the donor (Baba et al., 2006). Deletion of *sgrS* from M2-2 was achieved with λ-Red mediated recombination (Datsenko and Wanner, 2000) using pKD46recA (Solomon et al., 2013). PCR primers pKD13_sgrS_fwd and pKD_sgrS_rev (Table 2.2) were used to amplify the recombination cassette from pKD13 (Datsenko and Wanner, 2000), and strain M2-2 harboring pKD46recA was transformed with this PCR product. The *kan* selection cassette was cured from successful deletion mutants using FLP recombinase expressed from pCP20.

pWW308 was kindly provided by Prof. John Dueber (University of California, Berkeley). To remove a SpeI restriction site contained in tetR, the tetR-P_{tet} cassette was amplified in two fragments with primers BBa_tet_fwd1/tetR_B and tetR_A/BBa_tet_rev1. These two fragments were then combined into one larger fragment via Splicing by Overlap Extension PCR (SOEing PCR) using primers BBa_tet_fwd1 and BBa_tet_rev1

(Horton et al., 2013). The resulting fragment was digested with EcoRI and SpeI and inserted into the EcoRI and SpeI sites of pSB1A7 to generate pSB1A7-tetR1.

The plasmid pRSFD-MI was constructed by digesting pJ2-MIOX with EcoRI and HindIII and inserting the MIOX-containing fragment between the EcoRI and HindIII sites of pRSFDuet-1. To create pSB1A7-tetR1-MIOX, MIOX was first amplified from pRSFD-MI in two fragments with primers pRSFD_fwd1/MIOX_B and MIOX_A/BBa_MIOX_rev2 to remove an internal PstI restriction site. These two fragments were combined into a single fragment via SOEing PCR using primers pRSFD_fwd1 and BBa_MIOX_rev2. The resulting fragment was digested with XbaI and PstI and then cloned into the SpeI and PstI sites of pSB1A7-tetR1. To facilitate future cloning, an EcoRI restriction site directly 5′ to the MIOX coding sequence was removed via insertion of a single base pair using SOEing PCR. Fragments surrounding the mutation site were first generated using primers BBa_tet_fwd1/pSB_B1 and BBa_MIOX_rev3/pSB_A1, then spliced into a single fragment using primers BBa_tet_fwd1 and BBa_MIOX_rev3. The resulting fragment was digested with EcoRI and SpeI and cloned into the EcoRI and SpeI sites of pSB1A7 to generate pSB1A7-tetR1-MIOX1. In addition to removing the internal EcoRI restriction site, the single base pair mutation also shifted the MIOX coding sequence in-frame with the ribosome binding site.

pTrc-MIOX-1 was constructed by amplifying pSB1A7-tetR1-MIOX1 with BBa_MIOX_fwd and VR and cloning upstream of the double terminator (Part BBa_B0015, Registry of Standard Biological Parts) contained in pSB1AK3-B0015 using BioBricks Assembly Standard RFC 10 (Knight, 2003) to generate pSB1AK3-MIOX-B0015. This construct was then inserted downstream of a strong RBS (Part BBa_B0034, Registry of Standard Biological Parts) contained in pSB1A2-B0034 to generate pSB1A2-B0034-MIOX-B0015. Finally, pSB1A2-B0034-MIOX-B0015 was digested with EcoRI and PstI and the resulting fragment inserted into the EcoRI and PstI sites of pTrc99A to generate pTrc-MIOX-1. pE-SUMO-MIOX was generated by amplifying MIOX from pTrc-MIOX-1 using primers MI_SUMO_fwd and MI_SUMO_rev, then digesting with BsaI and ligating into BsaI-digested pE-SUMO. pRSFD-SUMO-MIOX was then constructed by digesting pE-SUMO-MIOX with NcoI and HindIII and ligating the resulting fragment into similarly digested pRSFDuet-1. pTrc-SUMO-MIOX was also constructed by digesting pE-SUMO-MIOX with NcoI and HindIII, then ligating the resulting fragment into similarly digested pTrc99A.

pTrc-MIOX-DE was constructed and isolated using the directed evolution procedure described in Sec. 2.2.4. pTrc-MIOX-DE-noins was produced by digesting pTrc-MIOX-DE with PstI, separating the DNA fragments on a 1% agarose gel to remove the 941 bp band, and self-ligating the 5.2 kb fragment. pTrc-MIOX-insert and

pTrc-insert were created by inserting the 941 bp DNA fragment isolated from the directed evolution screen into the PstI sites of pTrc-MIOX-1 and pTrc99A, respectively. Correct orientation of the fragment was verified by sequencing.

pTrc-ins-manX and pTrc-ins-yoaE were constructed via PCR amplification of pTrc-insert with primers pTrc_fwd/manX_rev and yoaE_fwd/pTrc_rev, respectively; the ins-manX PCR product was cloned into pTrc99A following digestion with EcoRI, and the ins-yoaE PCR product was cloned into pTrc99A following digestion with XbaI and PstI. Primers manX_frag_fwd and manX_frag_rev were annealed, digested with EcoRI and PstI, and cloned into pTrc99A to generate pTrc-manX-frag. Full-length manX was amplified from *E. coli* genomic DNA using primers manX_A and manX_B, digested with EcoRI and XbaI, and cloned into pTrc99A to yield pTrc-manX-full. pTrc-manX-mut was generated via site-directed mutagenesis of pTrc-ins-manX using primers manX_QC_fwd and manX_QC_rev. Figure 2.1 provides schematics for the plasmids described in this paragraph.

pTrc-Udh-insert was created by digesting pTrc-ins-manX with SmaI and inserting the resulting ins-manX fragment into the SmaI site in pTrc-Udh. Correct orientation of the ins-manX fragment was verified by sequencing. pRSFD-Udh was generated by digesting pTrc-Udh with NcoI and HindIII and cloning the resulting fragment into pRSFDuet-1. To generate pTrc-MIOX-ins-manX, pTrc-MIOX-1 and pTrc-ins-manX were digested with SpeI and PciI, and the relevant fragments were ligated together. pTrc-SUMO-MIOX-ins-manX was constructed in a similar manner using SpeI/PciI digests of pTrc-SUMO-MIOX and pTrc-ins-manX.

Table 2.1 | *E. coli* strains and plasmids used in this chapter.

Name	Relevant Genotype	Reference
Strains		
DH10B	F^- mcrA Δ(mrr-hsdRMA-mcrBC) ϕ80lacZΔM15 ΔlacX72 recA1 endA1 araD139 Δ(ara,leu)7697 galU galK λ-rpsL nupG	Life Technologies (Carlsbad, CA)
JW1087-2	F^-, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ^-, ΔptsG763::kan, rph-1, Δ(rhaD-rhaB)568, hsdR514	CGSC #9031 (Baba et al., 2006)
MG1655	F^- λ^- ilvG$^-$ frb-50 rph-1	CGSC #6300
M2	MG1655(DE3) Δ*endA* Δ*recA*	This study
M2-2	MG1655(DE3) Δ*endA* Δ*recA* Δ*gudD* Δ*uxaC*	This study
M2-2 Δ*ptsG*	MG1655(DE3) Δ*endA* Δ*recA* Δ*gudD* Δ*uxaC* Δ*ptsG*	This study
M2-2 Δ*sgrS*	MG1655(DE3) Δ*endA* Δ*recA* Δ*gudD* Δ*uxaC* Δ*sgrS*	This study
Plasmids		
pRedET	pSC101-derived ori[a], Tet^R, *araC*, λ phage *red*$\gamma\beta\alpha$ under control of P_{BAD}	Gene Bridges (Heidelberg, Germany)
pCP20	Rep[a], Amp^R, Cm^R, FLP recombinase expressed by λp_r under control of λcI857	CGSC #7629
pKD13	R6Kγ *ori*, Amp^R, *kan*	CGSC #7633
pKD46	R101 *ori*, *repA101*[a], Amp^R, *araC*, *araBp*-γ_γ-γ_β-γ_{exo}	CGSC #7739
pKD46recA	R101 *ori*, *repA101*[a], Amp^R, *araC*, *araBp*-γ_γ-γ_β-γ_{exo}, *recA*	(Solomon et al., 2013)
pRSFDuet-1	pRSF1030 *ori*, lacI, Kan^R	EMD4Biosciences (Darmstadt, Germany)
pE-SUMO	pBR322 *ori*, Kan^R	LifeSensors, Inc. (Malvern, PA)
pTrc99A	pBR322 *ori*, Amp^R	(Amann and Brosius, 1985)
pWW308	ColE1 *ori*, Amp^R, tetR, mRFP under control of P_{tet}	Dueber Lab (University of California, Berkeley)
pSB1A7	ColE1 *ori*, Amp^R	Registry of Standard Biological Parts[b]

(Continued on next page)

Table 2.1 | *E. coli* strains and plasmids used in this chapter (continued)

Name	Relevant Genotype	Reference
Plasmids		
pSB1A2-B0034	ColE1 *ori*, Amp^R, strong RBS	Registry of Standard Biological Parts (Part BBa_B0034)[b]
pSB1AK3-B0015	ColE1 *ori*, Amp^R, Kan^R, double terminator	Registry of Standard Biological Parts (Part BBa_B0015)[b]
pJ2-MIOX	Codon-optimized *Mus musculus* MIOX with 5′ EcoRI and 3′ HindIII sites	(Moon et al., 2009b)
pRSFD-MI	pRSFDuet-1 with MIOX inserted into the EcoRI and HindIII sites	This study
pSB1A7-tetR1	ColE1 *ori*, Amp^R, tetR (internal SpeI site removed), P_{tet}	This study
pSB1A7-tetR1-MIOX	pSB1A7-tetR1 with MIOX inserted into SpeI and PstI sites	This study
pSB1A7-tetR1-MIOX1	ColE1 *ori*, Amp^R, tetR, MIOX expressed from P_{tet}	This study
pSB1AK3-MIOX-B0015	pSB1AK3-B0015 with MIOX inserted into the EcoRI and XbaI sites	This study
pSB1A2-B0034-MIOX-B0015	pSB1A2-B0034 with MIOX-B0015 inserted into the SpeI and PstI sites	This study
pTrc-MIOX-1	pTrc99A with MIOX and an additional strong RBS (Part BBa_B0034, Registry of Standard Biological Parts)	This study
pE-SUMO-MIOX	pE-SUMO with MIOX inserted into the BsaI site	This study
pRSFD-SUMO-MIOX	pRSFDuet-1 with SUMO-MIOX from pE-SUMO-MIOX inserted into the NcoI and HindIII sites	This study
pTrc-SUMO-MIOX	pTrc99A with SUMO-MIOX inserted into the NcoI and HindIII sites	This study
pTrc-MIOX-DE	pTrc99A with MIOX directed evolution variant (R58H, V91F) inserted into the EcoRI and PstI sites; contains 941 bp fragment insertion into the PstI site	This study (Fig. 2.1A)
pTrc-MIOX-DE-noins	pTrc-MIOX-DE with 941 bp DNA fragment removed	This study
pTrc-MIOX-insert	pTrc-MIOX-1 with 941 bp DNA fragment inserted into the PstI site	This study

(Continued on next page)

Table 2.1 | *E. coli* strains and plasmids used in this chapter (continued)

Name	Relevant Genotype	Reference
Plasmids		
pTrc-insert	pTrc99A with 941 bp DNA fragment inserted into the PstI site	This study
pTrc-ins-manX	pTrc99A with *manX* portion of 941 bp DNA fragment	This study (Fig. 2.1B)
pTrc-ins-yoaE	pTrc99A with *yoaE* portion of 941 bp DNA fragment	This study (Fig. 2.1B)
pTrc-manX-frag	pTrc99A with *manX* fragment	This study (Fig. 2.1C)
pTrc-manX-full	pTrc99A with full-length *manX*	This study (Fig. 2.1C)
pTrc-manX-mut	pTrc99A with *manX* portion of 941 bp DNA fragment, start codon mutation (GTG → TAA)	This study (Fig. 2.1C)
pTrc-Udh	pTrc99A with Udh from *Pseudomonas syringae* inserted into the NcoI and HindIII sites	(Moon et al., 2009b, Yoon et al., 2009)
pTrc-Udh-ins	pTrc-Udh with 941 bp DNA fragment inserted into the PstI site	This study
pRSFD-Udh	pRSFDuet-1 with Udh inserted into the NcoI and HindIII sites	This study
pTrc-MIOX-ins-manX	pTrc-MIOX-1 with *manX* portion of 941 bp DNA fragment	This study
pTrc-SUMO-MIOX-ins-manX	pTrc-SUMO-MIOX with *manX* portion of 941 bp DNA fragment	This study

[a]Temperature-sensitive

[b]http://partsregistry.org

Table 2.2 | Oligonucleotides used in this chapter.

Name	5′ → 3′ Sequence[a]
KOrecA-F	*CAGAACATATTGACTATCCGGTATTACCCGGCATGACAGGAGTAAAAATG*ATTCCGGGGATCCGTCGACC
KOrecA-R	*ATGCGACCCTTGTGTATCAAACAAGACGATTAAAAATCTTCGTTAGTTTCTG*TGTAGGCTGGAGCTGCTTCG
KOendA-F	*AAACAGCTTTCGCTACGTTGCTGGCTCGTTTTAACACGGAGTAAGTGATG*ATTCCGGGGATCCGTCGACC
KOendA-R	*GTTAACAAAAAGAATCCCGCTAGTGTAGGTTAGCTCTTTCGCGCCTGGCA*TGTAGGCTGGAGCTGCTTCG
KOgudD-F	*AAACGTCCCGTTTTCGGCCGTCATTGATTCTGAAAAAGGACATAAATATGA*ATTCCGGGGATCCGTCGACC
KOgudD-R	*CCAGATAGAGCCGGTTTTGGTTTTCTGTCTTAACGCACCATGCACGGGCG*TGTAGGCTGGAGCTGCTTCG
KOuxaC-F	*CATCGCACCATAAGCAAGCTAGCTCACTCGTTGAGAGGAAGACGAAAATGA*ATTCCGGGGATCCGTCGACC
KOuxaC-R	*CTTGATGTATTGCATATCAACCCCAGACCTTAGTTCAGTTCAATGGCGAA*TGTAGGCTGGAGCTGCTTCG
pKD13_sgrS_fwd	*CATAAAAGGGGAACTCCTGTGCAAAAGACAGCAATTTTATTTTCCCTATATTAAGTCAATAATTCCTAAC*GTGTAGGCTGGAGCTGCTTC
pKD13_sgrS_rev	*AAACACCGTTCATACGGCGAGCCATCGTCATTATCCAGATCATACGTTCCCTTTTTAGCGCGGCGAGAAT*CTGTCAAACATGAGAATTAA
BBa_tet_fwd1	AATCGAATTCGCGGCCGCTTCTAGAGCGTGAAGTTACCATC
BBa_tet_rev1	CTGCAGCGGCCGCTACTAGTAGTGCTCAGTATCTC
tetR_A	GACTAGCAGATCCACTAGAG
tetR_B	CTCTAGTGGATCTGCTAGTC
pRSFD_fwd1	AATCGAATTCGCGGCCGCTTCTAGAGGTTTAACTTTAATAAGGAG
BBa_MIOX_rev2	CTGCAGCGGCCGCTACTAGTAGCTTTTACCAGGAC
MIOX_A	GGATAACCCGGATCTCCAGG
MIOX_B	CCTGGAGATCCGGGTTATCC
BBa_MIOX_rev3	TATCTGCAGCGGCCGCTACTAGTAGCTTTTACCAGGAC
pSB_A1	CAGCCAGGATCCGAATTACATGA
pSB_B1	TCATGTAATTCGGATCCTGGCTG
BBa_MIOX_fwd	GAATTCGCGGCCGCTTCTAGATGAAAGTTGATGTTGGTCC
VR	ATTACCGCCTTTGAGTGAGC
MI_SUMO_fwd	GGTCTCTAGGTATGAAAGTTGATGTTGGTCC
MI_SUMO_rev	GGTCTCTCTAGACCGCTACTAGTATATAAACGCAG

(Continued on next page)

Table 2.2 | Oligonucleotides used in this chapter (continued)

Name	5′ → 3′ Sequence[a]
pTrc_fwd	GCTCGTATAATGTGTGGAATTG
manX_rev	CATTAAGAATTCCATGACAA
yoaE_fwd	AATCTCTAGAATTGCTACCTCCTTTATTATC
pTrc_rev	CAGGCTGAAAATCTTCTCTC
manX_frag_fwd	AATCGAATTCGTGACCATTGCTATTGTTATAGGCACACATGGTTGGGCTGCTGCGGCCGCTACTAGTATATAACTGCAGAATC
manX_frag_rev	GATTCTGCAGTTATATACTAGTAGCGGCCGCAGCAGCCCAACCATGTGTGCCTATAACAATAGCAATGGTCACGAATTCGATT
manX_A	AATCGAATTCGTGACCATTGCTATTGTTATAGG
manX_B	AATCTCTAGACCAACACAATACGTTACTTATCG
manX_QC_fwd	ATGTGTGCCTATAACAATAGCAATGGTTTATTGCTACCTCCTTTATTATCGTTAACA
manX_QC_rev	TGTTAACGATAATAAAGGAGGTAGCAATAAACCATTGCTATTGTTATAGGCACACAT
DE_MIOX_fwd	ATCTCTAGAGAAAGAGGAGAAATACTAGATG
DE_MIOX_rev1	AAGCTTGCATGCCTGCAG
ptsG_fwd	CGGTTCCGCGAATTTCAG
ptsG_rev	CCGCCTGCTTCTGCCATA

[a]All oligonucleotides were purchased from Sigma-Genosys (St. Louis, MO). Homologous sequences used for recombination are *underlined and italicized*; restriction sites used for cloning are underlined.

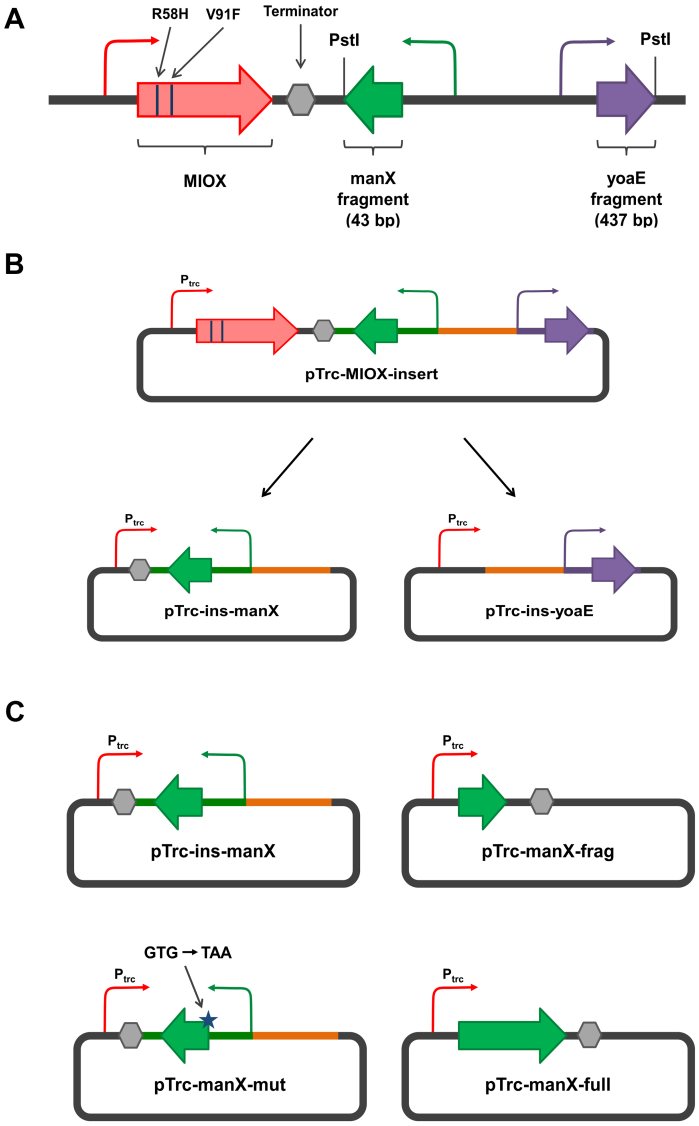


Figure 2.1 | Schematics of various plasmids used in this chapter. **(A)** Schematic of pTrc-MIOX-DE detailing the location of the DNA fragment. **(B)** Schematics of pTrc-ins-manX and pTrc-ins-yoaE. The intergenic region between and *manX* and *yoaE* promoters (orange) was retained in both constructs. Additionally, the 5′ UTR for *manX* and *yoaE* was retained in each of the respective constructs. **(C)** Schematics of pTrc-ins-manX, pTrc-manX-mut, pTrc-manX-frag, and pTrc-manX-full. In pTrc-manX-mut, the start codon of *manX* (GTG) has been mutated to a stop codon (TAA), and the native promoter and 5′ UTR have been retained. In pTrc-manX-frag and pTrc-manX-full, the native promoter and 5′ UTR have been replaced by the promoter and 5′ UTR contained within pTrc99A.

2.2.2 Culture and Assay Conditions for Organic Acid Production

For D-glucuronic and D-glucaric acid production, cultures were grown in 250 mL baffled shake flasks containing 50 mL LB medium supplemented with 60 mM (10.8 g/L) *myo*-inositol. Cultures were induced at inoculation with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and ampicillin (100 µg/mL) and kanamycin (30 µg/mL) were added as required. Seed cultures were grown overnight at 30°C in LB medium without *myo*-inositol and inoculated to an optical density at 600 nm (OD_{600}) of 0.01. Cultures were incubated at 30°C, 250 rpm, and 80% relative humidity for 72 hours. Samples were taken daily, centrifuged to remove cell debris, and the supernatants analyzed for metabolite concentrations.

Myo-inositol, D-glucuronic acid, and D-glucaric acid were quantified from culture supernatants with high performance liquid chromatography (HPLC) on an Agilent Series 1100 instrument equipped with an Aminex HPX-87H column (300 mm by 7.8 mm; Bio-Rad Laboratories, Hercules, CA). Sulfuric acid (5 mM) was used as the mobile phase at 55°C and a flow rate of 0.6 mL/min in isocratic mode. Compounds were detected and quantified from 10 µL sample injections using a refractive index detector. Reported metabolite concentrations are the average of triplicate samples, and error bars represent one standard deviation above and below the mean value.

2.2.3 Analysis of MIOX Activity and Expression Level

To prepare lysates for analysis of MIOX activity and expression level, cell pellets taken 12, 24, 48, and 72 hours post-inoculation were resuspended in sodium phosphate buffer (100 mM, pH 7.0) supplemented with an EDTA-free protease inhibitor cocktail (Roche Applied Science, Indianapolis, IN), then sonicated. The lysed samples were centrifuged to remove insoluble proteins, and the total protein concentration of the soluble fraction was determined using a modified Bradford assay (Zor and Selinger, 1996). Assays for MIOX activity were performed with appropriate no-lysate and no-substrate as described previously (Moon et al., 2009a), and measured activities were normalized by the measured total protein concentration. Reported activities are averages of triplicate samples, and error bars represent one standard deviation above and below the mean value.

For analysis of MIOX expression levels, 15 µg of total protein from each lysate was separated via SDS-PAGE and transferred onto nitrocellulose blotting membranes (Pall Life Sciences, Port Washington, NY). Following blocking, membranes were incubated overnight in a 1:200 dilution of anti-MIOX antibody (Santa Cruz

Biotechnology, Santa Cruz, CA). Immunodetection was performed using an anti-goat IgG-HRP antibody and Western Blotting Luminol Reagent (Santa Cruz Biotechnology, Santa Cruz, CA) according to the manufacturer's instructions.

2.2.4 Directed Evolution of MIOX

One round of directed evolution was performed to generate MIOX variants with improved productivity. Diversity was generated via error-prone PCR using the Genemorph Mutazyme II kit (Agilent Technologies, Santa Clara, CA) using primers DE_MIOX_fwd and DE_MIOX_rev1 and pTrc-MIOX-1 as a template. A template amount of 200 ng was used to provide medium mutagenesis rates (4.5 to 9 mutations/kb) as per manufacturer instructions. PCR products were subjected to DpnI treatment to eliminate parental DNA, then digested with XbaI and PstI and ligated into similarly digested pTrc99A. To maximize library size, commercial *E. coli* DH10B (Life Technologies, Carlsbad, CA) was transformed with the resulting ligation products and plated onto LB agar plates, resulting in approximately 10^4 colonies. Supercoiled plasmid DNA was harvested from this plate and used to transform MG1655(DE3) Δ*endA* Δ*recA* (strain M2) for screening on M9 minimal medium agar plates supplemented with 60 mM *myo*-inositol and 0.1 mM IPTG for induction of protein expression. Following incubation at 30°C, plasmid DNA was harvested from 15-30 large colonies, used to transform strain M2-2, and D-glucuronic production from *myo*-inositol was measured for cultures possessing each variant. Subsequent sequencing analysis of these colonies revealed an average mutation rate of 1.6 mutations/kb. We speculate that MIOX variants containing significantly more mutations were catalytically dead and unable to grow on *myo*-inositol as a sole carbon source.

2.2.5 Quantification of mRNA Levels

To quantify mRNA levels, samples of approximately 10^9 cells were taken 6 hours after inoculation. Total RNA was extracted from each of these samples using the illustra RNAspin Mini RNA Isolation Kit (GE Healthcare Bio-Sciences, Piscataway, NJ) with an on-column DNaseI treatment according to the manufacturer's instructions. Following an additional treatment to remove trace DNA contamination, 500 ng of total RNA was used to synthesize cDNA using random primers with the QuantiTect Reverse Transcription Kit (Qiagen, Valencia, CA). The synthesized cDNA was then amplified in a quantitative PCR (qPCR) reaction with primers ptsG_fwd and ptsG_rev for quantification of ptsG mRNA levels. qPCR reactions contained Brilliant II SYBR

Green High ROX QPCR Master Mix and were performed on an ABI 7300 Real Time PCR System Instrument (Applied Biosystems, Beverly, MA). Transcript levels were quantified in triplicate with appropriate no-template and no-RT (reverse transcriptase) controls and are relative to that of transcript levels found in strain M2-2 as determined from a standard curve. Dilution of cDNA samples was performed as necessary to keep C_t values within the linear range of the assay. Reported transcript levels are the averages of triplicate samples, each measured in triplicate. Error bars represent one standard deviation above and below the mean value.

2.3 Results

2.3.1 Protein Fusions for Increased Soluble Expression

Fusing difficult-to-express proteins to highly soluble affinity tags has been widely explored and has been reported to promote proper protein folding as well as increase protein expression, solubility, and half-life in a number of cases (Esposito and Chatterjee, 2006, Young et al., 2012a). A large number of fusion tags, including maltose binding protein (MBP), N-utilization substance A (NusA), glutathione S-transferase (GST), small ubiquitin-related modifier (SUMO), and protein disulfide isomerase I (DsbA) have been characterized and used in a variety of protein expression schemes (Esposito and Chatterjee, 2006, Young et al., 2012a). Green fluorescent protein (GFP) has also been used to increase soluble protein expression (Wu et al., 2009) and also as a reporter of soluble protein expression when fused to the C-terminus (Waldo et al., 1999). Despite the large amount of fusion tags which have been reported, predicting the fusion partner which optimizes expression for a particular protein *a priori* remains difficult.

A previous study showed MIOX stability to be a significant factor in limiting pathway productivity: although *myo*-inositol supplementation resulted in increased MIOX activity during the exponential phase, MIOX activity was quickly lost in stationary phase (Moon et al., 2009b). In addition, we have observed MIOX expression levels to be much lower than that of INO1 and Udh (data not shown). We hypothesized that increased soluble expression of MIOX would lead to increased D-glucaric acid titers and explored protein fusions as a method for increasing soluble protein expression. We fused MIOX to three fusion tags: MBP (N-terminal), GFP (C-terminal), and SUMO (N-terminal). Although MBP-MIOX and MIOX-GFP both expressed well, neither of these fusion proteins showed any activity *in vivo* (data not shown). The third fusion protein, SUMO-MIOX, displayed significantly increased productivity, resulting in a 125% increase in final D-glucuronic acid titers compared to

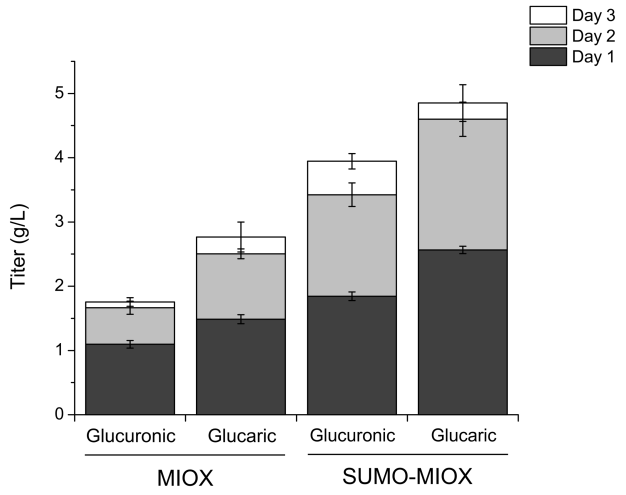


Figure 2.2 | Cumulative production of D-glucuronic and D-glucaric acids from *myo*-inositol with unfused and SUMO-fused MIOX. Productivity gains from the protein fusion were most significant during the second day of culture. For D-glucuronic acid production, cultures contained strain M2-2 harboring an empty pTrc99A vector and pRSFD-MI or pRSFD-SUMO-MIOX; cultures for D-glucaric acid production contained strain M2-2 harboring pTrc-Udh and pRSFD-MI or pRSFD-SUMO-MIOX. Error bars represent the standard deviation of biological triplicates.

an unfused control (Fig. 2.2). A similar boost in MIOX productivity was observed when SUMO-MIOX was combined with Udh for D-glucaric acid production: final D-glucaric acid titer with the fused protein was 4.85 g/L, 75% higher than with the unfused control (Fig. 2.2).

The productivity of unfused MIOX is moderate during Day 1 but is quickly lost in subsequent days. Fused MIOX, on the other hand, exhibits significantly higher productivity during Day 1 and is able to retain much of that productivity in subsequent days. There are several possible reasons for this observed gain in MIOX productivity: (1) improved MIOX specific activity, potentially due to active site stabilization via allosteric effects exerted by the SUMO fusion, (2) improved MIOX solubility, and (3) improved MIOX stability. To determine the underlying cause for the increase in MIOX productivity, we measured *in vitro* MIOX activity and analyzed MIOX protein levels using Western blots at several points during a typical three-day culture (Fig. 2.3). The results indicate that the observed boost in MIOX productivity is due to increased MIOX solubility and stability: SUMO-MIOX expression levels are much higher than unfused MIOX at every time point analyzed, and SUMO-MIOX levels remain relatively constant over time while unfused MIOX diminishes (Fig. 2.3A). Because SUMO-MIOX is present at much higher concentrations in the soluble fraction, *in vitro* activity in crude lysates is also higher (Fig. 2.3B). Additionally, specific activities for MIOX and SUMO-MIOX estimated by normalizing the *in vitro* activities in Fig. 2.3B by relative expression levels determined by spot densitometry were not

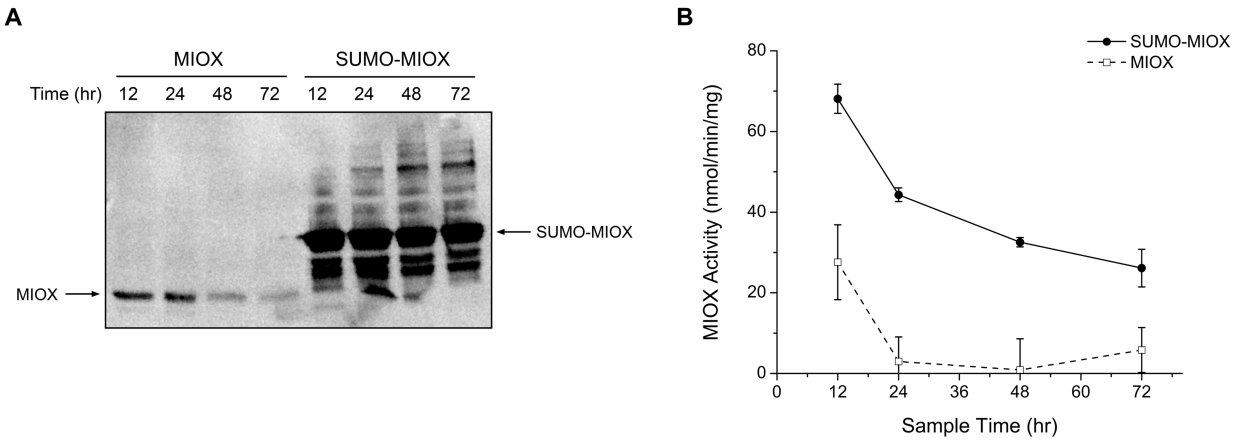


Figure 2.3 | Comparison of MIOX and SUMO-MIOX expression and activity. **(A)** MIOX and SUMO-MIOX expression levels at various times, as indicated. **(B)** *In vitro* MIOX activity for unfused MIOX and SUMO-MIOX in crude lysates. Cultures contained strain M2-2 harboring an empty pTrc99A vector and pRSFD-MI or pRSFD-SUMO-MIOX. Error bars represent the standard deviation of biological triplicates.

significantly different, suggesting that the SUMO fusion does not affect MIOX specific activity. Finally, while unfused MIOX has lost the majority of its activity after 24 hours of culture, SUMO-MIOX loses its activity more slowly, retaining nearly 40% of its activity measured at 12 hours after 72 hours of culture.

2.3.2 Directed Evolution of MIOX

Directed evolution is a powerful tool and has been used extensively to alter or improve various properties of a protein of interest (Collins et al., 2005, Hibbert and Dalby, 2005, Hawkins et al., 2007, Atsumi and Liao, 2008). In many cases, the lack of structural information necessitates a random approach, whereby sequence/structure diversity is generated via nonspecific methods such as error-prone PCR and whole-cell mutagenesis (Marcheschi et al., 2013). These methods are capable of generating large amounts of diversity in a short amount of time; however, the identification of protein variants with the desired properties from such diversity is a colossal task. The success of a random directed evolution approach therefore depends entirely on the availability of a suitable, high-throughput screen or selection for the desired protein property.

For directed evolution of MIOX, we took advantage of native *E. coli* metabolism to construct a screen for improved MIOX productivity. *E. coli* MG1655 cannot grow on *myo*-inositol as a sole carbon source, a trait which we verified in a minimal medium liquid culture supplemented with *myo*-inositol (data not shown). *E. coli* MG1655 can, however, grow on D-glucuronic acid as a sole carbon source (Ashwell, 1962). Consequently,

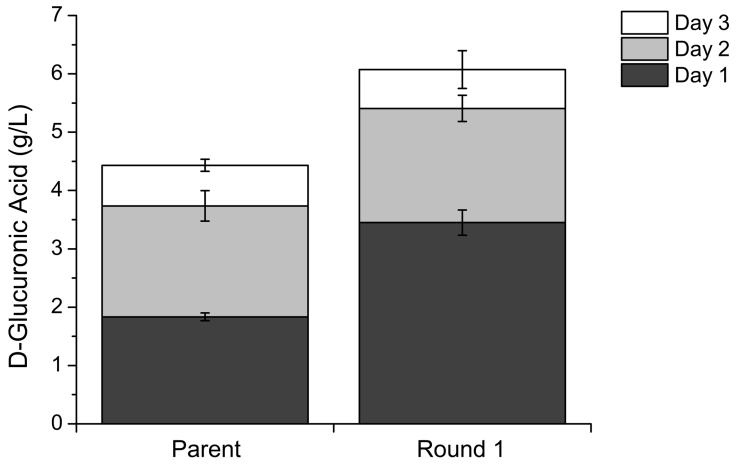


Figure 2.4 | D-glucuronic acid production from *myo*-inositol with unmutated MIOX (Parent) and the MIOX variant isolated from the first round of directed evolution (Round 1). Cultures contained strain M2-2 harboring pTrc-MIOX-1 (Parent) or pTrc-MIOX-1-DE (Round 1). Error bars represent the standard deviation of biological triplicates.

E. coli can grow on *myo*-inositol if supplied with MIOX, which converts *myo*-inositol to D-glucuronic acid. Furthermore, *E. coli* cultures which possess more active variants of MIOX should be able to grow at a faster rate. We employed a growth-based screen to identify improved MIOX variants whereby a library was plated onto minimal medium agar supplemented with *myo*-inositol. Plasmid DNA was isolated from colonies exhibiting increased growth, sequenced, then used to transform strain M2-2. The resulting clones were then tested for D-glucuronic acid production. After just one round of screening, we isolated what initially appeared to be a MIOX variant capable of producing significantly more D-glucuronic acid from *myo*-inositol, especially during the first day of cultivation. Final D-glucuronic acid titers in excess of 6 g/L were achieved with this MIOX variant (Fig. 2.4).

Sequencing of the newly isolated variant revealed two mutations in the MIOX coding sequence, R58H and V91F. Interestingly, sequencing also revealed a 941 bp DNA fragment inserted 3′ to the MIOX coding sequence and downstream of the terminator (Fig. 2.1A). We speculate that this fragment was inserted into the expression vector during the cloning procedure: following error-prone PCR amplification of the MIOX coding sequence, MIOX and the expression vector pTrc99A were both digested with XbaI and PstI and ligated. Because the insert is flanked by PstI sites (Fig. 2.1A), we believe that the insert originated from genomic DNA contamination in the pTrc99A miniprep. This insert contained fragments of the *manX* and *yoaE* genes (43 bp and 437 bp, respectively) and included the genes' native promoters from the *E. coli* genome. *manX*

encodes for $IIAB^{man}$, a membrane-associated phosphotransferase system (PTS) permease responsible for phosphoryl group transfer from phosphoenolpyruvate to incoming mannose molecules (Postma et al., 1993). The function of YoaE is not known; however, it has been identified as a putative inner membrane protein based on sequence homology (Serres et al., 2001). To test whether the observed increases in glucuronic acid productivity were due to the acquired mutations in MIOX or the 941 bp DNA fragment, we cloned the insert downstream of unmutated MIOX and removed the insert from the mutated MIOX variant. The results clearly indicate that the presence of the 941 bp DNA fragment is the cause of increased D-glucuronic acid titers (Fig. 2.5A). When the fragment is present, strains harboring unmutated MIOX and the MIOX variant demonstrate similar productivities, generating similar amounts of D-glucuronic acid at similar rates. When the fragment is absent, the MIOX variant performs slightly worse than its unmutated counterpart, indicating that the accumulated mutations may in fact be detrimental to MIOX productivity.

To determine the fragment's mechanism of action, we created two variants, *manX* only (pTrc-ins-manX) and *yoaE* only (pTrc-ins-yoaE) (Fig. 2.1B), and measured D-glucuronic acid production with MIOX expressed in *trans* from a separate plasmid to eliminate any potential *cis* effects (Fig.2.5B). The results indicate that the *manX* portion of the fragment increases productivity, not the *yoaE* fragment. However, overexpression of this small, 43 bp *manX* portion from an IPTG-inducible promoter did not result in increased D-glucuronic acid production compared to a control strain, nor did overexpression of full-length *manX* (Fig. 2.5B). In contrast, D-glucuronic acid production remained high when the start codon of the *manX* portion was removed from the original construct (manX-mut). Because a productivity enhancement was achieved even without translation of *manX*, we concluded that *manX* mRNA must be responsible for the observed effect. Additionally, the fact that modifications which altered the mRNA secondary structure (i.e., swapping of promoters and 5′ UTRs) resulted in lower D-glucuronic acid titers than that obtained with the native *manX* fragment supports this conclusion (Fig. 2.5B vs. Fig.2.5A).

manX has been reported to contain a binding site for *sgrS*, a *trans*-acting regulatory small RNA (Rice and Vanderpool, 2011). *sgrS* transcription is upregulated under conditions of sugar-phosphate stress (e.g., glucose 6-phosphate accumulation) and acts to downregulate expression of PTS permeases (e.g., PtsG, ManX), thereby alleviating that stress. Downregulation is achieved via base pairing interactions between *sgrS* and the target PTS permease mRNA, occluding the ribosome binding site and inhibiting translation. Additionally, the *sgrS*-PTS permease mRNA complex is quickly degraded by RNaseE, resulting in further downregulation of PTS permease expression (Rice and Vanderpool, 2011). The *sgrS*-PTS permease interaction has been

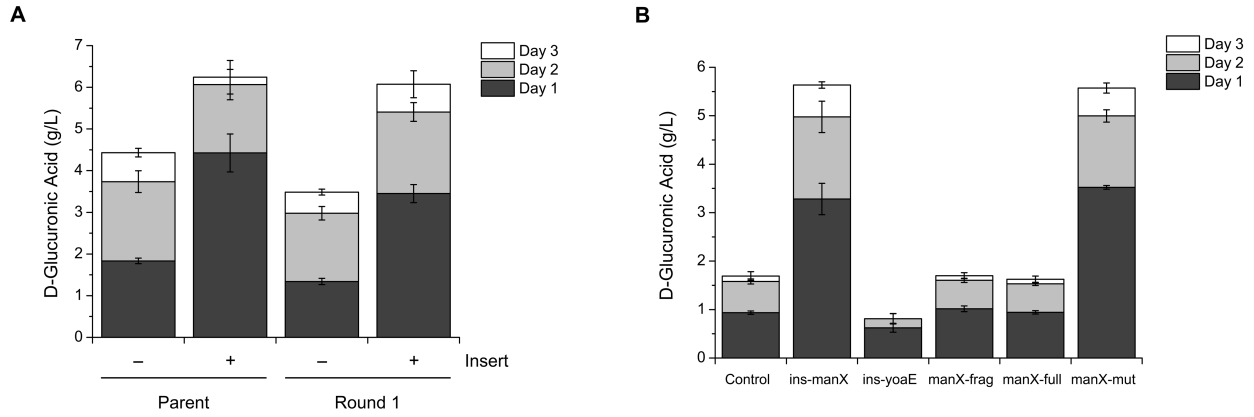


Figure 2.5 | Effect of the insert on D-glucuronic acid production from *myo*-inositol. **(A)** Presence of the insert resulted in increased productivity for both unmutated MIOX and the Round 1 mutant. Cultures contained strain M2-2 harboring pTrc-MIOX-1 (Parent, without insert), pTrc-MIOX-insert (Parent, with insert), pTrc-MIOX-DE-noins (Round 1, without insert), or pTrc-MIOX-DE (Round 1, with insert). Error bars represent the standard deviation of biological triplicates. **(B)** Presence of the *manX* portion of the insert resulted in increased D-glucuronic acid production, while the presence of the *yoaE* portion did not. Overexpression of neither the *manX* fragment (manX-frag) nor full-length *manX* (manX-full) resulted in increased D-glucuronic acid productivity, however. Because mutation of the *manX* start codon to a stop codon (manX-mut) did not affect D-glucuronic acid titers, improvement of D-glucuronic acid productivity must be dependent upon the presence of *manX* mRNA. Cultures contained strain M2-2 harboring pRSFD-MI and pTrc99A (Control), pTrc-ins-manX, pTrc-ins-yoaE, pTrc-manX-frag, pTrc-manX-full, or pTrc-manX-mut. See Figs. 2.1B and 2.1C for schematics of the above plasmids. Error bars represent the standard deviation of biological triplicates.

leveraged recently to control PtsG expression and glucose uptake in *E. coli*: by overexpressing *sgrS*, Negrete et al. were able to reduce PtsG expression and acetate secretion (a general indicator of glycolytic overflow) in *E. coli* K-12 (Negrete et al., 2013). We hypothesized that transcription of the *manX* fragment reduces intracellular levels of *sgrS* mRNA, leading to increased PtsG expression (Fig. 2.6). To validate this model, we measured *ptsG* mRNA transcript levels in cultures expressing the *manX* fragment as well as in a Δ*sgrS* mutant (Fig. 2.7). The results indicate that *ptsG* mRNA levels are significantly higher in cells harboring the *manX* fragment. *ptsG* mRNA levels are also higher in the Δ*sgrS* strain, albeit much lower than the case with insert present. Interestingly, D-glucuronic acid productivity in the Δ*sgrS* mutant is not significantly different from the *sgrS*$^+$ strain (Fig. 2.8). Because *ptsG* is highly regulated, we speculate that another regulatory process becomes dominant upon deletion of *sgrS*; additionally, this result suggests that the model depicted in Fig. 2.6 is not complete and that another interaction is involved or that the relationship between PtsG levels and pathway productivity is nonlinear.

The positive correlation between ptsG transcript level and D-glucuronic acid production suggests that

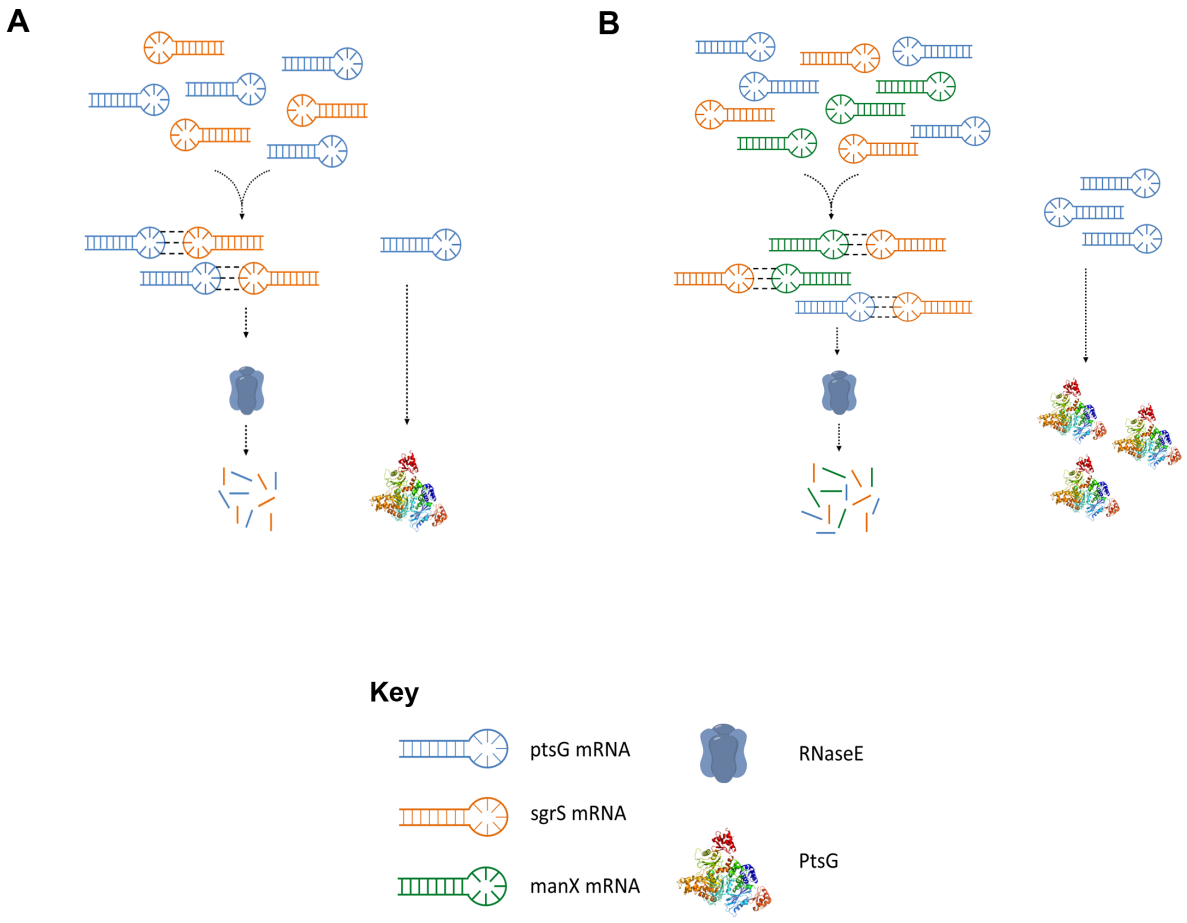


Figure 2.6 | Model for fragment-mediated increase in D-glucuronic acid productivity. **(A)** When the *manX* fragment (green) is absent, *sgrS* mRNA (orange) primarily binds to *ptsG* mRNA (blue), leading to degradation of *ptsG* mRNA and reduced PtsG expression. **(B)** When the *manX* fragment is present, both *sgrS-ptsG* and *sgrS-manX* binding interactions occur. Assuming that total (bound and unbound) *sgrS* and *ptsG* transcript levels remain unchanged in the presence of the *manX* fragment, this extra binding interaction leads to a net increase in the amount of unbound *ptsG* mRNA, leading to increased PtsG expression.

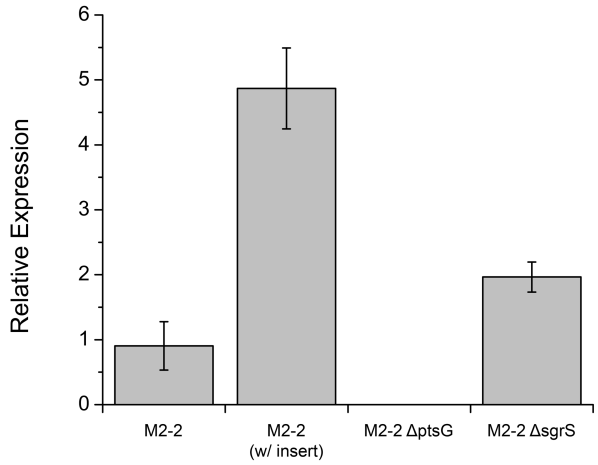


Figure 2.7 | Relative *ptsG* transcript levels in various strains. *ptsG* transcript levels were approximately five-fold higher than the control in the presence of the insert and nearly two-fold higher in a Δ*sgrS* strain. No transcript was detected in the Δ*ptsG* strain. Cultures contained either strain M2-2 with empty pTrc99A, strain M2-2 with pTrc-ins-manX, strain M2-2 Δ*ptsG* with empty pTrc99A, or strain M2-2 Δ*sgrS* with empty pTrc99A. All strains also contained pRSFD-MI. Error bars represent the standard deviation of biological triplicates.

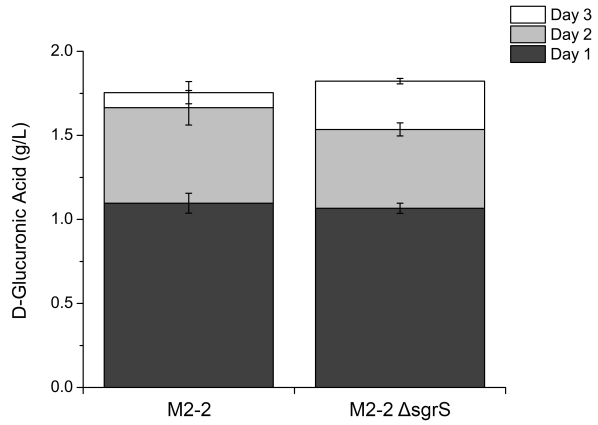


Figure 2.8 | D-glucuronic acid production in M2-2 vs. M2-2 Δ*sgrS*. Deletion of *sgrS* from M2-2 did not result in a significant increase in D-glucuronic acid titers. This result was somewhat surprising, given that a reduction in *sgrS* concentration via expression of the *manX* fragment resulted in a significant increase in D-glucuronic acid titers. We speculate that complete removal of *sgrS* leads to dominance of another *ptsG* regulatory mechanism, which lessens the effects of removing *sgrS*. Cultures contained strain M2-2 or M2-2 Δ*sgrS* harboring pRSFD-MI and empty pTrc99A. Error bars represent the standard deviation of biological triplicates.

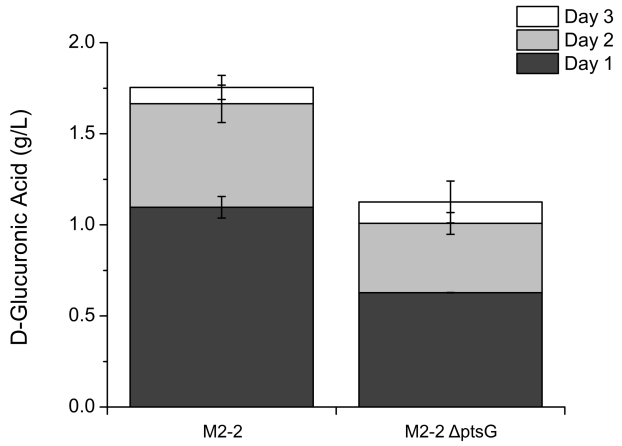


Figure 2.9 | D-glucuronic acid production in M2-2 vs. M2-2 Δ*ptsG*. Deletion of *ptsG* resulted in a 35% decrease in D-glucuronic acid production from *myo*-inositol, implicating PtsG in *myo*-inositol transport. Cultures contained strain M2-2 or M2-2 Δ*ptsG* with pRSFD-MI and empty pTrc99A. Error bars represent the standard deviation of biological triplicates.

PtsG functions as a transporter of *myo*-inositol in *E. coli* MG1655. Indeed, D-glucuronic acid production is reduced by 35% in the absence of ptsG (Fig. 2.9). The *manX* fragment, then, increases D-glucuronic acid production by increasing *myo*-inositol transport rates into the cell via increased expression of PtsG. This increased transport rate does not result in increased MIOX specific activity, however (Fig. 2.10). Instead, we hypothesize that the increased rate of transport leads to higher intracellular concentrations of *myo*-inositol, resulting in an *in vivo* MIOX activity much closer to the measured *in vitro* activity, which is measured at saturating *myo*-inositol concentrations. Finally, combination of the insert with MIOX and uronate dehydrogenase (Udh) yielded D-glucaric acid titers of 4.58 g/L, a 65% increase over a control strain without the insert (Fig. 2.11). Previously, we have observed D-glucaric acid titers equal to or greater than D-glucuronic acid titers when Udh is introduced (Fig. 2.2). In this case, D-glucaric acid titers were approximately 20% lower than the D-glucuronic acid titers obtained in the absence of Udh. We hypothesize that D-glucaric acid becomes inhibitory at concentrations approaching 5 g/L, effectively limiting D-glucaric acid titers to less than this threshold (Sec. 2.3.4).

2.3.3 Exploration of Synergistic Effects

Given that the two methods we investigated function in fundamentally disparate manners, we expected that additional gains in D-glucaric acid productivity could be realized by combining the two methods. However, no gains were observed with the original promoter configuration, with MIOX expressed from a T7 promoter

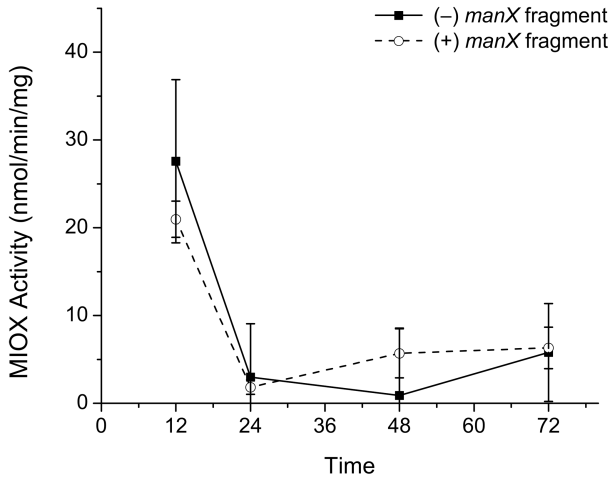


Figure 2.10 | MIOX activity in the absence and presence of the *manX* fragment. Activity of MIOX expressed in the presence of the insert was not significantly different from MIOX expressed in the absence of the insert. Cultures contained strain M2-2 harboring pRSFD-MI and pTrc99A or pTrc-ins-manX. Error bars represent the standard deviation of biological triplicates.

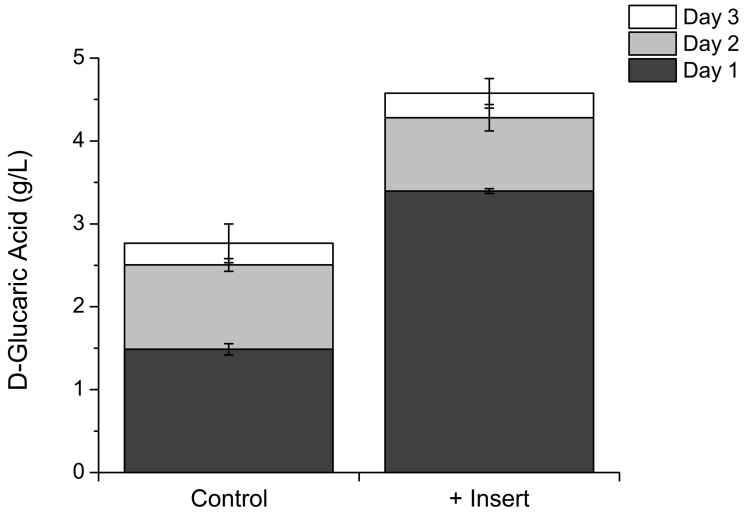


Figure 2.11 | D-glucaric acid production in the absence and presence of the *manX* fragment. Presence of the insert resulted in D-glucaric acid titers of 4.58 g/L, a 65% improvement over the control. Cultures contained strain M2-2 harboring pRSFD-MI and pTrc-Udh (Control) or pTrc-Udh-ins (+ Insert). Error bars represent the standard deviation of biological triplicates.

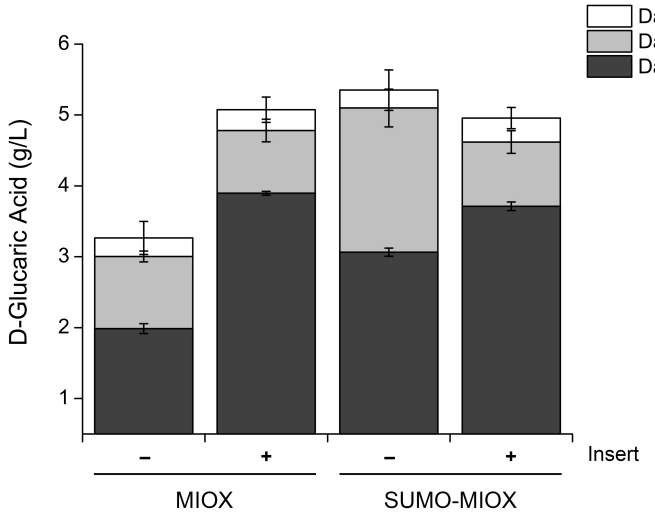


Figure 2.12 | D-glucaric acid production in the presence and absence of the 941 bp insert with MIOX and SUMO-MIOX expressed from pRSFDuet-1. The presence of the insert or SUMO fusion tag resulted in increased D-glucaric acid production, but these two effects are not synergistic. Cultures contained strain M2-2 harboring pRSFD-MI or pRSFD-SUMO-MIOX. Cultures without insert also contained pTrc-Udh, while cultures with insert contained pTrc-Udh-ins in addition to pRSFD-MI or pRSFD-SUMO-MIOX. Error bars represent the standard deviation of biological triplicates.

in pRSFDuet-1 and Udh expression from P_{trc} in pTrc99A (Fig. 2.12). A comparison of Figures 2.2 and 2.4 indicates that the expression system used for MIOX strongly influences its productivity: final D-glucuronic acid titers are 2.5 fold higher with pTrc-MIOX-1 compared to pRSFD-MI. To explore the effect of promoter configuration on D-glucaric acid production with the insert and SUMO fusion, we measured D-glucaric acid production in strains expressing MIOX and SUMO-MIOX from P_{trc} in pTrc99A and Udh from a T7 promoter in pRSFDuet-1 (Fig. 2.13). With this promoter configuration, synergistic effects become evident. The strain harboring SUMO-MIOX and the insert slightly outperformed the strain harboring the insert only and significantly outperformed the strain carrying SUMO-MIOX only during the first 24 hours of culture. This gain in productivity is lost after an additional 24 hours of culture, however, as final D-glucaric titers (4.67 g/L) are similar in all strains carrying either SUMO-MIOX, the insert, or both. Coupled with the maximum D-glucaric acid titer reported in Sec. 2.3.2 (4.58 g/L), this pattern hints at a product inhibition effect which limits pathway productivity.

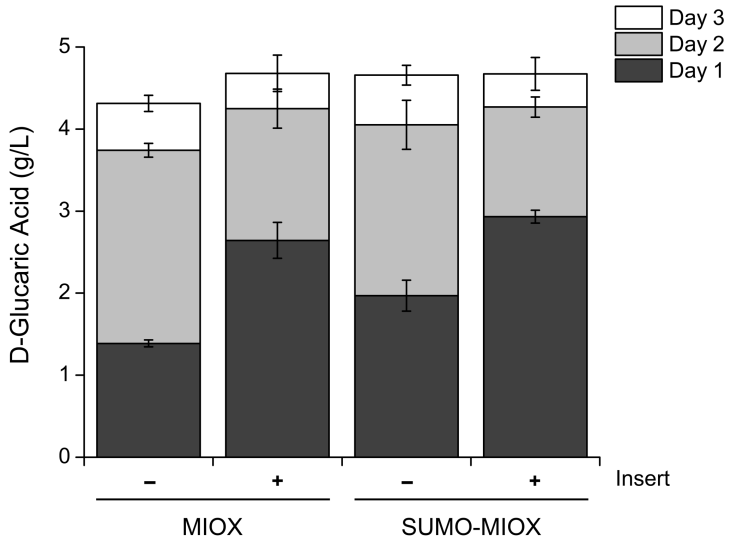


Figure 2.13 | D-glucaric acid production in the presence and absence of the 941 bp insert with MIOX and SUMO-MIOX. Combination of the insert with the SUMO fusion resulted in a slight increase in D-glucaric acid productivity over the insert alone in the initial 24 hours of culture; however, this effect is lost in subsequent days. Cultures contained strain M2-2 harboring pRSFD-Udh. Cultures without insert also contained pTrc-MIOX-1 or pTrc-SUMO-MIOX, while cultures with insert contained pTrc-MIOX-ins-manX or pTrc-SUMO-MIOX-ins-manX. Error bars represent the standard deviation of biological triplicates.

2.3.4 Inhibitory Effects of D-Glucaric Acid

Because we have not observed D-glucaric acid titers above 5 g/L, we hypothesize that D-glucaric acid at this concentration is inhibitory to further D-glucaric acid production. This hypothesis is supported by observed decreases in *in vitro* MIOX activity in the presence of increasing amounts of D-glucaric acid (Fig. 2.14); however, the relatively small extent to which MIOX is inhibited by D-glucaric acid cannot fully explain the apparent upper limit of D-glucaric acid titers. To further investigate this phenomenon, we supplemented cultures producing D-glucaric acid with 2 and 4 g/L D-glucaric acid at inoculation. In both cases, final D-glucaric acid titers did not exceed 5 g/L, providing further evidence for inhibition by D-glucaric acid at high concentrations (data not shown).

A potential explanation for the observed inhibitory effect of D-glucaric acid is pH: production of organic acids such as D-glucaric acid decreases culture pH, triggering *E. coli* stress response mechanisms which lead to physiological and metabolic changes within the cell (Warnecke and Gill, 2005). In the experiment described in the previous paragraph, cultures were not neutralized following the addition of D-glucaric acid.

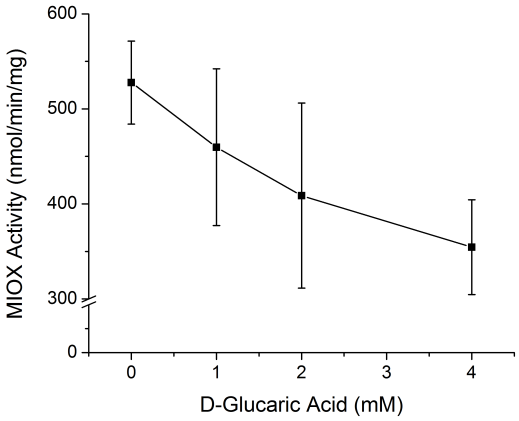
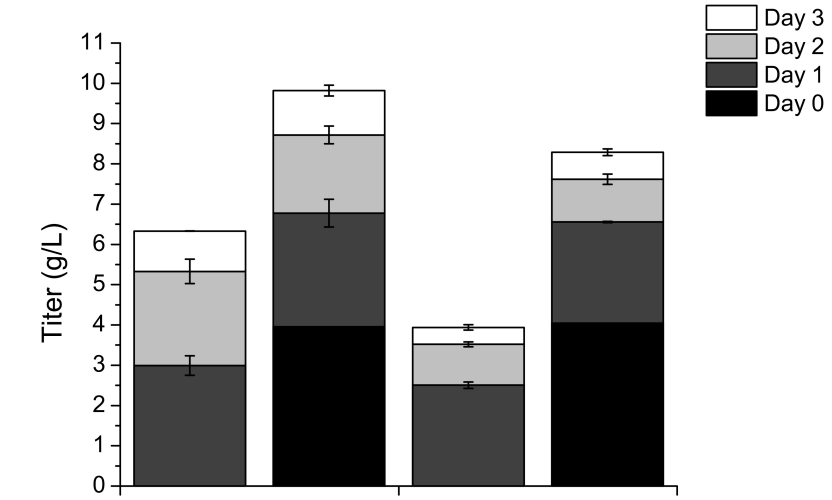


Figure 2.14 | *In vitro* MIOX activity in the presence of increasing amounts of D-glucaric acid. Measured MIOX activity decreases slightly as the amount of D-glucaric acid added to the assay reaction is increased. MIOX was expressed from pTrc-SUMO-MIOX-ins-manX in the presence of 60 mM *myo*-inositol. Crude MIOX-containing lysates were prepared from cultures grown for 24 hours at 30°C. Error bars represent the standard deviation of biological triplicates.

To test whether culture pH affects D-glucaric acid productivity, we performed a similar D-glucaric production experiment with the addition of approximately 4 g/L *neutralized* D-glucaric acid at inoculation (Fig. 2.15). D-glucuronic production in the presence of approximately 4 g/L neutralized D-glucuronic acid was also investigated for comparison. In this case, cultures which were supplemented with neutralized D-glucaric acid at inoculation produced similar amounts of D-glucaric acid (4.35 g/L) compared to a control which contained no D-glucaric acid at inoculation (3.94 g/L). Since the addition of neutralized D-glucaric acid did not significantly affect final titers, pH must play a significant role in limiting D-glucaric acid productivity.

Although we have not observed D-glucaric acid titers exceeding 5 g/L, we have observed D-glucuronic acid titers significantly higher than this threshold (Figs.2.4 and 2.5). Given that pH plays an important role in pathway productivity, we measured final pH values for each of the cultures described in the previous paragraph (Fig. 2.15). The pH of cultures producing D-glucuronic acid was approximately one pH unit higher that of cultures producing D-glucaric acid. D-glucuronic acid is a monocarboxylic acid with a pK_a of 3.30. D-glucaric acid, on the other hand, is a dicarboxylic acid with a pK_a of 2.99, indicating that a solution of D-glucaric acid will have a lower pH than an equimolar solution of D-glucuronic acid. Because we have observed inhibited product formation at low pH, we speculate that this chemical difference between D-glucuronic and D-glucaric acids is responsible for the observed titer differences between the two products.



Product	D-Glucuronic Acid		D-Glucaric Acid	
Initial Product (g/L)	0	3.95	0	4.05
Final pH	6.01 ± 0.09	6.34 ± 0.12	5.15 ± 0.05	5.29 ± 0.11

Figure 2.15 | D-glucuronic and D-glucaric acid production as well as final pH in cultures with and without approximately 4 g/L D-glucuronic or D-glucaric acid added at inoculation. For cultures to which additional D-glucuronic or D-glucaric acid was added, culture pH was adjusted to approximately 7.0 following addition of the acid. Final pH measurements were taken after 72 hours of culture. Cultures producing D-glucuronic acid contained strain M2-2 harboring pRSFD-SUMO-MIOX and pTrc-ins-manX, and cultures producing D-glucaric acid contained strain M2-2 harboring pRSFD-SUMO-MIOX and pTrc-Udh-insert. Error bars represent the standard deviation of biological triplicates.

2.4 Discussion

The elimination of pathway bottlenecks to maximize flux towards products of interest is a cornerstone of metabolic engineering. Overexpression of pathway genes is a common strategy for overcoming pathway limitations and has been applied to a number of metabolic pathways with great success (McKenna and Nielsen, 2011, Lütke-Eversloh and Stephanopoulos, 2008, Tseng et al., 2010). However, this strategy is ineffective for pathways in which enzyme solubility is an issue, as overexpression of weakly soluble proteins simply results in more insoluble protein. For these pathways, tools which increase protein solubility and/or activity are needed to achieve increases in productivity.

In this work, we studied soluble fusion partners as a method for increasing soluble expression of a pathway-limiting enzyme. While soluble fusion partners are commonly used to increase soluble protein yields in protein

purification applications (Young et al., 2012a), their use in metabolic engineering contexts remains limited, perhaps due to their unpredictable effect on protein activity. For example, MIOX fusions to maltose binding protein and green fluorescent protein were expressed at significantly higher levels than unfused MIOX but were completely inactive. A MIOX fusion to SUMO, however, resulted in a nearly twofold increase in D-glucaric acid productivity over an unfused control. Such disparity in the behavior of different fusion partners highlights the difficulty of predicting the optimal fusion partner for a particular protein *a priori*. On the other hand, the significant improvement in MIOX activity realized with the SUMO fusion highlights the power of soluble fusion partners for metabolic engineering applications. To our knowledge, this is the first successful application of soluble fusion partners towards an *in vivo* metabolic engineering problem. Further work to understand the interactions between fusion partners could facilitate increased use of soluble fusion partners for metabolic engineering applications.

Another strategy for improving protein properties is directed evolution. The directed evolution efforts in this study, however, did not result in increased protein solubility or activity but instead led to the isolation of an *E. coli* genomic DNA fragment which, when transcribed, resulted in increased D-glucuronic and D-glucaric acid productivities. Given that the directed evolution screen used in this study was designed to identify improved MIOX variants via improved growth on *myo*-inositol, this result was not altogether surprising, and this work stands as a prime example of the First Law of Directed Evolution; "you get what you screen for" (Schmidt-Dannert and Arnold, 1999). Further modifications to the screening method may be necessary to isolate MIOX mutants which exhibit increased specific activity, especially in the presence of *myo*-inositol concentrations more relevant in the context of D-glucaric acid synthesis from glucose.

Despite the failure to generate an improved MIOX variant, the directed evolution efforts in the study did lead to several interesting results. First, PtsG can function as a *myo*-inositol transporter in *E. coli*, although the PTS permease is not the sole transporter. While *myo*-inositol transporters have been identified in *Salmonella enterica* (Kröger et al., 2010), this is the first evidence of a *myo*-inositol transporter in *E. coli*. Second, *myo*-inositol transport is a limiting factor for D-glucaric acid production from *myo*-inositol. This result emphasizes the need to consider substrate transport into the cell for metabolic engineering problems. While substrate transport is often overlooked, tools to engineer transport have recently begun to emerge (Young et al., 2012b). Third, production of *manX* mRNA results in post-transcriptional upregulation of PtsG expression. While downregulation of PtsG expression through overexpression of *sgrS* has been demonstrated (Negrete et al., 2013), we report a method for post-transcriptional upregulation of PtsG expression. The combination

of these two methods represents a method for fine-tuned control of PtsG expression and glucose transport rate, a strategy which may be applied to any number of metabolic engineering problems where glucose is the primary substrate.

The identification of pH-related effects on D-glucaric acid productivity represents an interesting opportunity for process and strain improvements which may increase productivity by mitigating these effects. Mechanisms of organic acid toxicity and tolerance in *E. coli* have been reviewed previously (Warnecke and Gill, 2005), and engineering and/or adaptation of *E. coli* for improved acid tolerance could further improve D-glucaric acid productivity. The use of buffered media or online pH control in a bioreactor could also result in improved titers.

Ultimately, D-glucaric acid production from D-glucose is desired. However, only a small amount of D-glucaric acid was produced in a strain harboring INO1, Udh, SUMO-fused MIOX, and the *manX* insert fragment, and this amount was not significantly more than that produced in a control strain harboring INO1, Udh, and unfused MIOX (data not shown). The results of this experiment indicate that SUMO-fused MIOX still requires high concentrations of its substrate *myo*-inositol for activation. Further engineering will therefore be required to produce a system capable of generating large amounts of D-glucaric acid from D-glucose as a feedstock.

2.5 Conclusions

In this study, we report two parallel methods for improving D-glucaric acid production from *myo*-inositol in *E. coli*. An N-terminal SUMO fusion to MIOX resulted in a 75% increase in D-glucaric acid production, yielding final D-glucaric acid titers of 4.85 g/L from 10.8 g/L *myo*-inositol. Expression of a small fragment of *manX* mRNA, on the other hand, resulted in a 65% increase in D-glucaric acid production, yielding final D-glucaric acid titers of 4.58 g/L. A combination of these strategies increased initial D-glucaric acid productivity, but this effect was lost by the end of the culture. Additionally, pH was identified to have a major impact on D-glucaric acid productivity. Further work such as scale-up and process engineering will be necessary to continue to increase productivity towards titers relevant for commercialization of D-glucaric acid. However, this successful application of two parallel methods towards improved D-glucaric acid production bodes well for continued engineering efforts, as this expanded toolset provides a much greater potential for synergistic improvements to D-glucaric acid titers in the future.

Chapter 3

Application of Synthetic Biology Strategies

Abstract

Previous work has identified MIOX as the least active D-glucaric acid pathway enzyme and determined that MIOX activity can be increased significantly with the addition of *myo*-inositol to the culture medium (Moon et al., 2009b). These observations led to the idea of delaying MIOX expression relative to INO1, allowing *myo*-inositol to build up in the culture medium. Activation of MIOX expression after a threshold *myo*-inositol concentration is reached would ensure exposure of newly produced MIOX to activating concentrations of *myo*-inositol, resulting in increased concentrations of active MIOX and higher pathway productivity. A proof-of-concept experiment in which MIOX expression was delayed 10 hours relative to INO1 resulted in a five-fold increase in D-glucuronic acid production from D-glucose.

Although manual induction of MIOX via time-resolved addition of chemical inducers increased pathway productivity, the use of chemical inducers in industrial settings is generally undesired, as these molecules are often quite expensive. The use of chemical inducers is especially undesired in the case of D-glucaric acid, which is expected to be a low-margin, commodity-scale biochemical. This chapter describes attempts to develop "genetic timers" inspired by synthetic biology to autonomously delay MIOX expression relative to INO1. Such devices would not only eliminate the need for exogenous addition of chemical inducers but may also provide robustness against normal variations in biological processes.

3.1 Introduction

The observation that addition of *myo*-inositol to the culture medium significantly increased MIOX specific activity during exponential phase led to the desire to couple MIOX expression to *myo*-inositol production as a strategy to increase MIOX activity and D-glucaric acid titers. Such a coupling would effectively delay expression of MIOX, allowing *myo*-inositol to build up in the culture medium. After a threshold *myo*-inositol concentration is reached, MIOX expression would be activated, ensuring exposure of newly produced MIOX to high *myo*-inositol concentrations. This idea was tested in a simple proof-of-concept experiment where MIOX induction was manually delayed in a strain harboring plasmids for both INO1 and MIOX (Fig. 3.1). INO1, under control of an IPTG-inducible T7 promoter, was induced at culture inoculation. MIOX, under control of aTc-inducible P_{tet}, was expressed at varying times post-inoculation by the addition of aTc. It is clear from Fig. 3.1 that delayed expression of MIOX is beneficial for pathway productivity. However, a tradeoff between increased MIOX activity and productivity exists: the longer MIOX induction is delayed, the longer *myo*-inositol is allowed to accumulate in the culture broth, leading to increased activation of MIOX following induction. On the other hand, longer induction delays also reduce the amount of time available for MIOX expression, leading to reduced MIOX concentrations. We speculate that induction delays greater than 10 hours post-inoculation would lead to decreased product titers.

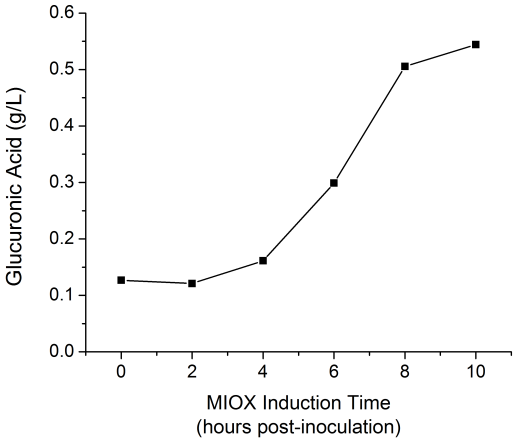


Figure 3.1 | Final D-glucuronic titers measured at 72 hours with delayed MIOX induction. INO1 was induced with 0.1 mM IPTG at inoculation, and MIOX was induced with 100 ng/mL aTc at the times indicated in the figure. Cultures contained strain M2-2 harboring pRSFD-IN and pSB1A7-tetR1-MIOX and were grown in LB + 1% D-glucose at 30°C.

Although manually delayed induction of MIOX is able to increase product titers significantly, the need for multiple chemical inducers is an enormous barrier to commercialization of biological D-glucaric acid production, as chemical inducers can be quite costly at industrial scales. In addition, a manually delayed induction scheme requires close monitoring of the fermentation, since normal biological variation and unforeseen process disturbances could affect cellular growth characteristics. To address both of these issues, we decided to explore synthetic biology devices capable of autonomously delaying the induction of gene expression. Two methods of delaying MIOX expression were investigated: (1) *myo*-inositol-responsive machinery for direct coupling of MIOX expression to intracellular *myo*-inositol levels, and (2) genetic "timers" for cell density-dependent induction of MIOX. While the applicability of method (1) is restricted to the D-glucaric acid pathway, it is hoped that the genetic "timers" in method (2) will be applicable to a wide variety of systems which require time-resolved induction of gene expression. The following section provides an overview of the systems which we attempted to utilize for time-delayed gene expression.

3.1.1 *Myo*-Inositol Responsive Promoters

Salmonella enterica serovar Typhimurium possesses an endogenous pathway for *myo*-inositol catabolism (Fig. 3.2) (Kröger and Fuchs, 2009). We hypothesized that the genes involved in *myo*-inositol utilization are activated in the presence of *myo*-inositol, since the expression of several sugar catabolism genes in *E. coli* (e.g., *araBAD*, *lacZYA*) is activated in the presence of that sugar (e.g., arabinose and lactose, respectively). In particular, *iolG*, which catalyzes the first step of the *myo*-inositol utilization pathway, may be subject to regulation by *myo*-inositol. If this is the case, then these *myo*-inositol responsive promoters can be recruited to activate MIOX expression once intracellular *myo*-inositol has accumulated above a particular threshold.

By cloning the promoters for the *myo*-inositol catabolism genes in front of a luminescence operon, Kröger et al. provide evidence that these promoters are responsive to *myo*-inositol: P_{iolI2}, the promoter for *iolI2*, was induced nearly 17,000-fold when grown in minimal media containing *myo*-inositol compared to a control grown in glucose (Kröger and Fuchs, 2009). P_{iolE} was induced over 3,000-fold in *myo*-inositol over glucose. Additionally, Kröger et al. demonstrate that *iolR* acts as a transcriptional repressor of the *myo*-inositol catabolic pathway: in a Δ*iolR* strain, all of the *Salmonella* promoters studied were induced, even in the absence of *myo*-inositol.

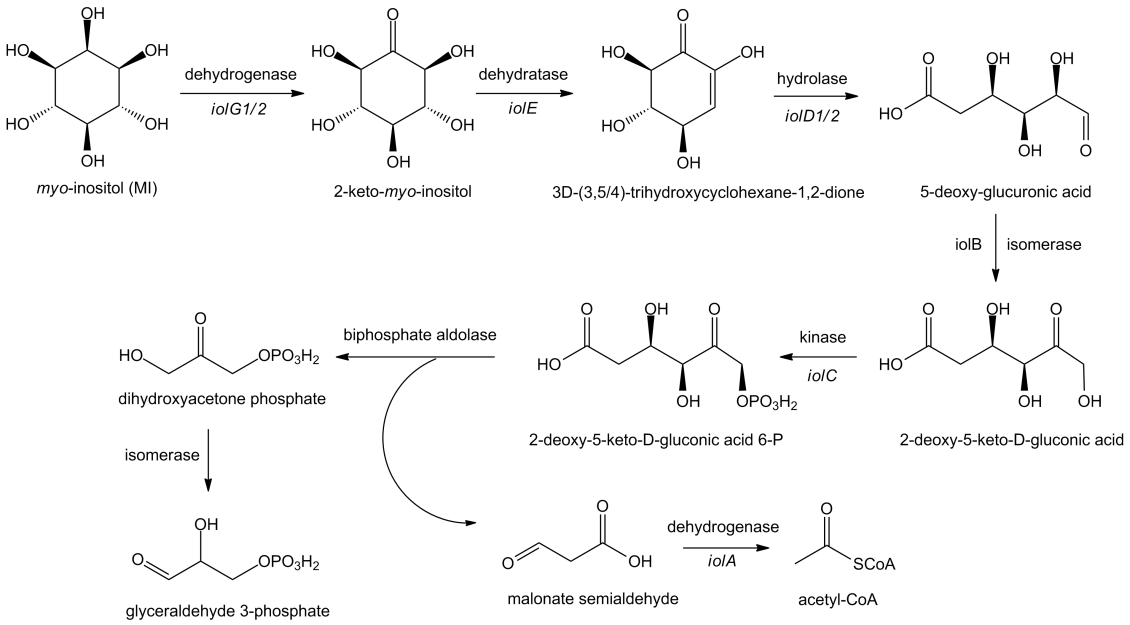


Figure 3.2 | *Myo*-inositol utilization pathway in *Salmonella enterica* serovar Typhimurium (Kröger and Fuchs, 2009).

3.1.2 Genetic Timers

Genetic timers, which delay gene expression autonomously, could have wide application to a myriad of systems. In the medical field, for example, cell density-dependent induction has been used for localized killing of tumor cells (Anderson et al., 2006). For this application, *E. coli* cells were engineered to invade cancerous cells only when high cell densities are present. This behavior mimics that of many plant pathogens, which delay expression of maceration enzymes until pathogen cell densities are high enough to overwhelm the plant. In the biotechnology field, cell density-dependent or metabolite-dependent expression systems can be used to maximize production of a product that competes directly with glycolysis, or to minimize metabolic burden from overexpression of heterologous enzymes. As more and more chemicals and proteins are being produced microbially, the potential of these genetic timers continues to grow.

Quroum Sensing

The phenomenon of quorum sensing, which involves cell density-dependent regulation, has been described in a wide variety of bacterial species (Uroz et al., 2009, Surette et al., 1999, Dunlap, 1999). Quorum sensing relies on the fast diffusion of signaling molecules (acyl homoserine lactones, AHLs) to achieve synchronized,

density-dependent gene expression. At low cell densities, newly synthesized AHLs diffuse out of the cell quickly and become diluted in the extracellular media. At such low concentrations, the AHLs do not bind to their cognate transcriptional activators at an appreciable rate, so expression of quorum-related genes is low. At high cell densities, a basal level of AHL has been established in the extracellular medium, so AHL buildup within the cell is able to occur, resulting in the expression of quorum-regulated genes. In 2004, it was shown that quorum sensing machinery could be used to control gene expression in a density-dependent manner (Kobayashi et al., 2004). Since then, quorum sensing machinery has been used to link a variety of cellular processes to cell density, including cell death (You et al., 2004), invasion (Anderson et al., 2006), and motility (Weiss et al., 2008). In this work, we explore quorum sensing systems from several organisms and attempt to construct "genetic timers" using this machinery.

Sigma/Anti-Sigma Factor Pairs

In prokaryotes, sigma factors initiate gene transcription by binding to promoters and recruiting RNA polymerase. Because sigma factors serve as "gatekeepers" of gene expression, these proteins are excellent control points for regulating gene expression. By controlling the binding of a sigma factor to a promoter, one can control the expression of any genes downstream of the promoter. For example, the catabolite repressor protein CRP prevents gene expression by binding to a particular DNA operator site, preventing sigma factor binding and RNA polymerase assembly. The presence of D-glucose leads to activation of CRP and subsequent downregulation of a whole host of catabolite-repressed genes. Another method of control is sequestration, where an anti-sigma factor binds to and sequesters its cognate sigma factor. When a sigma factor is sequestered by its anti-sigma factor, it is unable to activate transcription. It is this sigma/anti-sigma factor interaction that we wish to exploit to create a "genetic timer" capable of delaying gene expression.

3.2 Materials and Methods

3.2.1 *E. coli* Strains and Plasmids

E. coli strains and plasmids used in this study are listed in Table 3.1; oligonucleotides used are listed in Table 3.2. All molecular biology manipulations were performed according to standard practices (Sambrook and Russell, 2001). *E. coli* DH10B was used for transformation of cloning reactions and propagation of all

plasmids. Unless otherwise indicated, BioBricks Assembly Standard RFC 10 was used for all cloning procedures (Knight, 2003): to insert Part A upstream of Part B, Part A is digested with EcoRI and SpeI and purified away from its vector backbone. Part B is digested with EcoRI and XbaI, and the two parts are ligated. To insert Part A downstream of Part B, Part A is digested with XbaI and PstI, Part B is digested with SpeI and PstI, and the two parts are ligated. This assembly standard relies on the generation of compatible sticky ends by XbaI and SpeI which, upon ligation, generates a "scar" sequence that is not digestible by XbaI or SpeI.

To construct pSB1A7-PiolC1, pSB1A7-PiolD1, pSB1A7-PiolE, pSB1A7-PiolG2, pSB1A7-PiolI2, and pSB1A7-PiolR, promoter sequences P_{iolC1}, P_{iolD1}, P_{iolE}, P_{iolG2}, P_{iolI2}, and P_{iolR} as identified by Kröger et al. (2009) were first amplified from *Salmonella typhimurium* strain LT2 genomic DNA using primers BBa_iolC1_fwd1 and BBa_iolC1_rev1, BBa_iolD1_fwd1 and BBa_iolD1_rev1, BBa_iolE_fwd1 and BBa_iolE_rev1, BBa_iolG2_fwd1 and BBa_iolG2_rev1, BBa_iolI2_fwd1 and BBa_iolI2_rev1, and BBa_iolR_fwd1 and BBa_iolR1_rev1, respectively. To facilitate cloning, each forward primer contained a BioBricks prefix, and each reverse primer contained a BioBricks suffix. These PCR products were then digested with EcoRI and SpeI, then ligated into similarly digested pSB1A7, a plasmid backbone available from the Registry of Standard Biological Parts (http://partsregistry.org). These promoters abut the start codon of their respective open reading frames, so it is expected that they will contain a native ribosome binding site. pSB1A7-iolR0 was constructed by amplifying *iolR* with its native promoter from *Salmonella typhimurium* strain LT2 genomic DNA using primers BBa_iolR_fwd1 and BBa_iolR_rev1, digesting with EcoRI and SpeI, and ligating into similarly digested pSB1A7. pSB1A7-PiolC1-GFP, pSB1A7-PiolD1-GFP, pSB1A7-PiolE-GFP, pSB1A7-PiolG2-GFP, pSB1A7-PiolI2-GFP, and pSB1A7-PiolR-GFP were constructed by cloning GFP from pSB1A2-GFP downstream of the *Salmonella* promoters in pSB1A7-PiolC1, pSB1A7-PiolD1, pSB1A7-PiolE, pSB1A7-PiolG2, pSB1A7-PiolI2, and pSB1A7-PiolR using BioBricks assembly. Finally, *iolR* with its native promoter from pSB1A7-iolR0 was inserted upstream of PiolC1-GFP, PiolD1-GFP, PiolE-GFP, PiolG2-GFP, PiolI2-GFP, and PiolR-GFP using BioBricks Assembly Standard RFC 10 to generate pSB1A7-iolR0-PiolC1-GFP, pSB1A7-iolR0-PiolD1-GFP, pSB1A7-iolR0-PiolE-GFP, pSB1A7-iolR0-PiolG2-GFP, pSB1A7-iolR0-PiolI2-GFP, and pSB1A7-iolR0-PiolR-GFP.

For construction of the quorum sensing circuits, three constitutive promoters of varying strengths (relative strengths: 1.00, 0.56, and 0.10) from the Anderson promoter library were obtained from the Registry of Standard Biological Parts (http://partsregistry.org); these promoters are referred to individually as Pcon100, Pcon56, and Pcon10 in Table 3.1 and collectively as PconXX. Plasmids carrying AHL synthetases (*luxI*, *lasI*, and *rhlI*), quorum-responsive regulators (*luxR*, *lasR*, and *rhlR*), and quorum-controlled promoters (P_{lux}, P_{las},

and P_{rhl}) were also obtained from the Registry of Standard Biological Parts. Each of the AHL synthetases and quorum-responsive regulators were first cloned downstream of a strong RBS from the Registry of Standard Biological Parts (Part BBa_B0034); each RBS-CDS construct was then cloned upstream of a double terminator from the Registry of Standard Biological Parts (Part BBa_B0015). Each RBS-CDS-terminator was then cloned downstream of three different Anderson library promoters, yielding a total of 18 constructs: pSB1A2-PconXX-luxI, pSB1A2-PconXX-luxR, pSB1A2-PconXX-lasI, pSB1A2-PconXX-lasR, pSB1A2-PconXX-rhlI, and pSB1A2-PconXX-rhlR. Finally, inserts containing the promoter-AHL synthetase were subcloned into vector backbone pSB3C5 using EcoRI and PstI to avoid origin compatibility issues during coexpression of the AHL synthetase and quorum-responsive regulator.

To generate reporter devices for the quorum sensing circuits, GFP from pSB1A2-GFP was first cloned downstream of a strong RBS from the Registry of Standard of Biological Parts (Part BBa_B0034) to generate pSB1A2-B0034-GFP. The B0034-GFP insert from this plasmid was then cloned upstream of a double terminator from pSB1AK3-B0015 to yield pSB1AK3-B0034-GFP-B0015. This insert was then cloned downstream of each of the quorum-controlled promoters to generate pSB1A2-Plux-GFP, pSB1A2-Plas-GFP, and pSB1A2-Prhl-GFP. These promoter-GFP inserts were then cloned downstream of their cognate quorum-responsive regulators in pSB1A2-PconXX-luxR, pSB1A2-PconXX-lasR, and pSB1A2-PconXX-rhlR using standard BioBricks assembly to generate pSB1A2-PconXX-luxR-Plux-GFP, pSB1A2-PconXX-lasR-Plas-GFP, and pSB1A2-PconXX-rhlR-Prhl-GFP, respectively.

pDC296, which contains the sigma/anti-sigma factor pair *sigW*/*rsiW* from *Bacillus subtilis* was kindly provided by Prof. Adam Arkin (University of California, Berkeley). To generate a reporter plasmid for this circuit, the P_{sigW} promoter was amplified from pDC296 using primers Psig_fwd and Psig_rev and cloned upstream of GFP in pSB1AK3-B0034-GFP-B0015 using standard BioBricks assembly to produce pSB1AK3-PsigW-GFP. The promoter-GFP construct was then subcloned into pSB3C5 to generate pSB3C5-PsigW-GFP, again using standard BioBricks assembly.

To swap the P_{BAD} promoter of *rsiW* in pDC296 for an aTc-inducible promoter, the rsiW-RFP cassette was first amplified from pDC296 using the primers rsi_fwd1/rsi_rev1 and inserted downstream of the tetR cassette in pSB1A7-tetR1, generating pSB1A7-tetR1-rsiW-RFP. The P_{sigW}-sigW cassette was then amplified from pDC296 using the primers sig_fwd1/sig_rev1 and cloned into pSB1A7, yielding pSB1A7-PsigW-sigW. Finally, the P_{sigW}-sigW cassette in pSB1A7-PsigW-sigW was cloned downstream of RFP in pSB1A7-tetR1-rsiW-RFP to produce pECS01.

Table 3.1 | *E. coli* strains and plasmids used in this chapter.

Name	Relevant Genotype	Reference
Strains		
DH10B	F$^-$ mcrA Δ(mrr-hsdRMA-mcrBC) ϕ80lacZΔM15 ΔlacX72 recA1 endA1 araD139 Δ(ara,leu)7697 galU galK λ-rpsL nupG	Life Technologies (Carlsbad, CA)
M2	MG1655(DE3) ΔendA ΔrecA	This study (Sec. 2.2.1)
Plasmids		
pSB1A2	ColE1 *ori*, Amp^R	Registry of Standard Biological Parts[a]
pSB1A7	ColE1 *ori*, Amp^R	Registry of Standard Biological Parts[a]
pSB3C5	p15A *ori*, Cm^R	Registry of Standard Biological Parts[a]
pSB1A2-B0034	ColE1 *ori*, Amp^R, strong RBS	Registry of Standard Biological Parts (Part BBa_B0034)[a]
pSB1AK3-B0015	ColE1 *ori*, Amp^R, Kan^R, double terminator	Registry of Standard Biological Parts (Part BBa_B0015)[a]
pSB1A2-GFP	GFPmut3b in pSB1A2	Registry of Standard Biological Parts (Part BBa_E0040)[a]
Timed Induction - Proof of Concept		
pRSFD-IN	pRSFDuet-1 with INO1 from *Saccharomyces cerevisiae*	(Moon et al., 2009b)
pSB1A7-tetR1-MIOX1	ColE1 *ori*, Amp^R, tetR, MIOX expressed from P_{tet}	This study (Sec. 2.2.1)
Salmonella promoters		
pSB1A7-iolR0	pSB1A7 with *iolR* from *Salmonella typhimurium*	This study

(Continued on next page)

Table 3.1 | *E. coli* strains and plasmids used in this chapter (continued)

Name	Relevant Genotype	Reference
Plasmids		
pSB1A7-PiolC1	pSB1A7 with P_{iolC1} from *Salmonella typhimurium*	This study
pSB1A7-PiolD1	pSB1A7 with P_{iolD1} from *Salmonella typhimurium*	This study
pSB1A7-PiolE	pSB1A7 with P_{iolE} from *Salmonella typhimurium*	This study
pSB1A7-PiolG2	pSB1A7 with P_{iolG2} from *Salmonella typhimurium*	This study
pSB1A7-PiolI2	pSB1A7 with P_{iolI2} from *Salmonella typhimurium*	This study
pSB1A7-PiolR	pSB1A7 with P_{iolR} from *Salmonella typhimurium*	This study
pSB1A7-PiolC1-GFP	pSB1A7 with GFPmut3b under control of P_{iolC1} from *Salmonella typhimurium*	This study
pSB1A7-PiolD1-GFP	pSB1A7 with GFPmut3b under control of P_{iolD1} from *Salmonella typhimurium*	This study
pSB1A7-PiolE-GFP	pSB1A7 with GFPmut3b under control of P_{iolE} from *Salmonella typhimurium*	This study
pSB1A7-PiolG2-GFP	pSB1A7 with GFPmut3b under control of P_{iolG2} from *Salmonella typhimurium*	This study
pSB1A7-PiolI2-GFP	pSB1A7 with GFPmut3b under control of P_{iolI2} from *Salmonella typhimurium*	This study
pSB1A7-PiolR-GFP	pSB1A7 with GFPmut3b under control of P_{iolR} from *Salmonella typhimurium*	This study
pSB1A7-iolR0-PiolC1-GFP	pSB1A7-PiolC1-GFP with *iolR* from *Salmonella typhimurium* inserted upstream of P_{iolC1}-GFP	This study

(Continued on next page)

Table 3.1 | *E. coli* strains and plasmids used in this chapter (continued)

Name	Relevant Genotype	Reference
Plasmids		
pSB1A7-iolR0-PiolD1-GFP	pSB1A7-PiolD1-GFP with *iolR* from *Salmonella typhimurium* inserted upstream of P_{iolD1}-GFP	This study
pSB1A7-iolR0-PiolE-GFP	pSB1A7-PiolE-GFP with *iolR* from *Salmonella typhimurium* inserted upstream of P_{iolE}-GFP	This study
pSB1A7-iolR0-PiolG2-GFP	pSB1A7-PiolG2-GFP with *iolR* from *Salmonella typhimurium* inserted upstream of P_{iolG2}-GFP	This study
pSB1A7-iolR0-PiolI2-GFP	pSB1A7-PiolI2-GFP with *iolR* from *Salmonella typhimurium* inserted upstream of P_{iolI2}-GFP	This study
pSB1A7-iolR0-PiolR-GFP	pSB1A7-PiolR-GFP with *iolR* from *Salmonella typhimurium* inserted upstream of P_{iolR}-GFP	This study
Quorum sensing circuits		
J61002-Pcon100	ColE1 *ori*, Amp^R, mRFP under control of strong constitutive promoter	Registry of Standard Biological Parts (Part BBa_J23100)[a]
J61002-Pcon56	ColE1 *ori*, Amp^R, mRFP under control of medium strength constitutive promoter	Registry of Standard Biological Parts (Part BBa_J23118)[a]
J61002-Pcon10	ColE1 *ori*, Amp^R, mRFP under control of weak constitutive promoter	Registry of Standard Biological Parts (Part BBa_J23114)[a]
pSB1A2-luxI	pSB1A2 with *luxI* AHL synthetase from *Vibrio fischeri*	Registry of Standard Biological Parts (Part BBa_C0161)[a]
pSB1A2-luxR	pSB1A2 with *luxR* quorum-responsive regulator from *Vibrio fischeri*	Registry of Standard Biological Parts (Part BBa_C0062)[a]

(Continued on next page)

Table 3.1 | *E. coli* strains and plasmids used in this chapter (continued)

Name	Relevant Genotype	Reference
Plasmids		
pSB1A2-Plux	pSB1A2 with LuxR-controlled promoter from *Vibrio fischeri*	Registry of Standard Biological Parts (Part BBa_R0062)[a]
pSB1A2-lasI	pSB1A2 with *lasI* AHL synthetase from *Pseudomonas aeruginosa*	Registry of Standard Biological Parts (Part BBa_C0178)[a]
pSB1A2-lasR	pSB1A2 with *lasR* quorum-responsive regulator from *Pseudomonas aeruginosa*	Registry of Standard Biological Parts (Part BBa_C0179)[a]
pSB1A2-Plas	pSB1A2 with LasR-controlled promoter from *Pseudomonas aeruginosa*	Registry of Standard Biological Parts (Part BBa_R0079)[a]
pSB1A2-rhlI	pSB1A2 with *rhlI* AHL synthetase from *Pseudomonas aeruginosa*	Registry of Standard Biological Parts (Part BBa_C0170)[a]
pSB1A2-rhlR	pSB1A2 with *rhlR* quorum-responsive regulator from *Pseudomonas aeruginosa*	Registry of Standard Biological Parts (Part BBa_C0171)[a]
pSB1A2-Prhl	pSB1A2 with RhlR-controlled promoter from *Pseudomonas aeruginosa*	Registry of Standard Biological Parts (Part BBa_I14017)[a]
pSB1A2-B0034-GFP	pSB1A2-B0034 with GFPmut3b cloned downstream of B0034 RBS	This study
pSB1AK3-B0034-GFP-B0015	pSB1AK3-B0015 with B0034-GFP cloned upstream of B0015 double terminator	This study
pSB1A2-Plux-GFP	pSB1A2 with GFP under control of P_{lux}	This study
pSB1A2-Plas-GFP	pSB1A2 with GFP under control of P_{las}	This study
pSB1A2-Prhl-GFP	pSB1A2 with GFP under control of P_{rhl}	This study

(Continued on next page)

Table 3.1 | *E. coli* strains and plasmids used in this chapter (continued)

Name	Relevant Genotype	Reference
Plasmids		
pSB1A2-PconXX-luxI[b]	pSB1A2 with luxI under control a of constitutive Anderson promoter	This study
pSB1A2-PconXX-luxR[b]	pSB1A2 with luxR under control a of constitutive Anderson promoter	This study
pSB1A2-PconXX-lasI[b]	pSB1A2 with lasI under control a of constitutive Anderson promoter	This study
pSB1A2-PconXX-lasR[b]	pSB1A2 with lasR under control a of constitutive Anderson promoter	This study
pSB1A2-PconXX-rhlI[b]	pSB1A2 with rhlI under control a of constitutive Anderson promoter	This study
pSB1A2-PconXX-rhlR[b]	pSB1A2 with rhlR under control a of constitutive Anderson promoter	This study
pSB1A2-PconXX-luxR-Plux-GFP[b]	pSB1A2 with luxR under control of a constitutive Anderson promoter, GFPmut3b under control of P_{lux}	This study
pSB1A2-PconXX-lasR-Plas-GFP[b]	pSB1A2 with lasR under control of a constitutive Anderson promoter, GFPmut3b under control of P_{las}	This study
pSB1A2-PconXX-rhlR-Prhl-GFP[b]	pSB1A2 with rhlR under control of a constitutive Anderson promoter, GFPmut3b under control of P_{rhl}	This study
pSB3C5-PconXX-luxI	pSB3C5 with luxI under control of a constitutive Anderson promoter	This study
pSB3C5-PconXX-lasI	pSB3C5 with lasI under control of a constitutive Anderson promoter	This study
pSB3C5-PconXX-rhlI	pSB3C5 with rhlI under control of a constitutive Anderson promoter	This study
Sigma/anti-sigma factor circuits		
pDC296	ColE1 *ori*, Amp^R, *araC*, P_{BAD}-rsiW-RFP, P_{sigW}-sigW	Arkin Lab (University of California, Berkeley)
pSB1A7-tetR1	ColE1 *ori*, Amp^R, tetR (internal SpeI site removed), P_{tet}	This study (Sec. 2.2.1)

(Continued on next page)

Table 3.1 | *E. coli* strains and plasmids used in this chapter (continued)

Name	Relevant Genotype	Reference
Plasmids		
pSB1A7-PsigW-sigW	pSB1A7 with P_{sigW}-sigW from pDC296	This study
pSB1AK3-PsigW-sigW	pSB1AK3 with P_{sigW}-sigW from pDC296	This study
pSB1A7-tetR1-rsiW-RFP	pSB1A7-tetR1 with rsiW-RFP cassette from pDC296 cloned downstream	This study
pECS01	pSB1A7-tetR1-rsiW-RFP with P_{sigW}-sigW cloned downstream	This study

[a]http://partsregistry.org

[b]PconXX = Pcon100, Pcon56, or Pcon10

Table 3.2 | Oligonucleotides used in this chapter.

Name	5′ → 3′ Sequence[a]
BBa_iolC1_fwd1	AATCGAATTCGCGGCCGCTTCTAGAGTTCATTTATGGGAAGG
BBa_iolC1_rev1	CTGCAGCGGCCGCTACTAGTAATTCATTGACATATCC
BBa_iolD1_fwd1	AATCGAATTCGCGGCCGCTTCTAGAGCAAGATCACAGAAATG
BBa_iolD1_rev1	CTGCAGCGGCCGCTACTAGTATTTTCATGTACCCCACC
BBa_iolE_fwd1	AATCGAATTCGCGGCCGCTTCTAGAGTCAATATCGCAAGGACTATC
BBa_iolE_rev1	CTGCAGCGGCCGCTACTAGTATGGCTCCCACTTAATGAAAC
BBa_iolG2_fwd1	AATCGAATTCGCGGCCGCTTCTAGAGAAGTACCTGAGCTGGTG
BBa_iolG2_rev1	CTGCAGCGGCCGCTACTAGTAGCTTTTTCATTCGTACCTC
BBa_iolI2_fwd1	AATCGAATTCGCGGCCGCTTCTAGAGATTTCGTTGGGCCAGCG
BBa_iolI2_rev1	CTGCAGCGGCCGCTACTAGTACGATATTCATTATTTTCTCC
BBa_PiolR_fwd1	AATCGAATTCGCGGCCGCTTCTAGAGACGAAAAGAGCCAGTTCG
BBa_PiolR_rev1	ATCCTGCAGCGGCCGCTACTAGTAGTTTAGACATGCATGATAC
BBa_iolR_fwd1	AATCGAATTCGCGGCCGCTTCTAGAGGTGATTTATAAACGTCATC
BBa_iolR_rev1	CTGCAGCGGCCGCTACTAGTACTGATTAAGTTTCACCAC
Psig_fwd	GAATTCGCGGCCGCTTCTAGAGGAGATCTTCTACACCCTGCC
Psig_rev	CTGCAGCGGCCGCTACTAGTATCTGTATGTATACGAGCTTCG
rsi_fwd1	AATCGAATTCGCGGCCGCTTCTAGAGACCTCTCCCGACCCGAAAG
rsi_rev1	AATCCTGCAGCGGCCGCTACTAGTAACGAAAGGCCCAGTCTTTCG
sig_fwd1	AATCGAATTCGCGGCCGCTTCTAGAGTGAGATCTTCTACACCCTGC
sig_rev1	AATCCTGCAGCGGCCGCTACTAGTAAGTTCACCGACAAACAACAG

[a]All oligonucleotides were purchased from Sigma-Genosys (St. Louis, MO). BioBricks prefixes and suffixes are underlined.

3.2.2 Measurement of Fluorescent Protein Concentrations

The strain of *E. coli*, relevant plasmids, and culture conditions for each experiment are indicated in the figure captions. Cultures grown in LB medium were washed with TNG buffer (25 mM Tris-HCl, 0.2 M NaCl, 5% (v/v) glycerol, pH = 7.5) before fluorescence measurements to eliminate background fluorescence from the LB medium. Fluorescence and optical density measurements were taken on either a Thermo Scientific Varioskan Flash (fluorescence: 488 nm excitation/511 nm emission, optical density: 600 nm absorbance) or Tecan Infinite F200 Pro (fluorescence: 485 nm excitation/535 nm emission, optical density: 600 nm absorbance). Blank measurements were taken on uninoculated medium and subtracted from the sample measurements. Normalized fluorescence was calculated by dividing the measured fluorescence by the measured optical density at 600 nm.

3.3 Results

3.3.1 *Myo*-inositol Responsive Promoters

Although Kröger et al. observed extremely high levels of induction of several *iol* promoters in *Salmonella* in the presence of *myo*-inositol (Kröger and Fuchs, 2009), this observation is not sufficient evidence to suggest a direct interaction between *myo*-inositol and the *iol* promoters: because *Salmonella* possesses a catabolic pathway for *myo*-inositol (Fig. 3.2), it is possible that a *myo*-inositol degradation product is actually responsible for activating expression from the *iol* promoters. This type of feedback activation is a common network motif in biological circuits (Alon, 2006).

To test whether the *iol* promoters activate gene expression in response to *myo*-inositol, we cloned 300 bp sequences upstream of six *iol* genes (*iolC1*, *iolD1*, *iolE*, *iolG2*, and *iolI2*) from *Salmonella typhimurium* genomic DNA upstream of a GFP variant (Fig. 3.3). Because the mechanism of regulation was not fully understood, the gene for the regulatory protein *iolR* and its native promoter was also cloned from *Salmonella* genomic DNA. If *myo*-inositol is a direct activator of expression from the *iol* promoters, then IolR will not be necessary to observe activation. On the other hand, if expression is activated by binding of *myo*-inositol to the repressor protein, then IolR will be required for *myo*-inositol-dependent activation. GFP expression from these *iol* promoters was measured in minimal medium cultures in the presence and absence of *myo*-inositol at various temperatures (Fig. 3.4). Since these promoters are involved in carbon source utilization in *Salmonella*

Figure 3.3 | Circuit used to test *Salmonella* promoters for *myo*-inositol responsiveness.

typhimurium, catabolite repression by glucose is an important possibility. These promoters were therefore characterized in cultures containing either glucose or glycerol as the sole carbon source. In all cases, however, no significant differences in normalized fluorescence were observed upon addition of *myo*-inositol to the culture medium, regardless of the presence or absence of the regulator IolR.

It has been shown that 2-deoxy-5-keto-D-gluconic acid 6-phosphate (DKG-6P), an intermediate of *myo*-inositol degradation, prevents IolR binding to *iol* promoters in *Bacillus subtilis* (Yoshida et al., 2008). Kröger et al. therefore speculate DKG-6P could be responsible for activation of expression from the *iol* promoters in *Salmonella* as well (Kröger and Fuchs, 2009). Given that the *E. coli* strains in which these promoters were characterized do not possess genes for *myo*-inositol catabolism, this data provides further evidence for *iol* promoter activation via a *myo*-inositol catabolite rather than *myo*-inositol itself. Due to these observations, we did not continue to pursue the *Salmonella* promoters.

3.3.2 Genetic Timers

Quorum Sensing

Quorum sensing machinery has been discovered in several different organisms, including *Vibrio fischeri*, *Vibrio harveyi*, and *Pseudomonas aeruginosa* (Fuqua et al., 1994, Pesci et al., 1997). In *Vibrio fischeri*, the quorum sensing machinery consists of two components: a signaling molecule synthase (LuxI) and a transcriptional regulator (LuxR) which responds to the presence of the signaling molecule (Fig. 3.5). In the natural system, low-level expression of LuxI (red ovals) leads to production of the signaling molecule N-3-oxo-hexanoyl-homoserine lactone (3-oxo-C_6 HSL, yellow circles), which is capable of freely diffusing into and out of the cell. When cell concentration is low, 3-oxo-C_6 HSL diffuses out of the cell into the extracellular medium, resulting in a low intracellular concentration of 3-oxo-C_6 HSL. However, as cell concentration begins to rise, the level of 3-oxo-C_6 HSL begins to rise both intracellularly and in the extracellular medium. Eventually, buildup of intracellular 3-oxo-C_6 HSL will lead to dimerization of LuxR (blue ovals) and activation

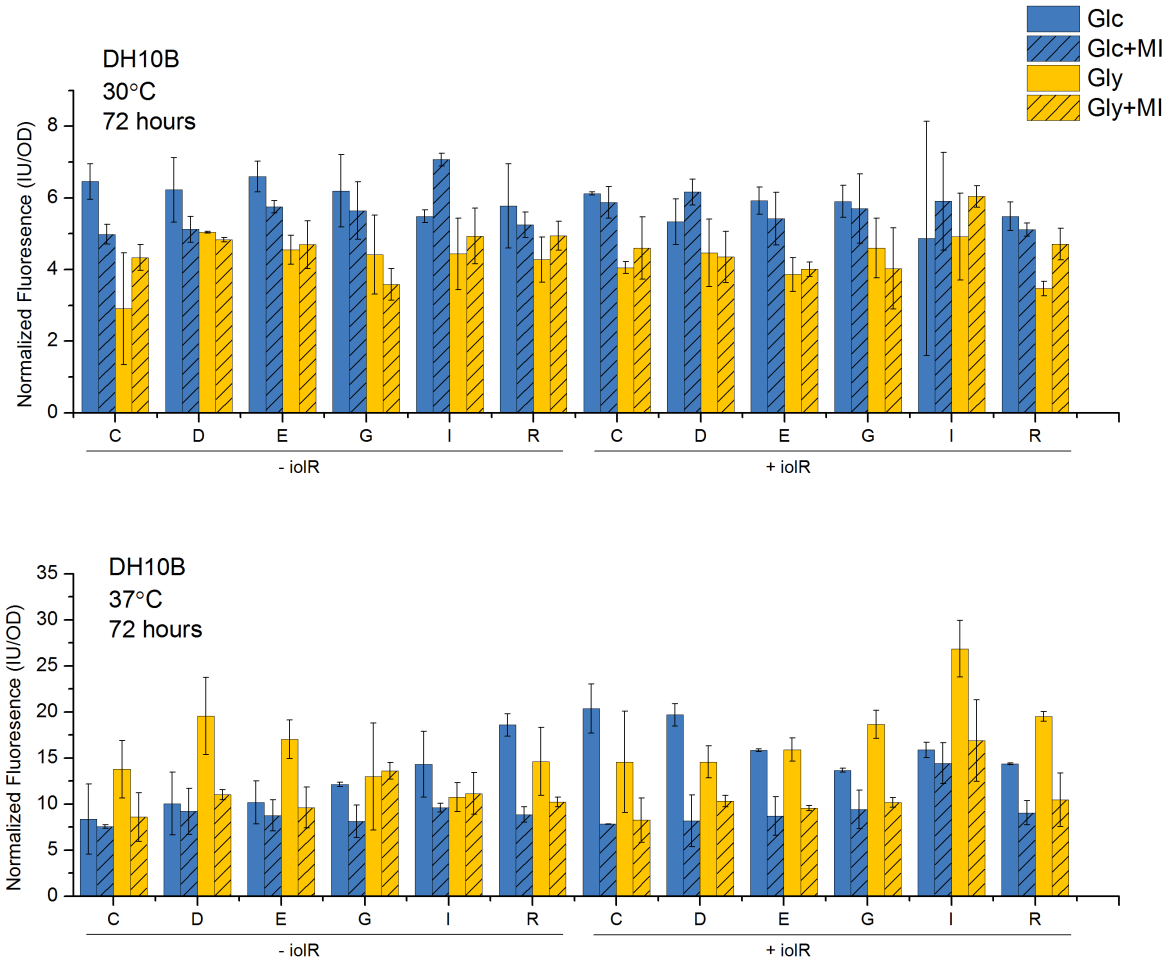


Figure 3.4 | Normalized GFP fluorescence from various *Salmonella* promoters. In all cases, expression is low, and there does not appear to be any significant effect from the presence of *myo*-inositol. Cultures contained strain DH10B harboring pSB1A2-PiolX-GFP (- iolR) and pSB1A2-iolR0-PiolX-GFP (+ iolR) and were grown in M9 minimal medium supplemented with glucose (Glc), glycerol (Gly), and *myo*-inositol (MI) as indicated. Error bars represent the standard deviation of biological triplicates.

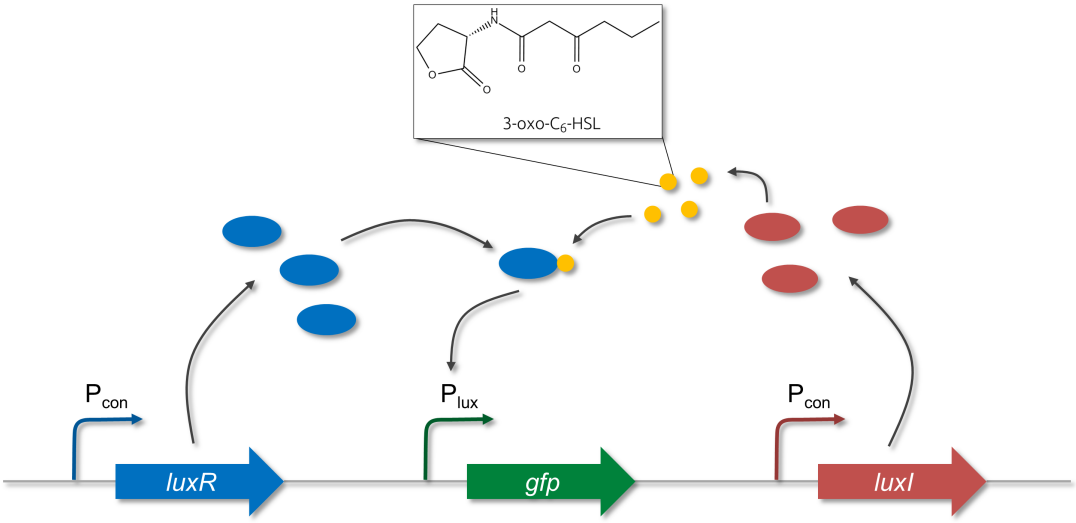


Figure 3.5 | Quorum sensing circuit tested for delayed induction behavior. P_{con}: constitutive promoter, P_{lux}: LuxR-responsive promoter.

of gene expression from a quorum-controlled promoter (P_{lux}).

We attempted to characterize the dynamics of the *Vibrio fischeri* quorum sensing circuit by expressing LuxI and LuxR from constitutive promoters and using this system to control GFP expression from a quorum-controlled promoter (Fig. 3.5). We also explored the *lasI/lasR* and *rhlI/rhlR* quorum sensing systems from *Pseudonomas aeruginosa* because these systems were easily obtainable from the Registry of Standard Biological Parts (http://partsregistry.org).

The necessity for 3-oxo-C_6 HSL to accumulate intracellularly in order to activate gene expression results in an inherent time delay in gene expression which is directly linked to the 3-oxo-C_6 HSL production rate. Therefore, it should be possible to increase or decrease the circuit's time delay simply by increasing or decreasing the expression level of the synthetase. We modified the homoserine lactone synthetase expression level using constitutive promoters from the Anderson promoter library (Registry of Standard Biological Parts) to determine whether the timing of induction can be controlled and measured GFP concentration as a function of time. A representative graph is shown in Fig. 3.6; the remaining graphs can be found in Appendix A.

From Fig. 3.6 and Appendix A, it is apparent that quorum sensing machinery is able to delay protein expression appreciably post-inoculation for a variety of synthetase and regulator expression levels. However, the trend is not what is expected: time delay increases as synthetase promoter strength is increased. This trend

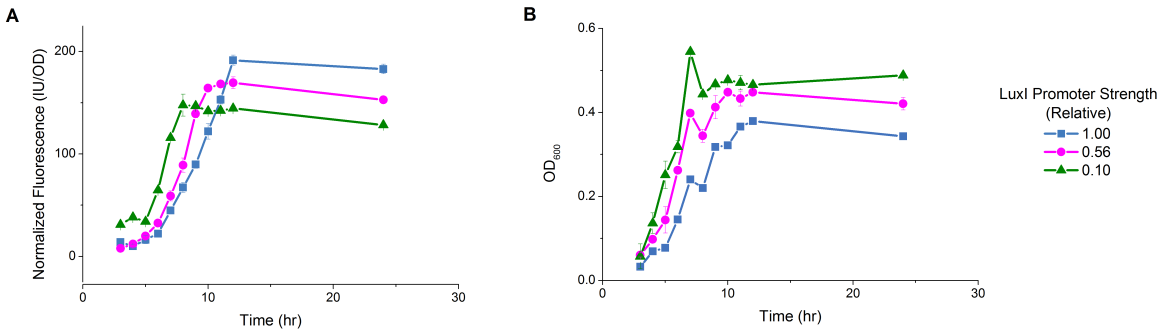


Figure 3.6 | Characterization of the *lux* quorum sensing circuit: **(A)** GFP concentration and **(B)** OD_{600} as a function of time for various LuxI expression levels. LuxR promoter strength: 0.56. Error bars represent the standard deviation of biological triplicates.

reversal is likely due to increased metabolic burden from expression of the synthetase, as the synthetase draws precursors from fatty acid biosynthesis. Metabolic burden is evident in the growth curves as well, since increasing expression of the synthetase results in a longer lag time, a slower growth rate, and a lower final cell density. In other cases (Appendix A), no trend is apparent; we believe that metabolic burden effects from overexpression of either the synthetase or the repressor protein are responsible for this unpredictable behavior.

To reduce metabolic burden, the AHL synthetase was integrated into the *E. coli* genome at the *gudD* locus. GFP and the transcriptional regulator were also moved to a lower copy plasmid (p15A *ori*) to further reduce metabolic burden. This integration was sufficient to eliminate reductions in growth rate due to overexpression of the synthetase; however, the reduction in copy number of the synthetase gene also significantly reduced expression of the reporter gene (Appendix A, Fig. A.9). Attempts to amplify this signal using T7 RNA polymerase (Fig. 3.7) were also unsuccessful (data not shown).

Sigma/Anti-Sigma Factor Pairs

The *sigW*/*rsiW* sigma/anti-sigma factor pair from *Bacillus subtilis* has been studied and used by the Arkin Lab at the University of California, Berkeley (Chen and Arkin, 2012). While *rsiW* has not been shown explicitly to sequester *sigW*, there is strong evidence that *rsiW* is the anti-sigma factor for *sigW* (Schöbel et al., 2004). This pair forms the basis of the genetic timer circuit shown in Fig. 3.8. Here, anti-sigma factor *rsiW* and reporter gene *rfp* are expressed from P_{BAD}, an L-arabinose-inducible promoter. Sigma factor *sigW* is expressed from P_{sig}, a *sigW*-activated promoter with a consensus *E. coli* UP element. Assuming similar degradation rates

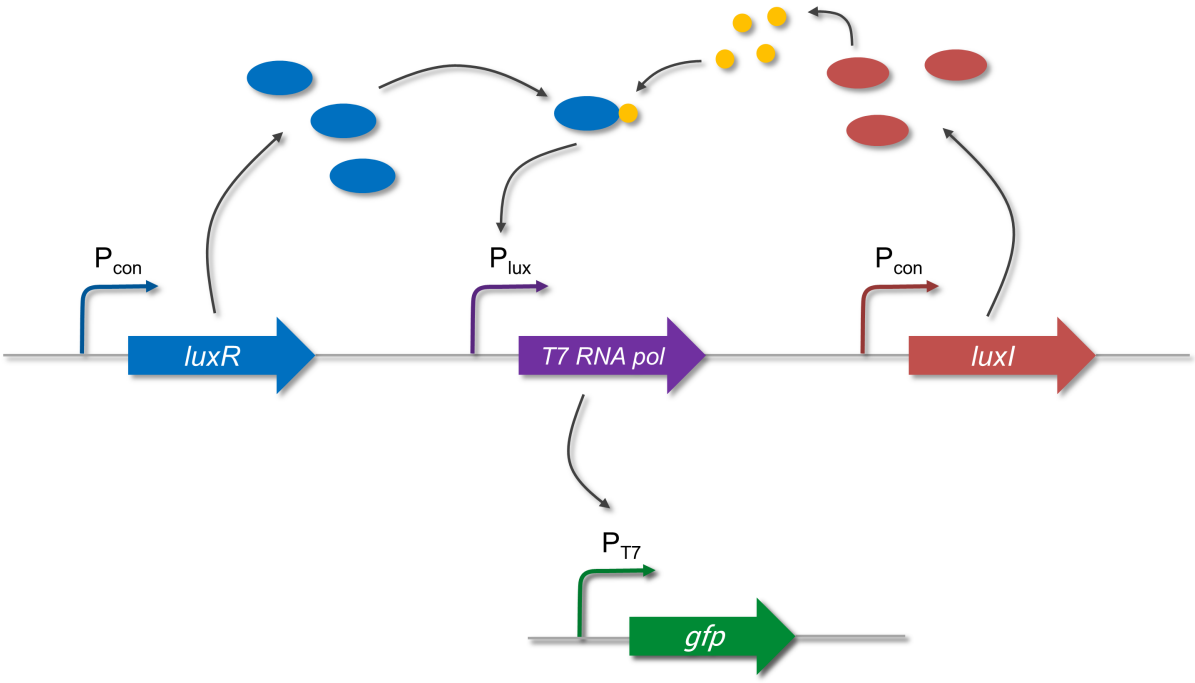


Figure 3.7 | Circuit for quorum sensing signal amplification through T7 RNA polymerase.

of RFP and RsiW, RFP serves as a readout of the RsiW concentration within the cell. The sensitivity of P_{sigW} to SigW creates a positive feedback loop, which ensures maintenance of the "ON" state once activated. The reporter gene (*gfp*) is expressed from the same SigW-sensitive P_{sigW} promoter and serves as a readout for circuit behavior. To "initialize" the circuit, starter cultures are grown in the presence of high levels of L-arabinose (0.4 wt%), resulting in high levels of RsiW and low levels of GFP, corresponding to an "OFF" state. The starter culture is then inoculated into flasks containing varying concentrations of L-arabinose (0 - 0.2 wt%). Because L-arabinose concentrations in the flask are lower than in the starter culture, there is a net decrease in intracellular RsiW concentration. Additionally, lower concentrations of L-arabinose result in slower rates of RsiW synthesis, leading to a faster decline in RsiW concentration. Once intracellular RsiW has decreased past a threshold concentration, SigW is freed to activate expression of the reporter protein. Timing of induction is thus controlled by the amount of L-arabinose added to the culture flasks: lower L-arabinose concentrations should translate to an earlier onset of induction.

As expected, the time delay of the circuit correlated well with L-arabinose concentration (Fig. 3.9A): as L-arabinose concentration was increased, the time delay was also increased. RFP concentrations, which

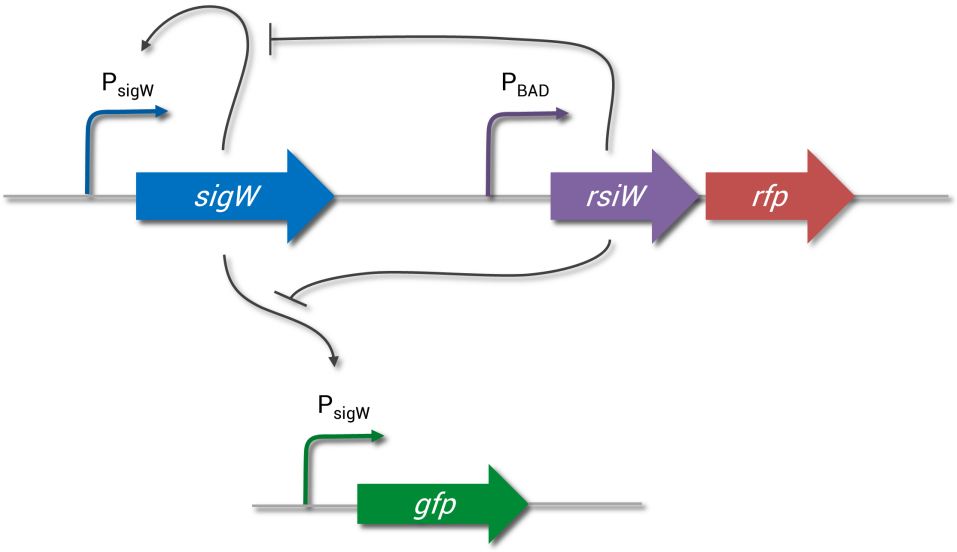


Figure 3.8 | Sigma/anti-sigma factor circuit tested for delayed induction behavior.

served as a proxy for RsiW concentrations, decayed as expected as well: as L-arabinose concentration was increased, the rate of decay decreased. Although this was an encouraging result, the conditions under which the circuit was tested (M9 minimal medium supplemented with glycerol, L-arabinose induction) do not match the typical conditions for D-glucaric acid production (LB medium supplemented with glucose, IPTG induction). In particular, the fact that glucose is required for D-glucaric acid production poses a challenge, since glucose-activated catabolite repression precludes L-arabinose uptake.

To address this issue, the P_{BAD} promoter was exchanged for P_{tet}, an aTc-inducible promoter and the circuit re-characterized in LB medium with aTc induction (Fig. 3.10). Unfortunately, these modifications eliminated both the induction and time delay behavior of the circuit. Because the original sigma/anti-sigma factor circuit was tuned for minimal medium, this result is not altogether surprising. We speculate that the increased growth rate in rich medium and the positive feedback loop controlling the sigma factor both contribute to the unpredictable behavior observed in Fig. 3.10.

3.4 Discussion

Autonomous, time-delayed induction of gene expression without the need for exogenous inducers would be an important tool for industrial bioproduction of commodity-scale chemicals, as profit margins are thin

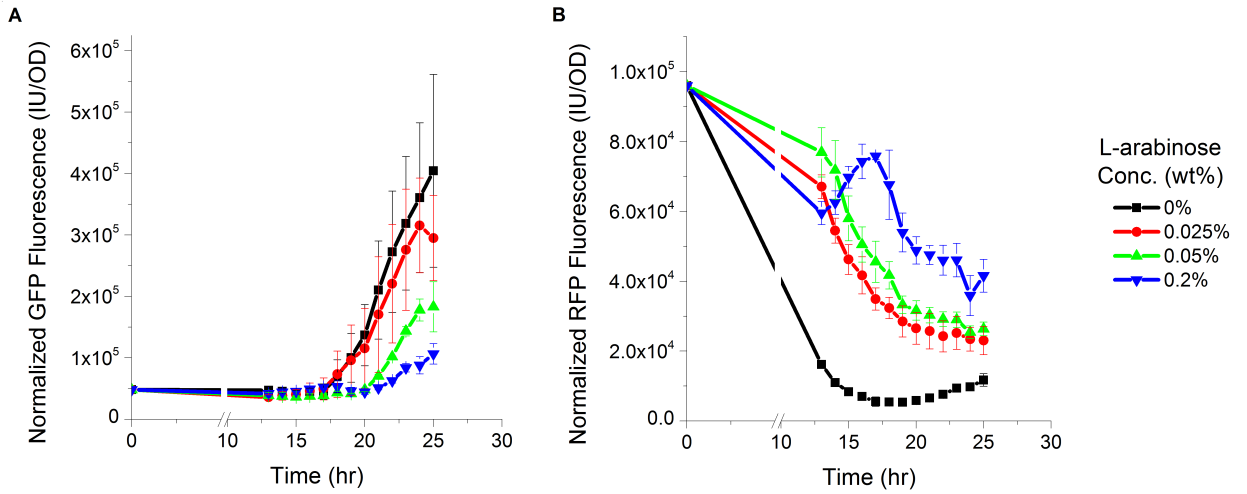


Figure 3.9 | Characterization of a sigma/anti-sigma factor circuit: **(A)** normalized GFP and **(B)** normalized RFP fluorescence as a function of time. Cultures were grown in M9 minimal media supplemented with 10 g/L glycerol and L-arabinose as indicated; starter cultures were grown in M9 minimal media with 10 g/L glycerol and 0.4 wt% L-arabinose for circuit "initialization." Error bars represent the standard deviation of biological triplicates.

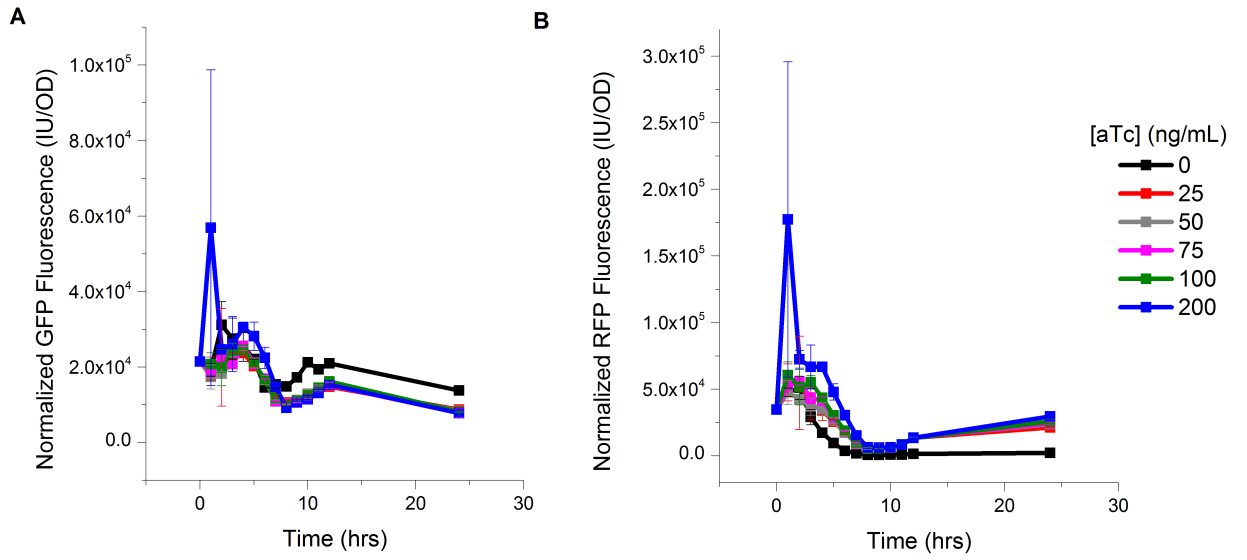


Figure 3.10 | Characterization of a sigma/anti-sigma factor circuit with aTc induction: **(A)** normalized GFP fluorescence and **(B)** normalized RFP fluorescence as a function of time. Cultures were grown in LB at 37°C with varying concentrations of aTc. Error bars represent the standard deviation of biological triplicates.

and inducers are generally an expensive medium component. In this work, we attempted to develop genetic timers using several different circuit architectures in order to improve D-glucaric acid production in *E. coli*. Three circuit architectures were explored:

1. *Myo*-inositol responsive promoters from *Salmonella typhimurium*
2. Quorum sensing machinery from *Vibrio fischeri* and *Pseudomonas aeruginosa*
3. A sigma/anti-sigma factor pair from *Bacillus subtilis*

Each of these circuit architectures did not perform as expected for different reasons. We believe that the *Salmonella* promoters do not respond directly to *myo*-inositol but instead to a downstream catabolite in the *myo*-inositol degradation pathway. This type of feedback circuit architecture is quite common in biological systems and serves to filter out short-term noise in the input signal: if gene expression were activated directly by *myo*-inositol, even transient increases in *myo*-inositol levels would activate gene expression, leading to wastage of cellular resources. On the other hand, the presence of a feedback activation loop requires that high levels of *myo*-inositol be maintained for an extended period of time, as the signal must travel through several intermediate steps before gene expression is upregulated.

Although this particular set of promoters was not responsive to *myo*-inositol, it is highly likely that a *myo*-inositol-responsive promoter and regulatory protein exists in nature. *Myo*-inositol is used as a building block for cell wall biosynthesis in a number of eukaryotic organisms, including *Saccharomyces cerevisiae*. Because of its function as a key metabolite, it is likely that *myo*-inositol levels are highly regulated, and it is reasonable to hypothesize that this regulation may be dependent on *myo*-inositol in some cases. Comparative transcriptomic analysis of these organisms in the presence and absence of *myo*-inositol could lead to the identification of promoters which are *myo*-inositol-responsive.

Metabolic burden and burden-related effects proved to be the limitation of the quorum sensing circuits. Overexpression of either the AHL synthetase or the quorum-responsive regulator led to slower growth rates and a lower final cell density, classical indicators of metabolic burden. Because the acyl group of the acyl homoserine lactone is derived from fatty acid biosynthesis, we believe that overexpression of the AHL synthetase deprives the cell of precursors for cell wall formation and forces the cell to expend more energy to maintain homeostasis. These metabolic burden effects then lead to the unpredictable behavior that we have observed.

If quorum sensing is to be applied to metabolic engineering problems, these metabolic burden issues must be addressed, as biological production of small molecules almost always requires overexpression of pathway enzymes in addition to the quorum sensing machinery. Efforts to eliminate metabolic burden effects by reducing expression level of the AHL synthetase and quorum-responsive regulator were relatively successful; however, these efforts also significantly reduced the expression level of the gene of interest. It is possible that, by reducing expression of the circuit components, the output range of the first circuit component no longer matched the input range of the second circuit component, and further impedance matching (Sec. 1.3.3) and circuit re-tuning may be required to generate a functional circuit.

The sigma/anti-sigma factor circuit was plagued by context dependence issues (Sec. 1.3.3): although the circuit performed as expected in minimal medium with arabinose induction, a simple switch to rich medium and aTc induction resulted in completely unpredictable circuit behavior. It is likely that the increased growth rate in rich medium leads to changes in the synthesis rate of circuit components, and these effects could be compounded by the presence of an autoregulatory positive feedback loop controlling *sigW* expression. This particular network motif has been shown to potentially lead to bistability (Alon, 2006). Elimination of this positive feedback loop did not restore time-delayed induction behavior, however (data not shown), so re-tuning of the circuit may be necessary.

3.5 Conclusions

In this work, we demonstrated the potential of time-delayed induction of MIOX for improving D-glucaric acid production through a simple proof-of-concept experiment involving delayed, manual addition of chemical inducers to the culture medium. Since chemical inducers are too costly for commodity-scale chemicals, we attempted to develop genetic circuits which would be capable of autonomously delaying the induction of gene expression in a controllable fashion. While unsuccessful, these attempts highlight the limitations which currently prevent widespread use of synthetic biology devices for metabolic engineering problems. It is hoped that the lessons learned here will help drive future development of synthetic biology devices for metabolic engineering applications.

Chapter 4

Improved Yield via Strain Engineering

Abstract

Chapters 2 and 3 focused on the application of metabolic engineering and synthetic biology strategies towards increasing productivity of D-glucaric acid. At the industrial scale, product yield is also an important process parameter, especially for low-margin products such as D-glucaric acid, where raw material costs can be a significant fraction of the overall manufacturing costs. Recent interest in using lignocellulosic biomass as a feedstock for microbial fermentation processes presents an opportunity for increasing the yield of bioproducts derived directly from D-glucose, such as D-glucaric acid. Lignocellulosic biomass consists of several fermentable sugars, including glucose, xylose, and arabinose. In this chapter, we investigate the ability of an *E. coli* Δ*pgi* Δ*zwf* mutant to consume alternative carbon sources (xylose, arabinose, and glycerol) for growth while reserving glucose for product formation. Deletion of *pgi* and *zwf* was found to eliminate catabolite repression as well as *E. coli*'s ability to consume glucose for biomass formation. In addition, productivity and yield of D-glucaric acid from D-glucose were significantly increased in a Δ*pgi* Δ*zwf* strain.

Portions of this chapter are being prepared for publication in:

Shiue, E and Prather, KLJ. Improving product yields on D-glucose in *Escherichia coli* via alternative carbon sources.

4.1 Introduction

In recent years, concerns over declining petroleum reserves and climate change due to atmospheric carbon dioxide accumulation have spurred significant interest in using alternative feedstocks for the manufacture of petroleum-derived products. Renewable and abundant, non-food plant (lignocellulosic) biomass represents a promising alternative feedstock to crude oil. Moreover, because plants take up carbon dioxide during growth, the use of plant-based feedstocks could potentially slow the accumulation of carbon dioxide in the atmosphere. Recent research has focused heavily on the identification of ideal plant biomass feedstocks (Joyce and Stewart, 2012), the determination of commercially valuable, biomass-derived products (Werpy and Petersen, 2004), and the development of processes for converting plant biomass into these products of interest.

One such process which has received heavy attention is microbial fermentation. A significant portion of lignocellulosic biomass consists of fermentable sugars such as glucose, xylose, arabinose, and galactose (Joyce and Stewart, 2012), and many microbes are naturally able to convert these sugars into products of interest such as biofuels (Jang et al., 2012) and biopolymers (Lee, 1996). In addition, microbes can be engineered to produce a wide array of non-natural products via recombinant DNA technology (Curran and Alper, 2012). A few products, including D-glucaric acid (Moon et al., 2009b) and D-gluconic acid (Rogers et al., 2006), are derived directly from glucose; however, microbial production of these products generally suffers from low yields, as a portion of the glucose feed is utilized for generation of cell biomass. To maximize yields for these glucose-derived products, we set out to design an *E. coli* production platform which utilizes an alternative carbon source such as arabinose or xylose for cell growth, reserving glucose solely for product generation.

The main pathways for D-glucose utilization in *E. coli* are depicted in Fig. 4.1. D-glucose enters the cell through the phosphotransferase system (PTS) encoded by *ptsG* and *ptsHI-crr* and is phosphorylated to glucose 6-phosphate in the process. Glucose 6-phosphate can then proceed through the Entner-Dudoroff Pathway via *zwf* or through the Embden-Meyerhoff-Parnas Pathway via *pgi*. The carbon in glucose-6 phosphate can also be stored as glycogen via *pgm*.

Previous work to engineer *E. coli* for coutilization of D-glucose and alternative carbon sources involved deletion of the PTS system (Solomon et al., 2013, Balderas-Hernández et al., 2011, Wang et al., 2011c). By eliminating the PTS system, catabolite repression can be eliminated, allowing simultaneous uptake of D-glucose and another carbon source. However, these strategies also eliminate *E. coli's* primary method of

glucose uptake, and the cell must rely on nonspecific transporters to import glucose into the cell. Subsequent phosphorylation of D-glucose to glucose 6-phosphate via ATP-dependent *glk* is also required for D-glucose metabolism in PTS-deficient E. coli. Overexpression of the galactose:H^+ symporter *galP* and upregulation of *glk* has been shown to recover wild-type growth rates in PTS-deficient strains of *E. coli* (Balderas-Hernández et al., 2011). However, because the glycolytic pathways remain in this strain, it is likely that product yields on D-glucose would remain low. In this study, we explore the behavior of an *E. coli* strain which lacks *pgi* and *zwf* and investigate its ability to produce a D-glucose-derived product when supplemented with L-arabinose and D-xylose, sugars which are readily available from biomass. Glycerol is also explored as an alternative carbon source, as its price has dropped greatly in recent years due to significant increases in biodiesel production (Johnson and Taconi, 2009). Previous work has demonstrated improved productivity and yield of D-glucose-derived products in a Δ*pgi* Δ*zwf* strain supplemented with mannitol (Kogure et al., 2007, Pandey et al., 2013); however, the price of mannitol remains high relative to glycerol and biomass-derived sugars.

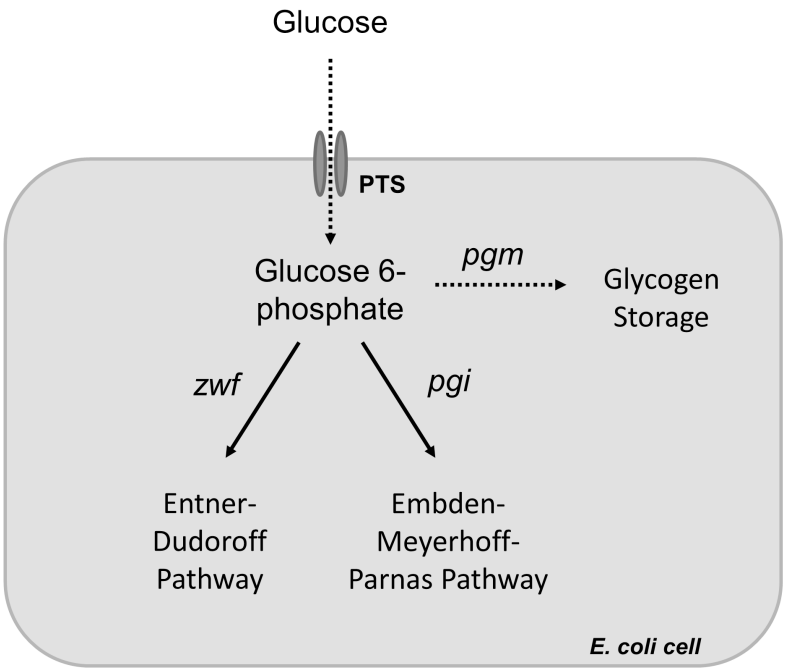


Figure 4.1 | D-glucose utilization pathways in *E. coli*.

4.2 Materials and Methods

4.2.1 *E. coli* Strains and Plasmids

E. coli strains and plasmids used in this study are listed in Table 4.1. All molecular biology manipulations were performed according to standard practices (Sambrook and Russell, 2001). *E. coli* DH10B was used for transformation of cloning reactions and propagation of all plasmids. Construction of strains M2 and M2-2 is described in Sec. 2.2.1. Strain M3 was constructed by our group previously (Gonçalves et al., 2013). Deletion of *zwf* from strain M3 was achieved by P1 transduction with Keio collection strain JW1841-1 as the donor (Baba et al., 2006). The λDE3 lysogen was then integrated site-specifically into this quadruple knockout strain using a λDE3 Lysogenization Kit (Novagen, Darmstadt, Germany), generating strain M4 (MG1655(DE3) Δ*endA* Δ*recA* Δ*pgi* Δ*zwf*. To eliminate the native ability of *E. coli* MG1655 to consume D-glucuronic or D-glucaric acids, *uxaC* and *gudD* were also deleted from the genome. Deletion of *uxaC* was performed with λ-Red mediated recombination (Datsenko and Wanner, 2000) using pKD46recA (Solomon et al., 2013). PCR primers pKD13_uxaC_fwd and pKD13_uxaC_rev (Table 4.2) were used to amplify the recombination cassette from pKD13 (Datsenko and Wanner, 2000), and strain M4 harboring pKD46recA was transformed with this PCR product. The *kan* selection cassette was cured from successful deletion mutants using FLP recombinase expressed from pCP20, generating strain M5. Finally, strain M6 (MG1655(DE3) Δ*endA* Δ*recA* Δ*pgi* Δ*zwf* Δ*uxaC* Δ*gudD*) was generated using the same λ-Red mediated recombination method described above; in this case, primers pKD13_gudD_fwd and pKD13_gudD_rev were used to amplify the recombination cassette from pKD13. To construct pRSFD-IN-Udh, pRSFD-IN was first with XhoI, end-filled with Klenow enzyme, then digested with EcoRI-compatible MfeI. pTrc-Udh was then digested with EcoRI and SmaI, and the Udh-containing fragment was ligated into digested pRSFD-IN to generate pRSFD-IN-Udh.

Table 4.1 | *E. coli* strains and plasmids used in this chapter.

Name	Relevant Genotype	Reference
Strains		
DH10B	F^- mcrA Δ(mrr-hsdRMA-mcrBC) ϕ80lacZΔM15 ΔlacX72 recA1 endA1 araD139 Δ(ara,leu)7697 galU galK λ-rpsL nupG	Life Technologies (Carlsbad, CA)
JW1841-1	F^-, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ^-, Δzwf777::kan, rph-1, Δ(rhaD-rhaB)568, hsdR514	CGSC #9031 (Baba et al., 2006)
MG1655	F^- λ^- ilvG$^-$ frb-50 rph-1	CGSC #6300
M2	MG1655(DE3) Δ*endA* Δ*recA*	This study (Sec. 2.2.1)
M2-2	MG1655(DE3) Δ*endA* Δ*recA* Δ*gudD* Δ*uxaC*	This study (Sec. 2.2.1)
M3	MG1655 Δ*endA* Δ*recA* Δ*pgi*	(Gonçalves et al., 2013)
M4	MG1655(DE3) Δ*endA* Δ*recA* Δ*pgi* Δ*zwf*	This study
M5	MG1655(DE3) Δ*endA* Δ*recA* Δ*pgi* Δ*zwf* Δ*uxaC*	This study
M6	MG1655(DE3) Δ*endA* Δ*recA* Δ*pgi* Δ*zwf* Δ*uxaC* Δ*gudD*	This study
Plasmids		
pCP20	Rep[a], Amp^R, Cm^R, FLP recombinase expressed by λp_r under control of λcI857	CGSC #7629
pKD13	R6Kγ *ori*, Amp^R, *kan*	CGSC #7633
pKD46	R101 *ori*, *repA101*[a], Amp^R, *araC*, *araBp*-γ_γ-γ_β-γ_{exo}	CGSC #7739
pKD46recA	R101 *ori*, *repA101*[a], Amp^R, *araC*, *araBp*-γ_γ-γ_β-γ_{exo}, *recA*	(Solomon et al., 2013)
pRSFDuet-1	pRSF1030 *ori*, lacI, Kan^R	EMD4Biosciences (Darmstadt, Germany)
pTrc99A	pBR322 *ori*, Amp^R	(Amann and Brosius, 1985)
pRSFD-IN	pRSFDuet-1 with INO1 inserted into the EcoRI and HindIII sites	(Moon et al., 2009b)

(Continued on next page)

Table 4.1 | *E. coli* strains and plasmids used in this study (continued)

Name	Relevant Genotype	Reference
Plasmids		
pTrc-Udh	pTrc99A with Udh from *Pseudomonas syringae* inserted into the NcoI and HindIII sites	(Moon et al., 2009b, Yoon et al., 2009)
pRSFD-IN-Udh	pRSFD-IN with Udh inserted into the MfeI and XhoI sites	This study
pTrc-SUMO-MIOX	pTrc99A with SUMO-MIOX	This study (Sec. 2.2.1)

[a]Temperature-sensitive

Table 4.2 | Oligonucleotides used in this chapter.

Name	5′ → 3′ Sequence[a]
pKD13_gudD_fwd	TCCCCGGCTGGACCTTTGACCGTAAACGTCCCGTTTTCGGCCGTCATTGATTCTGAAAAAGGACATAAATCTGTCAAACATGAGAATTAATTCC
pKD13_gudD_rev	CAACAGGCTATTTTGCGTTTAGCATCAGTCTCAAACCGGCTCCAGATAGAGCCGGTTTTGGTTTTCTGTCGTGTAGGCTGGAGCTGCTTC
pKD13_uxaC_fwd	AATTGGTGTGATAACTTTGTCAGCATCGCACCATAAGCAAGCTAGCTCACTCGTTGAGAGGAAGACGAAACTGTCAAACATGAGAATTAATTCC
pKD13_uxaC_rev	AAATCTGCTAAAGCGACCGCGACGTTATCCAGCGCATGGATCTTGATGTATTGCATATCAACCCCAGACCGTGTAGGCTGGAGCTGCTTC

[a]All oligonucleotides were purchased from Sigma-Genosys (St. Louis, MO). Homologous sequences used for recombination are underlined.

4.2.2 Culture Conditions

For determination of growth curves, cultures were grown in 250 mL baffled shake flasks containing 50 mL LB medium supplemented with approximately 10 g/L D-glucose, L-arabinose, glycerol, and/or D-xylose as indicated in Figs. 4.2-4.4. Seed cultures were grown overnight at 30°C and inoculated to an optical density at 600 nm (OD_{600}) of 0.005. Cultures were incubated at 30°C, 250 rpm, and 80% relative humidity for 72 hours. To construct a growth curve, cell density was measured at regular time intervals on a DU800 Spectrophotometer (Beckman Coulter, Pasadena, CA), with more frequent sampling during the exponential growth phase. For analysis of metabolite concentrations, samples were taken daily, centrifuged to remove cell debris, and the supernatants analyzed via high performance liquid chromatography as described in Sec. 4.2.3.

For D-glucaric acid production, cultures were grown in 250 mL baffled shake flasks containing 50 mL LB medium supplemented with 10 g/L D-glucose and 10 g/L L-arabinose, 10 g/L glycerol, or 10 g/L D-xylose. Cultures were induced at inoculation with 0.1 mM β-D-1-thiogalactopyranoside (IPTG). Ampicillin (100 µg/mL) and kanamycin (30 µg/mL) were added for plasmid maintenance. Seed cultures were grown overnight at 30°C in LB medium supplemented with 10 g/L D-glucose and 10 g/L L-arabinose, 10 g/L glycerol, or 10 g/L D-xylose and inoculated to an optical density at 600 nm (OD_{600}) of 0.005. Cultures were incubated at 30°C, 250 rpm, and 80% relative humidity for 72 hours. Samples were taken daily, centrifuged to remove cell debris, and the supernatants analyzed for metabolite concentrations as described in Sec. 4.2.3.

4.2.3 Determination of Metabolite Concentrations

D-glucose, L-arabinose, glycerol, D-xylose, and D-glucaric acid were quantified from culture supernatants using high performance liquid chromatography (HPLC) on an Agilent Series 1100 or Series 1200 instrument equipped with an Aminex HPX-87H column (300 mm by 7.8 mm; Bio-Rad Laboratories, Hercules, CA). Sulfuric acid (5 mM) was used as the mobile phase at 35°C and a flow rate of 0.6 mL/min in isocratic mode. Compounds were detected and quantified from 10 µL sample injections using a refractive index detector. Reported metabolite concentrations are the average of triplicate samples, and error bars represent one standard deviation above and below the mean value.

4.3 Results

In *E. coli*, glucose is imported into the cell and phosphorylated to glucose 6-phosphate (G6P) by the phosphotransferase system (PTS). Glucose metabolism then proceeds through two routes (Fig. 4.1): the Embden-Meyerhoff-Parnas Pathway via phosphoglucose isomerase (*pgi*) or the Entner-Dudoroff Pathway via glucose 6-phosphate dehydrogenase (*zwf*). A third route interconverts glucose 6-phosphate and glucose 1-phosphate via phosphoglucomutase (*pgm*) for glycogen storage and accumulation, but this route does not lead to glucose consumption. To eliminate native consumption of glucose, both *pgi* and *zwf* were deleted from an MG1655-derived strain.

4.3.1 Behavior of the Δ*pgi* Δ*zwf* Mutant

Cell growth was compared for the host strain MG1655(DE3) Δ*endA* Δ*recA* (strain M2) and the derived strain MG1655(DE3) Δ*endA* Δ*recA* Δ*pgi* Δ*zwf* (strain M4) (Fig. 4.2). Maximum specific growth rates for each combination of strain and carbon supplement were also calculated (Table 4.3). As expected, growth of strain M2 was similar for all conditions tested, with similar lag phases and maximum specific growth rates. Final cell densities were lower when strain M2 was fed D-glucose, likely due to increased production of acetate (Fig. 4.3), which has been shown to inhibit cell growth (Roe et al., 1998). Growth of strain M4 was also similar for all conditions tested with the exception of glycerol-supplemented cultures, which displayed a significant lag in growth around 24 hours of culture. We speculate that this lag corresponds to depletion of the metabolic precursors provided by the LB medium and a metabolic shift towards gluconeogenic metabolism for growth on glycerol. Maximum growth rate, lag time, and final cell densities are similar for strain M4 in the presence of either L-arabinose or D-xylose regardless of whether D-glucose was supplemented in the growth media, indicating that substrate consumption was similar regardless of the presence of D-glucose.

Concentrations of D-glucose, alternative carbon source, and acetate were measured for each strain/carbon supplement combination as a function of time (Fig. 4.3). As expected, the presence of D-glucose prevents consumption of the alternative carbon source in strain M2 via catabolite repression. In contrast, the deletion of *pgi* and *zwf* prevents consumption of D-glucose in strain M4. Interestingly, deletion of *pgi* and *zwf* appears to eliminate catabolite repression in strain M4, as the presence of D-glucose does not preclude consumption of the alternative carbon source in this strain. Additionally, acetate production by strain M4 is lower than strain M2 for all of the alternative carbon sources tested, likely due to reduced glycolytic flux.

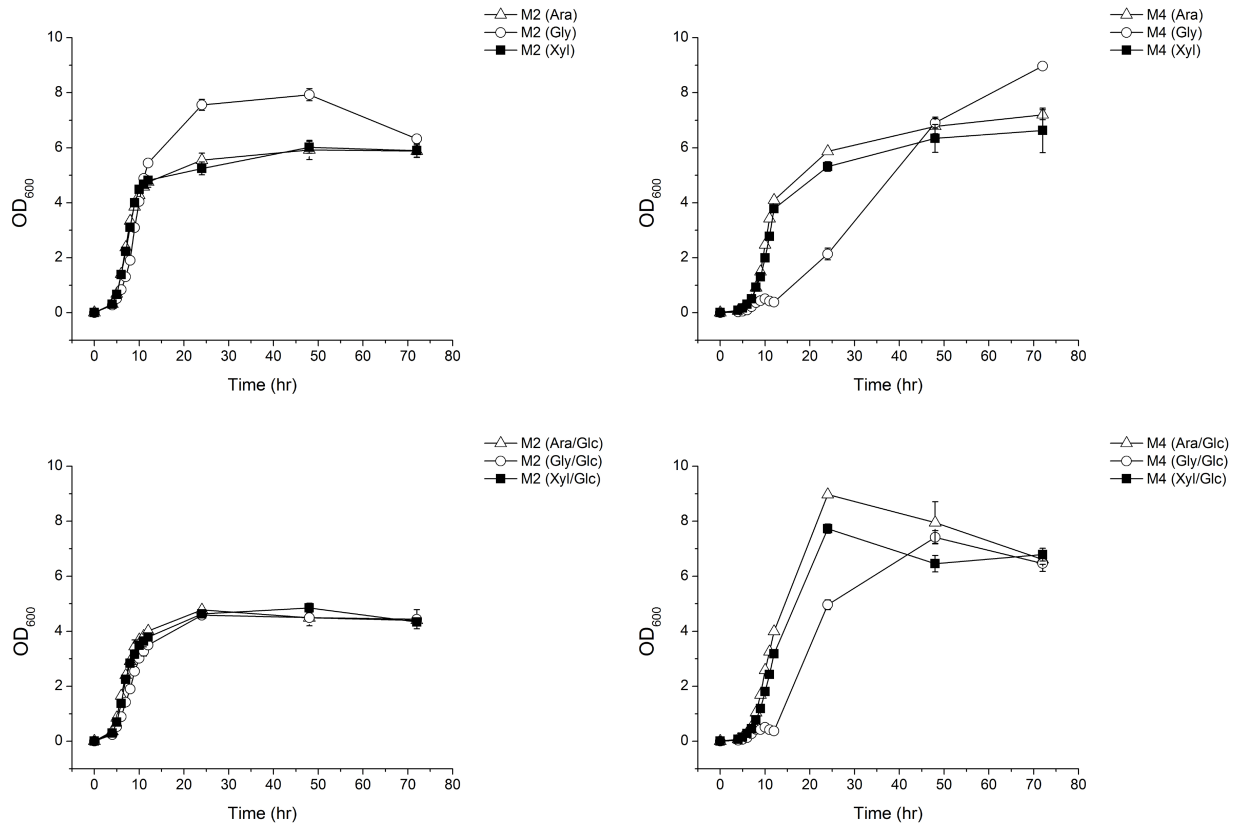


Figure 4.2 | Growth curves for strains M2 and M4 in LB medium supplemented with various carbon sources (Ara: L-arabinose, Gly: glycerol, Xyl: D-xylose, Glc: D-glucose). Error bars represent the standard deviation of biological triplicates.

We speculate that this phenomenon is due to intracellular buildup of glucose 6-phosphate. Catabolite repression is mediated by cyclic AMP (cAMP), which is synthesized by adenylate cylcase. Adenylate cylcase is activated via phosphorylation by $EIIA^{Glc}$, but this phosphorylation can only occur if $EIIA^{Glc}$ has a phosphate group to donate. When glucose is being actively imported through the PTS system, $EIIA^{Glc}$ donates its phosphate to the incoming glucose, resulting in a mostly unphosphorylated population of $EIIA^{Glc}$, inactive adenylate cyclase, and a low concentration of cAMP. The absence or depletion of glucose from the culture medium leads to a buildup of phosphorylated $EIIA^{Glc}$, activation of adenylate cyclase, and an increase in cAMP concentration, eventually leading to the expression of catabolite-repressed genes such as *araBAD* and *xylAB*. Thus, catabolite repression occurs when there is active flux of glucose into the cell, not simply when glucose is present in the medium. In a Δ*pgi* Δ*zwf* mutant, the accumulation of glucose 6-phosphate quickly eliminates glucose flux into the cell, resulting in derepression of genes normally repressed in the presence of glucose.

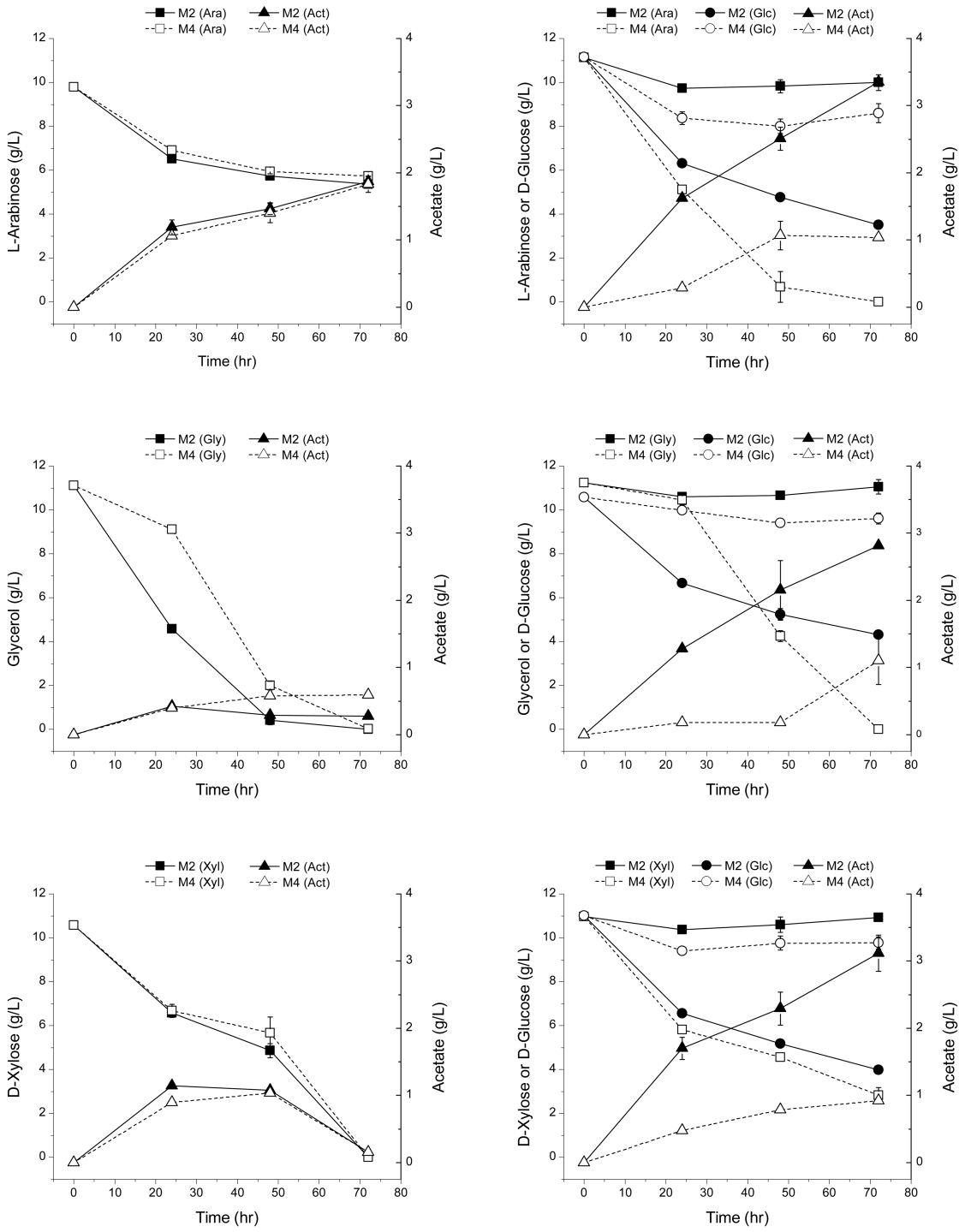


Figure 4.3 | Carbon source utilization and acetate production (Ara: L-arabinose, Gly: glycerol, Xyl: D-xylose, Glc: D-glucose, Act: acetate) as a function of time in strains M2 and M4. Error bars represent the standard deviation of biological triplicates.

Table 4.3 | Maximum growth rate μ_{max} of strains M2 and M4 in LB supplemented with various carbon substrates (10 g/L). Error represents the standard deviation of biological triplicates.

Strain	Alternative Carbon Supplemented	D-Glucose	μ_{max} (hr^{-1})
M2	L-Arabinose	−	1.00 ± 0.01
		+	1.01 ± 0.06
	Glycerol	−	0.98 ± 0.02
		+	0.97 ± 0.04
	D-Xylose	−	1.00 ± 0.03
		+	0.87 ± 0.20
M4	L-Arabinose	−	0.65 ± 0.01
		+	0.66 ± 0.02
	Glycerol	−	0.66 ± 0.05
		+	0.68 ± 0.01
	D-Xylose	−	0.70 ± 0.01
		+	0.67 ± 0.01

4.3.2 D-glucaric Acid Production in the Δ*pgi* Δ*zwf* Mutant

D-glucaric acid, a dicarboxylic organic acid, is a naturally occurring compound which has been investigated for a variety of potential applications, including cholesterol reduction and cancer treatment. A biosynthetic pathway to D-glucaric acid from D-glucose has been constructed in *E. coli* (Fig. 1.3) (Moon et al., 2009b). This pathway begins with glucose 6-phosphate, which is converted to *myo*-inositol-1-phosphate by *myo*-inositol-1-phosphate synthase (INO1). *Myo*-inositol-1-phosphate is then dephosphorylated by an endogenous phosphatase to yield *myo*-inositol, which is oxidized to D-glucuronic acid by *myo*-inositol oxygenase (MIOX). Finally, D-glucuronic acid is oxidized to D-glucaric acid by uronate dehydrogenase (Udh). Because production of D-glucaric acid requires glucose 6-phosphate, we hypothesized that the yield of D-glucaric acid could be increased significantly in a Δ*pgi* Δ*zwf* strain.

Production of D-glucaric acid in MG1655(DE3) Δ*endA* Δ*recA* Δ*gudD* Δ*uxaC* (strain M2-2) and MG1655(DE3)

Table 4.4 | D-glucaric acid titer and yield on D-glucose for strains M2-2 and M6 on various carbon sources. Error represents the standard deviation of biological triplicates.

Strain	Carbon	D-Glucaric Acid Titer (g/L)	Yield on D-glucose (g/g)
M2-2	L-Arabinose	0.13 ± 0.01	0.044 ± 0.002
	Glycerol	0.20 ± 0.02	0.052 ± 0.009
	D-Xylose	0.13 ± 0.01	0.039 ± 0.002
M6	L-Arabinose	0.50 ± 0.01	0.76 ± 0.13
	Glycerol	0.81 ± 0.10	0.44 ± 0.04
	D-Xylose	1.19 ± 0.08	0.73 ± 0.03

Δ*endA* Δ*recA* Δ*pgi* Δ*zwf* Δ*gudD* Δ*uxaC* (strain M6) was compared. Both *gudD* and *uxaC* were knocked out of both of these strains to prevent consumption of either D-glucuronic or D-glucaric acids. As expected, D-glucaric acid yield on D-glucose is increased in the Δ*pgi* Δ*zwf* mutant nearly 18-fold over an unmutated control supplemented with L-arabinose or D-xylose; yield is increased approximately 9-fold in the Δ*pgi* Δ*zwf* strain supplemented with glycerol (Table 4.4). Additionally, D-glucaric acid titers are significantly higher in the Δ*pgi* Δ*zwf* mutant (Fig. 4.4). We hypothesize that deletion of *pgi* and *zwf* results in higher glucose 6-phosphate pools, allowing INO1 to operate much closer to its maximum activity, leading to increased flux through the D-glucaric acid pathway.

4.4 Discussion

Traditionally, the main focus of metabolic engineering projects has been on increasing the productivity and final titer of a product of interest, and this approach has been widely successful for high-value compounds such as pharmaceutical intermediates and therapeutic proteins. However, simply increasing productivity and titers may not be sufficient for low-margin, high-volume bioproducts such as biofuels and commodity chemicals. In these cases, product yield becomes an important process consideration, as raw material costs can be a large percentage of the manufacturing costs. Strategies which are able to increase product yield without sacrificing productivity or titer would be valuable tools for the metabolic engineer.

Recent interest in the use of renewable feedstocks such as lignocelluosic biomass for biochemical pro-

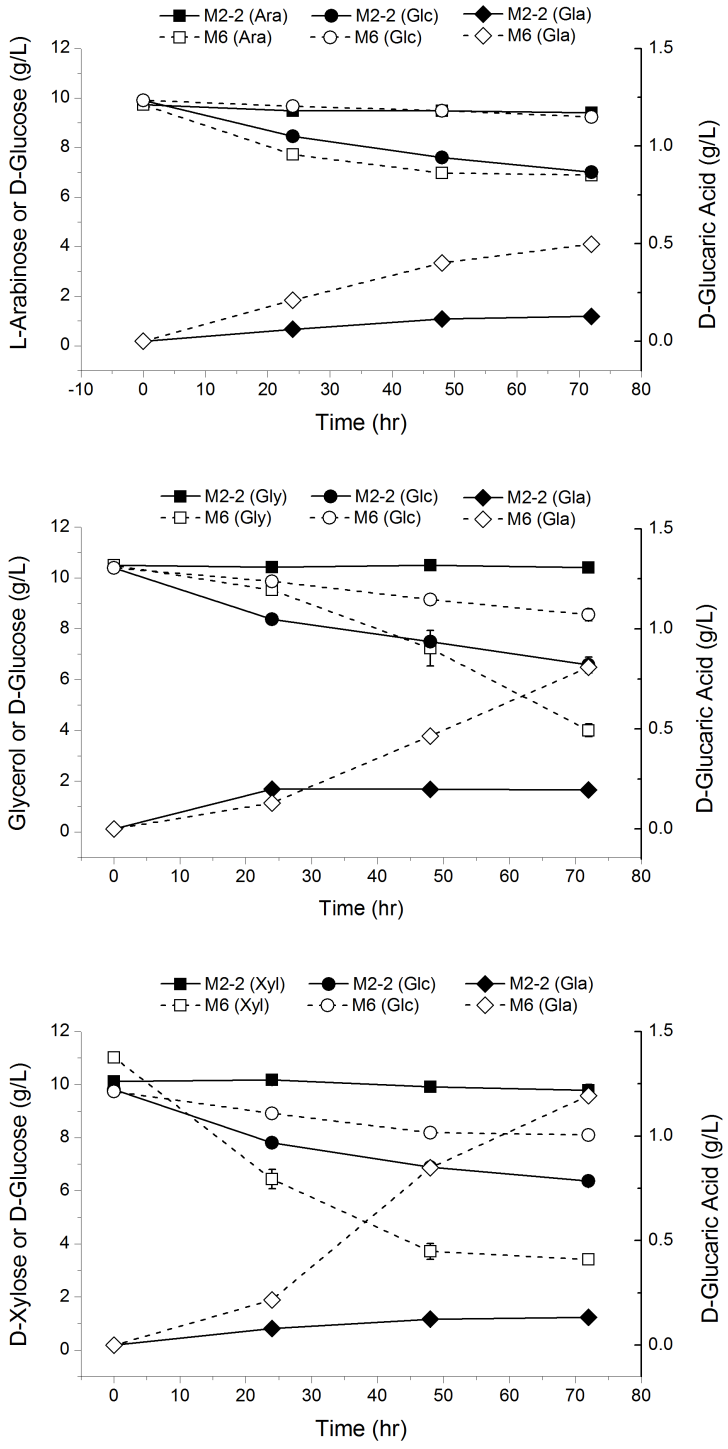


Figure 4.4 | D-glucaric acid production and carbon source utilization (Ara: L-arabinose, Gly: glycerol, Xyl: D-xylose, Glc: D-glucose, Gla: D-glucaric acid) as a function of time in strains M2-2 and M6. Both strains harbored pRSFD-IN-Udh and pTrc-SUMO-MIOX for production of D-glucaric acid from D-glucose. Error bars represent the standard deviation of biological triplicates.

duction presents an interesting opportunity for increasing the yield of biochemicals derived directly from glucose: in addition to glucose, lignocellulosic biomass contains several other fermentable sugars (e.g., xylose and arabinose) which may be used for biomass formation while reserving glucose solely for product generation. Because wild-type *E. coli* preferentially consumes glucose, strain engineering is necessary to shift the cell's preference towards alternative carbon sources. In this work, we characterized the carbon source preference of a Δ*pgi* Δ*zwf* mutant and explored its ability to improve the yield of D-glucaric acid on D-glucose.

As expected, deletion of *pgi* and *zwf* eliminates the cell's ability to consume D-glucose for biomass formation. Catabolite repression is eliminated in this strain as well, as the Δ*pgi* Δ*zwf* mutant is able to consume L-arabinose, glycerol, and D-xylose in the presence of D-glucose. Because glucose-mediated catabolite repression occurs when glucose transport into the cell is high, we believe that intracellular buildup of glucose 6-phosphate in the Δ*pgi* Δ*zwf* mutant leads to significantly reduced glucose transport, alleviating catabolite repression. Interestingly, introduction of the D-glucaric acid pathway, which should draw down intracellular glucose 6-phosphate pools, does not appear to affect the uptake of alternative carbon sources in the presence of D-glucose. We speculate that glucose influx in the presence of INO1 is not high enough to significantly reduce the levels of phosphorylated $EIIA^{Glc}$ to result in activation of catabolite repression. Because the threshold rate of glucose import necessary for activation of catabolite repression is unknown, efforts to further increase the activity of INO1 or to introduce more active glucose consumption pathways should proceed with caution to avoid reactivation of catabolite repression.

4.5 Conclusions

In this work, we investigated the behavior of a Δ*pgi* Δ*zwf* mutant and its ability to utilize alternative carbon sources for cell growth while reserving D-glucose for product formation. This strain was able to consume L-arabinose, glycerol, and D-xylose even in the presence of D-glucose, and yields of D-glucaric acid on D-glucose were increased 9- to 18-fold in the Δ*pgi* Δ*zwf* strain. Additionally, product titers were also increased, as the initial D-glucaric acid pathway enzyme was no longer in competition with glycolytic enzymes for glucose 6-phosphate. These gains in product yield and productivity should easily translate to other bioproducts derived from D-glucose, and it is hoped that this strain will help improve the process economics of these value-added biochemicals.

Chapter 5

Conclusions and Future Directions

Abstract

This thesis set out to improve the commercial prospects of biological D-glucaric acid production through the development and application of metabolic engineering and synthetic biology strategies towards improving pathway productivity and yield. This chapter summarizes the major findings and accomplishments of the thesis and discusses future directions for further improvements to biological production of D-glucaric acid.

5.1 Conclusions

D-glucaric acid, a hydroxylated dicarboxylic acid, has been reported to have a plethora of potential uses. Current production of D-glucaric acid, however, employs a chemical process which is nonselective and requires harsh reaction conditions. Given the commercial potential for D-glucaric acid, our group explored biological production of D-glucaric acid as a greener and potentially more economical process, leading to the construction of a heterologous pathway for D-glucaric acid production in *Escherichia coli* (Moon et al., 2009b).

The main goal of this thesis was to improve the economic viability of biological production of D-glucaric acid through improvements to process productivity and yield. Several metabolic engineering and synthetic biology strategies were explored over the course of this thesis, and the main findings, conclusions, and accomplishments are summarized below.

5.1.1 Metabolic Engineering for Improved Productivity

Previous work identified MIOX to be the least active pathway enzyme and demonstrated a linear correlation between *in vitro* MIOX specific activity and final D-glucaric acid titers (Moon et al., 2009b, Moon et al., 2010). In Chapter 2, I explored directed evolution and soluble protein fusion tags as methods for increasing MIOX activity, solubility, and stability. Both of these methods were very successful: a single round of directed evolution improved MIOX productivity 65%, and an N-terminal SUMO fusion to MIOX improved productivity 75%. However, both of these improvements were observed only when *myo*-inositol was supplied exogenously at high concentrations. Productivity improvements were not observed when D-glucose was fed, indicating that the *myo*-inositol activation effect was not removed by my protein engineering efforts.

Interestingly, the directed evolution efforts presented in Chapter 2 did not yield a MIOX variant with improved specific activity but instead isolated a 941 bp DNA fragment whose expression appears to increase *myo*-inositol transport into the cell. Expression of the *manX* portion of this 941 bp DNA fragment was shown to result in upregulation of *ptsG* transcription, likely via titration of *sgrS* mRNA away from *ptsG* interactions. This *sgrS*-*ptsG* interaction has been leveraged previously to downregulate expression of PtsG (Negrete et al., 2013). In contrast, our discovery of a method for titration of *sgrS* mRNA represents a strategy for upregulation of PtsG expression which does not require direct overexpression of PtsG. Additionally, PtsG was shown to function as a transporter of *myo*-inositol in *E. coli*, although it is apparent from the data that other *myo*-

inositol transporters exist in *E. coli*.

Finally, attempts to combine the DNA insert and the SUMO-MIOX fusion were not successful in increasing D-glucaric acid titers but did lead to the discovery of pH-related inhibition of D-glucaric acid productivity which limits productivity to approximately 5 g/L. A general pH-related inhibition effect also explains why it is possible to obtain higher titers of D-glucuronic acid than D-glucaric acid, since D-glucaric acid is more acidic than D-glucuronic acid.

5.1.2 Synthetic Biology for Improved Productivity

In addition to identifying MIOX as the least active pathway enzyme, Moon et al. also observed that addition of *myo*-inositol, the substrate for MIOX, to the culture medium significantly increased MIOX specific activity during exponential phase (Moon et al., 2009b). These observations led to the idea of delaying MIOX expression relative to INO1, which would allow *myo*-inositol to build up in the culture medium and intracellularly. Activation of MIOX expression some period of time after INO1 would ensure exposure of newly translated MIOX to activating concentrations of *myo*-inositol. A proof-of-concept experiment, in which MIOX expression was delayed 10 hours relative to INO1, resulted in a five-fold increase in D-glucuronic acid production from D-glucose.

Because D-glucaric acid is expected to be a commodity-scale biochemical, the use of exogenous chemical inducers in the production process is undesired, as chemical inducers are often quite expensive. Chapter 3 described attempts to develop genetic circuits which would be capable of autonomously performing time-delayed induction of MIOX expression. These circuits would not only have eliminated the need for costly chemical inducers but might also have provided robustness against the variations typical of biological processes.

Three separate circuits were explored: *myo*-inositol responsive promoters from *Salmonella typhimurium*, quorum sensing machinery from *Vibrio fischeri* and *Pseudomonas aeruginosa*, and a circuit based on a sigma/anti-sigma factor pair from *Bacillus subtilis*. The purported *myo*-inositol responsive promoters from *Salmonella* did not respond to *myo*-inositol, and it is believed that these promoters respond to a downstream catabolite instead. Overexpression of the quorum sensing machinery in *E. coli* led to undesired metabolic burden effects and unpredictable circuit behavior, and efforts to reduce metabolic burden effects also significantly reduced expression of the gene of interest. Finally, although the sigma/anti-sigma factor circuit behaved as expected in minimal medium, its behavior became unpredictable in rich medium.

The failure of these attempts to generate a controllable genetic timer highlights the current limitations of synthetic biology for metabolic engineering problems. The field of synthetic biology is predicated upon the availability of well-characterized genetic "parts" which can be composed to build genetic "devices" capable of performing complex biological tasks. In order for synthetic biology to be useful in metabolic engineering contexts, these genetic devices must be "modular," meaning that they must be able to perform their designed function regardless of host strain, medium composition, and/or the presence of other genetic devices. Unfortunately, not enough is currently understood about the relationship between context, cellular behavior, and circuit performance, and the identification of well-behaved genetic circuits remains a trial-and-error process.

Significant additional work will be necessary to fully characterize each genetic part, especially to understand each part's effect on its host cell and other parts within the cell. I believe that the nature of biological systems will ultimately prevent the creation of fully modular parts and devices: it is impossible to fully insulate a particular part from interactions with other cellular components. Gaining a deep understanding of the effect of context on part and device performance is therefore paramount to the success of synthetic biology, as such an understanding will hopefully allow for *a priori* design of devices which perform as expected under a given set of conditions.

5.1.3 Strain Engineering for Improved Yield

Chapters 2 and 3 focused on applying metabolic engineering and synthetic biology strategies to increase productivity of D-glucaric acid. At industrial scales, product yield is also an important process parameter, especially for low-margin products such as D-glucaric acid. To improve the yield of D-glucaric acid on D-glucose, *pgi* and *zwf* were deleted from an *E. coli* strain, effectively eliminating its ability to consume D-glucose for biomass formation. L-arabinose, glycerol, and D-xylose were investigated as alternative carbon sources which could be utilized for growth. In addition to eliminating D-glucose consumption, deletion of *pgi* and *zwf* was also found to eliminate catabolite repression in *E. coli*, allowing for consumption of alternative carbon sources even in the presence of D-glucose. Finally, productivity and yield of D-glucaric acid from D-glucose were shown to be significantly increased in the Δ*pgi* Δ*zwf* strain.

5.2 Future Directions

This thesis has made great strides in improving D-glucaric acid productivity and yield. Furthermore, this thesis has significantly improved our understanding of the D-glucaric acid pathway, its limitations, and its potential for further improvement. This section describes future research directions which leverage this improved understanding to further improve D-glucaric acid production.

5.2.1 Protein Engineering

The protein engineering efforts described in Chapter 2 of this thesis were very successful in increasing D-glucaric acid productivity from *myo*-inositol as a feedstock. However, our ultimate goal is high D-glucaric acid productivity from D-glucose, and further protein engineering of both INO1 and MIOX may help us achieve this goal.

INO1

Although MIOX was identified previously as the least active pathway enzyme, INO1 activity was also shown to be low relative to that of Udh, the final pathway enzyme (Moon et al., 2009b). Since *myo*-inositol is still required for activation of MIOX, improving INO1 activity could increase pathway productivity by boosting intracellular *myo*-inositol pools. Since a strain incapable of growth on D-glucose as a sole carbon source has been constructed, a simple growth-based screen should be possible: the Δ*pgi* Δ*zwf* mutant can be co-transformed with MIOX (or SUMO-fused MIOX) and a library of INO1 mutants, then plated on minimal medium plates with D-glucose as the sole carbon source. Colonies which exhibit improved growth should be better able to produce D-glucuronic acid from D-glucose. Ideally, this improved productivity would be due to improved activity of INO1.

One caveat of this screening method is that it will likely only isolate INO1 mutants with increased V_{max} rather than a lower K_M, since deletion of *pgi* and *zwf* likely results in significantly increased (and saturating) levels of intracellular glucose 6-phosphate. Because the INO1 mutants will not need to compete with glycolysis for glucose 6-phosphate, mutants which have a higher affinity for glucose 6-phosphate will not be selected for. This issue might be addressed by downregulating D-glucose import into the cell via overexpression of *sgrS* (Negrete et al., 2013) or by progressive reduction of D-glucose concentrations in the minimal medium screening plates.

An additional screening method revolves around the identification of *myo*-inositol responsive promoters (see Sec. 5.2.2). If such promoters can be found, a FACS screen could be used to identify improved INO1 variants: the *myo*-inositol responsive promoter can be used to activate expression of a fluorescent protein in a *myo*-inositol-dependent manner. Cells possessing more active INO1 variants should generate higher amounts of fluorescent protein, and these variants can then be isolated in a FACS machine.

MIOX

Although the directed evolution work described in Chapter 2 did not lead to a MIOX variant with improved activity, it did demonstrate the potential of a growth-based screen for isolating mutants with improved D-glucuronic acid productivity. In order to screen for MIOX mutants with improved activity, it might be necessary to screen under conditions more relevant to D-glucaric acid production, using a D-glucose feed instead of a *myo*-inositol feed. Such a screen should be possible in the Δ*pgi* Δ*zwf* mutant, and simultaneous directed evolution of INO1 and MIOX could speed the pathway improvement process.

An interesting undertaking could be attempting to evolve MIOX so that it no longer requires high concentrations of *myo*-inositol for high activity. This might be possible through a growth-based screen using progressively lower concentrations of *myo*-inositol as a feedstock. However, care should be taken to minimize the chances of once again recovering the 941 bp DNA fragment which increases *myo*-inositol transport into the cell. Screening in the presence of the 941 bp DNA fragment may avoid this problem and may also allow for lower extracellular concentrations of *myo*-inositol to be used in the screen.

5.2.2 Genetic Timers

Chapter 3 demonstrated that delayed induction of MIOX relative to INO1 could improve pathway productivity. However, the genetic circuits that were tested in this thesis failed to provide the desired time-delay effect. Continued development of these timers, including further characterization of individual circuit components and possible re-tuning of input/output relationships, could lead to more robust genetic timers. These timers would not only be applicable to D-glucaric acid production but could also find use in several other pathways currently being developed in the Prather Lab.

Given the importance of *myo*-inositol as an intermediate in various eukaryotic cell processes, intracellular *myo*-inositol levels are likely highly regulated. Comparative transcriptomic analysis of various eukaryotic organisms grown in the presence and absence of *myo*-inositol could identify promoters and regulatory proteins

which respond directly to *myo*-inositol instead of a downstream catabolite like the *Salmonella* promoters that were tested in Chapter 3. Successful identification of *myo*-inositol responsive promoters would allow for autonomous induction of MIOX in response to increasing *myo*-inositol levels as well as allow for a FACS-based screen of INO1 directed evolution mutants.

5.2.3 Process Engineering

Chapter 2 identified pH as a limiting factor for D-glucaric acid production, preventing production of more than approximately 5 g/L of product. The use of buffered media or online pH control in a bioreactor could alleviate this inhibition effect; additionally, exploration of D-glucaric acid production in a bioreactor would allow for identification of potential issues related to process scale-up. One potential issue is the increased aeration of bioreactors compared to shake flasks: although increased aeration often leads to increased growth rates and cell densities, it may also lead to increased inactivation of MIOX. MIOX contains a non-heme diiron cluster which must be in the (II/III) state to be catalytically active (Xing et al., 2006). The presence of excess amounts of oxygen could lead to oxidation of the diiron cluster to the catalytically inactive (III/III) state, leading to premature inactivation of MIOX and lower product titers as a result. However, oxygen cannot be eliminated completely from the system, since MIOX requires molecular oxygen to effect the conversion of *myo*-inositol to D-glucuronic acid. Therefore, it is likely that an optimal degree of oxygenation exists. Understanding the relationship between dissolved oxygen and MIOX activity/stability will be a key part of large-scale production of D-glucaric acid and perhaps allow process engineers to take advantage of the spatial inhomogeneities typically found in large-scale fermenters.

Chapter 4 described a strain capable of growth on an alternative carbon source (L-arabinose, glycerol, or D-xylose) in the presence of D-glucose, allowing D-glucose to be reserved for product formation. The use of this strain to produce D-glucaric acid resulted in significantly higher product yield and productivity. It may be possible to extend these gains in productivity and yield by using a fed-batch process to further increase cell densities. Variation of feed rate and feed composition could lead to the determination of an ideal feeding strategy which maximizes D-glucaric acid production in a large-scale fermenter.

5.2.4 Distributed Bioprocessing

The idea of "distributed bioprocessing," where tasks are distributed among several different strains or organisms, has garnered increasing attention in recent years. Distributed bioprocessing strategies can reduce metabolic burden associated with overexpression of heterologous pathway enzymes and could allow for expression of each individual pathway enzyme in its optimal host, potentially reducing the engineering required for high-level production. In the case of D-glucaric acid production, a distributed bioprocessing approach could leverage many of the advances described in this thesis without a significant amount of additional work.

One potential distributed bioprocessing scheme involves two strains of *E. coli*, both of which are unable to consume D-glucose, D-glucuronic acid, and D-glucaric acid (Δ*pgi* Δ*zwf* Δ*uxaC* Δ*gudD*) to maximize product yield. Additionally, one strain is deficient in L-arabinose consumption ("Strain A"), while the other strain is deficient in D-xylose consumption ("Strain B"), allowing both carbon sources to be fed. Co-culture of several *E. coli* strains on multiple carbon sources has been described previously (Xia et al., 2012).

Strain A would express INO1 and would solely be responsible for *myo*-inositol production. Additional work may be necessary to identify and overexpress a phosphatase which can dephosphorylate *myo*-inositol-1-phosphate to *myo*-inositol. Because INO1 is the only overexpressed protein in Strain A, INO1 levels are likely to be much higher in this strain compared to a strain expressing all three pathway enzymes. Additionally, inclusion of the 941 bp insert described in Chapter 2 may result in improved export of *myo*-inositol into the extracellular medium. Strain B would express MIOX and Udh to convert *myo*-inositol to D-glucaric acid. Inclusion of the 941 bp insert from Chapter 2 might improve *myo*-inositol uptake into Strain B depending on the effect of D-glucose on *myo*-inositol uptake through PtsG: in previous experiments, we have observed decreased MIOX activity in cultures grown in the presence of D-glucose and *myo*-inositol (data not shown), suggesting that the presence of D-glucose results in lower intracellular *myo*-inositol concentrations. If D-glucose blocks *myo*-inositol flux through PtsG through direct competition, then inclusion of the insert will not improve *myo*-inositol uptake rates. However, if D-glucose actually acts to downregulate PtsG expression (for example, via an *sgrS*-mediated mechanism), then inclusion of the insert could improve *myo*-inositol uptake in Strain B. Additionally, the SUMO-MIOX fusion could be used to increase the solubility and stability of MIOX. Furthermore, induction of MIOX could be delayed using a quorum sensing system: Strain A would possess the AHL synthetase and function as the "sender" strain, while Strain B would possess the quorum-responsive regulator and function as the "receiver" strain. Separated production of the two quorum sensing components

would likely reduce metabolic burden associated with expression of the circuit; however, it is unclear whether this reduction would be sufficient to restore predictable circuit behavior.

A distributed bioprocessing approach presents many interesting advantages over the traditional, single-strain approach. However, several issues must be solved before distributed bioproduction of D-glucaric acid can become a reality, including the removal of pH limitations to D-glucaric acid production and the development of methods for maintaining a co-culture of *E. coli* strains.

References

[Ajikumar et al., 2010] Ajikumar, P. K., Xiao, W. H., Tyo, K. E. J., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T. H., Pfeifer, B. and Stephanopoulos, G. (2010). Isoprenoid pathway optimization for taxol precursor overproduction in Escherichia coli. Science *330*, 70–74.

[Almaas et al., 2004] Almaas, E., Kovács, B., Vicsek, T., Oltvai, Z. N. and Barabási, A. L. (2004). Global organization of metabolic fluxes in the bacterium Escherichia coli. Nature *427*, 839–843.

[Alon, 2006] Alon, U. (2006). An introduction to systems biology: design principles of biological circuits. Chapman & Hall, Boca Raton, FL.

[Alper et al., 2005a] Alper, H., Fischer, C., Nevoigt, E. and Stephanopoulos, G. (2005a). Tuning genetic control through promoter engineering. Proc. Natl. Acad. Sci. U.S.A *102*, 12678–126783.

[Alper et al., 2005b] Alper, H., Miyaoku, K. and Stephanopoulos, G. (2005b). Construction of lycopene-overproducing E. coli strains by combining systematic and combinatorial gene knockout targets. Nat. Biotechnol. *23*, 612–616.

[Amann and Brosius, 1985] Amann, E. and Brosius, J. (1985). 'ATG vectors' for regulated high-level expression of cloned genes in Escherichia coli. Gene. *40*, 183–190.

[An and Chin, 2009] An, W. and Chin, J. W. (2009). Synthesis of orthogonal transcription-translation networks. Proc. Natl. Acad. Sci. U.S.A *106*, 8477–8482.

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

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

[Ashwell, 1962] Ashwell, G. (1962). Enzymes of glucuronic and galacturonic acid metabolism in bacteria. Methods Enzymol. *5*, 190–208.

[Atsumi and Liao, 2008] Atsumi, S. and Liao, J. C. (2008). Directed evolution of Methanococcus jannaschii citramalate synthase for biosynthesis of 1-propanol and 1-butanol by Escherichia coli. Appl. Environ. Microbiol. *74*, 7802–7808.

[Baba et al., 2006] Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L. and Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. *2*, 2006.0008.

[Balderas-Hernández et al., 2011] Balderas-Hernández, V. E., Hernández-Montalvo, V., Bolívar, F., Gosset, G. and Martínez, A. (2011). Adaptive evolution of Escherichia coli inactivated in the phosphotransferase system operon improves co-utilization of xylose and glucose under anaerobic conditions. Appl. Biochem. Biotechnol. *163*, 485–496.

[Bhattacharya et al., 2013] Bhattacharya, S., Manna, P., Gachhui, R. and Sil, P. C. (2013). D-saccharic acid 1,4-lactone protects diabetic rat kidney by ameliorating hyperglycemia-mediated oxidative stress and renal inflammatory cytokines via NF-κB and PKC signaling. Toxicol. Appl. Pharmacol. *267*, 16–29.

[Bond-Watts et al., 2011] Bond-Watts, B. B., Bellerose, R. J. and Chang, M. C. Y. (2011). Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. Nat. Chem. Biol. *7*, 222–227.

[Brekasis and Paget, 2003] Brekasis, D. and Paget, M. S. B. (2003). A novel sensor of NADH/NAD+ redox poise in Streptomyces coelicolor A3(2). EMBO J. *22*, 4856–4865.

[Burgard et al., 2003] Burgard, A. P., Pharkya, P. and Maranas, C. D. (2003). OptKnock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol. Bioeng. *84*, 647–657.

[Burrill and Silver, 2011] Burrill, D. R. and Silver, P. A. (2011). Synthetic circuit identifies subpopulations with sustained memory of DNA damage. Genes Dev. *25*, 434–439.

[Bylund et al., 1998] Bylund, F., Collet, E., Enfors, S. O. and Larsson, G. (1998). Substrate gradient formation in the large-scale bioreactor lowers cell yield and increases by-product formation. Bioprocess Eng. *18*, 171–180.

[Callura and Dwyer, 2010] Callura, J. and Dwyer, D. (2010). Tracking, tuning, and terminating microbial physiology using synthetic riboregulators. Proc. Natl. Acad. Sci. U.S.A *107*, 15898–15903.

[Canton et al., 2008] Canton, B., Labno, A. and Endy, D. (2008). Refinement and standardization of synthetic biological parts and devices. Nat. Biotechnol. *26*, 787–793.

[Carothers et al., 2011] Carothers, J. M., Goler, J. A., Juminaga, D. and Keasling, J. D. (2011). Model-driven engineering of RNA devices to quantitatively program gene expression. Science *334*, 1716–1719.

[Caspeta et al., 2009] Caspeta, L., Flores, N., Pérez, N. O., Bolívar, F. and Ramírez, O. T. (2009). The effect of heating rate on Escherichia coli metabolism, physiological stress, transcriptional response, and production of temperature-induced recombinant protein: a scale-down study. Biotechnol. Bioeng. *102*, 468–482.

[Chemler et al., 2010] Chemler, J. A., Fowler, Z. L., McHugh, K. P. and Koffas, M. A. G. (2010). Improving NADPH availability for natural product biosynthesis in Escherichia coli by metabolic engineering. Metab. Eng. *12*, 96–104.

[Chen and Arkin, 2012] Chen, D. and Arkin, A. P. (2012). Sequestration-based bistability enables tuning of the switching boundaries and design of a latch. Mol. Syst. Biol. *8*, 620.

[Cheol et al., 2010] Cheol, H., Sook, J., Jang, W. and Yong, S. (2010). High NADPH / NADP+ ratio improves thymidine production by a metabolically engineered Escherichia coli strain. J. Biotechnol. *149*, 24–32.

[Cohen et al., 2001] Cohen, B. D., Sertil, O., Abramova, N. E., Davies, K. J. and Lowry, C. V. (2001). Induction and repression of DAN1 and the family of anaerobic mannoprotein genes in Saccharomyces cerevisiae occurs through a complex array of regulatory sites. Nucleic Acids Res. *29*, 799–808.

[Collins et al., 2005] Collins, C. H., Arnold, F. H. and Leadbetter, J. R. (2005). Directed evolution of Vibrio fischeri LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones. Mol. Microbiol. *55*, 712–723.

[Cox et al., 2007] Cox, R. S., Surette, M. G. and Elowitz, M. B. (2007). Programming gene expression with combinatorial promoters. Mol. Syst. Biol. *3*, 145.

[Curran and Alper, 2012] Curran, K. A. and Alper, H. (2012). Expanding the chemical palate of cells by combining systems biology and metabolic engineering. Metab. Eng. *14*, 289–297.

[Dahl et al., 2013] Dahl, R. H., Zhang, F., Alonso-Gutierrez, J., Baidoo, E., Batth, T. S., Redding-Johanson, A. M., Petzold, C. J., Mukhopadhyay, A., Lee, T. S., Adams, P. D. and Keasling, J. D. (2013). Engineering dynamic pathway regulation using stress-response promoters. Nat. Biotechnol. *31*, 1039–1046.

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

[De Mey et al., 2007] De Mey, M., Maertens, J., Lequeux, G. J., Soetaert, W. K. and Vandamme, E. J. (2007). Construction and model-based analysis of a promoter library for E. coli: an indispensable tool for metabolic engineering. BMC Biotechnol. *7*, 34.

[de Smit and van Duin, 1990] de Smit, M. H. and van Duin, J. (1990). Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. Proc. Natl. Acad. Sci. U.S.A *87*, 7668–7672.

[Del Vecchio et al., 2008] Del Vecchio, D., Ninfa, A. J. and Sontag, E. D. (2008). Modular cell biology: retroactivity and insulation. Mol. Syst. Biol. *4*, 161.

[Desai and Gallivan, 2004] Desai, S. K. and Gallivan, J. P. (2004). Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation. J. Am. Chem. Soc. *126*, 13247–13254.

[Dijkgraaf et al., 1987] Dijkgraaf, P. J. M., Verkuylen, M. and Vanderwiele, K. (1987). Complexation of calcium ions by complexes of glucaric acid and boric acid. Carbohydr. Res. *163*, 127–131.

[Dueber et al., 2009] Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V., Prather, K. L. J. and Keasling, J. D. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. Nat. Biotechnol. *27*, 753–759.

[Dunlap, 1999] Dunlap, P. V. (1999). Quorum regulation of luminescence in Vibrio fischeri. J. Mol. Microbiol. Biotechnol. *1*, 5–12.

[Ellington and Szostak, 1990] Ellington, A. and Szostak, J. (1990). In vitro selection of RNA molecules that bind specific ligands. Nature *346*, 818–822.

[Ellis et al., 2009] Ellis, T., Wang, X. and Collins, J. J. (2009). Diversity-based, model-guided construction of synthetic gene networks with predicted functions. Nat. Biotechnol. *27*, 465–471.

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

[Esposito and Chatterjee, 2006] Esposito, D. and Chatterjee, D. K. (2006). Enhancement of soluble protein expression through the use of fusion tags. Curr. Opin. Biotechnol. *17*, 353–358.

[Farmer and Liao, 2000] Farmer, W. R. and Liao, J. C. (2000). Improving lycopene production in Escherichia coli by engineering metabolic control. Nat. Biotechnol. *18*, 533–537.

[Fedorov et al., 2003] Fedorov, R., Schlichting, I., Hartmann, E., Domratcheva, T., Fuhrmann, M. and Hegemann, P. (2003). Crystal structures and molecular mechanism of a light-induced signaling switch: the Phot-LOV1 domain from Chlamydomonas reinhardtii. Biophys. J. *84*, 2474–2482.

[Feist and Palsson, 2008] Feist, A. M. and Palsson, B. O. (2008). The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli. Nat. Biotechnol. *26*, 659–667.

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

[Fuqua et al., 1994] Fuqua, W. C., Winans, S. C. and Greenberg, E. P. (1994). Quorum sensing in bacteria: the luxR-luxI family of cell density-responsive transcriptional regulators. J. Bacteriol. *176*, 269–275.

[Ganduri et al., 2005] Ganduri, V. S., Ghosh, S. and Patnaik, P. R. (2005). Optimal mixing to improve the performance of batch and continuous fermentations with recombinant Escherichia coli. J. Chem. Technol. Biotechnol. *80*, 361–370.

[Glick, 1995] Glick, B. (1995). Metabolic load and heterologous gene expression. Biotechnol. Adv. *13*, 247–261.

[Gonçalves et al., 2013] Gonçalves, G. A. L., Prazeres, D. M. F., Monteiro, G. A. and Prather, K. L. J. (2013). De novo creation of MG1655-derived E. coli strains specifically designed for plasmid DNA production. Appl. Microbiol. Biotechnol. *97*, 611–620.

[Gupta and Singh, 2004] Gupta, K. P. and Singh, J. (2004). Modulation of carcinogen metabolism and DNA interaction by calcium glucarate in mouse skin. Toxicol. Sci. *79*, 47–55.

[Hansen et al., 1999] Hansen, C. A., Dean, A. B., Draths, K. M. and Frost, J. W. (1999). Synthesis of 1,2,3,4-tetrahydroxybenzene from D-glucose: exploiting myo-inositol as a precursor to aromatic chemicals. J. Am. Chem. Soc. *121*, 3799–3800.

[Hawkins et al., 2007] Hawkins, A. C., Arnold, F. H., Stuermer, R., Hauer, B. and Leadbetter, J. R. (2007). Directed evolution of Vibrio fischeri LuxR for improved response to butanoyl-homoserine lactone. Appl. Environ. Microbiol. *73*, 5775–5781.

[Hibbert and Dalby, 2005] Hibbert, E. G. and Dalby, P. A. (2005). Directed evolution strategies for improved enzymatic performance. Microb. Cell Fact. *4*, 29.

[Horton et al., 2013] Horton, R. M., Cai, Z., Ho, S. N. and Pease, L. R. (2013). Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. Biotechniques. *54*, 129–133.

[Jackson et al., 1972] Jackson, D. A., Symons, R. H. and Berg, P. (1972). Biochemical method for inserting new genetic information into DNA of simian virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli. Proc. Natl. Acad. Sci. U.S.A *69*, 2904–2909.

[Jang et al., 2012] Jang, Y. S., Malaviya, A., Cho, C., Lee, J. and Lee, S. Y. (2012). Butanol production from renewable biomass by clostridia. Bioresour. Technol. *123*, 653–663.

[Jensen and Hammer, 1998] Jensen, P. and Hammer, K. (1998). The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. Appl. Environ. Microbiol. *64*, 82–87.

[Johnson and Taconi, 2009] Johnson, D. T. and Taconi, K. A. (2009). The glycerin glut : options for the value-added conversion of crude glycerol resulting from biodiesel production. Environ. Prog. *26*, 338–348.

[Johnson, 1983] Johnson, I. S. (1983). Human insulin from recombinant DNA technology. Science *219*, 632–637.

[Jones et al., 2000] Jones, K. L., Kim, S. W. and Keasling, J. D. (2000). Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. Metab. Eng. *2*, 328–338.

[Joyce and Stewart, 2012] Joyce, B. L. and Stewart, C. N. (2012). Designing the perfect plant feedstock for biofuel production: using the whole buffalo to diversify fuels and products. Biotechnol. Adv. *30*, 1011–1022.

[Kaufman, 1991] Kaufman, R. (1991). Developing rDNA products for treatment of hemophilia A. Trends Biotechnol. *9*, 353–359.

[Khosla et al., 2009] Khosla, C., Kapur, S. and Cane, D. E. (2009). Revisiting the modularity of modular polyketide synthases. Curr. Opin. Chem. Biol. *13*, 135–143.

[Khosla and Keasling, 2003] Khosla, C. and Keasling, J. J. D. (2003). Metabolic engineering for drug discovery and development. Nat. Rev. Drug Discov. *2*, 1019–1025.

[Kiely et al., 1994] Kiely, D., Chen, L. and Lin, T. H. (1994). Simple preparation of hydroxylated nylons-polyamides derived from aldaric acids. ACS Symp. Ser. Am. Chem. Soc. *575*, 149–158.

[Klumpp et al., 2009] Klumpp, S., Zhang, Z. and Hwa, T. (2009). Growth rate-dependent global effects on gene expression in bacteria. Cell *139*, 1366–1375.

[Knight, 2003] Knight, T. (2003). Idempotent vector design for standard assembly of biobricks. MIT Synthetic Biology Working Group Technical Reports. http://hdl.handle.net/1721.1/21168.

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

[Kogure et al., 2007] Kogure, T., Wakisaka, N., Takaku, H. and Takagi, M. (2007). Efficient production of 2-deoxy-scyllo-inososе from d-glucose by metabolically engineered recombinant Escherichia coli. J. Biotechnol *129*, 502–509.

[Kotte et al., 2010] Kotte, O., Zaugg, J. B. and Heinemann, M. (2010). Bacterial adaptation through distributed sensing of metabolic fluxes. Mol. Syst. Biol. *6*, 1–9.

[Kröger and Fuchs, 2009] Kröger, C. and Fuchs, T. M. (2009). Characterization of the myo-inositol utilization island of Salmonella enterica serovar Typhimurium. J. Bacteriol. *191*, 545–554.

[Kröger et al., 2010] Kröger, C., Stolz, J. and Fuchs, T. M. (2010). Myo-inositol transport by Salmonella enterica serovar Typhimurium. Microbiology *156*, 128–138.

[Kurian, 2005] Kurian, J. V. (2005). A new polymer platform for the future — Sorona® from corn derived 1,3-Propanediol. J. Polym. Environ. *13*, 159–167.

[Lara et al., 2006] Lara, A. R., Galindo, E., Ramírez, O. T. and Palomares, L. A. (2006). Living with heterogeneities in bioreactors. Mol. Biotechnol. *34*, 355–381.

[Lee et al., 2008] Lee, S. H., Park, S. J., Lee, S. Y. and Hong, S. H. (2008). Biosynthesis of enantiopure (S)-3-hydroxybutyric acid in metabolically engineered Escherichia coli. Appl. Microbiol. Biotechnol. *79*, 633–641.

[Lee, 1996] Lee, S. Y. (1996). Bacterial polyhydroxyalkanoates. Biotechnol. Bioeng. *49*, 1–14.

[Leonard et al., 2010] Leonard, E., Ajikumar, P. and Thayer, K. (2010). Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. Proc. Natl. Acad. Sci. U.S.A *107*, 13654–13659.

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

[Lou et al., 2010] 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 Ouyang, Q. (2010). Synthesizing a novel genetic sequential logic circuit: a push-on push-off switch. Mol. Syst. Biol. *6*, 350.

[Lütke-Eversloh and Stephanopoulos, 2008] Lütke-Eversloh, T. and Stephanopoulos, G. (2008). Combinatorial pathway analysis for improved L-tyrosine production in Escherichia coli: identification of enzymatic bottlenecks by systematic gene overexpression. Metab. Eng. *10*, 69–77.

[Marcheschi et al., 2013] Marcheschi, R. J., Gronenberg, L. S. and Liao, J. C. (2013). Protein engineering for metabolic engineering: current and next-generation tools. Biotechnol. J. *8*, 545–555.

[McCormick et al., 1984] McCormick, F., Trahey, M., Innis, M., Dieckmann, B. and Ringold, G. (1984). Inducible expression of amplified human beta interferon genes in CHO cells. Mol. Cell. Biol. *4*, 166.

[McKenna and Nielsen, 2011] McKenna, R. and Nielsen, D. R. (2011). Styrene biosynthesis from glucose by engineered E. coli. Metab. Eng. *13*, 544–554.

[Mehltretter and Rist, 1953] Mehltretter, C. L. and Rist, C. E. (1953). Sugar oxidation - saccharic acid and oxalic acids by the nitric acid oxidation of glucose. J. Agric. Food Chem. *1*, 779–783.

[Merbouh et al., 2002] Merbouh, N., Bobbitt, J. M. and Brückner, C. (2002). 4-AcNH-tempo-catalyzed oxidation of aldoses to aldaric acids using chlorine or bromine as terminal oxidants. J. Carbohydr. Chem. *21*, 65–77.

[Michnick et al., 1997] Michnick, S., Roustan, J. L., Remize, F., Barre, P. and Dequin, S. (1997). Modulation of glycerol and ethanol yields during alcoholic fermentation in Saccharomyces cerevisiae strains overexpressed or disrupted for GPD1 encoding glycerol 3-phosphate dehydrogenase. Yeast *13*, 783–793.

[Milias-Argeitis et al., 2011] Milias-Argeitis, A., Summers, S., Stewart-Ornstein, J., Zuleta, I., Pincus, D., El-Samad, H., Khammash, M. and Lygeros, J. (2011). In silico feedback for in vivo regulation of a gene expression circuit. Nat. Biotechnol. *29*, 1114–1116.

[Möglich and Moffat, 2007] Möglich, A. and Moffat, K. (2007). Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA. J. Mol. Biol. *373*, 112–26.

[Moon, 2006] Moon, T. S. (2006). Retrobiosynthesis of D-glucaric acid in a metabolically engineered strain of Escherichia coli. MIT Thesis Proposal.

[Moon et al., 2010] Moon, T. S., Dueber, J. E., Shiue, E. and Prather, K. L. J. (2010). Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered E. coli. Metab. Eng. *12*, 298–305.

[Moon et al., 2009a] Moon, T. S., Yoon, S. H., Ching, M. J. T. M., Lanza, A. M. and Prather, K. L. J. (2009a). Enzymatic assay of D-glucuronate using uronate dehydrogenase. Anal. Biochem. *392*, 183–185.

[Moon et al., 2009b] Moon, T. S., Yoon, S. H., Lanza, A. M., Roy-Mayhew, J. D. and Prather, K. L. J. (2009b). Production of glucaric acid from a synthetic pathway in recombinant Escherichia coli. Appl. Environ. Microbiol. *75*, 589–595.

[Negrete et al., 2013] Negrete, A., Majdalani, N., Phue, J. N. and Shiloach, J. (2013). Reducing acetate excretion from E. coli K-12 by over-expressing the small RNA SgrS. N. Biotechnol. *30*, 269–273.

[Neubauer et al., 2003] Neubauer, P., Lin, H. Y. and Mathiszik, B. (2003). Metabolic load of recombinant protein production: inhibition of cellular capacities for glucose uptake and respiration after induction of a heterologous gene in Escherichia coli. Biotechnol. Bioeng. *83*, 53–64.

[Nevoigt et al., 2007] Nevoigt, E., Fischer, C., Mucha, O., Matthäus, F., Stahl, U. and Stephanopoulos, G. (2007). Engineering promoter regulation. Biotechnol. Bioeng. *96*, 550–558.

[Ni et al., 1999] Ni, M., Tepperman, J. M. and Quail, P. H. (1999). Binding of phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light. Nature *400*, 781–784.

[Nudler and Mironov, 2004] Nudler, E. and Mironov, A. (2004). The riboswitch control of bacterial metabolism. Trends Biochem. Sci. *29*, 11–17.

[Pamuk et al., 2001] Pamuk, V., Yilmaz, M. and Alicilar, A. (2001). The preparation of D-glucaric acid by oxidation of molasses in packed beds. J. Chem. Technol. Biotechnol. *76*, 186–190.

[Pandey et al., 2013] Pandey, R. P., Malla, S., Simkhada, D., Kim, B. G. and Sohng, J. K. (2013). Production of 3-O-xylosyl quercetin in Escherichia coli. Appl. Microbiol. Biotechnol. *97*, 1889–1901.

[Park et al., 2007] Park, J. H., Lee, K. H., Kim, T. Y. and Lee, S. Y. (2007). Metabolic engineering of Escherichia coli for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. Proc. Natl. Acad. Sci. U.S.A *104*, 7797–7802.

[Park et al., 2011] Park, S. J., Lee, T. W., Lim, S. C., Kim, T. W., Lee, H., Kim, M. K., Lee, S. H., Song, B. K. and Lee, S. Y. (2011). Biosynthesis of polyhydroxyalkanoates containing 2-hydroxybutyrate from unrelated carbon source by metabolically engineered Escherichia coli. Appl. Microbiol. Biotechnol. *93*, 273–283.

[Pasotti et al., 2011] Pasotti, L., Quattrocelli, M., Galli, D., De Angelis, M. G. C. and Magni, P. (2011). Multiplexing and demultiplexing logic functions for computing signal processing tasks in synthetic biology. Biotechnol. J. *6*, 784–795.

[Patnaik, 2002] Patnaik, P. R. (2002). Can imperfections help to improve bioreactor performance? Trends Biotechnol. *20*, 135–137.

[Perez et al., 2008] Perez, J. L., Jayaprakasha, G. K., Yoo, K. S. and Patil, B. S. (2008). Development of a method for the quantification of D-glucaric acid in different varieties of grapefruits by high-performance liquid chromatography and mass spectra. J. Chromatogr. A *1190*, 394–397.

[Pesci et al., 1997] Pesci, E. C., Pearson, J. P., Seed, P. C. and Iglewski, B. H. (1997). Regulation of las and rhl quorum sensing in Pseudomonas aeruginosa. J. Bacteriol. *179*, 3127–3132.

[Pfeifer et al., 2001] Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E. and Khosla, C. (2001). Biosynthesis of complex polyketides in a metabolically engineered strain of E. coli. Science *291*, 1790–1792.

[Pfleger et al., 2006] Pfleger, B. F., Pitera, D. J., Smolke, C. D. and Keasling, J. D. (2006). Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. Nat. Biotechnol. *24*, 1027–1032.

[Pharkya et al., 2004] Pharkya, P., Burgard, A. P. and Maranas, C. D. (2004). OptStrain: a computational framework for redesign of microbial production systems. Genome Res. *14*, 2367–2376.

[Piasecki et al., 2011] Piasecki, S. K., Taylor, C. A., Detelich, J. F., Liu, J., Zheng, J., Komsoukaniants, A., Siegel, D. R. and Keatinge-Clay, A. T. (2011). Employing modular polyketide synthase ketoreductases as biocatalysts in the preparative chemoenzymatic syntheses of diketide chiral building blocks. Chem. Biol. *18*, 1331–1340.

[Postma et al., 1993] Postma, P. W., Lengeler, J. W. and Jacobson, G. R. (1993). Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. Microbiol. Rev. *57*, 543–594.

[Qoronfleh et al., 1992] Qoronfleh, M., Debouck, C. and Keller, J. (1992). Identification and characterization of novel low-temperature-inducible promoters of Escherichia coli. J. Bacteriol. *174*, 7902–7909.

[Raab et al., 2005] Raab, R. M., Tyo, K. and Stephanopoulos, G. (2005). Metabolic engineering. Adv. Biochem. Eng. Biotechnol. *100*, 1–17.

[Radley et al., 2003] Radley, T. L., Markowska, A. I., Bettinger, B. T., Ha, J. H. and Loh, S. N. (2003). Allosteric switching by mutually exclusive folding of protein domains. J. Mol. Biol. *332*, 529–536.

[Ramalingam et al., 2009] Ramalingam, K. I., Tomshine, J. R., Maynard, J. A. and Kaznessis, Y. N. (2009). Forward engineering of synthetic bio-logical AND gates. Biochem. Eng. J. *47*, 38–47.

[Raman and Chandra, 2009] Raman, K. and Chandra, N. (2009). Flux balance analysis of biological systems: applications and challenges. Brief. Bioinform. *10*, 435–449.

[Rice and Vanderpool, 2011] Rice, J. B. and Vanderpool, C. K. (2011). The small RNA SgrS controls sugar-phosphate accumulation by regulating multiple PTS genes. Nucleic Acids Res. *39*, 3806–3819.

[Roe et al., 1998] Roe, A. J., Mclaggan, D., Davidson, I., Byrne, C. O., Booth, I. R., Laggan, D. M. C. and Davidson, I. A. N. (1998). Perturbation of anion balance during inhibition of growth of Escherichia coli by weak acids. J. Bacteriol. *180*, 767–772.

[Rogers et al., 2006] Rogers, P., Chen, J. S. and Zidwick, M. J. O. (2006). Prokaryotes. Springer New York, New York, NY.

[Salis, 2011] Salis, H. M. (2011). The ribosome binding site calculator. Methods Enzymol. *498*, 19–42.

[Salis et al., 2009] Salis, H. M., Mirsky, E. A. and Voigt, C. A. (2009). Automated design of synthetic ribosome binding sites to control protein expression. Nat Biotech *27*, 946–950.

[Sambrook and Russell, 2001] Sambrook, J. and Russell, D. W. (2001). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

[Schlabach et al., 2010] Schlabach, M. R., Hu, J. K., Li, M. and Elledge, S. J. (2010). Synthetic design of strong promoters. Proc. Natl. Acad. Sci. U.S.A *107*, 2538–2543.

[Schmidt-Dannert and Arnold, 1999] Schmidt-Dannert, C. and Arnold, F. H. (1999). Directed evolution of industrial enzymes. Trends Biotechnol. *17*, 135–136.

[Schöbel et al., 2004] Schöbel, S., Zellmeier, S., Schumann, W. and Wiegert, T. (2004). The Bacillus subtilis sigmaW anti-sigma factor RsiW is degraded by intramembrane proteolysis through YluC. Mol. Microbiol. *52*, 1091–1105.

[Serres et al., 2001] Serres, M. H., Gopal, S., Nahum, L. A., Liang, P., Gaasterland, T. and Riley, M. (2001). A functional update of the Escherichia coli K-12 genome. Genome Biol. *2*, research0035.1–0035.7.

[Solomon et al., 2013] Solomon, K. V., Moon, T. S., Ma, B., Sanders, T. M. and Prather, K. L. J. (2013). Tuning primary metabolism for heterologous pathway productivity. ACS Synth. Biol. *2*, 126–135.

[Solomon and Prather, 2011] Solomon, K. V. and Prather, K. L. J. (2011). The zero-sum game of pathway optimization: emerging paradigms for tuning gene expression. Biotechnol. J. *6*, 1064–1070.

[Song et al., 2011] Song, H., Payne, S., Tan, C. and You, L. (2011). Programming microbial population dynamics by engineered cell-cell communication. Biotechnol. J. *6*, 837–849.

[Stephanopoulos, 1999] Stephanopoulos, G. (1999). Metabolic fluxes and metabolic engineering. Metab. Eng. *1*, 1–11.

[Surette et al., 1999] Surette, M. G., Miller, M. B. and Bassler, B. L. (1999). Quorum sensing in Escherichia coli, Salmonella typhimurium, and Vibrio harveyi: a new family of genes responsible for autoinducer production. Proc. Natl. Acad. Sci. U.S.A *96*, 1639–1644.

[Tabor et al., 2011] Tabor, J. J., Levskaya, A. and Voigt, C. A. (2011). Multichromatic control of gene expression in Escherichia coli. J. Mol. Biol. *405*, 315–324.

[Topp and Gallivan, 2010] Topp, S. and Gallivan, J. P. (2010). Emerging applications of riboswitches in chemical biology. ACS Chem. Biol. *5*, 139–148.

[Tseng et al., 2010] Tseng, H. C., Harwell, C. L., Martin, C. H. and Prather, K. L. J. (2010). Biosynthesis of chiral 3-hydroxyvalerate from single propionate-unrelated carbon sources in metabolically engineered E. coli. Microb. Cell Fact. *9*, 96.

[Tseng et al., 2009] Tseng, H. C., Martin, C. H., Nielsen, D. R. and Prather, K. L. J. (2009). Metabolic engineering of Escherichia coli for enhanced production of (R)- and (S)-3-hydroxybutyrate. Appl. Environ. Microbiol. *75*, 3137–3145.

[Tu and Weissman, 2004] Tu, B. P. and Weissman, J. S. (2004). Oxidative protein folding in eukaryotes: mechanisms and consequences. J. Cell. Biol. *164*, 341–346.

[Tuerk and Gold, 1990] Tuerk, C. and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science *249*, 505–510.

[Uroz et al., 2009] Uroz, S., Dessaux, Y. and Oger, P. (2009). Quorum sensing and quorum quenching: the yin and yang of bacterial communication. Chembiochem *10*, 205–216.

[Walaszek et al., 1996] Walaszek, Z., Szemraj, J., Hanausek, M., Adams, A. K. and Sherman, U. (1996). Glucaric acid content of various fruits and vegetables and cholesterol-lowering effects of dietary glucarate in the rat. Nutr. Res. *16*, 673–681.

[Waldo et al., 1999] Waldo, G. S., Standish, B. M., Berendzen, J. and Terwilliger, T. C. (1999). Rapid protein-folding assay using green fluorescent protein. Nat. Biotechnol. *17*, 691–695.

[Wang et al., 2011a] Wang, B., Kitney, R. I., Joly, N. and Buck, M. (2011a). Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. Nat. Commun. *2*, 508.

[Wang et al., 2011b] Wang, C., Yoon, S. H., Jang, H. J., Chung, Y. R., Kim, J. Y., Choi, E. S. and Kim, S. W. (2011b). Metabolic engineering of Escherichia coli for α-farnesene production. Metab. Eng. *13*, 648–655.

[Wang et al., 2011c] Wang, D., Li, Q., Yang, M., Zhang, Y., Su, Z. and Xing, J. (2011c). Efficient production of succinic acid from corn stalk hydrolysates by a recombinant Escherichia coli with ptsG mutation. Process Biochem. *46*, 365–371.

[Warnecke and Gill, 2005] Warnecke, T. and Gill, R. T. (2005). Organic acid toxicity, tolerance, and production in Escherichia coli biorefining applications. Microb. Cell Fact. *4*, 25.

[Weber et al., 2006] Weber, W., Link, N. and Fussenegger, M. (2006). A genetic redox sensor for mammalian cells. Metab. Eng. *8*, 273–280.

[Weiss et al., 2008] Weiss, L. E., Badalamenti, J. P., Weaver, L. J., Tascone, A. R., Weiss, P. S., Richard, T. L. and Cirino, P. C. (2008). Engineering motility as a phenotypic response to luxI/R-dependent quorum sensing in Escherichia coli. Biotechnol. Bioeng. *100*, 1251–1255.

[Werpy and Petersen, 2004] Werpy, T. and Petersen, G. (2004). Top value added chemicals from biomass, volume 1: results of screening for potential candidates from sugars and synthesis gas. U.S. Department of Energy, Washington, DC.

[Wisselink and Mars, 2004] Wisselink, H. and Mars, A. (2004). Engineering of mannitol production in Lactococcus lactis: influence of overexpression of mannitol 1-phosphate dehydrogenase in different genetic backgrounds. Appl. Environ. Microbiol. *70*, 4286–4292.

[Wu et al., 2011] Wu, C. H., Lee, H. C. and Chen, B. S. (2011). Robust synthetic gene network design via library-based search method. Bioinformatics *27*, 2700–2706.

[Wu et al., 2009] Wu, X., Wu, D., Lu, Z., Chen, W., Hu, X. and Ding, Y. (2009). A novel method for high-level production of TEV protease by superfolder GFP tag. J. Biomed. Biotechnol. *2009*, 591923.

[Xia et al., 2012] Xia, T., Eiteman, M. A. and Altman, E. (2012). Simultaneous utilization of glucose, xylose and arabinose in the presence of acetate by a consortium of Escherichia coli strains. Microb. Cell Fact. *11*, 77.

[Xia et al., 2010] Xia, X. X., Qian, Z. G., Ki, C. S., Park, Y. H., Kaplan, D. L. and Lee, S. Y. (2010). Native-sized recombinant spider silk protein produced in metabolically engineered Escherichia coli results in a strong fiber. Proc. Natl. Acad. Sci. U.S.A *107*, 14059–14063.

[Xing et al., 2006] Xing, G., Barr, E. W., Diao, Y. H., Hoffart, L. M., Prabhu, K. S., Arner, R. J., Reddy, C. C., Krebs, C. and Bollinger, J. M. (2006). Oxygen activation by a mixed-valent, diiron(II/III) cluster in the glycol cleavage reaction catalyzed by myo-inositol oxygenase. Biochemistry *45*, 5402–5412.

[Xu et al., 2011] Xu, P., Ranganathan, S., Fowler, Z. L., Maranas, C. D. and Koffas, M. A. G. (2011). Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA. Metab. Eng. *13*, 578–587.

[Yamano et al., 1994] Yamano, S., Ishii, T., Nakagawa, M., Ikenaga, H. and Misawa, N. (1994). Metabolic engineering for production of beta-carotene and lycopene in Saccharomyces cerevisiae. Biosci. Biotechnol. Biochem. *58*, 1112–1114.

[Yeh et al., 1997] Yeh, K., Wu, S. H., Murphy, J. T. and Lagarias, J. C. (1997). A cyanobacterial phytochrome two-component light sensory system. Science *277*, 1505–1508.

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

[Yoon et al., 2009] Yoon, S. H., Moon, T. S., Iranpour, P., Lanza, A. M. and Prather, K. L. J. (2009). Cloning and characterization of uronate dehydrogenases from two pseudomonads and Agrobacterium tumefaciens strain C58. J. Bacteriol. *191*, 1565–1573.

[Yoshida et al., 2008] Yoshida, K., Yamaguchi, M., Morinaga, T., Kinehara, M., Ikeuchi, M., Ashida, H. and Fujita, Y. (2008). Myo-inositol catabolism in Bacillus subtilis. J. Biol. Chem. *283*, 10415–10424.

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

[Young et al., 2012a] Young, C. L., Britton, Z. T. and Robinson, A. S. (2012a). Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. Biotechnol. J. *7*, 620–634.

[Young et al., 2012b] Young, E. M., Comer, A. D., Huang, H. and Alper, H. (2012b). A molecular transporter engineering approach to improving xylose catabolism in Saccharomyces cerevisiae. Metab. Eng. *14*, 401–411.

[Zaslaver et al., 2006] Zaslaver, A., Mayo, A., Ronen, M. and Alon, U. (2006). Optimal gene partition into operons correlates with gene functional order. Phys. Biol. *3*, 183–189.

[Zaslaver et al., 2004] Zaslaver, A., Mayo, A. E., Rosenberg, R., Bashkin, P., Sberro, H., Tsalyuk, M., Surette, M. and Alon, U. (2004). Just-in-time transcription program in metabolic pathways. Nat. Genet. *36*, 486–491.

[Zhou et al., 2012] Zhou, K., Zou, R., Stephanopoulos, G. and Too, H. P. (2012). Enhancing solubility of deoxyxylulose phosphate pathway enzymes for microbial isoprenoid production. Microb. Cell Fact. *11*, 148.

[Zor and Selinger, 1996] Zor, T. and Selinger, Z. (1996). Linearization of the Bradford protein assay increases its sensitivity: theoretical and experimental studies. Anal. Biochem. *236*, 302–308.

Appendix A

Quorum Sensing Characterization

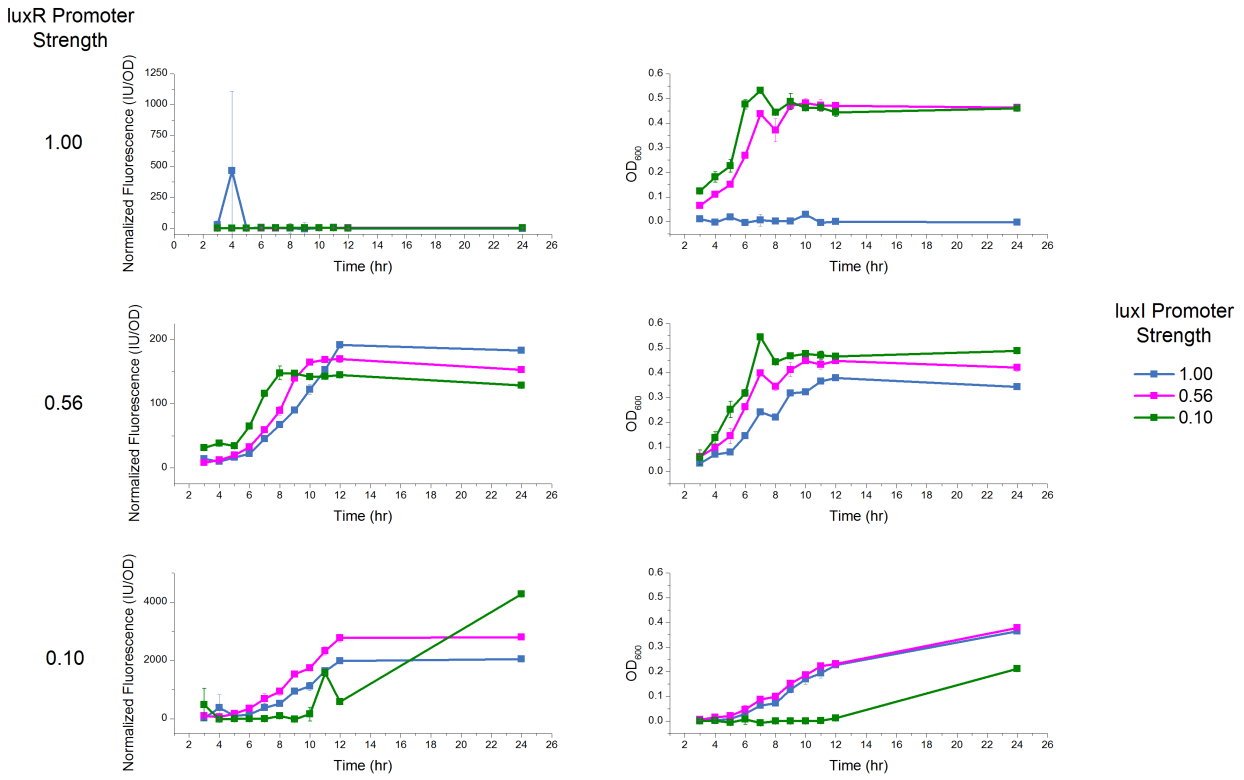


Figure A.1 | Characterization of the *lux* quorum sensing circuit from *Vibrio fischeri* for constant *luxR* expression level and various *luxI* expression levels. Cultures contained *E. coli* DH10B harboring pSB1A2-PconXX-luxR-Plux-GFP and pSB3C5-PconXX-luxI and were grown in 50 mL LB medium with ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL) at 37°C. Error bars represent the standard deviation of biological triplicates.

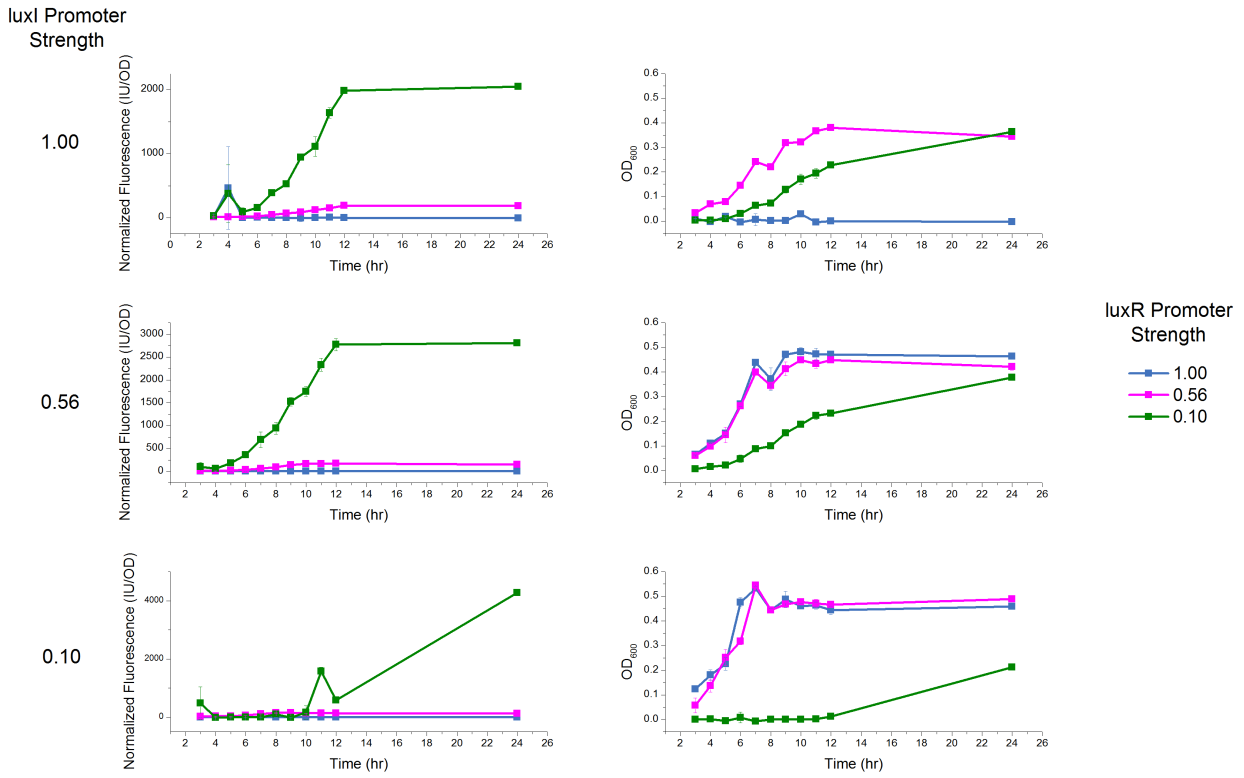


Figure A.2 | Characterization of the *lux* quorum sensing circuit from *Vibrio fischeri* for constant *luxI* expression level and various *luxR* expression levels. Cultures contained *E. coli* DH10B harboring pSB1A2-PconXX-luxR-Plux-GFP and pSB3C5-PconXX-luxI and were grown in 50 mL LB medium with ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL) at 37°C. Error bars represent the standard deviation of biological triplicates.

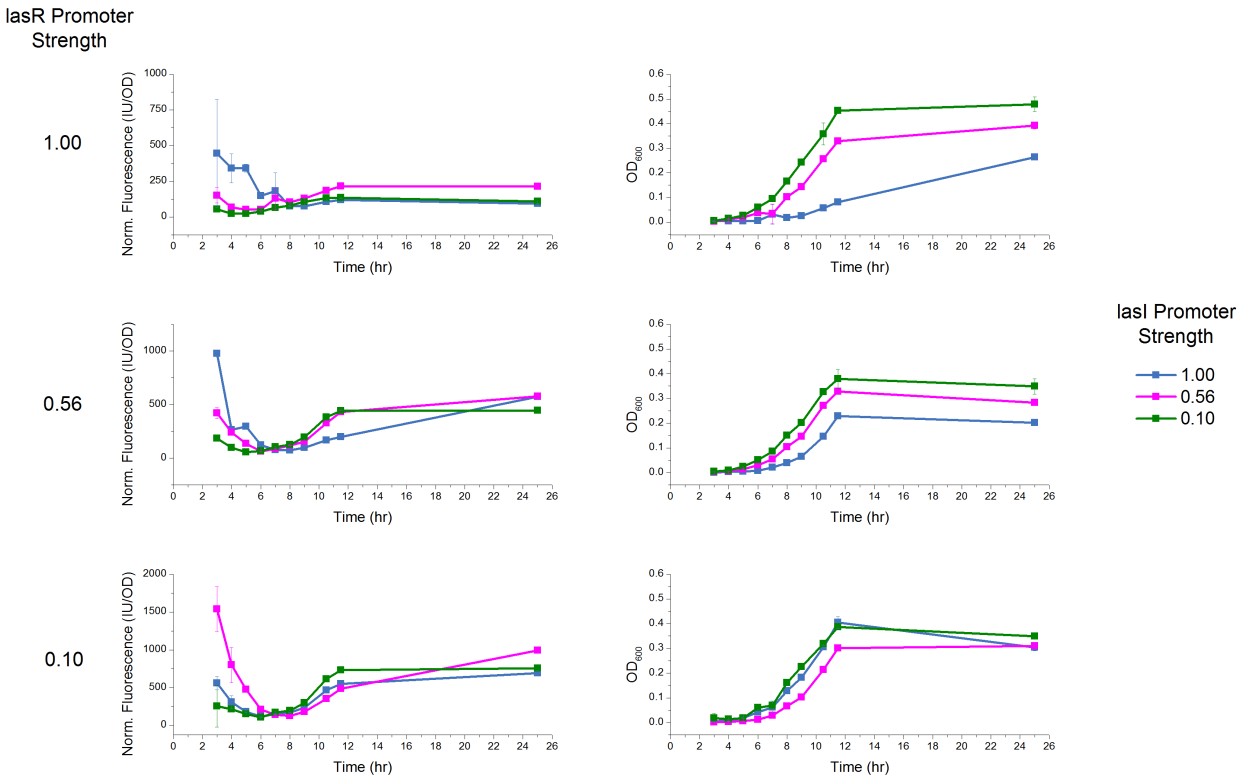


Figure A.3 | Characterization of the *las* quorum sensing circuit from *Pseudomonas aeruginosa* for constant *lasR* expression level and various *lasI* expression levels. Cultures contained *E. coli* DH10B harboring pSB1A2-PconXX-lasR-Plas-GFP and pSB3C5-PconXX-lasI and were grown in 50 mL LB medium with ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL) at 37°C. Error bars represent the standard deviation of biological triplicates.

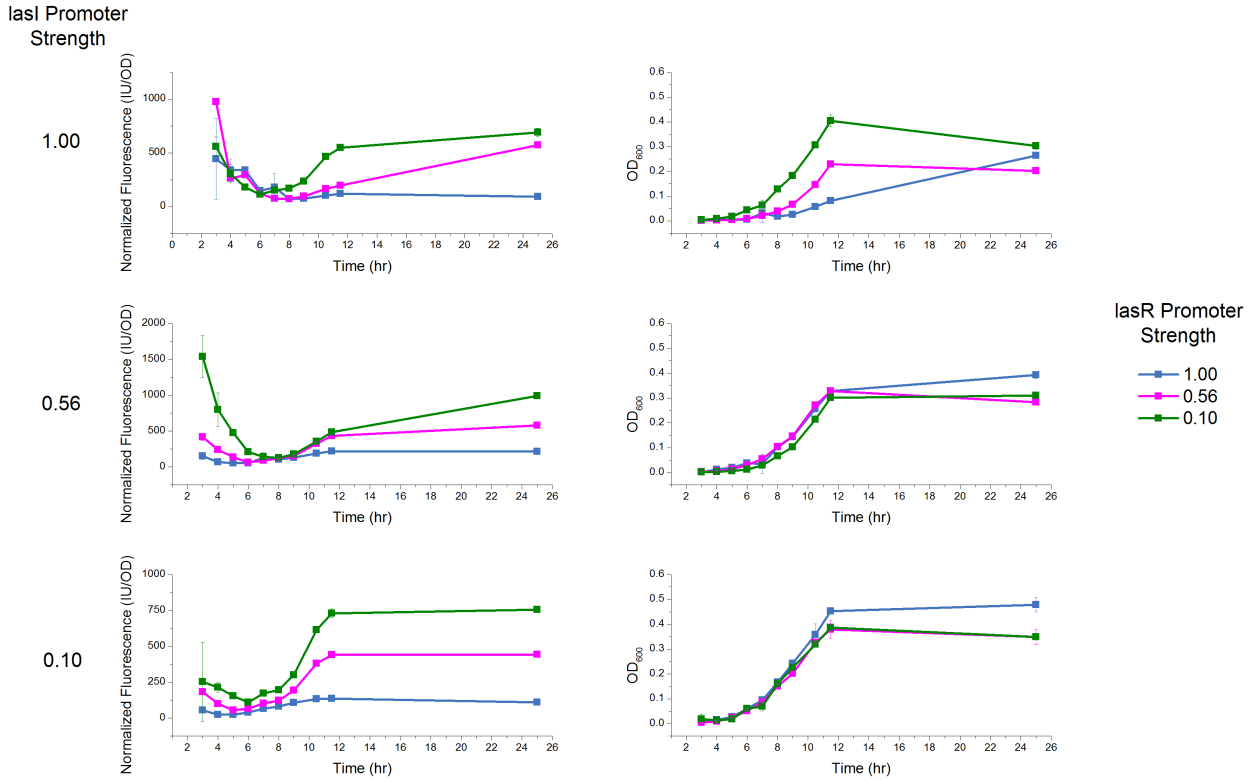


Figure A.4 | Characterization of the *las* quorum sensing circuit from *Pseudomonas aeruginosa* for constant *lasI* expression level and various *lasR* expression levels. Cultures contained *E. coli* DH10B harboring pSB1A2-PconXX-lasR-Plas-GFP and pSB3C5-PconXX-lasI and were grown in 50 mL LB medium with ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL) at 37°C. Error bars represent the standard deviation of biological triplicates.

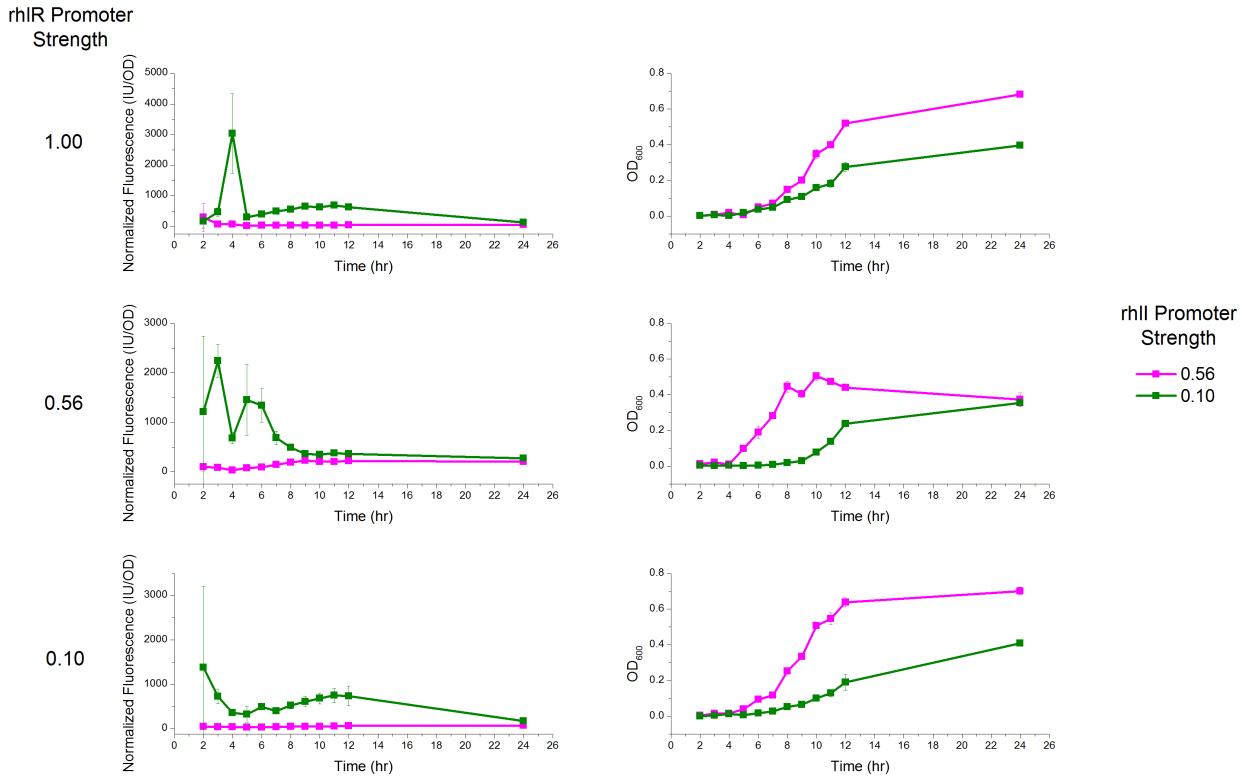


Figure A.5 | Characterization of the *rhl* quorum sensing circuit from *Pseudomonas aeruginosa* for constant *rhlR* expression level and various *rhlI* expression levels. Cultures contained *E. coli* DH10B harboring pSB1A2-PconXX-rhlR-Prhl-GFP and pSB3C5-PconXX-rhlI and were grown in 50 mL LB medium with ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL) at 37°C. Error bars represent the standard deviation of biological triplicates.

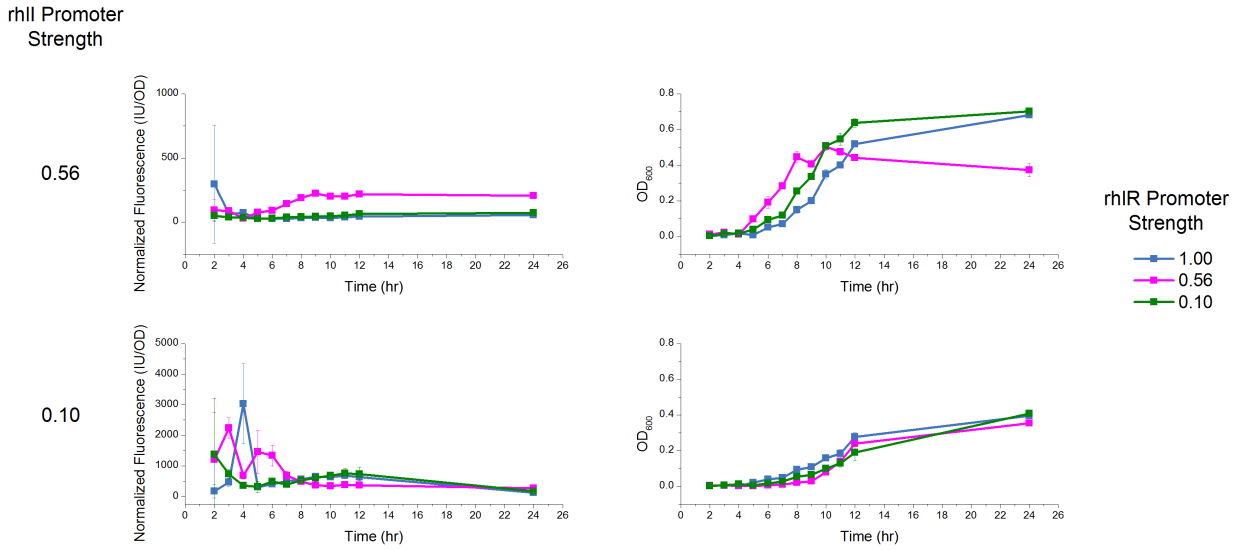


Figure A.6 | Characterization of the *rhl* quorum sensing circuit from *Pseudomonas aeruginosa* for constant *rhlI* expression level and various *rhlR* expression levels. Cultures contained *E. coli* DH10B harboring pSB1A2-PconXX-rhlR-PrhI-GFP and pSB3C5-PconXX-rhlI and were grown in 50 mL LB medium with ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL) at 37°C. Error bars represent the standard deviation of biological triplicates.

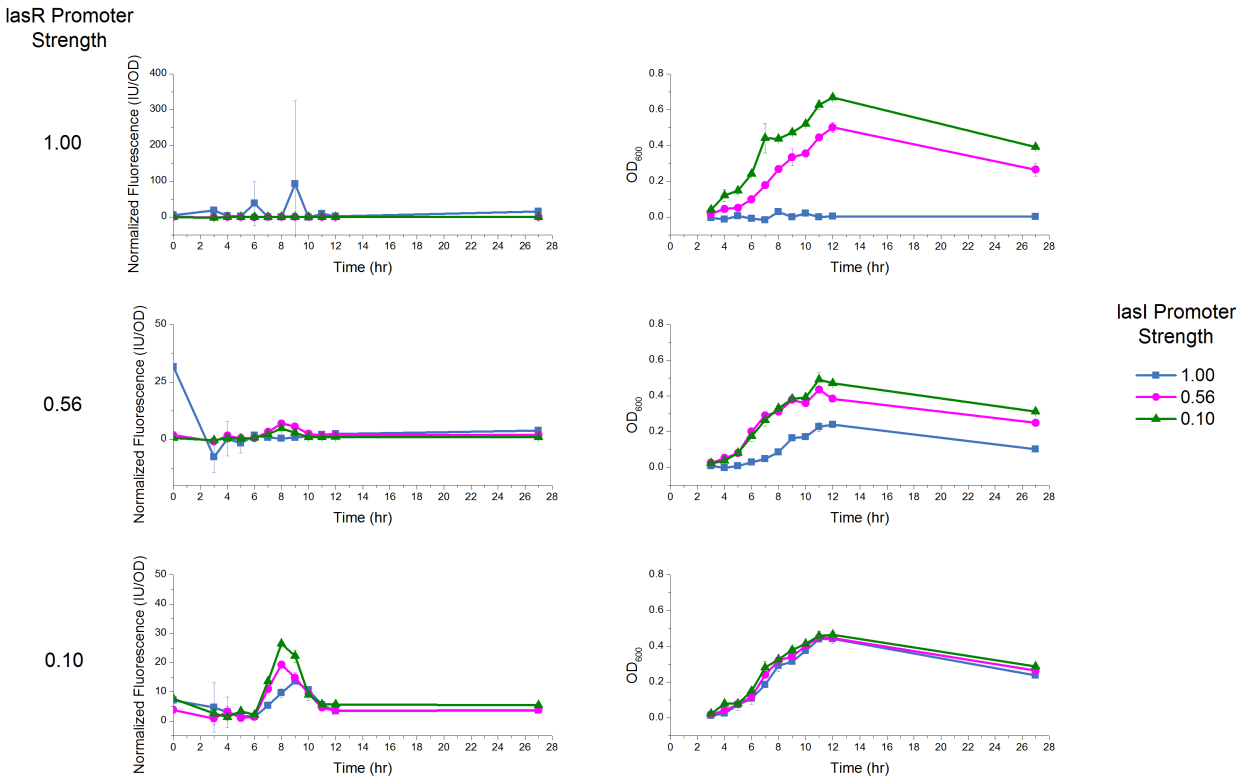


Figure A.7 | Characterization of the *las* quorum sensing circuit from *Pseudomonas aeruginosa* for constant *lasR* expression level and various *lasI* expression levels with LVA-tagged GFP. Cultures contained *E. coli* DH10B harboring pSB1A2-PconXX-lasR-Plas-GFP-LVA and pSB3C5-PconXX-lasI and were grown in 50 mL LB medium with ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL) at 37°C. Error bars represent the standard deviation of biological triplicates.

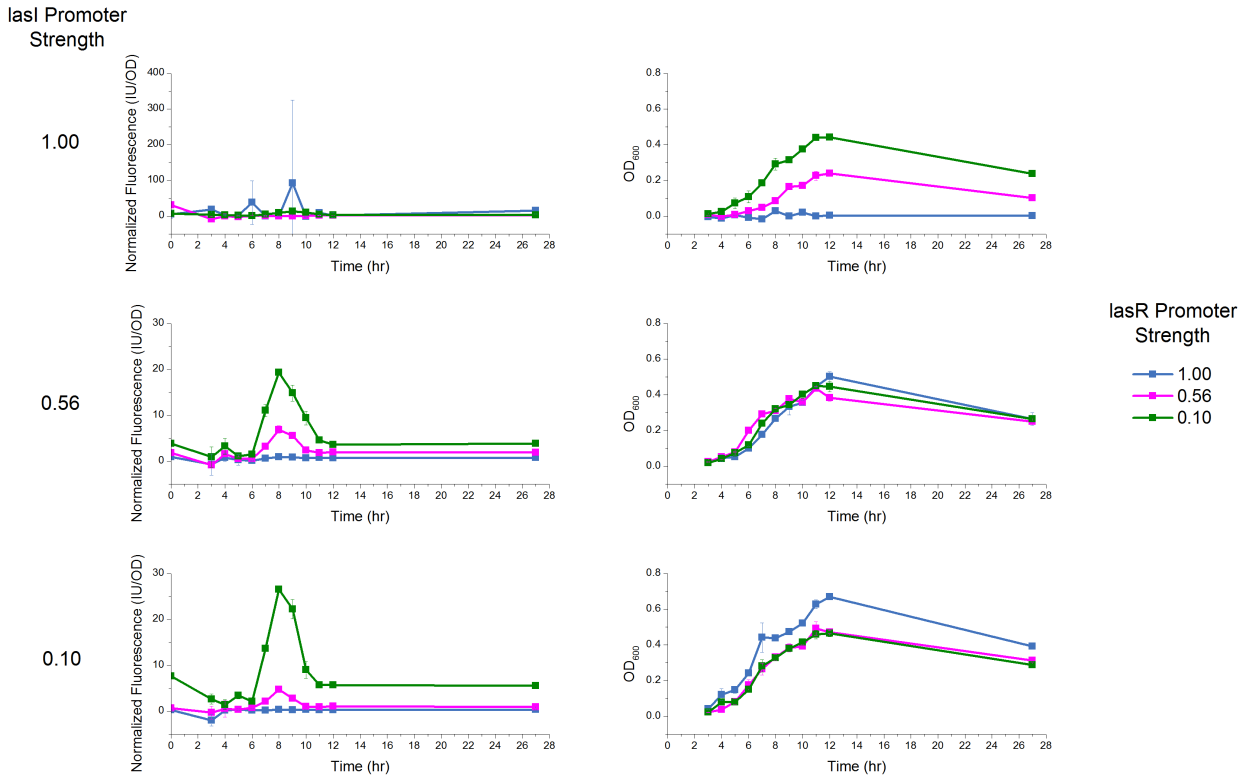


Figure A.8 | Characterization of the *las* quorum sensing circuit from *Pseudomonas aeruginosa* for constant *lasI* expression level and various *lasR* expression levels with LVA-tagged GFP. Cultures contained *E. coli* DH10B harboring pSB1A2-PconXX-lasR-Plas-GFP-LVA and pSB3C5-PconXX-lasI and were grown in 50 mL LB medium with ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL) at 37°C. Error bars represent the standard deviation of biological triplicates.

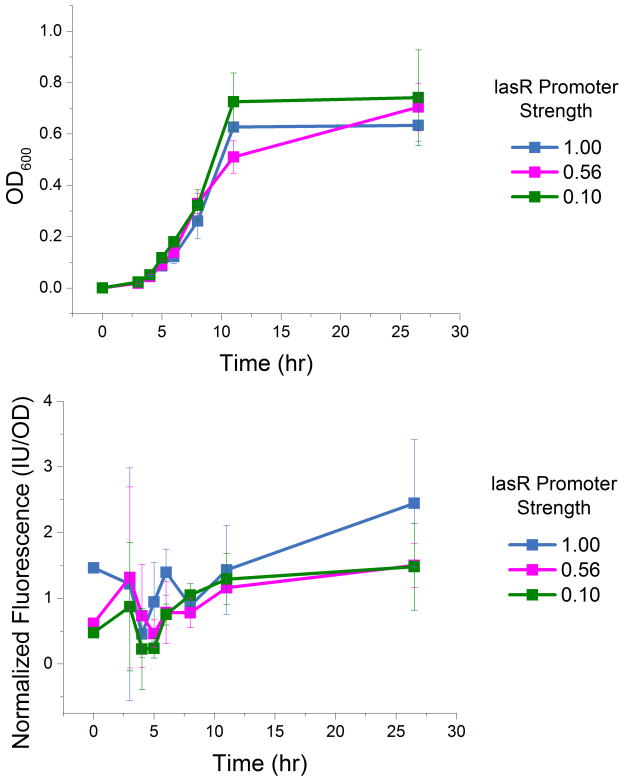


Figure A.9 | Characterization of the *las* quorum sensing circuit from *Pseudomonas aeruginosa* for constant *lasI* expression level and various *lasR* expression levels with *lasI* integrated into the *E. coli* genome. *lasI* is expressed from a strong constitutive promoter. Cultures contained *E. coli* MG1655(DE3) ΔendA ΔrecA gudD::Pcon100-lasI harboring pSB3C5-PconXX-lasR-Plas-GFP and were grown in 50 mL LB medium with 1% D-glucose and chloramphenicol (34 µg/mL) at 30°C. Error bars represent the standard deviation of biological triplicates.