Title
Transcription Factor-Based Small-Molecule Screens and Selections

Author
Dietrich, Jeffrey Allen

Publication Date
2011

Peer reviewed|Thesis/dissertation

Transcription Factor-Based Small-Molecule Screens and Selections

by

Jeffrey Allen Dietrich

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Joint Doctor of Philosophy

with University of California, San Francisco

in

Bioengineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jay D. Keasling, Chair

Professor Susan Marqusee

Professor Patricia C. Babbitt

Spring 2011

Transcription Factor-Based Small-Molecule Screens and Selections

by Jeffrey Allen Dietrich

Abstract

Transcription Factor-Based Small-Molecule Screens and Selections
by
Jeffrey Allen Dietrich
Joint Doctor of Philosophy
with University of California, San Francisco
in Bioengineering
University of California, Berkeley
Professor Jay D. Keasling, Chair

Directed evolution of *E. coli* for improved small-molecule production requires a combination of rational design and high-throughput screening technologies. Rational design-based directed evolution schemes use structural analyses and metabolic models to help identify targets for mutagenesis, thus improving the likelihood of identifying the desired phenotype. We used a strictly rational design-based approach to re-engineer cytochrome $P450_{BM3}$ for epoxidation of amorphadiene, developing a novel route for production of the anti-malarial compound artemisinin. A model structure of the lowest energy transition state complex for amorphadiene in the $P450_{BM3}$ active site was created using ROSETTA-based energy minimization. The resulting enzyme variant produced artemisinic-11S,12-epoxide at titers greater than 250 $mg \cdot l^{-1}$. Continued attempts to use ROSETTA and to either improve $P450_{BM3}$ epoxidase activity or introduce hydroxylase activity, however, proved unsuccessful. In the absence of a high-throughput screening approach, further improvement of the P450-based production system would be difficult.

As with most small-molecules, there exists no known high-throughput screen for artemisinic-11S-12-epoxide, amorphadiene, or any structurally-related compound. We hypothesized that a generalized method for high-throughput screen or selection design could be based on transcription factor-promoter pairs responding to the target small-molecule. Transcription factors have long been used to construct whole-cell biosensors for the detection of environmental small-molecule pollutants[1], but the work has remained largely un-translated toward screen development. While no known transcription factor binds artemisinic epoxide, a putative transcription factor-promoter pair responsive to 1-butanol, a biofuel molecule of interest in our laboratory, was recently reported[2]. The transcription factor, BmoR, and its cognate promoter, P_{BMO}, were used to build a short-chain alcohol biosensor for use as a genetic screen or selection. Following optimization of expression temperature, promoter, and reporter 5'-untranslated region, among other parameters, the BmoR-P_{BMO} system was shown to provide robust detection of 1-butanol in an *E. coli* host. The biosensor transfer function – relating input alcohol concentration to output fluorescent signal – was derived for 1-butanol and structurally related alcohols using the Hill Equation. The biosensor exhibited a linear response between 100 μM and 40mM 1-butanol, and a dynamic range of over 8000 GFP/OD_{600} unit. A 700 μM difference in 1-butanol concentration could be detected at 95% confidence. By replacing the GFP reporter with TetA, a tetracycline transporter, a 1-butanol selection was constructed; *E. coli* harboring the TetA-based biosensor exhibited 1-butanol dependent growth in the presence of tetracycline up to 40 mM exogenously added 1-butanol.

Demonstration of the biosensor in various high-throughput screening and selection applications first required construction of a 1-butanol production host. Studies have reporter 1-butanol production in *E. coli* through heterologous expression of either the *C. acetobutylicum* 1-butanol biosynthetic pathway[3], or a 2-keto acid-based pathway composed of a *L. lactis* 2-keto acid decarboxylase, KivD, and the *S. cerevisiae* alcohol reductase, ADH6[4]. In our hands, the *C. acetobutylicum* pathway proved non-robust and yielded low titers. In contrast, high-titer production of user-defined 2-keto acid derived alcohols was achieved by introduction of a *ΔilvDAYC* knockout in *E. coli* and expression of KivD and ADH6. The engineered strain is auxotrophic for 2-keto acids, and 1-butanol was produced by supplementing the growth medium with 2-oxopentanoate. A liquid culture screen was demonstrated using a 960-member KivD and ADH6 ribosome binding site library. Using the TetA-based biosensor, a strict cut-off between analyte 1-butanol concentration and biosensor output was observed. The assay led to the identification of a variant 2-keto acid-based alcohol production pathway exhibiting an approximately 20% increase in specific 1-butanol productivity.

Attempts to engineer concomitant 1-butanol production and selection in *E. coli* proved difficult. Both production and detection pathways functioned robustly when individually expressed in engineered *E. coli*; however, concomitant production and detection resulted in increased plasmid instability and cell death. We conclude by providing an analysis of observed cell stresses, generating negative 1-butanol selective pressures, and outline future strategies that can be used to address these hurdles.

Acknowledgements

I am first and foremost grateful to my parents, Steven and Marcia Dietrich, for their never ending support and love. My dad had the wisdom to allow me to explore the world unhindered, knowing that I would always find my way home, and perhaps learn something during the journey. My mom had the wisdom to remind me "to thy own self be true."
To Matt, I owe both perspective and understanding.
To Nature, I owe purpose.

If this
Be but a vain belief, yet, oh! how oft—
In darkness and amid the many shapes
Of joyless daylight; when the fretful stir
Unprofitable, and the fever of the world,
Have hung upon the beatings of my heart—
How oft, in spirit, have I turned to thee,
O sylvan Wye! thou wanderer thro' the woods,
How often has my spirit turned to thee!

–William Wordsworth, Tintern Abbey

Table of Contents

Chapter 1. High-Throughput Metabolic Engineering: Advances in Small-Molecule Screening and Selection[1]

1.1 The case for small-molecule screens

Natural selection, the force behind the amazing breadth of phenotypic variation in the living world, has long been a source of motivation for the engineering of synthetic biological systems. By mimicking the processes of mutation, recombination, and selection found in nature, directed evolution is used to impart industrial microbes with user-defined phenotypes. Frequently regarded as the First Law of Directed Evolution[1], the central tenet "you get what you screen for" draws attention to the paramount importance in finding an appropriate screen or selection assay to sift through vast libraries of variant hosts. A great body of work has been devoted to development of highly tailored assays specific for detection of single proteins or functions[2].

As applied to metabolic engineering, directed evolution is focused on improving small-molecule biosynthesis. Improving product yields or pathway efficiencies, however, can be a daunting task. Only a small subset of targeted compounds are natural chromophores or fluorophores that can be readily screened for using standard assay techniques. The majority of small-molecule targets for overproduction today do not illicit a conspicuous phenotype. For inconspicuous targets, chromatography-mass spectrometry methods have been the primary mode of detection; although they are nearly ubiquitously applied in small-molecule detection and quantification, these assays are inherently low throughput[3]. Screenable library sizes are generally limited to less than 10^3 variants. At this level of throughput, only a paltry number of rational modifications can be introduced into the panoply of host biosynthetic machinery, leaving the majority of sequence space untouched and unexplored.

Today, a discussion of improvements in small-molecule detection assays is set against the backdrop of increased focus in the field on microbial biosynthesis of commodity chemicals and fuels. Microbially produced alcohols, fatty acids, and alkanes are targeted for use as petroleum-derived fuel substitutes[4,5], aromatics[6], and diols[7], and polyhydroxyalkanoates[8] are being targeted for use as bioplastics. Spurred by an increased demand for renewable, green alternatives to petroleum-derived production routes, microbial production routes are competing economically against entrenched industry stakeholders[9]. Product yields and pathway efficiencies demanded from biosynthetic routes are pushing the limits of what can be achieved using existing metabolic engineering technologies. Efforts in the field are currently defined by the implementation of a series of rational design-based strategies to modify the host genome and heterologous pathway enzymes to achieve moderate product titers. However, metabolic engineering is a highly complex process, and product yields are dictated by a host of parameters. Biosynthetic pathways comprise multiple native and heterologous catalytic steps; each step is a potential bottleneck when directing carbon flux toward target small-molecule production. Furthermore, the host organism's native genetic network, regulation, and their interaction with the target pathway can all impact product yields. To achieve higher productivities, metabolic engineering must follow

[1] Reproduced with permission from JA Dietrich, AE McKee, JD Keasling. "High-Throughput Metabolic Engineering: Advances in Small-Molecule Screening and Selection." Annu. Rev. Biochem. **79**:563–590 **(**2010).

the decades-long trail of successes in protein engineering and develop more elegant approaches to high-throughput screening. Directed evolution through random and targeted mutagenesis of the host genome, overexpressed operon(s), and enzyme-encoded genes followed by high-throughput screening or selection is a requisite step in strain development.

In this review, we focus on advances in small-molecule screens using *Escherichia coli* and *Saccharomyces cerevisiae* as model prokaryotic and eukaryotic hosts. These organisms were selected both for their demonstrated application in industrial fermentation processes and because the vast majority of novel screening and selection processes are first demonstrated in these hosts. Given the immense number of mutable genetic elements to be targeted in metabolic pathway evolution, special consideration must be given to library design and sequence coverage. Screening efficacy has been demonstrated to increase in libraries that are maximally diverse, providing evidence for increased library size and mutation rate as methods for improving the diversity in the sample population[10]. Technical limitations imposed during library generation, transformation, and screening technologies are addressed below, and we focus in particular on screening and selection as rate-limiting steps in directed evolution efforts for small-molecule overproduction. Throughout our discussion, we highlight novel biosensor-based assays that enable inconspicuous small-molecule detection.

1.2. Statistical limitations on library size and sequence space coverage

Any directed evolution experiment, regardless of target, requires a thoughtful analysis of the library size required to gain significant coverage of the targeted sequence space. For good reason, most directed evolution efforts use focused libraries, mutating a relatively small number of preselected positions for saturation mutagenesis. However, even the most straightforward efforts to introduce a small number of substitutions are subject to harsh statistical realities. Simultaneous alteration of *n* selected positions in a given sequence necessitates the creation and screening a library of size *L*, according to Equation 1:

$$L = X^{n} \quad X = \begin{cases} 4 \text{ for, nucleotide substitution} \\ 20 \text{ for, amino acid substitution} \\ 2 \text{ for, genome (binary output)} \end{cases} \qquad \textit{eq 1}$$

Here, *X* corresponds to the number of possible genetic elements or states that may be present. The four naturally occurring nucleotides and the 20 naturally occurring amino acids provide the set of states for standard DNA and protein mutagenesis, respectively. The number of states for genome modifications, two, is modeled on deletion and insertion libraries. Although our discussion here has focused predominantly on randomization of existing genetic elements, a similar analysis can be applied to other, equally as important diversity-generating techniques, including insertions[11,12], deletions[11-13], and recombination[14,15], among others.

Exhaustive sequence coverage when randomization is not targeted to specific positions, but is instead applied uniformly over the full length of a target sequence, is a difficult prospect. When there exists little-to-no basis for rationalized substitutions using structure elucidation, or otherwise, the researcher may opt to introduce multiple, random mutations across a sequence of length *K*. The binomial coefficient describes the number of possible variants, *L,* given *N* randomizations with *X* genotypic states. Library sizes now scale according Equation 2:

$$L = X^{n} \frac{K!}{N!(K-N)!} \qquad \textit{eq 2}$$

Exhaustive library coverage for an indicated sequence space differs tremendously between the focused and random mutagenesis approaches. For example, mutation of 2 positions in a 100-amino acid protein (i.e., $X = 20$) yields a library of 400 unique members when the amino acid positions have been preselected and 1.98×10^{6} unique members when mutations are randomly incorporated over the length of the entire protein-coding sequence. The four order-of-magnitude difference in library size is of note, but underlying this analysis is the finding that most random mutations have neutral or deleterious effects on protein function[16]. Presupposing that some knowledge of protein structure or function can be used to guide a focused mutagenesis strategy, screening efficiency can be dramatically improved. These arguments exemplify why the vast majority of protein engineering efforts possess either a strong screening/selection assay or introduce focused mutations on a small subset of the total sequence space.

Given its small number of genotypic states, statistically speaking, targeted mutagenesis of the genome appears to provide the most straightforward approach. However, a number of caveats must be considered. First, the simplifying assumption was made that genomic elements have only on and off states, a route that ignores the importance of intermediate behaviors. For example, in the case of promoter insertions, induction profiles can range from all-or-none to graded responses[17]. Although library size is dependent only on the presence or absence of a promoter at a given position, assay size scales with the number of induction conditions tested. An additional caveat is that relatively little is known about the function of a large fraction of the genomes in experimental and industrial-use host microbes, making prediction of the effects of a targeted mutagenesis strategy difficult. For example, the genome of *E. coli* K-12 MG1655, the most well-studied microbe, contains approximately 4288 annotated protein-coding genes; of these, 19.5% remain of unknown function[18,19]. Without substantial a priori knowledge, the host genome, pathway operons, and its constituent enzymes all remain tenable targets for mutagenesis.

1.3. Technical limitations in library construction and screening

Even though statistical limitations establish a glass ceiling, hindering the exhaustive search of large sequence spaces, technical limitations in library generation and screening are the significant bottlenecks in practice. The workflow for a standard directed evolution assay can be divided into discrete segments: (*a*) *in vitro* diversity generation, (*b*) transformation and *in vivo* small-molecule production, and (*c*) screening and selection. Each step can impose significant technical limitations on the efficient exploration of sequence space.

1.3.1. Diversity generation

Diversity generation stands as perhaps the most robust step in mutant library screening. To a first approximation, technologies for diversity generation are agnostic to a genetic element's downstream, *in vivo* end function. For example, an experimentalist's success in introducing genotypic variability into a protein-coding sequence using error-prone polymerase chain reaction (PCR) is independent of the downstream function of the protein. Disconnect between diversity generation and end function is due, in part, to segregation of in vitro diversity generation from downstream *in vivo* expression. Creating genetic diversity in vitro allows for an extremely large number of variants to be created; plasmid libraries on the order of approximately 10^{14} molecules can be tractably prepared, amounting to 1 mg of plasmid DNA[20]. For everyday benchtop experiments, however, an upper limit of 10^{12} molecules is more commonly observed.

Until recently, a lack of computational estimates of library sequence diversity, using various methods, left experimentalists to wander through sequence space. Although still seemingly underutilized, in silico approaches to model diversity generation in error-prone PCR, gene shuffling, and oligonucleotide-directed randomization have been exhaustively reviewed[21-23]. When coupled with a readily accessible user interface, computational methods can be valuable guides to choose mutagenesis methods satisfying an experimentalist's library size, diversity, and coverage objectives. To this end, a suite of user-friendly diversity analysis software tools have been made readily available online[24]. The programs provide a useful statistical analysis of library diversity and sequence space coverage, which can guide library construction and screening when using error-prone PCR, site-directed mutagenesis, and in vitro recombination[10,25,26]. There still remains an upper limit on the sequence space that can be analyzed by computationally predictive methods; one recently published method for modeling random point mutagenesis establishes an upper limit of analyzing 2000 amino acids or 16,000 nucleotides, and 10^9--10^{10} individual sequences[27]. This level of computation power continues to meet or exceed the reasonable number of DNA variants that can be constructed using existing DNA synthesis technologies.

1.3.2. Transformation efficiency

The efficiency of introducing variability into the host to gain *in vivo* functionality provides the next significant bottleneck. The method most commonly employed for both *S. cerevisiae* and *E. coli* library incorporation is nucleotide transformation, our focus here. For *E. coli*, the maximum library size is estimated to be on the order of 10^{12} molecules[28], but libraries on the order of 10^9 transformants are more readily realized at the benchtop scale of everyday experiments. Library sizes in *S. cerevisiae* expression systems are subject to decreased transformation efficiencies, and

maximum library sizes are approximately an order-of-magnitude less than those witnessed for *E. coli*.

Transformation of an *in vitro* library into the *in vivo* screening context can entail significant losses in library size, and steps can be taken to circumvent transformation inefficiencies. One option is to generate and screen a library in vitro; the in vitro compartmentalization, mRNA display, and ribosome display methods have been reviewed elsewhere[29]. Commonly used for directed evolution of single proteins, completely in vitro assays are not generally applicable to metabolic engineering for small-molecule production and ignore *in vivo* biological context and regulation. The second, more commonly employed approach to avoiding library transformation inefficiencies is to develop the library diversity *in vivo*. Published methods include employing engineered strains of *E. coli* with increased mutation rates[30], an error-prone *E. coli* polymerase I (Pol I) for more targeted diversity generation in plasmids containing a ColE1 origin[31], and somatic hypermutation in mammalian expression hosts[32,33]. The general drawback to *in vivo* diversity generation is the introduction of untargeted, sometimes genome-wide, mutations. Circumventing a this issue, multiplex automated genome engineering (MAGE)[34], uses transformed ssDNA to target individual genomic regions for mutagenesis at high efficiency (>30%).

1.3.3. Small-molecule screening

Last in the directed evolution flowchart stand screening and selection. Technologies for small-molecule screening have long lagged behind those for diversity generation, predominantly owing to a need to independently tailor assay methods for application toward different target small molecules. For this reason, screening and selection processes are the most significant bottleneck in directed evolution efforts for small-molecule detection and quantification.

Although throughput can be generalized for specific screening technologies (**Figure 1.1**), small-molecule assays are burdened by additional, equally important parameters. Screen and selection strategies are only as good as their sensitivity and selectivity: A desired phenotype that remains undetected will not be captured, regardless of the number of variants analyzed. For this reason assay sensitivity, dynamic range, and linear range of detection – parameters commonly used to describe transfer functions in genetic circuit design[17,35-38] – are also apt for the characterization of small-molecule screens and selections (**Figure 1.2**). While virtually unreported in phenotypic assays described to date, this quantitative framework provides for a minimal set of parameters that enable more accurate comparison between different assay methodologies. The caveat being that, given the aforementioned molecule-tailored nature of most screens, codified rules regarding screen sensitivity and dynamic range can be ineffectual generalizations.

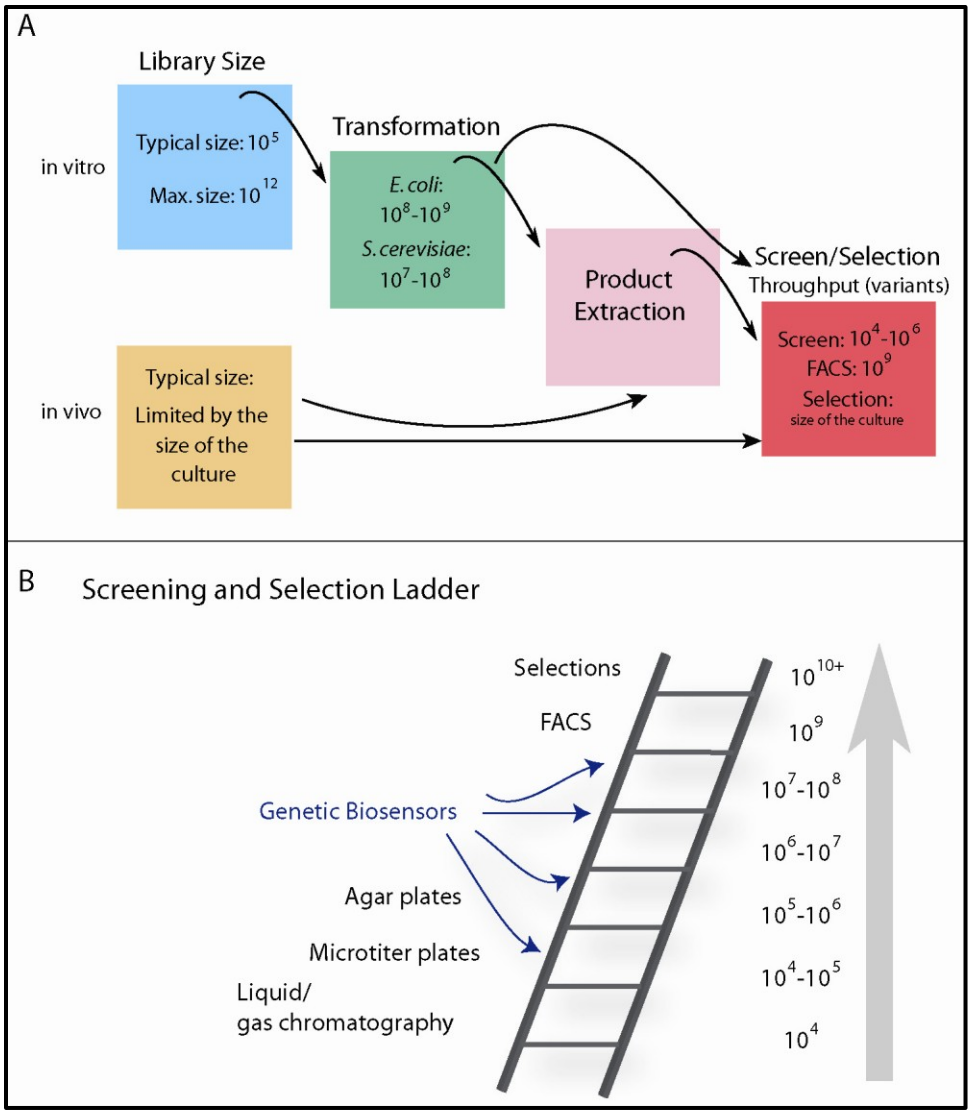


Figure 1.1: Inefficiencies in library screening. (*a*) The in vitro and *in vivo* steps associated with standard directed evolution dictate the scope of sequence space that can be explored. Moving between in vitro and *in vivo* compartments imposes significant losses in library size and diversity. *In vivo* methods of library generation avoid inefficiencies of transformation and product extraction; however, mutations are randomly inserted across a target sequence and cannot be focused to a few key positions. (*b*) The choice of screening assay will ultimately have the biggest impact on throughput; genetic biosensors converting small-molecule concentration into a readily detectable reporter molecule serve as one method to move away from low-throughput chromatography techniques. Abbreviation: FACS, fluorescence-activated cell sorting.

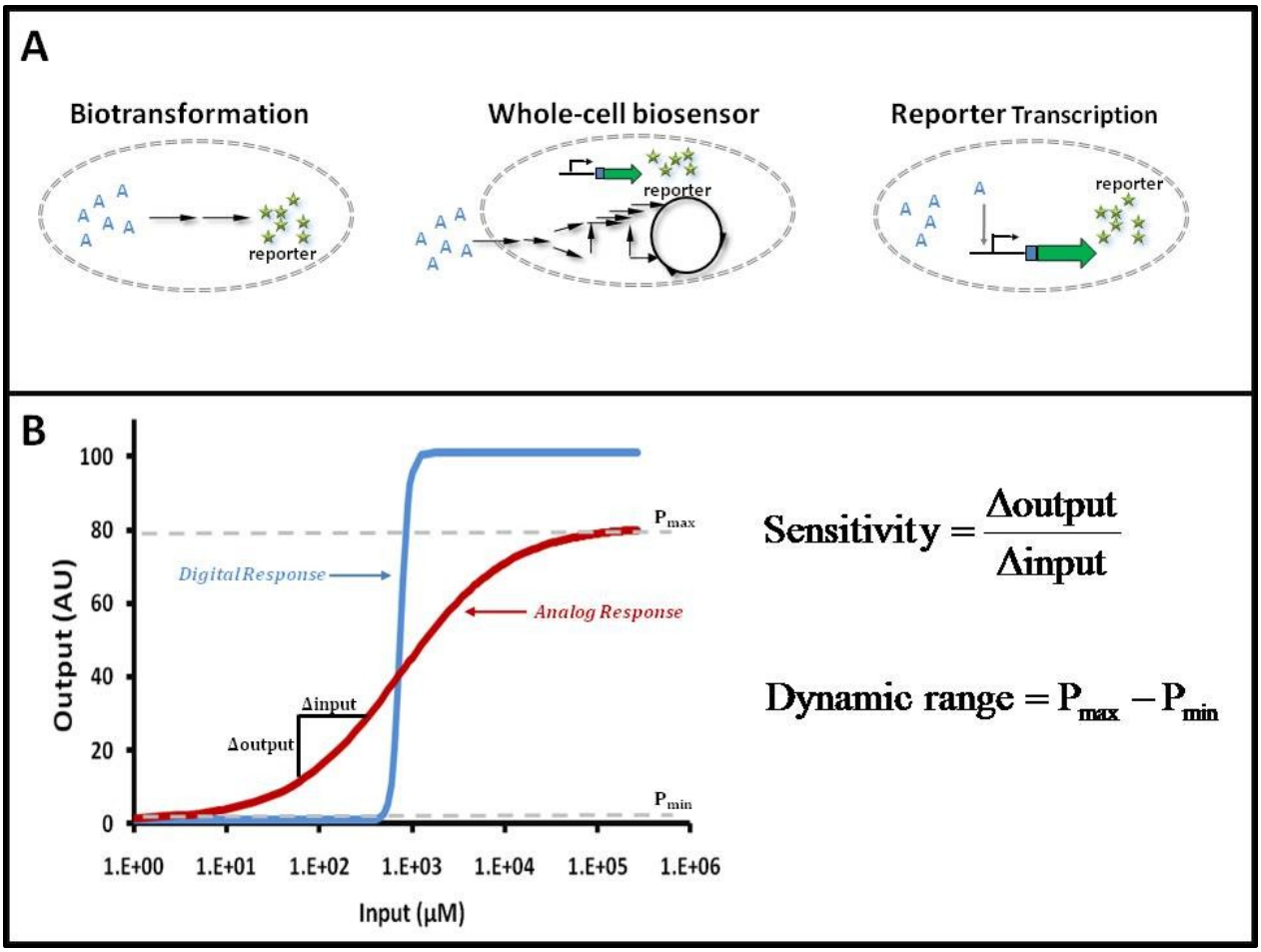


Figure 1.2: Transfer curves enable a quantitative characterization of performance features. (*a*) Biosynthesis of an inconspicuous small-molecule, A, results in a detectable signal by transformation using a series of biosensor constructs. Small-molecule A can be transformed into a detectable reporter molecule through the action of single- and multistep enzymatic pathways. Whole-cell biosensors couple growth (and a constitutively expressed reporter) with a host microbe's productivity. Lastly, reporter transcription can be achieved using classically regulated promoter systems induced by small-molecule A. (*b*) The correlation between input product concentration and output biosensor response (arbitrary reporter units) provides the biosensor transfer function. Transfer functions provide information on product sensitivity, the linear range of detection, and the detection threshold, which can guide assay design.

Foundational work in directed evolution for small-molecule production has been focused on conspicuous targets, those that can be optically detected. Without the need for additional synthetic chemistry or biotransformation, conspicuous products can be accurately measured using standard high-throughput colorimetric and fluorometric assays. In contrast, inconspicuous targets, a class to which the majority of small molecules belong, emit no spectral signature suitable for existing high-throughput screening technologies. Inconspicuous small molecules can be converted into detectable outputs through the activity of biosensors. Biosensors take form as single- or multistep enzymatic pathways, inducible expression systems, and entire host organisms; an accurate description of the correlation between a biosensor's inputs and outputs is provided by its transfer function (**Figure 1.2**)[38]. In the remainder of this review, we provide a discussion of small-molecule screens and selections and present methods for moving up the

screening and selection ladder (**Figure 1.1**). This qualitative metric seeks to incorporate throughput, sensitivity, and dynamic range into a generalized rank of assay strength and molecule applicability. Small-molecule screens exhibiting high-throughput and sensitive analyte detection for a broad spectrum of compounds rank higher than less-specific, lower-throughput assays. In our analysis of the various small-molecule screening methods commonly employed, we highlight the role of biosensor-driven assays that enable inconspicuous small-molecule targets to leapfrog rungs in the ladder.

1.4. The screening and selection ladder

We present here colorimetric, fluorometric, and growth-complementation assays as the foundation for small-molecule assays in metabolic engineering. Excluded from this review are gas and liquid chromatography and H^1- and C^{13}-based NMR techniques for small-molecule identification and quantification. Although ubiquitous in the field and offering unparalleled accuracy and precision, their extremely low throughput limits their application to assays with small sets ($< 10^3$) of modifications.

In engineering terms, we define our system as an individual microbe, with the target product being an output. As such, we focus only on whole-cell assays for target small-molecule production. End-product screens detect the desired output (i.e., improved molecule production) and are applicable with directed evolution of any upstream biosynthetic machinery. In this sense, small-molecule screens offer greater versatility and application than screens for intermediate functions. We leave untouched assays that analyze single or intermediate steps in the conversion of a carbon source to product.

For each assay technology, examples of conspicuous small-molecule detection are used to provide a backdrop for development of next-generation, biosensor-driven approaches. Screen throughput, sensitivity, and dynamic range, where available, are highlighted from exemplary studies in the field.

1.4.1. Colorimetric and fluorometric plate-based screens

The relative ease of conducting colorimetric and fluorometric assays has established their position as the a posteriori techniques of choice for proof-of-principle experiments in novel mutagenesis and directed evolution[39]. Assay sensitivity, linear range of detection, and, to some extent, throughput will vary depending on the method used and the compound being interrogated. Individual variants are monitored as liquid cultures on microtiter plates or as colonies on solid, agarose media. Photometric assays using microtiter plates are highly robust and provide the distinct advantage over other techniques in being able to broaden the linear range of detection by diluting or concentrating the sample. Assay throughput on microtiter plates is moderate (approximately 10^5 variants per experiment) and is strongly affected by an oft requisite in vitro product extraction step. In comparison, screening colonies on solid media provides increased throughput, and with modern robotics upward of 10^6 variants are screened[40]. The increase in throughput comes at the expense of a greatly diminished sensitivity, however, and small differences in intercolony productivities can be overlooked. Thus, the decision between liquid and solid media assays is largely dictated by available equipment and a compromise between throughput and sensitivity.

Detection of microbially produced photophores provides upper bounds with regard to throughput and sensitivity as no additional chemical transformations are required for product detection. From the expansive body of metabolic engineering work on production of natural photophores, we focus here on lessons learned during directed evolution for improved carotenoid biosynthesis. The carotenoids, including lycopene, β-carotene, and astaxanthin, were recognized early on as a high-valued neutraceuticals with remarkable antioxidant properties[41]. This feature, coupled with their relative ease in detection, has driven extensive research into establishing microbial production routes and, in large part, has laid the foundation for many aspects of metabolic

engineering. To date, lycopene has been the major carotenoid of focus for production in microbial hosts. The majority of titer-oriented research for lycopene has been conducted in *E. coli* and has highlighted a number of the performance features of colorimetric screens. Lycopene production can be monitored colorimetrically by measuring the absorbance at 470 nm following extraction into an organic solvent[42]. The assay is highly sensitive, differentiating between submilligram per liter differences in lycopene yields[43] and achieving a level of sensitivity more than adequate for directed evolution. Detracting from the lycopene screen, however, is a requisite organic solvent extraction that drives up assay price while decreasing throughput.

The success of plate-based photometric screening is exemplified by use of the carotenoids in various proof-of-principle experiments in organism and protein engineering. The broad spectrum of colored pigments found in the carotenoid family was utilized to demonstrate strategies in combinatorial biosynthesis[44,45], a powerful approach wherein promiscuous enzymes serve as molecular scaffolds on which nonnative small molecules are synthesized. Additionally, advances in mRNA-based regulation of pathway flux[46], model and combinatorial-driven methods in genome engineering[47], engineered metabolic control[48], and multiplex genome engineering[34] have been enabled.

In contrast to the carotenoids, the vast majority of primary and secondary metabolites targeted for overproduction in the laboratory are not natural chromo- or fluorophores. For these molecules, in vitro synthetic and enzyme-coupled catalyses may offer an alternative approach to direct product detection. Using synthetic chemistry, a target molecule-specific chemical moiety reacts with an exogenously added reagent to yield a detectable product. Fluorescent and colorimetric detection of specific chemical moieties and compound classes is an area of great interest in organic synthesis. A wealth of both demonstrated and potential high-throughput assays can be borrowed from the field, including thiols[49], cyclic and linear amines[50,51], carboxylic acids[52,53], alcohols[54], aldehydes[55], and glycosaminoglycans[56], among others[57,58]. Alternatively, indirect methods for product detection can also be employed, such as product-associated changes in pH[59,60] or coproduction of hydrogen peroxide[61]. The application of synthetic chemistry-based detection to metabolic engineering strategies, however, is not without limitation. Accurate product quantification in synthetic chemistry-based detection schemes may be impeded by a low signal-to-noise ratio, as the majority of moieties being targeted are naturally found at high abundance in the microbial intracellular environment. This issue may be minimized by increasing production titers. Lastly, reagent cost when incorporating in vitro synthetic chemistry may become a significant factor, a problem exacerbated with increasing library size.

When synthetic catalyses are either unavailable or cost prohibitive, or when higher substrate specificity is required to alleviate background noise, enzyme-coupled assays provide another approach. Coupling enzymes are added to the analyte solution to form single- or multistep pathways that yield detectable photophores or utilize traceable cofactors in their catalysis. The distinct advantage in cofactor-dependent assays is the broad range of enzymes and reaction types for which they are necessary; thus, cofactor monitoring can be viewed as a strategy more universally applicable than direct detection methods. Photometric assays have been described for quantification of many enzyme cofactors, ATP and ADP[62-64], reduced and oxidized states of NAD and NADP[65-67], and free coenzyme A[68]. When background noise from native, endogenous

small-molecule production is not an issue, using coupling enzymes with broad substrate acceptance is an interesting option. For example, substrate promiscuity is a hallmark characteristic of the P450 superfamily[69], and NADH or NADPH consumption in P450-catalyzed reactions is frequently used as an indirect measure for substrate oxidation[70,71]. Similarly, *S*-adenosyl-L-methionine is a cofactor for small-molecule methyltransferases and can be monitored following multienzyme biotransformations to homocysteine or hypoxanthine[49,72,73]. Although inherently indirect measures (and as such push the boundaries of the rule "you get what you screen for"), the cofactor assays nonetheless are often highly robust, accurate proxies for direct product detection[62,72,74].

Design of *in vitro* small-molecule assays using either synthetic or biocatalyzed transformations requires significant thought with regard to the reagents and enzymes used. Although not a prerequisite, driving a reaction to completion is highly desirable and will facilitate a more accurate back calculation of target small-molecule production titers. For this reason, use of irreversible enzymes without product inhibition is the preferred enzymatic route. The choice of catalyst also determines the reaction conditions. Under the best-case scenario, assays are performed directly in the growth medium without cell removal, product extraction, or pH adjustment; however, optimal reaction conditions always need to be determined experimentally. These considerations become increasingly important as library sizes grow, as each additional in vitro manipulation adds both significant consumable and throughput costs. Lastly, the synthetic reagents or coupling enzymes must also be economically synthesized, purified, or purchased to screen a complete set of library variants.

By shifting from an *in vitro*, enzyme-coupled product detection regime to an *in vivo* context, significant advantages can be garnered in terms of both throughput and cost. Extractions, enzyme purifications, and in vitro manipulations are minimized or eliminated. Catalytic biosensors, single and multistep *in vivo* biosynthetic pathways with inconspicuous substrate inputs and detectable product outputs, have just begun to be explored. Examples to date include assays for intermediates in an engineered *Pseudomonas putida* paraoxon catabolic pathway[75], strictosidine glucosidase-catalyzed transformation of tryptamine analogs[76], and a tyrosinase-catalyzed transformation of the amino acid L-tyrosine into melanin[77].

Directed evolution of *E. coli* for improved L-tyrosine production exemplifies the role enzyme-based biosensors can assume during inconspicuous small-molecule detection. Tyrosine is a colorless, essential metabolite of great import in the synthesis of a wide range of value-added pharmaceuticals and commodity chemicals, including the morphine alkaloids and *p*-hydroxystyrene[78]. Rational engineering strategies, targeting both tyrosine and its intermediates, have been extensively explored and reviewed[79,80]. Concomitant work in the Stephanopoulos lab[77,81] in both rational engineering and directed evolution of *E. coli* tyrosine biosynthesis enables a thorough comparison of the two approaches. Using strictly rational metabolic engineering, two tyrosine-overproducing strains of *E. coli* were reported[82]. The first, T1, incorporated feedback-inhibition-resistant derivatives of pathway enzymes and eliminated native pathway regulation. Building on T1, strain T2 increased the availability of the central metabolite precursors D-erythrose-4-phosphate and phosphoenolpyruvate necessary for tyrosine biosynthesis. Tyrosine production titers in minimal medium culture flasks reached 346±26 and 621±26 $mg \cdot L^{-1}$ for strains T1 and T2, respectively.

Before using random mutagenesis to build on their rationally designed strains, the authors developed a pair of high-throughput screening assays. The first, although not biosensor based, is an interesting application of an in vitro synthetic chemistry approach. It has long been known that 1-nitroso-2-naphthol reacts with parasubstituted phenols, such as tyrosine with high specificity, yielding a red-orange product[83]. By modifying the published assay methods the authors developed a high-throughput, microtiter-based fluorometric screen for L-tyrosine[81]. This assay was then used to screen a small, combinatorial library of amino acid biosynthetic genes overexpressed in *E. coli*[84]. The authors report that their screen accurately differentiates between product concentrations as low as 50 $mg \cdot L^{-1}$ over a linear range of 0.05--0.5 $g \cdot L^{-1}$ tyrosine[81]. As with lycopene, throughput is limited to the order of 10^5 variants per experiment owing to requisite use of microtiter plates and in vitro synthetic chemistry.

More recently, Santos & Stephanopoulos[77] describe a catalytic biosensor approach using an *in vivo* expressed tyrosinase to use melanin as a reporter of tyrosine productivity (**Figure 1.3**). A black, insoluble pigment, melanin, can be readily assayed in colonies grown on solid medium without need for additional in vitro synthetic chemistry or solvent extractions. Starting with a base *E. coli* strain producing 347 $mg \cdot L^{-1}$ tyrosine (similar to the above-mentioned rationally engineered strain T1), the authors constructed a transposon-mediated knockout library of the *E. coli* genome and screened 21,000 colonies on agar plates. Over two five-day rounds of screening, 30 variants were selected for a more rigorous characterization; two mutants were discovered that individually produced 57% and 71% higher L-tyrosine titers than the background strain. The associated chromosomal knockouts occurred in *dnaQ* and *ygdT*, whose gene products are part of the epsilon subunit of PolIII and a hypothetical protein, respectively. Under similar culture conditions, the *yadT* knockout strain exhibited product yields comparable to those measured in the highest producing rationally engineered strains[77,82]. Neither of these genomic knockouts could have been predicted using rational design or flux analysis approaches, a feature that draws attention to the potential value of a strong-screening assays. The performance features for this assay were not provided and thus inhibit comparison to their synthetic chemistry-based route; however, visual assessment of colonies is in general a much less sensitive detection method.

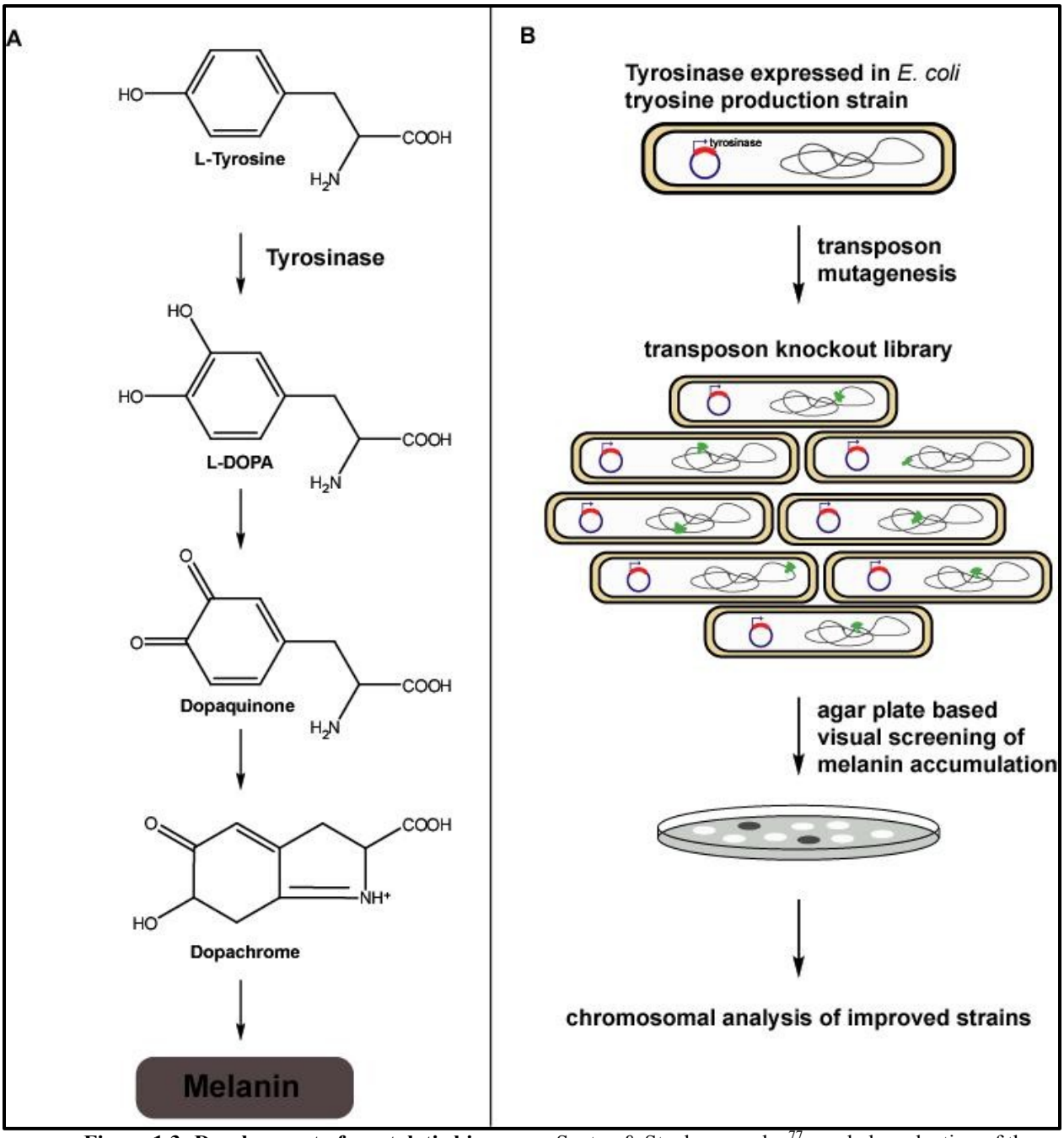


Figure 1.3: Development of a catalytic biosensor. Santos & Stephanopoulos[77] coupled production of the colorless amino acid l-tryosine to synthesis of the black, diffusible pigment melanin through heterologous expression of a tryosinase in *E. coli*. (*a*) Tyrosinases use molecular oxygen to catalyze the *ortho*-hydroxylation of l-tyrosine to l-DOPA, followed by its oxidation to dopachrome. The generated reactive quinones then polymerize to form melanin. (*b*) Expression of a bacterial tryosinase (*blue plasmid*) enabled a high-throughput screen, as melanin was used to report on the production of l-tyrosine. Agar plate-based screening of a transposon-mediated knockout library identified strains with increased tyrosine production. Analysis of the chromosomal mutations in the improved strains revealed interruptions in two genes, *dnaQ* and *yadT*. Neither gene would have been predicted to impact aromatic amino acid production through rational design or flux analysis strategies. This work illustrates the importance of strong screening assays in uncovering untapped genotypic improvements.

1.4.2. Growth complementation

Strains auxotrophic for essential small molecules are natural biosensors and have long been used to build strong selection assays[85]. These strains provide what is perhaps the most readily discernable phenotype, growth, if the auxotrophy is relieved by complementation of lost enzymatic function. Growth-complementation assays have been used extensively to engineer hosts for catabolism of nonnative or nonideal carbon, nitrogen, and phosphate sources[86,87]. Assay development is relatively straightforward: By supplementing the growth medium with a single molecular source of an essential element, only those host variants engineered or evolved for its catabolism survive.

Although highly successful in engineering novel catabolic activities, coupling anabolism to growth has proven to be a more difficult task. Oftentimes, an overproduction phenotype is deleterious to the host strain's reproductive fitness[88,89], a characteristic easily inferred by depressed growth curves in overproducing strains. To screen for anabolic activity, small-molecule production must complement the auxotrophy and be strongly correlated with the specific growth rate to enrich cultures for high producers.

Auxotrophies of components of amino acid biosynthesis pathways have been successfully utilized in metabolic engineering applications because the pathways are strongly coupled to growth. An early study using growth-complementation knocked out the gene encoding for branched chain amino acid aminotransferase *ilvE*; this strain was then used to evolve an aspartate aminotransferase to recognize branched amino acids[90]. Following five rounds of DNA shuffling, the catalytic efficiency for transamination between 2-oxovaline and aspartate was improved by five orders of magnitude. Further selection increased the catalytic efficiency of the final mutant variant by an additional order of magnitude[91]. Other examples from amino acid biosynthesis, which used similar strategies, include aminotransferases[92], an alanine racemase[93], and evolution of chorismate mutase[94,95].

Intermediates in amino acid biosynthesis, keto acids, have been more recently explored in a metabolic engineering context. Atsumi et al.[96] rationally engineered 1-propanol and 1-butanol production in *E. coli* using 2-ketobutyrate derived from threonine degradation. A more direct route to 2-ketobutyrate, however, is the citramalate pathway, identified in *Leptospira interrogans* and *Methanocaldococcus jannaschii* but not present in *E. coli*. An *E. coli* strain auxotrophic for isoleucine would require flux to be rerouted through a transformed citramalate pathway to restore growth. This strategy was employed in an *E. coli* host for functional expression and directed evolution of a heterologous citramalate biosynthetic pathway leading to 2-ketobutyrate[97]. The evolved strain exhibited 9- and 22-fold improvements in 1-propanol and 1-butanol productivities, respectively, over the wild-type citramalate base strain. Furthermore, when similar medium conditions and genetic backgrounds were employed, the evolved citramalate pathway provided greater than 35-fold improvement in 1-propanol yield and an eightfold improvement in 1-butanol yield. Given the large number of industrially important compounds that can be derived from amino acid biosynthetic pathways and the success of auxotrophic selection assays for these compounds, there is likely to be continued focus on these assays in the future.

A more generalized growth-complementation strategy is the use of an auxotrophic reporter strain. Here, a producer strain is engineered for overexpression of the target pathway, and a second reporter strain is constructed that constitutively expresses a detectable market (i.e., green fluorescent protein, GFP) and is auxotrophic for an essential metabolite found in the target pathway (**Figure 1.2**). When cocultured, the reporter strain's growth---and thus reporter output---is coupled to the metabolite yield achieved by the producer strain. This strategy was successfully employed in the development of a whole-cell biosensor for mevalonate (**Figure 1.4**)[98]. A 10% difference in mevalonate production could be discerned between liquid cell cultures with 95% confidence. When utilized in a spray-on technique for solid media screens, however, the assay was only effective in distinguishing mevalonate producers from nonproducers[98]; a finding that again highlights the higher sensitivity observed in liquid versus solid medium plate-based assays.

In general, auxotrophic strains provide a unique, high-throughput approach for the screening or selection of small-molecule overproducing strains. However, there remain some significant drawbacks when considering their broad-scale adoption. First, the essential small molecule must be native to the host or reporter microbe; thus, growth complementation is more suitable to building platform production strains for metabolite precursors than exotic secondary metabolites. Second, the dynamic range in growth-complementation assays can be limited. From a protein engineering perspective, even slightly functional biocatalysts can be sufficient to restore cell growth; in practice, this places a glass ceiling on the small-molecule product yields that can be accurately screened or selected. Addressing this challenge, user-defined transcription and enzyme degradation tags have been explored as means of increasing selective pressures and assay dynamic range[95]. Other readily accessible approaches for fine-tuning protein expression levels include decreasing plasmid copy number, modifying RNA stability, and modifying translation initiation efficiency[99].

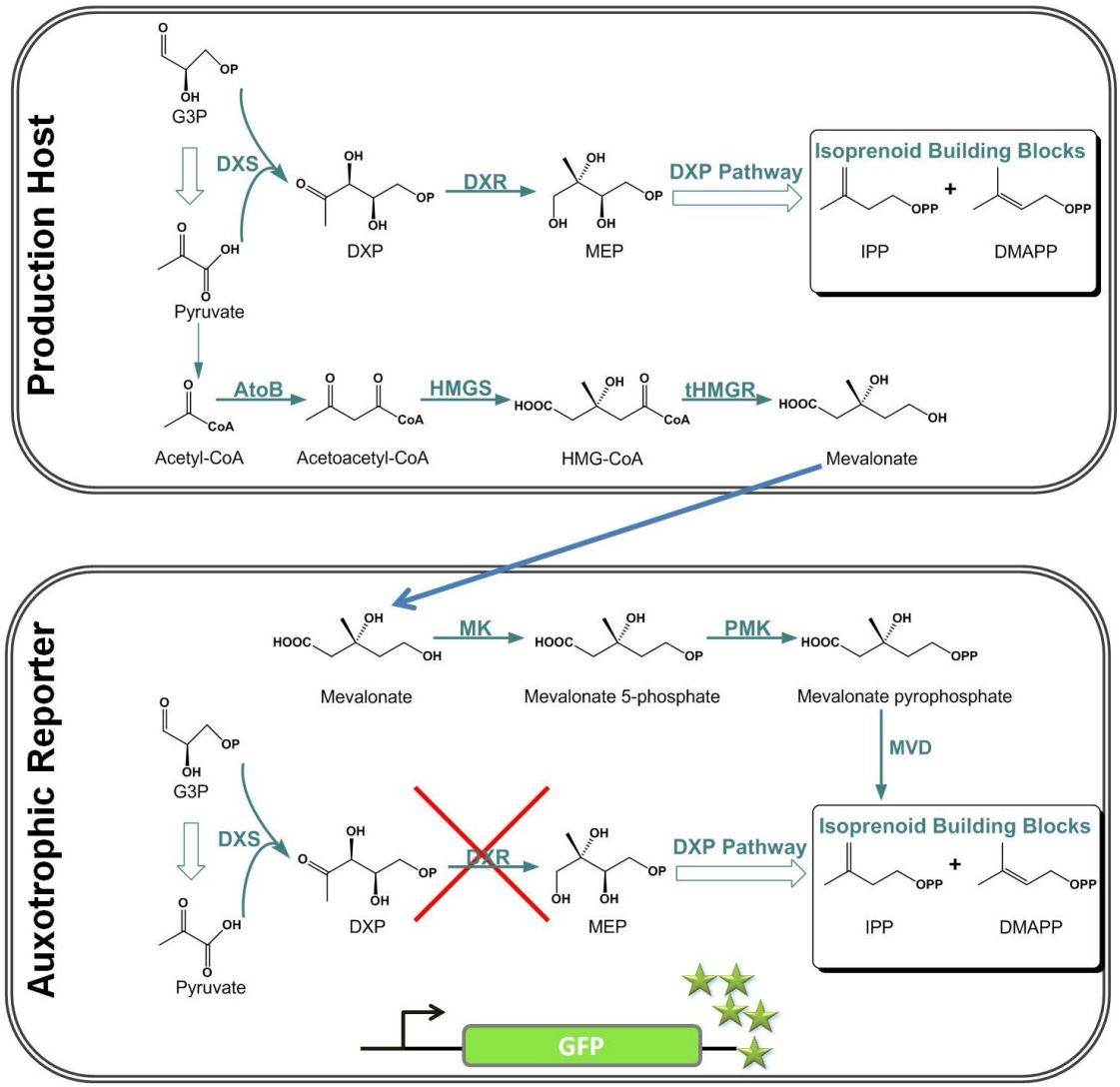


Figure 1.4: Development of a mevalonate biosensor. The isoprene subunits isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) are essential for *E. coli* cell growth. Natively, IPP and DMAPP are produced in *E. coli* through the 1-deoxy-D-xylulose-5-phosphate (DXP) pathway; however, expression of a heterologous mevalonate pathway provides an alternative route. By concomitantly knocking out the DXP pathway while overexpressing the enzymes necessary for mevalonate utilization, a mevalonate biosensor was constructed[98]. The mevalonate biosensor also contained a constitutively expressed gene for green fluorescent protein (GFP) to provide a fluorescent readout of cell growth. Mevalonate was supplied by a production host containing genes necessary for transformation of acetyl-CoA to mevalonate. Abbreviations: G3P, glyceraldehyde-3-phosphate; DXP, 1-deoxy-D-xylulose-5-phosphate; MEP, 2-C-methyl-D-erythritol-4-phosphate; IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; DXS, 1-deoxyxylulose-5-phosphate synthase; DXR, 1-deoxy-D-xylulose-5-phosphate reductoisomerase; AtoB, acetoacetyl-CoA thiolase; HMGS, HMG-CoA synthase; tHMGR, truncated HMG-CoA reductase; MK, mevalonate kinase; PMK, phosphomevalonate kinase; MVD, mevalonate pyrophosphate decarboxylase; GFP, green fluorescent protein.

1.4.3 Fluorescence-Activated Cell Sorting-Based Screening

Near the top of the phenotypic screening and selection ladder stands fluorescence-activated cell sorting (FACS). Through rapid analysis of size and fluorescence measurements of single cells, this technique possesses nearly ideal high-throughput screening characteristics[100]. Libraries are commonly first passed through a primary FACS screen, thereby enriching the population

between 80—5000-fold for the 0.5% to 1.0% cells exhibiting the highest measured fluorescence[101-105]. Library sizes between 10^8 and 10^9 variants have been readily screened in assays for modified protein specificity[101,106], achieving the realistic 10^9 variant cutoff incurred owing to low microbial transformation efficiencies[28]. Those cells surviving the primary screen are subjected to a secondary screen to provide phenotypic confirmation. In practice, between 10 and 10^3 clones enriched by the primary FACS screen are analyzed by gas chromatography-mass spectrometry (GC-MS) for identification or quantification[107]. Although FACS is ultra high throughput, a drawback is overlapping fluorescence profiles, and aberrant fluorescence can lead to a high rate of false positives during screening. This arises from the distribution of fluorescence values found in a population when single cells are analyzed, a problem not witnessed in plate-based screens because of population averaging (**Figure 1.5**). For this reason, FACS-based screens are used to enrich for a target population and are then followed by secondary, orthogonal screens to provide a more accurate validation.

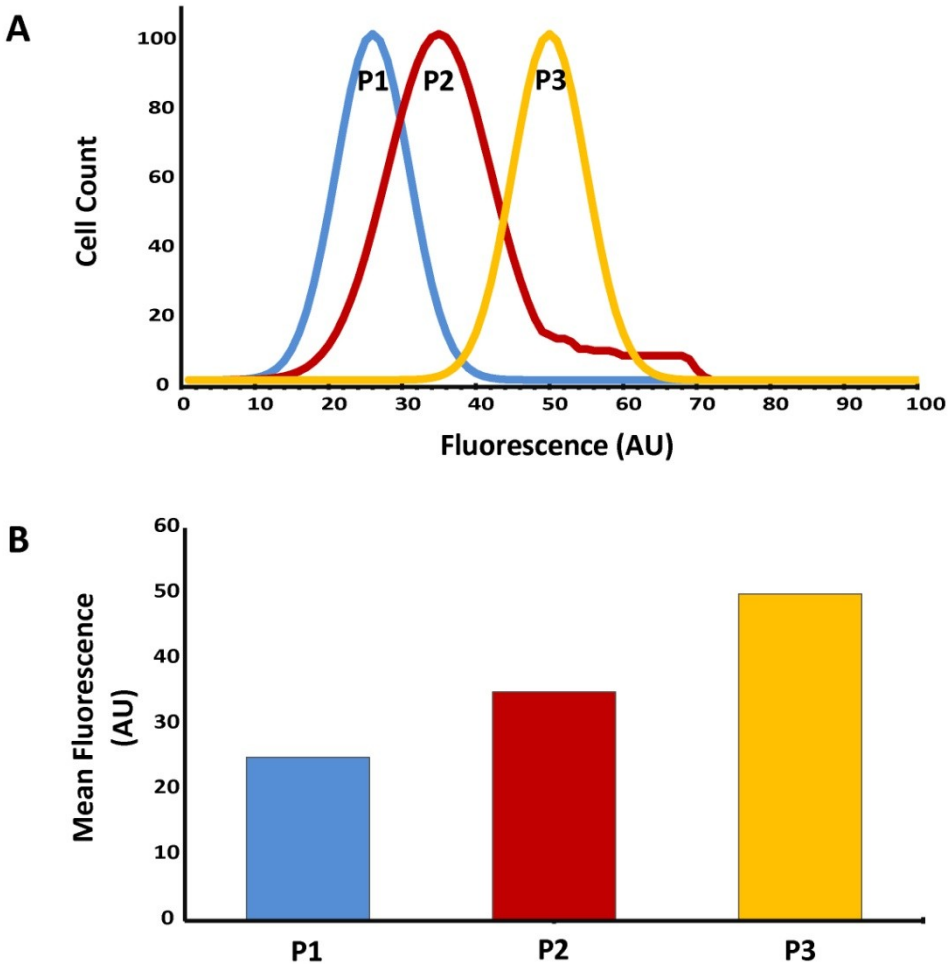


Figure 1.5: Population distributions can inhibit efficient cell sorting. (*a*) Representative data collected from fluorescence-activated cell sorting (FACS) and (*b*) fluorometer analyses of output reporter fluorescence highlight the difficulties in efficient screening. We examine the hypothetical strains P1, P2, and P3, displaying different levels of target small-molecule production. Although the mean fluorescence for all cell populations does not change between fluorometer and FACS instruments, population distributions inhibit efficient cell sorting by FACS. Overlapping population profiles (separating P2 from P1 and P3) and aberrant fluorescence (the right-hand

tail on P2) can decrease enrichment efficiencies. Lower-throughput, orthogonal secondary screens must be used to further purify the target population. Abbreviation: P1, 2, 3 hypothetical populations.

Direct detection of fluorescent proteins and metabolites using FACS is a straightforward application and has been well explored in both protein and metabolic engineering. In metabolic engineering, naturally fluorescent metabolites can be directly screened so long as they are intracellular and so long as their emission spectra are not masked by that of the native host. Again, the carotenoids have served as model small molecules for isolation of hyperproducing yeast strains. The inherent fluorescent property of astaxanthin was used to recover yeast strains at high sensitivity, capturing productivity improvements as low as 260 $\mu g \cdot g^{-1}$ dry cell weight[108,109]. A recent effort to improve astaxanthin production from *Xanthophyllomyces dendrorhous* achieved a 3.8-fold production improvement and reported fivefold improvement in screening efficiency as compared to plate-based methods[110]. Beyond the carotenoids, FACS screening has been restricted to a small number of natural and synthetically coupled fluorophores, including gramicidin S[111], polyhydroxyalkanoates[112], poly(β-hydroxybutyrate)[113], and lipids[114,115].

Another route for FACS-based screening is to assess those changes to cellular physiology induced by small-molecule production or cell stress. In this indirect method, fluorescent dyes act as indicators for changes in intracellular pH or viability[116]. Following this strategy, strains of *S. cerevisiae* capable of improved lactic acid tolerance and production have been isolated from a mutant library, using both ethidium bromide staining and a pH-dependent fluorescent probe[117]. The fluorescent probe used, cSNARF-4F, exhibits a shift in the spectral emission depending on intracellular pH; the ratio of the two local maxima indicate the pH[118]. Because the assay is physiology dependent and not product dependent, it can be extrapolated toward detection of a range of microbially synthesized weak organic acids and bases.

The approach of screening for a downstream effect of molecule overproduction, and not for the small molecule itself, may be a versatile strategy given that entire classes of molecules can produce similar changes in physiology. For instance, FACS-based screens could prove effective in cases of product-induced oxidative stress, a common side effect of P450 overexpression[119], using *in vivo* fluorescent probes for reactive oxygen species[120,121]. Similarly, membrane fluidity[122], intracellular magnesium[123], and intracellular calcium[124] can also be quantified using fluorescent dyes. Despite the straightforwardness of this screening approach, it is hampered by the dearth of available *in vivo* reporter probes. In addition, it remains unknown whether a strong, reproducible correlation can be drawn between small-molecule biosynthesis and a microbe's physiological state over a large set of molecule classes.

Genetic biosensors can take inconspicuous small-molecule inputs and output one of several available fluorescent proteins as reporters. In general, methods described to date have focused on transcriptional read through from product-encoding operons as a proxy for end function. One interesting approach has been construction of genetic traps, screens used to isolate genetic elements from metagenomic libraries that encode for user-desired phenotypes. SIGREX, or substrate-induced gene-expression screening, uses a fluorescent reporter housed behind a multiple cloning site into which a metagenomic library is inserted[125]. When a small-molecule substrate is exogenously added, only clones harboring metagenomic inserts whose transcription is activated by the substrate small molecule will express the downstream fluorescent reporter. A

second, negative screen under uninduced conditions can remove variants with constitutive fluorescence. For the insert to produce a fluorescent signal, it must encode for a responsive transcription factor-promoter system. The promoter must also be oriented in the correct direction to transcribe the downstream reporter gene. The success of this method hinges on the overarching assumption that the requisite transcription factor is housed proximal to its cognate promoter controlling GFP expression[126], a guilt-by-association strategy that is by no means a genotypic rule[127].

User-selected biosensors can also be applied for isolation of anabolic operons from metagenomic libraries encoding for biosynthesis of biosensor-responsive small molecules[128]. This was demonstrated using the well-characterized acyl-homoserine lactone-responsive LuxR- P_{luxI} genetic actuator isolated from *Vibrio fischeri*[129]. By placing GFP under control of P_{luxI}, a metagenomic library was transformed into *E. coli* and screened for novel inhibitors of the quorum-sensing response. Approximately 10,000 colonies survived the preliminary FACS assay by inhibiting GFP production in the presence of exogenously added acyl-homoserine lactone. Only a meager 0.2%, or two colonies, produced confirmable inhibition of transcription from P_{luxI}, a finding that draws light toward the high rate of false positives using FACS for phenotypic characterization. A similar approach was recently used to screen for alkane biosynthesis and transport in engineered strains of *E. coli* using the AlkR transcription factor isolated from *Acinetobacter*[130]. In cases where a target small molecule lacks no known cognate transcription factor, FACS-coupled directed evolution of the transcription factor's effector recognition domain can realize minor perturbations in ligand specificity[131]. Nearly all genetic biosensors described to date have been used to obtain an all-or-none fluorescent response profile, which is well suited for gene discovery but poorly suited for mechanistic studies where intermediate behaviors provide information equally as important as the final product. As discussed in detail below, however, genetic biosensors providing a linear response profile can have also been characterized and can be used to elucidate intermediate behaviors.

1.5. Future directions: transcription factor-based small-molecule screening

The screening strategies described above have demonstrated success in identifying proteins or strains capable of increased small-molecule production titers. Many small molecules, however, are not natively chromophoric, fluorescent, or essential for growth. Furthermore, only a limited few can be readily transformed into compounds possessing these properties. Thus, a universal high-throughput screening platform for small molecules remains out of reach. In spite of that, to expand our existing repertoire of detection techniques, we can turn to nature, the apt experimentalist, and borrow from the plethora of molecular devices evolved for small-molecule recognition.

Nature has evolved RNA, DNA, and protein devices for the binding of small molecules and for activation of a downstream transcriptional response. Some molecular devices, including the aptamer[132] and FRET biosensors[133], can be engineered for small-molecule detection and deserve further exploration. One strategy deserving special attention is the use of transcription factors, proteins that regulate a promoter's transcriptional output in response to a small-molecule ligand, to report on *in vivo* small-molecule production. Bioengineers have long used transcription factors to construct whole-cell biosensors for the detection of environmental small-molecule pollutants[134]. However, this same approach has remained largely untranslated toward library screening and directed evolution purposes.

1.5.1. Biosensor response characterization

Screening methods utilizing transcription factors possess many characteristics ideal for developing small-molecule screens and selections. The reporter gene housed downstream of an inducible promoter can be user selected to encode for colorimetric, fluorescent, or growth-coupled responses. Additionally, because the transcription factor-promoter pair is responsible for direct, *in vivo* detection of the target analyte, the need for downstream synthetic chemistry or *in vitro* manipulation is eliminated.

The first task when designing a screen is to understand the relationship between small-molecule input and reporter protein outputs. This process is greatly facilitated in transcription factor--based detection technologies by continued work in genetic circuit design. In part, this is due to a trend in the synthetic biology community toward better characterization of standardized hosts and biological parts[17,135,136]. For example, a thorough characterization of the frequently used LuxR-P_{LuxR} expression system for *E. coli* provides a solid framework for biosensor design and characterization[37]. The mathematical description of an expression system's transfer function can elucidate the conditions under which a biosensor can provide robust, reproducible function. As with optical detection techniques, biosensor transfer functions are characterized by their dynamic range**,** sensitivity, small-molecule specificity, and range of linear analyte detection (**Figure 1.2**)[38]. All of the relevant parameters can be readily calculated using existing models that describe a range of regulatory schemes[137-139].

Even in *E. coli*, unfortunately, standardized promoter characterization remains scant to date and remains virtually nonexistent for yeast expression systems. We provide here a retrospective analysis of published response curves and provide estimates for a minimum set of governing parameters for many of the most commonly used biosensor-ligand induction systems (**Table 1.1**). Shown are the sensitivity, dynamic range, and linear detection range for a number of

promoter-induction systems when fit using the Hill equation. Few of the small-molecule ligands for which these systems were designed to detect are industrially relevant; thus, their cognate biosensors are unlikely to be used for small-molecule assay design. However, we can still glean useful information on generalized trends that will facilitate biosensor-based assay design.

Induction system	Ligand	Sensitivity (n_{app})	Dynamic Range (fold increase)	Linear Detection range	Reference
$P_{LtetO-1}$-TetR	Anhydrotetracycline	10.43	2535	≈0--4.5 nM	[140]
P_{TETO7}-rtTA[a]	Doxycycline	1.12	28	0.01--2.5 μM	[141]
P_{TETO2}-rtTA[a]	Doxycycline	0.96	15	0.05--11 μM	[141]
P_{TETO1}-rtTA[a]	Doxycycline	0.70	15	0.25--35 μM	[141]
$P_{LlacO-1}$-LacR	Isopropyl β-D-1-thiogalacto-pyranoside (IPTG)	1.14	620	0.01--1 mM	[140]
P_{BAD}-AraC	Arabinose	0.43	5.5	0.08--20 mM	[142]
P_{prpB}-PrpR	Propionate	2.25	6.5	0.2--12.6 mM	[142]
P_{Trc}-lacI	Isopropyl β-D-1-thiogalacto-pyranoside (IPTG)	0.52	5	4--6.3 μM	[142]
P_{LuxR}-LuxR	3-oxohexanoyl-L-homoserine lactone ($3OC_6HSL$)	1.6	500	0.5--10 nM	[37]
P_{pu}-XylR	Toluene	2.3	16	50--300 μM	[143]
P_{pu}-XylR	Benzene	3.2	9.8	NA	[143]
P_{pu}-XylR	4-Xylene	2.0	26	NA	[143]
P_{pu}-XylR	O-xylene	0.79	20.3±1.6	0.05--5 mM	[144]
$P_{po'}$-XylR	O-xylene	[b]	3151±156	0.1--10 mM	[144]
P_{ps}-XylR	O-xylene	1.41	32.5±6.1	0.1--5 mM	[144]

Table 1.1: Transfer function--derived performance parameters for common induction systems. [a]The O# in P_{TETO} indicates the number of operator sites included in the promoter design. [b]Curve could not be accurately fit using the provided data points, however an $n_{app} > 10$ is estimated

From this data there emerge a number of key trends that can help guide more efficient biosensor design. Perhaps the most important performance feature describing a transfer function is the apparent Hill coefficient, n_{app}. By definition n_{aoo} describes the assay sensitivity, but information on the linear range of detection, the detection threshold, and the dynamic range can be inferred. For example, a large n_{app} indicates a high slope within the linear portion of the transfer curve and a sensitive assay, but it is also associated with a small-linear induction window, a large dynamic range, and a low detection threshold. In electric and genetic circuit design, a large n_{app} describes digital-like behavior. The tight regulatory control displayed by digital logic gates has made them a staple for *in vivo* genetic circuit design; the P_{Tet}-TetR and P_{LuxR}-LuxR systems, in particular, are archetypal model systems providing robust, user-defined behaviors[145-148]. Analog-like behavior is also witnessed in our selection of expression systems. For example, the P_{BAD} and P_{PRO} systems surveyed here have low sensitivity and a broad window for which the output response is linearly correlated to the input concentration. In general, analog-type responses are more in line with established screening assays for small-molecule detection. Along a similar vein, analog logic gates are suitable for use in directed evolution assays, where phenotypic improvements are typically small and incremental.

Although broad-reaching conclusions remain elusive, evolutionary arguments support some general trends seen among the transfer functions between various expression systems. Evolution has tailored the regulation of each promoter to function in a specific environmental context. Digital behavior is more often witnessed in genetic systems evolved for antibiotic, quorum sensing, and sugar utilization. In contrast, analog behavior is more typical of nonideal carbon source utilization. Of course, when the system is removed from its original genetic context (i.e., removing a transcription factor from control by its native promoter), these observations may not hold. Transcriptional logic gates have proven to be readily engineered genetic parts, and transfer functions can be altered for more digital- or analog-like profiles, as needed for the application at hand. A shift toward a more digital response was observed by increasing the number of operator sites on a doxycycline-induced P_{Tet} promoter[141], which manifested as an increase in n_{app}, a lower detection threshold, and an increase in dynamic range (**Table 1.1**). Similarly, for transcription factors exhibiting a poor binding profile to the target ligand, directed protein evolution has been demonstrated to increase ligand specificity and biosensor sensitivity[149]. Perhaps more often than realized, response profiles can be quickly adjusted without the need for laborious protein or promoter engineering. Poor ligand-transcription factor binding is often a characteristic of an analog response, suggesting transcription factors with promiscuous ligand binding could serve as ideal analog devices using weak inducers. Lastly, in the case of the transcription factor XylR, just moving between one of its three native promoter systems dramatically changed the response profiles in an o-xylene-induced system[144]. Whether analog or digital, both types of behaviors can be either pilfered from nature's abundant stock or engineered from preexisting, standardized biological parts.

1.5.2. Applications using digital and analog biosensor response profiles

Digital and analog biosensor response modes are both equally important for *in vivo* small-molecule assays; digital behavior has demonstrated a value for construction of genetic traps, whereas analog behavior is well poised for applications in directed evolution. Mohn et al.[150] describe both digital and analog transcription factor--based biosensors in their detection of the biotransformation of linade into 1,2,4-trichlorobenzene. The authors first construct their

biosensor using an evolved mutant variant of the σ^{54}-dependent transcription factor XylR, which is sensitive to 1,2,4-trichlorobenzene but not to lindane. Using this system, three reporter strains were created: two *P. putida* strains harboring luciferase or LacZ reporters and an *E. coli* strain harboring a LacZ fusion protein reporter. To achieve a strong digital response, the biosensor's dynamic range was increased by eliminating sources of background noise caused by leaky transcription from their promoter. In the luciferase reporter strain, the authors couple their evolved XylR protein with the *Po* promoter, which exhibits a substantially lower level of leaky expression as compared to the more commonly used *Pu* promoter[144,151]. In the *E. coli* reporter strain, LacZ was fused to XylU, the first native gene product downstream of the *Pu* promoter. Strains harboring the fused reporter construct were reported to exhibit background expression levels below the detectable limit of a β-galactosidase assay[150]. When cotransformed with a lindane dehydrochlorinase from *Sphingomonas paucimobilis* UT26, the strains provided proof-of-concept demonstration of digital biosensor response using both colorimetric and growth-complementation assays. Although not employed for purposes of directed evolution, the *E. coli* biosensor system yielded a near-linear relationship between output β-galactosidase activity and concentration of exogenously added lindane. Linearity was held over a two orders-of-magnitude concentration range, providing strong evidence this strain could have a use as an *in vivo* screen for improved lindane or 1,2,4-trichlorobenzene biosynthesis.

1.6. Conclusions

Metabolic engineers require high-throughput screens to facilitate the directed evolution of microbial strains for small-molecule overproduction. The sheer number of variables that can be modified within either the genome and heterologous pathway are astounding; furthermore, biosynthetic pathways comprise multiple enzymes subject to regulation, which is too often poorly understood. Without detailed experimental evidence, there is little hope for rational mutagenesis strategies to substantially improve small-molecule yields and pathway efficiencies. When coupled with a high-throughput screen, random mutagenesis can be a powerful method to explore sequence space with little-to-no a priori knowledge of the subject's structure or function. The walk through sequence space, however, is limited by the underlying statistics behind library generation and screening assay throughput. For the vast majority of small-molecule targets, screening throughput stands as the rate-limiting step, with low-throughput liquid and gas chromatography being the modus operandi in the field today. Recent innovations are enabling experimentalists to "climb the screening and selection ladder," the underlying principle being the design of biosensors to transform inconspicuous small molecules into more readily detectable reporters.

2. A Semi-biosynthetic Route for Artemisinin Production Using Engineered Substrate-Promiscuous P450-BM3[2]

2.1 Introduction

Gas and liquid chromatography-mass spectrometry (GC-MS and LC-MS, respectively) are the standard detection methods utilized in the laboratory today. Their role in *de novo* pathway construction and enzyme directed evolution is invaluable. In addition to offering highly sensitive compound detection and quantification, mass spectra can be used to assign putative functions to previously unknown or altered enzyme activities. However, the fundamental disadvantage of both GC- and LC-MS methods is their exceedingly low-throughput. Typically no more than 10^2 samples can be analyzed per day, making these methods untenable for large-scale library screening. Furthermore, for the majority of compounds being targeted for over-production today there exist no readily available, high-throughput screening methods.

In the absence of a high-throughput screening or selection approach, and when faced with a native biosynthetic pathway or constituent enzyme with intransigent expression or catalytic activity, a different strategy is required. One alternative paradigm to using native pathways and their cognate enzymes is *de novo* pathway construction through incorporation of engineered substrate-promiscuous enzymes – enzymes capable of acting upon a broad range of substrates. Substrate and catalytic promiscuities are believed to be hallmark characteristics of primitive enzymes, serving as evolutionary starting points from which greater specificity is acquired following application of selective pressures[152,153,154]. In this respect, substrate promiscuous enzymes are logical starting points for *de novo* metabolic pathway design. This process functionally mimics one method for how nature is believed to evolve novel biosynthetic routes. Following demonstration of a desired enzyme activity, computational, rational, or directed evolution engineering strategies can be used to further tailor a promiscuous enzyme toward greater specificity. An approach that becomes highly attractive if high-throughput screening techniques are available to monitor production of the target molecule or the associated enzymatic activity.

As an example of using substrate-promiscuous enzymes in engineered biosynthetic pathways, we sought to develop a novel semi-biosynthetic route for production of the sesquiterpene lactone endoperoxide artemisinin. Naturally derived from the plant *Artemisia annua*, artemisinin is a highly important antimalarial pharmaceutical. Artemisinin-based combination therapies are currently being recommended by the World Health Organization as the first-line malaria treatment[155]. However, under the production regime in which artemisinin is harvested from natural sources, ACTs are 10–20 times more expensive than existing alternatives[156]. Paucity of supply and high cost currently make ACTs prohibitively expensive in areas to which malaria is endemic. In light of this – and other – difficulties, malaria remains one of the most debilitating and prevalent infectious diseases; between 300 and 500 million people are infected annually, resulting in more than one million deaths[157]. High-level production of a precursor to artemisinin

[2] Reproduced with permission from JA Dietrich, Y Yoshikuni, FK Fisher, FX Woolard, D Ockey, DJ McPhee, NS Renninger, MCY Chang, D Baker, JD Keasling. "A novel semi-biosynthetic route for artemisinin production using engineered substrate-promiscuous $P450_{BM3}$." ACS Chemical Biology **4**(4):261-267 (2009).

in a microbial host could significantly reduce production costs and enable broader distribution of ACTs.

Although there exist multiple routes for both the complete and partial chemical synthesis of artemisinin[158], they are all too low yielding and complicated to enable production of low-cost artemisinin. When examining the native artemisinin biosynthetic route, two distinct challenges emerge: synthesis of the complicated terpene olefin precursor, amorphadiene, and selective oxidation and reduction of amorphadiene to produce dihydroartemisinic acid (**Figure 2.1**). Dihydroartemisinic acid is the immediate precursor to artemisinin and has been shown to be readily transformed[159,160]. The first complication can be overcome by producing amorphadiene microbially. To this end, we have previously engineered both *Escherichia coli* and *Saccharomyces cerevisiae* for the high-level production of amorphadiene[161,162,163,164].

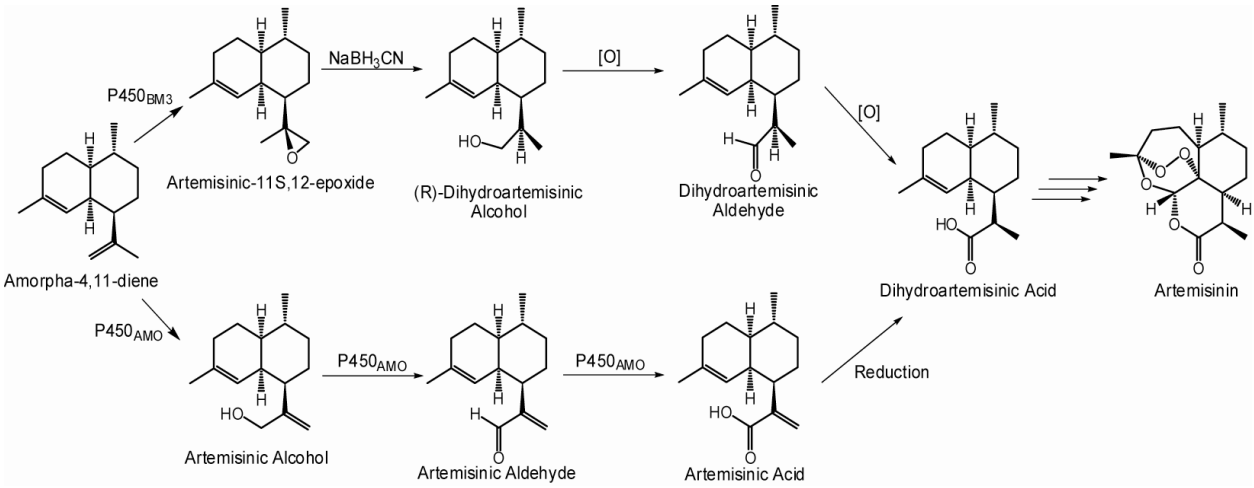


Figure 2.1: Semi-biosynthetic strategies for production of artemisinin. A comparison of the routes catalyzed by $P450_{AMO}$ and $P450_{BM3}$ for the production of dihydroartemisinic acid, from which artemisinin can be readily synthesized. Both routes necessitate additional synthetic chemistry following P450-catalyzed functionalization of amorphadiene.

Selective oxidation and reduction of the undecorated amorphadiene skeleton remains a difficult task. There exist multiple routes for oxidizing amorphadiene: chemical oxidations, biological oxidations using native enzymes, or biological oxidations using non-native enzymes. Oxidation of amorphadiene using chemical catalysts poses numerous problems, including high expense, low product yield, and poor regio- and stereo-selectivity[158,165]. Use of the native *A. annua* cytochrome P450 monooxygenase (*CYP71AV1*, referred to herein as $P450_{AMO}$) and oxidoreductase provides a biological alternative to chemical synthesis. Compared to synthetic chemistry-based oxidations, biocatalyst-based approaches are often able to circumvent problems with regio- and stereo-selectivity pervasive to synthetic chemistry[166,167]. However, eukaryotic P450s have their own share of difficulties and are notoriously problematic to express in heterologous hosts such as *Escherichia coli*. Difficulties with protein-membrane association, folding, post-translational modification, and cofactor integration all hinder successful integration of native enzymes into heterologous pathways expressed in *E. coli*[168]. Although we previously engineered both *E. coli* and *S. cerevisiae* for production of artemisinic acid[164,169], *E. coli* cultures harboring the $P450_{AMO}$-catalyzed biosynthetic route exhibit numerous handicaps. Mixed oxidation products with a relatively low artemisinic acid yield (≈100 mg L^{-1}), 7-day culture

periods at 20°C to obtain functional $P450_{AMO}$ activity, and the need for selective reduction of the terminal olefin to produce dihydroartemisinic acid all inhibit industrial scale-up of this route. In particular, selective reduction of amorphadiene's terminal olefin over the endocylic olefin using synthetic chemistry is difficult to achieve at high yields[170,171].

Here we demonstrate an alternate semi-biosynthetic route, producing dihydroartemisinic acid using an engineered substrate-promiscuous $P450_{BM3}$ in conjunction with high-yielding, selective synthetic chemistry. We selected $P450_{BM3}$ derived from *Bacillus megaterium* for this study. Wild-type $P450_{BM3}$ is known to catalyze the hydroxylation of long chain, saturated fatty acids. $P450_{BM3}$ possesses the highest turnover rate of any known P450, approaching 17,000 min^{-1} in the case of arachidonate[172], and has a demonstrated capacity to be re-engineered[173,174,175].

Minimizing the number of side reactions and products is a difficult task when re-engineering enzymes for non-native substrates, but is a near requirement in metabolic engineering applications. In the case of $P450_{BM3}$, when using amorphadiene as a substrate we believed that the lower transition state energy of C=C epoxidation compared to C-H hydroxylation would favor production of a few, or at best one, product species. Artemisinic-11S,12-epoxide (artemisinic epoxide), the result of terminal olefin oxidation, could serve as an intermediate in our novel semi-biosynthetic route toward artemisinin (**Figure 2.1**). Epoxide cleavage to form dihydroartemisinic alcohol necessitates two oxidations that are not required in the native biosynthetic pathway to form dihydroartemisinic acid; however, this route negates the need for the difficult regio-selective reduction of the terminal olefin of artemisinic acid to form dihydroartemisinic acid, as is required when using $P450_{AMO}$.

2.2 Rational design-based engineering of substrate-promiscuous $P450_{BM3}$

2.2.1 Proof-of-principle demonstration of downstream synthetic chemistry

Technical feasibility of our proposed synthetic chemistry steps was demonstrated by conversion of artemisinic epoxide to dihydroartemisinic acid. First, artemisinic acid was reduced to dihydroartemisinic alcohol (81% yield based on reacted epoxide); subsequently, dihydroartemisinic alcohol was oxidized to dihydroartemisinic aldehyde and dihydroartemisinic acid with yields of 76% and 78%, respectively. Identification of the intermediates along the above route were confirmed by gas chromatography-mass spectrometry (GC-MS) and ^{1}H-NMR spectrometry comparison to both an authentic standard and published results[176,177]. The yields witnessed at the bench scale indicate our proposed synthetic route is competitive with the native $P450_{AMO}$-catalyzed route for artemisinin production.

2.2.2 ROSETTA-based design of $P450_{BM3}$ epoxidase activity

The initial over-expression of wild-type $P450_{BM3}$ in *E. coli* engineered to produce amorphadiene demonstrated no detectable oxidation activity toward the substrate as indicated by GC-MS analysis. Purified wild-type enzyme also did not show any detectable activity during *in vitro* assays using amorphadiene as a substrate. A model structure of the lowest energy transition state complex for amorphadiene in the $P450_{BM3}$ active site was created using ROSETTA-based energy minimization[178]. This model clearly illustrates the steric hindrance imposed on amorphadiene by residue Phe87 in the wild-type enzyme (**Figure 2.2).** This analysis was supported by previous work demonstrating mutations at position Phe87 strongly dictate alternate substrate promiscuity[172,174,175].

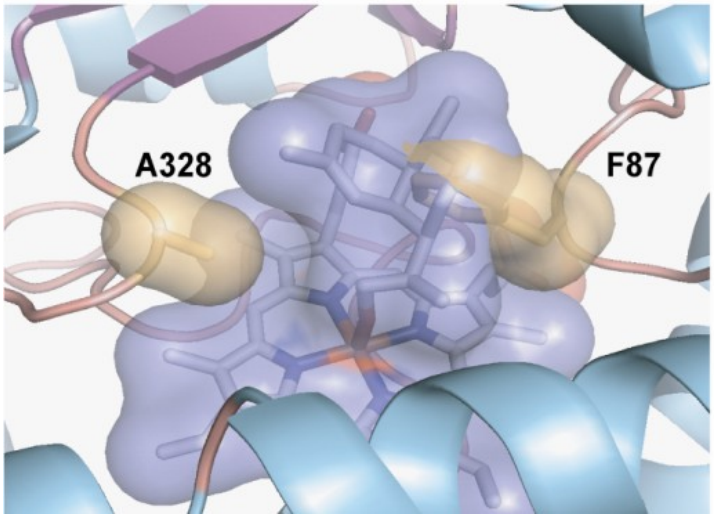


Figure 2.2: **Transition state structure of wild-type $P450_{BM3}$.** Lowest energy transition state complexes were built using the Rosetta algorithm for the wild-type active site of $P450_{BM3}$ with amorphadiene as a substrate. Residues F87 and A328 sterically hinder or fail to hold amorphadiene in close proximity to the heme group, respectively.

Our first goal was to increase the size of the active site binding pocket, which we believed would impart broad substrate promiscuity on $P450_{BM3}$. To this end, we incorporated mutation F87A into the wild-type enzyme (referred to as G1 herein). *E. coli* engineered to produce

amorphadiene and harboring $P450_{BM3}$ variant G1 produced a single compound exhibiting identical retention time and electron-impact (EI) mass spectrum to an artemisinic epoxide standard when analyzed by GC-MS (**Figure 2.3** and **Figure 2.4**)[165].

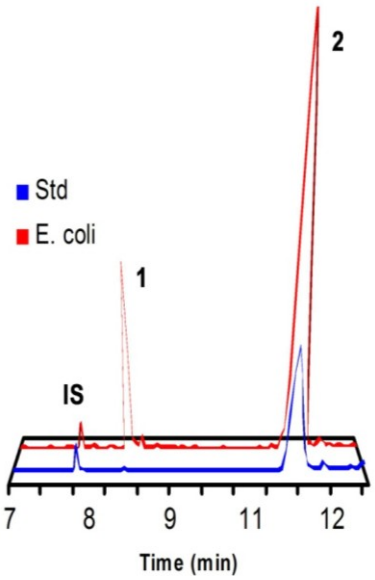


Figure 2.3: GC trace demonstrating *E. coli* produced artemisinic epoxide. Full scan GC-MS trace of authentic artemisinic-11S,12-epoxide (2) standard compared to an *E. coli* DH1 strain harboring plasmids pAM92 and pTrc14-BM3 G4, producing amorpha-4,11-diene (1) and artemisinic-11S,12-epoxide using a caryophyllene internal standard (IS).

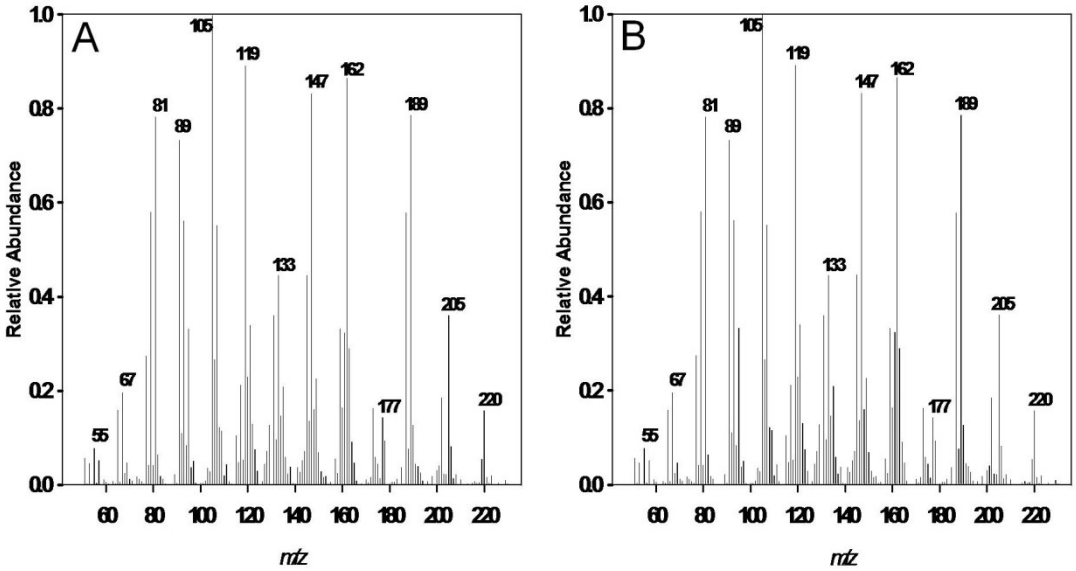


Figure 2.4: EI mass spectra demonstrating *E. coli* produced artemisinic epoxide. EI mass spectra of authentic artemisinic-11S,12-epoxide standard (**A**) and *E. coli* produced artemisinic-11S,12-epoxide (**B**).

While the promiscuity of $P450_{BM3}$ variant G1 enabled production of our desired product, the *in vivo* expression of this enzyme could potentially have the negative effect of oxidizing a wide range of intracellular metabolites. Fatty acids, the native substrates for wild-type $P450_{BM3}$, were identified as the most likely intracellular candidates for oxidation. In an attempt to address this

issue, we created an additional variant containing the mutations G1+R47L/Y51F (referred to herein as G3). Residues Arg47 and Tyr51 are found along the entrance of the substrate access channel. These residues interact with the carboxy group of fatty acids, stabilizing the molecule in the channel during oxidation, and the combination of both mutations has been demonstrated to greatly reduce fatty acid oxidation[179]. We hypothesized the incorporation of $P450_{BM3}$ variant G3 into our biosynthetic pathway would increase product titers due to the elimination of potential sources of competitive inhibition. Production titers for mutation G3 were approximately 0.25-fold increased over mutant variant G1 following 48 hours of culture.

To investigate the possibility of further improving product titers we selected four residues supported by our structured-based models in the $P450_{BM3}$ active site on which saturation mutagenesis was performed. Position Phe87 was selected based upon the previous results demonstrating a dramatic impact upon amorphadiene epoxidation. Positions Ile263, Ala264, and Ala328 were also selected; all three residues lie near the active site and were believed to potentially affect substrate access to the prosthetic heme group. Using $P450_{BM3}$ variant G3 as the starting template, each position was individually mutagenized, the resulting mutants were screened by sequencing for all possible amino acid mutations, and the resultant oxidized amorphadiene metabolite distribution and production of artemisinic epoxide for each mutant was analyzed by GC-MS (**Figure 2.5**). $P450_{BM3}$ variants G3+A328L and G3+A328N were identified to each improve production of artemisinic epoxide by greater than 2.5-fold and 3-fold over that of G3, respectively. Again, no other oxidized amorphadiene products were witnessed.

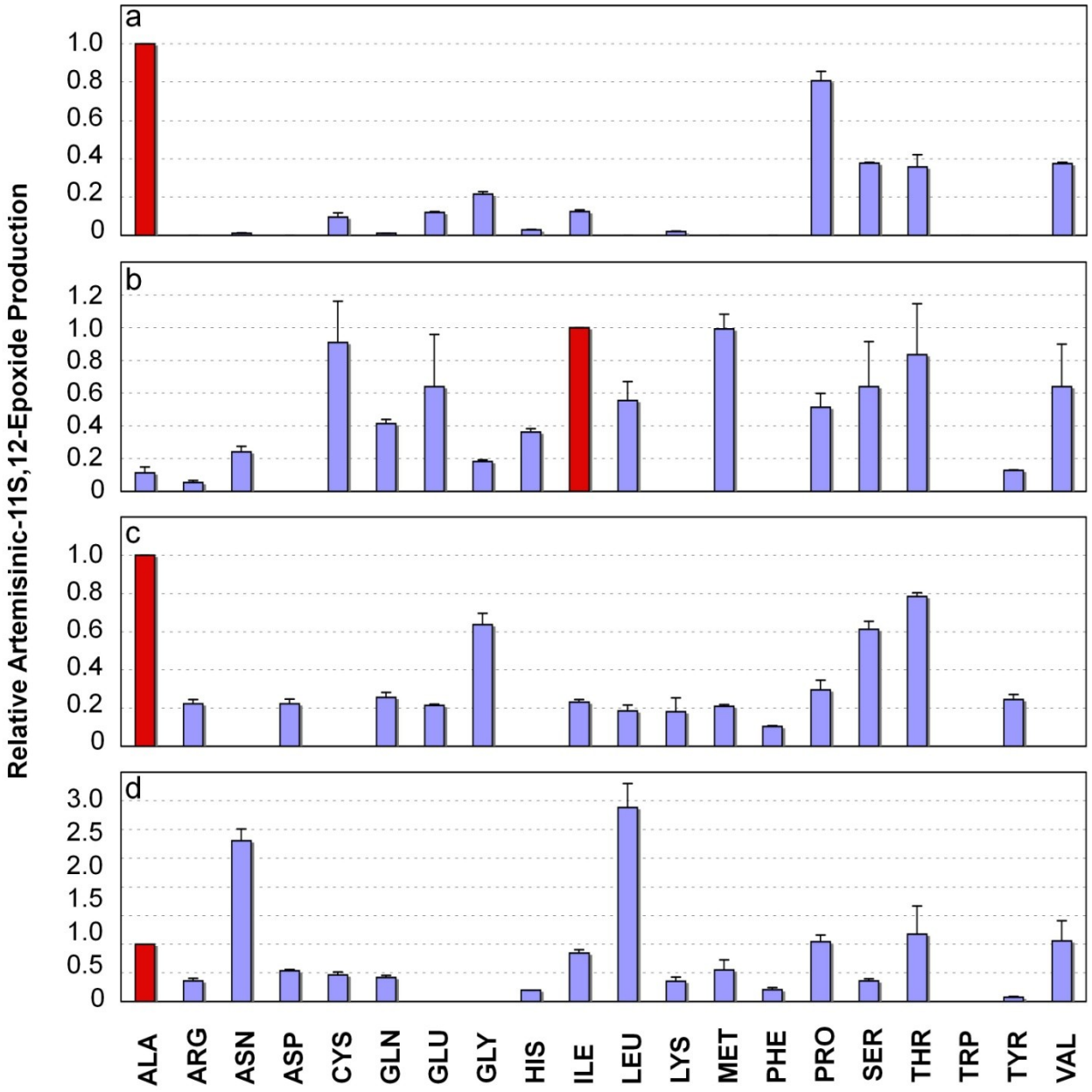


Figure 2.5: $P450_{BM3}$ active site mutation library. Using $P450_{BM3}$ variant G3 as the starting template, residues A87 (**a**), I263 (**b**), A264 (**c**), and A328 (**d**) were targeted for saturation mutagenesis. Artemisinic-11S-12-epoxide production from all possible variants (blue) was quantified relative to G3 (red) by GC-MS analysis (mean ± S.D., $n \geq 3$). Mutant variant G3+A328L (G4) was selected for further study.

2.2.3 Optimization of promoter and induction conditions

Appropriate promoter selection governing expression of P450-BM3 was a critical parameter in achieving high-level production of artemisinic epoxide. The isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible P_{TRC} promoter was being used extensively in the lab for expression of terpene synthases and downstream tailoring enzymes, and was thus also selected for the initial $P450_{BM3}$ expression experiments. It was quickly discovered that inducer concentration and timing both strongly dictated final artemisinic epoxide production titers (**Figure 2.6**). Cultures induced during early log-phase growth (OD_{600}=0.25) and at low IPTG

concentrations (<0.05mM) yielded a nearly 4-fold increase in artemisinic epoxide production titers as compared to those cultures induced in mid-exponential phase or at high IPTG concentrations.

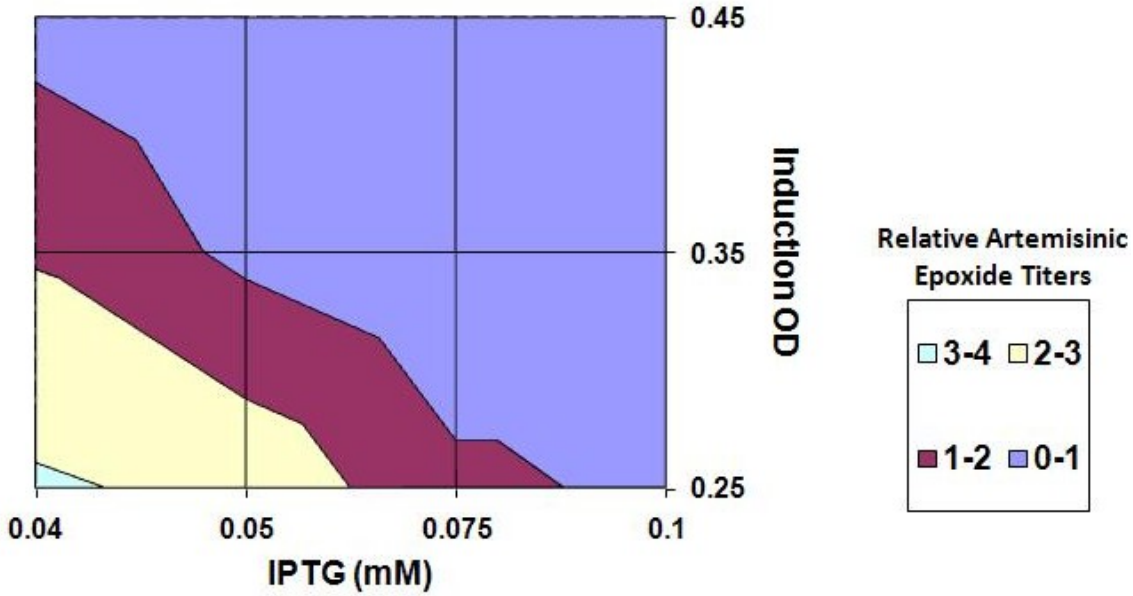


Figure 2.6: Effect of induction strength and timing on artemisinic epoxide titers. Both $P450_{BM3}$ induction time and IPTG concentration had an effect on artemisinic epoxide titers, as measured by the relative artemisinic epoxide GC-MS peak area normalized to a caryophyllene internal standard. The extremely low concentrations of IPTG required to achieve maximal production titers suggest the P_{TRC} promoter is too strong a choice for this system (n=3, coefficient of variation < 15%).

We hypothesized the P_{TRC} promoter was too strong a choice for $P450_{BM3}$ over-expression. Promoter strength was believed to be of particular importance given the optimal IPTG concentration for induction of producing amorphadiene from the pAM92 plasmid co-expressed in our host was 1mM, an order-of-magnitude higher than observed for optimal $P450_{BM3}$ over-expression. Concomitant work in our lab on promoter and inducer optimization of the $P450_{AMO}$ expression system found that the pCWori vector – containing two P_{TAC} promoters, referred to herein as the P_{CWori} promoter – improved both P450 solubility and production titers[180]. We compared artemisinic epoxide titers from *E. coli* expressing $P450_{BM3}$ variant G4 with either a P_{TRC} or P_{CWori} promoter (**Figure 2.7**). At high concentrations IPTG the P_{CWori} promoter was found to significantly ($p<0.05$) improve artemisinic epoxide production titers as compared to the P_{TRC} promoter. Furthermore, the highest production titers were observed using 1mM IPTG induction, the optimal IPTG concentration for amorphadiene biosynthetic pathway overexpression.

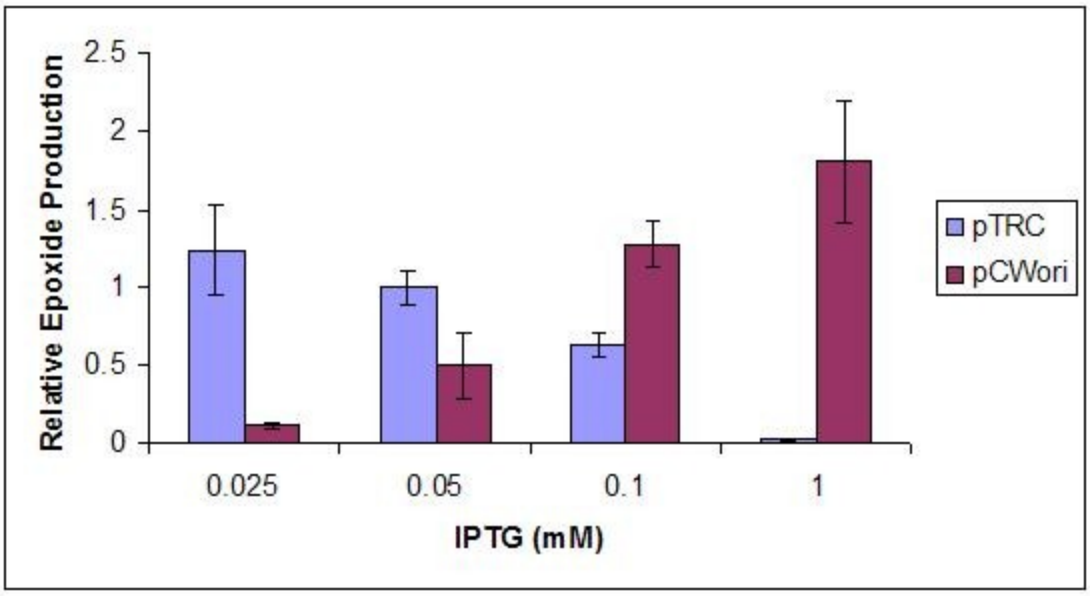


Figure 2.7: Promoter and IPTG inducer optimization. Inducer (IPTG) concentration and artemisinic epoxide production titers were positively correlated when $P450_{BM3}$ variant G4 expression was being driving by the P_{CWori} promoter; in contrast, an inverse relationship was witnessed in the case of the P_{TRC} promoter. The slight increase in final product titer using the P_{CWori} promoter, as well an increase in production assay robustness, prompted the adoption of the P_{CWori} expression system for all final production experiments.

2.2.4 Elimination of byproducts

In addition to improved artemisinic epoxide production, *E. coli* strains harboring the $P450_{BM3}$ variant G4 expressed on the P_{CWori} vector also catalyzed the production of a blue-green pigment when grown in rich (LB or TB) medium; this product was not observed in cultures harboring the $P450_{BM3}$ variants G1 or G3. Following extraction of the cell culture broth with ethyl acetate, the organic solvent acquired a rich, dark blue color. A previous saturation mutagenesis study identified an array of independent mutations in the $P450_{BM3}$ active site that impart hydroxylase activity toward indole[181]. More specifically, the hydroxylation of indole to 3-hydroxyindole is followed by a second, spontaneous oxidation to 3-oxoindole, which is subsequently dimerised to yield the insoluble pigment indigo (**Figure 2.8**). Interestingly, the mutation F87V was essential for indigo production in all active site designs reported in the Literature. F87V is sterically similar to the F87A mutation used in our $P450_{BM3}$ active site re-design, and mutation of phenylalanine to a less bulky amino acid appears to be a key determinant in enabling aromatic substrates greater access to the prosthetic heme group (**Figure 2.2**). Position A328, however, was not explored in the previous study. The finding that numerous, independent mutations are capable of imparting hydroxylase activity toward indole demonstrates the high degree of plasticity inherent to the $P450_{BM3}$ active site.

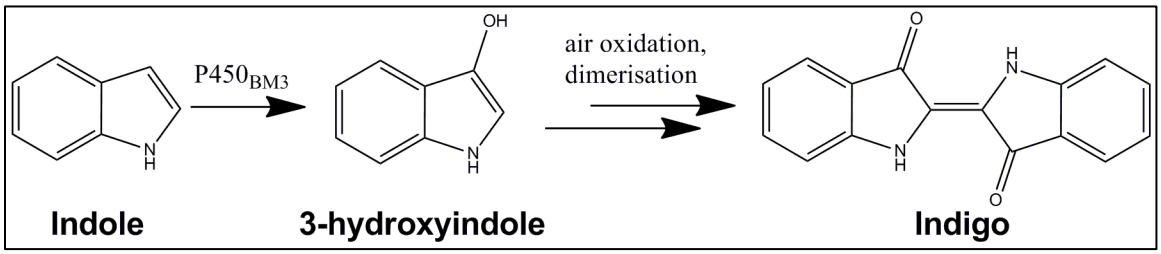


Figure 2.8: $P450_{BM3}$ catalyzed route from indole to indigo. Indigo was observed in the cell cultures following incorporation of mutation A328L into the $P450_{BM3}$ variant G3 active site design. Production of indigo was traced to $P450_{BM3}$-catalyzed hydroxylation of indole, yielding 3-hydroxyindole.

Indigo production diverts a significant fraction of the total $P450_{BM3}$ turnovers, and the host cell's reductive capacity, away from artemisinic epoxide production. Indigo titers were 30±5 mg/L after 48 hours of culture in rich medium (n=3, mean±s.d.), equivalent to 50 mg/L artemisinic epoxide potential based on enzyme turnover. Additionally, NADPH uncoupling (enzyme turnovers that do not result in successful substrate oxidation) using $P450_{BM3}$ to oxidize non-native, aromatic substrates is often greater than 50%. The host cell's total oxidative capacity is therefore diminished by an amount greater than is accounted for based on a molecule-to-molecule comparison of the two oxidized products.

The indigo byproduct could be mitigated in two ways: through $P450_{BM3}$ active site engineering to increase specificity toward amorphadiene, or through host engineering eliminate indole production and remove the source of competitive inhibition. The former option, while perhaps a more elegant solution from an engineering perspective, would be difficult to accomplish in the absence of a high-throughput screening approach. By comparison, eliminating endogenous indole production is straight-forward. Trytophanase, encoded by the gene *tnaA*, catalyzes the hydrolysis of tryptophan to indole, pyruvate, and water[182]; production of the enzyme is stimulated by the presence of tryptophan, which explains why indigo was observed only in experiments conducted using rich, undefined mediums (LB or TB), but not in a minimal medium. *E. coli tnaA* was thus knocked out, and the resulting strain, DH1($TnaA^-$) was tested for its ability to produce indigo in TB medium when harboring $P450_{BM3}$ variant G4. Elimination of indigo production was readily observed by examining the culture broth following 48 hours of growth (**Figure 2.9**), and was also confirmed by GC-MS analysis of ethyl acetate extracted cultures.

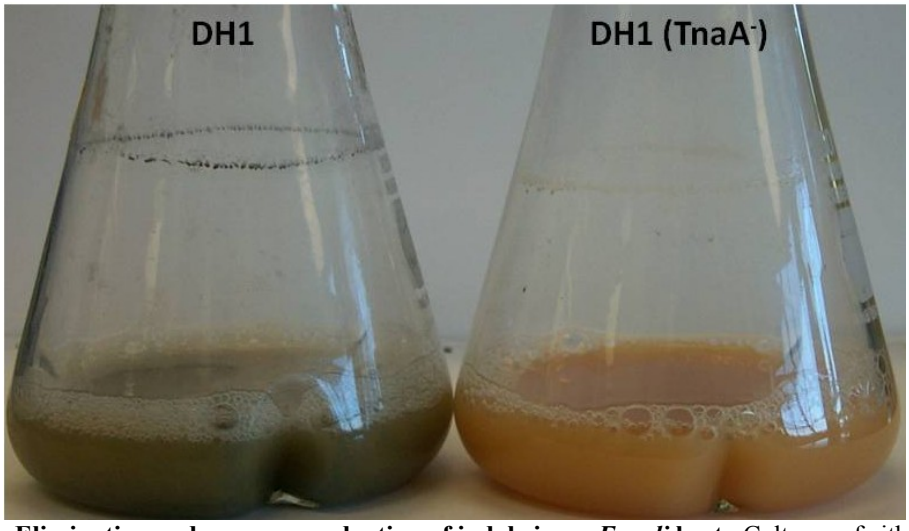


Figure 2.9: Elimination endogenous production of indole in an *E. coli* host. Cultures of either wild-type *E. coli* DH1 or a strain containing a *tnaA* gene deletion, DH1 ($TnaA^-$) were transformed with the $P450_{BM3}$ variant G4 production plasmid and grown in TB (2% glycerol) for 48 hours. The wild-type DH1 strain clearly demonstrated production of the insoluble indigo pigment; the indigo production phenotype was not observed in the TnaA knockout strain.

During more recent attempts to engineer the native E. coli isoprenoid biosynthetic pathway (DXP pathway) for production of taxadiene and taxadiene-5α-ol, intermediates in the taxol biosynthetic pathway, it was discovered that indole concentrations in the culture medium greater than 100 mg/L resulted in decreased carbon flux through the DXP pathway[183]. While the mechanism of inhibition was not elucidated, indole's inhibitory effects were avoided through use of a defined, rich medium. In comparison, our TnaA– strain displayed a statistically significant (t-test, $p<0.05$) decrease in growth rates and artemisinic epoxide production (Figure 2-10). Diminished growth was observed in all P450BM3 mutant variants as well as wild-type P450BM3 and an empty vector control, albeit to a lesser degree in the empty vector control. Thus, the decreased in specific growth rates observed using the TnaA– strain is independent of P450BM3 activity or artemisinic epoxide production. However, synergistic effects with the amorphadiene biosynthetic pathway cannot be ruled out from this experiment.

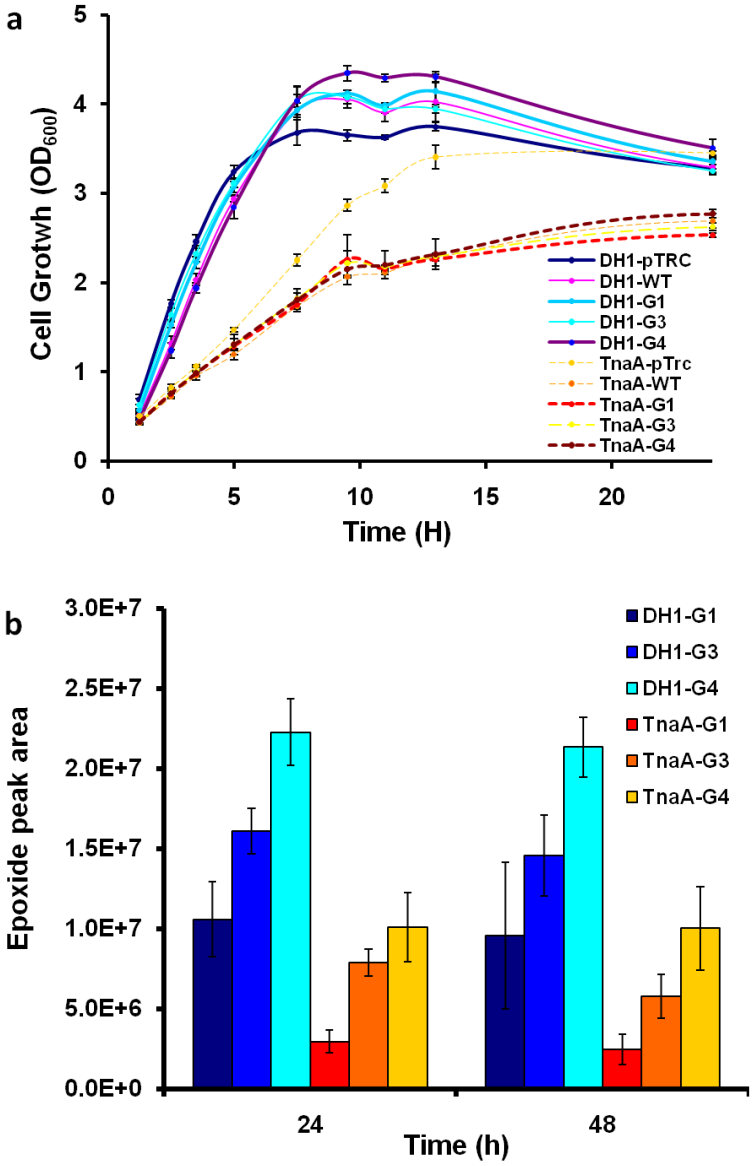


Figure 2.10: Comparison of wild-type *E. coli* DH1 and TnaA⁻ knockout strain. (**a**) All TnaA⁻ samples exhibited decreased specific growth rates as compared to the wild-type DH1 strain. The empty vector control

performed better as compared to those samples expressing wild-type $P450_{BM3}$ and its associated mutant variants in the $TnaA^-$ strain; thus, the additional metabolic load from protein overexpression imposed on the host exacerbates the decreased growth phenotype. (**b**) All $P450_{BM3}$ mutant variants produced significantly (t-test, $p<0.05$) less artemisinic epoxide at both 24 and 48 hours as compared to their corresponding mutations expressed in a wild-type DH1 *E. coli* host (n=3, mean ± s.d.).

In light of both the decrease in specific growth rates and the approximately 50% decrease in total artemisinic epoxide titers we decided not to incorporate the *tnaA* gene deletion in the final production strain. It currently remains unclear to what extent the presence of indole inhibits flux through our heterologously expressed mevalonate pathway. In our artemisinic epoxide production strain deleting *tnaA* had the intended effect of eliminating indole formation in cultures grown in rich medium; however, there was also an unanticipated decrease in cellular fitness as manifested in both the TnaA knockout strain growth curves and artemisinic epoxide production titers. If the decrease in cellular fitness is a characteristic inherent to this knockout strain it may be possible to adapt the strain using a continuous culture and thereby restore a wild-type growth phenotype.

2.2.5 Characterization of optimized production host

$P450_{BM3}$ mutant variants G1, G3 and G4 were cloned into the pCWori expression vector, transformed into wilt type *E. coli* host, and final product titers quantified by comparison to a purified sample of microbially-produced artemisinic epoxide of known concentration. $P450_{BM3}$ variant G1 yielded 110 ± 8 mg L^{-1} of artemisinic epoxide; as expected, variant G3 provided a modest improvement in production over G1, to levels of 140 ± 3 mg L^{-1} (mean ± s.d., n=3). Mutant variant G3+A328L (G4) demonstrated titers of artemisinic epoxide at levels greater than 250 ± 8 mg L^{-1} following 48 hrs of culture (**Figure 2.11**).

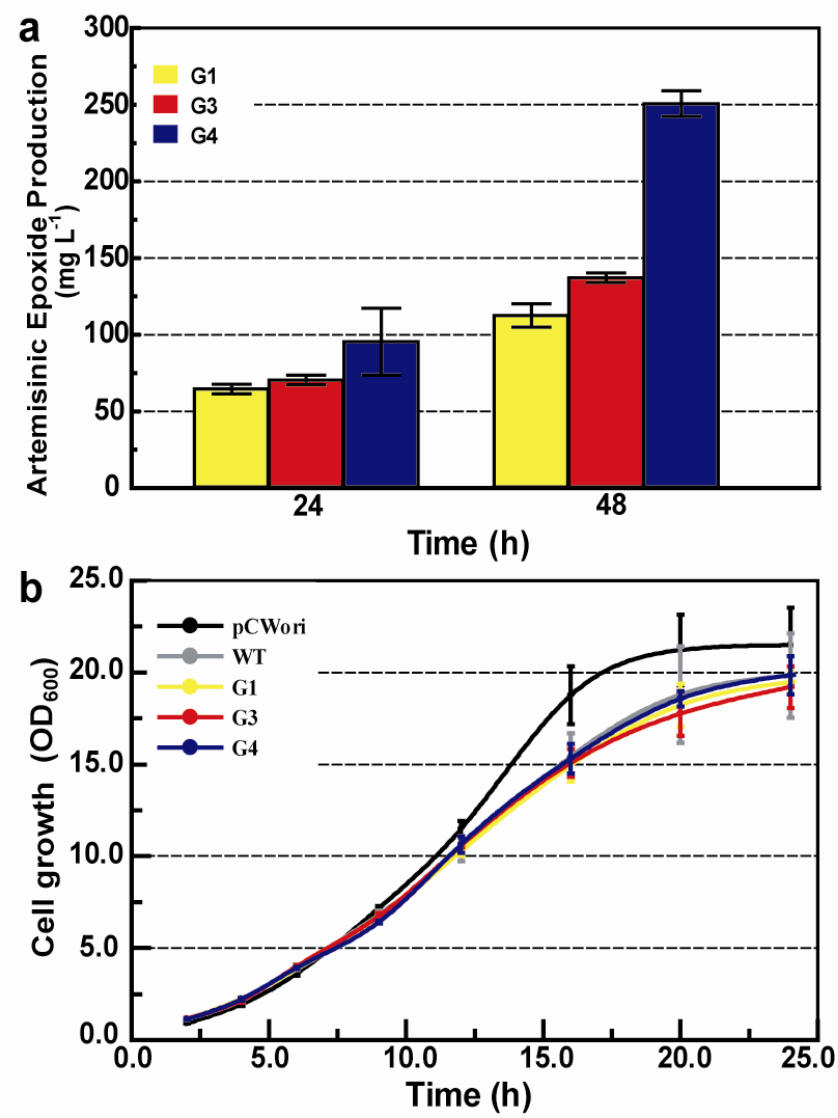


Figure 2.11: Improved *in vivo* production of artemisinic epoxide from $P450_{BM3}$ variants**.** *Escherichia coli* DH1 harboring pAM92 and a $P450_{BM3}$ variant or the empty vector control plasmid pCWori were used for *in vivo* artemisinic epoxide production. Production levels at 24 and 48 hours post inoculation (**a**) and the growth curve (**b**) are shown (all data represent mean ± S.D., n = 3). Strains harboring pCWori or WT $P450_{BM3}$ demonstrated no detectable production of artemisinic epoxide. No adverse physiological effects from $P450_{BM3}$ overexpression or artemisinic epoxide production were observed as demonstrated by the similar growth characteristics between all strains.

A structure-based interpretation of the active site mutations using ROSETTA-based energy minimization enabled a prediction of the amorphadiene epoxidation transition state complex (**Figure 2.12**). Introduction of mutation F87A opened up the active site and relieved the steric hindrance the phenylalanine residue imposed upon the ring structure of amorphadiene. Mutation A328L decreased the size of the $P450_{BM3}$ binding pocket, restricting the mobility of amorphadiene and promoting access of the terminal olefin to the heme group of $P450_{BM3}$ in the correct orientation. Thus, the predicted transition-state structures are in agreement with the observed *in vivo* production titers.

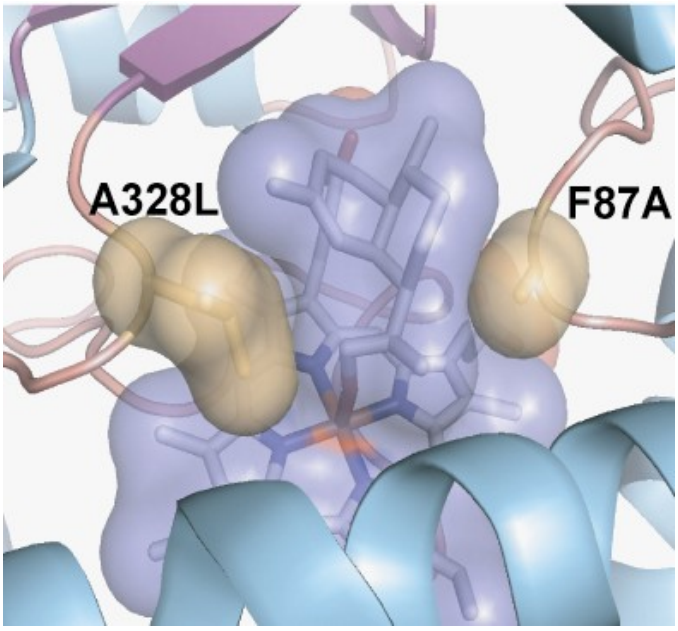


Figure 2.12: **Transition state structures of active site mutations.** Lowest energy transition state complexes were built using the Rosetta algorithm for the mutated active site of $P450_{BM3}$ with amorphadiene as a substrate. Mutation F87A appears to relieve the steric hindrance imposed upon amorphadiene and allows for increased access to the heme group. Mutation A328L appears to decrease the mobility of amorphadiene in the active site, promoting epoxidation.

Of particular importance for *in vivo* expression of an engineered cytochrome P450 is the uncoupling of reduced nicotinamide adenine dinucleotide phosphate (NADPH) consumption to oxidized products (e.g., superoxide radicals and hydrogen peroxide), resulting in increased oxidative stress. Engineered $P450_{BM3}$ variants often exhibit extremely low coupling of NADPH to product formation[174,175], potentially limiting the *in vivo* application of these engineered variants. To determine the extent to which decoupling could hinder *in vivo* performance, purified, wild-type $P450_{BM3}$ and variants G1, G3, and G4 were incubated with amorphadiene, and NADPH consumption was monitored spectrophotometrically. Production of artemisinic epoxide was measured via GC-MS following complete consumption of NADPH, allowing for determination of uncoupling and amorphadiene oxidation rates (**Table 2.1**). While the *in vitro* NADPH consumption rates between mutant variants followed a trend similar to the *in vivo* artemisinic epoxide production measurements, the amorphadiene oxidation rates plateau and incomplete conversion of added amorphadiene was witnessed. This finding suggests NADPH availability may be a limiting factor during *in vivo* production; a hypothesis supported by the observation that significant quantities of amorphadiene is also witnessed at the endpoints of the *in vivo* culture experiments. Increasing the *in vivo* NADPH supply to increase enzymatic activity has been successfully employed in *E. coli* previously by increasing carbon flux through the pentose phosphate pathway[184,185]. A logical next step in improving artemisinic epoxide titers would explore this and other potential routes.

$P450_{BM3}$ variant	NADPH consumption rate	Amorphadiene epoxidation rate	Coupling
WT	N.D.	N.D.	N.D.
G1	22.20 ± 4.10	7.77 ± 1.4	35%
G3	48.23 ± 3.61	30.38 ± 2.27	63%
G4	60.90 ± 3.84	30.45 ± 3.84	50%

Table 2.1: Amorphadiene epoxidation rates of $P450_{BM3}$ variants. NADPH consumption rates and amorphadiene epoxidation rates are given in nmol (nmol P450)$^{-1}$ min^{-1} (mean ± S.D., n=3). Coupling is the ratio of the amorphadiene epoxidation rate to the NADPH consumption rate. No epoxidation of amorphadiene was detected using wild-type $P450_{BM3}$. Observed increases in amorphadiene epoxidation rates between G1 and G3 match well with *in vivo* production data; G4 exhibited a higher NADPH consumption rate than G3, but this did not manifest itself as a higher amorphadiene oxidation rate. This may be explained, in part, by the low coupling efficiency measured *in vitro* for variant G4.

During *in vitro* experiments, all $P450_{BM3}$ variants exhibited NADPH coupling efficiencies of approximately 50%, suggesting that amorphadiene remains a difficult substrate for the studied $P450_{BM3}$ variants to epoxidize. However, as exhibited by the growth curves (**Figure 2.11**), production of artemisinic epoxide did not appear to elicit adverse physiological effects. We believe differences in cell growth between $P450_{BM3}$-producing strains and the empty vector control were likely due to increased cell stress from $P450_{BM3}$ over-production.

The oxidation rate of palmitic acid with $P450_{BM3}$ variant G4 was measured to estimate residual activity of our evolved P450 toward native substrates. We measured the palmitic acid oxidation rate to be 153.5 ± 7 nmol (nmol P450)$^{-1}$ min^{-1} (mean ± S.D., n=3), compared to an oxidation rate of approximately 2,600 nmol (nmol P450)$^{-1}$ min^{-1} with wild-type $P450_{BM3}$[173]. Thus, there is over a 15-fold decrease in activity toward a fatty acid substrate using the evolved mutant variant G4. However, variant G4's activity toward amorphadiene is a fraction (86-fold less) of wild-type $P450_{BM3}$ toward its fatty acid substrates. In light of this, further directed evolution of $P450_{BM3}$ variant G4 is warranted and holds high potential to increase both substrate selectivity and activity toward amorphadiene over fatty acid substrates.

2.3 ROSETTA-based engineering of $P450_{BM3}$ hydroxylase activity

Having demonstrated the capacity of ROSETTA to assist in engineering a promiscuous $P450_{BM3}$, resulting in epoxidase activity, we became interested in re-engineering the enzyme for amorphadiene hydroxylase activity, akin to $P450_{AMO}$, producing artemisinic alcohol (**Figure 2.1**).

2.3.1 Introduction of spacing mutations in $P450_{BM3}$ active site

Beginning with a $P450_{BM3}$ template containing mutations F87A, A328L, and I263G, we planned to first enlarge the P450-BM3 active site through introduction of the spacing mutations L437A and V78A; subsequently, mutations L75F and A82L would be introduced as packing residues to reorient amorphadiene in a position more favorable for hydroxylation of the terminal carbon. All possible combinations of these mutations were tested for activity – as measured by production of an oxidized amorphadiene product – and, in the event that activity was lost, a carbon monoxide difference spectrum assay was used to determine if the heme prosthetic group was intact.

During initial testing, introduction of mutation L437A or mutations V78A/A82L eliminated production of artemisinic epoxide, and no other oxidized products were observed. A Carbon monoxide spectrum, however, depicted a clear shift from 420nm to 450nm upon reduction with sodium dithionite, indicating retention of the heme prosthetic group in the protein active site (**Figure 2.13** and **Figure 2.14**). Thus, these mutants are likely to retain some level of activity toward amorphadiene.

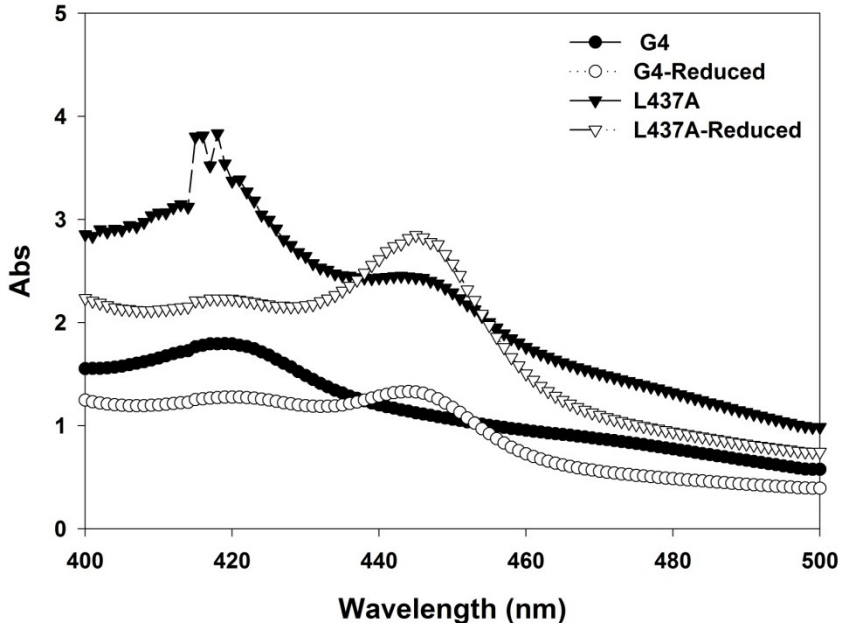


Figure 2.13: $P450_{BM3}$ L437A carbon monoxide difference spectrum. An absorbance shift from 420 to 450nm upon reduction of carbon monoxide bound $P450_{BM3}$ demonstrates retention of the heme prosthetic group, an indication of proper protein folding. $P450_{BM3}$ mutant variant G4 was included as a control.

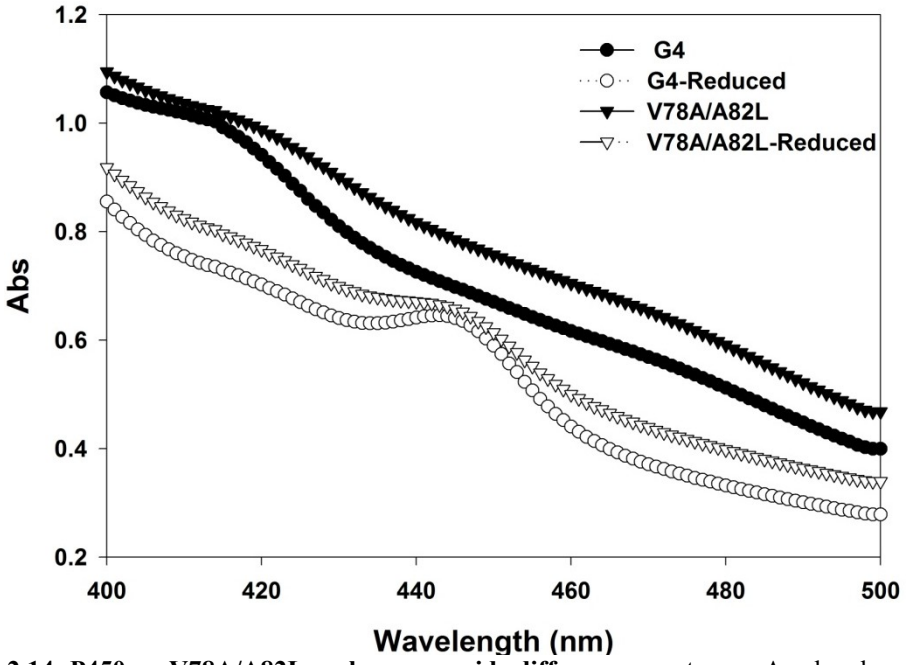


Figure 2.14: $P450_{BM3}$ V78A/A82L carbon monoxide difference spectrum. An absorbance shift from 420 to 450nm upon reduction of carbon monoxide bound $P450_{BM3}$ demonstrates retention of the heme prosthetic group, an indication of proper protein folding. $P450_{BM3}$ mutant variant G4 was included as a control.

Introduction of the spacing mutations, V78A and L437A, were predicted to completely eliminate epoxidase activity. Of these two mutations only L437A, as mentioned above, resulted in a loss of activity. Mutation V78A alone or in conjunction with L437A produced 0.66 ± 0.6 and 0.89 ± 0.17 fold artemisinic epoxide titers, respectively, relative to mutant variant G4 (n=3, mean ± s.d.; **Figure 2-15**).

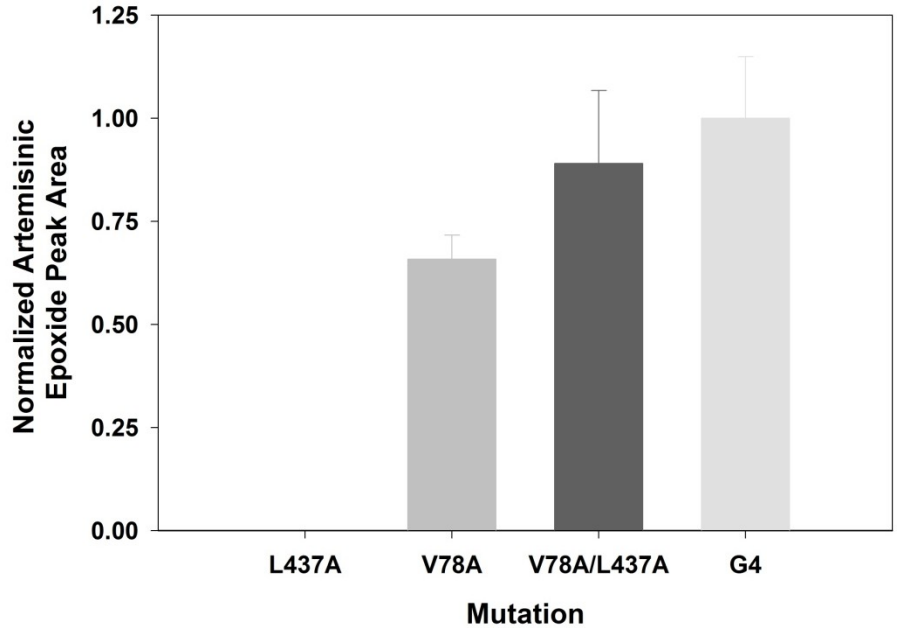


Figure 2.15: **Introduction of spacing mutations into $P450_{BM3}$ active site.** Starting with $P450_{BM3}$ template F87A/A328L/I263G mutations L437A and V78A were introduced into the active site; mutation L437A

resulted in loss of activity toward amorphadiene, and mutations V78A and V78A/L437A resulted in production of artemisinic epoxide, al beit at lower titers than the G4 control strain.

2.3.2 Introduction of packing mutations into P450$_{BM3}$ active site

Despite the model's inability to accurately predict spacing mutations which would result in elimination of epoxidase activity, we moved forward with the addition of the two packing mutations L75F and A82L (**Figure 2.16**). However, all mutants would end up producing artemisinic epoxide as the sole product; none of the mutants yielded detectable levels of artemisinic alcohol or downstream aldehyde or acid products.

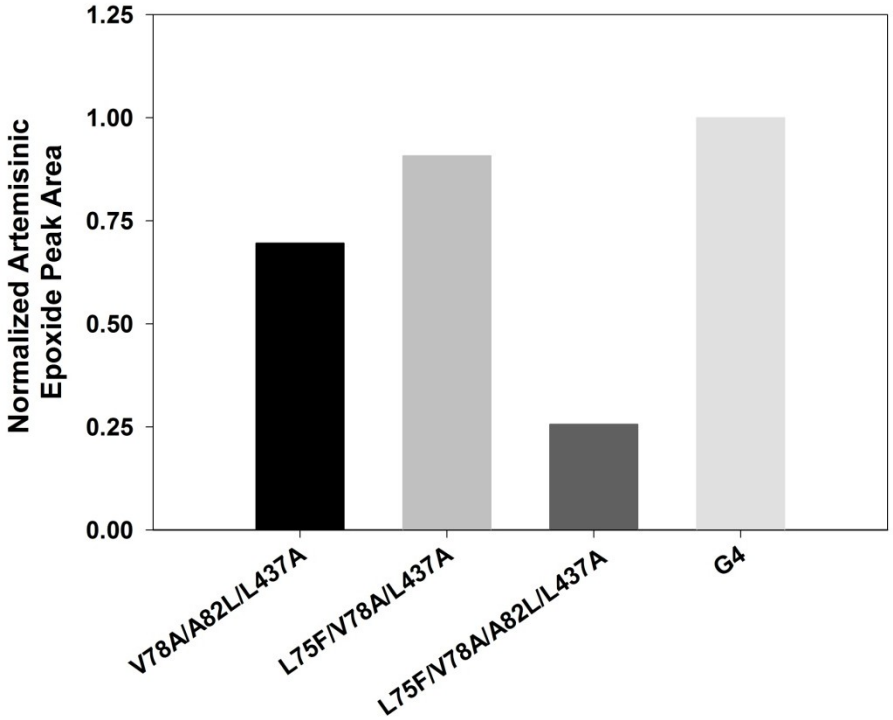


Figure 2.16: Introduction of packing mutations into P450$_{BM3}$ active site for production of artemisinic alcohol. Duplicate measurements of all mutant variants containing both spacking and packing mutations indicated production of artemisinic epoxide; the final quadruple mutant V78A/L437A/L75F/A82L, predicted to possess only hydroxylase activity, produced only the oxidized species as detected by GC-MS.

While there is a wealth of knowledge regarding the directed evolution of P450$_{BM3}$ for non-native enzyme specificity, this work demonstrates the difficulty in achieving a desired regio-selectivity and activity using a rational design-based engineering approach. Designing and successfully implementing a P450-amorphadiene transition state complex that favored hydroxylation over epoxidation was a much more difficult task than originally anticipated. This result, however, should not be completely unsurprising given the lower transition state energy required for epoxidation as compared to hydroxylation.

2.4 Conclusions

Production of artemisinin at low cost is highly dependent upon the capacity to successfully complete what is both relatively expensive and low-yielding synthetic chemistry (i.e. oxygenation and reduction) in a microbial host. Three selective oxidations at the terminal alkane and reduction of the terminal olefin are required for the conversion of amorphadiene to dihydroartemisinic acid; conversion of dihydroartemisinic acid to artemisinin is well established[159,160]. Here, we present a novel route for the selective oxidation of amorphadiene, yielding artemisinic-11S,12-epoxide at titers greater than 250 mg L^{-1}. Microbial production of artemisinic-11S,12-epoxide can be followed by a high-yielding synthetic chemistry route to dihydroartemisinic acid and onward to artemisinin. The strategy outlined here demonstrates improved production titers in *E. coli* compared to those previously achieved using the native $P450_{AMO}$-based pathway. Additionally, the $P450_{BM3}$ route reduces culture time from 7 days to 2, is optimal at 30°C, and does not require reduction of the terminal olefin, as is the case when using $P450_{AMO}$. Thus, the strategy outlined here is a viable alternative to the native $P450_{AMO}$-catalzyed route. This semi-biosynthetic approach outlines a metabolic engineering paradigm for pathway design based on incorporation of enzymes with broad substrate promiscuity that can be well expressed in a heterologous host. The insertion of a minimum number of active site mutations can enable catalysis of a wide range of non-native substrates, making this a highly robust approach in metabolic engineering applications.

There remains, however, an abundant opportunity and need to further improve upon our $P450_{BM3}$ biosynthetic route. Our rationally engineered $P450_{BM3}$ variant G4 exhibits both high NADPH uncoupling and remained active toward non-amorphadiene substrates, including indole. Beyond the $P450_{BM3}$ variant G4 reported here, further ROSETTA-based engineering of $P450_{BM3}$ for either improved amorphadiene epoxidase or hydroxylase activity was met with no success. In general, enzyme engineering for increased substrate promiscuity – particularly in the presence of a substrate-bound crystal structure – is a more straight-forward process than rationally engineering specificity for a single substrate or reaction. Promiscuity is imparted through introduction of mutations that eliminate sources of steric hindrance and allow active site access to a wider range of substrates (and also substrate-protein conformations). In contrast, substrate specificity – and perhaps more importantly, enzyme regio- and stereo-specificity – require introduction of packing mutations that correctly orient the ligand within the active site. Our efforts demonstrate that this remains a highly difficult task, even when provided with an enzyme crystal structure and using well-established computational methods.

Our $P450_{BM3}$ rational engineering efforts highlight the need for high-throughput small-molecule screening or selection techniques. Our efforts were limited by the number of samples that could be screened by gas chromatography-mass spectrometry (GC-MS) to approximately 50 mutants per day. A dedicated GC-MS would be required just to screen samples from one saturation mutagenesis experiment targeting one amino acid residue at a time. As described previously (**Chapter 1: High-throughput metabolic engineering: advances in small-molecule screening and selection**), GC-MS is an intolerable screening method when screening large libraries in which combinatorial mutations are anticipated to lead to the desired phenotype. A potential high-throughput approach would be to monitor the rate of NADPH consumption associated with

$P450_{BM3}$ turnover, similar to a NADPH coupling assay described previously. The requisite mutant $P450_{BM3}$ library could be assembled and cloned into an *E. coli* expression and a crude lysate used to assay for NADPH consumption in the presence of an amorphadiene substrate. The disadvantages of this method, however, are numerous. The screening method is indirect and could provide highly skewed results based on differences amongst samples with regard to protein expression, cell lysis efficiency, mutant reactivity to endogenous *E. coli* metabolites present in the crude lysate, and NADPH uncoupling ratios. At best, perhaps, this method would be an appropriate screening approach to remove the null mutants that either miss-fold or have lost activity toward amorphadiene. Returning to the central tenet of directed evolution, a better screening approach would be based on direct detection of the artemisinic epoxide product.

2.5 Methods

2.5.1 Reagents and equipment

All enzymes and chemicals were purchased from New England Biolabs and Sigma-Aldrich Co., respectively, unless otherwise indicated. Gas chromatography was conducted on a Polaris Q (Thermo Electron Corporation) gas chromatograph, a DB5 capillary column (30 m × 250 μm internal diameter, 0.25 μm film thickness, Agilent), and a TriPlus auto sample-injector (Thermo Electron Corporation). ^{1}H-NMR was collected in $CDCl_3$ (Cambridge Isotope Laboratories) at 25°C on a Bruker AV-500 or AV-400 spectrometer at the University of California, Berkeley, College of Chemistry NMR Facility.

2.5.2 Strains and Plasmids

Genbank files for all plasmids constructed during this study are provided (**Appendix 1**); sequences for DNA primers used in construction of strains are provided below (**Table 2.2**).

Escherichia coli strain DH10b was used for all molecular cloning and saturation mutagenesis screening; strain DH1 was used for *in vivo* artemisinic-11S,12-epoxide production assays and protein over-expression and purification. Strain DH1 (ΔTnaA) is unable to catalyze the conversion of tryptophan to indole, pyruvate, and ammonia; the strain was constructed using the λ-red mediated gene knockout system.[186] Briefly, primers tnaA-F and tnaA-R (**Table 2.2**) were used to amplify a kanamycin resistance cassette flanked by FRT sites from plasmid pKD13. Following confirmation of the *tnaA*::*neo* allelic replacement, the *neo* marker was excised by transformation of plasmid pCP20 and induction of the FLP recombinase. All plasmids exhibit heat-sensitive replication and strains were cured pKD46 and pCP20 after each step.

For amorphadiene production, plasmid pAM92 was obtained from Amyris Biotechnologies Inc. and effectively combines plasmids pMevT, pMBIS, and pADS onto a chloramphenicol resistant plasmid with p15a origin, IPTG inducible lacUV5 and pTRC promoters, and LacIQ[161]. The genes responsible for production of mevalonate from acetyl-CoA (AtoB, HMGS, tHMGR) are housed as a single operon under the control of a lacUV5 promoter. A second operon, also under the control of a lacUV5 promoter, is responsible for the conversion of mevalonate to the isoprenoid precursor farnesyl pyrophosphate (containing MK, PMK, PMD, IDI, IspA). Lastly, the cyclization of farnesyl pyrophosphate to form amorpha-4,11-diene is accomplished by amorphadiene synthase (ADS) under control of a pTRC promoter. Codon-optimized $P450_{BM3}$ was obtained from DNA2.0 and inserted into the NcoI-HindIII sites of pTrc99a or the NdeI-HindIII sites to form pTrcBM3 and pCWoriBM3, respectively.

Rational design mutations F87A (G1), and G1+R47L/Y51F (G3) were constructed via overlap polymerase chain reaction (PCR). Two DNA fragments were created: one encoding the N-terminus of the P450 domain (using primer BM3:NcoI-F) to the C-terminus of the desired mutation, and a second encoding the N-terminus of the desired mutation to the C-terminus of the P450 domain (ending at the SacI site, primer BM3:SacI-R). Both fragments were amplified by PCR: 98°C for 30 sec, 50°C for 45 sec, 72°C for 60 sec, repeated 30 times. The reaction mixture contained 1X Phusion buffer, 0.2 mM dNTP, 0.5 μM forward and reverse primers, 2.5 U Phusion DNA polymerase, and 50 ng pTrcBM3 as a template in a final volume of 100 μl. Amplified DNA was gel purified using a gel purification kit (Quiagen). The two amplified DNA

fragments were spliced via overlap PCR using the same PCR conditions as above. The final amplified $P450_{BM3}$ fragment was digested with NcoI/SacI and cloned into the corresponding site of pTrcBM3.

Reduced expression of $P450_{BM3}$ variants improved artemisinic epoxide production, and an additional six base pairs were introduced between the ribosome binding site (RBS) and the start codon at the NcoI site of pTrcBM3. The RBS region was amplified by PCR: 98°C for 30 sec, 55°C for 30 sec, and 72°C for 30°C sec, repeated 30 times. The reaction mixture contained 1X Phusion buffer, 0.2 mM dNTP, 0.5 μM forward and reverse primers (pTrc99a:RBS_6-F and pTrc99a:RBS_6-R, see **Table 2**, below), 2.5 U Phusion DNA polymerase, and 50 ng pTrcBM3 as a template in a total volume of 100 μl. The amplified fragments were then digested with EcoRV/NcoI and inserted into the corresponding site of pTrcBM3 to form pTrcBM3-14. pTrcBM3 variants were digested with NcoI/HindIII and ligated to form the pTrcBM3-14 variants.

pTrcSBM3-15 was constructed for $P450_{BM3}$ overexpression and S-tag purification, containing an S-tag fused to the N-terminus of $P450_{BM3}$. pTrcSBM3-15 was constructed based on the previously constructed pTrcSHUM15 vector[187]. The pTrcSHUM15 and pTrcBM3-14 (and its mutant variants) were digested with NcoI/HindIII and ligated to form pTrcSBM3-15 and the associated mutant variants.

pCWoriBM3 was constructed using the SLIC method and protocols[188]. Briefly, pCWori empty vector was digested with NdeI and HindIII. Using the PCR protocol described above, $P450_{BM3}$ variants were amplified using the forward (pCWori-1) and reverse (pCWori-1') primers with overhanging regions complementary to the pCWori cut sites. PCR products were gel extracted, treated with T4 DNA polymerase, annealed with the vector, and transformed into chemically competent DH10b.

2.5.3 Site directed mutagenesis by overlap PCR

Site-directed mutagenesis was performed using overlap PCR (see **Table** 2, below). Two DNA fragments were created: one encoding the N-terminus of the P450 domain (using primer BM3:NcoI-F) to the C-terminus of the desired mutation, and a second encoding the N-terminus of the desired mutation to the C-terminus of the P450 domain (ending at the SacI site, primer BM3:SacI-R). Both fragments were amplified by PCR: 98°C for 30 sec, 50°C for 45 sec, 72°C for 60 sec, repeated 30 times. The reaction mixture contained 1X Phusion buffer, 0.2 mM dNTP, 0.5 μM forward and reverse primers, 2.5 U Phusion DNA polymerase, and 50 ng pTrcBM3 as a template in a final volume of 100 μl. Amplified DNA was gel purified using a gel purification kit (Qiagen). The two amplified DNA fragments were spliced via overlap PCR using the same PCR conditions as above. The final amplified $P450_{BM3}$ fragment was digested with NcoI/SacI and cloned into the corresponding site of pTrcBM3-14. For those positions undergoing saturation mutagenesis, 120 colonies from the resulting transformation were screened by DNA sequencing to obtain all 20 possible amino acids at each position.

2.5.4 Synthesis of artemisinic-11.12-epoxide

A mixture of the (R) and (S) configurations of artemisinic-11,12-epoxide was obtained from Amyris Biotechnologies; the methods used for the synthesis of this sample have been reported

previously[165]. Briefly, two diastereomers of the artemisinic-11,12-epoxide were obtained in a 2:1 ratio. ^{1}H-NMR yielded a spectrum in which the signal for the terminal olefin at δ=4.6 and δ=4.9 ppm was absent, and the allylic C6-H remains, indicating epoxidation of the terminal olefin. ^{1}H-NMR ($CDCl_3$) (with minor diastereomer in brackets) δ: 5.17 [5.50] (br 2, 1H), 2.60 [2.40] (d, J=4.5, 1H), 2.83 [2.75] (d, J=4.5, 2H) 2.60 [2.50] (s br, 1H).

2.5.5 Experimental determination of artemisinic-11,12-epoxide stereochemistry

DH1 strains harboring pAM92 and pTrcBM3-14 (G4) used in production assays were extracted into an equal volume of ethyl acetate. After drying, the crude epoxide was purified by silica gel chromatography using 5% ethyl acetate in hexanes as eluent. The mixture was dried *in vacuo* yielding impure epoxide (5.6 mg, ca. 75% pure, 0.019 mmol, contains amorphadiene). The mixture was dissolved in 0.40 mL of tetrahydrofuran and solid sodium cyanoborohydride (27.4 mg, 0.44 mmol) was added, followed by 5 mL of bromocresol green indicator solution. Five drops of 0.15 mL boron trifluoride in 1.0 mL tetrahydrofuran was added, causing the blue color to discharge. After stirring for 112 hours, an additional portion of sodium cyanoborohydride (26.6 mg, 0.423 mmol) was added followed by an additional 5 mL of the indicator followed by 5 drops of the boron trifluoride solution. The mixture was stirred an additional 48 hours and then dried *in vacuo*. The residue was dissolved in a mixture of 1 mL ethyl acetate and 1 mL water. The layers were separated and the organic phase was concentrated. The oil was purified by silica gel chromatography using 10% ethyl acetate as eluant to give 2.7 mg recovered epoxide, along with 2.2 mg dihydroartemisinic alcohol (39% yield, or 81% based on recovered epoxide). Stereochemistry of the purified dihydroartemisinic alcohol was confirmed to be (R) by comparison to published ^{1}H-NMR results[176]. Hyride attack to produce the (R) stereochemistry in the alcohol under these conditions necessitates that artemisinic-11S,12-epoxide be the substrate[189].

2.5.6 Oxidation of dihydroartemisinic alcohol to dihydroartemisinic aldehyde

Using previously described synthetic chemistry[190], (R)-dihydroartemisinic alcohol was oxidized to the aldehyde. To a 50 mL round-bottomed flask equipped with a magnetic stirrer, were added 0.444 g (2.00 mmol) of (R)-dihydroartemisinic alcohol, 1.12 mL (0.808 g, 8.00 mmol) of triethylamine and 10.4 ml of a 5:1 mixture (V/V) CH_2Cl_2 and DMSO. The mixture was stirred and the flask immersed in an ice salt bath at -10 ^oC. Sulfurtrioxide pyridine complex (0.796 g, 5.00 mmol) was then added in three portions over 20 minutes. The ice bath was allowed to come to ambient temperature and the stirring continued for an additional 15 hours at which time GC-MS and TLC (silica gel, EtOAc/hexane) indicated complete conversion to dihydroartemisinic aldehyde. The reaction mixture was poured into 10 mL of 10% aqueous citric acid solution and stirred for 10 minutes. The layers were separated and the organic phase was washed with 10 mL of citric acid solution, 10 mL of saturated $NaHCO_3$ solution, 10 mL of saturated NaCl solution, dried ($MgSO_4$) and the solvent removed under reduced pressure to afford 0.447 g of pale yellow oil. The oil was passed through a plug (5 X 1 cm) of silica gel with 20% EtOAc/hexane 0.377 g (76.4%) of artemisinic aldehyde as determined by GC-MS and ^{1}H-NMR comparison to an authentic standard and previously published results[176].

2.5.7 Oxidation of dihydroartemisinic aldehyde to dihydroartemisinic acid

Using previously described synthetic chemistry[191,192], dihydroartemisinic aldehyde was oxidized to dihydroartemisinic acid. To a 100 mL Ace Glass two-piece reactor equipped with a

mechanical stirrer, were added 0.281 g (1.28 mmol) of dihydroartemisinic aldehyde and 24 mL of DMSO. The mixture was stirred and a solution of 0.172 g (1.90 mmol) of $NaOCl_2$ and 0.966 g (8.30 mmol) of NaH_2PO_4 in 12 mL of H_2O was added at ambient temperature via syringe pump over four hours. After an additional hour GC-MS analysis showed the reaction to be complete. The reaction was diluted with 40 mL of H_2O and acidified to pH 2 with conc. H_3PO_4. Vacuum filtration of the resulting suspension afforded 0.234 g (77.7%) of dihydroartemisinic acid as very fine white needles as determined by GC-MS and ^{1}H-NMR comparison an authentic standard and previously published results[177].

2.5.8 Transition state complex structural predictions

The transition state complex for the epoxidation of amorpha-4,11-diene was constructed based on previously performed energy density calculations on propene hydroxylation and epoxidation[193]. ROSETTA based energy minimization was carried our based on previously described methodologies[194]. The resulting model was visualized using PYMOL[195].

2.5.9 Purification of $P450_{BM3}$ variants

DH1 strains harboring a pTrcSBM3-15 variant (WT, G1, G3, G4) were inoculated into 5 ml TB containing CB[50] and grown overnight at 30°C. 500 ml of fresh TB containing Cb[50] was inoculated using the overnight cultures to an OD_{600}=0.05 and grown at 30°C. Approximately one hour prior to induction cultures were further supplemented with 65 mg l^{-1} ALA. Upon reaching an OD_{600}=0.60 cultures were induced with 0.05 mM IPTG and grown for an additional 15 hours. Cultures were then centrifuged (5000 × g, 4°C, 15 min) and resuspended in 10 ml S-tag™ purification kit (Novagen) wash/bind buffer containing 20 U Dnase I and bacterial protease inhibitor. The suspension was then sonicated (VirTis) on ice, centrifuged (15000 × g, 4°C, 15 min), and the resulting supernatant was passed through a 0.45 μm filter. S-tag purification was used following the recommended protocol with the exception that the thrombin cleavage step was extended to 4 hours. The eluted protein was concentrated using a centrifugal spin filter (Millipore, MWCO 10000). $P450_{BM3}$ concentration was measured by its carbon monoxide difference spectra[196].

2.5.10 In vitro $P450_{BM3}$ characterization

The NADPH turnover rate was determined by incubation of a purified $P450_{BM3}$ variant in the presence of amorphadiene or palmitic acid and NADPH. A 1-ml reaction volume containing 1 μM purified $P450_{BM3}$, 500 μM substrate in 100 mM potassium phosphate buffer (pH=7.5) was equilibrated to 30°C. To initiate the reaction, 250 μM NADPH was added to the solution and the absorbance at 340 nm was monitored. NADPH turnover rates were calculated with a ε_{340}=6.22 mM^{-1} cm^{-1}.

For amorphadiene samples, following complete consumption of NADPH, 900 μl of the reaction volume was taken and added to 500 μl ethyl acetate containing 5 μg ml^{-1} caryophyllene for use as the internal standard in GC-MS analysis. The mixture was vortexed, centrifuged (5000 × g, 25°C, 1 min), and the organic layer was sampled. Samples were then analyzed by GC-MS using the method indicated previously. Coupling of NADPH turnover to epoxidation of amorphadiene was calculated by measuring the decrease in amorphadiene peak area. The apparent initial rate of amorphadiene epoxidation was then obtained by multiplying the coupling efficiency by the NADPH consumption rate.

For palmitic acid samples, 1 ml of the reaction volume was derivatized 50 µl of 2M TMS-diazomethane with 10% methanol. The samples were than analyzed by GC-MS using the previously described method. Coupling of NADPH turnover to hydroxylation of palmitic acid was calculated by measuring the decrease in palmitic acid peak area. The apparent initial rate of palmitic acid hydroxylation was then obtained by multiplying the coupling efficiency by the NADPH consumption rate.

2.5.11 In vivo production, purification, and chemical analysis of artemisinic-11S,12-epoxide

Pre-cultured *Escherichia coli* DH1 transformed with pAM92 and either a pTrcBM3-14 or pCWoriBM3 variant (pTRC for initial screening assays, and pCWori for final production assays) were inoculated into fresh Terrific Broth (TB) supplemented with 2% glycerol (% v/v), 65 mg L^{-1} δ-aminolevulinic acid hydrochloride (ALA), and 50 µg ml^{-1} each of carbenicillin (Cb^{50}) and chloramphenicol (Cm^{50}). All cultures were inoculated at an optical density at a wavelength of 600 nm (OD_{600}) of 0.05. Cultures were induced with IPTG (0.05mM with pTRC and 1 mM with pCWori) upon reaching an OD_{600}=0.25. After 24 and 48 hours of culture at 30°C, 250 µl of culture was extracted with 750 µl ethyl acetate spiked with caryophyllene (15 µg ml^{-1}) as an internal standard. The organic layer was then sampled and analyzed by GC-MS (70 eV, Thermo Electron) equipped with a DB5 capillary column (30 m × 0.25 mm internal diameter, 0.25 µm film thickness, Agilent Technologies). The gas chromatography program used was 100°C for 5 min, then ramping 30°C min^{-1} to 150°C, 5°C min^{-1} to 180°C, and 50°C min^{-1} to 300°C. Identification and quantification of microbially-produced artemisinic-11S,12-epoxide was carried by GC/MS using authentic artemisinic epoxide standards (obtained from Amyris Biotechnologies) of known concentration. ^{1}H-Nuclear magnetic resonance (^{1}H-NMR) spectroscopy was used to confirm the GC-MS identification.

2.5.12 Experimental determination of microbially produced artemisinic-11S,12-epoxide stereochemistry

DH1 strains harboring pAM92 and pTrcBM3-14 (G4) used in production assays were extracted into an equal volume of ethyl acetate. After drying, the crude epoxide was purified by silica gel chromatography using 5% ethyl acetate in hexanes as eluent. The mixture was dried *in vacuo* yielding impure epoxide (5.6 mg, ca. 75% pure, 0.019 mmol, contains amorphadiene). The mixture was dissolved in 0.40 mL of tetrahydrofuran and solid sodium cyanoborohydride (27.4 mg, 0.44 mmol) was added, followed by 5 mL of bromocresol green indicator solution. Five drops of 0.15 mL boron trifluoride in 1.0 mL tetrahydrofuran was added, causing the blue color to discharge. After stirring for 112 hours, an additional portion of sodium cyanoborohydride (26.6 mg, 0.423 mmol) was added followed by an additional 5 mL of the indicator followed by 5 drops of the boron trifluoride solution. The mixture was stirred an additional 48 hours and then dried *in vacuo*. The residue was dissolved in a mixture of 1 mL ethyl acetate and 1 mL water. The layers were separated and the organic phase was concentrated. The oil was purified by silica gel chromatography using 10% ethyl acetate as eluant to give 2.7 mg recovered epoxide, along with 2.2 mg dihydroartemisinic alcohol (39% yield, or 81% based on recovered epoxide). Stereochemistry of the purified dihydroartemisinic alcohol was confirmed to be (R) by comparison to published ^{1}H-NMR results (*25*). Hyride attack to produce the (R) stereochemistry in the alcohol under these conditions necessitates that artemisinic-11S,12-epoxide be the substrate[189].

Table 2.2. Primers used in this study	
Primer Name	**Sequence (5' to 3')**
Cloning primers	
BM3_NcoI-F:	catgccatgacaattaaagaaatgcc
BM3_HindIII-R:	gaagcttttacccagcccacacgtcttttg
BM3_SacI-R:	agaaccatgagctcgtcgccttttttctaaaggatat
pTrc99a:RBS_6-F	gcgcgttggtgcggatatc
pTrc99a:RBS_6-R	attgccatgggcttattctgtttcctgtgtgaaattg
pCWori-1	catcgatgcttaggaggtcatatggcgattaaagaaatgc
pCWori-1'	cgtttgttttcgtcatacgccggatcatccgggttagcgc
Rational design mutations	
F87A-F:	tttgcaggagacgggttaGCTacaagctggacgcatgaa
F87A-R:	ttcatgcgtccagcttgtAGCtaacccgtctcctgcaaa
R47L/Y51F-F:	ttcgaggcgcctggTCTggtaacgcgcTTCttatcaagtcagcgt
R57L/Y51F-R:	acgctgacttgataaGAAgcgcgttaccAGAccaggcgcctcgaa
Saturation mutagenesis	
87F:	tttgcaggagacgggttannsacaagctggacgcatgaa
87R:	ttcatgcgtccagcttgtsnntaacccgtctcctgcaaa
263F:	caaattattacattcttannsgcgggacacgaaacaaca
263R:	tgttgtttcgtgtcccgcsnntaagaatgtaataatttg
264F:	attattacattcttaattnnsggacacgaaacaacaagt
264R:	acttgttgtttcgtgtccsnnaattaagaatgtaataat
328F:	ctgcgcttatggccaactnnscctgcgttttccctatat
328R:	atatagggaaaacaggsnnagttggccataagcgcag
***tnaA* knockout primers**	
tnaA-F	acagggatcactgtaattaaaataaatgaaggattatgtagtgtaggctggagctgcttc
tnaA-R	caccccaaaatgcagagtgctttttttcagcttgatcagtattccggggatccgtcgacc

Chapter 3. Construction a Short-Chain Alcohol Responsive Biosensor in Engineered *E. coli*

3.1 Introduction

Our rational design-based approach to engineering $P450_{BM3}$ for increased substrate promiscuity was highly successful: epoxidase activity toward amorphadiene was observed and an alternative route to artemisinin was demonstrated. However, the limits of our rational design approach also became evident. Amorphadiene was still detected at appreciable levels (>100 mg/L) following 48 hours of culture, indicating that $P450_{BM3}$ activity or expression levels were limiting pathway productivity. Second, the NADPH uncoupling rate remained high (50%) in the final $P450_{BM3}$ design (G4). A high uncoupling rate both wastes cellular resources and leads to generation of damaging free-radical species. Lastly, continued attempts to garner further improvements in $P450_{BM3}$ epoxidase – or detection of novel hydroxylase – activity using ROSETTA were unsuccessful.

In this light, and as outlined previously in Chapter 1, metabolic engineering efforts can benefit greatly from high-throughput screening and selection techniques. Rational design-based approaches are highly useful, and indeed are the norm, during proof-of-principle demonstration of a novel biosynthetic pathway. The number of parameters that can be altered, and the complexity of the interactions between these parameters, however, rapidly becomes prohibitively large. Realizing further pathway improvements requires coupling a directed evolution strategy with a high-throughput screening or selection technique.

Directed evolution, in which a synthetic selective pressure is applied on a diverse pool target sequences to identify a desired trait, is a hallmark of metabolic engineering and biotechnology efforts[197,198]. The success of any directed evolution strategy is contingent upon the effectiveness of two key technologies. First, generating large, diverse genotypic libraries, and second, effectively screening or selecting for the desired phenotype. To date, the capacity to generate genotypic diversity far outstrips our ability to efficiently and effectively interrogate a library. *In vitro* methods for the incorporation of both targeted and random mutations into user-specified DNA sequences are numerous and well-explored[199]. These *in vitro* approaches are complemented by a number of *in vivo*, advanced genome engineering techniques, including multiplex genome engineering (MAGE)[200], global transcription machinery engineering (gTEM)[201,202] and multiscale analysis of library enrichment (SCALEs)[203,204]. These *in vivo* technologies have greatly expanded our capacity to generate diversity at the genome level and investigate the role of individual genes as well as combinatorial effects witnessed when numerous loci are altered simultaneously. However, the above techniques' full potential as applied toward improved microbial production processes remains unrealized in the absence of a high-throughput screening or selection strategy.

Recent efforts to overproduce medium-chain (C4-C6) linear alcohols exemplify this problem. *E. coli* engineered for butanol biosynthesis has been extensively investigated through heterologous expression of the *Clostridium acetobutylicum* pathway[205-209], or through decarboxylation and reduction of 2-ketoacids, the precursors to amino acids [210,211]. While there exist no known native pentanol or hexanol biosynthetic pathways, low level titers (< 1 $g \cdot l^{-1}$) have been demonstrated in

E. coli expressing a promiscuous 2-ketoacid decarboxylase and alcohol dehydrogenase[210]. All engineered hosts reported to date are the product of rational design, and possess suboptimal productivities, titers, and yields (**Table 3.1**). Medium-chain alcohol biosynthesis is not required for *E. coli* growth and is thus not naturally selected for; in fact, alcohol toxicity selects against high-titer alcohol production. There also exists no high-throughput photometric screens for the direct detection of alcohols; indirect assays based on oxidizing an alcohol to its corresponding aldehyde have been used[212-215], but the assays require purified protein, are costly, and are subject to high error rates. In more detail, aldehydes are the immediate precursors to biologically derived alcohols, and aldehyde-based screens are all unable to effectively discriminate between the desired product and the penultimate intermediate. In this light, gas and liquid chromatography have remained the *de facto* screening methods reported in the Literature.

Table 3.1 Performance metrics for reported alcohol biosynthetic pathways

Target Molecule (organism)	Yield (% Theoretical)	Productivity ($g \cdot L^{-1} \cdot hr^{-1}$)	Product Titers ($g \cdot L^{-1}$)	Ref.
Butanol (*E. coli*)	7-12	≈0.015-0.05‡	0.37-1.2	205,206
Butanol (*C. acetobutylicum*)	≈20	0.19‡	13.9	216
Butanol (*C. beijerinckii*)	≈35	0.34-0.50	≈23-27	217
Isobutanol (*E. coli*)	86†	0.18-0.33‡	≈20	218,219
Isopentanol (*E. coli*)	33	0.07-0.12	1.28	220

† Experiments conducted in rich medium (+ yeast extract), yields are overestimated under these conditions
‡ Productivities were not provided, and estimates are based on products titer and total fermentation times

In an effort to develop a high-throughput alcohol screen, we explored the development of an alcohol-responsive transcription factor-based biosensor. As detailed previously, transcription factor-based biosensors combine the exquisite specificity of protein-ligand binding with quantifiable measurement of ligand concentration based on expression of a reporter protein. From a design perspective, a biosensor for detection of an exogenously added small-molecule ligand is similar to reported efforts on construction and characterization of *E. coli* promoter systems[221-224]. From a screening perspective, an alcohol-responsive biosensor can be readily implemented as a liquid- or solid-medium plate screen (**Chapter 1, Figure 1.1**). Briefly, the biosensor strain is cultured either in microtiter or in solid medium plates. In a liquid culture format the spent medium resulting from culture of a production strain (and containing the desired alcohol) is titrated into the biosensor medium (**Figure 3.1**). The biosensor output (i.e. GFP, OD_{600}) is correlated to small-molecule concentration in the liquid culture medium. An alcohol screen exhibits improved throughputs over current GC-MS methods (10^3 to 10^4 samples/day versus 10^2 samples per day, respectively), and is significantly less expensive for large scale library analysis. While not explored here, the biosensor strain could also be included in solid-medium plates and used to analyze alcohol production resulting from individual colonies on the plate surface.

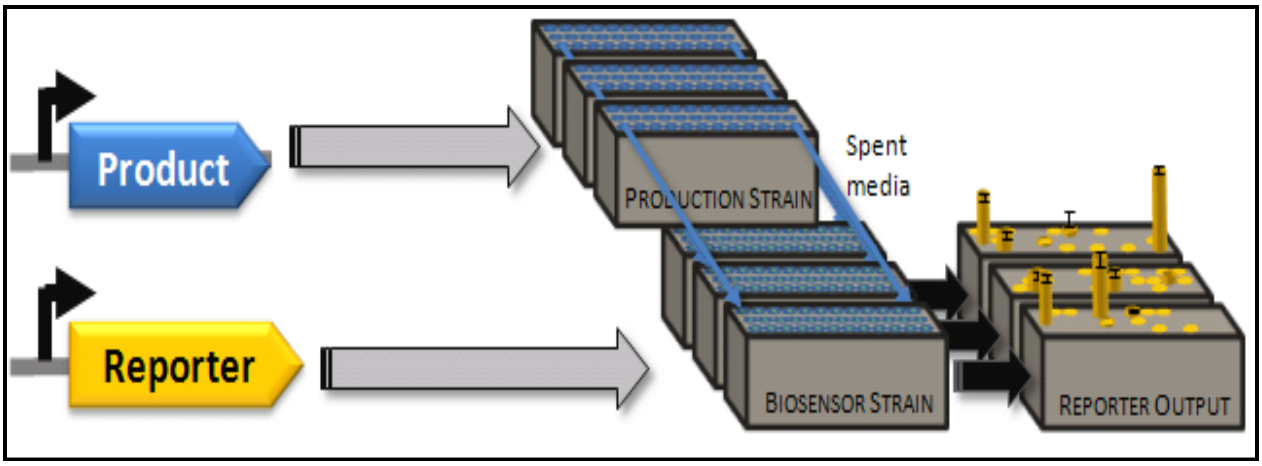


Figure 3.1: Biosensor-based high-throughput liquid culture screens. An engineered host harboring a production cassette is grown in microtiter plates, leading to accumulation of the target analyte in the spent medium. The spent medium is subsequently analyzed using the biosensor strain; the reporter output corresponds to the concentration of analyte in the sample.

3.2 Biosensor design and construction

3.2.1 Transcription factor sourcing

The most straight-forward approach to sourcing an alcohol-responsive transcription factor was to examine those found in *E. coli*. The primary advantage to this strategy is *a priori* knowledge the transcription factor can be functionally expressed in an *E. coli* host, and the biosensor may require little optimization. Under the assumption an alcohol-responsive transcription factor would regulate expression of an alcohol dehydrogenase, we first investigated regulation of the *E. coli* alcohol dehydrogenase genes *adhE* and *yqhD*. YqhD, in particular, has been demonstrated to possess activity toward butanol; however, the high Km value (≈36mM) indicates butanol is not the native substrate[225]. Transcriptional regulation of both genes, however, has been shown to be alcohol-independent, and there are no known alcohol-responsive transcription factors regulating these genes.

We then expanded our search to include related gram-negative microbes, looking specifically towards alcohol (excluding ethanol) catabolic operons. A putative σ^{54}-transcriptional regulator (BmoR) and a σ^{54}-dependent, alcohol-regulated promoter (P_{BMO}) were reported in *Pseudomonas butanovora* (later reclassified as *Thauera butanivorans* sp. nov.[226]) upstream of an n-alkane catabolic operon[227] (**Figure 3.2**).

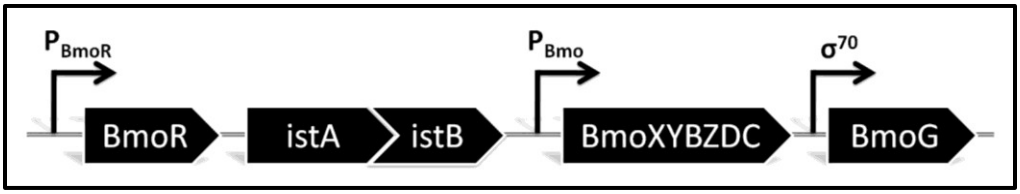


Figure 3.2. **Organization of *P. butanovora* n-alkane catabolic gene cluster**. The *P. butanovora* catabolic operon BmoXYBZDC is responsible for growth on C2-C9 alkanes, with each n-alkane proceeding through a corresponding terminal alcohol intermediate. The σ^{54}-dependent promoter, P_{BMO}, governs transcription of the BmoXYBZDC operon[228]; a σ^{54}-dependent transcription factor, BmoR, is proximally located to the P_{BMO} promoter and catabolic operon[227].

The function of σ^{54} transcription factors in their native bacterial hosts makes these proteins ideally suited for biosensor applications. The activation of σ^{54}-RNA polymerase, and subsequent promoter melting, requires nucleotide hydrolysis by an associated σ^{54} transcription factor[229]; thus, transcription initiation rates are tightly regulated and exhibit low levels of basal expression[230,231]. These features translate well into biosensor design and implementation, serving to decrease background noise and increase biosensor dynamic range over a wide, linear analyte concentration range.

As a family, σ^{54}-transcription factors are activated by a diverse range of small-molecule ligands and protein kinases[232]. Transcription initiation can proceed directly through transcription factor-ligand binding or through phosphorylation by a histidine kinase partner as part of a two-component system. In the later case, the sensor histidine kinase binds an extracellularly localized ligand and subsequently phosphorylates the transcription factor. While 1-butanol readily diffuses across the cell membrane, and thus should not require the presence of a sensor histidine kinase in the signal transduction pathway, a closer analysis was warranted before proceeding with *in vivo* experiments. The domain structure of σ^{54}-transcription factors has been

extensively reviewed[233], and two highly conserved phosphorylation residues have been identified on the N-terminus of σ^{54}-transcription factors that are members of two-component systems[234]. An amino acid sequence alignment of BmoR against protypical two-component system σ^{54}-transcriptional activators showed an absence of the requisite phosphorylation residues (**Figure 3.3**); furthermore, alignment against NCBI's conserved domain database[235] also did not indicate the presence of a phosphorylation motif.

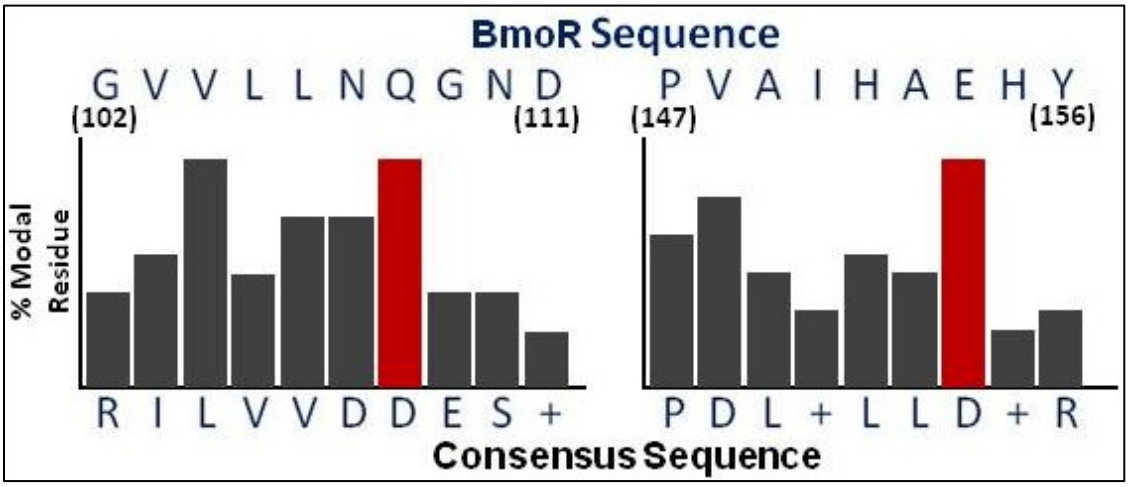


Figure 3.3: Sequence-based analysis of potential BmoR N-terminal phosphorylation residues. The BmoR protein sequence was aligned against a range of two-component response regulators (*E. coli* CheY, Genbank: AAA23577.1; CheB, Genbank: AAA23569.1; NarL, Genbank: AAA24199.1; GlnG, Genbank: AAN83246.1; PhoB, Genbank: ACJ50526.1; AtoC, Genbank: AAA60332.1; ZraR, Genbank: CAP78460.1; and *Pseudomonas aeruginosa* AlgB, Genbank: AAA25700.1). Sequence alignment was poor and highly conserved consensus phosphoacceptor residues (red bars) are absent in BmoR.

3.2.2 Confirmation of BmoR function in E. coli

A two-plasmid system was constructed to obtain preliminary data on biosensor function in *E. coli*. Plasmid pBMO#1 contains the *gfp* gene under transcriptional control of the putative *P. butanovora* butanol-responsive promoter P_{BMO} (device P_{BMO}:*gfp*). Because the location of the BmoR operator site was unknown, 525 base pairs upstream of the σ^{54}-RNA polymerase binding site were included in the promoter design. Plasmid pBMO#6 contains an arabinose-responsive P_{BAD} promoter controlling transcription of BmoR (device P_{BAD}:*bmoR*). The P_{BAD} promoter was selected due to low levels of leaky transcription in the absence of inducer, large dynamic range, and linear expression when expressed in *E. coli* strain BW27783 harboring a constitutively expressed AraE arabinose transporter[236].

Initial experiments using the pBMO#1/pBMO#6 two-plasmid system demonstrated no 1-butanol based induction of GFP fluorescence from the biosensor (Data not shown).

3.3 Biosensor optimization

In an effort to better understand the different biosensor failure modes, we explored several aspects of the two-plasmid biosensor design in greater detail, including temperature and arabinose concentration (for P_{BAD} based induction of BmoR), GFP ribosome binding site and 5-prime untranslated region (5'-UTR), induction timing, P_{BMO} promoter upstream activating sequences, promoter choice for BmoR overexpression, and carbon source.

3.3.1 Temperature and arabinose concentration optimization

A potential explanation for the initial negative results was an inherent inability for heterologously expressed BmoR to function in an *E. coli* host. For example, poor BmoR-σ^{54} RNA polymerase recognition, or poor BmoR-P_{BMO} binding (resulting from the absence of a BmoR operator site in our cloned promoter construct) would both yield little-to-no GFP transcription. There is evidence in the literature[237], however, for functional expression of *Pseudomonas* σ^{54}-dependent transcription factors in *E. coli*; thus, we hypothesized that BmoR was either not expressed, was being localized to inclusion bodies, or the operator site was not included in the P_{BMO} design.

We first over-expressed BmoR from the original arabinose-inducible system and looked for protein expression by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE; **Figure 3.4**). Because the native *P. butanovora* host is a soil bacterium, growing at 25°C, we also examined protein expression at 25°C, 30°C, and 37°C and looked for protein in both the soluble and insoluble protein fractions. None of the arabinose conditions tested using the two-plasmid system produced a detectable BmoR band in the soluble protein fraction as analyzed by SDS-PAGE. Below 0.05 mM arabinose, neither an insoluble nor soluble BmoR band could be detected at all three temperatures tested (data not shown).

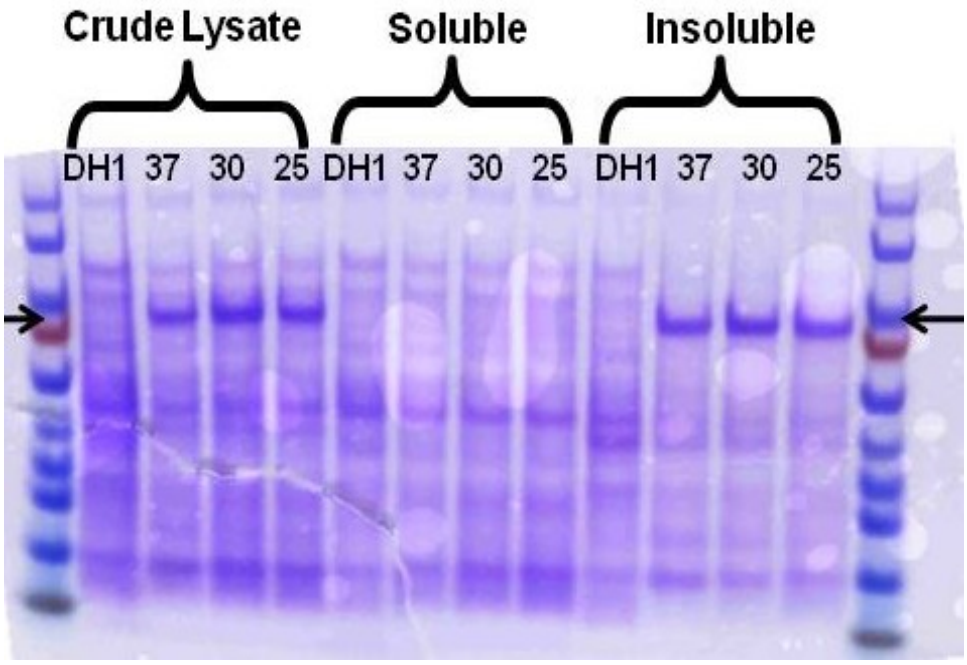


Figure 3.4: SDS-PAGE analysis of BmoR insolubility. BmoR Overexpression from the P_{BAD} promoter (1mM arabinose) resulted in formation of inclusion bodies at all temperatures tested (37°C, 30°C, and 25°C). Arrows indicate location of 72KD BmoR band.

Heterologous protein insolubility is a frequent occurrence in industrial biotechnology, and can often be readily addressed by directed evolution methods[238]; indeed, we previously developed a strategy based on protein multiple sequence alignments to predict amino acid mutations that would impart improved *in vivo* properties, including protein solubility[239]. In the absence of a functional BmoR-P_{BMO} positive control, however, we wished to avoid incorporating mutations into the BmoR protein coding sequence. A more straight-forward approach is to concomitantly decrease both the expression temperature and transcript levels. This strategy proved critical in obtaining high-level expression of both $P450_{BM3}$ and $P450_{AMO}$. As applied to biosensor optimization, we controlled for BmoR transcript level by testing a range of arabinose concentrations (0 to 2.5 mM) at three different culture temperatures (25°C, 30°C and 37°C). 1-butanol concentration was varied (0-100 mM) to provide a measurement of dynamic range and fold-induction after overnight growth in 1-butanol induction medium. Of the conditions tested, optimal biosensor performance was witnessed at low BmoR expression levels (0.05 mM arabinose) and at temperatures less than 37°C (**Figure 3.5**).

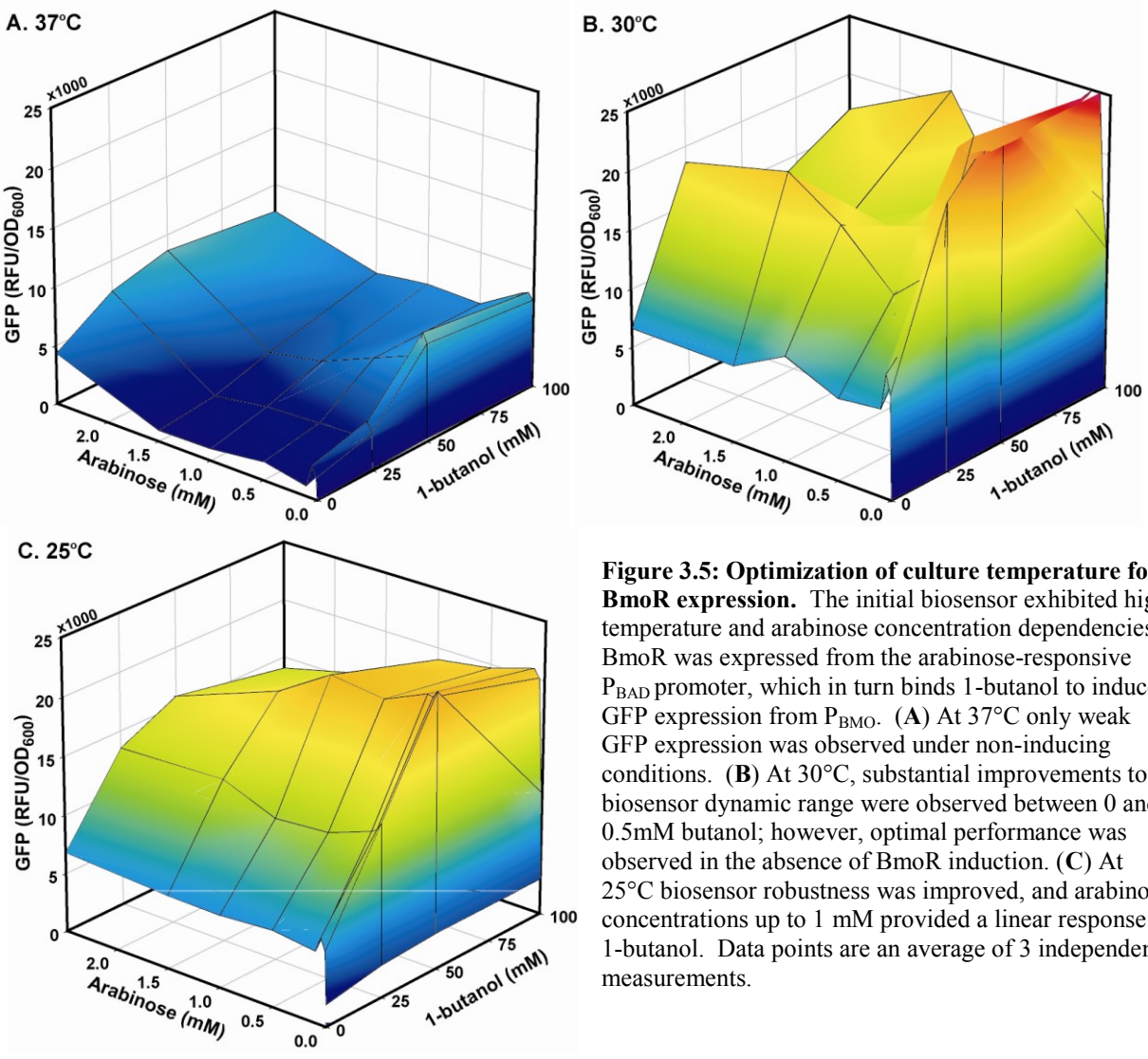


Figure 3.5: Optimization of culture temperature for BmoR expression. The initial biosensor exhibited high temperature and arabinose concentration dependencies. BmoR was expressed from the arabinose-responsive P_{BAD} promoter, which in turn binds 1-butanol to induce GFP expression from P_{BMO}. (**A**) At 37°C only weak GFP expression was observed under non-inducing conditions. (**B**) At 30°C, substantial improvements to biosensor dynamic range were observed between 0 and 0.5mM butanol; however, optimal performance was observed in the absence of BmoR induction. (**C**) At 25°C biosensor robustness was improved, and arabinose concentrations up to 1 mM provided a linear response to 1-butanol. Data points are an average of 3 independent measurements.

Biosensor robustness increased dramatically at lower temperatures. The percent coefficient of variation (%CV) witnessed for samples exhibiting the highest GFP expression (50mM 1-butanol and 0.05mM arabinose) at 37°C was 41%. The %CV was decreased to 4.6% and 2.6% at 30°C and 25°C, respectively, at the same 1-butanol and arabinose concentrations. Thus, the biosensor exhibited substantially improved robustness at lower induction temperatures, a characteristic that serves to minimize the frequency of false hits during screens in addition to reducing variation between multiple, independent assays.

In all cases, the linear range of induction lay between 0 and 40 mM 1-butanol; at concentrations greater than 40 mM, alcohol-induced toxicity resulted in decreased OD_{600} values following overnight growth. At 37°C and with 0.05 mM arabinose, the dynamic range of the biosensor was measured using the minimum and maximum GFP expression values, occurring at 0 and 40 mM 1-butanol, respectively. While a higher average GFP fluorescence was observed at 40 mM, the high degree of variation observed at 37°C resulted in no statistical difference between the two sample groups ($p>0.05$, t-test; $n=3$); thus, for practical purposes, the biosensor dynamic range at 37°C is zero. At both 30°C and 25°C the local minima and maxima were significantly different ($p<0.05$, t-test; $n=3$), and yielded dynamic ranges of 24,100 RFU/OD_{600} and 15,300 RFU/OD_{600}, respectively. These dynamic ranges correspond to a 4.3-fold induction at 30°C, and a 3.1-fold induction at 25°C. While a vast improvement over the preliminary biosensor experiments, the fold-induction values remained low; for comparison, a 5000-fold induction was reported for a GFP reporter under the control of the $P_{LTETO-1}$ promoter system commonly used in synthetic biology applications[240].

3.3.2 GFP ribosome binding site and 5'-untranslated region optimization

Having demonstrated a functional BmoR-P_{BMO} biosensor in *E. coli* we next worked on improving biosensor dynamic range. Both ribosome binding site (RBS) sequence, and more generally the 5'-untranslated region (5'UTR), strongly dictate protein levels. At the primary sequence level, the importance of a standard ATG initiation codon[241], along with both canonical Shine-Dalgarno sequence and spacing relative to the initiation codon[242,243], are well known. The RNA secondary structure of the Shine-Dalgarno region has been shown to be as important as the primary sequence in dictating translation initiation rates[244-247]. The P_{BMO}:*gfp* construct is predicted to exhibit high GFP expression based on the primary sequence; however, the presence of a strong hairpin loop (**Appendix 1, Figure A1.1**) locking up the RBS-initiation codon spacer sequence may lead to decreased GFP translation initiation rates.

To address this issue we employed the RBS calculator[244] to design two synthetic RBS sequences with strong predicted translation initiation rates (50,000 and 100,000 relative units, FLU). Both designs eliminated the presence of the hairpin occluding ribosome access to the Shine-Dalgarno and surrounding sequence space (**Appendix 1, Figure A1.2 and A1.3**). The redesigned reporter constructs were tested over the predicted BmoR 1-butanol linear response range (0-40 mM) using the previously optimized culture and induction conditions (**Figure 3.6**). In this experiment, the wild type (WT) RBS exhibited less than 2-fold induction over the 1-butanol concentrations tested. By comparison, the GFP expression levels measured for both the 50K FLU and 100K FLU constructs were significantly higher than the WT sequence at all 1-butanol

concentrations tested (n=3; t-test, $p<0.05$); the fold-induction increased to 2.73 and 4.26, respectively (**Table 3.2**).

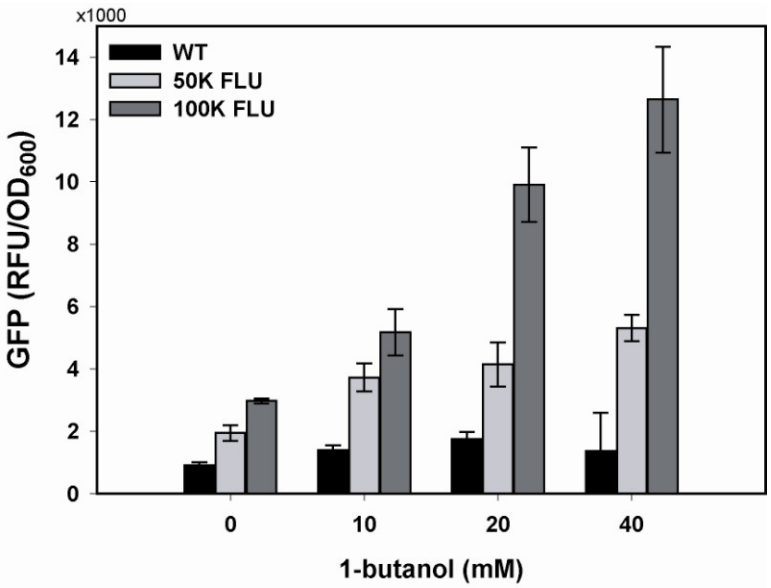


Figure 3.6: GFP ribosome binding site optimization. Increasing the biosensor dynamic range would facilitate identification of positive hits from background GFP expression. The wild type RBS construct (WT) was compared against two synthetic RBS sequences designed for high level translation initiation rates in *E. coli*[244]. Synthetic RBS sequences were designed to yield 50,000 and 100,000 relative fluorescence units (50K FLU and 100K FLU, respectively). n=3; mean±s.d.

Table 3.2: Performance features for ribosome binding site test constructs

Construct	Fold Induction	Dynamic Range
WT	1.52	467
50K FLU	2.73	3367
100K FLU	4.26	9668

n=3; mean±s.d.

Interestingly, while an improvement in dynamic range was expected, the change in fold-induction was not anticipated. Fold-induction should be static for constructs with the same biosensor architecture, and neither a change in translation initiation rate nor RNA stability should affect this performance feature. Non-linear behavior may be observed in the system as the relative GFP expression levels are increased.

3.3.3 Timing of BmoR induction optimization

From our experience engineering biosynthetic pathways there is an optimal induction time – as measured by OD_{600} starting from a 1% inoculation – that produces the highest titer. With this in mind, we also assumed the same would be true for expression of BmoR. A range of induction OD_{600} values covering cell growth from lag through late-exponential phases were tested; 0.05 mM arabinose and 1-butanol were concomitantly added to the growth medium, and GFP expression was measured after overnight growth (**Figure 3.7**).

GFP expression levels following induction with either 1 mM or 40 mM exogenously added 1-butanol were significantly higher than the negative control samples at all time points tested (n=3, t-test, $p<0.05$). Induction in early exponential phase ($OD_{600}\approx0.25$) yielded the highest dynamic range and the lowest percent coefficient of variation (4.5%). Subsequent experiments followed these results as a guideline.

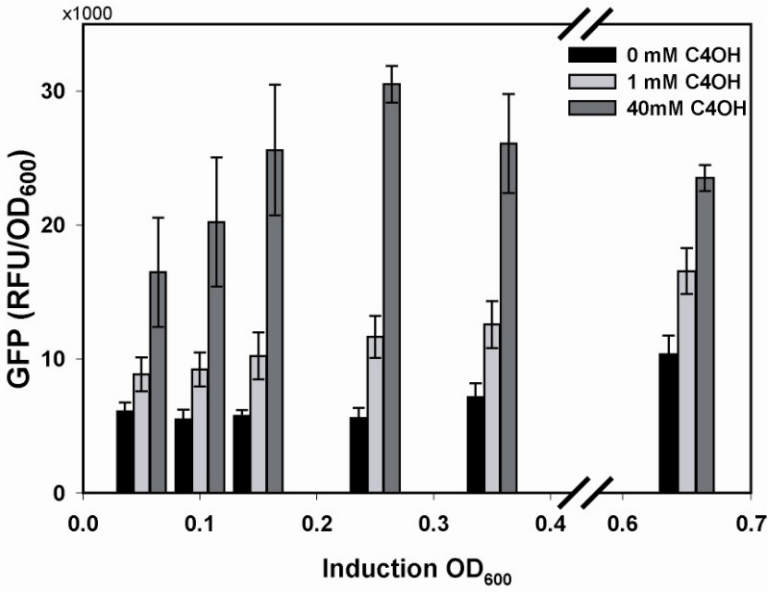


Figure 3.7: Optimization of BmoR and GFP induction timing. The 2-plasmid biosensor construct was induced at various time-points ranging from lag phase (OD_{600}=0.05), through early exponential phase (OD_{600}=0.25), and into late exponential phase (OD_{600}=0.65). Both low (1mM) and high (40mM) concentrations 1-butanol (C4OH) were tested in addition to the negative control (no 1-butanol added). n=3; mean±s.d.

3.3.5 Carbon source and BmoR promoter optimization

The initial characterization of the *P. butanovora* P_{BMO} promoter system reported both alcohol- and carbon source-dependent induction of downstream genes[227]. Furthermore, over a 300-fold induction ratio was witnessed in the native system following induction with 1-butanol. By comparison, our two-plasmid biosensor system in *E. coli* exhibited a less than 10-fold induction ratio.

One of the original motivations for building the *E. coli* biosensor with a *Pseudomonas* transcription factor-promoter pair was the oft relied upon assumption that a heterologous system will be orthogonal to the native *E. coli* regulatory network. Refactoring[248] helps address this problem during biosynthetic pathway construction. Removing native promoters, incorporating synthetic RBS sites, and removing intragenic regulatory elements by scrambling protein-coding DNA sequences (while maintaining the amino acid primary sequence) all help in the construction of a user-controlled, orthogonal system. This task is made more difficult when applied to transcription factor-promoter systems because the desired, internal regulation must be preserved. Using P_{BAD} to drive BmoR eliminated native regulation over transcription factor expression; however, absent further characterization of P_{BMO}, we shied away from making significant modification to the primary sequence in an attempt to eliminate potential internal regulation.

In an effort to better understand, and potentially eliminate, P_{BMO} internal regulatory elements, a computational model of *E. coli* transcription factor DNA-binding motifs was used to identify putative regulatory sequences within the P_{BMO} sequence. In light of reported catabolite repression of the P_{BMO}-governed BmoXYBZDC operon in a *P. butanovora* host[227], we initially focused on identification of potential cyclic AMP receptor protein (CRP) binding sites. *Pseudomonas putida* contains a CRP homologous to *E. coli* CRP, and has also been shown to involve in catabolite repression[249]. The consensus CRP binding site sequences between *E. coli* and *P. putida* are similar (**Figure 3.8**), suggesting that a *P. butanovora* CRP binding site within the native P_{BMO} sequence would cross-talk with the *E. coli* CRP regulatory network. Two hits for potential CRP binding sites were identified approximately 91 bp and 329 bp upstream of the σ^{54}-RNA polymerase subunit recognition sequence. Comparison of the putative CRP operator sites on P_{BMO} to the consensus sequences from both *E. coli* and *P. putida* revealed weak alignment with one of the two core-regions of the CRP consensus sequences for both hosts (**Figure 3.8**). The spacer region between the core binding sites in the putative CRP hit at -329 basepairs is extended, decreasing the probability it acts as a functional CRP operator site.

Consensus *E. coli* CRP	**5'-**	N	T	G	T	G	A	N	N	N	N	N	N	T	C	A	C	A	N												
-329 PBMO	**5'-**	C	A	G	T	G	A	C	A	C	C	G	G	T	C	G	C	C	T												
-91 PBMO	**5'-**	A	G	A	T	G	T	A	C	C	C	T	T	C	T	T	T	A	A	C	G	T	G	T	A	A	C	A	C	A	C
Consensus *P. putida* CRP	**5'-**	W	T	G	N	G	A	W	N	Y	A	W	W	T	C	A	C	A	T												

Figure 3.8: Putative P_{BMO} CRP binding sites. Using TRANSFAC[250], two putative CRP binding sites were identified in the P_{BMO} sequence. The core CRP sites (yellow) possess weak alignment with the consensus *E. coli* and *P. putida* CRP recognition motifs; furthermore, the putative CRP site found at -329 is not spaced properly. W=A,T; Y=T,C.

To further investigate the role catabolite repression plays in the P_{BAD}:*bmoR* and P_{BMO}:*gfp* two-plasmid system a series of experiments were conducted with different carbon sources (**Figure 3.9**). All experiments were conducted in Luria-Bertani (LB) broth alone or supplemented with glucose or propionate as carbon sources; propionate was selected because it elicited high background expression in the *P. butanovora*[227]. Catabolite repression was observed under high glucose conditions (0.2% and 0.4% w/v), but was not present at a low concentration glucose (0.1% w/v) or when LB medium was supplemented with propionate.

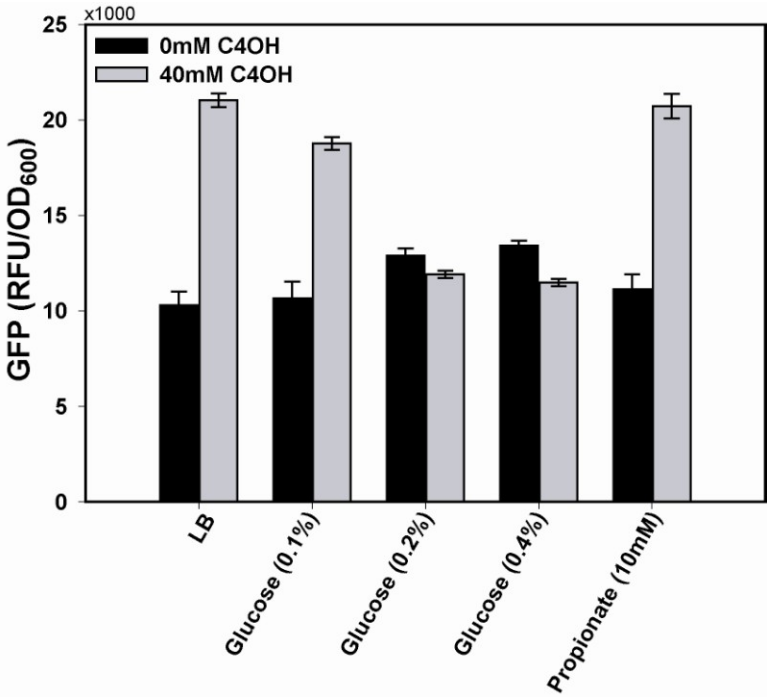


Figure 3.9: Carbon source dependence on BmoR-P_{BMO} biosensor response. GFP expression levels were measured using 0mM (leaky expression) and 40mM 1-butanol (C4OH) and a range of carbon sources. Butanol dependent GFP expression was observed using LB without additional carbon supplementation or with propionate supplemented medium. Butanol dependent GFP expression was observed at a low glucose concentration (0.1% w/v), but was eliminated at high glucose concentration (0.2-0.4% w/v). n=3, mean±s.d.

While these results make clear the biosensor is non-functional at high glucose concentrations, because our initial biosensor design incorporated a P_{BAD} promoter – which is subject to glucose repression[251] – we were unable to determine the exact mechanism of repression. To better address this finding we re-designed the two-plasmid biosensor system to replace P_{BAD} with the anhydrotetracycline-responsive P_{ZT} promoter (plasmid pBMO#7). Glycerol, which elicits a different catabolite repression profile from glucose in *E. coli*[252,253], was also included in the experimental design.

The P_{ZT}- and P_{BAD}-based biosensors exhibited divergent behavior with respect to the effect of carbon source on GFP output (**Figure 3.10**). Using medium supplemented with 0.1% (w/v) glucose, the P_{BAD}-based biosensor exhibited a linear response to 1-butanol while the P_{ZT}-based biosensor was non-functional. The inverse behavior was observed when the growth medium was supplemented with 0.2% (v/v) glycerol. This finding was both surprising and difficult to rectify with our previous results. Transcription from the P_{ZT} promoter is reportedly unaffected by either glucose or glycerol supplementation[254], a finding that stands in contrast to our observed biosensor behavior (albeit in a more complex system).

From our work optimizing the culture temperature and arabinose concentrations (**Figure 3.5**) it was known that high BmoR expression levels led to formation of inclusion bodies and loss of biosensor function. Low-level glucose supplementation may lead to weak repression of P_{BAD}, and having the unintended effect of avoiding high-level production of insoluble BmoR. Glycerol

supplementation, in contrast, would have the opposite effect: relieving P_{BAD} repression and leading to formation of inclusion bodies.

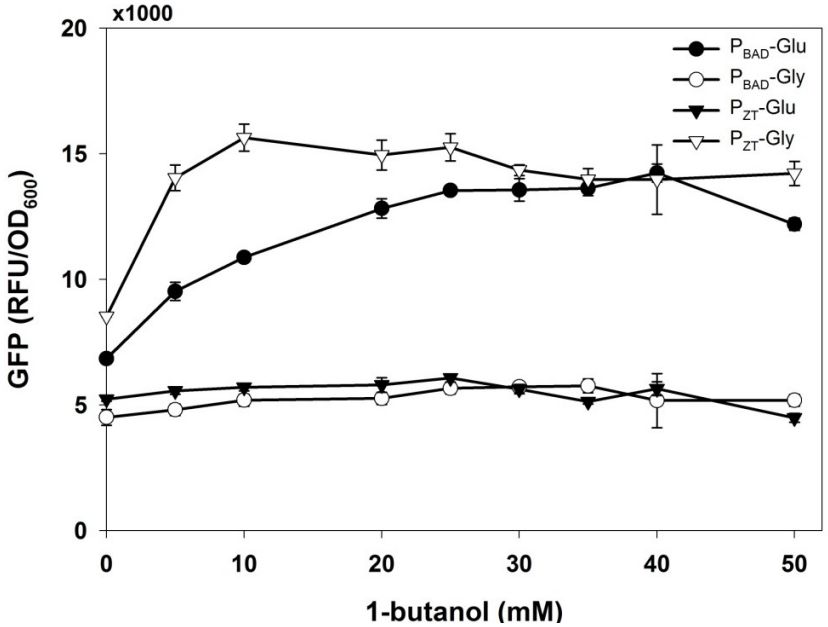


Figure 3.10: Comparison of PBAD and PBAD promoters driving expression of BmoR following medium supplementation with glucose or glycerol. Constructs pBMO#6 (P_{BAD}:b*moR*) and pBMO#7 (P_{ZT}:*bmoR*) were co-transformed with plasmid pBMO#1 (P_{BMO}:*gfp*) and GFP expression monitored following medium supplementation with glucose (Glu, 0.1% w/v) or glycerol (Gly, 0.2% v/v). The two systems exhibited divergent behavior with respect to carbon-source supplementation.

During optimization of the two-plasmid biosensor system it became clear that obtaining low-level, reproducible expression of soluble BmoR was essential for improving biosensor robustness. A two-plasmid biosensor design, while enabling rapid adjustment of the component parts, was a likely source of variability. Construction of similar heterologous promoters for protein over-expression in *E. coli* were based on single-plasmid designs[222,255]; furthermore, in these systems, transcription factor expression was driven by its native promoter. A series of plasmids based on P_{BmoR} promoter driving expression of BmoR were constructed. pBMO#35 is a medium copy (p15a origin) plasmid housing the P_{BmoR}:*bmoR* device, and is designed for use with pBMO#1 as part of a two-plasmid system. Plasmids pBMO#36 and pBMO#40 are single-plasmid systems that house both the P_{BmoR}:*bmoR* and P_{BMO}:*gfp* devices. The plasmids differ with respect to their origin of replication; pBMO#40 contains a pSC101 origin while pBMO#36 contains a ColE1 origin. Both constructs were tested in a DH1 background, and the high-copy number pBMO#36 plasmid was also tested in a DH1 (Δ*adhE*) host. The Δ*adhE* knockout strain is unable to produce ethanol, which we hypothesized was a source of background transcription in our system. All constructs were assayed over a four order-of-magnitude range in 1-butanol concentration (10 μM to 50 mM) in order to better define the biosensor linear range of induction (**Figure 3.11**)

The two-plasmid pBMO#35/pBMO#1 system exhibited a GFP expression profile similar to that witnessed for the two-plasmid system utilizing a P_{BAD} promoter. Of the single-plasmid systems – pBMO #40 (pSC101 origin) and pBMO#36 (ColE1 origin) – tested here, only the high-copy

pBMO#36 plasmid yielded detectable GFP expression. The low-copy pBMO#40 plasmid may not yield enough active BmoR to catalyze transcription from P_{BMO}, or GFP expression may be too low to be detected above background. The pBMO#36 system provided a higher dynamic range (≈61,000 GFP RFU/OD_{600}) and fold-induction (8.4) relative to the two-plasmid pBMO#35/pBMO#1 design. Over multiple experiments, the P_{BMOR}:*bmoR* based biosensor systems also proved more robust, exhibiting less variability between experiments.

When tested in the *E. coli* DH1 (Δ*adhE*) knockout strain, the background GFP expression levels decreased significantly relative to a wild type DH1 host (n=3; t-test, $p<0.05$). As expected, the decrease in background expression levels had a strong impact on fold-induction (39.4-fold) even though dynamic range decreased slightly. Although pBMO#36 performed well in both wild-type DH1 and the DH1 (Δ*adhE*) – and both strains are likely suitable candidates for implementation as a high-throughput screen –subsequent experiments were conducted in DH1 (Δ*adhE*) to take advantage of the decrease in background fluorescence.

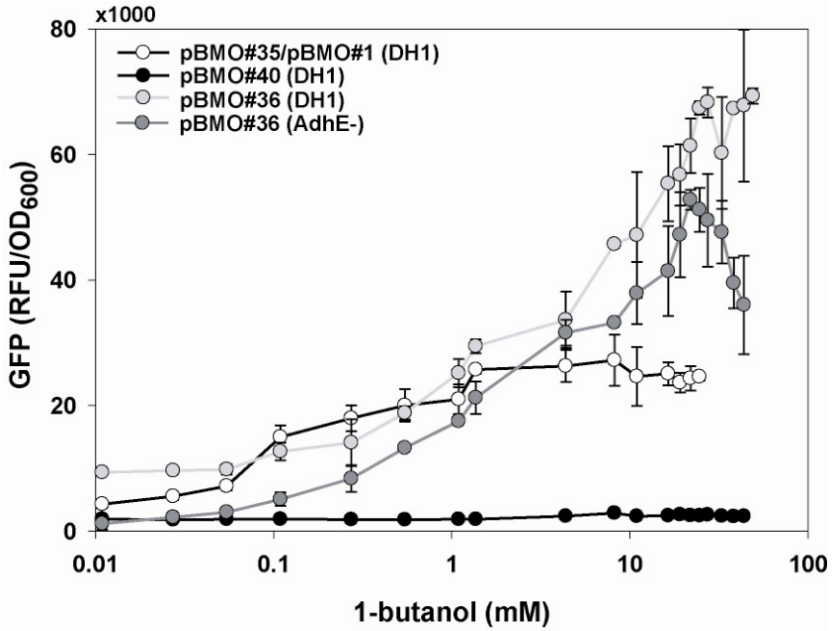


Figure 3.11: Comparison of P_{BmoR}:BmoR-based biosensor designs in wild type and engineered *E. coli*. Plasmid pBMO#35 (p15a origin; P_{BmoR}:*bmoR*) was co-transformed with plasmid pBMO#1 (ColE1 origin; P_{BMO}:GFP). Plasmids pBMO#40 (pSC101 origin; P_{BmoR}: *bmoR*, P_{BMO}:*gfp*) and pBMO#36 (ColE1 origin; P_{BmoR}:*bmoR*, P_{BMO}:*gfp*) were designed as single-plasmid biosensors. A linear response to exogenously added 1-buanol was observed in strains housing either the pBMO#35/pBMO#1 or pBMO#36 plasmids. pBMO#36 in a DH1 (Δ*adhE*) host demonstrated the highest fold-induction (greater than 39-fold) of all constructs tested. n=3; mean±s.d

3.4 Biosensor characterization

3.4.1 Characterization of BmoR operator site in promoter P_{BMO}

Having demonstrated robust 1-butanol induced GFP expression in *E. coli* we next turned our attention towards achieving a more accurate characterization of the P_{BMO} promoter sequence. σ^{54}-dependent transcription factors have been shown to bind several hundred base pairs upstream of the -24 and -12 consensus promoter region[256]; with this in mind, we included the 525 bp upstream of the transcription +1 site in the initial promoter design. To narrow our focus, we constructed and tested a series of truncated P_{BMO} promoter constructs driving GFP expression (**Figure 3.12**). Under inducing conditions, only background levels of GFP fluorescence were observed in constructs truncated prior to the -24/-12 consensus σ^{54}-RNA polymerase subunit binding site. Low-level fluorescence was observed in constructs truncated 175 bp and 225 bp upstream of the +1 transcription start site; similar fluorescence levels were observed in un-induced cell cultures, suggesting fluorescence is due to leaky, or butanol-independent transcription from P_{BMO}. A 3-fold increase in GFP fluorescence was observed in the two constructs housing either 325 bp or 425 bp upstream of the +1 transcription start site. Closer inspection of this region led to identification of an inverted repeat sequence spanning 227–264 base pairs upstream of the transcription start site (**Figure 3.13**).

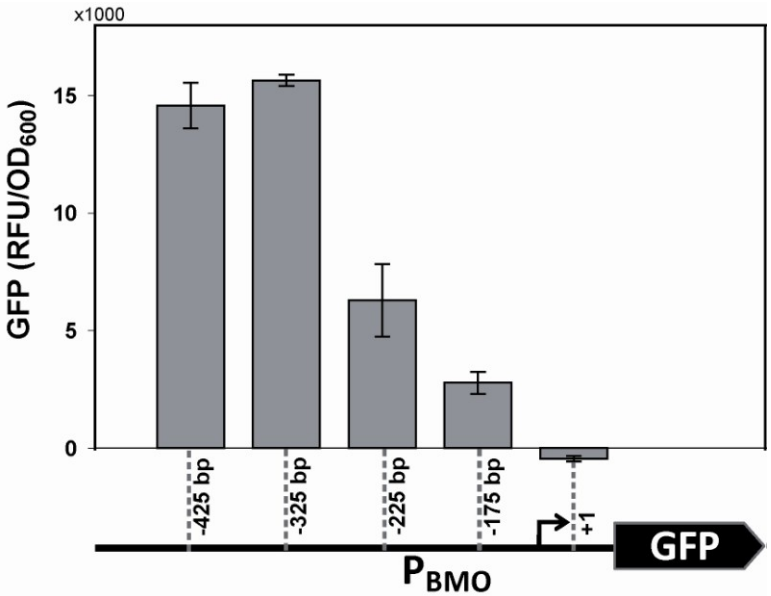


Figure 3.12: Essential elements of P_{BMO} promoter. σ^{54}-dependent transcriptional activators typically bind several hundred base pairs (bp) upstream of the +1 transcription start site; and because the BmoR operator site on P_{BMO} was not yet elucidated, the initial promoter design included 425 bp upstream of the +1 transcription start site. By sequentially shortening the total promoter length it was found that only 325 bp are necessary to have full promoter activation in the presence of 40mM 1-butanol. n=3; mean±s.d.

Figure 3.13: Putative BmoR operator site on P_{BMO} promoter. The C-terminal of BmoR contains a helix-turn-helix domain, a hallmark characteristic of transcription factors. An inverted repeat identified at 227–264 base pairs upstream of the transcription start site was identified as the putative BmoR operator. The location of the operator site corresponds well with the promoter truncation assay results.

3.4.2 Characterization of biosensor with GFP reporter

Construct pBMO#36 was used during whole-cell bioassay response characterization toward a range of short-chain, hydroxylated small molecule inducers (**Figure 3.14**). GFP fluorescence normalized to cell density were collected up to the IC_{75} – the concentration of inducing agent resulting in 75% of maximum cell density measured in the negative control after overnight growth – or up to 1M exogenously added inducing agent**.** In line with preliminary standards for characterization of biological devices[257] we measured a range of biosensor performance features. The Hill Equation (eq 3.1), commonly used to describe promoter activation, provided an accurate measurement of biosensor sensitivity, dynamic range, and switching-point.

$$GFP = GFP_0 + \left(\frac{GFP_{max} \bullet [I]^n}{K_m^n + [I]^n} \right) \qquad \text{eqt 3.1}$$

Where GFP_0 is the background GFP expression level, GFP_{max} is the maximum observed GFP expression, [*I*] is the inducer concentration, K_m is the inducer concentration resulting in half-maximal induction, and *n* is the Hill coefficient describing biosensor sensitivity. Biosensor performance features derived from the response curves are also presented (**Table 3.3**)
The biosensor exhibited a strong response to C4-C6 linear alcohols while 1-propanol and 1-heptanol elicited low or undetectable GFP signals over the assayed concentration range. Ethanol elicited a detectable GFP signal only when concentrations approached 1M. Previous experiments demonstrated a decrease in background GFP expression after knocking out the native *E. coli* ethanol dehydrogenase, AdhE. These results are suggestive of increased biosensor sensitivity to endogenously produced alcohols over those exogenously added. Interestingly, only 1-butanol exhibited a Hill coefficient less than 1, resulting in a much broader linear range of induction as compared to the other alcohols tested in the assay.

The biosensor was less responsive to C3-C5 branched-chain alcohols and aldehydes as compared to butanol; and the dynamic range for all branched-chain alcohols tested was 2-4 fold lower than observed with 1-butanol. Lastly, a slight, but statistically significant (t-test, $p<0.05$) increase in GFP/OD signal was obtained for 1,4-butanediol (BDO), but not 1,5-pentandiol.
In general, trends between biosensor performance characteristics and inducers were more complex than anticipated. For example, with the linear alcohols we expected to see a stronger trend emerge between alcohol chain length and both K_m and GFP_{max} values. Conversely, such a dramatic difference in biosensor sensitivity between various alcohols was unanticipated. The ambiguity in results can be explained, in part, by the complexity of an *in vivo* biosensor as compared to standard *in vitro* models of protein-ligand binding and promoter activation. First,

the inducing agents we explored have different volatilities and toxicities (**Figure 3.15**), and the *E. coli* host is placed under significant alcohol-induced cell stress at the concentrations tested. Second, all the small-molecule inducers tested enter *E. coli* by passive diffusion, and the partition coefficient between the hydrophobic membrane and the cytosol or growth medium will be different for all compounds tested. These factors, among others, complicate the analysis of trends in biosensor performance between the different inducers.

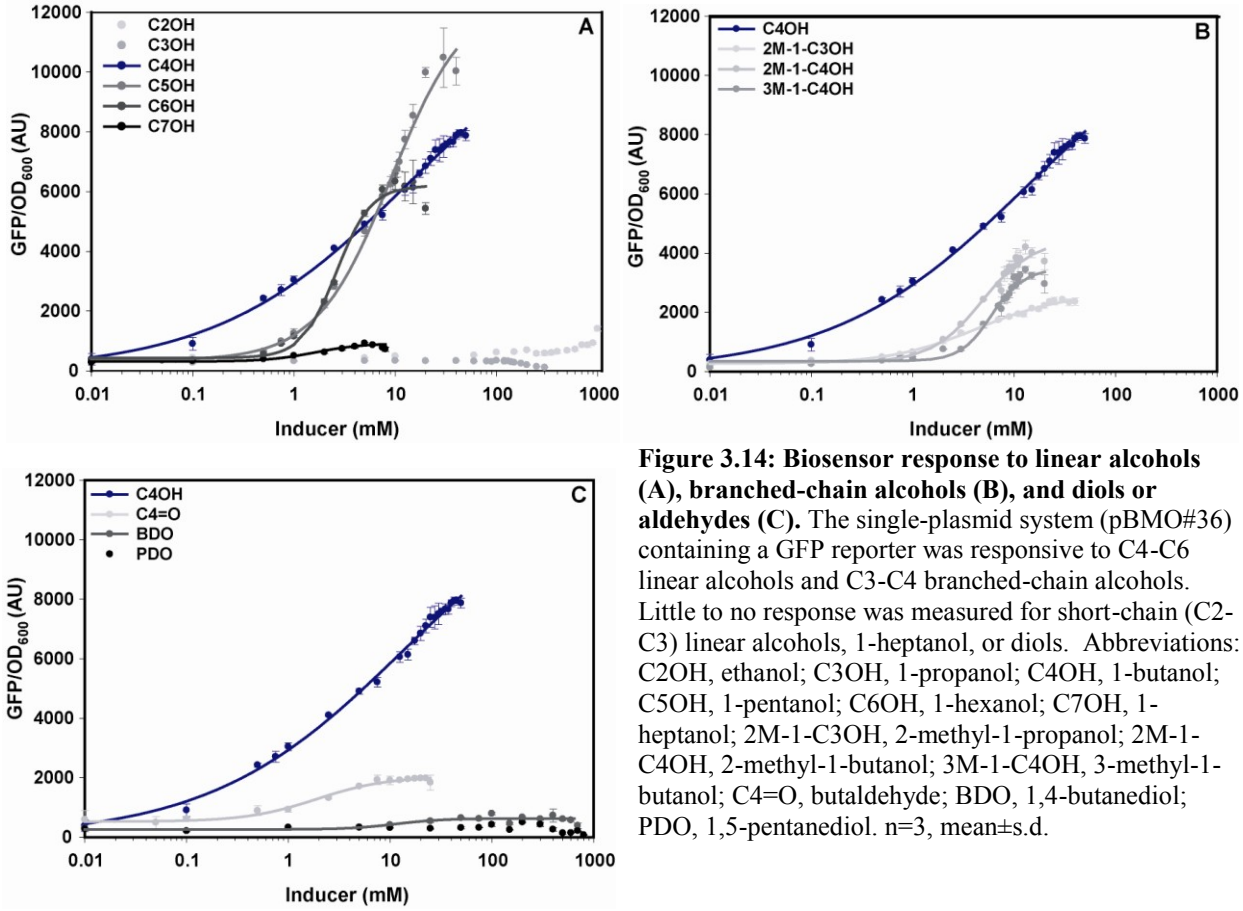


Figure 3.14: Biosensor response to linear alcohols (A), branched-chain alcohols (B), and diols or aldehydes (C). The single-plasmid system (pBMO#36) containing a GFP reporter was responsive to C4-C6 linear alcohols and C3-C4 branched-chain alcohols. Little to no response was measured for short-chain (C2-C3) linear alcohols, 1-heptanol, or diols. Abbreviations: C2OH, ethanol; C3OH, 1-propanol; C4OH, 1-butanol; C5OH, 1-pentanol; C6OH, 1-hexanol; C7OH, 1-heptanol; 2M-1-C3OH, 2-methyl-1-propanol; 2M-1-C4OH, 2-methyl-1-butanol; 3M-1-C4OH, 3-methyl-1-butanol; C4=O, butaldehyde; BDO, 1,4-butanediol; PDO, 1,5-pentanediol. n=3, mean±s.d.

Table 3.3: Biosensor performance features

Target	Sensitivity (n)	Dynamic Range (GFP_{MAX})	K_m (mM)	Selectivity (K_{C40H}/K_x)
C4OH	0.78	8000	3.83	1.00
C5OH	1.18	10,200	9.13	0.42
C6OH	2.54	5,700	2.63	1.45
C7OH	1.86	600	1.62	2.36
2M-1-C3OH	1.01	2,400	4.34	0.88
2M-1-C4OH	1.74	4,100	4.82	0.79
3M-1-C4OH	2.84	3,100	6.00	0.64
C5=O	1.90	1,500	1.90	2.01
PDO	1.87	400	11.49	0.33

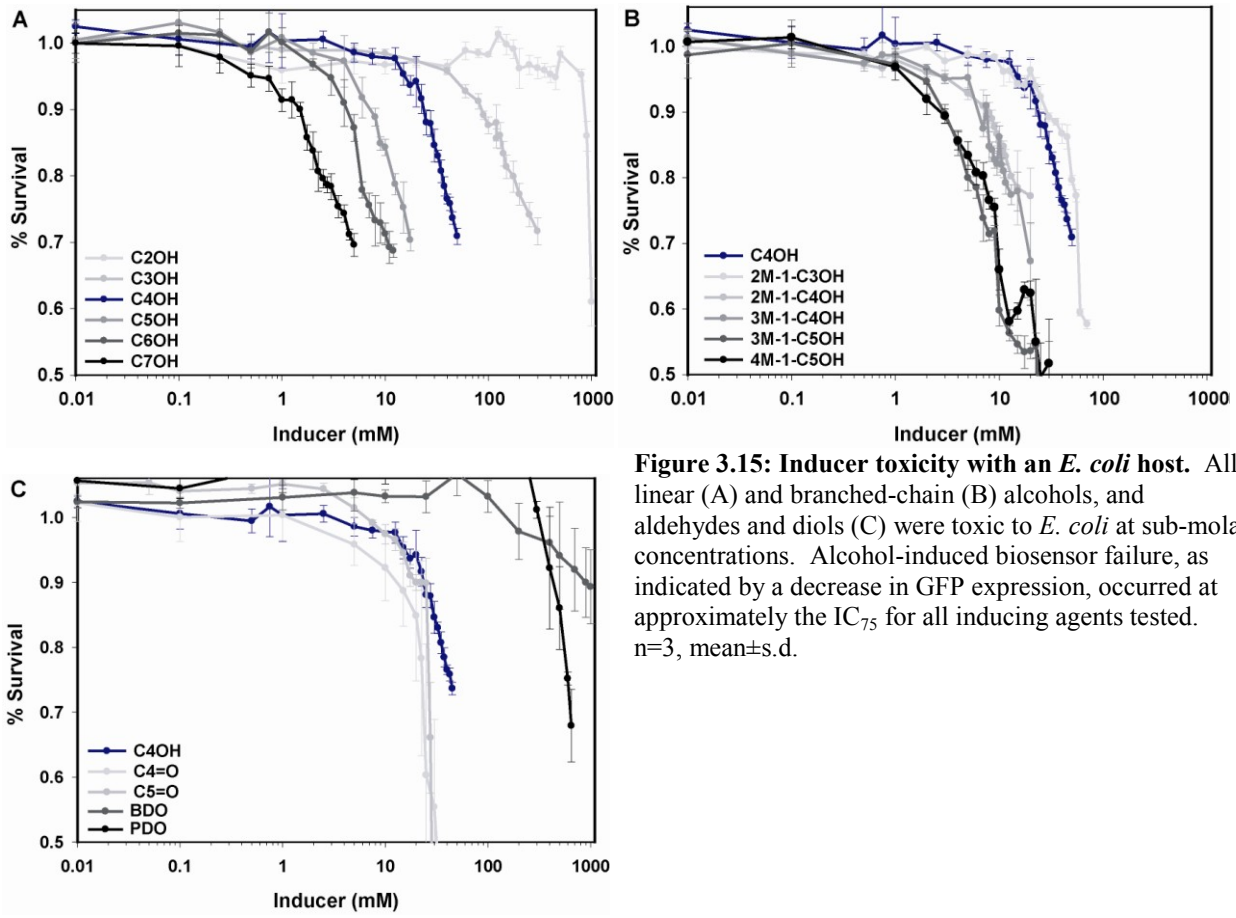


Figure 3.15: Inducer toxicity with an *E. coli* host. All linear (A) and branched-chain (B) alcohols, and aldehydes and diols (C) were toxic to *E. coli* at sub-molar concentrations. Alcohol-induced biosensor failure, as indicated by a decrease in GFP expression, occurred at approximately the IC_{75} for all inducing agents tested. n=3, mean±s.d.

While more conclusive results regarding BmoR-inducer binding kinetics require *in vitro* experimentation, a number of conclusions can be derived from the biosensor performance features. For exogenously added 1-butanol, the biosensor response was linear over nearly a three order-of-magnitude range inducer concentration, and the dynamic range is well suited for differentiating positive hits from a background, leaky GFP signal. In more detail, the whole-cell bioassay exhibits a z-score[258] of 0.93 over the linear range of detection; a z-score greater than 0.90 is generally regarded as an excellent high-throughput screening assay. As applied toward 2-methyl-1-propanol and 3-methyl-1-butanol, the major branched-chain alcohols being targeted for over-production in *E. coli*[218,259,260], biosensor responds weakly and only at high alcohol concentrations. As applied toward a branched-chain alcohol assay, a small dynamic range and shortened linear range of induction increase decrease the accuracy in which incremental increases in strain productivity can be identified.

3.4.3 Characterization of 1-butanol biosensor with TetA-GFP reporter

The choice of biosensor reporter strongly dictates the high-throughput screening application space (**Chapter 1, Figure 1.1**). Our P_{BMO}:*gfp*-based device is well suited for high-throughput liquid culture screening; spent production strain medium, containing the target alcohol, is titrated into a biosensor strain culture and fluorescence readout of alcohol concentration is readily obtained. A GFP reporter-based construct can also be used with higher-throughput screening

applications – including, solid-medium plates and fluorescence activated-cell sorting – but is not suitable for development of transcription factor-based selections.

A single reporter, capable of both screening and selection would be ideal. In addition to broadening the biosensor application space, incorporating a screening reporter in the design could provide confirmation of any positive hits resulting from the selection. A dual screening-selection reporter, as compared to a polycistronic reporter operon, would provide a more direct, accurate measure of selection protein concentration. For polycistronic operons, it is well established that differences in gene order affect transcription and translation efficiencies. And at the posttranslational level, different protein degradation rates would add further error.

A previously reported TetA-GFP fusion protein with a $(Gly\text{-}Gly\text{-}Gly\text{-}Ser)_4$ linker[261] appeared ideal for this application. The *tetA* gene encodes for a tetracycline/H^+ antiporter; a transport protein, as opposed to an enzyme, was chosen to maintain high-selective pressure against cells displaying poor tetracycline resistant in a heterogeneous population. Furthermore, a negative selection against alcohol-independent TetA-GFP transcription can be performed by adding nickel-chloride to the medium[262]. The resulting construct, pBMO#41, is identical to the reporter plasmid pBMO#36, with exception of a TetA-GFP fusion protein in place of the GFP reporter.

E. coli transformed with pBMO#41 exhibited butanol-dependent growth when tetracycline was supplemented to the growth medium (**Figure 3.16**). Growth rates were dependent on the presence of 1-butanol in the medium up to a concentration of 40 mM, at which point alcohol toxicity exceeded the positive tetracycline selective pressure. A negative control consisting of *E. coli* transformed with an empty vector grew only in the absence of tetracycline (data not shown). In negative selection mode, the correlation between *E. coli* growth rate and 1-butanol concentration was inverted.

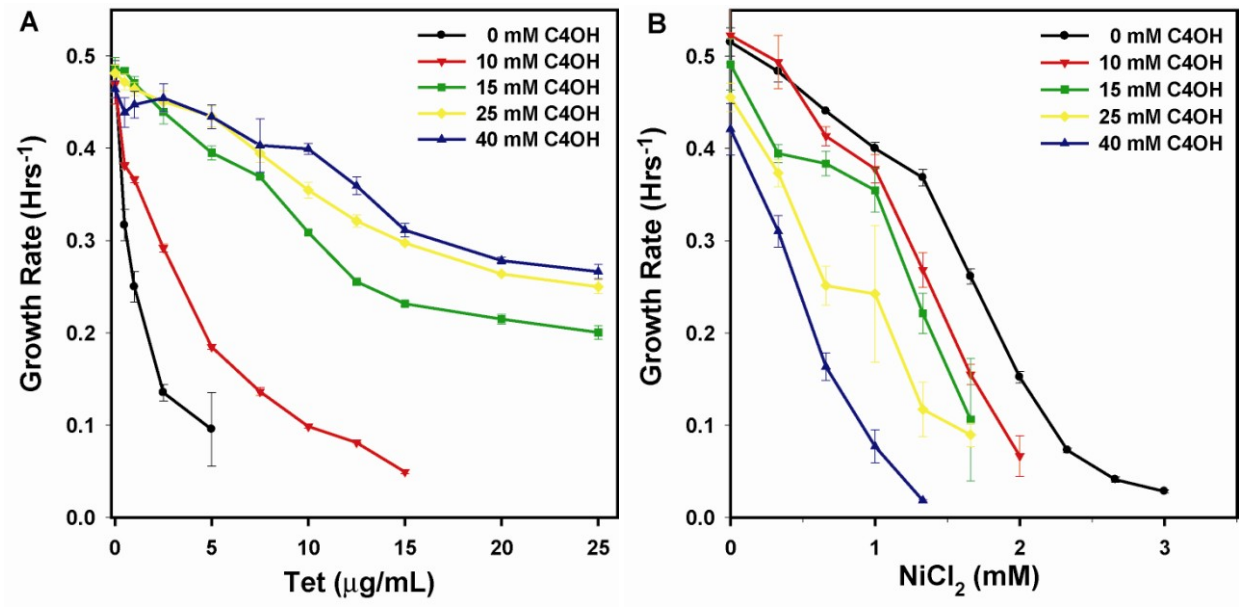


Figure 3.16: Positive and negative selection modes for presence of 1-butanol. (A) *E. coli* strain DH1 (Δ*adhE*) harboring the 1-butanol responsive biosensor plasmid pBMO#41 exhibited butanol-dependent growth upon addition of tetracycline up to a 40mM exogenously added 1-butanol. Control cultures lacking the biosensor grew only under conditions with no tetracycline added to the culture medium. (**B**) Addition of $NiCl_2$ to *E. coli* cell

cultures harboring plasmid pBMO#41 exhibited a negative correlation between the growth rate and the concentration of exogenously added 1-butanol. n=4, mean±s.d.

GFP fluorescence from the TetA-GFP fusion protein on pBMO#41 was also measured in these experiments (**Figure 3.17**). While a detectable increase in GFP signal was observed with increasing 1-butanol concentration, it was an order-of-magnitude lower than measured using GFP alone in construct pBMO#36. Under non-selective conditions, an approximately 5-fold induction was measured between 10 µM and 10 mM 1-butanol. The addition of tetracycline to the culture medium above 2.5 µg/mL, however, eliminated this trend. Given the relationship between *E. coli* growth rate and exogenously added 1-butanol, positive selection assays will be conducted at tetracycline concentrations higher than can be used to obtain a detectable GFP signal.

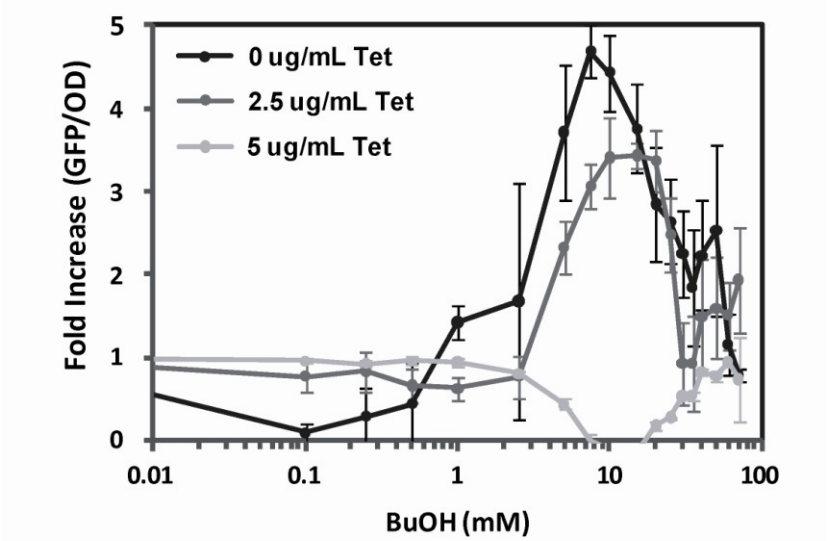


Figure 3.17: GFP fluorescence from pBMO#41 in presence of tetracycline. A butanol-dependent increase in fluorescence from the TetA-GFP fusion protein was detected only at tetracycline concentrations below 2.5 µg/mL. Higher tetracycline concentrations eliminated this trend, and resulted in decreased fluorescence between 1 and 40 mM exogenously added 1-butanol. At the tetracycline concentrations necessary to achieve a positive selection for 1-butanol a dual GFP screen is non-functional.

3.5 Conclusions

We presented here the construction and characterization of a short-chain alcohol responsive biosensor. The optimized biosensor incorporates two component parts, the BmoR transcription factor auto-regulated by its native promoter (P_{BmoR}:*bmoR*), and a reporter gene housed under transcriptional control of the P_{BMO} promoter (P_{BMO}:*gfp* or P_{BMO}:*tetA-gfp*). Using a GFP reporter, the biosensor performance features for a range of linear- and branched-chain alcohols, aldehydes, and diols were determined. The biosensor exhibited a linear response between 100 μM and 40 mM 1-butanol, and a dynamic range of over 8000 GFP/OD_{600} units; a 700 μM difference in 1-butanol concentration could be detected at 95% confidence. By replacing the GFP reporter with TetA, a tetracycline transporter, a 1-butanol selection was constructed; *E. coli* harboring the TetA-based biosensor exhibited 1-butanol dependent growth in the presence of tetracycline up to 40mM exogenously added 1-butanol.

The two optimized biosensor constructs can be directly implemented in high-throughput liquid culture or fluorescence-activated cell sorting screens, and as a selection for 1-butanol producing *E. coli*.

3.6 Materials and methods

3.6.1 Reagents

All enzymes and chemicals were purchased from Fermentas and Sigma-Aldrich Co., respectively, unless otherwise indicated. DNA oligomers were ordered from Integrated DNA Technologies (Coralville, IA).

3.6.2 Strains and Plasmids

Genbank files for all plasmids are included in **Appendix 2**. All plasmids were assembled from PCR product by sequence ligation independent cloning (SLIC)[263], unless otherwise indicated, using Phusion DNA polymerase (New England Biolabs). The *Pseudomonas butanovora* (ATCC# 43655) genes and promoters were cloned from genomic DNA.

Escherichia coli strain DH10b was used for all molecular cloning; all engineered *E. coli* strains were based on a DH1 or MG1655 background, as indicated. Introduction of an alcohol dehydrogenase knockout (Δ*adhE*) in a DH1 background was performed by λ Red-mediated gene deletion[264].

Table 3.4: Plasmids used in this study

Plasmid	Description	Source
pBMO#1	P_{BMO}:*gfp*, Amp^R, ColE1	This study
pBMO#6	P_{BAD}:*bmoR*, Cm^R, p15a	This study
pBMO#7	P_{ZT}:*bmoR*, Cm^R, p15a	This study
pBMO#35	P_{BmoR}:*bmoR*, Cm^R, p15a	This study
pBMO#36	P_{BmoR}:*bmoR*, P_{BMO}:*gfp*, Amp^R, ColE1	This study
pBMO#40	P_{BmoR}:*bmoR*, P_{BMO}:*gfp*, Amp^R, pSC101	This study
pBMO#41	P_{BmoR}:*bmoR*, P_{BMO}:*tetA*, Amp^R, ColE1	This study

3.6.3 Protein Purification

Batch purification of a 6X-his tagged version of BmoR was achieved through a nickel-affinity column. The over-expression *E. coli* strain BLR (De3) was transformed with the plasmid pET29b–His-BmoR. 5 mL LB broth supplemented with Cb^{50} was inoculated with a single colony of BLR(dE3) harboring the plasmid pET29b:His-BmoR and grown overnight at 37 °C. The full overnight culture was used to inoculate 500 mL of fresh Terrific Broth supplemented with 0.4% v/v glycerol and Cb^{50}. Cells were initially grown at 37°C shaking at 150 rpm until an OD_{600}=0.25, cultures were then induced with 0.05 mM IPTG and grown for an additional 12 hours at 20°C. Cultures were centrifuged (10 min, 4°C, 8000 rcf) and the pellet resuspended in 10 mL Buffer A (50 mM Tris (pH8.0) 50 mM KCl) supplemented with a 20 mM imidazole and protease inhibitor cocktail (EMD Biosciences). Cell membranes were sonicated and the soluble fraction separated by centrifugation (10 min, 4°C, 10000 rcf) and then filtered through a 0.45-µm filter. Purification was done using a Ni+-agarose matrix (Quiagen) according to the manufacture's protocols. Briefly, 5 mL resuspended slurry was equilibrated with 25 mL buffer A supplemented with 20 mM imidazole at 4°C. Following, the soluble protein fraction was bound the matrix for 15 min, and then washed with 25 mL wash solutions (2 x 50 mM imidazole, 2 X 100 mM imidazole, 1 X 150 mM imidazole) and then eluted with 10 mL 500 mM imidazole. The eluent was concentrated via xxx (Ambion, 32K MWCO) and analyzed by SDS-PAGE to

estimate protein purity. The BmoR solution was dialyzed against Buffer A to remove any imidazole; for long term use the solution was dialyzed against Buffer A supplemented with 30% glycerol and stored at -20°C

3.6.4 Gel mobility shift assays

Gel mobility shift assays were performed to further localize the DNA-binding region using a DIG-Gel Shift Kit (Roche Applied Science) following the manufacturer's protocol. The pBMO promoter was divided into 5x125 bp segments with approximately 60-bp overlaps covering 350 bp upstream of the pBMO +1 transcription start site. 10 mM BmoR was incubated with 30 fmol DIG-labeled PCR-products for 15 min at 25°C. The solutions were then run on a 12% Tris-glycine gel (Novex) in 50 mM Tris-HCl, 50 mM KCl buffer at 150V. DNA products were then transferred to nylon membrane at 25V for 45 min.

3.6.5 96-well plate biosensor characterization

E. coli strain DH1 (Δ*adhE*) harboring either biosensor plasmid pBMO#36 or pBMO#41 was cultured overnight in LB medium (Cb^{50}, 200 rpm, 30°C). Cultures were then inoculated 1% v/v into fresh EZ-rich medium (Teknova) supplemented with antibiotic (Cb^{50}), and grown until final cell densities reached an OD_{600}=0.20 (200 rpm, 30°C). Biosensor culture were diluted 1:4 in fresh EZ-rich medium (0.5% w/v glucose, Cb^{50}) supplemented with a known concentration alcohol in 96 deep-well plates (2-mL total capacity, polypropylene, square-bottomed; Corning). Cultures were incubated for 16 hrs (200 rpm, 30°C). Both fluorescence and absorbance measurements were performed on dual spectrophotometer-fluorometer (Spectromax M2, Molecular Devices). GFP fluorescence was measured using an excitation wavelength of 400 nm and an emission wavelength of 510 nm. Optical density measurements were monitored at 600 nm (OD_{600}). GFP fluorescence values were first normalized to OD_{600} (GFP/OD_{600}). *E. coli* auto-fluorescence was subtracted using a standard curve of GFP fluorescence from wild type *E. coli* optical density. Fold-induction was calculated as the difference between the averages of the induced and un-induced GFP fluorescence measurements normalized to the un-induced GFP measurement.

3.6.6 96-well plate growth assays and growth rate calculations

Selective pressure, as measured by cell growth rates, was determined in 96-well plates. *E. coli* strain DH1 (Δ*adhE*) harboring plasmid pBMO#41 was cultured overnight in LB medium supplemented with 0.5% w/v glucose (Cb^{50}, 200 rpm, 30°C). Cultures were then inoculated 1% v/v into fresh EZ-rich medium (0.5% w/v glucose, Cb^{50}), grown until final cell densities reached an OD_{600}=0.20 (200 rpm, 30°C), and subsequently diluted 1:4 in inducing medium to a final volume of 150μL in 96-well plates (2-mL total capacity, polypropylene, square-bottomed; Corning). The cultures were then grown for 1 hr before addition of tetracycline or $NiCl_2$ at the indicated concentrations; cultures were incubated in a 96-well plate reader (30°C, Tecan) and OD_{600} measurements taken every 15 minutes for a total culture time of 20 hours. Growth rates were determined by first normalizing each curves to the starting cell density $\ln\left(\frac{\quad}{\quad}\right)$ and fitting to a modified Gompertz equation for microbial growth[265]:

$$y = Ae^{-e^{\left(\left(\frac{\mu_m \cdot e}{A}\right)\cdot(\lambda - t)+1\right)}}$$

Where A is the maximum cell density $\ln(N/N_o)$, λ is the lag period (hrs), t is the time (hrs), and μ_m is the maximum specific growth rate (hrs^{-1}). Growth rates are all reported as mean±s.d. (n=3).

3.6.7 Z-score Calculation

Z-scores for 1-butanol induced GFP expression from pBMO#36 were calculated based on results obtained from 96-well plate response curves as described above. The limits of linear range of detection were established at 0 and 40mM 1-butanol. Z-scores were calculated as follows:

$$Z = 1 - \left(\frac{(3\sigma_s + 3\sigma_c)}{|\mu_s - \mu_c|}\right)$$

Where μ and σ are the mean and standard deviation, respectively, and *s* refers to sample and *c* refers to control.

Chapter 4. High-Throughput Screens and Selections Using an Alcohol-Responsive Transcription Factor-Promoter Pair

4.1 Expression of heterologous 1-butanol biosynthetic pathways in engineered *E. coli*

A series of 1-butanol production plasmids and strains (see **4.4 Materials and methods**, **Table 4.2**) were constructed for use as positive controls in proof-of-principle screens and selections. All plasmids were reconstructions of published work on heterologous expression of either the *C. acetobutylicum* 1-butanol biosynthetic pathway[205], or a 2-keto acid decarboxylation and reduction pathway for mixed alcohol biosynthesis[218,266].

4.1.1 Expression of C. acetobutylicum 1-butanol biosynthetic pathway in engineered E. coli

Plasmid pBUT#50, harboring the *C. acetobutylicum* 1-butanol biosynthetic genes, is composed of two operons (*crt.bcd.etfBA.hbd* and *atoB.adhE2*) under control of IPTG inducible P_{TRC} promoters (**Figure 4.1**). While the gene products required for 1-butanol biosynthesis are identical to those previously published[205], the pBUT#50 plasmid design differs significantly. In more detail, Atsumi et al. constructed a two-plasmid system using P_{LacO1} promoters and native RBS sequences. In contrast, transcriptional control in our one-plasmid design is performed by two P_{TRC} promoters, and our design incorporates both synthetic ribosome binding sites and a 5'-untranslated region (5'-UTR).

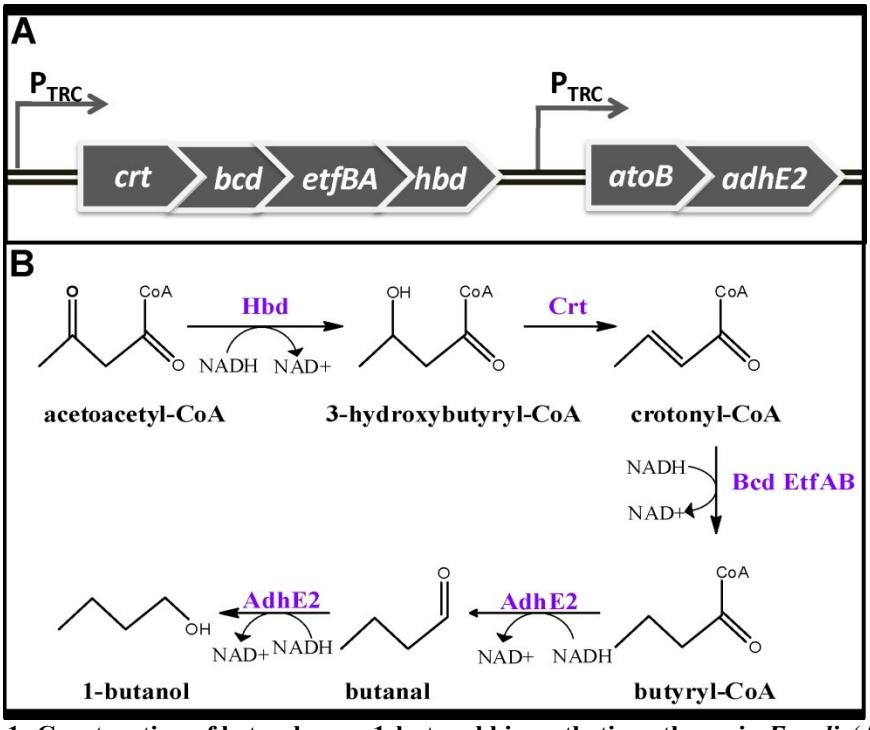


Figure 4.1: Construction of heterologous 1-butanol biosynthetic pathway in *E. coli*. (**A**) The *C. acetobutylicum* biosynthetic genes were constructed as a two-operon system under the control of IPTG inducible P_{TRC} promoters. Genes encoding for *E. coli* acetyl-CoA acetyltransferase, *atoB*, and *C. acetobutylicum* alcohol dehydrogenase, *adhE2*, were housed on a single operon. Genes encoding for *C. acetobutylicum* crotonase, *crt*, butyryl-CoA dehydrogenase, *bcd*, two-electron transferring flavoproteins, *etfAB*, and 3-hydroxybutyryl-CoA dehydrogenase, *hbd*, were housed on a second operons. (**B**) The 1-butanol biosynthetic pathway from acetoacetyl-

CoA; 1-butanol biosynthesis is highly NADH dependent, requiring 4 molecules of NADH per molecule 1-butanol.

1-butanol biosynthesis branches from central metabolism at acetyl-CoA, and thus competes with native *E. coli* fermentation pathways for carbon and NADH (**Figure 4.2**). Engineering *E. coli* for homobutanologenic fermentation depends on constructing a redox balanced, anaerobic pathway. In addition to heterologous pathway overexpression, a number of chromosomal modifications are necessary to increase NADH availability and eliminate native fermentation byproducts. Atsumi et al.[205] demonstrated production of 1-butanol titers of over 500 mg/L in rich medium by knocking out a series of *E. coli* fermentation pathways and anaerobic regulatory genes; the knockouts included the *adhE*, *ldhA*, *frdBC*, *fnr* and *pta* genes. The resulting homo-butanologenic strain, however, is still not redox balanced under anaerobic conditions. Native *E. coli* routes carbon through pyruvate formate lyase during anaerobic growth – which yields no NADH – and through the pyruvate dehydrogenase complex during aerobic growth – which yields 1 NADH per acetyl-CoA molecule formed. By knocking out *pflB*, Atsumi et al. attempted to reroute carbon through the pyruvate dehydrogenase complex under anaerobic conditions; however, the resulting strains failed to grow under these conditions. In line with this finding, optimal production was observed under micro-aerobic growth conditions, presumably because this maximized the redox capacity in engineered *E. coli*.

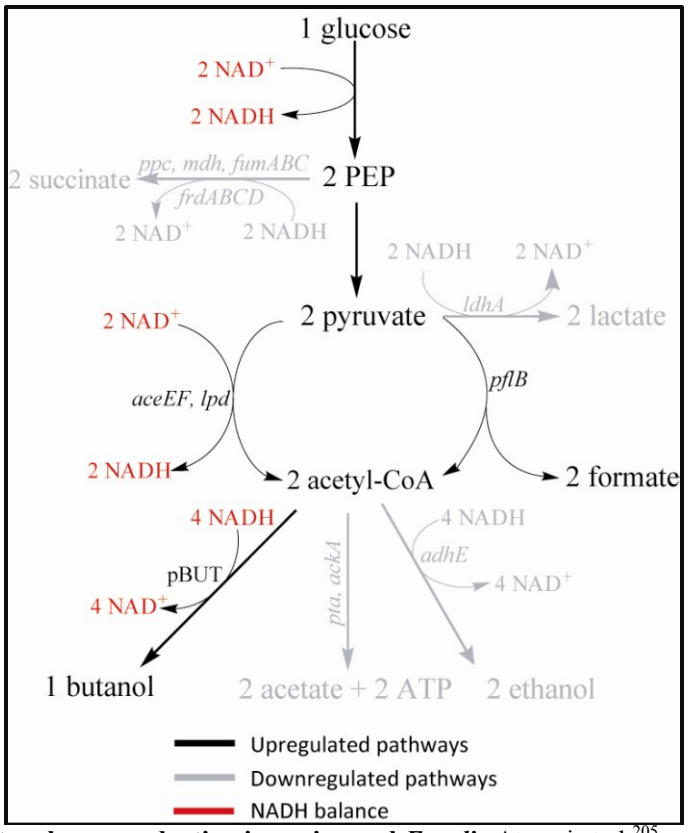


Figure 4.2: 1-butanol overproduction in engineered *E. coli*. Atsumi et al.[205] achieved high-level 1-butanol production in E. coli under micro-aerobic conditions be overexpressing the *C. acetobutylicum* pathway genes (pBUT pathway) and deleting native fermentation pathways competing for NADH.

Along these lines, our proof-of-principle demonstration of the biosensor-based screening strategy required a series of positive control strains exhibiting different 1-butanol productivities. We first

compared 1-butanol titers in wild type DH1 or DH1 (Δ*adhE*) transformed with the 1-butanol production plasmid pBUT#50 (**Figure 4.3A**). The two strains were grown at 5-mL culture scale under micro-aerobic conditions in either luria-bertani broth (LB) or terrific broth (TB). After 24 hours, only minute quantities (<5 mg/L) of 1-butanol were detected; after 48 hours, the DH1 (Δ*adhE*) strain exhibited significantly improved 1-butanol titers as compared to wild type DH1 (t-test, $p<0.05$). This result suggests that increased acetyl-CoA and NADH availability in the engineered strain improved titers. 1-butanol titers under all conditions, however, were less than 15 mg/L (≈250 μM).

We hypothesized that either IPTG induction timing or strength contributed to low product titers, and we tested both strains in undefined medium containing autoinduction sugars[267]; both strains were also tested under aerobic, anaerobic, and micro-aerobic growth conditions to confirm the importance of oxygen availability in our system (**Figure 4.3B**). Autoinduction of pathway expression resulted in over 10-fold higher 1-butanol titers as compared to IPTG-based induction; the ratio of 1-butanol titers between the wild type and Δ*adhE* knockout strain remained the same in autoinduction medium. Micro-aerobic growth resulted in an approximately 3-fold improvement in 1-butanol titers as compared to aerobic growth conditions; however, DH1 (Δ*adhE*) also performed well – albeit with a high level of variability – under anaerobic growth conditions, and 1-butanol titers ranged from 36.5 to 202 mg/L. The high degree of variability may suggest problems obtaining homogenous protein expression across biological replicates under these conditions.

The results from the 5-mL culture experiments in autoinduction medium suggested the pBUT#50 plasmid system was well suited for use in proof-of-concept testing of our biosensor-based liquid culture screen. Alcohol titers ranged between 10-200 mg/L – depending on oxygen availability, medium type, and strain – and the ratio of 1-butanol between the supernatant and cell pellet was 1.02 ± 0.14 (n=15, mean ± s.d.), an indication that endogenously produced 1-butanol readily diffuses across the *E. coli* cell membrane. Lastly, the use of autoinduction medium eliminates the requirement of adding an exogenous inducer, a particular unwieldy step during scale-up for high-throughput screening assays.

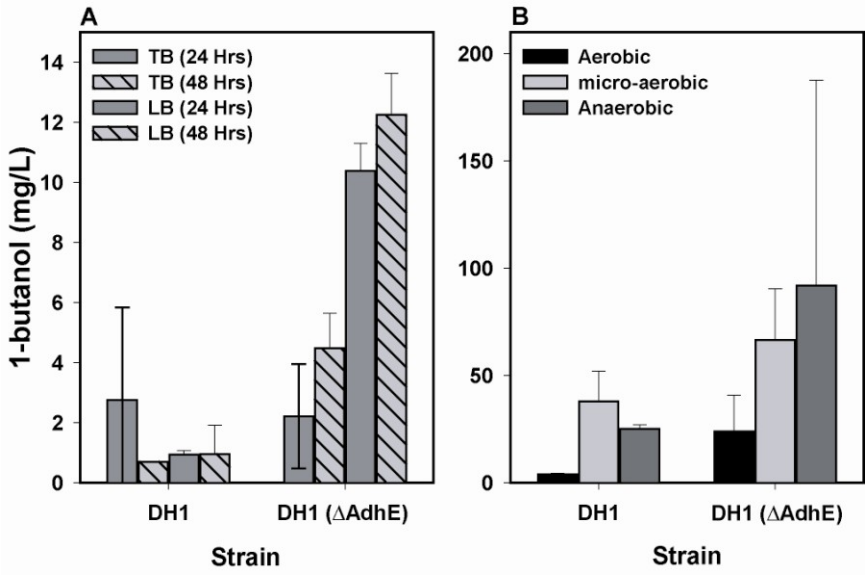


Figure 4.3: 1-butanol titers from engineered E. coli. Wild type DH1 and engineered DH1 (Δ*adhE*) were transformed with 1-butanol production plasmid pBUT#50, and 1-butanol titers measured at 24 and 48 hours under different medium and oxygenation conditions. (**A**) *E. coli* strains DH1 and DH1 (Δ*adhE*) were grown under micro-aerobic conditions in luria-bertani (LB) and terrific broth (TB) mediums following induction with 1mM IPTG. (**B**) Wild type DH1 and engineered strain DH1 (Δa*dhE*) were grown in autoinduction medium[267] for 48 hours under aerobic, micro-aerobic, and anaerobic growth conditions. n=3, mean ± s.d.

In an effort to further increase 1-butanol titers we tested a series of additional knockout mutations. Lactate dehydrogenase, the protein product of *ldhA*, catalyzes reduction of pyruvate to lactate under anaerobic conditions yielding one NADH in the process. As with AdhE-derived ethanol, knocking out lactate production should increase the pool of NADH available for reduction of butyryl-CoA to 1-butanol. We also targeted FNR, the primary transcription factor regulating the shift from aerobic to anaerobic growth in *E. coli*[268], in an attempt to increase anaerobic expression of the pyruvate dehydrogenase complex and create a redox balanced 1-butanol biosynthetic pathway. Lastly, to generate a more stable production host we integrated the heterologous 1-butanol pathway onto the *E. coli* chromosome. The resulting strains were tested under both micro-aerobic and anaerobic growth conditions (**Figure 4.4**).

Under micro-aerobic growth conditions, all knockout strains exhibited similar production titers; however, because cell growth was low (<1.0 OD_{600} units) specific production from the Δ*ldhA*/Δ*adhE* double mutant was improved over both the Δ*ldhA* single mutant and the Δ*ldhA* /Δ*adhE*/Δ*fnr* triple mutant. Under anaerobic growth conditions, the Δ*ldhA* single knockout exhibited the highest absolute, but lowest specific, production titers. In general, strain β (strain MAL with a chromosomal copy of the P_{TRC}:*crt.bcd.etfBA.hbd* and P_{TRC}:*atoB.adhE2* operons) performed similarly to engineered strains transformed with the pBUT#50 alcohol production plasmid. As compared to the plasmid-based expression systems, β exhibited a 2-fold increase in 1-butanol titers under micro-aerobic growth conditions. This finding suggests that plasmid-based pathway expression is a non-optimal production method.

It was also discovered, however, that after multiple passages β would spontaneously lose the ability to produce 1-butanol. Interestingly, the antibiotic resistance cassette remained present on the chromosome, but the P_{TRC}:*crt.bcd.etfBA.hbd* operon had been excised, potentially by recombination between the two P_{TRC} promoters in the device architecture. The P_{TRC} promoter is known to exhibit high levels of leaky expression[269], and it is probable that stress from leaky pathway expression generated strong selective pressure against pathway maintenance.

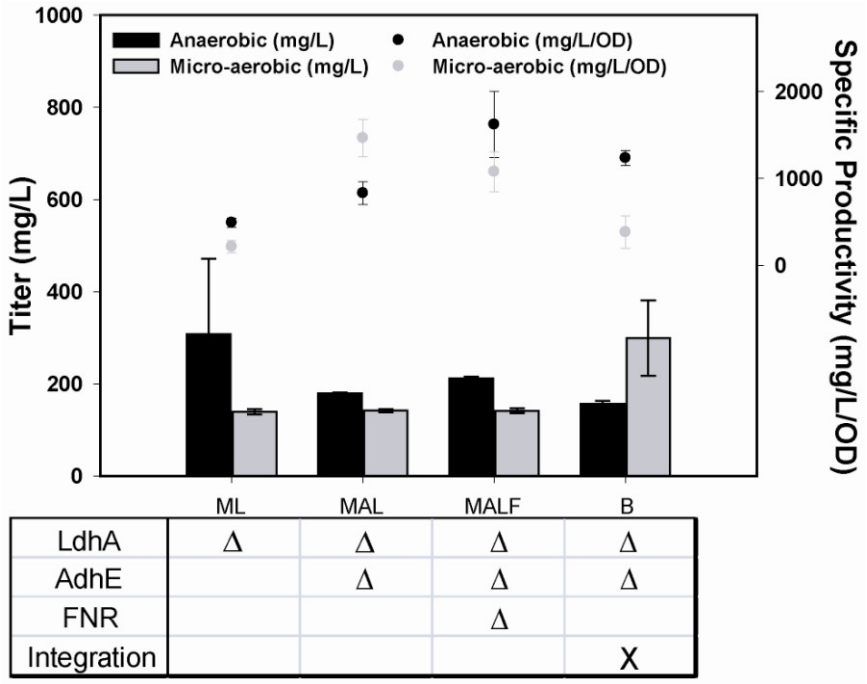


Figure 4.4: Production of 1-butanol from *E. coli* knockout strains under anaerobic and micro-aerobic conditions. A series of knockout strains were first constructed in an MG1655 background (ML: MG1655(Δ*ldhA*); MAL: MG1655(Δ*ldhA*, Δ*adhE*); MALF: MG1655(Δ*ldhA*, Δ*adhE*, Δ*fnr*)) and transformed with 1-butanol production plasmid pBUT#50. Strain β (MG1655(Δ*ldhA*, Δ*adhE*, *intA*::pBUT) was constructed using the lambda-red recombinase system[264] to integrate the butanol biosynthesis genes and antibiotic resistance marker derived from pBUT#50 at the *intA* gene locus. n=3; mean ± s.d.

While our engineered strains were demonstrating 1-butanol titers on par with previous reports, we did not witness any knockout-dependent improvements in titer. To better account for carbon in our engineered *E. coli* strains, we measured the major fermentation byproducts with and without expression of our 1-butanol biosynthetic pathway (**Table 4.1**).

Lactate and formate (not shown) were not observed under any of the conditions tested. Acetate, succinate, and ethanol were present at approximately the same concentration, and account for between 7.5-25% of potential 1-butanol theoretical yield from engineered *E. coli*. Ethanol production was observed in strain ML, which does not possess the Δ*adhE* knockout, and in strains expressing the heterologous AdhE2 alcohol dehydrogenase. *In vitro* assays with AdhE2 have described activity toward acetyl-CoA[205,206], highlighting ethanol production as a major hurdle in efforts to engineer a homo-butanologenic strain of *E. coli*.

Table 4.1 Mixed acid fermentation products from engineered *E. coli* strains

With 1-butanol production (g/L)								
	ML		MAL		MALF		Beta	
	0	1	0	1	0	1	0	1
Succinate	0.27±0.087	0.24±0.003	0.10±0.01	0.11±0.08	0.06±0.006	0.14±0.02	0.09±0.01	0.08±0.04
Acetate	0.17±0.063	0.16±0.031	0.09±0.02	0.28±0.26	0.09±0.02	0.12±0.02	005±0.022	0.13±0.04
Ethanol	0.35±0.063	0.35±0.052	0.12±0.036	0.20±0.03	0.14±0.02	0.14±0.02	0.096±0.03	0.16±0.04
Empty vector control (g/L)								
	ML		MAL		MALF			
	O	1	O	1	O	1		
Succinate	0.24±0.008	0.17±0.08	0.07±0.004	0.06±0.01	0.04±0.011	0.05±0.036		
Acetate	0.19±0.052	0.26±0.12	0.05±0.014	0.12±0.02	0.04±0.007	0.35±0.23		
Ethanol	0.25±0.027	0.23±0.03	ND	ND	ND	ND		

0 corresponds to anaerobic growth, and 1 corresponds to micro-aerobic growth; n=3, mean±std dev

All strain testing and culture optimization up this point were performed in 5-mL culture tubes. A high-throughput liquid culture screen, however, necessitates production be performed in 96-well plate format. Scaling culture size down from 5-mL tubes to 96-well plates (containing less than 1-mL total culture volume) typically results in decreased oxygen availability. In our previous work with the isoprenoid pathway, scaled-down cultures exhibited lower cell densities and production titers. We did not predict oxygen availability to be an issue for 1-butanol production based on results from 5-mL cultures under micro-aerobic and anaerobic growth.

Lastly, the culture conditions should be optimized to reduce variability among biological replicates; a difficult to achieve goal, but one that improves the accuracy of the screening assay. We attempted to address this issue by measuring 1-butanol titers for biological (n=10) and technical replicates (n=3) across a 96 deep-well plate (**Figure 4.5**). Additional variables included growth medium (undefined autoinduction[267] versus defined EZ-rich with IPTG induction) and oxygen availability.

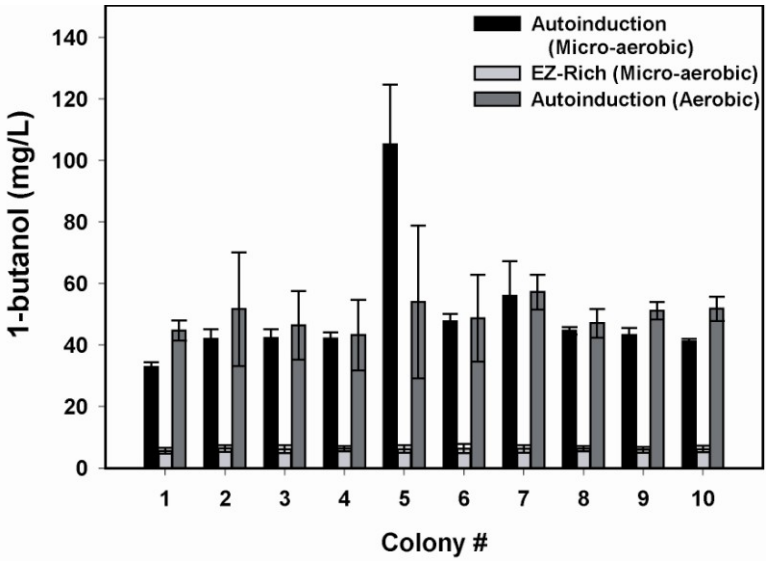


Figure 4.5: Analysis of 96-well plate 1-butanol titers and variability. Variability in 1-butanol titers from 10 colonies (3 biological replicates) were analyzed in undefined rich (autoinduction[267]) and defined rich (EZ-rich[270]) mediums was analyzed under micro-aerobic conditions. Autoinduction medium was tested under both micro-aerobic and aerobic growth conditions. n=3, mean ± s.d.

EZ-rich medium with IPTG induced pathway expression yielded significantly lower 1-butanol titers as compared to autoinduction medium (t-test, $p<0.05$). No statistical significance was found between micro-aerobic and anaerobic growth in autoinduction medium (t-test, $p>0.05$), suggesting that oxygen remains limiting even under conditions promoting aerobic growth in 96-well plate format. In autoinduction medium under micro-aerobic growth, the coefficient of variance in 1-butanol titers was 41.4% (15.8% excluding outlier colony #5). A coefficient of variance less than 10% for biological replicates is preferred for high-throughput screening assays. To decrease the variance in 1-butanol titers from cultures grown in 96 deep-well plates we investigated multiple plate sealing, incubation temperatures, and aeration strategies. In brief, using a plate sealer quipped with gas-impermeable film provided the most uniform results and lowest well-to-well variability in 1-butanol titers. Decreasing the temperature to 25°C also decreased the variance, but at the expense of cell density and 1-butanol titer. Increased culture aeration by running experiments at higher rpm's on various shaker platforms increased the variance. The final, optimized format used vacuum-sealed film, culture growth at 25°C, and with low aeration (200 rpm). Through these efforts the percent coefficient of variation was reproducibly decreased to 7-8%.

Numerous hurdles were encountered when developing and optimizing the heterologous 1-butanol biosynthetic pathway in engineered *E. coli*. Robust, high-titer production was never achieved using our system, and we believed the culture conditions required to achieve high-level 1-butanol titers would be difficult to translate to a high-throughput screen or selection strategy. First, optimal 1-butanol production was witnessed under micro-aerobic conditions in undefined medium; by contrast, the biosensor was optimized for performance in a defined medium under aerobic conditions. Second, plasmid-based 1-butanol production was highly variable between strains and experiments. No difference in 1-butanol titers was frequently observed between wild type and engineered *E. coli*, even though the engineered strains were reported in the Literature to display superior production. All efforts to improve 1-butanol titers by introduction of additional

knockout mutations were unsuccessful. In the absence of a set of robust control strains exhibiting different 1-butanol productivities, we concluded it would be difficult to integrate our heterologously expressed *C. acetobutylicum* pathway with our biosensor screening and selection strategy. A small number of preliminary screening experiments were conducted using the *C. acetobutylicum* pathway, and are highlighted below (**4.2 High-throughput liquid culture screening**).

4.1.2 Expression of L. lactis KivD and S. cerevisiae ADH6 in engineered E. coli

A smaller, more well-defined alcohol production system would facilitate proof-of-principle demonstration of a biosensor-based screening and selection strategy. We believed that a 2-keto acid-based alcohol biosynthetic pathway could better address this criterion. High level production of mixed 2-keto acid-derived alcohols in non-fermentative, aerobic growth[218,220,259,266,271,272]. Through the activity of a promiscuous *L. lactis* 2-keto acid decarboxylase (KivD) and *S. cerevisiae* alcohol dehydrogenase (ADH6), engineered *E. coli* produces three biosensor responsive alcohols: 2-methyl-1-propanol, 1-butanol and 3-methyl-1-butanol (**Figure 4.6**).

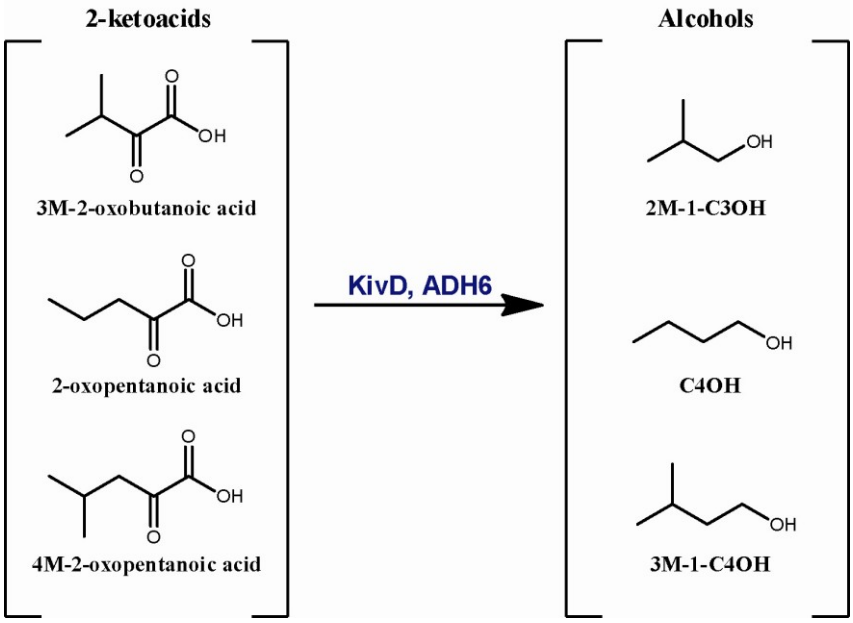


Figure 4.6: Production of biosensor responsive alcohols from 2-keto acids. Expression of *L. lactis* KivD and *S. cerevisiae* ADH6 in *E. coli* leads to production of a broad range of mixed alcohols[218,266]. The BmoR-based biosensor responds to 2-methyl-1-propanol (2M-1-C3OH), 1-butanol (C4OH) and 3-methyl-1-butanol (3M-1-C4OH).

While a biosynthetic pathway composed of a substrate promiscuous 2-keto acid decarboxylase and alcohol dehydrogenase has the capability of producing a wide variety of alcohols, only a small subset is produced in wild type *E. coli*. 2-keto acids are intermediates in amino acid biosynthesis, thus the measured titer for any particular alcohol is reflective of the carbon flux through the cognate amino acid pathway. For example, the major alcohols measured in wild type *E. coli* are 1-propanol and 2M-1-butanol, derived from isoleucine biosynthesis, 2M-1-propanol, derived from valine biosynthesis, and 3M-1-butanol, derived from leucine biosynthesis (**Figure 4.7**). 1-butanol is derived from an intermediate (2-oxopentanoate) in norvaline

biosynthesis, a minor offshoot of the isoleucine pathway[273]. High titer (over 600 mg/L) 1-butanol production was reported in engineered *E. coli* with ilvD knocked out, overexpressing *ilvA* and *leuABCD* operon, and feeding 8 g/L threonine[218].

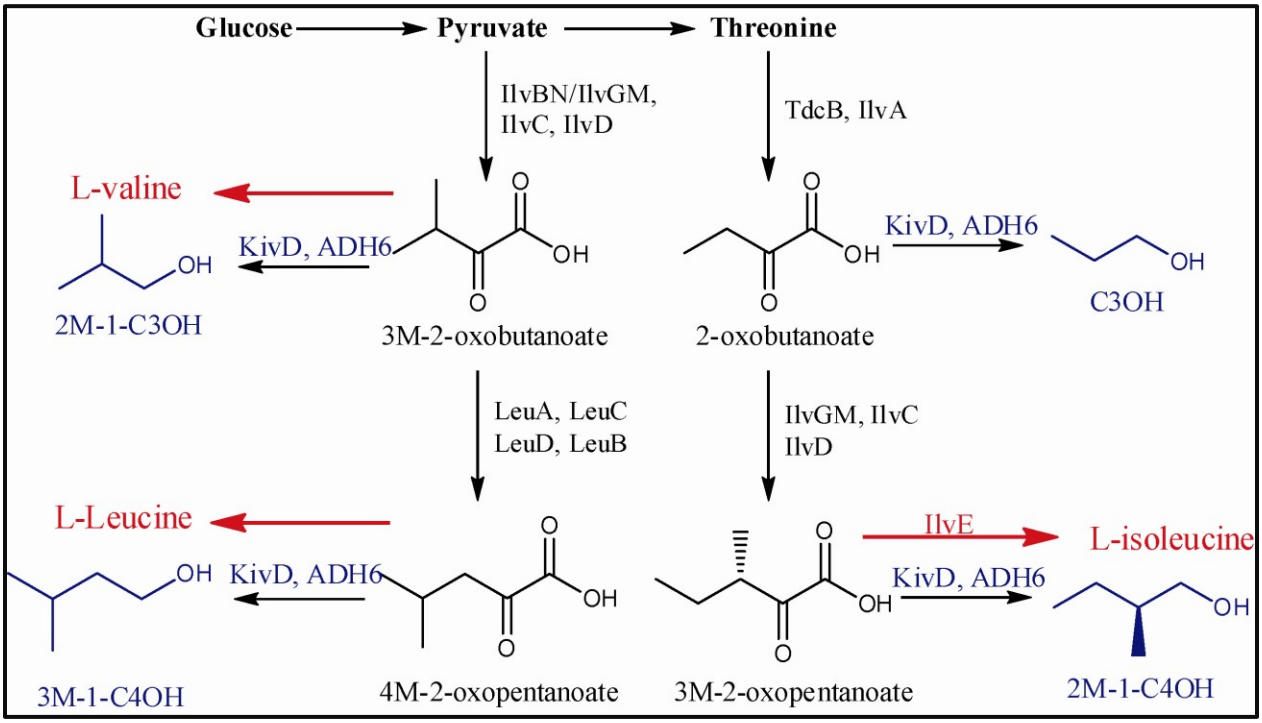


Figure 4.7: Production of mixed alcohols in wild type *E. coli*. Production of a wide variety of 2-keto acid-derived alcohols has been demonstrated in engineered *E. coli*[218,266]. Native *E. coli*, however, is more limited in this capacity. The major 2-keto acid-derived alcohols (blue) produced in wild type *E. coli* are rooted in high flux amino acid biosynthetic pathways (red). Production of all alcohols in this system can be traced to pyruvate and threonine.

A single plasmid, pBUT#52, was constructed for overexpression of *L. lactis* KivD and *S. cerevisiae* ADH6 using a P_{TRC} promoter (P_{TRC}:KivD.ADH6). When transformed into an *E. coli* AdhE knockout strain, high titers of 2-methyl-1-propanol and 3M-1-butanol were observed, and a trace of 1-butanol was measured (**Figure 4.8**). Production was tested in M9 minimal medium containing an autoinduction carbon source[267], and in undefined autoinduction medium; production from pBUT#52 was also measured when co-transformed with the biosensor plasmid pBMO#41 to determine if concomitant alcohol production and biosensor activity decreased host fitness.

M9 minimal medium supplemented with an autoinduction carbon source proved superior to undefined, autoinduction medium; the absence of amino acid supplementation in the minimal medium formulation likely increases carbon flux through the 2-keto acid intermediates in amino acid biosynthesis. Total mixed alcohol titers were 265±59 mg/L (n=3, mean±s.d.) after 48 hours growth, on par or superior to titers realized through overexpression of the heterologous *C. acetobutylicum* pathway in *E. coli*. As expected, 1-butanol was a minor product in this system, comprising less than 1% of all alcohols measured. Most importantly, production proved to be highly robust from experiment to experiment, and no statistically significant (t-test, $p>0.05$) change in alcohol titers was observed during concomitant alcohol production and sensing –

although alcohol titer variability did increase somewhat. Interestingly, 2-methyl-1-butanol was not observed in our production host, which may be due to either co-elution with 3-methyl-1-butanol or absence of production in our host strain. As discussed below, this detail did not alter the assay methodology.

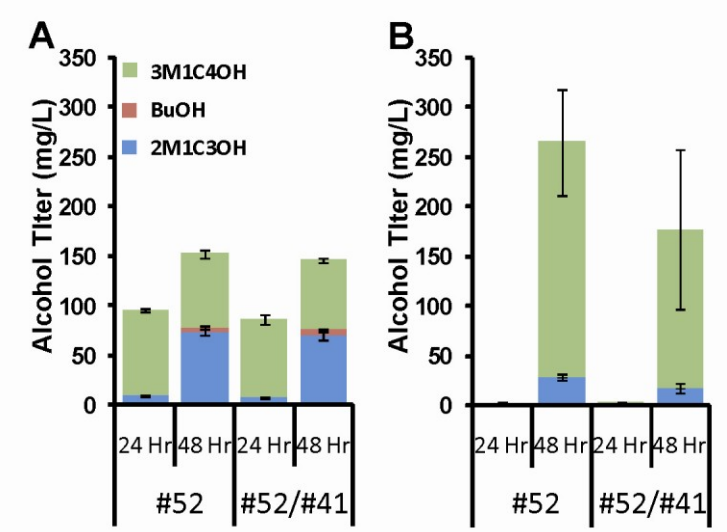


Figure 4.8: 2-keto acid-derived alcohol production from plasmid pBMO#52. High-titer alcohol production was observed in autoinduction (**A**), and M9 minimal medium supplemented with an autoinduction carbon source (**B**). A defined, minimal medium proved superior to an undefined autoinduction medium at the 48 hour timepoint. Co-transformation of the alcohol production plasmid (#52) with the biosensor plasmid (#41) did not have a statistically significant effect on total alcohol titers (t-test, $p>0.05$). n=3, mean±s.d.

Mixed alcohol production is not highly amenable to our high-throughput screening or selection strategy. Strain-to-strain and culture-to-culture differences in the relative ratios of the alcohols could skew results, particularly in light of variation in the biosensor's responsiveness 2-methyl-1-propanol, 1-butanol, and 3-methyl-1-butanol. Atsumi et al. demonstrated that a specific target alcohol could become a major product from their system if the *E. coli* growth medium was supplemented with the corresponding 2-keto acid[218]. We hypothesized a strain auxotrophic for the 2-keto acids leading to the observed alcohols could be utilized to produce single alcohols when cultured in minimal medium supplemented with user-defined 2-keto acid substrates. All 2-keto acid-derived alcohols measured in the wild type *E. coli* host are derived from pyruvate and threonine (**Figure 4.7**); by knocking out *tdcB*, encoding for threonine dehydratase, and the *ilvDAYC* isoleucine biosynthesis operon, a 2-keto acid auxotroph can be constructed (**Figure 4.9**).

A DH1 (Δ*adhE*) background strain transformed with pBMO#52 grew on minimal medium and produced three products co-eluting with authentic standards of 2M-1-C3OH, C4OH, and 3M-1-C4OH. Deletion of TdcB did not affect cell growth in M9 minimal medium. Deletion of ilvDAYC eliminated growth in M9 minimal medium, but was rescued by additional supplementation with valine, isoleucine and leucine. Addition of over 4.5 mg/L exogenously added amino acid supplementation, however, yielded trace (<5 mg/L) 2-keto acid-derived alcohol products. Presumably alcohols production proceeded through amino acid degradation, although the presence of trace 2-keto acids in the amino acid supplement cannot be ruled out. Unexpectedly, the single Δ*ilvDAYC* knockout was sufficient to eliminate production of our target

alcohols; because *tdcB* is predominantly expressed in stationary phase and in the presence of excess threonine, transcription may not be activated under our experimental conditions.

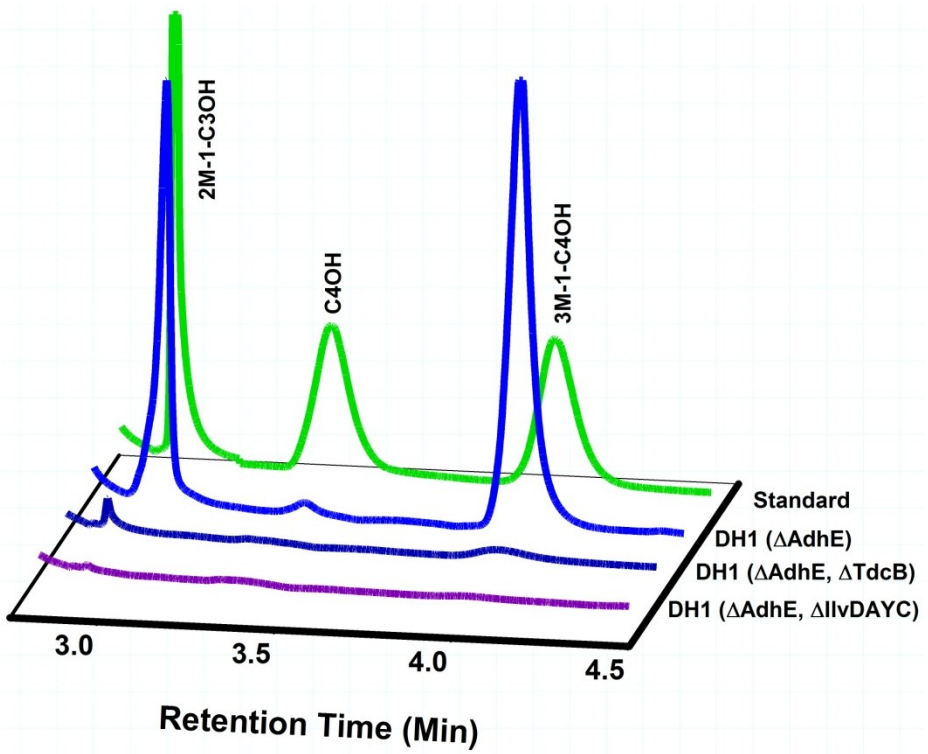


Figure 4.9: Construction of an *E. coli* 2-keto acid auxotroph. Introduction of either a Δ*tdcB* knockout or a Δ*ilvDAYC* knockout in an *E. coli* Δ*adhE* background strain eliminated 2-keto acid-derived alcohol production. Abbreviations: 2-methyl-1-propanol (2M-1-C3OH), 1-butanol (C4OH), 3-methyl-1-butanol (3M-1-C4OH).

The DH1 (Δ*adhE*, Δ*ilvDAYC*) mutant proved highly robust. Following autoinduction of both KivD and ADH6, near 100% conversion of 1 g/L exogenously added 2-oxopentanoate or 4-methyl-2-oxopentanoate to 1-butanol and 3-methyl-1-butanol, respectively, was observed in under 12 hours. This strain serves as a simple, robust system for user-defined production of 2-keto acid-derived alcohols.

4.2 High-throughput liquid culture screening

We first sought proof-of-principle demonstration of our alcohol-responsive biosensor as a high-throughput liquid culture screen. As discussed previously, this assay format is relatively straight-forward since production and detection are performed in two independent steps. Each step is optimized individually, and only alcohol containing medium is passed between the two strains. Because alcohol detection is done in a separate host, this assay format has the additional advantage of being production strain agnostic.

First, we implemented the 1-butanol production plasmid pBUT#50 (housing the operons: P_{TRC}:*crt.bcd.etfBA.hbd* and P_{TRC}:*atoB.adhE2*) in conjunction with the reporter plasmid pBMO#36 (housing the operons: P_{BMOR}:*bmoR* and P_{BMO}:GFP). Using the optimized 96 deep-well production system (see ***4.1.1 Expression of C. acetobutylicum 1-butanol biosynthetic pathway in engineered E. coli***), strains DH1 and DH1 (Δ*adhE*) were grown micro-aerobically in undefined autoinduction medium (**Figure 4.10A**). Over multiple production assays a number of samples exhibiting a range of 1-butanol concentrations (≈50μM to 550μM) were obtained; the spent medium was analyzed by both the biosensor strain and by gas chromatography-mass spectrometry (**Figure 4.10B**).

When pBUT#50 was grown in our final 96 deep-well plate format, optimized for decreased biological and technical replicate variability, there was no statistically significant (t-test, $p>0.05$) difference in 1-butanol titers between wild type DH1 and the DH1 (Δ*adhE*). This result strongly suggested it would be difficult to screen a library of genetic mutants in 96 deep-well plate format and identify those with improved 1-butanol productivities. Furthermore, the 1-butanol concentrations were near the lower end of the biosensor linear range of detection (**Figure 3.14**). When the resulting spent medium was analyzed using pBMO#36, the GFP-based biosensor, only those samples possessing 1-butanol concentrations less than 50 μM could be distinguished from those with higher 1-butanol concentrations.

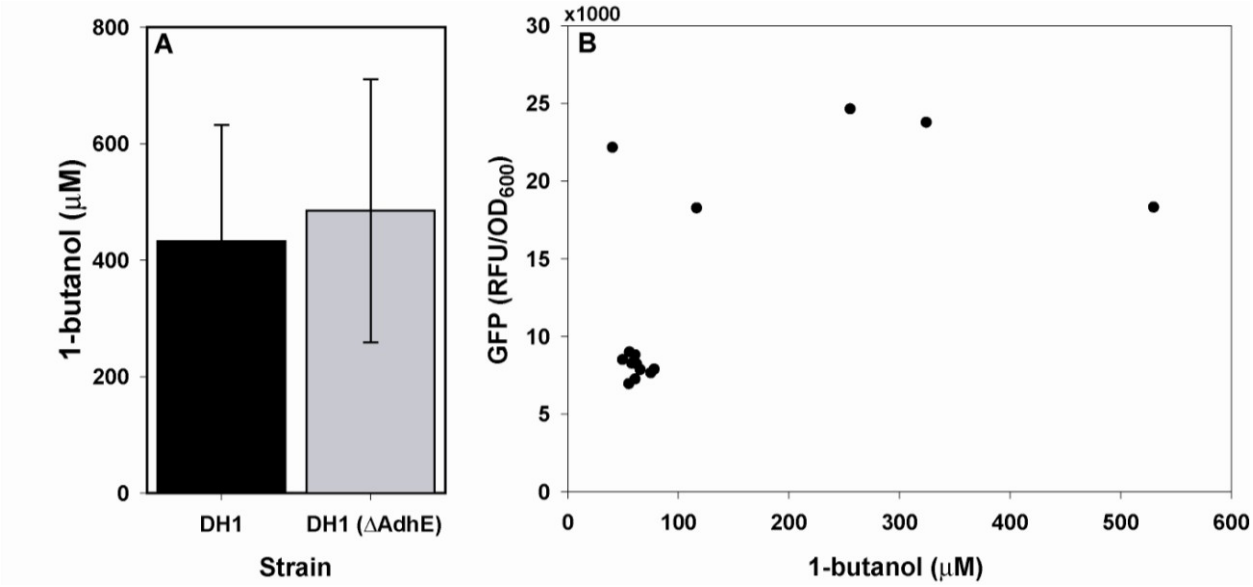


Figure 4.10: pBUT#50-based production and pBMO#36-based detection of 1-butanol. (**A**) 1-butanol production from DH1 and DH1 (Δ*adhE*) grown in 96 deep-well plates under autoinducing, micro-aerobic conditions. (**B**) GFP response from pBMO#36 following induction by pBUT#50-derived spent medium containing different concentrations 1-butanol.

The more straight-forward, accurate approach utilizes the pBUT#52-based 2-keto acid pathway; here the 1-butanol concentration can be readily altered by adjusting the concentration of 2-oxopentanoate supplementation in the growth medium. We first characterized the growth advantage 1-butanol concentration imparted to engineered *E. coli* harboring pBMO#41 following selection with tetracycline (**Figure 4.11**).

A combined log-logistic mathematical model[274] was used to fit the biphasic dose-response curves and derive values for the concentration of 1-butanol resulting in maximal ($C4OH_{max}$) and half-maximal (IC_{50}) *E. coli* growth.

$$OD_{600} = \omega + \frac{\alpha - \omega + \frac{OD_{max} - \alpha}{1 + \left(IC_{50}^{Tet} / [C4OH]\right)^{\beta_{Tet}}}}{1 + \left(IC_{50}^{C4OH} / [C4OH]\right)^{\beta_{C4OH}}} \qquad \textit{eq 4.1}$$

Here, the OD_{600} at a given 1-butanol concentration ([C4OH]) is described by the parameters α and ω, the horizontal asymptotes as the butanol concentration approaches 0 and positive infinity, respectively. Additional parameters include the slopes of the rising (β_{Tet}) and falling (β_{C4OH}) sides of the biphasic relationship as well as the half-maximal response due to tetracycline-induced (IC_{50}^{Tet}) and butanol-induced (IC_{50}^{C4OH}) toxicities.

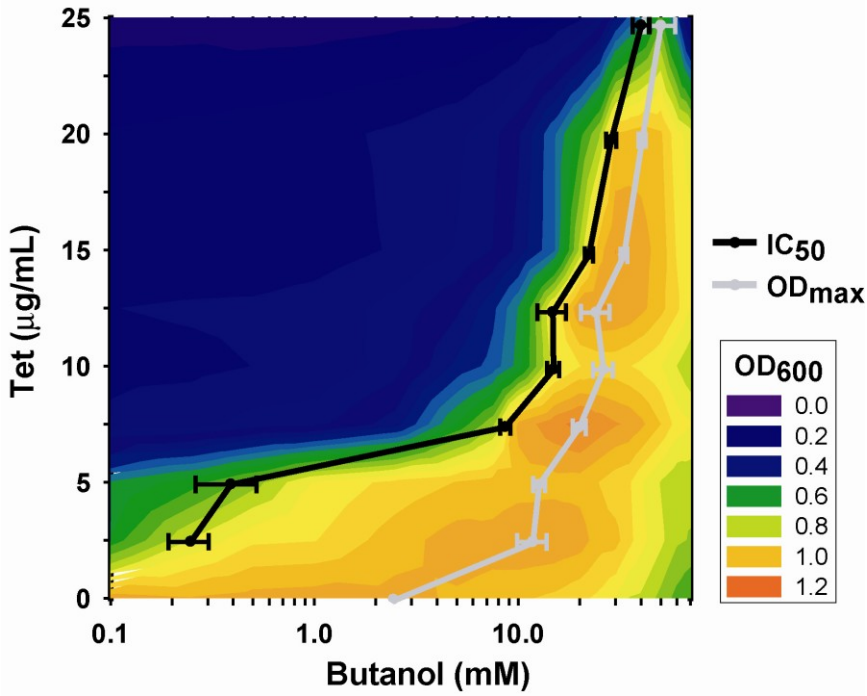


Figure 4.11: 96-well plate liquid culture screen. A 96-well plate, liquid culture screen was characterized, showing an increasing in IC_{50} (black) and maximum cell density (grey) as selective pressure was increased by addition of tetracycline to the growth medium. IC_{50} and OD_{max} curves: n=4, mean±s.d.; OD_{600} heat plot is an average of 4 replicate cultures, %CV<10%.

The assay proved ideally suited to the range of 1-butanol concentrations from 2-oxopentanoate supplemented cultures, and the stringency of the selection could be user-controlled by modulating the concentration tetracycline supplemented to the medium. The was controlled over a three order-of-magnitude range, between 0.240 ± 0.05 mM and 38.0 ± 3.5 mM 1-butanol (n=3, mean±s.d.), by increasing the tetracycline concentration in the culture broth. Similarly, the OD_{max} values increased from 11.9±2.0mM 1-butanol in the control culture lacking tetracycline to 50.5±8.7mM 1-butanol under 25μg/mL tetracycline selective pressure. Screening stringency – defined here as the 1-butanol concentration difference between the OD_{max} and IC_{50} curves – was observed to increase dramatically, as measured by convergence of the two curves with increasing tetracycline concentration. Additional control over the selection was realized by altering the time between addition of 1-butanol and tetracycline to the culture medium as well as the total incubation time (**Appendix 1, Figure A1.4**).

We next applied our liquid culture screening strategy toward endogenously produced 1-butanol. To generate a population with diversity in alcohol productivity and titer we mutated the *kivD* and *ADH6* ribosome binding site sequences on pBUT#52. When transformed into DH1 (Δ*adhE*) total alcohol titers – including 2M-1-propanol, 1-butanol and 3M-1-butanol – ranged between 32 and 713 mg/L (450±178 mg/L; median of 497 mg/L); comparatively, mixed alcohol titers using the wild type pBMO#52 plasmid ranged between 120 and 320 mg/L (227±54 mg/L; median of 233 mg/L; **Figure 4.12A**). The increase in the median 1-butanol titer between the wild type and library populations indicate the initial synthetic RBS designs for our pathway were non-optimal.

The pBMO#52 RBS library was transformed into DH1 (Δ*ilvDAYC*) for proof-of-principle demonstration of the genetic screen. Briefly 960 colonies were grown in 96 deep-well plates in autoinduction medium. Upon reaching stationary phase, 1 g/L of 2-oxopentanoate was added per well and cultured for 6 hours. The supernatant was then assayed for 1-butanol concentration using the biosensor strain harboring plasmid pBUT#41 placed under selective pressure with 7.5 µg/mL tetracycline (**Figure 4.12B**).

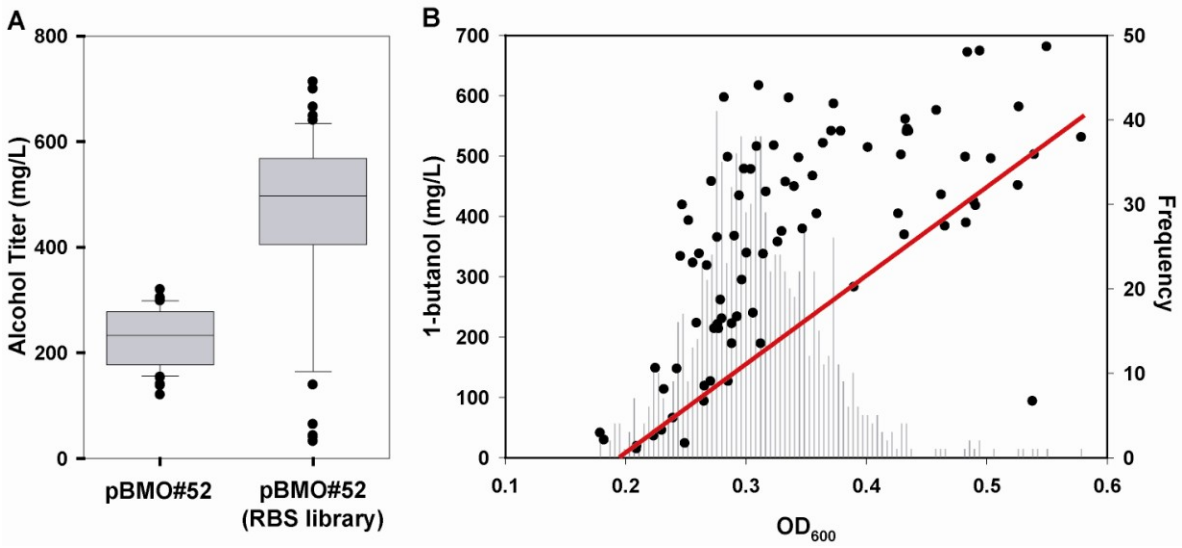


Figure 4.12: Proof-of-principle demonstration of biosensor-based liquid culture screen. (**A**) Total mixed alcohol titer resulting from plasmid pBMO#52 expression in DH1 (Δ*adhE*) was significantly lower (t-test, $p<0.05$) as compared to a heterogeneous population containing mutated *kivD* and *ADH6* ribosome binding site (RBS) sequences. Member of the pBMO#52 (RBS library) population produced a broad range of alcohol titers (n=50; box and whisker plot depicts 10th, 25th, median, 75th and 90th percentiles). (**B**) The biosensor response (x-axis, OD_{600}) to a 960-member library of mutated *kivD* and *ADH6* RBS sequences was normally distributed around OD_{600}=0.31 with a right-hand tail. GC-MS was used to confirm 1-butanol titers for 10% of the sample population; a strict threshold (red line; y=1466.5x - 283.01) described the lower limit of 1-butanol concentration and biosensor output OD_{600} value.

Biosensor output was normally distributed around OD_{600}=0.31, and thirteen samples (1.35% of the library population) exhibited a z-score greater than three. The average 1-butanol concentration for this sample subset (493±156 mg/L) was significantly higher ($p<0.005$) than those samples exhibiting a z-score of ±1 (345±146 mg/L). Thus, the assay can accurately identify samples with high concentration 1-butanol.

There exists only a weak linear relationship (R^2= 0.41) between biosensor output and 1-butanol concentration. However, the assay was highly accurate when screening against samples possessing low concentration 1-butanol. For example, from the 10% of the population whose 1-butanol titers were confirmed by GC-MS, only one sample fell below a threshold cutoff (this sample exhibited an OD_{600} of 0.55, but only 100mg/L 1-butanol). This outlier may be explained by carryover of production strain cell material into the assay well, resulting in an artificial increase the biosensor output signal.

While the liquid culture assay is adept at screening against poor butanol producers, a large fraction of high 1-butanol producers yielded average OD_{600} values. One straight-forward

explanation is the difference in assay growth medium between the initial biosensor characterization (**Figure 3.14**) and the final library screening experiment. Characterization assays were performed using fresh, defined medium; by comparison, the library screening assay used spent medium potentially containing cellular byproducts, residual antibiotic and production strain cell material. For example, supernatant from a library mutant containing a weak *ADH6* RBS – resulting in toxic butanal accumulation – might inhibit biosensor growth. All things considered, using tetracycline and cell growth as a biosensor output is a straight-forward, tunable method, but the presence of unknown factors in the sample medium may increase the frequency of false negatives.

As indicated previously, the synthetic RBS sequences used in the original pBMO#52 design were non-optimal. The average mixed alcohol titer from pBMO#52 transformed *E. coli* fell below the 25th percentile of alcohol titers from the mutated RBS library (**Figure 4.12A**). The top five positive hits from the RBS library screening were sequenced and analyzed in greater detail. One sequence was wild type pBMO#52, two were identical mutant variants (6G8), and two were unique mutant variants (7A12 and S71). In general, mutations were strongly skewed towards the *ADH6* RBS, and yielded more consensus-like Shine-Dalgarno sequences. This may be an indication that poor ADH6 expression in the original plasmid design created an unintended bottleneck in the pathway. Mutant S71, in particular, was highly divergent (**Figure 4.13**), and also included an unexpected mutation proximal to the *lac* operator. This mutation may have arisen during PCR amplification of the vector backbone or circular polymerase extension cloning (CPEC)[275].

		KivD start codon
pBMO#52	**C C A A T A T A T A A T A A A A T A T G G A G G A A T G C G A T G**	
pBMO#52-S71	**C C A A T A T A T A A T A A A A T A T G G A G G A A A G C G A T G**	
		ADH6 start codon
pBMO#52	**T C C A C G A G T T A A G G A G A G G G G G T T C C A A T G**	
pBMO#52-S71	**T C C A C G A G T T A A G G C G A A G A G G T T C C A A T G**	

Figure 4.13: *kivD* and *ADH6* RBS sequence mutations in plasmid pBMO#52-S71. Mutations (yellow) were identified in both the kivD and ADH6 RBS sequences; the ADH6 mutations, in particular, formed a more consensus Shine-Dalgarno sequence.

To confirm the improvement in pathway expression, and thus alcohol productivity, we repeated the 2-keto acid feeding assay in DH1 (Δ*adhE*, Δ*ilvDAYC*). The 3-methyl-1-butanol and 1-butanol specific productivities were monitored by GC-MS following supplementation of the corresponding 2-keto acid (**Figure 4.14**). The results were compared to a negative control strain harboring a P_{TRC}:RFP device. The negative control strain failed to demonstrate conversion of either 2 keto acid substrate to the corresponding alcohol. Plasmids pBMO#52 and pBMO#52-S71 exhibited 3-methyl-1-butanol specific productivities of 33.8±1.7 mg/L/hr/OD and 40.3±3.5 mg/L/hr/OD, respectively (n=3; mean±s.d.). Comparatively, conversion of 2-oxopentanoate to 1-butanol occurred more slowly, at rates of 13.5±1.3 mg/L/hr/OD and 18.6±1.0 mg/L/hr/OD for pBMO#52 and pBMO#52–S71, respectively (n=3; mean±s.d.). Thus, our high-throughput liquid culture screen was able to identify a mutant variant with significantly improved specific alcohol productivity as compared to the original plasmid design (t-test, $p<0.05$).

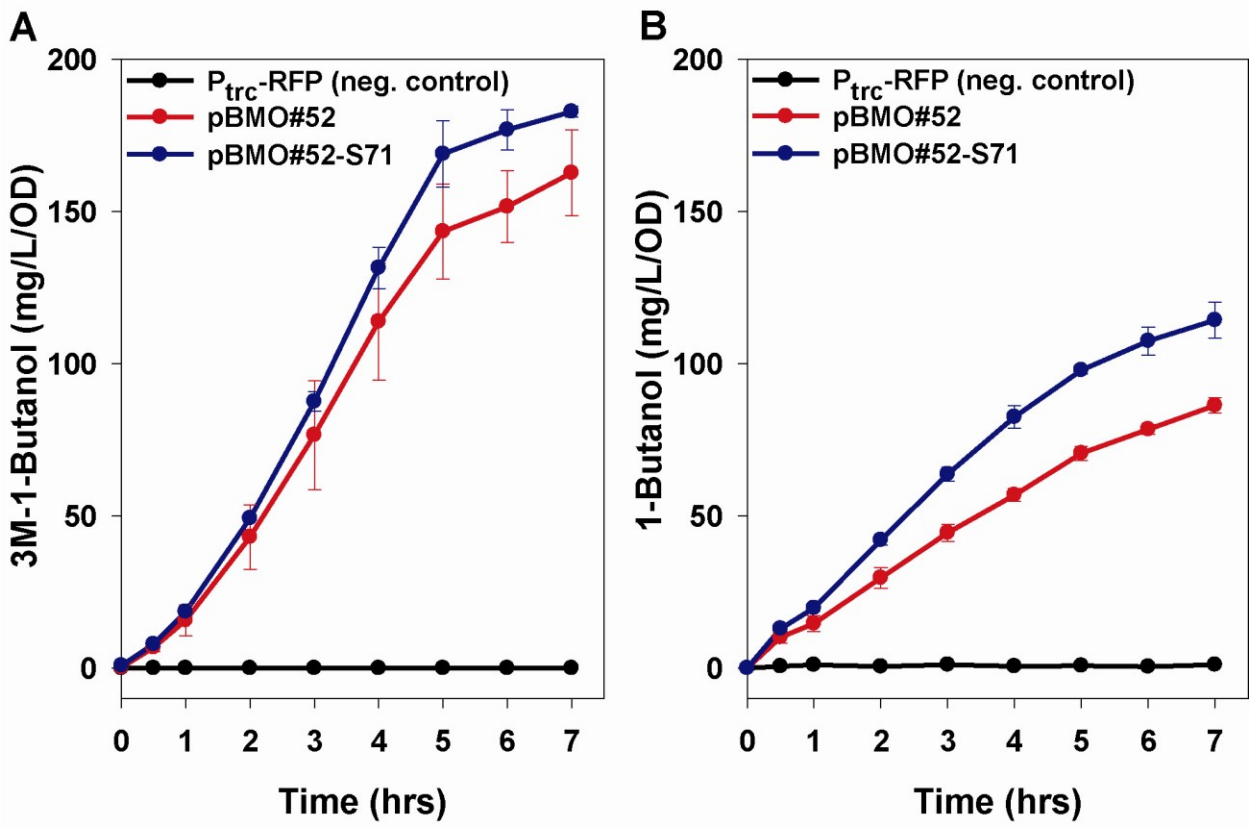


Figure 4.14: Confirmation of improved alcohol productivity in pBUT#52-S71. The rate of (**A**) 3-methyl-1-butanol (3M-1-butanol) and (**B**) 1-butanol formation was monitored by GC-MS in a DH1 (Δ*adhE*, Δ*ilvDAYC*) background strain grown in M9 minimal autoinduction medium supplemented with either 4-methyl-2-oxopentanoate or 2-oxopentanoate. Productivities were also obtained for the original pBUT#52 plasmid and a negative control strain harboring a P_{trc}:RFP device. n=3, mean±s.d.

There remains ample opportunity for future characterization and refinement of the 2-keto acid-derived alcohol biosynthesis pathway. A number of the positive hits identified in the RBS screen were not characterized further, and it remains possible that one or more of these variants is superior to pBUT#52-S71. Additional rounds of RBS mutagenesis can be performed using pBUT#52-S71 as a template, and potentially lead to further improvements in pathway flux. However, because the alcohol biosynthesis pathway is being optimized out of the context of the host metabolism (i.e. 2-keto acid substrates are being exogenously supplied), the same improvements in titer, productivity, and yield may not be witnessed when the plasmid is expressed in wild type *E. coli*. The endogenous supply of 2-keto acids will eventually limit alcohol production and require host strain engineering.

Lastly, a substrate promiscuous enzyme, producing a range of 2-keto acid-derived products is not ideally suited for metabolic engineering applications, where single pathways and products are often preferred. There remains abundant opportunity to use the screening method to construct several KivD mutants, with each variant exhibiting a high specificity toward a different 2-keto acid substrate. These variants could be identified through extensive active site mutation and screening.

4.3 Concomitant alcohol production and sensing

It would be ideal to obtain single-cell measurements of alcohol concentration using the biosensor in a high-throughput screen or selection. Both fluorescence activated-cell sorting (FACS) and selections are performed at the single-cell level, and thus production and detection will occur near simultaneously. For a FACS assay, the transcription factor-based biosensor output is a fluorescent protein; for a selection, the biosensor output is a protein imparting the host with either antibiotic resistance or improved specific growth rate (i.e. relieving an auxotrophy). A user-induced selection requires high temporal resolution; the target biosynthetic pathway must yield a high enough intracellular concentration of the desired compound to trigger sufficient expression of the selection marker that the host is conferred with a growth advantage. Initiating the selection prematurely results in rapid host death. Using either the heterologous *C. acetobutylicum* 1-butanol pathway or the 2-keto acid-derived alcohol pathway we sought to explore these concepts in greater detail.

4.3.1 FACS-based detection of 1-butanol production

Prior to constructing our 2-keto acid-based alcohol production strain, we attempted to use the *C. acetobutylicium* 1-butanol production pathway in a FACS screen for 1-butanol titers. pBUT#50 was co-transformed with pBMO#36 in a DH1 or DH1 (Δ*adhE*) background strain. Endogenously produced 1-butanol should induce GFP expression in these strains. We hypothesized that a DH1(Δ*adhE*) background – which produces higher 1-butanol titers (**Figure 4.3**) – would exhibit higher fluorescence.

Achieving concomitant alcohol production and detection proved highly difficult using the heterologously expressed *C. acetobutylicum* pathway. First, the *C. acetobutylicum* pathway performs best under micro-aerobic conditions, GFP, however, requires oxygen be present in order to fluoresce[276]. While fluorescence can be restored following re-introduction of oxygen into the system, this step imparts additional complexity to the screen. Of greater issue were obtaining positive control strains exhibiting different 1-butanol productivities while also combating plasmid instability. When expressed in the absence of the biosensor, the *C. acetobutylicum* pathway in DH1 and DH1 (Δ*adhE*) backgrounds produced 38±14 mg/L and 67±24 mg/L, respectively (n=3; mean±s.d.). Although we hypothesized that the biosensor would be sensitive to endogenously produced alcohol (as compared to the biosensor characterization experiments using exogenously added alcohol), it would be ideal if we possessed a set of strains exhibiting a range of 1-butanol productivities.

An initial flow cytometry characterization of pBMO#36 was completed using exogenously added 1-butanol as an inducer (**Figure 4.15**). In the absence of an alcohol inducer, *E. coli* harboring biosensor plasmid pBMO#36 exhibited increased GFP fluorescence as compared to background DH1. A bimodal population distribution was observed with a small fraction of the population possessing approximately 50-fold higher fluorescence as compared to the median fluorescence in the background strain, an indication of high-level, alcohol independent induction. The population distribution was also shifted slightly towards high fluorescence, an indication of leaky, basal transcription from P_{BMO}.

Exogenous addition of increasing concentrations of 1-butanol shifted the population distribution toward increased GFP fluorescence. In general, the distributions were wide, spanning a three order-of-magnitude range in the case of induction with 11 mM 1-butanol, and most of the distributions were also bimodal. The high level of heterogeneity in these results can be traced to several sources. Induced *E. coli* cultures were grown for a period of 12-16 hours – reaching stationary phase – before fluorescence was measured. Flow cytometry, however, is best performed on *E. coli* in exponential growth phase, and a much broader distribution in *E. coli* cell size and fluorescence signal is observed in during stationary phase. Second, the pBMO#36 biosensor plasmid uses a high copy number ColE1 origin of replication; high-copy number plasmids, in general, exhibit a broader distribution in protein expression levels. During 96-well plate measurement of GFP, in which a population averaged GFP fluorescence signal, the ColE1 origin resulted in an improved biosensor dynamic range, and provided high resolution dose-response curves. In contrast, a FACS-based application could benefit from a lower copy number plasmid. Unfortunately, pSC101 is the only other readily available origin (the alcohol production plasmids are housed on a p15A origin vector), and no detectable GFP signal was observed when using these constructs (data not shown).

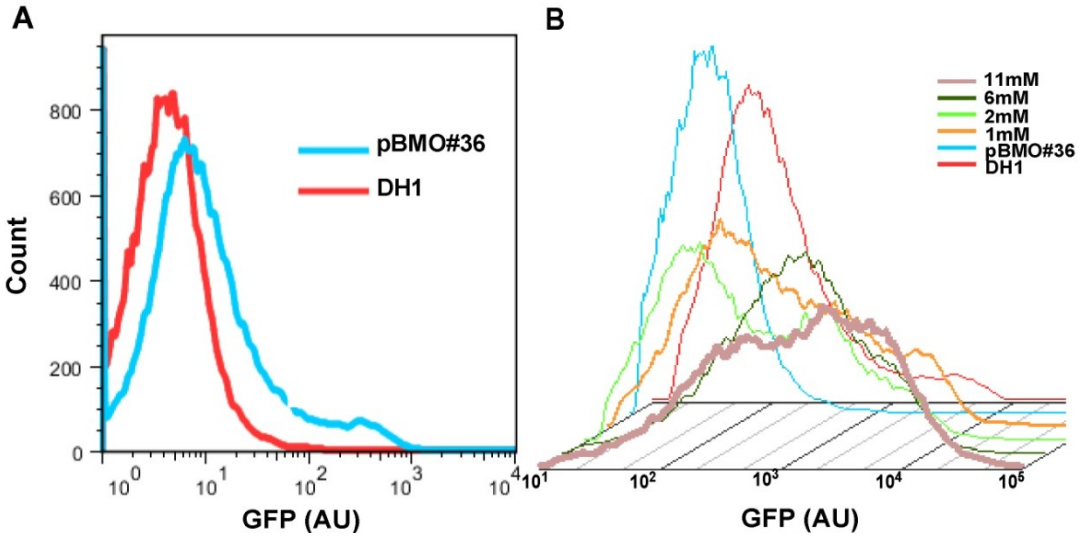


Figure 4.15: Flow cytometry characterization of pBMO#36 induced by exogenous 1-butanol addition. **(A)** Comparison of wild type E. coli with and with the biosensor plasmid pBMO#36, demonstrating leaky and alcohol-independent expression of GFP. (**B**) Representative population distributions for 1-butanol induced cultures (indicated as mM exogenously added 1-butanol; pBMO#36 and DH1 are uninduced controls) harboring biosensor plasmid pBMO#36.

Even with broad population distributions it remains possible to screen for variants with increased GFP fluorescence by FACS (**Figure 4.16**). By mixing two populations, exhibiting either high or low GFP fluorescence, we demonstrated how the FACS cutoff can be set so as to enrich for the only those members originally sourced from the cell population displaying high GFP expression.

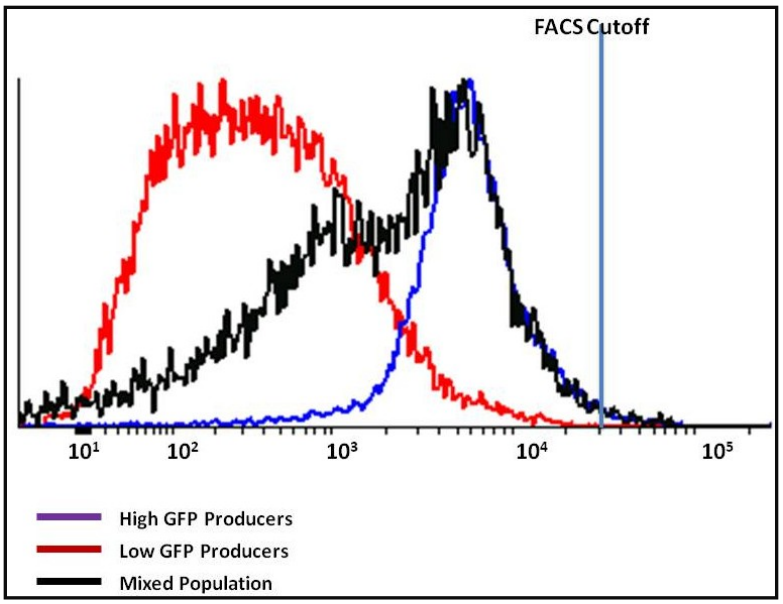


Figure 4.16: Visualization of FACS enrichment for high GFP expression. Populations exhibiting high (blue) and low (red) GFP expression were analyzed by flow cytometry; the populations were subsequently mixed (black) and reanalyzed. A FACS cutoff can be established to preferentially enrich for the high GFP producers. The two populations were not mixed in a one-to-one ratio, explaining the relative decrease in low-level GFP producers in the mixed population.

We next analyzed the concomitant production and detection of 1-butanol produced from a heterologously expressed *C. acetobutylicum* biosynthetic pathway. 1-butanol production plasmid pBUT#50 was co-transformed with biosensor plasmid pBMO#36 into both DH1 and DH1 (Δ*adhE*) background strains. Engineered strains were induced with IPTG and grown in micro-aerobic conditions for a period of 24 hours, at which point in time the 1-butanol concentration was measured by GC-MS and single-cell fluorescence measured by flow cytometry. 1-butanol titers from the two plasmid system dropped by over an order of magnitude as compared to an *E. coli* host harboring only the butanol production plasmid. In most samples, 1-butanol titers were under 10 mg/L. No statistically significant difference in 1-butanol titers was measured between a DH1 versus DH1 (Δ*adhE*) background under these conditions. Stability of the 1-butanol production plasmid was isolated as the primary problem in this system (**Figure 4.17**), and less than 5% of cells maintained the pBUT#50 plasmid after 24 hours growth. pBUT#50 was stably expressed, however, when expressed alone. In contrast, the reporter plasmid pBMO#36 was stable in the co-expression regime over the time course measured.

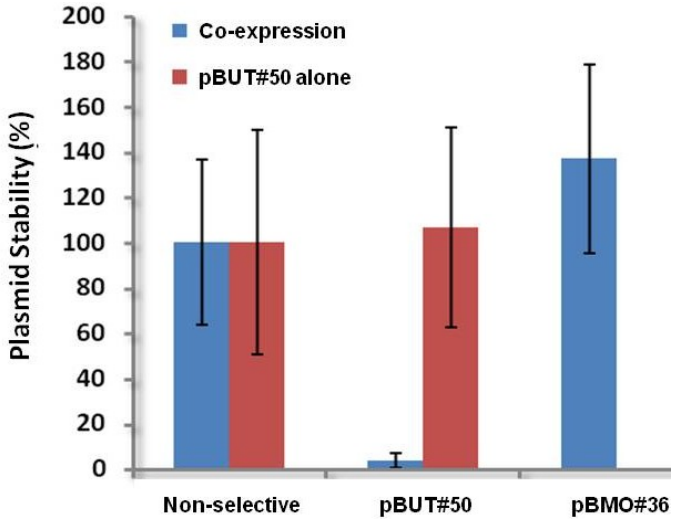


Figure 4.17: Plasmid stability during co-expression of pBUT#50 and pBMO#36. pBUT#50 plasmid stability was measured when expressed both alone and in conjunction with pBMO#36; pBMO#36 plasmid stability was only measured during the co-expression regime, and was stably maintained. pBUT#50 was stable when expressed alone, but near complete plasmid loss was witnessed after 24 hours upon co-expression with pBMO#36.

Plasmid instability greatly hindered the development of a FACS-based screening strategy for this strain. No differences in 1-butanol titers were observed between either strains or following induction with varying concentrations IPTG (data not shown). This effectively eliminated the positive control in the experimental setup and severely inhibited efforts to characterize the system by flow cytometry. From a directed evolution standpoint, plasmid instability limits the available mutation targets to only the host genome. Given our previous difficulty obtaining improved 1-butanol titers when replicating published reports (refer to **Section 4.1**), this route appeared unlikely to yield positive results.

As indicated, flow cytometry analysis and FACS testing were completed prior to construction of the 2-keto acid-based alcohol production pathway. Thus, there remains opportunity to complete proof-of-principle flow cytometry analyses of concomitant alcohol production and detection. First, the alcohol production plasmid pBUT#52 can be expressed stably with a reporter plasmid, and no statistically significant decrease ($p>0.05$) in alcohol titers was observed (**Figure 4.8**). Second, alcohol productivity can be readily altered by changing the 2-keto acid substrate concentration or by utilizing the previously constructed *KivD-ADH6* RBS library.

Based on our initial flow cytometry findings using exogenously added 1-butanol, a number of plasmid modifications can also be suggested to increase the assay resolution. For example, a relatively high level of *E. coli* auto-fluorescence noise was observed when using a GFPuv reporter. A red fluorescent protein (RFP) reporter is a superior alternative in this regard. A fast-folding, high quantum efficiency fluorescent reporter (i.e. visgreenGFP[277]) may also tighten the population distribution. Similarly, using a lower copy number origin of replication could reduce the background fluorescence levels resulting from leaky transcription from P_{BMO}.

4.3.2 Transcription factor-based selection for 1-butanol production

A biosensor-based selection offers the highest possible throughput, effectively scaling with the size of the population be assayed. The primary disadvantage is that selections only provide a

live-or-die output, necessitating a highly robust approach to ensure selection of the target phenotype. To date, selections have only been able to target phenotypes that are directly correlated with growth; improved biosynthesis of essential metabolites and increased resistance to harsher environmental conditions are standard examples. In these cases, the phenotype of interest is directly selected for.

In contrast, biosensor-based selections proceed through an intermediate step, and as such are indirect selections. The selectable phenotype, tetracycline resistance, is linked to the desired phenotype, 1-butanol biosynthesis, by the biosensor transfer function. The strength of the selection is dependent upon biosensor performance features. The linear range of induction, dynamic range, level of basal – or leaky – transcription and biosensor robustness will all impact the selection application space. While the BmoR-$_{PBMO}$ alcohol responsive biosensor exhibited a highly linear transfer function when 1-butanol was exogenously supplemented (**Figure 3.14**), a higher sensitivity is anticipated with intracellularly produced alcohol. The output of the biosensor, the TetA tetracycline transporter, also complicates the selection. Even low-level, background expression of TetA will result in some degree of tetracycline resistance. The resistance may increase if the rate of leaky expression is below the TetA protein half-life. In light of these arguments, biosensor-based selections are likely better suited for detecting novel activities; improvement to an existing activity is anticipated to be a much more difficult prospect. As discussed in greater detail in Chapter 5, additional biosensor engineering may be available to address some of these issues.

We sought to demonstrate a proof-of-principle selection strategy using a series of 2-keto acid-derived alcohols. Two negative control strains were designed for these experiments. A P_{TRC}:RFP plasmid was used to control for vector background. The second control replaces the *kivD* gene, encoding for the 2-keto acid decarboxylase, with *PDC*, encoding for *Zymomonas mobilis* pyruvate decarboxylase. When PDC is co-expressed with ADH6, the two enzyme pathway catalyzes the production of ethanol from pyruvate, but does not catalyze production of longer-chain 2-keto acid-derived alcohols[278]. Ethanol does not elicit a biosensor response below 1 mM (**Figure 3.14**). The resulting plasmid, pBUT#61 (containing a P_{TRC}:*PDC.ADH6* device on a p15A origin backbone), controls for the negative fitness observed from overexpressing proteins in the alcohol biosynthetic pathway. Both control strains were compared against pBUT#52 (containing the P_{TRC}:*kivD.ADH6*).

Control and experimental plasmids were co-transformed with biosensor plasmid pBMO#41 into an *E. coli* DH1 (Δ*adhE*) background, and tested for resistance to tetracycline after inducing alcohol production by addition of 2-oxopentanoate (**Figure 4.18**). Under non-selective conditions, *E. coli* harboring either the P_{TRC}:*RFP* and P_{TRC}:*PDC.ADH6* controls grew better than *E. coli* harboring the alcohol production plasmid housing a P_{TRC}:*kivD.ADH6* device. The P_{TRC}:*RFP* control strain also grew under selective conditions; in contrast, strains housing either the P_{TRC}:*PDC.ADH6* or P_{TRC}:k*ivD.ADH6* devices exhibited no growth under selective conditions.

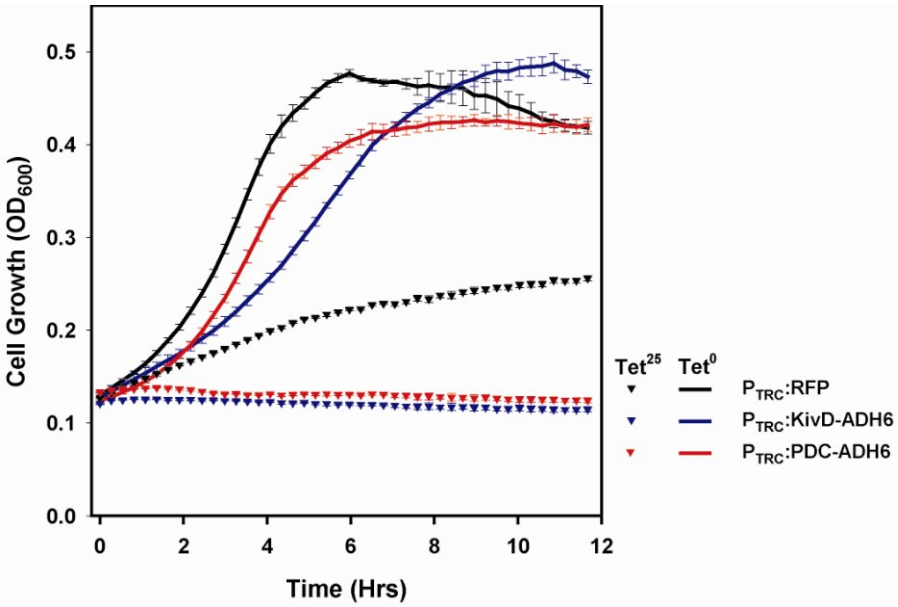


Figure 4.18: Biosensor-based selection for endogenously produced 1-butanol. The butanol selection plasmid pBMO#41 was co-transformed with pBUT#52 (P_{TRC}:*KivD.ADH6*), pBUT#61 (P_{TRC}:*PDC.ADH6*), or pBUT#63 (P_{TRC}:*RFP*). Strains were grown under either non-selective (no tetracycline, Tet^0) or selective (25 µg/mL tetracycline, Tet^{25}) conditions following induction of alcohol production and supplementation with the 1-butanol precursor 2-oxopentanoate. n=3, mean±s.d.

Under all conditions tested strains harboring the P_{TRC}–*RFP* negative control exhibited superior growth over strains harboring the KivD- and PDC-based alcohol production pathways. A simple mathematical framework describing the selective pressures in the system was used to help interpret these results. The cumulative host fitness, *F*, is described as the sum of numerous independent positive and negative selective pressures imparted by various components of our engineered system.

$$\mathbf{F} = f_{\mathbf{Tet^R}} - f_{\mathbf{2\text{-}ketoacid}} - f_{\mathbf{aldhyde}} - f_{\mathbf{alcohol}} - \sum_{x=1}^{n} f_{\mathbf{protein}}(x)$$

$$x = [\mathbf{TetA\text{-}GFP, RFP, BmoR, KivD, PDC, ADH6}]$$

eq 4.2

The only positive fitness in equation 4.2 is described f_{TetR}, the increase in tetracycline resistance due to expression of the TetA transporter. The other terms in equation 4.2 all assume negative fitness values and describe additional stresses imposed on the host cell by the alcohol production and detection devices. $f_{protein}$ describes stress due to PDC, KivD, ADH6, BmoR, and TetA-GFP overexpression. $f_{2\text{-ketoacid}}$, $f_{aldehyde}$ and $f_{alcohol}$ describe the negative fitness effects of small-molecule substrates, intermediates, and products found in the system.

Based on experimental results we developed a series of inequalities as to the importance and relative strengths of individual selective pressures. Growth is observed in strains expressing BmoR and TetA-GFP in the presence of up to 40 mM exogenously added 1-butanol and 25 µg/mL tetracycline (**Figure 4.11**). Similarly, growth inhibition was not observed in strains expressing BmoR and GFP when using up to approximately 10mM 1-butanol or 5mM butaldehdye (**Figure 3.15**). Endogenously produced aldehyde or alcohol, however, may be toxic

at a lower concentration than measured during exogenous addition assays. None-the-less, cultures supplemented with up to 1 g/L (≈8.6mM) 2-oxopentanoate showed both cell growth and full conversion of the 2-keto acid substrate to 1-butanol.

Thus, to a first approximation the $f_{protein}$ and f_{TetR} terms appear to be the dominant terms in equation 4.3, and the $f_{2\text{-ketoacid}}$, $f_{aldehyde}$ and $f_{alcohol}$ are subordinate. The results from the selection assay with a P_{TRC}:*PDC.ADH6* construct support our analysis of the relative importance of the various fitness factors. PDC is unable to catalyze the production of 1-butanol from 2-oxopentanoate, thus the butanal and 1-butanol stresses are replaced with acetaldehyde and ethanol. Both acetaldehyde and ethanol are considerably less toxic compared to their C4 counterparts, and are expected to be observed at a low concentration in our system. The common terms that describe both the PDC- and KivD-based experimental groups are the f_{TetR} and $f_{protein}$ terms, and the following inequalities are drawn:

$$0 < F_{P_{TRC}:GFP} = f_{Tet^R} - f_{protein}(\text{BmoR, TetA-GFP})$$

$$0 \geq F_{P_{TRC}:PDC\_ADH6} = f_{Tet^R} - f_{protein}(\text{PDC, ADH6, BmoR, TetA-GFP})$$

$$0 \geq F_{P_{TRC}:KivD\_ADH6} = f_{Tet^R} - f_{protein}(\text{KivD, ADH6, BmoR, TetA-GFP})$$

$$\left|f_{Tet^R}\right| < \left|f_{protein}(\text{PDC, ADH6, BmoR, TetA-GFP})\right| \leq \left|f_{protein}(\text{KivD, ADH6, BmoR, TetA-GFP})\right|$$

Future work developing a genetic selection based on the BmoR-P_{BMO} biosensor with a 2-keto acid-based alcohol biosynthetic pathway will need to address the relative magnitudes of the f_{TetR} and $f_{protein}$ positive and negative selective pressures, respectively.

Multiple experiments can be outlined to test the above hypotheses. First, the TetA-GFP fusion protein may lead to significant membrane stress as compared to TetA expression alone. The fluorescence signal from TetA-GFP is an order of magnitude lower than GFP alone, and is decreased further by addition of tetracycline to the growth medium. These results indicate GFP is misfolding when expressed as a fusion protein. Expression of TetA as a single protein product should help address this issue; and if a dual screen-selection is desired, the *gfp* gene can be housed downstream of *tetA*. Even when expressed alone, however, TetA may not be the ideal reporter for a selection device, and non-membrane associated resistance gene may be a superior option. TetA overexpression has been shown to result in a decrease in cell membrane potential, leading to decreased cellular fitness or death[279]. Decreasing the biosensor copy number or the strength of the TetA ribosome binding site could mitigate this deleterious phenotype.

Increasing f_{TetR} addresses only part of the equation; the negative fitness value of $f_{protein}$ must also be addressed. Fortunately, there exist a number of straight-forward strategies to improve protein solubility. The assay temperature can be decreased from 30°C to 25°C, which is supported by our analysis of P_{BAD}-based BmoR expression (**Figure 3.5**). At the vector level, the plasmid origin of replication and promoter can be altered. Changing from a p15A to pSC101 origin will result in an order-of-magnitude lower DNA transcript in the host, and changing from a P_{TRC} to P_{LAC01} or P_{LACUV5} promoter will provide more user-defined control over protein expression levels. While decreasing the concentration of alcohol biosynthesis enzymes in the host cell will

decrease 1-butanol productivity, introduction of the selective pressure can be delayed to enable 1-butanol-induced expression of TetA. Lastly, expression of *E. coli* chaperon proteins, GroEL and GroES, following induction of the alcohol pathway proteins may improve protein solubility and increase cellular fitness.

4.4 Materials and methods

4.4.1 Reagents

All enzymes and chemicals were purchased from Fermentas and Sigma-Aldrich Co., respectively, unless otherwise indicated. DNA oligomers were ordered from Integrated DNA Technologies (Coralville, IA).

4.4.2 Strains and Plasmids

Genbank files for all plasmids are included in **Appendix 2**. All plasmids were assembled from PCR product by sequence ligation independent cloning (SLIC)[263], unless otherwise indicated, using Phusion DNA polymerase (New England Biolabs). The *Clostridium acetobutylicum* (ATCC# 824) and *Pseudomonas butanovora* (ATCC# 43655) genes and promoters were cloned from genomic DNA. The *L. lactis kivD* gene and the *S. cerevisiae* ADH6 gene were synthesized (DNA 2.0). The *pdc* gene was amplified from plasmid pKS13 described elsewhere[280]. Mutagenesis of *kivD* and *ADH6* ribosome binding sites was performed by amplification of *kivD* from pBUT#52 using primers DC133 (5'- TAACAATTAGATCTCCAATATATAATAAAA $N_4N_1N_4N_3N_3N_1N_3N_3N_1N_1N_4$GCGATGTATACAGTAGGAG-3') and DC134 (5'- CTTCAAATT TCTCAGGATAAGACATTGGAAC$N_2N_4N_2N_2N_4N_2N_4N_2N_2N_4N_4N_1N_1N_2$TCGTGGATTATGA TT-3'). N1:[A_{96},C_2,G_2,T_2], N2:[A_2,C_{96},G_2,T_2], N3:[A_2,C_2,G_{96},T_2], N4:[A_2,C_2,G_2,T_{96}] where the subscript for each deoxyribonucleotide denotes percentage of each base included in final population. The resulting PCR product was cloned into the pBUT#52 vector by circular polymerase extension cloning (CPEC)[275].

Escherichia coli strain DH10b was used for all molecular cloning; all engineered *E. coli* strains were based on a DH1 or MG1655 background, as indicated. Deletion of alcohol dehydrogenase (Δ*adhE*), lactate dehydrogenase (Δ*ldhA*), threonine dehydratase (Δ*tdcB*), and the isoleucine biosynthetic operons (Δ*ilvDAYC*) were achieved by λ Red-mediated gene deletion[264]. *E. coli* Strain β was constructed in an MG1655 (Δ*adhE*, Δ*ldhA*) background; using the pBUT#50 plasmid as a template the 1-butanol biosynthetic operon was amplified along with a chloramphenicol resistance *cat* gene using PCR primers containing 30-bp homology to the *E. coli intA* gene. The cassette was inserted into the *intA* gene locus by λ Red-mediated homologous recombination.

Table 4.2: Plasmids used in this study

Plasmid	Description	Source
pBUT#50	P_{TRC}:*crt.bcd.etfBA.hbd*, P_{TRC}:*atoB.adhE2*, Cm^R, p15a	This study
pBUT#52	P_{TRC}:*kivD.ADH6*, Cm^R, p15a	This study
pBUT#61	P_{TRC}:*PDC.ADH6*, Cm^R, p15a	This study
pBMO#36	P_{BmoR}:*bmoR*, P_{BMO}:*gfp*, Amp^R, ColE1	This study
pBMO#41	P_{BmoR}:*bmoR*, P_{BMO}:*tetA*, Amp^R, ColE1	This study

4.4.3 Metabolite quantification

1-butanol, 3-methyl-1-butanol and 2-methyl-1-propanol were extracted from cell cultures with ethyl acetate. Equal volumes of culture broth and ethyl acetate (containing 0.01 % v/v 1-hexanol as internal standard) were vortexed for 5 min. The ethyl acetate was recovered and applied to a

gas chromatograph (Focus GC, Thermo Scientific) equipped with autosampler (TriPlus, Thermo Scientific), TR-wax column (30 m x 0.25 mm x 0.25 µm;Thermo Scientific), and flame ionization detector (Agilent). The samples were run on the GC with the following program: initial temperature, 40°C for 2 min, ramped to 120°C at 15°C/min. Authentic standards were prepared by titrating each alcohol into a 1 mL aqueous solution, vortexed for 5 min (creating a 10 g/L stock solution) and serially diluted to the working concentration. Standards were than extracted with ethyl acetate as described previously.

Quantification of succinate, formate, ethanol, acetate, and 1-butanol was accomplished by liquid chromatography-mass spectrometry (LC-MS). 1 mL cultures were filtered and separated on a Zorbax 300SB-C1 8 column (Agilent; 2.1 mm i.d. × 10 cm length) using an Agilent 1100 series HPLC at a flow rate of 200 µL/min and 50°C running temperature. Samples were run in a 4 mM H_2SO4 buffer for 45 min. The LC system was interfaced to a refractive index detector (1200 Series, Agilent Technologies).

4.4.3 Biosensor reporter quantification

Both fluorescence and absorbance measurements were performed on dual spectrophotometer-fluorometer (Spectromax M2, Molecular Devices). GFP fluorescence was measured using an excitation wavelength of 400 nm and an emission wavelength of 510 nm. Optical density measurements were monitored at 600 nm (OD_{600}). GFP fluorescence values were first normalized to OD_{600} (GFP/OD_{600}). *E. coli* auto-fluorescence was subtracted using a standard curve of GFP fluorescence from wild type *E. coli* optical density. Fold-induction was calculated as the difference between the averages of the induced and un-induced GFP fluorescence measurements normalized to the un-induced GFP measurement.

Single cell fluorescence measurements were performed on a flow cytometer (FACSAria II, BD Biosciences) equipped with an argon laser (emission at 488 nm/20 mW), a 530/30nm bandpass filter, and 70µm nozzle. 1 mL cell culture was first centrifuged (6000 x g) and washed with phosphate buffered saline (PBS) at pH 7.4. Cultures were diluted 50-fold prior to flow cytometry analysis. Data was analyzed

4.4.4 Production of 1-butanol and 2-keto acid-derived alcohols

Colonies of engineered *E. coli* harboring plasmids pBUT#50 or pBUT#52 were inoculated into 5-mL Luria-Bertani (LB) medium supplemented with glucose (2% v/v) and chloramphenicol (50 µg/mL, Cm^{50}) and grown overnight (200 rpm, 37°C). Strains were the sub-cultured (1% v/v) into fresh medium (Cm^{50}); mediums included terrific broth (TB) supplemented with glycerol (2% v/v), LB medium, defined rich medium[270] supplemented with glucose (2% v/v), M9 minimal medium supplemented with glucose (2% w/v), or autoinduction medium[267]. For M9 minimal medium autoinduction experiments glucose was replaced with 2x autoinduction sugars (1.0% w/v glycerol, 0.1% w/v glucose, 0.4% w/v lactose) was used. When required, pathway induction was achieved by addition of isopropylthiogalactoside (IPTG) at an OD_{600} absorbance of 0.25. All production experiments were carried out at 30°C.

1-butanol production experiments were conducted in 16 mm x 125 mm culture tubes with PTFE-faced rubber-lined caps (Kimble Chase). During anaerobic growth experiments tubes were filled

such that no headspace was present; during micro-aerobic and aerobic growth experiments cultures were prepared in 10 mL medium, with the caps sealed for micro-aerobic samples.

96 deep-well plate assays were performed similarly to those described above. 0.6 mL medium was used per well, and plates were sealed with adhesive PCR film using an automated adhesive sealer (Seal-it 100, Thermo Scientific) prior to incubation (30°C, 250 rpm).

4.4.5 Liquid culture screening assay

E. coli strain DH1 (Δ*adhE*, Δ*ilvDAYC*) harboring the alcohol production plasmid pBUT#52 was cultured overnight in LB medium supplemented with 0.5% w/v glucose (Cm^{50}, 200 rpm, 30°C). Cultures were then inoculated 1% v/v into 0.6-mL fresh M9 minimal medium in 96 deep-well plates (2 mL total capacity, polypropylene, square-bottomed; Corning); M9 medium was supplemented with antibiotic (Cm^{50}) and 2x autoinduction sugars (1.0% w/v glycerol, 0.1% w/v glucose, 0.4% w/v lactose). Cultures were grown for 24 hours (30°C, 300rpm), centrifuged (x3000g, 4 min), and resuspended into fresh EZ-rich medium (Teknova) supplemented with 0.5% w/v glucose, antibiotic (Cb^{50}), and 1 g/L 2-oxopentanoate. After six hours incubation (30°C, 300 rpm), plates were centrifuged (3000 x g, 4 min) and the supernatant collected for analysis.

E. coli strain DH1 (Δ*adhE*) harboring either biosensor plasmid pBMO#36 or pBMO#41 was cultured overnight in LB medium (Cb^{50}, 200 rpm, 30°C). Cultures were then inoculated 1% v/v into fresh EZ-rich medium (Teknova) supplemented with antibiotic (Cb^{50}), and grown until final cell densities reached an OD_{600}=0.20 (200 rpm, 30°C). For biosensor characterization experiments, biosensor culture was diluted 1:4 in fresh EZ-rich medium (0.5% w/v glucose, Cb^{50}) supplemented with a known concentration alcohol. When assaying 1-butanol concentrations in spent production medium (described in the preceding paragraph), 150 µL of biosensor culture was added to 150 µL spent production medium and 300 µL 2X EZ-rich medium (0.5% w/v glucose, Cb^{50}) in 96 deep-well plates (2 mL total capacity, polypropylene, square-bottomed; Corning). For pBMO#41 selections, assay samples were incubated for 0 – 2 hrs, supplemented with either nickel chloride or tetracycline, and grown for an additional 16 hrs (200 rpm, 30°C). Fluorescence and cell density were measured as described above.

A mathematical model based on a combined log-logistic function was used to describe the biphasic butanol-response curves observed[274]:

$$OD_{600} = \omega + \frac{\alpha - \omega + \frac{OD_{max} - \alpha}{1 + \left(IC_{50}^{Up} / [C4OH] \right)^{\beta_{Up}}}}{1 + \left(IC_{50}^{Dn} / [C4OH] \right)^{\beta_{Dn}}}$$

Here OD_{600} at a given concentration of butanol ([C4OH]) is described by the parameters α and ω describing the horizontal asymptotes as the butanol concentration approaches 0 and positive infinity, respectively. Additional parameters include the slopes of the rising (β_{Up}) and falling (β_{Dn}) sides of the biphasic relationship as well as the half-maximal response due to tetracycline-induced (IC_{50}^{Up}) and butanol-induced (IC_{50}^{Dn}) toxicities.

Chapter 5. Conclusions and Future Directions

We began this study by postulating that a native transcription factor-promoter pair can be identified in Nature that is capable of detecting all industrially produced small-molecules. Supporting this statement is the theory that long-term, large-scale release of anthropogenic compounds into the environment results in selective pressure for the evolution of catabolic pathways to eliminate or utilize these compounds. Transcription factor-promoter pairs ensure the pathway is upregulated only when required.

While we have focused on the *P. butanovora* BmoR-P_{BMO} system for detection of 1-butanol and structurally similar terminal alcohols, a wide number of transcription factors responding to industrially important small-molecules have been reported in the literature. For example, the well-characterized XylR and XylS transcription factors from *Pseudomonas putida* regulate a toluene-xylene catabolic pathway, and both bind substrates and intermediates in this pathway[281]. Similarly, transcription factor-promoter pairs have been identified for detection of ε-caprolactam[282], succinate[283], adipate[284,285] and tetrahydrofuran[286], among others. The BmoR-P_{BMO} promoter system was used to demonstrate the potential application of transcription factor-promoter pairs as high-throughput screening and selection devices.

5.1 Biosensor construction and testing

Construction a functional biosensor in *E. coli* using the BmoR-P_{BMO} transcription factor-promoter pair proved to be a difficult task. As detailed in Chapter 3, BmoR heterologously expressed in an *E. coli* host was localized in the insoluble protein fraction when cultured at 37°C, and the biosensor produced no detectable fluorescent signal in the presence of 1-butanol. Decreasing the incubation temperature to between 25-30°C produced a functional biosensor, but the biosensor exhibited poor dynamic range and was non-robust. We anticipate a low expression temperature being a common requirement for biosensors constructed from transcription factors responding to industrial chemicals; the native hosts are often soil bacteria that have been selected for growth, and hence protein expression, at atmospheric temperatures. While decreasing the expression temperature improved the dynamic range, it did not impart the biosensor with robust behavior. Optimization of the promoter driving expression of BmoR, the GFP ribosome binding site, 5'-untranslated region and induction timing were all required before achieving reproducible behavior.

A generalized heuristics-based approach to obtaining functional heterologous protein expression in *E. coli* has been reported elsewhere[287], and we expand upon these rules as applied to biosensor design. The biosensor transfer function, describing the relationship between small-molecule input and reporter output, is strongly affected by biosensor component parts. For this reason, use of well-characterized genetic parts[288] – including, if possible, the transcription factor-promoter pair – can greatly simplify troubleshooting efforts. However, steps can also be taken to mitigate the failure rate when using uncharacterized parts; for example, optimizing the BmoR protein coding sequence for expression in *E. coli* may garner further improvement in biosensor performance features.

The reporter genes used in our constructs, either GFPuv[289] or TetA, strongly affected the background signal and biosensor dynamic range. The peak GFPuv excitation/emission wavelengths (395/508nm) also produce a strong auto-fluorescence signal in *E. coli*. The *E. coli* background signal from RFP (variant mCherry[290]; excitation/emission wavelengths at 587/610)

is comparatively weaker, and is a superior fluorescent reporter in biosensor applications. While mathematically decreasing the background fluorescence will have no effect on biosensor dynamic range, it does improve the ability to discriminate between positive and negative hits. In addition fluorescent protein choice, background noise from leaky promoter expression can be mitigated by decreasing the reporter mRNA or protein half-life, or the translation initiation rate.

In contrast to a fluorescent protein, the background signal using the TetA tetracycline-resistance reporter was eliminated at high tetracycline concentration (**Figure 4.11**). The disadvantage, however, is a decrease in linear range of induction and concomitant increase in biosensor sensitivity. Implementing a TetA-based biosensor as a liquid culture screen may require the analyte be concentrated (or diluted) to ensure the target small-molecule concentration falls within the narrowed linear range.

5.2 Biosensor Implementation

Application of a transcription factor-based biosensor as a liquid culture screen has the advantage of minimizing interaction between production and detection strains, and as such, each system can be optimized independently. However, there also exists an inverse relationship between system (production or detection) interaction and assay throughput. The higher the assay throughput the more intertwined production and detection must become. For example, in an outline of a plate-based screening approach, the biosensor strain is incorporated into a solid-medium (M9 medium-agar) and the production strain colonies are overlaid on top of the biosensor layer. Alcohol produced from individual colonies then diffuses into the surrounding solid medium and induces reporter expression in the biosensor strain. As compared to a liquid culture screen, in which production and detection are physically separated and there is no interaction between the two strains, a solid-medium screen requires communication between the production colony and the underlying biosensor strain.

In vivo, concomitant small-molecule production and detection greatly improves assay throughput, but requires a high-degree of coordination between the production and sensing functions; thus, it can be difficult to successfully implement. The production and detection devices may behave differently when co-expressed as compared to their individual behaviors. In the case of the heterologously expressed *C. acetobutylicum* 1-butanol pathway, a decrease in plasmid stability was observed upon co-expression with the biosensor device; instability was not observed, however, when the alcohol production pathway was expressed alone. Similarly, concomitant production and selection for 2-keto acid-derived alcohols resulted in an unanticipated increase in host stress that decreased host strain growth rates.

The case for implementing an *in vivo* biosensor remains strong. Two reports recently described successful concomitant small-molecule production and sensing. The first study screened 53,000 members of a metagenomic library by FACS or fluorescence microscopy to identify clones harboring a genetic cassettes encoding for production of acyl-homo-serine lactone that activate a LuxR-P_{LuxI} biosensor[291]. The second study used a mutant variant of the AraC-P_{BAD} biosensor system responsive to mevalonate to screen a pathway enzyme RBS library using a plate-based β-galactosidase screen[292]. Both studies demonstrate the technical feasibility of the approach, but pathway- and biosensor-specific idiosyncrasies had to be addressed in each case. Achieving a genetic selection, however, remains undemonstrated to date.

5.3 Future work

From this study, the most robust application of the BmoR-P_{BMO} based biosensor was as a liquid culture screen, and the methodologies developed here can be readily extrapolated toward other transcription factor-based biosensors. There remains ample opportunity for continued characterization of the BmoR-P_{BMO} promoter system and optimization of biosensor performance features. Dynamic range and sensitivity dictate how accurately an assay can discriminate between incremental increases in small-molecule concentration, and further improvements in biosensor sensitivity and dynamic range would enable more accurate analyte quantification.

The BmoR-P_{BMO} system can be used as a testbed to explore development of generalized methods for improving biosensor performance features. Depending on the desired application space, there may be need to either increase or decrease biosensor sensitivity. The BmoR-P_{BMO} biosensor exhibited a linear response to exogenously added alcohols (**Figure 3.15**). A greater than 10-fold dynamic range over a 10 μM – 40 mM linear response window was observed using a wild-type *E. coli* host, and a 700 μM concentration difference can be distinguished with 95% confidence over the linear response range. If dynamic range and sensitivity were increased, the biosensor could better discriminate between smaller increases in analyte concentration. Conversely, during direct sensing of endogenously produced alcohols the biosensor may be too sensitive, and more robust, accurate analyte detection may be achieved with a less sensitive system.

A number of control points can be targeted for engineering transcription factor (TF)-based biosensors, including TF-ligand binding site (protein-ligand) and activated TF-promoter binding (protein-operator). The TF-ligand binding constant ($K_D^{TF \cdot Lig}$) [293-295], TF-operator binding constant($K_D^{TF \cdot DNA}$)[296], operator architecture, and operator strength all contribute to the shape and location of the dose-response curve with respect to input ligand (**Figure 5.1**). Theoretical[296-298] and experimental[296] evidence support the use of multiple, cooperative operator sites to improve ligand sensitivities by up to 4-fold. Dynamic ranges can increase multiplicatively with the number of added operator sites (assuming reporter saturation does not occur)[299]. Adding additional operator sites is referred to herein as altering biosensor architecture. A complimentary approach to altering biosensor architecture is to engineer a single operator for an altered ($K_D^{TF \cdot DNA}$) compared to the wild-type system; a weaker TF binding results in decreased ligand sensitivity and vice versa for stronger TF binding.

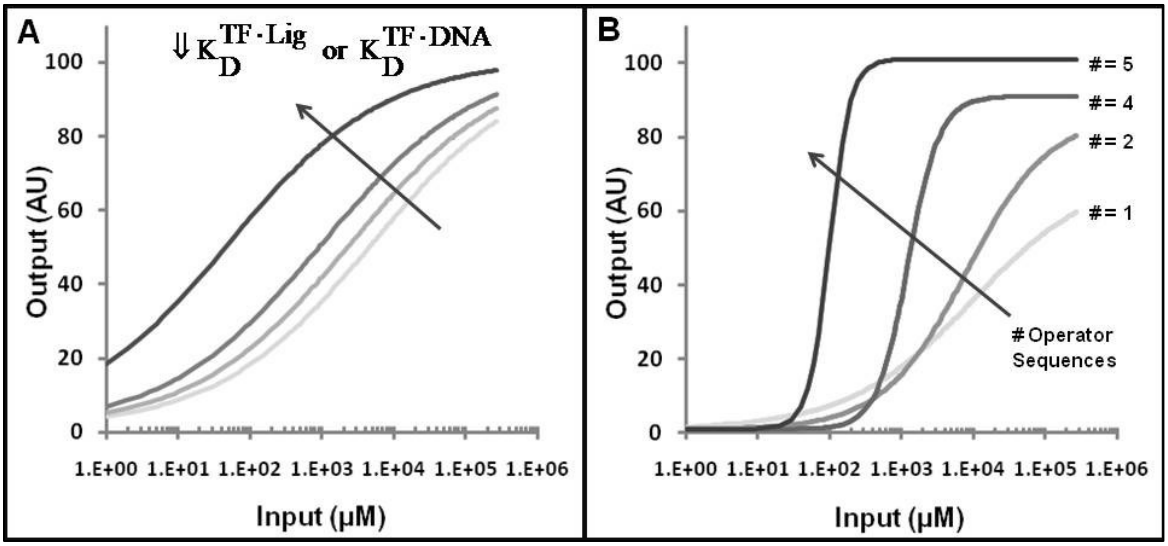


Figure 5.1: Anticipated results from modified TF-Ligand and TF-DNA binding. A) Improving TF-Ligand or TF-DNA binding shifts the dose-response curve to the right. **B**) Incorporating multiple, cooperative operator sequences narrows the linear response window, but increases biosensor sensitivity and dynamic range. Hybrid curves covering even greater dose-response space can be obtained by combining variant TFs, operators, and biosensor architectures.

References

1. Schmidt-Dannert, C. & Arnold, F.H. Directed evolution of industrial enzymes. *Trends Biotechnol.* **17**, 135-136 (1999).
2. Arnold, F.H. & Georgiou, G. *Directed Enzyme Evolution: Screening and Selection Methods*. (Humana Press: Totowa, NJ, 2003).
3. Niessen, W.M.A. *Liquid Chromatography – Mass Spectrometry*. (CRC Press: Boca Raton, 2006).
4. Fortman, J.L. et al. Biofuel alternatives to ethanol: pumping the microbial well. *Trends Biotechnol.* **26**, 375-381 (2008).
5. Atsumi, S. & Liao, J.C. Metabolic engineering for advanced biofuels production from Escherichia coli. *Curr. Opin. Biotech.* **19**, 414-419 (2008).
6. Barker, J.L. & Frost, J.W. Microbial synthesis of p-hydroxybenzoic acid from glucose. *Biotechnol. Bioeng.* **76**, 376 - 390 (2001).
7. Nakamura, C.E. & Whited, G.M. Metabolic engineering for the microbial production of 1,3-propanediol. *Curr. Opin. Biotech.* **14**, 454-459 (2003).
8. Aldor, I.S. & Keasling, J.D. Process design for microbial plastic factories: metabolic engineering of polyhydroxyalkanoates. *Curr. Opin. Biotech.* **14**, 475-483 (2003).
9. Dale, B.E. "Greening" the chemical industry: research and development priorities for biobased industrial products. *J. Chem. Technol. Biot.* **78**, 1093-1103 (2003).
10. Patrick, W.M., Firth, A.E. & Blackburn, J.M. User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries. *Protein Eng.* **16**, 451-457 (2003).
11. Murakami, H., Hohsaka, T. & Sisido, M. Random insertion and deletion of arbitrary number of bases for codon-based random mutation of DNAs. *Nat. Biotechnol.* **20**, 76-81 (2002).
12. Pikkemaat, M.G. & Janssen, D.B. Generating segmental mutations in haloalkane dehalogenase: a novel part in the directed evolution toolbox. *Nucleic Acids Res.* **30**, e35 (2002).
13. Osuna, J., Yáñez, J., Soberón, X. & Gaytán, P. Protein evolution by codon-based random deletions. *Nucleic Acids Res.* **32**, e136 (2004).
14. Stemmer, W.P.C. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389 - 391 (1994).
15. Zhao, H., Giver, L., Shao, Z., Affholter, J.A. & Arnold, F.H. Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.* **16**, 258-261 (1998).
16. Voigt, C., Kauffman, S. & Wang, Z. Rational evolutionary design: the theory of in vitro protein evolution. *Adv. Protein Chem.* **55**, 79-160 (2000).
17. Voigt, C.A. Genetic parts to program bacteria. *Curr. Opin. Biotech.* **17**, 548-557 (2006).
18. Blattner, F.R. et al. The complete genome sequence of Escherichia coli K-12. *Science* **277**, 1453-1462 (1997).
19. Serres, M.H. et al. A functional update of the Escherichia coli K-12 genome. *Genome Biol.* **2**, research0035.1–0035.7 (2001).
20. Lin, H. & Cornish, V.W. Screening and selection methods for large-scale analysis of protein function. *Angew. Chem. Int. Ed.* **41**, 4402-4425 (2002).

21. Wong, T.S., Roccatano, D. & Schwaneberg, U. Steering directed protein evolution: strategies to manage combinatorial complexity of mutant libraries. *Environ. Microbiol.* **9**, 2645-2659 (2007).
22. Shivange, A.V., Marienhagen, J., Mundhada, H., Schenk, A. & Schwaneberg, U. Advances in generating functional diversity for directed protein evolution. *Curr. Opin. Chem. Biol.* **13**, 19-25 (2009).
23. Patrick, W.M. & Firth, A.E. Strategies and computational tools for improving randomized protein libraries. *Biomol. Eng.* **22**, 105-112 (2005).
24. Firth, A.E. Statistics of randomized library construction. (2005).at <http://guinevere.otago.ac.nz/aef/STATS/index.html>
25. Firth, A.E. & Patrick, W.M. GLUE-IT and PEDEL-AA: new programmes for analyzing protein diversity in randomized libraries. *Nucleic Acids Res.* **35**, W281-W285 (2008).
26. Firth, A.E. & Patrick, W.M. Statistics of protein library construction. *Bioinformatics* **21**, 3314-3315 (2005).
27. Volles, M.J. & Lansbury, P.T. A computer program for the estimation of protein and nucleic acid sequence diversity in random point mutagenesis libraries. *Nucleic Acids Res.* **33**, 3667-3677 (2005).
28. Sidhu, S.S., Lowman, H.B., Cunningham, B.C. & Wells, J.A. Phage display for selection of novel binding peptides. *Methods Enzymol.* **328**, 333-358 (2000).
29. Leemhuis, H., Stein, V., Griffiths, A.D. & Hollfelder, F. New genotype–phenotype linkages for directed evolution of functional proteins. *Curr. Opin. Struct. Biol.* **15**, 472-478 (2005).
30. Greener, A., Callahan, M. & Jerpseth, B. An efficient random mutagenesis technique using an E. coli mutator strain. *Mol. Biotechnol.* **7**, 189-195 (1997).
31. Camps, M., Naukkarinen, J., Johnson, B.P. & Loeb, L.A. Targeted gene evolution in Escherichia coli using a highly error-prone DNA polymerase I. *P. Natl. Acad. Sci. USA* **100**, 9727-9732 (2003).
32. Wang, L., Jackson, W.C., Steinbach, P.A. & Tsien, R.Y. Evolution of new nonantibody proteins via iterative somatic hypermutation. *P. Natl. Acad. Sci. USA* **101**, 16745-16749 (2004).
33. Wang, L. & Tsien, R.Y. Evolving proteins in mammalian cells using somatic hypermutation. *Nat. Protoc.* **1**, 1346-1350 (2006).
34. Wang, H.H. et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894-898 (2009).
35. Yokobayashi, Y., Weiss, R. & Arnold, F.H. Directed evolution of a genetic circuit. *P. Natl. Acad. Sci. USA* **99**, 16587-16591 (2002).
36. Weiss, R., Homsy, G.E. & Knight, T.F. Toward in vivo digital circuits. *DIMACS Workshop on Evolution as Computation* **1**, 1-18 (1999).
37. Canton, B., Labno, A. & Endy, D. Refinement and standardization of synthetic biological parts and devices. *Nat. Biotechnol.* **26**, 787-793 (2008).
38. Tabor, J.J., Groban, E.S. & Voigt, C.A. Performance characteristics for sensors and circuits used to program E. coli. *Systems Biology and Biotechnology of Escherichia coli* 466 (2009).
39. Sylvestre, J., Chautard, H., Cedrone, F. & Delcourt, M. Directed evolution of biocatalysts. *Org. Process Res. Dev.* **10**, 562–571 (2006).
40. Turner, N.J. Agar plate-based assays. *Enzyme Assays: High-throughput Screening, Genetic Selection and Fingerprinting* (2006).

41. Mascio, P.D., Kaiser, S. & Sies, H. Lycopene as the most efficient biological carotenoid singlet oxygen quencher. *Arch. Biochem. Biophys.* **274**, 532-538 (1989).
42. Kim, S.-W. & Keasling, J.D. Metabolic engineering of the nonmevalonate isopentenyl diphosphate synthesis pathway in Escherichia coli enhances lycopene production. *Biotechnol. Bioeng.* **72**, 408-415 (2000).
43. Harker, M. & Bramley, P.M. Expression of prokaryotic 1-deoxy-D-xylulose-5-phosphatases in Escherichia coli increases carotenoid and ubiquinone biosynthesis. *FEBS Lett.* **448**, 115-119 (1999).
44. Albrecht, M., Takaichi, S., Steiger, S., Wang, Z.-Y. & Sandmann, G. Novel hydroxycarotenoids with improved antioxidative properties produced by gene combination in Escherichia coli. *Nat. Biotechnol.* **18**, 843-846 (2000).
45. Sandmann, G. Combinatorial biosynthesis of carotenoids in a heterologous host: a powerful approach for the biosynthesis of novel structures. *ChemBioChem* **3**, 629-635 (2002).
46. Smolke, C.D., Martin, V.J.J. & Keasling, J.D. Controlling the metabolic flux through the carotenoid pathway using directed mRNA processing and stabilization. *Metab. Eng.* **3**, 313-321 (2001).
47. Alper, H., Miyaoku, K. & Stephanopoulos, G. Construction of lycopene-overproducing E. coli strains by combining systematic and combinatorial gene knockout targets. *Nat. Biotechnol.* **23**, 612-616 (2005).
48. Farmer, W.R. & Liao, J.C. Improving lycopene production in Escherichi coli by engineering metabolic control. *Nat. Biotechnol.* **18**, 533-537 (2000).
49. Wang, C., Leffler, S., Thompson, D.H. & Hrycyna, C.A. A general fluorescence-based coupled assay for S-adenosylmethionine-dependent methyltransferases. *Biochem. Biophys. Res. Commun.* **331**, 351-356 (2005).
50. Zatar, N.A., Abu-Zuhri, A.Z. & Abu-Shaweesh, A.A. Spectrophotometric determination of some aromatic amines. *Talanta* **47**, 883-890 (1998).
51. Benson, J.R. & Hare, P.E. o-phthalaldehyde: fluorogenic detection of primary amines in the picomole range. Comparison with fluorescamine and ninhydrin. *P. Natl. Acad. Sci. USA* **72**, 619-622 (1975).
52. Mukherjee, P.S. & Karnes, H.T. Ultraviolet and fluorescence derivatization reagents for carboxylic acids suitable for high performance liquid chromatography: a review. *Biomed. Chromatogr.* **10**, 193-204 (1998).
53. Toyo'oka, T. Fluorescent tagging of physiologically important carboxylic acids, including fatty acids, for their detection in liquid chromatography. *Anal. Chim. Acta* **465**, 111-130 (2002).
54. Takadate, A., Rikura, M., Suehiro, T., Fujino, H. & Goya, S. New labeling reagents for alcohols in fluorescence high-performance liquid chromatography. *Chem. Pharm. Bull.* **33**, 1164-1169 (1985).
55. Vogel, M., Büldt, A. & Karst, U. Hydrazine reagents as derivatizing agents in environmental analysis – a critical review. *Z. Anal. Chem. Freseniu.* **366**, 781-791 (2000).
56. Yu, H., Tyo, K., Alper, H., Klein-Marcuschamer, D. & Stephanopoulos, G. A high-throughput screen for hyaluronic acid accumulation in recombinant Escherichia coli transformed by libraries of engineered sigma factors. *Biotechnol. Bioeng.* **101**, 788-796 (2008).

57. Danielson, N.D., Targove, M.A. & Miller, B.E. Pre- and postcolumn derivatization chemistry in conjunction with HPLC for pharmaceutical analysis. *J. Chromatogr. Sci.* **26**, 362-371 (1988).
58. Alcalde, M., Bulter, T. & Arnold, F.H. Colorimetric assays for biodegradation of polycyclic aromatic hydrocarbons by fungal laccases. *J. Biomol. Screen.* **7**, 547-553 (2002).
59. Miesenböck, G., Angelis, D.A. d & Rothman, J.E. Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature* **394**, 192-195 (1998).
60. Bornscheuer, U.T., Altenbuchner, J. & Meyer, H.H. Directed evolution of an esterase: screening of enzyme libraries based on pH-indicators and a growth assay. *Bioorgan. Med. Chem.* **7**, 2169-2173 (1999).
61. Zhou, M., Diwu, Z., Panchuk-Voloshina, N. & Haugland, R.P. A stable nonfluorescent derivative of resorufin for the fluorometric determination of trace hydrogen peroxide: applications in detecting the activity of phagocyte NADPH oxidase and other oxidases. *Anal. Biochem.* **253**, 162-8 (1997).
62. Charter, N.W., Kauffman, L., Singh, R. & Eglen, R.M. A generic, homogenous method for measuring kinase and inhibitor activity via adenosine 5'-diphosphate accumulation. *J. Biomol. Screen.* **11**, 390-399 (2006).
63. Koresawa, M. & Okabe, T. High-throughput screening with quantitation of ATP consumption: a universal non-radioisotope, homogeneous assay for protein kinase. *Assay Drug Dev. Techn.* **2**, 153-160 (2004).
64. Kleman-Leyer, K.M. et al. Characterization and Optimization of a Red-Shifted Fluorescence Polarization ADP Detection Assay. *Assay Drug Dev. Techn.* **7**, 56-67 (2009).
65. Klingenberg, M. Nicotinamid-adenine-dinucleotides (NAD, NADP, NADH, NADPH). 2045-2059 (1974).
66. Molnos, J., Gardiner, R., Dale, G.E. & Lange, R. A continuous coupled enzyme assay for bacterial malonyl–CoA:acyl carrier protein transacylase (FabD). *Anal. Biochem.* **319**, 171-176 (2003).
67. Smith, B.C., Hallows, W.C. & Denu, J.M. A continuous microplate assay for sirtuins and nicotinamide-producing enzymes. *Anal. Biochem.* in press (2009).
68. Hulcher, F.H. & Oleson, W.H. Simplified spectrophotometric assay for microsomal 3-hydroxy-3-methylglutaryl CoA reductase by measurement of coenzyme A. *J. Lipid Res.* **14**, 625-631 (1973).
69. Nath, A. & Atkins, W.M. A quantitative index of substrate promiscuity. *Biochemistry* **47**, 157–166 (2008).
70. Glieder, A., Farinas, E.T. & Arnold, F.H. Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. *Nat. Biotechnol.* **20**, 1135-1139 (2002).
71. Dietrich, J. et al. A novel semi-biosynthetic route for artemisinin production using engineered substrate-promiscuous P450BM3. *ACS Chem. Biol.* **4**, 261–267 (2009).
72. Hendricks, C.L., Ross, J.R., Pichersky, E., Noel, J.P. & Zhou, Z.S. An enzyme-coupled colorimetric assay for S-adenosylmethionine-dependent methyltransferases. *Anal. Biochem.* **326**, 100-105 (2004).
73. Dorgan, K.M. et al. An enzyme-coupled continuous spectrophotometric assay for S-adenosylmethionine-dependent methyltransferases. *Anal. Biochem.* **350**, 249-255 (2006).
74. Wagschal, K., Franqui-Espiet, D., Lee, C.C., Robertson, G.H. & Wong, D.W.S. Enzyme-coupled assay for β-xylosidase hydrolysis of natural substrates. *Appl. Environ. Microbiol.* **71**, 5318–5323 (2005).

75. Mattozzi, M., Tehara, S.K., Hong, T. & Keasling, J.D. Mineralization of paraoxon and its use as a sole C and P source by a rationally designed catabolic pathway in Pseudomonas putida. *Appl. Environ. Microb.* **72**, 6699-6706 (2006).
76. Bernhardt, P., McCoy, E. & O'Connor, S.E. Rapid identification of enzyme variants for reengineered alkaloid biosynthesis in periwinkle. *Chem. Biol.* **14**, 888-897 (2007).
77. Santos, C.N.S. & Stephanopoulos, G. Melanin-based high-throughput screen for L-tyrosine production in Escherichia coli. *Appl. Environ. Microb.* **74**, 1190-1197 (2008).
78. Sariaslani, F.S. Development of a combined biological and chemical process for production of industrial aromatics from renewable resources. *Annu. Rev. Microbiol.* **61**, 51-69 (2007).
79. Lütke-Eversloh, T., Nicole, C., Santos, S. & Stephanopoulos, G. Perspectives of biotechnological production of L-tyrosine and its applications. *Appl. Microbiol. and Biot.* **77**, 751-762 (2007).
80. Sprenger, G.A. From scratch to value: engineering Escherichia coli wild type cells to the production of L-phenylalanine and other fine chemicals derived from chorismate. *Appl. Microbiol. and Biot.* **75**, 739-749 (2007).
81. Lütke-Eversloh, T. & Stephanopoulos, G. A semi-quantitative high-throughput screening method for microbial L-tyrosine production in microtiter plates. *J. Ind. Microbiol. Biotechnol.* **34**, 807-811 (2007).
82. Lütke-Eversloh, T. & Stephanopoulos, G. L-Tyrosine production by deregulated strains of Escherichia coli. *Appl. Microbiol. and Biot.* **75**, 103-110 (2007).
83. Knight, J.A., Robertson, G. & Wu, J.T. The chemical basis and specificity of the nitrosonaphthol Reaction. *Clin. Chem.* **29**, 1969-1971 (1983).
84. Lütke-Eversloh, T. & Stephanopoulos, G. Combinatorial pathway analysis for improved L-tyrosine production in Escherichia coli: identification of enzymatic bottlenecks by systematic gene overexpression. *Metab. Eng.* **10**, 69-77 (2008).
85. Hall, B.G. Changes in the Substrate Specificities of an Enzyme during Directed Evolution of New Functions. *Biochemistry* **20**, 4042-4049 (1981).
86. Parales, R.E., Bruce, N.C., Schmid, A. & Wackett, L.P. Biodegradation, biotransformation, and biocatalysis (B3). *Appl. Environ. Microb.* **68**, 4699-4709 (2002).
87. Singh, S., Kang, S.H., Mulchandani, A. & Chen, W. Bioremediation: environmental clean-up through pathway engineering. *Curr. Opin. Biotech.* **19**, 437-444 (2008).
88. Martin, V.J.J., Pitera, D.J., Withers, S.T., Newman, J.D. & Keasling, J.D. Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nat. Biotechnol.* **21**, 796-802 (2003).
89. Pitera, D.J., Paddon, C.J., Newman, J.D. & Keasling, J.D. Balancing a heterologous mevalonate pathway for improved isoprenoid production in Escherichia coli. *Metab. Eng.* **9**, 193-207 (2007).
90. Yano, T., Oue, S. & Kagamiyama, H. Directed evolution of an aspartate aminotransferase with new substrate specificities. *P. Natl. Acad. Sci. USA* **95**, 5511-5515 (1998).
91. Oue, S., Okamoto, A., Yano, T. & Kagamiyama, H. Redesigning the substrate specificity of an enzyme by cumulative effects of the mutations of non-active site residues. *J. Biol. Chem.* **274**, 2344-2349 (1999).
92. Rothman, S.C. & Kirsch, J.F. How does an enzyme evolved in vitro compare to naturally occurring homologs possessing the targeted function? Tyrosine aminotransferase from aspartate aminotransferase. *J. Mol. Biol.* **327**, 593–608 (2003).

93. Jua, J., Misonob, H. & Ohnishic, K. Directed evolution of bacterial alanine racemases with higher expression level. *J. Biosci. Bioeng.* **100**, 246-254 (2005).
94. MacBeath G, Kast P, , and , Hilvert D Exploring sequence constraints on an interhelical turn using in vivo selection for catalytic activity. *Protein Sci.* **7**, 325-335 (1998).
95. Neuenschwander, M., Butz, M., Heintz, C., Kast, P. & Hilvert, D. A simple selection strategy for evolving highly efficient enzymes. *Nat. Biotechnol.* **25**, 1145-1147 (2007).
96. Atsumi, S., Hanai, T. & Liao, J.C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-90 (2008).
97. Atsumi, S. & Liao, J.C. Directed evolution of Methanococcus jannaschii citramalate synthase for biosynthesis of 1-propanol and 1-butanol by Escherichia coli. *Appl. Environ. Microb.* **2008**, 7802–7808 (2008).
98. Pfleger, B.F., Pitera, D.J., Newmana, J.D., Martin, V.J.J. & Keasling, J.D. Microbial sensors for small molecules: Development of a mevalonate biosensor. *Metab. Eng.* **9**, 30-38 (2007).
99. Kudla, G., Murray, A.W., Tollervey, D. & Plotkin, J.B. Coding-sequence determinants of gene expression in Escherichia coli. *Science* **324**, 255-258 (2009).
100. Becker, S., Schmoldt, H.-U., Adams, T.M., Wilhelm, S. & Kolmar, H. Ultra-high-throughput screening based on cell-surface display and fluorescence-activated cell sorting for the identification of novel biocatalysts. *Curr. Opin. Biotech.* **15**, 323-329 (2004).
101. Santoro, S.W. & Schultz, P.G. Directed evolution of the site specificity of Cre recombinase. *P. Natl. Acad. Sci. USA* **99**, 4185-4190 (2002).
102. Olsen, M.J. et al. Function-based isolation of novel enzymes from a large library. *Nat. Biotechnol.* **18**, 1071-1074 (2000).
103. Aharoni, A. et al. High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nat. Methods* **3**, 609-614 (2006).
104. van den Berg, S., Löfdahl, P.-Å., Härd, T. & Berglund, H. Improved solubility of TEV protease by directed evolution. *J. Biotechnol.* **121**, 291-298 (2006).
105. Cormack, B.P., Valdivia, R.H. & Falkow, S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33-38 (1996).
106. Santoro, S.W., Wang, L., Herberich, B., King, D.S. & Schultz, P.G. An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nat. Biotechnol.* **20**, 1044-1048 (2002).
107. Olsen, M.J., Gam, J., Iverson, B.L. & Georgiou, G. High-throughput FACS method for directed evolution of substrate specificity. *Directed Enzyme Evolution* **230**, 329-342 (2003).
108. Nonomura, A.M. & Coder, D.M. Improved phycocatalysis of carotene production by flow cytometry and cell sorting. *Biocatal. Biotransform.* **1**, 333-338 (1988).
109. An, G.-H., Bielich, J., Auerbach, R. & Johnson, E.A. Isolation and characterization of carotenoid hyperproducing mutants of yeast by flow cytometry and cell sorting. *Nat. Biotechnol.* **9**, 70-73 (1991).
110. Ukibe, K., Katsuragi, T., Tani, Y. & Takagi, H. Efficient screening for astaxanthin-overproducing mutants of the yeast Xanthophyllomyces dendrorhous by flow cytometry. *FEMS Microbiol. Lett.* **286**, 241-248 (2008).
111. Azuma, T., Harrison, G. & Demain, A. Isolation of a gramicidin S hyperproducing strain of Bacillus brevis by use of a flourescence activated cell sorting system. *Appl. Microbiol. and Biot.* **38**, 173-178 (1992).

112. Vidal-Mas, J. et al. Rapid flow cytometry – nile red assessment of PHA cellular content and heterogeneity in cultures of Pseudomonas aeruginosa 47T2 (NCIB 40044) grown in waste frying oil. *Anton. Leeuw.* **80**, 57-63 (2001).
113. Fouchet, P. et al. Quantitative single-cell detection of poly(β-hydroxybutyrate) accumulation in Rhizobium meliloti by flow cytometry. *FEMS Microbiol. Lett.* **126**, 31-35 (2006).
114. Silva, T.L., Reis, A., Medeiros, R., Oliveira, A.C. & Gouveia, L. Oil production towards biofuel from autotrophic microalgae semicontinuous cultivations monitorized by flow cytometry. *Appl. Biochem. Biotechnol.* **159**, 568-578 (2008).
115. Gouveia, L., Marques, A.E., Silva, T.L. & Reis, A. Neochloris oleabundans UTEX #1185: a suitable renewable lipid source for biofuel production. *J. Ind. Microbiol. Biotechnol.* **36**, 1367-5435 (2009).
1. Nebe-von-Caron, G., Stephens, P., Hewitt, C., Powell, J. & Badley, R. Analysis of bacterial function by multi-colour fluorescence flow cytometry and single cell sorting. *J. Microbiol. Methods* **42**, 97-114 (2000).
117. Valli, M. et al. Improvement of lactic acid production in Saccharomyces cerevisiae by cell sorting for high intracellular pH. *Appl. Environ. Microb.* **72**, 5492-5499 (2006).
118. Valli, M. et al. Intracellular pH distribution in Saccharomyces cerevisiae cell populations, analyzed by flow cytometry. *Appl. Environ. Microb.* **71**, 1515-1521 (2005).
119. Lewis, D.F.V. Oxidative stress: the role of cytochromes P450 in oxygen activation. *J. Chem. Technol. Biotechnol.* **77**, 1095-1100 (2002).
120. Halliwell, B. & Whiteman, M. Measuring reactive species and oxidative damage in vivo and in cell culture: how should you do it and what do the results mean? *Br. J. Pharmacol.* **142**, 231–255 (2009).
121. Soh, N. Recent advances in fluorescent probes for the detection of reactive oxygen species. *Anal. Bioanal. Chem.* **386**, 532–543 (2006).
122. Laroche, C., Beney, L., Marechal, P.A. & Gervais, P. The effect of osmotic pressure on the membrane fluidity of Saccharomyces cerevisiae at different physiological temperatures. *Appl. Microbiol. Biotechnol.* **56**, 249-254 (2001).
123. Kolisek, M. et al. Mrs2p is an essential component of the major electrophoretic Mg2+ influx system in mitochondria. *EMBO J.* **22**, 1235-1244 (2003).
124. Gangola, P. & Rosen, B. Maintenance of intracellular calcium in Escherichia coli. *J. Biol. Chem.* **262**, 12570-12574 (1987).
125. Uchiyama, T., Abe, T., Ikemura, T. & Watanabe, K. Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes. *Nat. Biotechnol.* **23**, 88-93 (2005).
126. Uchiyama, T. & Watanabe, K. Substrate-induced gene expression (SIGEX) screening of metagenome libraries. *Nat. Protoc.* **3**, 1202-1212 (2008).
127. Galvão, T.C., Mohn, W.W. & Lorenzo, V. d Exploring the microbial biodegradation and biotransformation gene pool. *Trends Biotechnol.* **23**, 497-506 (2005).
128. Williamson, L.L. et al. Intracellular screen to identify metagenomic clones that induce or inhibit a quorum-sensing biosensor. *Appl. Environ. Microb.* **71**, 6335-6344 (2005).
129. Sitnikov, D.M., Schineller, J.B. & Baldwin, T.O. Transcriptional regulation of bioluminesence genes from Vibrio fischeri. *Mol. Microbiol.* **17**, 801-812 (1995).
130. Schirmer, A., Hu, Z. & Da Costa, B. Methods and compositions for identification of hydrocarbon response, transport and biosynthesis genes. (2008).

131. Beggah, S., Vogne, C., Zenaro, E. & Meer, J.R. v d Mutant HbpR transcription activator isolation for 2-chlorobiphenyl via green fluorescent protein-based flow cytometry and cell sorting. *Microb. Biotechnol.* **1**, 68-78 (2008).
132. Breaker, R.R. Natural and engineered nucleic acids as tools to explore biology. *Nature* **432**, 838-845 (2004).
133. Lalonde, S., Ehrhardt, D.W. & Frommer, W.B. Shining light on signaling and metabolic networks by genetically encoded biosensors. *Curr. Opin. Plant Biol.* **8**, 574-581 (2005).
134. Gu, M.B., Mitchell, R.J. & Kim, B.C. Whole-cell-based biosensors for environmental biomonitoring and application. *Biomanufacturing* **87**, 269-305 (2004).
135. Endy, D. Foundations for engineering biology. *Nature* **438**, 449-453 (2005).
136. Ellis, T., Wang, X. & Collins, J.J. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat. Biotechnol.* **27**, 465-471 (2009).
137. Bintu, L. et al. Transcriptional regulation by the numbers: applications. *Curr. Opin. Genet. Dev.* **15**, 125-135 (2005).
138. Bintu, L. et al. Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.* **15**, 116-124 (2005).
139. Levitzki, A. & Schlessinger, J. Cooperativity in associating proteins. Monomer-dimer equilibrium coupled to ligand binding. *Biochemistry* **13**, 5214-5219 (1974).
140. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res.* **25**, 1203-1210 (1997).
141. Becskei, A., Kaufmann, B.B. & Oudenaarden, A. v Contributions of low molecule number and chromosomal positioning to stochastic gene expression. *Nat. Genet.* **37**, 937-944 (2005).
142. Lee, S. & Keasling, J.D. Propionate-regulated high-yield protein production in Escherichia coli. *Biotechnol. Bioeng.* **93**, 912 - 918 (2006).
143. Willardson, B.M. et al. Development and testing of a bacterial biosensor for toluene-based environmental contaminants. *Appl. Environ. Microbiol.* **64**, 1006-1012 (1998).
144. Kim, M.N., Park, H.H., Lim, W.K. & Shin, H.J. Construction and comparison of Escherichia coli whole-cell biosensors capable of detecting aromatic compounds. *J. Microbiol. Methods* **60**, 235-245 (2005).
145. Gardner, T.S., Cantor, C.R. & Collins, J.J. Construction of a genetic toggle switch in Escherichia coli. *Nature* **403**, 339-342 (2000).
146. Guet, C., Elowitz, M., Hsing, W. & Leibler, S. Combinatorial synthesis of genetic networks. *Science* **296**, 1466-1470 (2002).
1. Basu, S., Gerchman, Y., Collins, C., Arnold, F. & Weiss, R. A synthetic multicellular system for programmed pattern formation. *Nature* **434**, 1130-1134 (2005).
148. Basu, S., Mehreja, R., Thiberge, S., Chen, M.-T. & Weiss, R. Spatiotemporal control of gene expression with pulse-generating networks. *P. Natl. Acad. Sci. USA* **101**, 6355-6360 (2004).
149. Collins, C.H., Leadbetter, J.R. & Arnold, F.H. Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR. *Nat. Biotechnol.* **24**, 708-712 (2006).
150. Mohn, W.W., Garmendia, J., Galvao, T.C. & Lorenzo, V. d Surveying biotransformations with à la carte genetic traps: translating dehydrochlorination of lindane (gamma-

hexachlorocyclohexane) into lacZ-based phenotypes. *Environ. Microbiol.* **8**, 546-555 (2008).
151. Fernández, S., Shingler, V. & Lorenzo, V.D. Cross-regulation by XylR and DmpR activators of Pseudomonas putida suggests that transcriptional control of biodegradative operons evolves independently of catabolic genes. *J. Bacteriol.* **176**, 5052-5058 (1994).
152. Jensen, R.A. Enzyme recruitment in evolution of new function. *Annual Review of Microbiology* **30**, 409-425 (1976).
153. Khersonsky, O., Roodveldt, C. & Tawfik, D.S. Enzyme promiscuity: evolutionary and mechanistic aspects. *Curr. Opin. Biotech.* **10**, 498-508 (2006TY - JOUR).
154. Hult, K. & Berglund, P. Enzyme promiscuity: mechanism and applications. *Trends Biotechnol.* **25**, 231-238 (2007).
155. Olumeses, P. *Guidelines for the treatment of malaria*. (World Health Organization: Geneva, 2006).
156. Hale, V., Keasling, J.D., Renninger, N. & Diagana, T.T. Microbially derived artemisinin: a biotechnology solution to the global problem of access to affordable antimalarial drugs. *Am. J. Trop. Med. Hyg.* **77**, 198-202 (2007).
157. Korenromp, E., Miller, J., Nahlen, B., Wardlaw, T. & Young, M. *World Malaria Report 2005*. (World Health Organization, Roll Back Malaria: Geneva, 2005).
158. Covello, P.S. Making artemisinin. *Phytochemistry* **69**, 2881-2885 (2008).
159. Acton, N. & Roth, R.J. On the conversion of dihydroartemisinic acid into artemisinin. *J. Org. Chem.* **57**, 3610-3614 (1992).
160. Haynes, R.K. & Vonwiller, S.C. Cyclic peroxyacetal lactone, lactol and ether compounds. *US Patent 5310946* (1994).
161. Martin, V.J.J., Pitera, D.J., Withers, S.T., Newman, J.D. & Keasling, J.D. Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nat Biotech* **21**, 796-802 (2003).
162. Newman, J.D. et al., Jay D. Keasling High-level production of amorpha-4,11-diene in a two-phase partitioning bioreactor of metabolically engineered Escherichia coli. *Biotechnology and Bioengineering* **95**, 684-691 (2006).
163. Pitera, D.J., Paddon, C.J., Newman, J.D. & Keasling, J.D. Balancing a heterologous mevalonate pathway for improved isoprenoid production in Escherichia coli. *Metab. Eng.* **9**, 193-207 (2007).
164. Ro, D.-K. et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940-943 (2006).
165. Reiling, K.K., Renninger, N.S., McPhee, D.J., Fisher, K.J. & Ockey, D.A. Conversion of amorpha-4,11-diene to artemisinin and artemisinin precursors. *US Patent Application 20060270863* (2006).
166. Schmid, A. et al. Industrial biocatalysis today and tomorrow. *Nature* **409**, 258-268 (2001).
167. Schoemaker, H.E., Mink, D. & Wubbolts, M.G. Dispelling the myths-biocatalysis in industrial synthesis. *Science* **299**, 1694-1697 (2003).
168. Carter, O.A., Peters, R.J. & Croteau, R. Monoterpene biosynthesis pathway construction in Escherichia coli. *Phytochemistry* **64**, 425-433 (2003).
169. Chang, M.C.Y., Eachus, R.A., Trieu, W., Ro, D.-K. & Keasling, J.D. Engineering Escherichia coli for production of functionalized terpenoids using plant P450s. *Nat Chem Biol* **3**, 274-277 (2007).
170. Thomas, A.F. & Bessiere, Y. Limonene. *Nat. Prod. Rep.* **6**, 291-309 (1989).

171. Bhyrappa, P., Young, J.K., Moore, J.S. & Suslick, K.S. Dendrimer-metalloporphyrins: synthesis and catalysis. *J. Am. Chem. Soc.* **118**, 5708-5711 (1996).
172. Noble, M.A. et al. Roles of key active-site residues in flavocytochrome P450 BM3. *Biochem. J.* **339**, 371-9 (1999).
173. Glieder, A., Farinas, E.T. & Arnold, F.H. Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. *Nat Biotech* **20**, 1135-1139 (2002).
174. Carmichael, A.B. & Wong, L.-L. Protein engineering of Bacillus megaterium CYP102. The oxidation of polycyclic aromatic hydrocarbons. *European Journal of Biochemistry* **268**, 3117-3125 (2001).
175. Sowden, R.J., Yasmin, S., Rees, N.H., Bell, S.G. & Wong, L.-L. Biotransformation of the sesquiterpene (+)-valencene by cytochrome P450cam and P450BM3. *ChemInform* **3**, 57-64 (2005).
176. Bertea, C.M. et al. Identification of intermediates and enzymes involved in the early steps of artemisinin biosynthesis in Artemisia annua. *Planta Med.* **71**, 40-47 (2005).
177. Wallaart, T.E. et al. Isolation and identification of dihydroartemisinic acid from Artemisia annua and its possible role in the biosynthesis of artemisinin. *J. Nat. Prod.* **62**, 430-433 (1999).
178. Simons, K.T., Kooperberg, C., Huang, E. & Baker, D. Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and bayesian scoring functions. *J. Mol. Biol.* **268**, 209-225 (1997).
179. Ost, T.W.B. et al. Rational re-design of the substrate binding site of flavocytochrome P450 BM3. *FEBS Lett.* **486**, 173-177 (2000).
180. Chang, M.C.Y., Eachus, R.A., Trieu, W., Ro, D.-K. & Keasling, J.D. Engineering Escherichia coli for production of functionalized terpenoids using plant P450s. *Nat. Chem. Biol.* **3**, 274-277 (2007).
181. Lu, Y. & Mei, L. Co-expression of P450 BM3 and glucose dehydrogenase by recombinant Escherichia coli and its application in an NADPH-dependent indigo production system. *J Ind Microbiol Biotechnol* **34**, 247-253 (2006).
182. Watanabe, T. & Snell, E.E. Reversibility of the Tryptophanase Reaction: Synthesis of Tryptophan from Indole, Pyruvate, and Ammonia. *Proc Natl Acad Sci U S A* **69**, 1086-1090 (1972).
183. Ajikumar, P.K. et al. Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in Escherichia coli. *Science* **330**, 70 -74 (2010).
184. Lee, W.-H., Park, J.-B.P., Kim, M.-D. & Seo, J.-H. Enhanced production of e-caprolactone by overexpression of NADPH-regenerating glucose 6-phosphate dehydrogenase in recombinant Escherichia coli harboring cyclohexanone monooxygenase gene. *Appl. Microbiol. Biot.* **76**, 329-338 (2007).
185. Lim, S.-J., Jung, Y.-M., Shin, H.-D. & Lee, Y.-H. Amplification of the NADPH-related genes zwf and gnd for the oddball biosynthesis of PHB in an E. coli transformant harboring a cloned phbCAB operon. *J. Biosci. Bioeng.* **93**, 543-549 (2002).
186. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 6640 -6645 (2000).
187. Yoshikuni, Y., Ferrin, T.E. & Keasling, J.D. Designed divergent evolution of enzyme function. *Nature* **440**, 1078-1082 (2006).

188. Li, M.Z. & Elledge, S.J. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat Meth* **4**, 251-256 (2007).
189. Hutchins, R.O., Taffer, I.M. & Burgoyne, W. Regio- and stereoselective cleavage of epoxides with cyanoborohydride and boron trifluoride etherate. *J. Org. Chem.* **46**, 5214-5215 (1981).
190. Krysan, D.J., Haight, A.R., Menzia, J.A. & Welch, N. A stereoselective synthesis of the dihydroxyethylene dipeptide isostere, A82768. *Tetrahedron* **50**, 6163-6172 (1994).
191. Bal, B.S., Childers, W.E. & Pinnick, H.W. Oxidation of a, b-un saturated aldehydes. *Tetrahedron* **37**, 2091-2096 (1981).
192. Dalcanale E, M.F. Selective oxidation of aldehydes to carboxylic acids with sodium chlorite-hydrogen peroxide. *J. Org. Chem.* **51**, 567 (1986).
193. deVisser, S.P. What factors influence the ratio of C-H hydroxylation versus C=C epoxidation by a nonheme cytochrome P450 biomimetic? *J. Am. Chem. Soc.* **128**, 15809-15818 (2006).
194. Zanghellini, A. et al. New algorithms and an in silico benchmark for computational enzyme design. *Protein Sci* **15**, 2785-2794 (2006).
195. DeLano, W.L. The PyMOL molecular graphics system, http://www.pymol.org. (2002).at <http://www.pymol.org>
196. Omura, T. & Sato, R. The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. *J. Biol. Chem.* **239**, 2370-8 (1964).
197. Jäckel, C., Kast, P. & Hilvert, D. Protein design by directed evolution. *Annual Review of Biophysics* **37**, 153-173 (2008).
198. Arnold, F.H. & Volkov, A.A. Directed evolution of biocatalysts. *Current Opinion in Chemical Biology* **3**, 54-59 (1999).
199. Shivange, A.V., Marienhagen, J., Mundhada, H., Schenk, A. & Schwaneberg, U. Advances in generating functional diversity for directed protein evolution. *Current Opinion in Chemical Biology* **13**, 19-25 (2009).
200. Wang, H.H. et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894-898 (2009).
201. Alper, H. & Stephanopoulos, G. Global transcription machinery engineering: A new approach for improving cellular phenotype. *Metabolic Engineering* **9**, 258-267 (2007).
202. Alper, H., Moxley, J., Nevoigt, E., Fink, G.R. & Stephanopoulos, G. Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science* **314**, 1565-1568 (2006).
203. Lynch, M.D., Warnecke, T. & Gill, R.T. SCALEs: multiscale analysis of library enrichment. *Nature Methods* **4**, (2007).
204. Bonomo, J., Lynch, M.D., Warnecke, T., Price, J.V. & Gill, R.T. Genome-scale analysis of anti-metabolite directed strain engineering. *Metabolic Engineering* **10**, 109-120 (2008).
205. Atsumi, S. et al. Metabolic engineering of Escherichia coli for 1-butanol production. *Metab. Eng.* **10**, 305-311 (2008).
206. Inui, M. et al. Expression of Clostridium acetobutylicum butanol synthetic genes in Escherichia coli. *Appl. Microbiol. Biotechnol.* **77**, 1305-1316 (2008).
207. Nielsen, D.R. et al. Engineering alternative butanol production platforms in heterologous bacteria. *Metabolic Engineering* **11**, 262-273 (2009).
208. Steen, E.J. et al. Metabolic engineering of Saccharomyces cerevisiae for the production of n-butanol. *Microbial Cell Factories* **7**, 36 (2008).

209. Berezina, O.V. et al. Reconstructing the clostridial n-butanol metabolic pathway in Lactobacillus brevis. *Applied Microbiology and Biotechnology* **Online**, (2010).
210. Zhang, K., Sawaya, M.R., Eisenberg, D.S. & Liao, J.C. Expanding metabolism for biosynthesis of nonnatural alcohols. *Proceedings of the National Academy of Sciences USA* **105**, 20653-20658 (2008).
211. Atsumi, S., Hanai, T. & Liao, J.C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-89 (2008).
212. Sawicki, E., Hauser, T.R. & Fox, F.T. Spectrophotometric determination of aliphatic aldehyde 2,4-dinitrophenylhydrazones with 3-methyl-2-benzothiazolinone hydrazone. *Analytica Chimica Acta* **26**, 229-234 (1962).
213. Dickinson, R.G. & Jacobsen, N.W. A new sensitive and specific test for the detection of aldehydes: formation of 6-mercapto-3-substituted-s-triazolo[4,3-b]-s-tetrazines. *Chemical Communications* 1719-1720 (1970).
214. Minak-Bernero, V., Bare, R.E., Haith, C.E. & Grossman, M.J. Detection of alkanes, alcohols, and aldehydes using bioluminescence. *Biotechnology and Bioengineering* **87**, 170-177 (2004).
215. Tsotsou, G.E., Cass, A.E.G. & Gilardi, G. High throughput assay for cytochrome P450 BM3 for screening libraries of substrates and combinatorial mutants. *Biosensors and Bioelectronics* **17**, 119-131 (2002).
216. Monot, F., Martin, J.-R., Petitdemange, H. & Gay, R. Acetone and Butanol Production by Clostridium acetobutylicum in a Synthetic Medium. *Appl Environ Microbiol* **44**, 1318-1324 (1982).
217. Qureshi, N. & Blaschek, H.P. Production of Acetone Butanol Ethanol (ABE) by a Hyper-Producing Mutant Strain of Clostridium beijerinckii BA101 and Recovery by Pervaporation. *Biotechnology Progress* **15**, 594-602 (1999).
218. Atsumi, S., Hanai, T. & Liao, J.C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-90 (2008).
219. Atsumi, S. et al. Engineering the isobutanol biosynthetic pathway in Escherichia coli by comparison of three aldehyde reductase/alcohol dehydrogenase genes. *Applied Genetics and Molecular Biotechnology* **online**, DOI 10.1007/s00253-009-2085-6 (2009).
220. Connor, M.R. & Liao, J.C. Engineering Escherichia coli for the production of 3-methyl-1-butanol. *Appl. Environ. Microbiol.* AEM.00468-08 (2008).doi:10.1128/AEM.00468-08
221. Canton, B., Labno, A. & Endy, D. Refinement and standardization of synthetic biological parts and devices. *Nature Biotechnology* **26**, 787-793 (2008).
222. Lee, S.K. et al. Directed Evolution of AraC for Improved Compatibility of Arabinose- and Lactose-Inducible Promoters. *Appl. Environ. Microbiol.* **73**, 5711-5715 (2007).
223. Lee, S. & Keasling, J.D. Propionate-regulated high-yield protein production in Escherichia coli. *Biotechnol. Bioeng.* **93**, 912 - 918 (2006).
224. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res.* **25**, 1203-1210 (1997).
225. Sulzenbacher, G. et al. Crystal Structure of E. coli Alcohol Dehydrogenase YqhD: Evidence of a Covalently Modified NADP Coenzyme. *Journal of Molecular Biology* **342**, 489-502 (2004).

226. Dubbels, B.L., Sayavedra-Soto, L.A., Bottomley, P.J. & Arp, D.J. Thauera butanivorans sp. nov., a C2–C9 alkane-oxidizing bacterium previously referred to as "Pseudomonas butanovora."*Int J Syst Evol Microbiol* **59**, 1576-1578 (2009).
227. Kurth, E.G., Doughty, D.M., Bottomley, P.J., Arp, D.J. & Sayavedra-Soto, L.A. Involvement of BmoR and BmoG in n-alkane metabolism in "Pseudomonas butanovora."*Microbiology* **154**, 139-147 (2008).
228. Takahashi, J. et al. Isolation and identification of n-butane-assimilating bacterium. *Agric. Biol. Chem.* **44**, 1835-1840 (1980).
229. Popham, D., Szeto, D., Keener, J. & Kustu, S. Function of a bacterial activator protein that binds to transcriptional enhancers. *Science* **243**, 629 -635 (1989).
230. Collado-Vides, J., Magasanik, B. & Gralla, J.D. Control site location and transcriptional regulation in Escherichia coli. *Microbiol. Mol. Biol. Rev.* **55**, 371-394 (1991).
231. Wang, L. & Gralla, J.D. Multiple in vivo roles for the -12-region elements of sigma 54 promoters. *J. Bacteriol* **180**, 5626-5631 (1998).
232. Shingler, V. Signal sensing by sigma54-dependent regulators: derepression as a control mechanism. *Mol Microbiol* **19**, 409-416 (1996).
233. Studholme, D.J. & Dixon, R. Domain Architectures of {sigma}54-Dependent Transcriptional Activators. *J. Bacteriol.* **185**, 1757-1767 (2003).
234. Goldberg, J.B. & Dahnke, T. Pseudomonas aeruginosa AlgB, which modulates the expression of alginate, is a member of the NtrC subclass of prokaryotic regulators. *Mol Microbiol* **6**, 59-66 (1992).
235. Marchler-Bauer, A. et al. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Research* **39**, D225-D229 (2010).
236. Khlebnikov, A., Datsenko, K.A., Skaug, T., Wanner, B.L. & Keasling, J.D. Homogeneous expression of the PBAD promoter in Escherichia coli by constitutive expression of the low-affinity high-capacity AraE transporter. *Microbiology* **147**, 3241-3247 (2001).
237. Hugouvieux-Cotte-Pattat, N., Kohler, T., Rekik, M. & Harayama, S. Growth-phase-dependent expression of the Pseudomonas putida TOL plasmid pWW0 catabolic genes. *J. Bacteriol.* **172**, 6651-6660 (1990).
238. Waldo, G.S. Genetic screens and directed evolution for protein solubility. *Current Opinion in Chemical Biology* **7**, 33-38 (2003).
239. Yoshikuni, Y., Dietrich, J.A., Nowroozi, F.F., Babbitt, P.C. & Keasling, J.D. Redesigning Enzymes Based on Adaptive Evolution for Optimal Function in Synthetic Metabolic Pathways. *Chemistry & Biology* **15**, 607-618 (2008).
240. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res.* **25**, 1203-1210 (1997).
241. O'Donnell, S.M. & Janssen, G.R. The Initiation Codon Affects Ribosome Binding and Translational Efficiency in Escherichia coli of cI mRNA with or without the 5′ Untranslated Leader. *J Bacteriol* **183**, 1277-1283 (2001).
242. Chen, H., Bjerknes, M., Kumar, R. & Jay, E. Determination of the optimal aligned spacing between the Shine – Dalgarno sequence and the translation initiation codon of Escherichia coli m RNAs. *Nucleic Acids Research* **22**, 4953 -4957 (1994).
243. Shultzaberger, R.K., Bucheimer, R.E., Rudd, K.E. & Schneider, T.D. Anatomy of Escherichia coli ribosome binding sites. *Journal of Molecular Biology* **313**, 215-228 (2001).

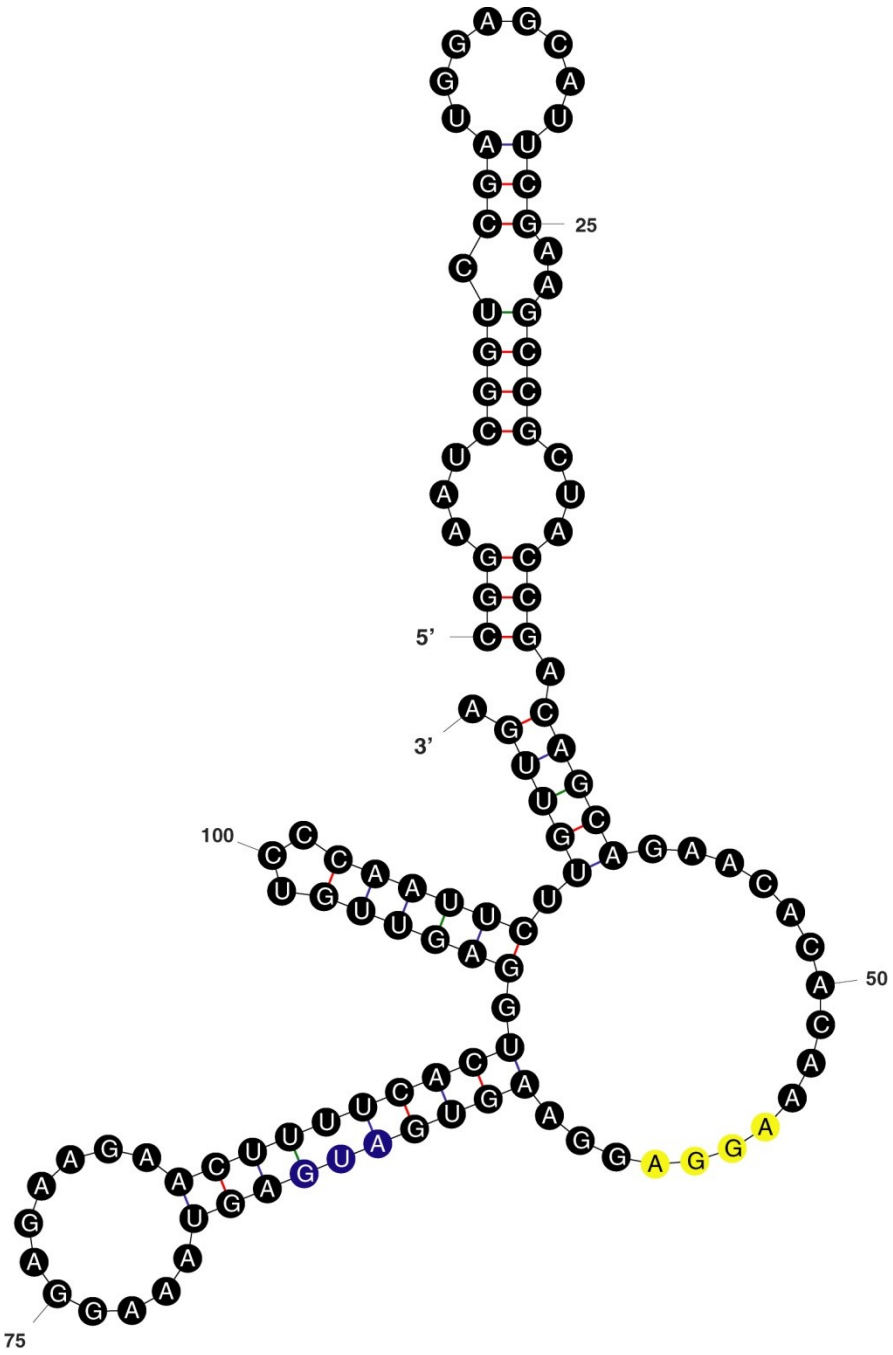
244. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946-950 (2009).
245. de Smit, M.H. & van Duin, J. Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 7668 -7672 (1990).
246. Studer, S.M. & Joseph, S. Unfolding of mRNA Secondary Structure by the Bacterial Translation Initiation Complex. *Molecular Cell* **22**, 105-115 (2006).
247. Kudla, G., Murray, A.W., Tollervey, D. & Plotkin, J.B. Coding-Sequence Determinants of Gene Expression in Escherichia coli. *Science* **324**, 255 -258 (2009).
248. Chan, L.Y., Kosuri, S. & Endy, D. Refactoring bacteriophage T7. *Mol Syst Biol* **1**, E1-E10 (2005).
249. Milanesio, P., Arce-Rodríguez, A., Muñoz, A., Calles, B. & de Lorenzo, V. Regulatory exaptation of the catabolite repression protein (Crp)-cAMP system in Pseudomonas putida. *Environmental Microbiology* **13**, 324-339 (2011).
250. Matys, V. et al. TRANSFAC®: transcriptional regulation, from patterns to profiles. *Nucleic Acids Research* **31**, 374 -378 (2003).
251. Miyada, C.G., Stoltzfus, L. & Wilcox, G. Regulation of the araC gene of Escherichia coli: catabolite repression, autoregulation, and effect on araBAD expression. *Proceedings of the National Academy of Sciences of the United States of America* **81**, 4120 -4124 (1984).
252. Eppler, T. & Boos, W. Glycerol-3-phosphate-mediated repression of malT in Escherichia coli does not require metabolism, depends on enzyme IIAGlc and is mediated by cAMP levels. *Molecular Microbiology* **33**, 1221-1231 (2002).
253. Eppler, T., Postma, P., Schutz, A., Volker, U. & Boos, W. Glycerol-3-Phosphate-Induced Catabolite Repression in Escherichia coli. *J. Bacteriol.* **184**, 3044-3052 (2002).
254. Skerra, A. Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in Escherichia coli. *Gene* **151**, 131-135 (1994).
255. Lee, S. & Keasling, J.D. Propionate-regulated high-yield protein production in Escherichia coli. *Biotechnol. Bioeng.* **93**, 912 - 918 (2006).
256. Buck, M., Gallegos, M.-T., Studholme, D.J., Guo, Y. & Gralla, J.D. The Bacterial Enhancer-Dependent sigma 54 (sigma N) Transcription Factor. *J. Bacteriol.* **182**, 4129-4136 (2000).
257. Canton, B., Labno, A. & Endy, D. Refinement and standardization of synthetic biological parts and devices. *Nat. Biotechnol.* **26**, 787-793 (2008).
258. Zhang, J.-H., Chung, T.D.Y. & Oldenburg, K.R. A simple statistical parameter for use in evaluation and validation of high throughput screening assay. *J Biomol Screen 1999; 4; 67* **4**, 67-73 (1999).
259. Atsumi, S. & Liao, J.C. Directed evolution of Methanococcus jannaschii citramalate synthase for biosynthesis of 1-propanol and 1-butanol by Escherichia coli. *Appl. Environ. Microb.* **2008**, 7802–7808 (2008).
260. Connor, M.R. & Liao, J.C. Engineering of an Escherichia coli strain for the production of 3-Methyl-1-Butanol. *Appl. Environ. Microbiol.* **74**, 5769-5775 (2008).
261. Muranaka, N., Sharma, V., Nomura, Y. & Yokobayashi, Y. An efficient platform for genetic selection and screening of gene switches in Escherichia coli. *Nucleic Acids Research* **7**, e39 (2009).
262. Podolsky, T., Fong, S.-T. & Lee, B.T.O. Direct selection of tetracycline-sensitive Escherichia coli cells using nickel salts. *Plasmid* **36**, 112-115 (1996).

263. Li, M.Z. & Elledge, S.J. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat Meth* **4**, 251-256 (2007).
264. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *P. Natl. Acad. Sci. USA* **97**, 6640-6645 (2000).
265. Zwietering, M.H., Jongenburger, I., Rombouts, F.M. & Riet, K. v t Modeling of the bacterial growth curve. *Appl Environ Microbiol* **56**, 1875-1881 (1990).
266. Zhang, K., Sawaya, M.R., Eisenberg, D.S. & Liao, J.C. Expanding metabolism for biosynthesis of nonnatural alcohols. *Proceedings of the National Academy of Sciences* doi:10.1073/pnas.0807157106
267. Studier, F.W. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* **41**, 207-234 (2005).
268. Salmon, K. et al. Global Gene Expression Profiling in Escherichia coli K12. *Journal of Biological Chemistry* **278**, 29837 -29855 (2003).
269. Huang, H.-H., Camsund, D., Lindblad, P. & Heidorn, T. Design and characterization of molecular tools for a Synthetic Biology approach towards developing cyanobacterial biotechnology. *Nucleic Acids Research* **38**, 2577 -2593 (2010).
270. Neidhardt, F.C., Bloch, P.L. & Smith, D.F. Culture Medium for Enterobacteria. *J Bacteriol* **119**, 736-747 (1974).
271. Hanai, T., Atsumi, S. & Liao, J.C. Engineered synthetic pathway for isopropanol production in Escherichia coli. *Appl. Environ. Microb.* **73**, 7814-7818 (2007).
272. Shena, C.R. & Liao, J.C. Metabolic engineering of Escherichia coli for 1-butanol and 1-propanol production via the keto-acid pathways. *Metab. Eng.* **10**, 312-320 (2008).
273. Soini, J. et al. Norvaline is accumulated after a down-shift of oxygen in Escherichia coli W3110. *Microbial Cell Factories* **7**, 30 (2008).
274. Parkins, C., Beckon, A.V., Beckon, W.N. & Maximovich, A. A General Approach to Modeling Biphasic Relationships. *Environmental Science & Technology* **42**, 1308-1314 (2008).
275. Quan, J. & Tian, J. Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways. *PLoS ONE* **4**, e6441 (2009).
276. Heim, R., Prasher, D.C. & Tsien, R.Y. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 12501 -12504 (1994).
277. Teerawanichpan, P., Hoffman, T., Ashe, P., Datla, R. & Selvaraj, G. Investigations of combinations of mutations in the jellyfish green fluorescent protein (GFP) that afford brighter fluorescence, and use of a version (VisGreen) in plant, bacterial, and animal cells. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1770**, 1360-1368 (2007).
278. Siegert, P. et al. Exchanging the substrate specificities of pyruvate decarboxylase from Zymomonas mobilis and benzoylformate decarboxylase from Pseudomonas putida. *Protein Engineering Design and Selection* **18**, 345 -357 (2005).
279. Eckert, B. & Beck, C.F. Overproduction of transposon Tn10-encoded tetracycline resistance protein results in cell death and loss of membrane potential. *J. Bacteriol.* **171**, 3557-3559 (1989).
280. Steen, E.J. et al. Microbial production of fatty acid-derived fuels and chemicals from plant biomass. *Nature* **463**, 559-562 (2010).

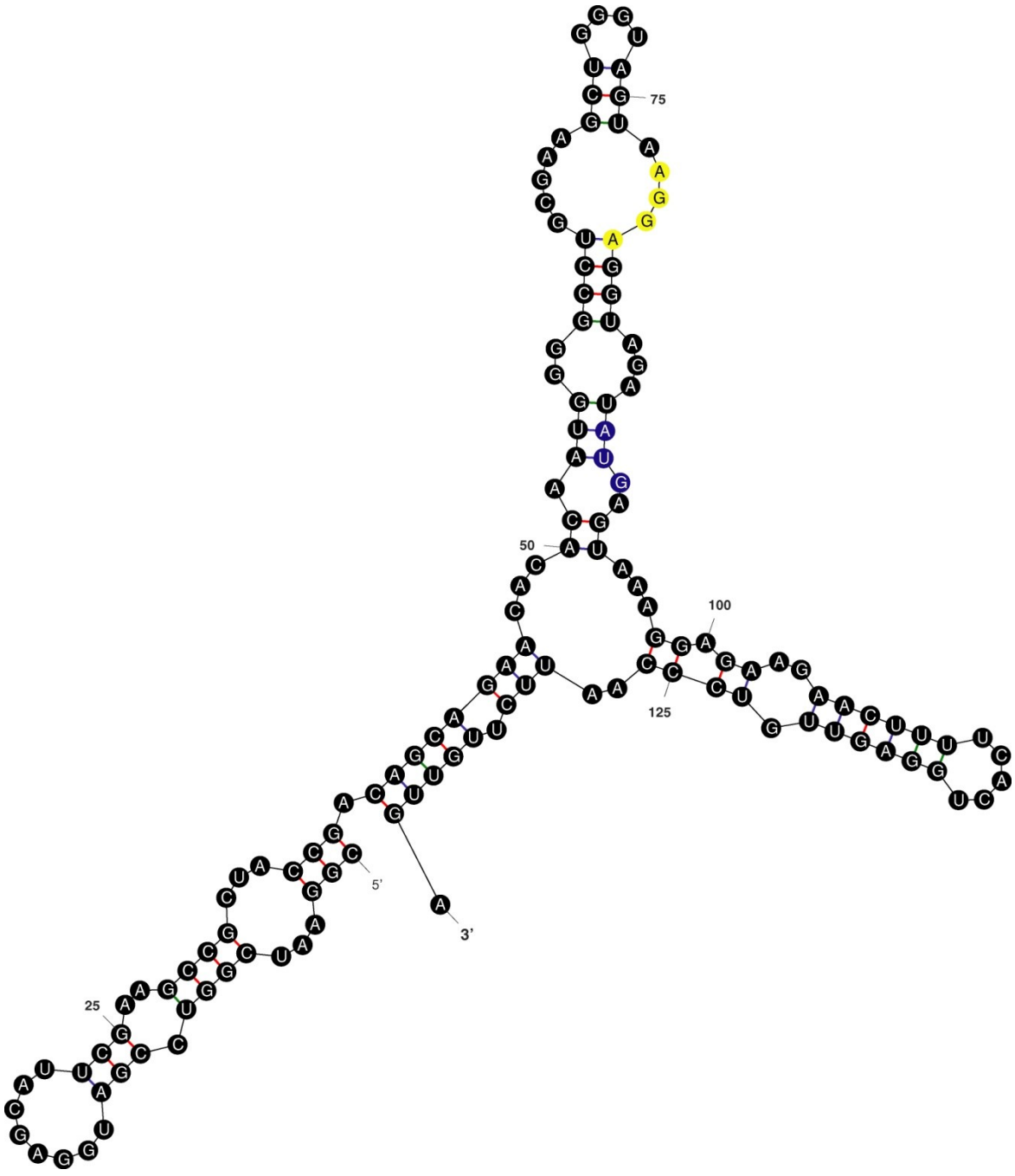
281. Abril, M.A., Michan, C., Timmis, K.N. & Ramos, J.L. Regulator and enzyme specificities of the TOL plasmid-encoded upper pathway for degradation of aromatic hydrocarbons and expansion of the substrate range of the pathway. *J Bacteriol* **171**, 6782-6790 (1989).
282. Herai, S. et al. Hyper-inducible expression system for streptomycetes. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14031 -14035 (2004).
283. Davies, S. et al. Inactivation and regulation of the aerobic C(4)-dicarboxylate transport (dctA) gene of Escherichia coli. *Journal of Bacteriology* **181**, 5624-5635 (1999).
284. Parales, R. & Harwood, C. Regulation of the pcaIJ genes for aromatic acid degradation in Pseudomonas putida. *Journal of Bacteriology* **175**, 5829-5838 (1993).
285. Guo, Z. & Houghton, J. PcaR-mediated activation and repression of pca genes from Pseudomonas putida are propagated by its binding to both the -35 and -10 promoter elements. *Molecular microbiology* **32**, 253-263 (1999).
286. Thimer, B., Andreesen, J. & Schrader, T. Cloning and characterization of a gene cluster involved in tetrahydrofuran degradation in Pseudonocarida sp. strain K1. *Archives of Microbiology* **179**, 266-277 (2003).
287. Donovan, R.S., Robinson, C.W. & Glick, B.R. Review: Optimizing inducer and culture conditions for expression of foreign proteins under the control of thelac promoter. *Journal of Industrial Microbiology* **16**, 145-154 (1996).
288. Voigt, C.A. Genetic parts to program bacteria. *Curr. Opin. Biotech.* **17**, 548-557 (2006).
289. Crameri, A., Whitehorn, E.A., Tate, E. & Stemmer, W.P.C. Improved Green Fluorescent Protein by Molecular Evolution Using DNA Shuffling. *Nat Biotech* **14**, 315-319 (1996).
290. Shaner, N.C. et al. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat Biotech* **22**, 1567-1572 (2004).
291. Williamson, L.L. et al. Intracellular screen to identify metagenomic clones that induce or inhibit a quorum-sensing biosensor. *Appl. Environ. Microb.* **71**, 6335-6344 (2005).
292. Tang, S.-Y. & Cirino, P. Design and Application of a Mevalonate-Responsive Regulatory Protein. *Angewandte Chemie International Edition* **49**, 1-4 (2010).
293. Collins, C.H., Arnold, F.H. & Leadbetter, J.R. Directed evolution of Vibrio fischeri LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones. *Mol. Microbiol.* **55**, 712–723 (2005).
294. Collins, C.H., Leadbetter, J.R. & Arnold, F.H. Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR. *Nat. Biotechnol.* **24**, 708-712 (2006).
295. Hawkins, A.C., Arnold, F.H., Hauer, R.S.B. & Leadbetter, J.R. Directed evolution of Vibrio fischeri LuxR for improved response to butanoyl-homoserine lactone. *Appl. Environ. Microbiol.* **73**, 5775-5781 (2007).
296. Becskei, A., Kaufmann, B.B. & Oudenaarden, A. v Contributions of low molecule number and chromosomal positioning to stochastic gene expression. *Nat. Genet.* **37**, 937-944 (2005).
297. Bintu, L. et al. Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.* **15**, 116-124 (2005).
298. Bintu, L. et al. Transcriptional regulation by the numbers: applications. *Curr. Opin. Genet. Dev.* **15**, 125-135 (2005).
299. Tabor, J.J., Groban, E.S. & Voigt, C.A. Performance characteristics for sensors and circuits used to program E. coli. *Systems Biology and Biotechnology of Escherichia coli* 466 (2009).

Appendix 1. Additional Figures

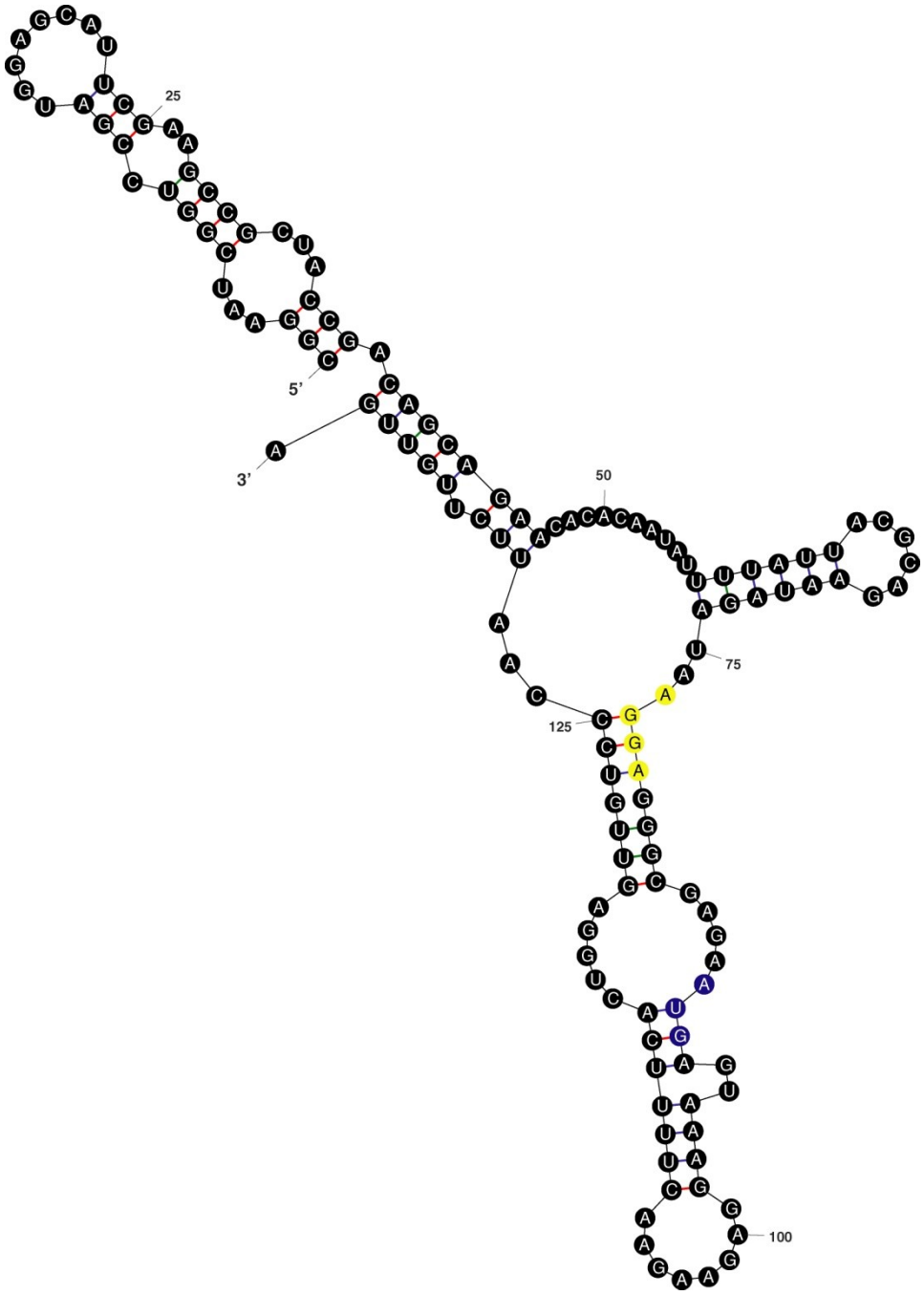
A1.1 Preliminary green fluorescent protein reporter ribosome binding site and 5'-untranslated region secondary structure. At 30°C, the preliminary reporter construct exhibited minimal dynamic range and fold-induction; a thermodynamic model of RNA folding depicts the presence of a hairpin loop in the first 50 bases of GFP[1], and we hypothesized that removing this hairpin loop would improve GFP translation. The canonical Shine-Dalgarno sequence (yellow) and GFP start site (blue) are indicated. ΔG=-24.4 kcal/mol.



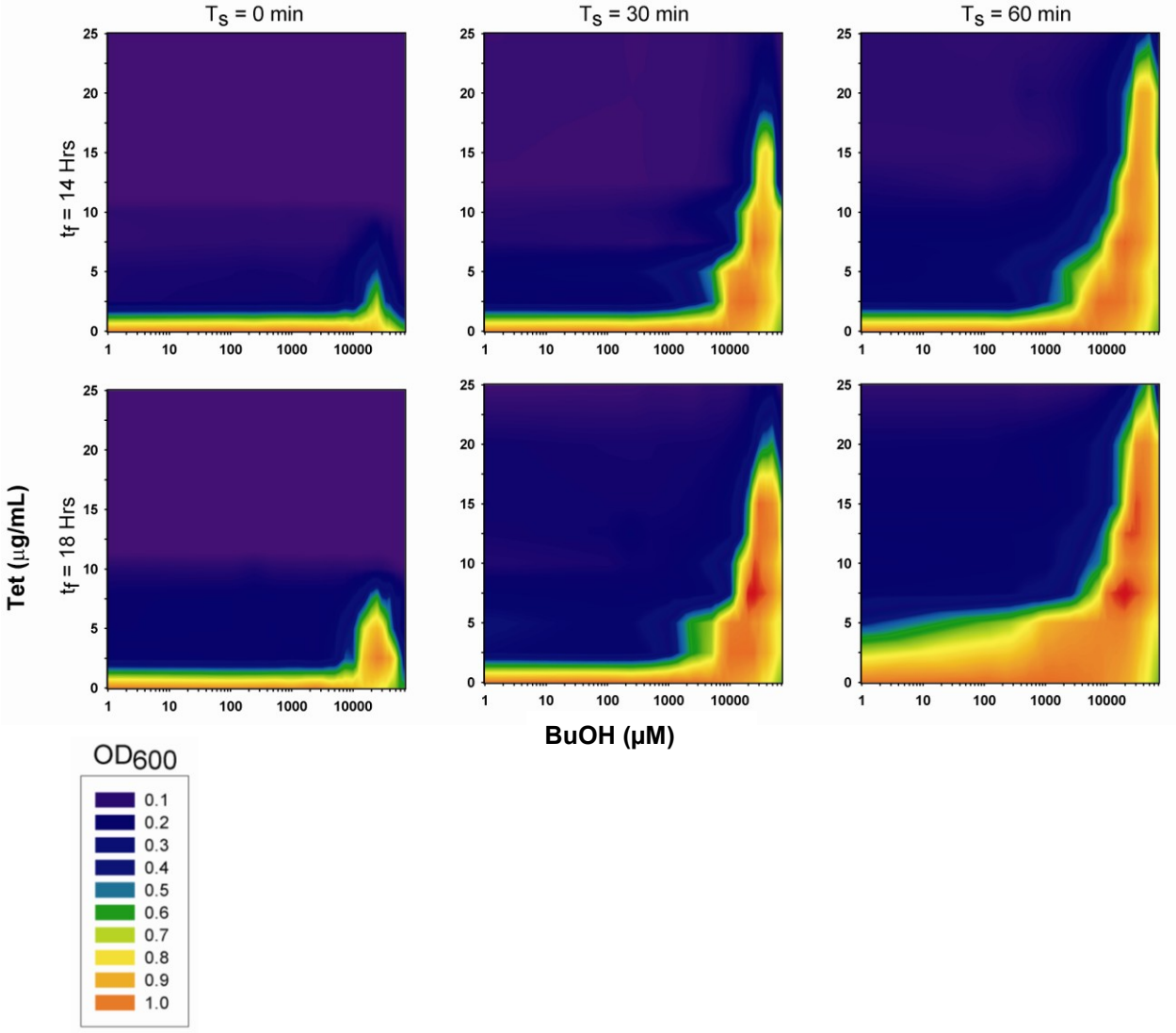
A1.2 *50K FLU* Synthetic ribosome binding site and 5'- untranslated region secondary structure. The Ribosome Binding Calculator, based on a thermodynamic model of binding between the mRNA transcript and the 30S ribosome complex, was used to design a synthetic ribosome binding site with improved translation initiation rate[2]. A thermodynamic model of RNA folding demonstrates removal of the hairpin at the 5' terminus of GFP[1], and substantial secondary structure remains present in the 5'-UTR. The canonical Shine-Dalgarno sequence (yellow) and GFP start site (blue) are indicated. ΔG=-27.2 kcal/mol.



A1.3 ***100K FLU*** **Synthetic ribosome binding site and 5'- untranslated region secondary structure.** The Ribosome Binding Calculator, based on a thermodynamic model of binding between the mRNA transcript and the 30S ribosome complex, was used to design a synthetic ribosome binding site with improved translation initiation rate[2]. As found with construct ***50K FLU*** (**see Figure A1.2**), a thermodynamic model of RNA folding demonstrates removal of the hairpin at the 5' terminus of GFP[1], and substantial secondary structure remains present in the 5'-UTR. The canonical Shine-Dalgarno sequence (yellow) and GFP start site (blue) are indicated. ΔG=-26.0 kcal/mol.



A1.4 Modifying 1-butanol selection parameters. The time between addition of 1-butanol and tetracycline, T_s, to cultures harboring pBMO#41 transformed *E. coli*, and the total assay length, t_f, control selection performance features.



Appendix 2. GenBank Files for Referenced Plasmids

A2.1. pCWori:BM3

```
LOCUS       pCWori:BM3             8113 bp    DNA     circular     11-MAY-2011
SOURCE
  ORGANISM
COMMENT     This file is created by Vector NTI
            http://www.invitrogen.com/
COMMENT     VNTDATE|589932576|
COMMENT     VNTDBDATE|589932576|
COMMENT     LSOWNER|
COMMENT     VNTNAME|pCWori:BM3|
COMMENT     VNTAUTHORNAME|Jeffrey Dietrich|
COMMENT     VNTAUTHOREML|jadietrich@berkeley.edu|
FEATURES             Location/Qualifiers
     promoter        7904..7998
                     /vntifkey="30"
                     /label=TAC\promoter
                     /note="TAC promoter"
     promoter        7999..8093
                     /vntifkey="30"
                     /label=TAC\promoter
                     /note="TAC promoter"
     misc_marker     3916..4779
                     /vntifkey="22"
                     /label=AmpR
                     /note="Ampicillin resistance gene"
     misc_feature    2..3151
                     /vntifkey="21"
                     /label=BM3
     mutation        143..145
                     /vntifkey="62"
                     /label=R47L
     mutation        155..157
                     /vntifkey="62"
                     /label=Y51F
     mutation        263..265
                     /vntifkey="62"
                     /label=F87A
     mutation        986..988
                     /vntifkey="62"
                     /label=A328L
BASE COUNT     1984 a     2098 c     2122 g     1909 t
ORIGIN
        1 tatggcgatt aaagaaatgc ctcaacctaa aaccttcggt gaactgaaaa acctgccgct
       61 gctgaacacc gacaagccag ttcaggcact gatgaaaatt gccgacgagc tcggcgaaat
```

121 tttcaaattc gaagccccag gcctggtgac ccgttttctg agcagccagc gtctgattaa
181 agaggcatgc gacgaatcta gatttgataa aaacctgtct caggccctga aattcgtgcg
241 tgatttcgca ggtgacggtc tggcgacttc ttggacccac gaaaagaatt ggaaaaaggc
301 ccacaacatt ctgctgcctt ctttctctca acaggcaatg aaaggttatc atgcaatgat
361 ggttgacatc gctgtccagc tggtccagaa atgggagcgt ctgaacgcgg atgaacacat
421 tgaagttcct gaagatatga cccgcctgac tctggacacc attggcctgt gtggtttcaa
481 ctaccggttc aacagcttct accgcgacca gccgcatccg ttcatcacca gcatggtgcg
541 tgctctggac gaagcaatga ataagctgca gcgcgctaac ccggatgatc cggcgtatga
601 cgaaaacaaa cgtcaattcc aggaagatat taaagtaatg aacgatctgg tagataagat
661 catcgcggac cgtaaggcta gcggtgagca aagcgacgac ctgctgacgc acatgctgaa
721 cggcaaagac ccggaaacgg gtgagccgct ggatgacgaa aatatccgtt atcagattat
781 tacctttctg attgcaggtc acgagactac tagcggtctg ctgtccttcg cgctgtactt
841 cctggtgaaa aatccacatg tgctgcagaa ggcggcggaa gaagccgcgc gtgtgctggt
901 tgacccggtg ccgtcctata aacaggtcaa acagctgaaa tatgtaggta tggttctgaa
961 cgaggccttg cgcctgtggc cgactctgcc ggcgttctct ctgtatgcga aggaagatac
1021 tgttctgggc ggtgaatacc cgctcgagaa aggtgatgaa ctgatggtcc tgattccgca
1081 gctgcaccgt gataagacga tttggggcga cgacgtagaa gaattccgtc cggagcgttt
1141 cgaaaatcct tccgctatcc cgcagcacgc cttcaaaccg tttggtaacg gtcaacgtgc
1201 ttgcattggc cagcaattcg ccctgcacga agctacgctg gtgctgggta tgatgctgaa
1261 gcacttcgac ttcgaggacc atactaacta cgagctggac atcaaagaaa ccctgactct
1321 gaagccggag ggtttcgttg ttaaagctaa atccaagaaa attccgctgg gtggtatccc
1381 ttctccttct acggaacaga gcgccaagaa agttcgtaaa aaggcggaaa acgcgcataa
1441 cacgccgctg ctggtactgt acggttctaa catgggtact gcggagggca ccgcccgtga
1501 tctggcggac atcgcaatgt ccaaaggctt cgccccgcaa gttgccaccc tggactccca
1561 tgcgggcaac ctgccgcgtg aaggtgccgt tctgatcgtt accgcatcct ataacggcca
1621 tccgccggat aatgcgaaac agtttgtaga ctggctggac caggcttctg cggatgaagt
1681 gaaaggtgtt cgctatagcg ttttcggttg cggtgacaaa aactgggcaa ctacctacca
1741 gaaagtacct gccttcatcg acgaaaccct ggccgctaaa ggtgctgaaa acattgcaga
1801 tcgtggtgaa gctgatgcgt ccgacgattt tgaaggtacc tacgaggaat ggcgtgaaca
1861 catgtggtct gatgtggctg cctatttcaa cctggacatc gaaaactctg aagacaacaa
1921 aagcactctg tccctgcagt ttgttgattc tgcggcggat atgccgctgg cgaaaatgca
1981 cggcgcgttc agcaccaatg tggttgcgtc caaggaactg caacagccgg gttctgcacg
2041 ctccacccgc cacctggaaa tcgaactgcc taaagaagcg agctaccagg aaggtgacca
2101 tctgggtgtc atcccgcgta actacgaagg tatcgtgaac cgtgtgactg ctcgttttgg
2161 cctggatgca agccagcaga ttcgcctgga agccgaagag gaaaaactgg ctcatctgcc
2221 gctggctaaa actgtaagcg tagaagaact gctgcagtat gtggaactgc aggacccggt
2281 tactcgcact caactgcgtg ctatggccgc gaaaaccgta tgtccgccgc acaaagttga
2341 actggaagcg ctgctggaga aacaggcata caaagaacag gtactggcca aacgtctgac
2401 catgctggaa ctgctggaaa aatatccggc gtgcgaaatg aaattctctg agttcattgc
2461 cctgctgccg tccatccgtc cgcgttacta ctccatcagc tcttcccctc gtgttgacga
2521 aaaacaggca agcattactg tatccgtggt ttccggcgaa gcgtggtctg gttacggcga
2581 atataagggc atcgcgagca actacctggc tgaactgcaa gaaggtgata ccatcacctg
2641 cttcatttct accccgcagt ccgaatttac cctgccgaaa gacccagaga ctccgctgat
2701 catggtcggt ccgggcaccg gcgttgcacc gttccgcggt tttgtacaag cacgtaagca
2761 gctgaaagag cagggccagt ccctgggtga agcgcacctg tacttcggtt gtcgttctcc
2821 gcatgaagac tacctgtacc aggaagaact ggagaacgcc cagagcgagg gtattattac

2881 cctgcatacc gctttctctc gtatgccgaa ccagccgaag acctacgtgc agcatgttat
2941 ggaacaggat ggcaagaaac tgatcgaact gctggaccag ggcgctcact tctatatctg
3001 cggtgatggt agccaaatgg caccggcggt cgaagcgacg ctgatgaaaa gctacgcaga
3061 cgtgcaccag gttagcgagg ctgacgcgcg tctgtggctg cagcagctgg aggagaaagg
3121 tcgttacgcg aaagatgtat gggccggtta aaagcttatc gatgataagc tgtcaaacat
3181 gagcagatct gagcccgcct aatgagcggg ctttttttc agatctgctt gaagacgaaa
3241 gggcctcgtg atacgcctat ttttataggt taatgtcatg ataataatgg tttcttagac
3301 gatgcgtcaa agcaaccata gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg
3361 gtggttacgc gcagcgtgac cgctacactt gccagcgccc tagcgcccgc tcctttcgct
3421 ttcttccctt cctttctcgc cacgttcgcc ggctttcccc gtcaagctct aaatcggggg
3481 ctccctttag ggttccgatt tagagcttta cggcacctcg accccaaaaa acttgatttg
3541 ggtgatggtt cacgtagtgg gccatcgccc tgatagacgg tttttcgccc tttgacgttg
3601 gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact caaccctatc
3661 tcgggctatt cttttgattt ataagggatt ttgccgattt cggcctattg gttaaaaaat
3721 gagctgattt aacaaaaatt taacgcgaat tttaacaaaa tattaacgtt tacaattcat
3781 cgtcaggtgg cacctttttcg gggaaatgtg cgcggaaccc ctatttgttt atttttctaa
3841 atacattcaa atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat
3901 tgaaaaagga agagtatgag tattcaacat ttccgtgtcg cccttattcc ctttttgcg
3961 gcattttgcc ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa
4021 gatcagttgg gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt
4081 gagagttttc gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt
4141 ggcgcggtat tatcccgtgt tgacgccggg caagagcaac tcggtcgccg catacactat
4201 tctcagaatg acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg
4261 acagtaagag aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta
4321 cttctgacaa cgatcggagg accgaaggag ctaaccgctt ttttgcacaa catgggggat
4381 catgtaactc gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag
4441 cgtgacacca cgatgcctgc agcaatggca acaacgttgc gcaaactatt aactggcgaa
4501 ctacttactc tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca
4561 ggaccacttc tgcgctcggc ccttccggct ggctggttta ttgctgataa atctggagcc
4621 ggtgagcgtg ggtctcgcgg tatcattgca gcactggggc cagatggtaa gccctcccgt
4681 atcgtagtta tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc
4741 gctgagatag gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat
4801 atactttaga ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt
4861 tttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac
4921 cccgtagaaa agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc
4981 ttgcaaacaa aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca
5041 actcttttc cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgtccttcta
5101 gtgtagccgt agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct
5161 ctgctaatcc tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg
5221 gactcaagac gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc
5281 acacagccca gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta
5341 tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg
5401 gtcggaacag gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt
5461 cctgtcgggt ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcaggggggg
5521 cggagcctat ggaaaaacgc cagcaacgcg gccttttac ggttcctggc cttttgctgg
5581 cctttgctc acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtattacc

5641 gcctttgagt gagctgatac cgctcgccgc agccgaacga ccgagcgcag cgagtcagtg
5701 agcgaggaag cggaagagcg cctgatgcgg tattttctcc ttacgcatct gtgcggtatt
5761 tcacaccgca tatatggtgc actctcagta caatctgctc tgatgccgca tagttaagcc
5821 agtatacact ccgctatcgc tacgtgactg ggtcatggct gcgccccgac acccgccaac
5881 acccgctgac gcgccctgac gggcttgtct gctcccggca tccgcttaca gacaagctgt
5941 gaccgtctcc gggagctgca tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag
6001 gcagaacgcc atcaaaaata attcgcgtct ggccttcctg tagccagctt tcatcaacat
6061 taaatgtgag cgagtaacaa cccgtcggat tctccgtggg aacaaacggc ggattgaccg
6121 taatgggata ggttacgttg gtgtagatgg gcgcatcgta accgtgcatc tgccagtttg
6181 aggggacgac gacagtatcg gcctcaggaa gatcgcactc cagccagctt tccggcaccg
6241 cttctggtgc cggaaaccag gcaaagcgcc attcgccatt caggctgcgc aactgttggg
6301 aagggcgatc ggtgcgggcc tcttcgctat tacgccagct ggcgaaaggg ggatgtgctg
6361 caaggcgatt aagttgggta acgccagggt tttcccagtc acgacgttgt aaaacgacgg
6421 ccagtgaatc cgtaatcatg gtcatagctg tttcctgtgt gaaattgtta tccgctcaca
6481 attccacaca acatacgagc cggaagcata aagtgtaaag cctggggtgc ctaatgagtg
6541 agctaactca cattaattgc gttgcgctca ctgcccgctt tccagtcggg aaacctgtcg
6601 tgccagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggcgc
6661 cagggtggtt tttcttttca ccagtgagac gggcaacagc tgattgccct tcaccgcctg
6721 gccctgagag agttgcagca agcggtccac gctggtttgc cccagcaggc gaaaatcctg
6781 tttgatggtg gttgacggcg ggatataaca tgagctgtct tcggtatcgt cgtatcccac
6841 taccgagata tccgcaccaa cgcgcagccc ggactcggta atggcgcgca ttgcgcccag
6901 cgccatctga tcgttggcaa ccagcatcgc agtgggaacg atgccctcat tcagcatttg
6961 catggtttgt tgaaaaccgg acatggcact ccagtcgcct tcccgttccg ctatcggctg
7021 aatttgattg cgagtgagat atttatgcca gccagccaga cgcagacgcg ccgagacaga
7081 acttaatggg cccgctaaca gcgcgatttg ctggtgaccc aatgcgacca gatgctccac
7141 gcccagtcgc gtaccgtctt catgggagaa aataatactg ttgatgggtg tctggtcaga
7201 gacatcaaga aataacgccg gaacattagt gcaggcagct tccacagcaa tggcatcctg
7261 gtcatccagc ggatagttaa tgatcagccc actgacgcgt tgcgcgagaa gattgtgcac
7321 cgccgcttta caggcttcga cgccgcttcg ttctaccatc gacaccacca cgctggcacc
7381 cagttgatcg gcgcgagatt taatcgccgc gacaatttgc gacggcgcgt gcagggccag
7441 actggaggtg gcaacgccaa tcagcaacga ctgtttgccc gccagttgtt gtgccacgcg
7501 gttgggaatg taattcagct ccgccatcgc cgcttccact ttttcccgcg ttttcgcaga
7561 aacgtggctg gcctggttca ccacgcggga aacggtctga taagagacac cggcatactc
7621 tgcgacatcg tataacgtta ctggtttcac attcaccacc ctgaattgac tctcttccgg
7681 gcgctatcat gccataccgc gaaaggtttt gcaccattcg atggtgtcct ggcacgacag
7741 gtttcccgac tggaaagcgg gcagtgagcg caacgcaatt aatgtgagtt agctcactca
7801 ttaggcaccc caggctttac actttatgct tccggctcgt ataatgtgtg gaattgtgag
7861 cggataacaa tttcacacag gaaacaggat cgatccatcg atgagcttac tccccatccc
7921 cctgttgaca attaatcatc ggctcgtata atgtgtggaa ttgtgagcgg ataacaattt
7981 cacacaggaa acaggatcag cttactcccc atccccctgt tgacaattaa tcatcggctc
8041 gtataatgtg tggaattgtg agcggataac aatttcacac aggaaacagg atccatcgat
8101 gcttaggagg tca
//

A2.2. pBMO#1

```
LOCUS       (Bmo#1)\pBMOE1-v       3338 bp    DNA     circular    27-AUG-2008
SOURCE
  ORGANISM
COMMENT     This file is created by Vector NTI
            http://www.invitrogen.com/
COMMENT     VNTDATE|489865550|
COMMENT     VNTDBDATE|491755219|
COMMENT     LSOWNER|
COMMENT     VNTNAME|(Bmo#1) pBMOE1-v1.1|
COMMENT     VNTAUTHORNAME|Jeffrey Dietrich|
COMMENT     VNTAUTHOREML|jadietrich@berkeley.edu|
FEATURES             Location/Qualifiers
     misc_feature    1257..1262
                     /vntifkey="21"
                     /label=XbaI\(1770)
     misc_feature    525..1241
                     /vntifkey="21"
                     /label=GFP
     terminator      1263..1391
                     /vntifkey="43"
                     /label=dblTerm
     rep_origin      complement(1524..2206)
                     /ApEinfo_fwdcolor=gray50
                     /ApEinfo_revcolor=gray50
                     /vntifkey="33"
                     /label=ColE1\origin
     CDS             complement(2345..3004)
                     /ApEinfo_fwdcolor=yellow
                     /ApEinfo_revcolor=yellow
                     /vntifkey="4"
                     /label=AmpR
     misc_feature    1..524
                     /vntifkey="21"
                     /label=pBMO
     misc_feature    432..437
                     /vntifkey="21"
                     /label=-24
     misc_feature    443..449
                     /vntifkey="21"
                     /label=-12
     misc_feature    461..461
                     /vntifkey="21"
                     /label=Approx\+1
     RBS             514..517
                     /vntifkey="32"
```

/label=RBS
misc_feature 4..22
/vntifkey="21"
/label=D24
misc_feature 505..524
/vntifkey="21"
/label=D27-R
misc_feature 224..243
/vntifkey="21"
/label=D26
misc_feature 283..302
/vntifkey="21"
/label=D25
primer_bind 1224..1279
/vntifkey="28"
/label=C22-R
misc_feature 440..453
/vntifkey="21"
/label=putative\C/EBP\factor\binding
misc_feature 503..507
/vntifkey="21"
/label=putatitve\HSF
misc_feature 71..96
/vntifkey="21"
/label=putative\CRP\binding\site
BASE COUNT 844 a 862 c 814 g 818 t
ORIGIN
1 ctgcccaacg acgtccgtca gagcccggtt cgagtggctt ctatatgccg atcatcggtg
61 gctctattgt ggcggtcagt gacaccggtc gccttcaccc ccacagatag taggtgctgc
121 ggctgctcat gctcctgtcg cggtagcgcg ctgttacgcg accgcccccg gacctcggcg
181 gacagcgcgg aagattggaa acagcccgag cgtgcgtgcc tcgggctgca tccttgccac
241 acccaaccgg attcgtcgga ccgctcgaca ttcgcgttcg ctcccgcggc gccgcgggtg
301 taccgttgcg ttacagatgt acccttcttt aacgtgtaac acacgcctgg agcggccaag
361 agccccgcac cttgcggcgc gtcttcccca ggggcccacc ggttgcggcc ttttgctgcg
421 accgtccatg ctggcacgac acttgctgaa agcgttagag cggaatcggt ccgatggagc
481 attcgaagcc gctaccgaca gcagaacaca caaaggagga agtgatgagt aaaggagaag
541 aacttttcac tggagttgtc ccaattcttg ttgaattaga tggtgatgtt aatgggcaca
601 aattttctgt cagtggagag ggtgaaggtg atgcaacata cggaaaactt acccttaaat
661 ttatttgcac tactggaaaa ctacctgttc cgtggccaac acttgtcact actttctctt
721 atggtgttca atgcttttcc cgttatccgg atcacatgaa acggcatgac tttttcaaga
781 gtgccatgcc cgaaggttat gtacaggaac gcactatatc tttcaaagat gacgggaact
841 acaagacgcg tgctgaagtc aagtttgaag gtgataccct tgttaatcgt atcgagttaa
901 aaggtattga ttttaaagaa gatggaaaca ttctcggaca caaactggag tacaactata
961 actcacacaa tgtatacatc acggcagaca aacaaaagaa tggaatcaaa gctaacttca
1021 aaattcgcca caacattgaa gatggctccg ttcaactagc agaccattat caacaaaata
1081 ctccaattgg cgatggccct gtccttttac cagacaacca ttacctgtcc acacaatctg

1141 ccctttcgaa agatcccaac gaaaagcgtg accacatggt ccttcttgag tttgtaactg
1201 ctgctgggat tacacatggc atggatgagc tctacaaata aggatcctaa ctcgagtcta
1261 gaccaggcat caaataaaac gaaaggctca gtcgaaagac tgggcctttc gttttatctg
1321 ttgtttgtcg gtgaacgctc tctactagag tcacactggc tcaccttcgg gtgggccttt
1381 ctgcgtttat acctaggcgt tcggctgcgg cgagcggtat cagctcactc aaaggcggta
1441 atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag
1501 caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc
1561 cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc gacaggacta
1621 taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt tccgaccctg
1681 ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct ttctcaatgc
1741 tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac
1801 gaaccccccg ttcagcccga ccgctgcgcc ttatccggta actatcgtct tgagtccaac
1861 ccggtaagac acgacttatc gccactggca gcagccactg gtaacaggat tagcagagcg
1921 aggtatgtag gcggtgctac agagttcttg aagtggtggc ctaactacgg ctacactaga
1981 aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagttggt
2041 agctcttgat ccggcaaaca aaccaccgct ggtagcggtg gttttttgt ttgcaagcag
2101 cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct
2161 gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgactag tgcttggatt
2221 ctcaccaata aaaaacgccc ggcggcaacc gagcgttctg aacaaatcca gatggagttc
2281 tgaggtcatt actggatcta tcaacaggag tccaagcgag ctcgtaaact tggtctgaca
2341 gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca
2401 tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc
2461 ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa
2521 accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc
2581 agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca
2641 acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat
2701 tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag
2761 cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac
2821 tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt
2881 ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt
2941 gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc
3001 tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat
3061 ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca
3121 gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga
3181 cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg
3241 gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg
3301 ttccgcgcac atttccccga aaagtgccac ctgacgtc
//

A2.3. pBMO#6

```
LOCUS       (Bmo#06)\pBadA2:       5795 bp    DNA     circular     6-OCT-2009
SOURCE
  ORGANISM
COMMENT
          ApEinfo:methylated:1
COMMENT     This file is created by Vector NTI
          http://www.invitrogen.com/
COMMENT     ORIGDB|GenBank
COMMENT     VNTDATE|531414884|
COMMENT     VNTDBDATE|531414884|
COMMENT     LSOWNER|
COMMENT     VNTNAME|(Bmo#06) pBadA2:BmoR|
COMMENT     VNTAUTHORNAME|Jeffrey Dietrich|
COMMENT     VNTAUTHOREML|jadietrich@berkeley.edu|
COMMENT     VNTOAUTHORNAME|UNKNOWN|
FEATURES             Location/Qualifiers
     misc_feature    complement(4747..4752)
                     /ApEinfo_label=rbs
                     /ApEinfo_fwdcolor=#fcc466
                     /ApEinfo_revcolor=#fcc466
                     /vntifkey="21"
                     /label=rbs
     misc_feature    4805..5649
                     /ApEinfo_label=p15A, OripACYC
                     /ApEinfo_fwdcolor=#feffb1
                     /ApEinfo_revcolor=#feffb1
                     /vntifkey="21"
                     /label=p15A,\OripACYC
     misc_feature    3250..3617
                     /ApEinfo_label=TrrnB
                     /ApEinfo_fwdcolor=#9191ff
                     /ApEinfo_revcolor=#9191ff
                     /vntifkey="21"
                     /label=TrrnB
     misc_feature    3740..4396
                     /ApEinfo_label=CmR (EcoRI-KO)
                     /ApEinfo_fwdcolor=#fff54c
                     /ApEinfo_revcolor=#fff54c
                     /vntifkey="21"
                     /label=CmR\(EcoRI-KO)
     CDS             complement(7..885)
                     /vntifkey="4"
                     /label=araC
     promoter        complement(1036..1064)
                     /vntifkey="30"
```

/label=Pc
misc_feature 914..931
/vntifkey="21"
/label=O2
misc_feature 1072..1093
/vntifkey="21"
/label=O1
misc_feature 1115..1128
/vntifkey="21"
/label=CAP\site
misc_feature 1124..1162
/vntifkey="21"
/label=I2\+\I1
promoter 1161..1188
/vntifkey="30"
/label=PBAD
misc_feature 1211..1233
/vntifkey="21"
/label=RBS
misc_feature 3244..3249
/vntifkey="21"
/label=Scar
primer_bind 5780..30
/vntifkey="28"
/label=C49-F
primer_bind 5780..30
/vntifkey="28"
/label=C50-R
misc_feature 1234..3243
/vntifkey="21"
/label=BmoR
primer_bind 3224..3271
/vntifkey="28"
/label=D64
primer_bind complement(3215..3263)
/vntifkey="28"
/label=D63
primer_bind complement(1206..1253)
/vntifkey="28"
/label=D62
primer_bind 1214..1262
/vntifkey="28"
/label=D61
primer_bind 1234..1256
/vntifkey="28"
/label=D57

```
     primer_bind     complement(3221..3243)
                     /vntifkey="28"
                     /label=D58
BASE COUNT     1338 a      1576 c      1551 g      1330 t
ORIGIN
        1 cctaggttat gacaacttga cggctacatc attcactttt tcttcacaac cggcacggaa
       61 ctcgctcggg ctggccccgg tgcatttttt aaatacccgc gagaaataga gttgatcgtc
      121 aaaaccaaca ttgcgaccga cggtggcgat aggcatccgg gtggtgctca aaagcagctt
      181 cgcctggctg atacgttggt cctcgcgcca gcttaagacg ctaatcccta actgctggcg
      241 gaaaagatgt gacagacgcg acggcgacaa gcaaacatgc tgtgcgacgc tggcgatatc
      301 aaaattgctg tctgccaggt gatcgctgat gtactgacaa gcctcgcgta cccgattatc
      361 catcggtgga tggagcgact cgttaatcgc ttccatgcgc cgcagtaaca attgctcaag
      421 cagatttatc gccagcagct ccgaatagcg cccttcccct tgcccggcgt taatgatttg
      481 cccaaacagg tcgctgaaat gcggctggtg cgcttcatcc gggcgaaaga accccgtatt
      541 ggcaaatatt gacggccagt taagccattc atgccagtag gcgcgcggac gaaagtaaac
      601 ccactggtga taccattcgc gagcctccgg atgacgaccg tagtgatgaa tctctcctgg
      661 cgggaacagc aaaatatcac ccggtcggca aacaaattct cgtccctgat ttttcaccac
      721 cccctgaccg cgaatggtga gattgagaat ataacctttc attcccagcg gtcggtcgat
      781 aaaaaaatcg agataaccgt tggcctcaat cggcgttaaa cccgccacca gatgggcatt
      841 aaacgagtat cccggcagca ggggatcatt ttgcgcttca gccatacttt tcatactccc
      901 gccattcaga gaagaaacca attgtccata ttgcatcaga cattgccgtc actgcgtctt
      961 ttactggctc ttctcgctaa ccaaaccggt aaccccgctt attaaaagca ttctgtaaca
     1021 aagcgggacc aaagccatga caaaaacgcg taacaaaagt gtctataatc acggcagaaa
     1081 agtccacatt gattatttgc acggcgtcac actttgctat gccatagcat ttttatccat
     1141 aagattagcg gatcctacct gacgcttttt atcgcaactc tctactgttt ctccataccc
     1201 gtttttttgg tagagaaaga ggagaaatac tagatgtcca agatgcaaga gttcgcgcgg
     1261 ctggagacag tcgcgtcgat gcgcagagcg gtctgggacg gcaacgagtg tcagccgggg
     1321 aaagtggctg atgtcgtttt gcgctcgtgg acccggtgtc gtgctgaagg tgtcgttccc
     1381 aatgcccgcc aggagttcga cccgatcccg cgaacggcgc ttgacgaaac ggttgaggcc
     1441 aagcgggcgc tgatccttgc tgccgagccg gtcgtcgacg cgttgatgga gcagatgaac
     1501 gacgccccca ggatgatcat cctgaacgac gaacggggcg tcgtgctgct gaaccaggga
     1561 aacgacaccc tccttgaaga cgcccgccgc cgggccgtgc gggtgggcgt ctgctgggac
     1621 gaacacgccc gaggcaccaa tgccatggga accgcgctcg cggagaggag gcccgtagcg
     1681 atccacggcg cagagcacta cctcgagtcg aatacgattt tcacctgcac cgcggcgccg
     1741 atctacgatc cgttcggcga gttcaccgga attctggata tcagcggata tgcgggggac
     1801 atgggcccgg ttccgattcc ctttgttcag atggcggtgc aattcatcga gaatcagttg
     1861 ttccgccaga cctttgccga ttgcattctg ctgcactttc atgtgcgccc cgacttcgtc
     1921 ggaacgatgc gcgaagggat agccgtgctg tcgcgcgagg gaaccatcgt ctcgatgaac
     1981 cgtgctgggc tcaagatcgc agggctcaac ctggaggccg tcgccgatca ccgtttcgat
     2041 tccgtcttcg acttgaattt cggggccttt ctcgaccacg tgcggcagtc cgccttcggt
     2101 ctcgtccgcg tctcgctcta cggcggcgtt caggtctacg cccgagtgga accgggcctg
     2161 cgtgttccgc cacgtccggc cgcccacgcc cgccctcctc ggccggcacc gcggcctctg
     2221 gattcgctgg acacgggcga cgcagcagtc cgcctcgcga ttgaccgcgc ccgccgcgcg
     2281 atcggccgca acctcagcat cctcatccag ggcgagacgg gtgccggcaa ggaagtgttc
     2341 gccaagcatc tgcatgccga gagcccgaga agcaaggggc cgttcgttgc cgtcaattgc
     2401 gccgccatac ccgagggttt gatcgagtcc gagcttttcg gatacgaaga aggggccttc
```

2461 actggcggac ggcgcaaagg caacatcggc aaggtcgccc aggcccacgg cggcacgttg
2521 ttcctcgacg agatcggaga catggcgccg gggctgcaga ccagactgct gcgggttcta
2581 caggaccggg cagtgatgcc cctcggcggt cgcgaaccta tgccggtcga cattgcgctg
2641 gtatgcgcga cgcaccgcaa tttgcgcagc ctgatagcgc agggccagtt ccgcgaagac
2701 ctctactacc gcctgaacgg gctggcgatc tcgctgccac ccctgcgtca acgtagcgac
2761 cttgccgccc tggttaacca tatcctcttt cagtgttgcg ggggcgagcc gcactacagc
2821 gtgagcccag aagtgatgac gctcttcaag cggcatgcgt ggcccggcaa cctacgccaa
2881 ctacataacg ttctcgatgc agcgcttgcc atgctcgacg acggccatgt cattgagccc
2941 catcacctcc ccgaagactt cgtcatggag gtcgattcgg gcctccgacc gatcgaggaa
3001 gacggttcga cggcggcgca tcgcgcgcga cagccggcgt cgggaagcgg tcctgccaaa
3061 aagttgcaag atctcgcgtt ggatgccatc gagcaggcga tcgagcaaaa cgagggaaat
3121 atatcggtcg ccgcacggca gttgggggtc agccggacca cgatctaccg caagctgagg
3181 caactttcac caaccggttg tcaccgaccg gcacattgga gccagtcgcg gatcggcaca
3241 tagggatctg aagcttgggc ccgaacaaaa actcatctca gaagaggatc tgaatagcgc
3301 cgtcgaccat catcatcatc atcattgagt ttaaacggtc tccagcttgg ctgttttggc
3361 ggatgagaga agattttcag cctgatacag attaaatcag aacgcagaag cggtctgata
3421 aaacagaatt tgcctggcgg cagtagcgcg gtggtcccac ctgaccccat gccgaactca
3481 gaagtgaaac gccgtagcgc cgatggtagt gtggggtctc cccatgcgag agtagggaac
3541 tgccaggcat caaataaaac gaaaggctca gtcgaaagac tgggcctttc gttttatctg
3601 ttgtttgtcg gtgaactaat tatctagact gcagttgatc gggcacgtaa gaggttccaa
3661 ctttcaccat aatgaaataa gatcactacc gggcgtattt tttgagttat cgagattttc
3721 aggagctaag gaagctaaaa tggagaaaaa aatcactgga tataccaccg ttgatatatc
3781 ccaatggcat cgtaaagaac attttgaggc atttcagtca gttgctcaat gtacctataa
3841 ccagaccgtt cagctggata ttacggcctt tttaaagacc gtaaagaaaa ataagcacaa
3901 gttttatccg gcctttattc acattcttgc ccgcctgatg aatgctcatc cggaatttcg
3961 tatggcaatg aaagacggtg agctggtgat atgggatagt gttcaccctt gttacaccgt
4021 tttccatgag caaactgaaa cgttttcatc gctctggagt gaataccacg acgatttccg
4081 gcagtttcta cacatatatt cgcaagatgt ggcgtgttac ggtgaaaacc tggcctattt
4141 ccctaaaggg tttattgaga atatgttttt cgtctcagcc aatccctggg tgagtttcac
4201 cagttttgat ttaaacgtgg ccaatatgga caacttcttc gcccccgttt tcaccatggg
4261 caaatattat acgcaaggcg acaaggtgct gatgccgctg gcgattcagg ttcatcatgc
4321 cgtttgtgat ggcttccatg tcggcagaat gcttaatgaa ttacaacagt actgcgatga
4381 gtggcagggc ggggcgtaat ttgatatcga gctcgcttgg actcctgttg atagatccag
4441 taatgacctc agaactccat ctggatttgt tcagaacgct cggttgccgc cgggcgtttt
4501 ttattggtga gaatccaagc ctcggtgaga atccaagcct cgatcaacgt ctcattttcg
4561 ccaaaagttg gcccagggct tcccggtatc aacagggaca ccaggattta tttattctgc
4621 gaagtgatct tccgtcacag gtatttattc ggcgcaaagt gcgtcgggtg atgctgccaa
4681 cttactgatt tagtgtatga tggtgttttt gaggtgctcc agtggcttct gtttctatca
4741 gctgtccctc ctgttcagct actgacgggg tggtgcgtaa cggcaaaagc accgccggac
4801 atcagcgcta gcggagtgta tactggctta ctatgttggc actgatgagg gtgtcagtga
4861 agtgcttcat gtggcaggag aaaaaaggct gcaccggtgc gtcagcagaa tatgtgatac
4921 aggatatatt ccgcttcctc gctcactgac tcgctacgct cggtcgttcg actgcggcga
4981 gcggaaatgg cttacgaacg gggcggagat ttcctggaag atgccaggaa gatacttaac
5041 agggaagtga gagggccgcg gcaaagccgt ttttccatag gctccgcccc cctgacaagc
5101 atcacgaaat ctgacgctca aatcagtggt ggcgaaaccc gacaggacta taaagatacc
5161 aggcgtttcc ccctggcggc tccctcgtgc gctctcctgt tcctgccttt cggtttaccg

5221 gtgtcattcc gctgttatgg ccgcgtttgt ctcattccac gcctgacact cagttccggg
5281 taggcagttc gctccaagct ggactgtatg cacgaacccc ccgttcagtc cgaccgctgc
5341 gccttatccg gtaactatcg tcttgagtcc aacccggaaa gacatgcaaa agcaccactg
5401 gcagcagcca ctggtaattg atttagagga gttagtcttg aagtcatgcg ccggttaagg
5461 ctaaactgaa aggacaagtt ttggtgactg cgctcctcca agccagttac ctcggttcaa
5521 agagttggta gctcagagaa ccttcgaaaa accgccctgc aaggcggttt tttcgttttc
5581 agagcaagag attacgcgca gaccaaaacg atctcaagaa gatcatctta ttaatcagat
5641 aaaatatttc tagatttcag tgcaatttat ctcttcaaat gtagcacctg aagtcagccc
5701 catacgatat aagttgtaat tctcatgttt gacagcttat catcgataag cttccgatgg
5761 cgcgccgaga ggctttacac tttatgcttc cggct
//

A2.4. pBMO#7

```
LOCUS       (Bmo#07)\pZTA2:B       5502 bp    DNA     circular     13-OCT-2008
SOURCE
  ORGANISM
COMMENT
         ApEinfo:methylated:1
COMMENT     This file is created by Vector NTI
         http://www.invitrogen.com/
COMMENT     ORIGDB|GenBank
COMMENT     VNTDATE|496076509|
COMMENT     VNTDBDATE|496076849|
COMMENT     LSOWNER|
COMMENT     VNTNAME|(Bmo#07) pZTA2:BmoR|
COMMENT     VNTAUTHORNAME|Jeffrey Dietrich|
COMMENT     VNTAUTHOREML|jadietrich@berkeley.edu|
COMMENT     VNTOAUTHORNAME|UNKNOWN|
FEATURES             Location/Qualifiers
     misc_feature    complement(4454..4459)
                     /ApEinfo_label=rbs
                     /ApEinfo_fwdcolor=#fcc466
                     /ApEinfo_revcolor=#fcc466
                     /vntifkey="21"
                     /label=rbs
     misc_feature    4512..5356
                     /ApEinfo_label=p15A, OripACYC
                     /ApEinfo_fwdcolor=#feffb1
                     /ApEinfo_revcolor=#feffb1
                     /vntifkey="21"
                     /label=p15A,\OripACYC
     misc_feature    2957..3324
                     /ApEinfo_label=TrrnB
                     /ApEinfo_fwdcolor=#9191ff
                     /ApEinfo_revcolor=#9191ff
                     /vntifkey="21"
                     /label=TrrnB
     misc_feature    3447..4103
                     /ApEinfo_label=CmR (EcoRI-KO)
                     /ApEinfo_fwdcolor=#fff54c
                     /ApEinfo_revcolor=#fff54c
                     /vntifkey="21"
                     /label=CmR\(EcoRI-KO)
     misc_feature    2951..2956
                     /vntifkey="21"
                     /label=Scar
     misc_feature    941..2950
                     /vntifkey="21"
```

/label=BmoR
primer_bind 2931..2978
/vntifkey="28"
/label=D64
primer_bind complement(2922..2970)
/vntifkey="28"
/label=D63
primer_bind 941..963
/vntifkey="28"
/label=D57
primer_bind complement(2928..2950)
/vntifkey="28"
/label=D58
primer_bind 840..857
/vntifkey="28"
/label=ZBZF
misc_feature complement(7..633)
/vntifkey="21"
/label=TetR
promoter 745..814
/vntifkey="30"
/label=Pzt1
primer_bind 5485..30
/vntifkey="28"
/label=D103-F
primer_bind complement(5485..30)
/vntifkey="28"
/label=D104-R
primer_bind 916..963
/vntifkey="28"
/label=D106-F
primer_bind complement(916..963)
/vntifkey="28"
/label=D107-R
BASE COUNT 1290 a 1438 c 1440 g 1334 t
ORIGIN
1 cctaggttaa gacccacttt cacatttaag ttgtttttct aatccgcata tgatcaattc
61 aaggccgaat aagaaggctg gctctgcacc ttggtgatca aataattcga tagcttgtcg
121 taataatggc ggcatactat cagtagtagg tgtttccctt tcttctttag cgacttgatg
181 ctcttgatct tccaatacgc aacctaaagt aaaatgcccc acagcgctga gtgcatataa
241 tgcattctct agtgaaaaac cttgttggca taaaaaggct aattgatttt cgagagtttc
301 atactgtttt tctgtaggcc gtgtacctaa atgtactttt gctccatcgc gatgacttag
361 taaagcacat ctaaaacttt tagcgttatt acgtaaaaaa tcttgccagc tttccccttc
421 taaagggcaa aagtgagtat ggtgcctatc taacatctca atggctaagg cgtcgagcaa
481 agcccgctta ttttttacat gccaatacaa tgtaggctgc tctacaccta gcttctgggc
541 gagtttacgg gttgttaaac cttcgattcc gacctcatta agcagctcta atgcgctgtt

601 aatcacttta cttttatcta atctagacat cattaattcc taatttttgt tgacactcta
661 tcgttgatag agttatttta ccactcccta tcagtgatag agaaaagtga aaatccatat
721 gactagtaga tcctctagag tcgactaaga aaccattatt atcatgacat taacctataa
781 aaataggcgt atcacgaggc cctttcgtct tcacctcgag aaatcataaa aaatttattt
841 gcttccctat cagtgataga gtataataga gtcgaattgt tagcggagaa gaatttcaca
901 cagaattcat tctgcagcag gacctaggag gaggaattca atgtccaaga tgcaagagtt
961 cgcgcggctg gagacagtcg cgtcgatgcg cagagcggtc tgggacggca acgagtgtca
1021 gccggggaaa gtggctgatg tcgttttgcg ctcgtggacc cggtgtcgtg ctgaaggtgt
1081 cgttcccaat gcccgccagg agttcgaccc gatcccgcga acggcgcttg acgaaacggt
1141 tgaggccaag cgggcgctga tccttgctgc cgagccggtc gtcgacgcgt tgatggagca
1201 gatgaacgac gcccccagga tgatcatcct gaacgacgaa cggggcgtcg tgctgctgaa
1261 ccagggaaac gacaccctcc ttgaagacgc ccgccgccgg gccgtgcggg tgggcgtctg
1321 ctgggacgaa cacgcccgag gcaccaatgc catgggaacc gcgctcgcgg agaggaggcc
1381 cgtagcgatc cacggcgcag agcactacct cgagtcgaat acgattttca cctgcaccgc
1441 ggcgccgatc tacgatccgt tcggcgagtt caccggaatt ctggatatca gcggatatgc
1501 gggggacatg ggcccggttc cgattccctt tgttcagatg gcggtgcaat tcatcgagaa
1561 tcagttgttc cgccagacct ttgccgattg cattctgctg cactttcatg tgcgccccga
1621 cttcgtcgga acgatgcgcg aagggatagc cgtgctgtcg cgcgagggaa ccatcgtctc
1681 gatgaaccgt gctgggctca agatcgcagg gctcaacctg gaggccgtcg ccgatcaccg
1741 tttcgattcc gtcttcgact tgaatttcgg ggcctttctc gaccacgtgc ggcagtccgc
1801 cttcggtctc gtccgcgtct cgctctacgg cggcgttcag gtctacgccc gagtggaacc
1861 gggcctgcgt gttccgccac gtccggccgc ccacgcccgc cctcctcggc cggcaccgcg
1921 gcctctggat tcgctggaca cgggcgacgc agcagtccgc ctcgcgattg accgcgcccg
1981 ccgcgcgatc ggccgcaacc tcagcatcct catccagggc gagacgggtg ccggcaagga
2041 agtgttcgcc aagcatctgc atgccgagag cccgagaagc aaggggccgt tcgttgccgt
2101 caattgcgcc gccatacccg agggtttgat cgagtccgag cttttcggat acgaagaagg
2161 ggccttcact ggcggacggc gcaaaggcaa catcggcaag gtcgcccagg cccacggcgg
2221 cacgttgttc ctcgacgaga tcggagacat ggcgccgggg ctgcagacca gactgctgcg
2281 ggttctacag gaccgggcag tgatgcccct cggcggtcgc gaacctatgc cggtcgacat
2341 tgcgctggta tgcgcgacgc accgcaattt gcgcagcctg atagcgcagg gccagttccg
2401 cgaagacctc tactaccgcc tgaacgggct ggcgatctcg ctgccacccc tgcgtcaacg
2461 tagcgacctt gccgccctgg ttaaccatat cctctttcag tgttgcgggg gcgagccgca
2521 ctacagcgtg agcccagaag tgatgacgct cttcaagcgg catgcgtggc ccggcaacct
2581 acgccaacta cataacgttc tcgatgcagc gcttgccatg ctcgacgacg gccatgtcat
2641 tgagccccat cacctccccg aagacttcgt catggaggtc gattcgggcc tccgaccgat
2701 cgaggaagac ggttcgacgg cggcgcatcg cgcgcgacag ccggcgtcgg gaagcggtcc
2761 tgccaaaaag ttgcaagatc tcgcgttgga tgccatcgag caggcgatcg agcaaaacga
2821 gggaaatata tcggtcgccg cacggcagtt gggggtcagc cggaccacga tctaccgcaa
2881 gctgaggcaa ctttcaccaa ccggttgtca ccgaccggca cattggagcc agtcgcggat
2941 cggcacatag ggatctgaag cttgggcccg aacaaaaact catctcagaa gaggatctga
3001 atagcgccgt cgaccatcat catcatcatc attgagttta aacggtctcc agcttggctg
3061 ttttggcgga tgagagaaga ttttcagcct gatacagatt aaatcagaac gcagaagcgg
3121 tctgataaaa cagaatttgc ctggcggcag tagcgcggtg gtcccacctg accccatgcc
3181 gaactcagaa gtgaaacgcc gtagcgccga tggtagtgtg gggtctcccc atgcgagagt
3241 agggaactgc caggcatcaa ataaaacgaa aggctcagtc gaaagactgg gcctttcgtt
3301 ttatctgttg tttgtcggtg aactaattat ctagactgca gttgatcggg cacgtaagag

3361 gttccaactt tcaccataat gaaataagat cactaccggg cgtatttttt gagttatcga
3421 gattttcagg agctaaggaa gctaaaatgg agaaaaaaat cactggatat accaccgttg
3481 atatatccca atggcatcgt aaagaacatt ttgaggcatt tcagtcagtt gctcaatgta
3541 cctataacca gaccgttcag ctggatatta cggccttttt aaagaccgta aagaaaaata
3601 agcacaagtt ttatccggcc tttattcaca ttcttgcccg cctgatgaat gctcatccgg
3661 aatttcgtat ggcaatgaaa gacggtgagc tggtgatatg ggatagtgtt cacccttgtt
3721 acaccgtttt ccatgagcaa actgaaacgt tttcatcgct ctggagtgaa taccacgacg
3781 atttccggca gtttctacac atatattcgc aagatgtggc gtgttacggt gaaaacctgg
3841 cctatttccc taaagggttt attgagaata tgtttttcgt ctcagccaat ccctgggtga
3901 gtttcaccag ttttgattta aacgtggcca atatggacaa cttcttcgcc cccgttttca
3961 ccatgggcaa atattatacg caaggcgaca aggtgctgat gccgctggcg attcaggttc
4021 atcatgccgt ttgtgatggc ttccatgtcg gcagaatgct taatgaatta caacagtact
4081 gcgatgagtg gcagggcggg gcgtaatttg atatcgagct cgcttggact cctgttgata
4141 gatccagtaa tgacctcaga actccatctg gatttgttca gaacgctcgg ttgccgccgg
4201 gcgtttttta ttggtgagaa tccaagcctc ggtgagaatc caagcctcga tcaacgtctc
4261 attttcgcca aaagttggcc cagggcttcc cggtatcaac agggacacca ggatttattt
4321 attctgcgaa gtgatcttcc gtcacaggta tttattcggc gcaaagtgcg tcgggtgatg
4381 ctgccaactt actgatttag tgtatgatgg tgtttttgag gtgctccagt ggcttctgtt
4441 tctatcagct gtccctcctg ttcagctact gacggggtgg tgcgtaacgg caaaagcacc
4501 gccggacatc agcgctagcg gagtgtatac tggcttacta tgttggcact gatgagggtg
4561 tcagtgaagt gcttcatgtg gcaggagaaa aaaggctgca ccggtgcgtc agcagaatat
4621 gtgatacagg atatattccg cttcctcgct cactgactcg ctacgctcgg tcgttcgact
4681 gcggcgagcg gaaatggctt acgaacgggg cggagatttc ctggaagatg ccaggaagat
4741 acttaacagg gaagtgagag ggccgcggca aagccgtttt tccataggct ccgccccccct
4801 gacaagcatc acgaaatctg acgctcaaat cagtggtggc gaaacccgac aggactataa
4861 agataccagg cgtttccccc tggcggctcc ctcgtgcgct ctcctgttcc tgcctttcgg
4921 tttaccggtg tcattccgct gttatggccg cgtttgtctc attccacgcc tgacactcag
4981 ttccgggtag gcagttcgct ccaagctgga ctgtatgcac gaacccccccg ttcagtccga
5041 ccgctgcgcc ttatccggta actatcgtct tgagtccaac ccggaaagac atgcaaaagc
5101 accactggca gcagccactg gtaattgatt tagaggagtt agtcttgaag tcatgcgccg
5161 gttaaggcta aactgaaagg acaagttttg gtgactgcgc tcctccaagc cagttacctc
5221 ggttcaaaga gttggtagct cagagaacct tcgaaaaacc gccctgcaag gcggtttttt
5281 cgttttcaga gcaagagatt acgcgcagac caaaacgatc tcaagaagat catcttatta
5341 atcagataaa atatttctag atttcagtgc aatttatctc ttcaaatgta gcacctgaag
5401 tcagccccat acgatataag ttgtaattct catgtttgac agcttatcat cgataagctt
5461 ccgatggcgc gccgagaggc tttacacttt atgcttccgg ct
//

A2.5. pBMO#36

```
LOCUS       (Bmo#36)\pBMOE1:       5778 bp    DNA     circular     21-FEB-2011
SOURCE
  ORGANISM
COMMENT     This file is created by Vector NTI
            http://www.invitrogen.com/
COMMENT     VNTDATE|510745607|
COMMENT     VNTDBDATE|582462583|
COMMENT     LSOWNER|
COMMENT     VNTNAME|(Bmo#36) pBMOE1:V2|
COMMENT     VNTAUTHORNAME|Jeffrey Dietrich|
COMMENT     VNTAUTHOREML|jadietrich@berkeley.edu|
FEATURES             Location/Qualifiers
     misc_feature    3703..3708
                     /vntifkey="21"
                     /label=XbaI\(1770)
     misc_feature    2971..3687
                     /vntifkey="21"
                     /label=GFP
     terminator      3709..3837
                     /vntifkey="43"
                     /label=dblTerm
     rep_origin      complement(3970..4652)
                     /ApEinfo_fwdcolor=gray50
                     /ApEinfo_revcolor=gray50
                     /vntifkey="33"
                     /label=ColE1
     CDS             complement(4791..5450)
                     /ApEinfo_fwdcolor=yellow
                     /ApEinfo_revcolor=yellow
                     /vntifkey="4"
                     /label=AmpR
     promoter        2523..2935
                     /vntifkey="30"
                     /label=pBMO_v1.5.2
     misc_feature    2854..2859
                     /vntifkey="21"
                     /label=-24
     misc_feature    2865..2871
                     /vntifkey="21"
                     /label=-12
     misc_feature    2883..2883
                     /vntifkey="21"
                     /label=Approx\+1
     promoter        2849..2871
                     /vntifkey="30"
```

/label=Sigma-54
misc_feature 2936..2970
/vntifkey="21"
/label=Syn\RBS-100000
misc_feature complement(375..2384)
/vntifkey="21"
/label=BmoR
terminator complement(1..368)
/vntifkey="43"
/label=TrrnB
primer_bind 5754..25
/vntifkey="28"
/label=D143/D144
promoter complement(2385..2522)
/vntifkey="30"
/label=Sigma70
primer_bind 350..399
/vntifkey="28"
/label=D149F/D150R
primer_bind 2498..2547
/vntifkey="28"
/label=D147F/D148R
primer 2356..2410
/vntifkey="27"
/label=D159
primer complement(2356..2410)
/vntifkey="27"
/label=D160R
primer 1569..1623
/vntifkey="27"
/label=D161
primer 1569..1623
/vntifkey="27"
/label=D162R
stem_loop 2619..2655
/vntifkey="39"
/label=Palindromic\sequence
stem_loop 2619..2632
/vntifkey="39"
/label=O1
stem_loop 2642..2655
/vntifkey="39"
/label=O2
BASE COUNT 1311 a 1609 c 1543 g 1315 t
ORIGIN
1 agttcaccga caaacaacag ataaaacgaa aggcccagtc tttcgactga gcctttcgtt

61 ttatttgatg cctggcagtt ccctactctc gcatggggag accccacact accatcggcg
121 ctacggcgtt tcacttctga gttcggcatg gggtcaggtg ggaccaccgc gctactgccg
181 ccaggcaaat tctgttttat cagaccgctt ctgcgttctg atttaatctg tatcaggctg
241 aaaatcttct ctcatccgcc aaaacagcca agctggagac cgtttaaact caatgatgat
301 gatgatgatg gtcgacggcg ctattcagat cctcttctga gatgagtttt tgttcgggcc
361 caagcttcag atccctatgt gccgatccgc gactggctcc aatgtgccgg tcggtgacaa
421 ccggttggtg aaagttgcct cagcttgcgg tagatcgtgg tccggctgac ccccaactgc
481 cgtgcggcga ccgatatatt tccctcgttt tgctcgatcg cctgctcgat ggcatccaac
541 gcgagatctt gcaacttttt ggcaggaccg cttcccgacg ccggctgtcg cgcgcgatgc
601 gccgccgtcg aaccgtcttc ctcgatcggt cggaggcccg aatcgacctc catgacgaag
661 tcttcgggga ggtgatgggg ctcaatgaca tggccgtcgt cgagcatggc aagcgctgca
721 tcgagaacgt tatgtagttg gcgtaggttg ccgggccacg catgccgctt gaagagcgtc
781 atcacttctg ggctcacgct gtagtgcggc tcgcccccgc aacactgaaa gaggatatgg
841 ttaaccaggg cggcaaggtc gctacgttga cgcaggggtg gcagcgagat cgccagcccg
901 ttcaggcggt agtagaggtc ttcgcggaac tggccctgcg ctatcaggct gcgcaaattg
961 cggtgcgtcg cgcataccag cgcaatgtcg accggcatag gttcgcgacc gccgaggggc
1021 atcactgccc ggtcctgtag aacccgcagc agtctggtct gcagccccgg cgccatgtct
1081 ccgatctcgt cgaggaacaa cgtgccgccg tgggcctggg cgaccttgcc gatgttgcct
1141 ttgcgccgtc cgccagtgaa ggccccttct tcgtatccga aaagctcgga ctcgatcaaa
1201 ccctcgggta tggcggcgca attgacggca acgaacggcc ccttgcttct cgggctctcg
1261 gcatgcagat gcttggcgaa cacttccttg ccggcacccg tctcgccctg gatgaggatg
1321 ctgaggttgc ggccgatcgc gcggcgggcg cggtcaatcg cgaggcggac tgctgcgtcg
1381 cccgtgtcca gcgaatccag aggccgcggt gccggccgag gagggcgggc gtgggcggcc
1441 ggacgtggcg gaacacgcag gcccggttcc actcgggcgt agacctgaac gccgccgtag
1501 agcgagacgc ggacgagacc gaaggcggac tgccgcacgt ggtcgagaaa ggccccgaaa
1561 ttcaagtcga agacggaatc gaaacggtga tcggcgacgg cctccaggtt gagccctgcg
1621 atcttgagcc cagcacggtt catcgagacg atggttccct cgcgcgacag cacggctatc
1681 ccttcgcgca tcgttccgac gaagtcgggg cgcacatgaa agtgcagcag aatgcaatcg
1741 gcaaaggtct ggcggaacaa ctgattctcg atgaattgca ccgccatctg aacaaaggga
1801 atcggaaccg ggcccatgtc ccccgcatat ccgctgatat ccagaattcc ggtgaactcg
1861 ccgaacggat cgtagatcgg cgccgcggtg caggtgaaaa tcgtattcga ctcgaggtag
1921 tgctctgcgc cgtggatcgc tacgggcctc ctctccgcga gcgcggttcc catggcattg
1981 gtgcctcggg cgtgttcgtc ccagcagacg cccacccgca cggcccggcg gcgggcgtct
2041 tcaaggaggg tgtcgtttcc ctggttcagc agcacgacgc cccgttcgtc gttcaggatg
2101 atcatcctgg gggcgtcgtt catctgctcc atcaacgcgt cgacgaccgg ctcggcagca
2161 aggatcagcg cccgcttggc ctcaaccgtt tcgtcaagcg ccgttcgcgg gatcgggtcg
2221 aactcctggc gggcattggg aacgacacct tcagcacgac accgggtcca cgagcgcaaa
2281 acgacatcag ccactttccc cggctgacac tcgttgccgt cccagaccgc tctgcgcatc
2341 gacgcgactg tctccagccg cgcgaactct tgcatcttgg acataccgtc tcctcattca
2401 cctctgtcgg gcgaattgct tatcttcgcg agtctctcct cgctgtcgag agctttagca
2461 aaaccataat agcacgtggg aaattttggt ggtatctgcc cgctcaaggt cacctcaagg
2521 tcccacagat agtaggtgct gcggctgctc atgctcctgt cgcggtagcg cgctgttacg
2581 cgaccgcccc cggacctcgg cggacagcgc ggaagattgg aaacagcccg agcgtgcgtg
2641 cctcgggctg catccttgcc acacccaacc ggattcgtcg gaccgctcga cattcgcgtt
2701 cgctcccgcg gcgccgcggg tgtaccgttg cgttacagat gtacccttct ttaacgtgta
2761 acacacgcct ggagcggcca agagccccgc accttgcggc gcgtcttccc caggggccca

2821 ccggttgcgg ccttttgctg cgaccgtcca tgctggcacg acacttgctg aaagcgttag
2881 agcggaatcg gtccgatgga gcattcgaag ccgctaccga cagcagaaca cacaatattt
2941 tattacgcag aatagataag gagggcgaga atgagtaaag gagaagaact tttcactgga
3001 gttgtcccaa ttcttgttga attagatggt gatgttaatg ggcacaaatt ttctgtcagt
3061 ggagagggtg aaggtgatgc aacatacgga aaacttaccc ttaaatttat ttgcactact
3121 ggaaaactac ctgttccgtg gccaacactt gtcactactt tctcttatgg tgttcaatgc
3181 ttttcccgtt atccggatca catgaaacgg catgactttt tcaagagtgc catgcccgaa
3241 ggttatgtac aggaacgcac tatatctttc aaagatgacg ggaactacaa gacgcgtgct
3301 gaagtcaagt ttgaaggtga taccccttgtt aatcgtatcg agttaaaagg tattgatttt
3361 aaagaagatg gaaacattct cggacacaaa ctggagtaca actataactc acacaatgta
3421 tacatcacgg cagacaaaca aaagaatgga atcaaagcta acttcaaaat tcgccacaac
3481 attgaagatg gctccgttca actagcagac cattatcaac aaaatactcc aattggcgat
3541 ggccctgtcc ttttaccaga caaccattac ctgtccacac aatctgccct ttcgaaagat
3601 cccaacgaaa agcgtgacca catggtcctt cttgagtttg taactgctgc tgggattaca
3661 catggcatgg atgagctcta caaataagga tcctaactcg agtctagacc aggcatcaaa
3721 taaaacgaaa ggctcagtcg aaagactggg cctttcgttt tatctgttgt ttgtcggtga
3781 acgctctcta ctagagtcac actggctcac cttcgggtgg gcctttctgc gtttatacct
3841 aggcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac
3901 agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa
3961 ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg acgagcatca
4021 caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc
4081 gtttcccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata
4141 cctgtccgcc tttctccctt cgggaagcgt ggcgctttct caatgctcac gctgtaggta
4201 tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca
4261 gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga
4321 cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg
4381 tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg
4441 tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg
4501 caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag
4561 aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa
4621 cgaaaactca cgttaaggga ttttggtcat gactagtgct tggattctca ccaataaaaa
4681 acgcccggcg gcaaccgagc gttctgaaca aatccagatg gagttctgag gtcattactg
4741 gatctatcaa caggagtcca agcgagctcg taaacttggt ctgacagtta ccaatgctta
4801 atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc
4861 cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg
4921 ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga
4981 agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt
5041 tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt
5101 gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc
5161 caacgatcaa ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc
5221 ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca
5281 gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag
5341 tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg
5401 tcaatacggg ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa
5461 cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa
5521 cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga

5581 gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga
5641 atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg
5701 agcggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt
5761 ccccgaaaag tgccacct
//

A2.6. pBMO#40

```
LOCUS       (Bmo#40)\pBMOS1:        7194 bp    DNA     circular     9-JUN-2009
SOURCE
  ORGANISM
COMMENT     This file is created by Vector NTI
            http://www.invitrogen.com/
COMMENT     VNTDATE|520598609|
COMMENT     VNTDBDATE|520601468|
COMMENT     LSOWNER|
COMMENT     VNTNAME|(Bmo#40) pBMOS1:V2|
COMMENT     VNTAUTHORNAME|Jeffrey Dietrich|
COMMENT     VNTAUTHOREML|jadietrich@berkeley.edu|
FEATURES             Location/Qualifiers
     misc_feature    3703..3708
                     /vntifkey="21"
                     /label=XbaI\(1770)
     misc_feature    2971..3687
                     /vntifkey="21"
                     /label=GFP
     terminator      3709..3837
                     /vntifkey="43"
                     /label=dblTerm
     CDS             complement(6207..6866)
                     /ApEinfo_fwdcolor=yellow
                     /ApEinfo_revcolor=yellow
                     /vntifkey="4"
                     /label=AmpR
     promoter        2523..2935
                     /vntifkey="30"
                     /label=pBMO_v1.5.2
     misc_feature    2854..2859
                     /vntifkey="21"
                     /label=-24
     misc_feature    2865..2871
                     /vntifkey="21"
                     /label=-12
     misc_feature    2883..2883
                     /vntifkey="21"
                     /label=Approx\+1
     promoter        2849..2871
                     /vntifkey="30"
                     /label=Sigma-54
     misc_feature    2936..2970
                     /vntifkey="21"
                     /label=Syn\RBS-100000
     misc_feature    complement(375..2384)
```

/vntifkey="21"
/label=BmoR
terminator complement(1..368)
/vntifkey="43"
/label=TrrnB
primer_bind 7170..25
/vntifkey="28"
/label=D143/D144
promoter complement(2385..2522)
/vntifkey="30"
/label=Sigma70
primer_bind 350..399
/vntifkey="28"
/label=D149F/D150R
primer_bind 2498..2547
/vntifkey="28"
/label=D147F/D148R
rep_origin complement(3839..6067)
/vntifkey="33"
/label=pSC101**\ori
misc_feature complement(5305..5312)
/vntifkey="21"
/label=Origin\start\codon
/note="See paper for description of hairpins and protein coding region. SpeI lies just before one of the hairpins at positin 400 (in paper) http://www.pnas.org/content/80/21/6557.full.pdf"
mutation 5381..5381
/vntifkey="62"
/label=T>A
primer_bind 3834..3863
/vntifkey="28"
/label=F99
primer_bind complement(6046..6078)
/vntifkey="28"
/label=F100
primer_bind 5369..5408
/vntifkey="28"
/label=F101F
misc_feature complement(5358..5394)
/vntifkey="21"
/label=F102R
BASE COUNT 1668 a 1855 c 1766 g 1905 t
ORIGIN
1 agttcaccga caaacaacag ataaaacgaa aggcccagtc tttcgactga gcctttcgtt
61 ttatttgatg cctggcagtt ccctactctc gcatggggag accccacact accatcggcg
121 ctacggcgtt tcacttctga gttcggcatg gggtcaggtg ggaccaccgc gctactgccg

181 ccaggcaaat tctgttttat cagaccgctt ctgcgttctg atttaatctg tatcaggctg
241 aaaatcttct ctcatccgcc aaaacagcca agctggagac cgtttaaact caatgatgat
301 gatgatgatg gtcgacggcg ctattcagat cctcttctga gatgagtttt tgttcgggcc
361 caagcttcag atccctatgt gccgatccgc gactggctcc aatgtgccgg tcggtgacaa
421 ccggttggtg aaagttgcct cagcttgcgg tagatcgtgg tccggctgac ccccaactgc
481 cgtgcggcga ccgatatatt tccctcgttt tgctcgatcg cctgctcgat ggcatccaac
541 gcgagatctt gcaacttttt ggcaggaccg cttcccgacg ccggctgtcg cgcgcgatgc
601 gccgccgtcg aaccgtcttc ctcgatcggt cggaggcccg aatcgacctc catgacgaag
661 tcttcgggga ggtgatgggg ctcaatgaca tggccgtcgt cgagcatggc aagcgctgca
721 tcgagaacgt tatgtagttg gcgtaggttg ccgggccacg catgccgctt gaagagcgtc
781 atcacttctg ggctcacgct gtagtgcggc tcgccccccgc aacactgaaa gaggatatgg
841 ttaaccaggg cggcaaggtc gctacgttga cgcaggggtg gcagcgagat cgccagcccg
901 ttcaggcggt agtagaggtc ttcgcggaac tggccctgcg ctatcaggct gcgcaaattg
961 cggtgcgtcg cgcataccag cgcaatgtcg accggcatag gttcgcgacc gccgaggggc
1021 atcactgccc ggtcctgtag aacccgcagc agtctggtct gcagccccgg cgccatgtct
1081 ccgatctcgt cgaggaacaa cgtgccgccg tgggcctggg cgaccttgcc gatgttgcct
1141 ttgcgccgtc cgccagtgaa ggccccttct tcgtatccga aaagctcgga ctcgatcaaa
1201 ccctcgggta tggcggcgca attgacggca acgaacggcc ccttgcttct cgggctctcg
1261 gcatgcagat gcttggcgaa cacttccttg ccggcacccg tctcgccctg gatgaggatg
1321 ctgaggttgc ggccgatcgc gcggcgggcg cggtcaatcg cgaggcggac tgctgcgtcg
1381 cccgtgtcca gcgaatccag aggccgcggt gccggccgag gagggcgggc gtgggcggcc
1441 ggacgtggcg gaacacgcag gcccggttcc actcgggcgt agacctgaac gccgccgtag
1501 agcgagacgc ggacgagacc gaaggcggac tgccgcacgt ggtcgagaaa ggccccgaaa
1561 ttcaagtcga agacggaatc gaaacggtga tcggcgacgg cctccaggtt gagccctgcg
1621 atcttgagcc cagcacggtt catcgagacg atggttccct cgcgcgacag cacggctatc
1681 ccttcgcgca tcgttccgac gaagtcgggg cgcacatgaa agtgcagcag aatgcaatcg
1741 gcaaaggtct ggcggaacaa ctgattctcg atgaattgca ccgccatctg aacaaaggga
1801 atcggaaccg ggcccatgtc ccccgcatat ccgctgatat ccagaattcc ggtgaactcg
1861 ccgaacggat cgtagatcgg cgccgcggtg caggtgaaaa tcgtattcga ctcgaggtag
1921 tgctctgcgc cgtggatcgc tacgggcctc ctctccgcga gcgcggttcc catggcattg
1981 gtgcctcggg cgtgttcgtc ccagcagacg cccacccgca cggcccggcg gcgggcgtct
2041 tcaaggaggg tgtcgtttcc ctggttcagc agcacgacgc cccgttcgtc gttcaggatg
2101 atcatcctgg gggcgtcgtt catctgctcc atcaacgcgt cgacgaccgg ctcggcagca
2161 aggatcagcg cccgcttggc ctcaaccgtt tcgtcaagcg ccgttcgcgg gatcgggtcg
2221 aactcctggc gggcattggg aacgacacct tcagcacgac accgggtcca cgagcgcaaa
2281 acgacatcag ccactttccc cggctgacac tcgttgccgt cccagaccgc tctgcgcatc
2341 gacgcgactg tctccagccg cgcgaactct tgcatcttgg acataccgtc tcctcattca
2401 cctctgtcgg gcgaattgct tatcttcgcg agtctctcct cgctgtcgag agctttagca
2461 aaaccataat agcacgtggg aaattttggt ggtatctgcc cgctcaaggt cacctcaagg
2521 tcccacagat agtaggtgct gcggctgctc atgctcctgt cgcggtagcg cgctgttacg
2581 cgaccgcccc cggacctcgg cggacagcgc ggaagattgg aaacagcccg agcgtgcgtg
2641 cctcgggctg catccttgcc acacccaacc ggattcgtcg gaccgctcga cattcgcgtt
2701 cgctcccgcg gcgccgcggg tgtaccgttg cgttacagat gtaccttct ttaacgtgta
2761 acacacgcct ggagcggcca agagccccgc accttgcggc gcgtcttccc caggggccca
2821 ccggttgcgg ccttttgctg cgaccgtcca tgctggcacg acacttgctg aaagcgttag
2881 agcggaatcg gtccgatgga gcattcgaag ccgctaccga cagcagaaca cacaatattt

2941 tattacgcag aatagataag gagggcgaga atgagtaaag gagaagaact tttcactgga
3001 gttgtcccaa ttcttgttga attagatggt gatgttaatg ggcacaaatt ttctgtcagt
3061 ggagagggtg aaggtgatgc aacatacgga aaacttaccc ttaaatttat ttgcactact
3121 ggaaaactac ctgttccgtg gccaacactt gtcactactt tctcttatgg tgttcaatgc
3181 ttttcccgtt atccggatca catgaaacgg catgactttt tcaagagtgc catgcccgaa
3241 ggttatgtac aggaacgcac tatatctttc aaagatgacg ggaactacaa gacgcgtgct
3301 gaagtcaagt ttgaaggtga taccttgtt aatcgtatcg agttaaaagg tattgatttt
3361 aaagaagatg gaaacattct cggacacaaa ctggagtaca actataactc acacaatgta
3421 tacatcacgg cagacaaaca aaagaatgga atcaaagcta acttcaaaat tcgccacaac
3481 attgaagatg gctccgttca actagcagac cattatcaac aaaatactcc aattggcgat
3541 ggccctgtcc ttttaccaga caaccattac ctgtccacac aatctgccct ttcgaaagat
3601 cccaacgaaa agcgtgacca catggtcctt cttgagtttg taactgctgc tgggattaca
3661 catggcatgg atgagctcta caaataagga tcctaactcg agtctagacc aggcatcaaa
3721 taaaacgaaa ggctcagtcg aaagactggg cctttcgttt tatctgttgt ttgtcggtga
3781 acgctctcta ctagagtcac actggctcac cttcgggtgg gcctttctgc gtttatacct
3841 agggtacggg ttttgctgcc cgcaaacggg ctgttctggt gttgctagtt tgttatcaga
3901 atcgcagatc cggcttcagc cggtttgccg gctgaaagcg ctatttcttc cagaattgcc
3961 atgatttttt ccccacggga ggcgtcactg gctcccgtgt tgtcggcagc tttgattcga
4021 taagcagcat cgcctgtttc aggctgtcta tgtgtgactg ttgagctgta acaagttgtc
4081 tcaggtgttc aatttcatgt tctagttgct ttgttttact ggtttcacct gttctattag
4141 gtgttacatg ctgttcatct gttacattgt cgatctgttc atggtgaaca gctttgaatg
4201 caccaaaaac tcgtaaaagc tctgatgtat ctatcttttt tacaccgttt tcatctgtgc
4261 atatggacag ttttcccttt gatatgtaac ggtgaacagt tgttctactt ttgtttgtta
4321 gtcttgatgc ttcactgata gatacaagag ccataagaac ctcagatcct tccgtattta
4381 gccagtatgt tctctagtgt ggttcgttgt ttttgcgtga gccatgagaa cgaaccattg
4441 agatcatact tactttgcat gtcactcaaa aattttgcct caaaactggt gagctgaatt
4501 tttgcagtta aagcatcgtg tagtgttttt cttagtccgt tatgtaggta ggaatctgat
4561 gtaatggttg ttggtatttt gtcaccattc atttttatct ggttgttctc aagttcggtt
4621 acgagatcca tttgtctatc tagttcaact tggaaaatca acgtatcagt cgggcggcct
4681 cgcttatcaa ccaccaattt catattgctg taagtgttta aatctttact tattggtttc
4741 aaaacccatt ggttaagcct tttaaactca tggtagttat tttcaagcat taacatgaac
4801 ttaaattcat caaggctaat ctctatattt gccttgtgag ttttcttttg tgttagttct
4861 tttaataacc actcataaat cctcatagag tatttgtttt caaaagactt aacatgttcc
4921 agattatatt ttatgaattt ttttaactgg aaaagataag gcaatatctc ttcactaaaa
4981 actaattcta atttttcgct tgagaacttg gcatagtttg tccactggaa aatctcaaag
5041 cctttaacca aaggattcct gatttccaca gttctcgtca tcagctctct ggttgcttta
5101 gctaatacac cataagcatt ttccctactg atgttcatca tctgagcgta ttggttataa
5161 gtgaacgata ccgtccgttc tttccttgta gggttttcaa tcgtggggtt gagtagtgcc
5221 acacagcata aaattagctt ggtttcatgc tccgttaagt catagcgact aatcgctagt
5281 tcatttgctt tgaaaacaac taattcagac atacatctca attggtctag gtgattttaa
5341 tcactatacc aattgagatg ggctagtcaa tgataattac aagtcctttt cccgggtgat
5401 ctgggtatct gtaaattctg ctagaccttt gctggaaaac ttgtaaattc tgctagaccc
5461 tctgtaaatt ccgctagacc tttgtgtgtt ttttttgttt atattcaagt ggttataatt
5521 tatagaataa agaaagaata aaaaaagata aaaagaatag atcccagccc tgtgtataac
5581 tcactacttt agtcagttcc gcagtattac aaaaggatgt cgcaaacgct gtttgctcct
5641 ctacaaaaca gaccttaaaa ccctaaaggc ttaagtagca ccctcgcaag ctcgggcaaa

5701 tcgctgaata ttccttttgt ctccgaccat caggcacctg agtcgctgtc tttttcgtga
5761 cattcagttc gctgcgctca cggctctggc agtgaatggg ggtaaatggc actacaggcg
5821 ccttttatgg attcatgcaa ggaaactacc cataatacaa gaaaagcccg tcacgggctt
5881 ctcagggcgt tttatggcgg gtctgctatg tggtgctatc tgactttttg ctgttcagca
5941 gttcctgccc tctgattttc cagtctgacc acttcggatt atcccgtgac aggtcattca
6001 gactggctaa tgcacccagt aaggcagcgg tatcatcaac aggcttaccc gtcttactgt
6061 ccctagtact agtgcttgga ttctcaccaa taaaaaacgc ccggcggcaa ccgagcgttc
6121 tgaacaaatc cagatggagt tctgaggtca ttactggatc tatcaacagg agtccaagcg
6181 agctcgtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg
6241 atctgtctat ttcgttcatc catagttgcc tgactccccg tcgtgtagat aactacgata
6301 cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc acgctcaccg
6361 gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggtcct
6421 gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag agtaagtagt
6481 tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt ggtgtcacgc
6541 tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga
6601 tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt
6661 aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc
6721 atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa
6781 tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca
6841 catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca
6901 aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc caactgatct
6961 tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc
7021 gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa
7081 tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt
7141 tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acct
//

A2.7. pBMO#41

```
LOCUS       (Bmo#41)\pBMOE1:       7023 bp    DNA     circular     10-MAR-2010
SOURCE
  ORGANISM
COMMENT     This file is created by Vector NTI
            http://www.invitrogen.com/
COMMENT     VNTDATE|544800387|
COMMENT     VNTDBDATE|548329116|
COMMENT     LSOWNER|
COMMENT     VNTNAME|(Bmo#41) pBMOE1:V3|
COMMENT     VNTAUTHORNAME|Jeffrey Dietrich|
COMMENT     VNTAUTHOREML|jadietrich@berkeley.edu|
FEATURES             Location/Qualifiers
     N_region        4948..4953
                     /vntifkey="24"
                     /label=XbaI\(1770)
     misc_feature    4216..4932
                     /vntifkey="21"
                     /label=GFP
     terminator      4954..5082
                     /vntifkey="43"
                     /label=dblTerm
     rep_origin      complement(5215..5897)
                     /ApEinfo_fwdcolor=gray50
                     /ApEinfo_revcolor=gray50
                     /vntifkey="33"
                     /label=ColE1
     CDS             complement(6036..6695)
                     /ApEinfo_fwdcolor=yellow
                     /ApEinfo_revcolor=yellow
                     /vntifkey="4"
                     /label=AmpR
     promoter        2523..2935
                     /vntifkey="30"
                     /label=pBMO_v1.5.2
     misc_feature    2854..2859
                     /vntifkey="21"
                     /label=-24
     misc_feature    2865..2871
                     /vntifkey="21"
                     /label=-12
     misc_feature    2883..2883
                     /vntifkey="21"
                     /label=Approx\+1
     promoter        2849..2871
                     /vntifkey="30"
```

/label=Sigma-54
misc_feature 2936..2970
/vntifkey="21"
/label=Syn\RBS-100000
misc_feature complement(375..2384)
/vntifkey="21"
/label=BmoR
terminator complement(1..368)
/vntifkey="43"
/label=TrrnB
primer_bind 6999..25
/vntifkey="28"
/label=D143F/D144R
promoter complement(2385..2522)
/vntifkey="30"
/label=Sigma70
primer_bind 350..399
/vntifkey="28"
/label=D149F/D150R
primer_bind 2498..2547
/vntifkey="28"
/label=D147F/D148R
CDS 2971..4167
/gene="tetA"
/product="TetA"
/SECDrawAs="Gene"
/vntifkey="4"
/label=tetA
primer_bind 3830..3859
/vntifkey="28"
/label=Z38
repeat_region 4168..4215
/vntifkey="34"
/label=Linker
primer 4196..4245
/vntifkey="27"
/label=D163
primer complement(4166..4215)
/vntifkey="27"
/label=D164R
primer 4131..4190
/vntifkey="27"
/label=D165F
/note="D165F"
primer complement(2912..2961)
/vntifkey="27"

/label=D89R
primer complement(4101..4160)
/vntifkey="27"
/label=D166R
primer 2941..3000
/vntifkey="27"
/label=D167F/D168R
primer 4903..4962
/vntifkey="27"
/label=D169/D170
primer complement(2624..2646)
/vntifkey="27"
/label=D84R
misc_feature 4933..4938
/vntifkey="21"
/label=BamHI
primer 2523..2545
/vntifkey="27"
/label=D78
primer_bind complement(2989..3015)
/vntifkey="28"
/label=Z101R
primer_bind 6974..7023
/vntifkey="28"
/label=D180F
primer_bind complement(6949..6998)
/vntifkey="28"
/label=D181R
primer_bind 6924..6973
/vntifkey="28"
/label=D182F
primer_bind complement(2911..2970)
/vntifkey="28"
/label=D183R
primer 2881..2940
/vntifkey="27"
/label=D184F
primer_bind 2971..3030
/vntifkey="28"
/label=D185F
primer complement(3001..3060)
/vntifkey="27"
/label=D186R
iDNA 3900..4585
/vntifkey="13"
/label=Z38

```
     iDNA          6700..550
                   /vntifkey="13"
                   /label=Seq\S1
     iDNA          1985..2940
                   /vntifkey="13"
                   /label=Seq\Z101
     iDNA          4425..5150
                   /vntifkey="13"
                   /label=S5
     iDNA          1610..2475
                   /vntifkey="13"
                   /label=Seq\Z34
BASE COUNT     1483 a      1997 c      1949 g      1594 t
ORIGIN
        1 agttcaccga caaacaacag ataaaacgaa aggcccagtc tttcgactga gcctttcgtt
       61 ttatttgatg cctggcagtt ccctactctc gcatggggag accccacact accatcggcg
      121 ctacggcgtt tcacttctga gttcggcatg gggtcaggtg ggaccaccgc gctactgccg
      181 ccaggcaaat tctgttttat cagaccgctt ctgcgttctg atttaatctg tatcaggctg
      241 aaaatcttct ctcatccgcc aaaacagcca agctggagac cgtttaaact caatgatgat
      301 gatgatgatg gtcgacggcg ctattcagat cctcttctga gatgagtttt tgttcgggcc
      361 caagcttcag atccctatgt gccgatccgc gactggctcc aatgtgccgg tcggtgacaa
      421 ccggttggtg aaagttgcct cagcttgcgg tagatcgtgg tccggctgac ccccaactgc
      481 cgtgcggcga ccgatatatt tccctcgttt tgctcgatcg cctgctcgat ggcatccaac
      541 gcgagatctt gcaacttttt ggcaggaccg cttcccgacg ccggctgtcg cgcgcgatgc
      601 gccgccgtcg aaccgtcttc ctcgatcggt cggaggcccg aatcgacctc catgacgaag
      661 tcttcgggga ggtgatgggg ctcaatgaca tggccgtcgt cgagcatggc aagcgctgca
      721 tcgagaacgt tatgtagttg gcgtaggttg ccgggccacg catgccgctt gaagagcgtc
      781 atcacttctg ggctcacgct gtagtgcggc tcgcccccgc aacactgaaa gaggatatgg
      841 ttaaccaggg cggcaaggtc gctacgttga cgcaggggtg gcagcgagat cgccagcccg
      901 ttcaggcggt agtagaggtc ttcgcggaac tggccctgcg ctatcaggct gcgcaaattg
      961 cggtgcgtcg cgcataccag cgcaatgtcg accggcatag gttcgcgacc gccgaggggc
     1021 atcactgccc ggtcctgtag aacccgcagc agtctggtct gcagccccgg cgccatgtct
     1081 ccgatctcgt cgaggaacaa cgtgccgccg tgggcctggg cgaccttgcc gatgttgcct
     1141 ttgcgccgtc cgccagtgaa ggccccttct tcgtatccga aaagctcgga ctcgatcaaa
     1201 ccctcgggta tggcggcgca attgacggca acgaacggcc ccttgcttct cgggctctcg
     1261 gcatgcagat gcttggcgaa cacttccttg ccggcacccg tctcgccctg gatgaggatg
     1321 ctgaggttgc ggccgatcgc gcggcgggcg cggtcaatcg cgaggcggac tgctgcgtcg
     1381 cccgtgtcca gcgaatccag aggccgcggt gccggccgag gagggcgggc gtgggcggcc
     1441 ggacgtggcg gaacacgcag gcccggttcc actcgggcgt agacctgaac gccgccgtag
     1501 agcgagacgc ggacgagacc gaaggcggac tgccgcacgt ggtcgagaaa ggccccgaaa
     1561 ttcaagtcga agacggaatc gaaacggtga tcggcgacgg cctccaggtt gagccctgcg
     1621 atcttgagcc cagcacggtt catcgagacg atggttccct cgcgcgacag cacggctatc
     1681 ccttcgcgca tcgttccgac gaagtcgggg cgcacatgaa agtgcagcag aatgcaatcg
     1741 gcaaaggtct ggcggaacaa ctgattctcg atgaattgca ccgccatctg aacaaaggga
     1801 atcggaaccg ggcccatgtc ccccgcatat ccgctgatat ccagaattcc ggtgaactcg
     1861 ccgaacggat cgtagatcgg cgccgcggtg caggtgaaaa tcgtattcga ctcgaggtag
```

1921 tgctctgcgc cgtggatcgc tacgggcctc ctctccgcga gcgcggttcc catggcattg
1981 gtgcctcggg cgtgttcgtc ccagcagacg cccacccgca cggcccggcg gcgggcgtct
2041 tcaaggaggg tgtcgtttcc ctggttcagc agcacgacgc cccgttcgtc gttcaggatg
2101 atcatcctgg gggcgtcgtt catctgctcc atcaacgcgt cgacgaccgg ctcggcagca
2161 aggatcagcg cccgcttggc ctcaaccgtt tcgtcaagcg ccgttcgcgg gatcgggtcg
2221 aactcctggc gggcattggg aacgacacct tcagcacgac accgggtcca cgagcgcaaa
2281 acgacatcag ccactttccc cggctgacac tcgttgccgt cccagaccgc tctgcgcatc
2341 gacgcgactg tctccagccg cgcgaactct tgcatcttgg acataccgtc tcctcattca
2401 cctctgtcgg gcgaattgct tatcttcgcg agtctctcct cgctgtcgag agctttagca
2461 aaaccataat agcacgtggg aaattttggt ggtatctgcc cgctcaaggt cacctcaagg
2521 tcccacagat agtaggtgct gcggctgctc atgctcctgt cgcggtagcg cgctgttacg
2581 cgaccgcccc cggacctcgg cggacagcgc ggaagattgg aaacagcccg agcgtgcgtg
2641 cctcgggctg catccttgcc acacccaacc ggattcgtcg gaccgctcga cattcgcgtt
2701 cgctcccgcg gcgccgcggg tgtaccgttg cgttacagat gtacccttct ttaacgtgta
2761 acacacgcct ggagcggcca agagccccgc accttgcggc gcgtcttccc caggggccca
2821 ccggttgcgg ccttttgctg cgaccgtcca tgctggcacg acacttgctg aaagcgttag
2881 agcggaatcg gtccgatgga gcattcgaag ccgctaccga cagcagaaca cacaatattt
2941 tattacgcag aatagataag gagggcgaga atgaaaccca acataccct gatcgtaatt
3001 ctgagcactg tcgcgctcga cgctgtcggc atcggcctga ttatgccggt gctgccgggc
3061 ctcctgcgcg atctggttca ctcgaacgac gtcaccgccc actatggcat tctgctggcg
3121 ctgtatgcgt tggtgcaatt tgcctgcgca cctgtgctgg gcgcgctgtc ggatcgtttc
3181 gggcggcggc caatcttgct cgtctcgctg gccggcgcca ctgtcgacta cgccatcatg
3241 gcgacagcgc ctttcctttg ggttctctat atcgggcgga tcgtggccgg catcaccggg
3301 gcgactgggg cggtagccgg cgcttatatt gccgatatca ctgatggcga tgagcgcgcg
3361 cggcacttcg gcttcatgag cgcctgtttc gggttcggga tggtcgcggg acctgtgctc
3421 ggtgggctga tgggcggttt ctccccccac gctccgttct tcgccgcggc agccttgaac
3481 ggcctcaatt tcctgacggg ctgtttcctt ttgccggagt cgcacaaagg cgaacgccgg
3541 ccgttacgcc gggaggctct caacccgctc gcttcgttcc ggtgggcccg gggcatgacc
3601 gtcgtcgccg ccctgatggc ggtcttcttc atcatgcaac ttgtcggaca ggtgccggcc
3661 gcgctttggg tcattttcgg cgaggatcgc tttcactggg acgcgaccac gatcggcatt
3721 tcgcttgccg catttggcat tctgcattca ctcgcccagg caatgatcac cggccctgta
3781 gccgcccggc tcggcgaaag gcgggcactc atgctcggaa tgattgccga cggcacaggc
3841 tacatcctgc ttgccttcgc gacacgggga tggatggcgt tcccgatcat ggtcctgctt
3901 gcttcgggtg gcatcggaat gccggcgctg caagcaatgt tgtccaggca ggtggatgag
3961 gaacgtcagg ggcagctgca aggctcactg gcggcgctca ccagcctgac ctcgatcgtc
4021 ggacccctcc tcttcacggc gatctatgcg gcttctataa caacgtggaa cgggtgggca
4081 tggattgcag gcgctgccct ctacttgctc tgcctgccgg cgctgcgtcg cgggctttgg
4141 agcggcgcag ggcaacgagc cgatcgcggc ggtggaagcg gcggcggctc cggtggtggt
4201 tctggaggcg gttctatgag taaaggagaa gaacttttca ctggagttgt cccaattctt
4261 gttgaattag atggtgatgt taatgggcac aaattttctg tcagtggaga gggtgaaggt
4321 gatgcaacat acggaaaact tacccttaaa tttatttgca ctactggaaa actacctgtt
4381 ccgtggccaa cacttgtcac tactttctct tatggtgttc aatgcttttc ccgttatccg
4441 gatcacatga aacggcatga ctttttcaag agtgccatgc ccgaaggtta tgtacaggaa
4501 cgcactatat ctttcaaaga tgacgggaac tacaagacgc gtgctgaagt caagtttgaa
4561 ggtgataccc ttgttaatcg tatcgagtta aaaggtattg attttaaaga agatggaaac
4621 attctcggac acaaactgga gtacaactat aactcacaca atgtatacat cacggcagac

4681 aaacaaaaga atggaatcaa agctaacttc aaaattcgcc acaacattga agatggctcc
4741 gttcaactag cagaccatta tcaacaaaat actccaattg gcgatggccc tgtcctttta
4801 ccagacaacc attacctgtc cacacaatct gccctttcga aagatcccaa cgaaaagcgt
4861 gaccacatgg tccttcttga gtttgtaact gctgctggga ttacacatgg catggatgag
4921 ctctacaaat aaggatccta actcgagtct agaccaggca tcaaataaaa cgaaaggctc
4981 agtcgaaaga ctgggccttt cgttttatct gttgtttgtc ggtgaacgct ctctactaga
5041 gtcacactgg ctcaccttcg ggtgggcctt tctgcgttta tacctaggcg ttcggctgcg
5101 gcgagcggta tcagctcact caaaggcggt aatacggtta tccacagaat caggggataa
5161 cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc
5221 gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc
5281 aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag
5341 ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct
5401 cccttcggga agcgtggcgc tttctcaatg ctcacgctgt aggtatctca gttcggtgta
5461 ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccccc gttcagcccg accgctgcgc
5521 cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc
5581 agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt
5641 gaagtggtgg cctaactacg gctacactag aaggacagta tttggtatct gcgctctgct
5701 gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc
5761 tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca
5821 agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta
5881 agggattttg gtcatgacta gtgcttggat tctcaccaat aaaaaacgcc cggcggcaac
5941 cgagcgttct gaacaaatcc agatggagtt ctgaggtcat tactggatct atcaacagga
6001 gtccaagcga gctcgtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct
6061 atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata
6121 actacgatac gggagggctt accatctggc cccagtgctg caatgatacc gcgagaccca
6181 cgctcaccgg ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga
6241 agtggtcctg caactttatc cgcctccatc cagtctatta attgttgccg ggaagctaga
6301 gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg ccattgctac aggcatcgtg
6361 gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga
6421 gttacatgat cccccatgtt gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt
6481 gtcagaagta agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct
6541 cttactgtca tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca
6601 ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat
6661 accgcgccac atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga
6721 aaactctcaa ggatcttacc gctgttgaga tccagttcga tgtaacccac tcgtgcaccc
6781 aactgatctt cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg
6841 caaaatgccg caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc
6901 ctttttcaat attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt
6961 gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcca
7021 cct
//

A2.8. pBMO#50

```
LOCUS       (Bmo#50)\pTrcA2:      12871 bp    DNA     circular     1-MAY-2009
SOURCE
  ORGANISM
COMMENT
          ApEinfo:methylated:1
COMMENT     This file is created by Vector NTI
          http://www.invitrogen.com/
COMMENT     ORIGDB|GenBank
COMMENT     VNTDATE|517147829|
COMMENT     VNTDBDATE|517158080|
COMMENT     LSOWNER|
COMMENT     VNTNAME|(Bmo#50) pTrcA2:BUT_V3|
COMMENT     VNTAUTHORNAME|Jeffrey Dietrich|
COMMENT     VNTAUTHOREML|jadietrich@berkeley.edu|
COMMENT     VNTOAUTHORNAME|UNKNOWN|
FEATURES             Location/Qualifiers
     misc_feature    complement(11823..11828)
                     /ApEinfo_label=rbs
                     /ApEinfo_fwdcolor=#fcc466
                     /ApEinfo_revcolor=#fcc466
                     /vntifkey="21"
                     /label=rbs
     misc_feature    complement(137..1219)
                     /ApEinfo_label=LacIq
                     /ApEinfo_fwdcolor=#910b6e
                     /ApEinfo_revcolor=#910b6e
                     /vntifkey="21"
                     /label=LacIq
     misc_feature    11881..12725
                     /ApEinfo_label=p15A, OripACYC
                     /ApEinfo_fwdcolor=#feffb1
                     /ApEinfo_revcolor=#feffb1
                     /vntifkey="21"
                     /label=p15A,\OripACYC
     misc_binding    1484..1506
                     /ApEinfo_label=LacO
                     /ApEinfo_fwdcolor=cornflower blue
                     /ApEinfo_revcolor=cornflower blue
                     /vntifkey="20"
                     /label=LacO
     misc_feature    10705..10710
                     /ApEinfo_label=PstI
                     /ApEinfo_fwdcolor=#0080ff
                     /ApEinfo_revcolor=#0080ff
                     /vntifkey="21"
```

/label=PstI
misc_feature 10326..10693
/ApEinfo_label=TrrnB
/ApEinfo_fwdcolor=#9191ff
/ApEinfo_revcolor=#9191ff
/vntifkey="21"
/label=TrrnB
misc_feature 1276..1486
/ApEinfo_label=pTRC
/ApEinfo_fwdcolor=#2b2bff
/ApEinfo_revcolor=#2b2bff
/vntifkey="21"
/label=pTRC
misc_feature 10816..11472
/ApEinfo_label=CmR (EcoRI-KO)
/ApEinfo_fwdcolor=#fff54c
/ApEinfo_revcolor=#fff54c
/vntifkey="21"
/label=CmR\(EcoRI-KO)
misc_feature 1517..1520
/vntifkey="21"
/label=RBS
insertion_seq 1535..6253
/vntifkey="14"
/label=insert
misc_feature 1535..2320
/vntifkey="21"
/label=crt
misc_feature 2334..3473
/vntifkey="21"
/label=bcd
misc_feature 3492..4271
/vntifkey="21"
/label=etfB
misc_feature 4290..5300
/vntifkey="21"
/label=etfA
misc_feature 5405..6253
/vntifkey="21"
/label=hbd
misc_feature 7743..10319
/vntifkey="21"
/label=AdhE2
primer 7713..7771
/vntifkey="27"
/label=F72

```
     primer_bind     10299..10348
                     /vntifkey="28"
                     /label=F74
     primer_bind     complement(10299..10348)
                     /vntifkey="28"
                     /label=F75
     primer_bind     1488..1537
                     /vntifkey="28"
                     /label=F76F/F77R
     misc_binding    6468..6490
                     /ApEinfo_label=LacO
                     /ApEinfo_fwdcolor=cornflower blue
                     /ApEinfo_revcolor=cornflower blue
                     /vntifkey="20"
                     /label=LacO
     misc_feature    6260..6470
                     /ApEinfo_label=pTRC
                     /ApEinfo_fwdcolor=#2b2bff
                     /ApEinfo_revcolor=#2b2bff
                     /vntifkey="21"
                     /label=pTRC
     RBS             7710..7742
                     /vntifkey="32"
                     /label=20K
     misc_feature    6525..7709
                     /vntifkey="21"
                     /label=AtoB
     RBS             6491..6524
                     /vntifkey="32"
                     /label=RBS_100k
     primer_bind     6232..6289
                     /vntifkey="28"
                     /label=F90
     primer_bind     complement(6461..6520)
                     /vntifkey="28"
                     /label=F91R
     primer_bind     6498..6556
                     /vntifkey="28"
                     /label=F92
     primer_bind     complement(7680..7737)
                     /vntifkey="28"
                     /label=F93R
BASE COUNT     3890 a      2530 c      3209 g      3242 t
ORIGIN
        1 gaattcgcgg ccgcttctag agttcgcgcg cgaaggcgaa gcggcatgca tttacgttga
       61 caccatcgaa tggtgcaaaa cctttcgcgg tatggcatga tagcgcccgg aagagagtca
```

121 attcagggtg gtgaatgtga aaccagtaac gttatacgat gtcgcagagt atgccggtgt
181 ctcttatcag accgtttccc gcgtggtgaa ccaggccagc cacgtttctg cgaaaacgcg
241 ggaaaaagtg gaagcggcga tggcggagct gaattacatt cccaaccgcg tggcacaaca
301 actggcgggc aaacagtcgt tgctgattgg cgttgccacc tccagtctgg ccctgcacgc
361 gccgtcgcaa attgtcgcgg cgattaaatc tcgcgccgat caactgggtg ccagcgtggt
421 ggtgtcgatg gtagaacgaa gcggcgtcga agcctgtaaa gcggcggtgc acaatcttct
481 cgcgcaacgc gtcagtgggc tgatcattaa ctatccgctg gatgaccagg atgccattgc
541 tgtggaagct gcctgcacta atgttccggc gttatttctt gatgtctctg accagacacc
601 catcaacagt attattttct cccatgaaga cggtacgcga ctgggcgtgg agcatctggt
661 cgcattgggt caccagcaaa tcgcgctgtt agcgggccca ttaagttctg tctcggcgcg
721 tctgcgtctg gctggctggc ataaatatct cactcgcaat caaattcagc cgatagcgga
781 acgggaaggc gactggagtg ccatgtccgg ttttcaacaa accatgcaaa tgctgaatga
841 gggcatcgtt cccactgcga tgctggttgc caacgatcag atggcgctgg gcgcaatgcg
901 cgccattacc gagtccgggc tgcgcgttgg tgcggatatc tcggtagtgg gatacgacga
961 taccgaagac agctcatgtt atatcccgcc gtcaaccacc atcaaacagg attttcgcct
1021 gctggggcaa accagcgtgg accgcttgct gcaactctct cagggccagg cggtgaaggg
1081 caatcagctg ttgcccgtct cactggtgaa aagaaaaacc accctggcgc ccaatacgca
1141 aaccgcctct ccccgcgcgt tggccgattc attaatgcag ctggcacgac aggtttcccg
1201 actggaaagc gggcagtgag cgcaacgcaa ttaatgtgag ttagcgcgaa ttgatctggt
1261 ttgacagctt atcatcgact gcacggtgca ccaatgcttc tggcgtcagg cagccatcgg
1321 aagctgtggt atggctgtgc aggtcgtaaa tcactgcata attcgtgtcg ctcaaggcgc
1381 actcccgttc tggataatgt tttttgcgcc gacatcataa cggttctggc aaatattctg
1441 aaatgagctg ttgacaatta atcatccggc tcgtataatg tgtggaattg tgagcggata
1501 acaatttcac accagcagga cgcactgagg gcccatggaa ctaaacaatg tcatccttga
1561 aaaggaaggt aaagttgctg tagttaccat taacagacct aaagcattaa atgcgttaaa
1621 tagtgataca ctaaaagaaa tggattatgt tataggtgaa attgaaaatg atagcgaagt
1681 acttgcagta attttaactg gagcaggaga aaaatcattt gtagcaggag cagatatttc
1741 tgagatgaag gaaatgaata ccattgaagg tagaaaattc gggatacttg gaaataaagt
1801 gtttagaaga ttagaacttc ttgaaaagcc tgtaatagca gctgttaatg gttttgcttt
1861 aggaggcgga tgcgaaatag ctatgtcttg tgatataaga atagcttcaa gcaacgcaag
1921 atttggtcaa ccagaagtag gtctcggaat aacacctggt tttggtggta cacaaagact
1981 ttcaagatta gttggaatgg gcatggcaaa gcagcttata tttactgcac aaaatataaa
2041 ggcagatgaa gcattaagaa tcggacttgt aaataaggta gtagaaccta gtgaattaat
2101 gaatacagca aaagaaattg caaacaaaat tgtgagcaat gctccagtag ctgttaagtt
2161 aagcaaacag gctattaata gaggaatgca gtgtgatatt gatactgctt tagcatttga
2221 atcagaagca tttggagaat gcttttcaac agaggatcaa aaggatgcaa tgacagcttt
2281 catagagaaa agaaaaattg aaggcttcaa aaatagatag gaggtaagtt tatatggatt
2341 ttaatttaac aagagaacaa gaattagtaa gacagatggt tagagaattt gctgaaaatg
2401 aagttaaacc tatagcagca gaaattgatg aaacagaaag atttccaatg gaaaatgtaa
2461 agaaaatggg tcagtatggt atgatgggaa ttccattttc aaaagagtat ggtggcgcag
2521 gtggagatgt attatcttat ataatcgccg ttgaggaatt atcaaaggtt tgcggtacta
2581 caggagttat tctttcagca catacatcac tttgtgcttc attaataaat gaacatggta
2641 cagaagaaca aaaacaaaaa tatttagtac ctttagctaa aggtgaaaaa ataggtgctt
2701 atggattgac tgagccaaat gcaggaacag attctggagc acaacaaaca gtagctgtac
2761 ttgaaggaga tcattatgta attaatggtt caaaaatatt cataactaat ggaggagttg
2821 cagatacttt tgttatattt gcaatgactg acagaactaa aggaacaaaa ggtatatcag

2881 catttataat agaaaaaggc ttcaaaggtt tctctattgg taaagttgaa caaaagcttg
2941 gaataagagc ttcatcaaca actgaacttg tatttgaaga tatgatagta ccagtagaaa
3001 acatgattgg taaagaagga aaaggcttcc ctatagcaat gaaaactctt gatggaggaa
3061 gaattggtat agcagctcaa gctttaggta tagctgaagg tgctttcaac gaagcaagag
3121 cttacatgaa ggagagaaaa caatttggaa gaagccttga caaattccaa ggtcttgcat
3181 ggatgatggc agatatggat gtagctatag aatcagctag atatttagta tataaagcag
3241 catatcttaa acaagcagga cttccataca cagttgatgc tgcaagagct aagcttcatg
3301 ctgcaaatgt agcaatggat gtaacaacta aggcagtaca attatttggt ggatacggat
3361 atacaaaaga ttatccagtt gaaagaatga tgagagatgc taagataact gaaatatatg
3421 aaggaacttc agaagttcag aaattagtta tttcaggaaa aatttttaga taatttaagg
3481 aggttaagag gatgaatata gttgtttgtt taaaacaagt tccagataca gcggaagtta
3541 gaatagatcc agttaaggga acacttataa gagaaggagt tccatcaata ataaatccag
3601 atgataaaaa cgcacttgag gaagctttag tattaaaaga taattatggt gcacatgtaa
3661 cagttataag tatgggacct ccacaagcta aaaatgcttt agtagaagct ttggctatgg
3721 gtgctgatga agctgtactt ttaacagata gagcatttgg aggagcagat acacttgcga
3781 cttcacatac aattgcagca ggaattaaga agctaaaata tgatatagtt tttgctggaa
3841 ggcaggctat agatggagat acagctcagg ttggaccaga aatagctgag catcttggaa
3901 tacctcaagt aacttatgtt gagaaagttg aagttgatgg agatacttta aagattagaa
3961 aagcttggga agatggatat gaagttgttg aagttaagac accagttctt ttaacagcaa
4021 ttaaagaatt aaatgttcca agatatatga gtgtagaaaa aatattcgga gcatttgata
4081 aagaagtaaa aatgtggact gccgatgata tagatgtaga taaggctaat ttaggtctta
4141 aaggttcacc aactaaagtt aagaagtcat caactaaaga agttaaagga cagggagaag
4201 ttattgataa gcctgttaag gaagcagctg catatgttgt ctcaaaatta aaagaagaac
4261 actatattta agttaggagg gatttttcaa tgaataaagc agattacaag ggcgtatggg
4321 tgtttgctga acaaagagac ggagaattac aaaaggtatc attggaatta ttaggtaaag
4381 gtaaggaaat ggctgagaaa ttaggcgttg aattaacagc tgttttactt ggacataata
4441 ctgaaaaaat gtcaaaggat ttattatctc atggagcaga taaggtttta gcagcagata
4501 atgaactttt agcacatttt tcaacagatg gatatgctaa agttatatgt gatttagtta
4561 atgaaagaaa gccagaaata ttattcatag gagctacttt cataggaaga gatttaggac
4621 caagaatagc agcaagactt tctactggtt taactgctga ttgtacatca cttgacatag
4681 atgtagaaaa tagagattta ttggctacaa gaccagcgtt tggtggaaat ttgatagcta
4741 caatagtttg ttcagaccac agaccacaaa tggctacagt aagacctggt gtgtttgaaa
4801 aattacctgt taatgatgca aatgtttctg atgataaaat agaaaaagtt gcaattaaat
4861 taacagcatc agacataaga acaaaagttt caaaagttgt taagcttgct aaagatattg
4921 cagatatcgg agaagctaag gtattagttg ctggtggtag aggagttgga agcaaagaaa
4981 actttgaaaa acttgaagag ttagcaagtt tacttggtgg aacaatagcc gcttcaagag
5041 cagcaataga aaaagaatgg gttgataagg accttcaagt aggtcaaact ggtaaaactg
5101 taagaccaac tctttatatt gcatgtggta tatcaggagc tatccagcat ttagcaggta
5161 tgcaagattc agattacata attgctataa ataaagatgt agaagcccca ataatgaagg
5221 tagcagattt ggctatagtt ggtgatgtaa ataaagttgt accagaatta atagctcaag
5281 ttaaagctgc taataattaa gataaataaa aagaattatt taaagcttat tatgccaaaa
5341 tacttatata gtattttggt gtaaatgcat tgatagtttc tttaaattta gggaggtctg
5401 tttaatgaaa aaggtatgtg ttataggtgc aggtactatg ggttcaggaa ttgctcaggc
5461 atttgcagct aaaggatttg aagtagtatt aagagatatt aaagatgaat ttgttgatag
5521 aggattagat tttatcaata aaaatctttc taaattagtt aaaaaaggaa agatagaaga
5581 agctactaaa gttgaaatct taactagaat ttccggaaca gttgacctta atatggcagc

5641 tgattgcgat ttagttatag aagcagctgt tgaaagaatg gatattaaaa agcagatttt
5701 tgctgactta gacaatatat gcaagccaga aacaattctt gcatcaaata catcatcact
5761 ttcaataaca gaagtggcat cagcaactaa aagacctgat aaggttatag gtatgcattt
5821 ctttaatcca gctcctgtta tgaagcttgt agaggtaata agaggaatag ctacatcaca
5881 agaaactttt gatgcagtta aagagacatc tatagcaata ggaaaagatc ctgtagaagt
5941 agcagaagca ccaggatttg ttgtaaatag aatattaata ccaatgatta atgaagcagt
6001 tggtatatta gcagaaggaa tagcttcagt agaagacata gataaagcta tgaaacttgg
6061 agctaatcac ccaatgggac cattagaatt aggtgatttt ataggtcttg atatatgtct
6121 tgctataatg gatgttttat actcagaaac tggagattct aagtatagac cacatacatt
6181 acttaagaag tatgtaagag caggatggct tggaagaaaa tcaggaaaag gtttctacga
6241 ttattcaaaa taaggatccc gactgcacgg tgcaccaatg cttctggcgt caggcagcca
6301 tcggaagctg tggtatggct gtgcaggtcg taaatcactg cataattcgt gtcgctcaag
6361 gcgcactccc gttctggata atgtttttg cgccgacatc ataacggttc tggcaaatat
6421 tctgaaatga gctgttgaca attaatcatc cggctcgtat aatgtgtgga attgtgagcg
6481 gataacaatt catcgaaaaa attgaaataa ggaggtaggt agtaatgaaa aattgtgtca
6541 tcgtcagtgc ggtacgtact gctatcggta gttttaacgg ttcactcgct tccaccagcg
6601 ccatcgacct gggggcgaca gtaattaaag ccgccattga acgtgcaaaa atcgattcac
6661 aacacgttga tgaagtgatt atgggtaacg tgttacaagc cgggctgggg caaaatccgg
6721 cgcgtcaggc actgttaaaa agcgggctgg cagaaacggt gtgcggattc acggtcaata
6781 aagtatgtgg ttcgggtctt aaaagtgtgg cgcttgccgc ccaggccatt caggcaggtc
6841 aggcgcagag cattgtggcg gggggtatgg aaaatatgag tttagccccc tacttactcg
6901 atgcaaaagc acgctctggt tatcgtcttg gagacggaca ggtttatgac gtaatcctgc
6961 gcgatggcct gatgtgcgcc acccatggtt atcatatggg gattaccgcc gaaaacgtgg
7021 ctaaagagta cggaattacc cgtgaaatgc aggatgaact ggcgctacat tcacagcgta
7081 aagcggcagc cgcaattgag tccggtgctt ttacagccga aatcgtcccg gtaaatgttg
7141 tcactcgaaa gaaaaccttc gtcttcagtc aagacgaatt cccgaaagcg aattcaacgg
7201 ctgaagcgtt aggtgcattg cgcccggcct tcgataaagc aggaacagtc accgctggga
7261 acgcgtctgg tattaacgac ggtgctgccg ctctggtgat tatggaagaa tctgcggcgc
7321 tggcagcagg ccttaccccc ctggctcgca ttaaaagtta tgccagcggt ggcgtgcccc
7381 ccgcattgat gggtatgggg ccagtacctg ccacgcaaaa agcgttacaa ctggcggggc
7441 tgcaactggc ggatattgat ctcattgagg ctaatgaagc atttgctgca cagttccttg
7501 ccgttgggaa aaacctgggc tttgattctg agaaagtgaa tgtcaacggc ggggccatcg
7561 cgctcgggca tcctatcggt gccagtggtg ctcgtattct ggtcacacta ttacatgcca
7621 tgcaggcacg cgataaaacg ctggggctgg caacactgtg cattggcggc ggtcagggaa
7681 ttgcgatggt gattgaacgg ttgaattaac aaattaatac gcgagttaag gagtacagga
7741 gcatgaaggt gaccaatcag aaagaactga aacaaaagct gaatgagctg cgtgaggccc
7801 aaaagaagtt tgctacctat acccaagagc aggttgacaa aatcttcaaa cagtgcgcga
7861 tcgctgcagc aaaagaacgc attaacctgg cgaaactggc agtcgaagag actggcattg
7921 gtctggtgga ggacaaaatc atcaaaaacc actttgctgc ggagtacatt tacaataagt
7981 acaaaaacga gaaaacgtgt ggtatcattg atcacgacga ttctctgggt atcacgaagg
8041 tcgcggaacc gatcggcatc gttgccgcaa ttgttccgac caccaatccg accagcaccg
8101 ctatcttcaa aagcttgatt agcctgaaaa ctcgcaacgc gattttcttc agcccgcacc
8161 cgcgtgccaa aaagagcacc attgcggcag cgaaactgat tctggatgca gcggtgaagg
8221 cgggtgcccc gaagaacatc attggttgga ttgacgaacc gtctattgaa ctgagccaag
8281 acttgatgtc ggaggccgac atcatcctgg ctaccggtgg tccgagcatg gttaaggcag
8341 cgtactccag cggcaagcca gccatcggcg ttggcgcagg caacaccccg gcaatcattg

8401 atgagagcgc agacatcgac atggcagtca gcagcatcat tctgtccaaa acctatgaca
8461 atggcgtgat ttgcgcgagc gagcagagca tcctggtcat gaatagcatc tacgaaaaag
8521 ttaaagaaga gtttgtgaag cgtggtagct acattctgaa tcaaaacgaa atcgctaaga
8581 tcaaagaaac gatgttcaag aatggtgcga ttaacgccga catcgtcggc aagtctgcct
8641 atatcatcgc gaagatggcg ggtatcgagg tccctcagac caccaaaatc ctgatcggtg
8701 aggtgcagtc ggtggagaag tcggaactgt ttagccatga aaagctgagc ccggttctgg
8761 ctatgtacaa ggttaaagat tttgatgagg cgttgaagaa ggctcaacgt ctgattgaac
8821 tgggtggtag cggccacacg agcagcttgt acatcgattc ccaaaacaat aaggacaaag
8881 tgaaagagtt cggcctggcg atgaaaacca gccgcacgtt catcaatatg ccgagctccc
8941 agggtgcaag cggtgatctg tataactttg caattgcgcc gtccttcacg ctgggttgcg
9001 gcacttgggg tggcaattct gtttctcaaa atgtggaacc aaagcacttg ctgaatatca
9061 aatccgttgc ggaacgccgc gaaaacatgc tgtggttcaa ggtgccgcag aagatttact
9121 ttaagtatgg ctgcctgcgc ttcgcactga aagagctgaa ggatatgaat aagaaacgtg
9181 cgttcattgt tactgacaag gacctgttta agctgggtta tgttaacaag attaccaagg
9241 tcctggacga gattgacatt aagtatagca ttttcaccga tatcaaaagc gatccgacga
9301 tcgatagcgt gaagaaaggt gccaaagaga tgctgaattt cgaacctgat acgatcattt
9361 ccatcggtgg cggttccccg atggatgcgg cgaaggtcat gcacctgctg tacgagtacc
9421 cggaagcaga gattgaaaac ctggcaatca actttatgga tattcgcaag cgtatctgca
9481 acttcccgaa actgggcacc aaggcgatta gcgtcgctat tccgacgacc gcaggtacgg
9541 gtagcgaggc gacccctttt gcagtcatca cgaatgacga aacgggcatg aaatacccgc
9601 tgacgagcta tgaattgacg ccgaatatgg ccatcattga taccgagctg atgctgaaca
9661 tgccgcgtaa actgaccgca gcgacgggca ttgatgcgct ggtccatgcc atcgaggcct
9721 atgtgagcgt catggcgacg gactataccg atgagctggc actgcgtgca atcaaaatga
9781 ttttcaaata cctgccgcgt gcgtacaaaa acggtacgaa tgacattgag gctcgtgaga
9841 aaatggccca tgccagcaat atcgcgggca tggcatttgc caacgcgttc ttgggtgttt
9901 gtcacagcat ggcgcataag ttgggtgcga tgcatcacgt gccacatggt attgcgtgtg
9961 cagttctgat cgaagaagtg attaagtaca atgccacgga ctgtccgacc aaacagaccg
10021 cgtttccgca gtataagagc ccaaacgcca aacgtaagta cgccgagatc gcggagtatt
10081 tgaatctgaa aggtaccagc gacaccgaaa aagtgactgc gttgattgag gcgatcagca
10141 aattgaagat tgatctgtcc atcccgcaaa acatctctgc tgccggcatc aacaaaaagg
10201 acttctacaa caccctggat aagatgagcg agctggcgtt tgatgaccag tgcaccaccg
10261 cgaacccgcg ttatccgctg attagcgaac tgaaagacat ctacattaag agcttttgaa
10321 gatctgaagc ttgggcccga acaaaaactc atctcagaag aggatctgaa tagcgccgtc
10381 gaccatcatc atcatcatca ttgagtttaa acggtctcca gcttggctgt tttggcggat
10441 gagagaagat tttcagcctg atacagatta aatcagaacg cagaagcggt ctgataaaac
10501 agaatttgcc tggcggcagt agcgcggtgg tcccacctga ccccatgccg aactcagaag
10561 tgaaacgccg tagcgccgat ggtagtgtgg ggtctcccca tgcgagagta gggaactgcc
10621 aggcatcaaa taaaacgaaa ggctcagtcg aaagactggg cctttcgttt tatctgttgt
10681 ttgtcggtga actaattatc tagactgcag ttgatcgggc acgtaagagg ttccaacttt
10741 caccataatg aaataagatc actaccgggc gtatttttg agttatcgag attttcagga
10801 gctaaggaag ctaaaatgga gaaaaaaatc actggatata ccaccgttga tatatcccaa
10861 tggcatcgta aagaacattt tgaggcattt cagtcagttg ctcaatgtac ctataaccag
10921 accgttcagc tggatattac ggccttttta aagaccgtaa agaaaaataa gcacaagttt
10981 tatccggcct ttattcacat tcttgcccgc ctgatgaatg ctcatccgga atttcgtatg
11041 gcaatgaaag acggtgagct ggtgatatgg gatagtgttc accttgtta caccgttttc
11101 catgagcaaa ctgaaacgtt ttcatcgctc tggagtgaat accacgacga tttccggcag

11161 tttctacaca tatattcgca agatgtggcg tgttacggtg aaaacctggc ctatttccct
11221 aaagggttta ttgagaatat gtttttcgtc tcagccaatc cctgggtgag tttcaccagt
11281 tttgatttaa acgtggccaa tatggacaac ttcttcgccc ccgttttcac catgggcaaa
11341 tattatacgc aaggcgacaa ggtgctgatg ccgctggcga ttcaggttca tcatgccgtt
11401 tgtgatggct tccatgtcgg cagaatgctt aatgaattac aacagtactg cgatgagtgg
11461 cagggcgggg cgtaatttga tatcgagctc gcttggactc ctgttgatag atccagtaat
11521 gacctcagaa ctccatctgg atttgttcag aacgctcggt tgccgccggg cgtttttat
11581 tggtgagaat ccaagcctcg gtgagaatcc aagcctcgat caacgtctca ttttcgccaa
11641 aagttggccc agggcttccc ggtatcaaca gggacaccag gatttattta ttctgcgaag
11701 tgatcttccg tcacaggtat ttattcggcg caaagtgcgt cgggtgatgc tgccaactta
11761 ctgatttagt gtatgatggt gtttttgagg tgctccagtg gcttctgttt ctatcagctg
11821 tccctcctgt tcagctactg acggggtggt gcgtaacggc aaaagcaccg ccggacatca
11881 gcgctagcgg agtgtatact ggcttactat gttggcactg atgagggtgt cagtgaagtg
11941 cttcatgtgg caggagaaaa aaggctgcac cggtgcgtca gcagaatatg tgatacagga
12001 tatattccgc ttcctcgctc actgactcgc tacgctcggt cgttcgactg cggcgagcgg
12061 aaatggctta cgaacggggc ggagatttcc tggaagatgc caggaagata cttaacaggg
12121 aagtgagagg gccgcggcaa agccgttttt ccataggctc cgccccccctg acaagcatca
12181 cgaaatctga cgctcaaatc agtggtggcg aaacccgaca ggactataaa gataccaggc
12241 gtttcccccct ggcggctccc tcgtgcgctc tcctgttcct gcctttcggt ttaccggtgt
12301 cattccgctg ttatggccgc gtttgtctca ttccacgcct gacactcagt tccgggtagg
12361 cagttcgctc caagctggac tgtatgcacg aacccccccgt tcagtccgac cgctgcgcct
12421 tatccggtaa ctatcgtctt gagtccaacc cggaaagaca tgcaaaagca ccactggcag
12481 cagccactgg taattgattt agaggagtta gtcttgaagt catgcgccgg ttaaggctaa
12541 actgaaagga caagttttgg tgactgcgct cctccaagcc agttacctcg gttcaaagag
12601 ttggtagctc agagaacctt cgaaaaaccg ccctgcaagg cggttttttc gttttcagag
12661 caagagatta cgcgcagacc aaaacgatct caagaagatc atcttattaa tcagataaaa
12721 tatttctaga tttcagtgca atttatctct tcaaatgtag cacctgaagt cagccccata
12781 cgatataagt tgtaattctc atgtttgaca gcttatcatc gataagcttc cgatggcgcg
12841 ccgagaggct ttacacttta tgcttccggc t
//

A2.9. pBMO#52

```
LOCUS       (Bmo#52)\pTRC:ki       6857 bp    DNA     circular     2-NOV-2010
SOURCE
  ORGANISM
COMMENT     This file is created by Vector NTI
            http://www.invitrogen.com/
COMMENT     VNTDATE|558204163|
COMMENT     VNTDBDATE|569764081|
COMMENT     LSOWNER|
COMMENT     VNTNAME|(Bmo#52) pTRC:kivD,ADH6|
COMMENT     VNTAUTHORNAME|Jeffrey Dietrich|
COMMENT     VNTAUTHOREML|jadietrich@berkeley.edu|
FEATURES             Location/Qualifiers
     misc_feature    complement(4303..4308)
                     /ApEinfo_label=rbs
                     /ApEinfo_fwdcolor=#fcc466
                     /ApEinfo_revcolor=#fcc466
                     /vntifkey="21"
                     /label=rbs
     misc_feature    complement(5488..6570)
                     /ApEinfo_label=LacIq
                     /ApEinfo_fwdcolor=#910b6e
                     /ApEinfo_revcolor=#910b6e
                     /vntifkey="21"
                     /label=LacIq
     misc_feature    4361..5205
                     /ApEinfo_label=p15A, OripACYC
                     /ApEinfo_fwdcolor=#feffb1
                     /ApEinfo_revcolor=#feffb1
                     /vntifkey="21"
                     /label=p15A,\OripACYC
     misc_binding    6835..6857
                     /ApEinfo_label=LacO
                     /ApEinfo_fwdcolor=cornflower blue
                     /ApEinfo_revcolor=cornflower blue
                     /vntifkey="20"
                     /label=LacO
     misc_feature    3185..3190
                     /ApEinfo_label=PstI
                     /ApEinfo_fwdcolor=#0080ff
                     /ApEinfo_revcolor=#0080ff
                     /vntifkey="21"
                     /label=PstI
     misc_feature    2806..3173
                     /ApEinfo_label=TrrnB
                     /ApEinfo_fwdcolor=#9191ff
```

/ApEinfo_revcolor=#9191ff
/vntifkey="21"
/label=TrrnB
misc_feature 6627..6837
/ApEinfo_label=pTRC
/ApEinfo_fwdcolor=#2b2bff
/ApEinfo_revcolor=#2b2bff
/vntifkey="21"
/label=pTRC
misc_feature 3296..3952
/ApEinfo_label=CmR (EcoRI-KO)
/ApEinfo_fwdcolor=#fff54c
/ApEinfo_revcolor=#fff54c
/vntifkey="21"
/label=CmR\(EcoRI-KO)
RBS 7..36
/vntifkey="32"
/label=RBS
/note="Predicted strength of 8333... I don't know where I derived this RBS from"
misc_feature 37..1683
/vntifkey="21"
/label=KivD
misc_feature 1711..2793
/vntifkey="21"
/label=ADH6
RBS 1684..1710
/vntifkey="32"
/label=RBS_50K
/note="synthetic RBS (50K)_RBS calculator"
insertion_seq 6798..2864
/vntifkey="14"
/label=DNA2.0
/note="DNA2.0 order"
primer 2806..2865
/vntifkey="27"
/label=F132/F133
primer 6798..6857
/vntifkey="27"
/label=F130/F131
primer 1851..1876
/vntifkey="27"
/label=S36F/S37R
primer 2744..2769
/vntifkey="27"
/label=S38F/S39R
primer 131..156

```
                /vntifkey="27"
                /label=S40F/S41R
     primer     1550..1576
                /vntifkey="27"
                /label=S42F/S43R
     primer     1654..1713
                /vntifkey="27"
                /label=F136F/F137R
     primer     2776..2835
                /vntifkey="27"
                /label=F138F/F139R
     primer     6850..52
                /vntifkey="27"
                /label=DC133
     primer     1676..1735
                /vntifkey="27"
                /label=DC134R
     primer     complement(6820..22)
                /vntifkey="27"
                /label=DC145R
     primer     1708..1767
                /vntifkey="27"
                /label=DC146F
BASE COUNT     1935 a     1493 c     1631 g     1798 t
ORIGIN
        1 agatctccaa tatataataa aatatggagg aatgcgatgt atacagtagg agattaccta
       61 ttagaccgat tacacgagtt aggaattgaa gaaatttttg gagtccctgg agactataac
      121 ttacaatttt tagatcaaat tatttcccac aaggatatga aatgggtcgg aaatgctaat
      181 gaattaaatg cttcatatat ggctgatggc tatgctcgta ctaaaaaagc tgccgcattt
      241 cttacaacct ttggagtagg tgaattgagt gcagttaatg gattagcagg aagttacgcc
      301 gaaaatttac cagtagtaga aatagtggga tcacctacat caaaagttca aaatgaagga
      361 aaatttgttc atcatacgct ggctgacggt gattttaaac actttatgaa aatgcacgaa
      421 cctgttacag cagctcgaac tttactgaca gcagaaaatg caaccgttga aattgaccga
      481 gtactttctg cactattaaa agaaagaaaa cctgtctata tcaacttacc agttgatgtt
      541 gctgctgcaa aagcagagaa accctcactc cctttgaaaa aggaaaactc aacttcaaat
      601 acaagtgacc aagaaatttt gaacaaaatt caagaaagct tgaaaaatgc caaaaaacca
      661 atcgtgatta caggacatga aataattagt tttggcttag aaaaaacagt cactcaattt
      721 atttcaaaga caaaactacc tattacgaca ttaaactttg gtaaaagttc agttgatgaa
      781 gccctccctt catttttagg aatctataat ggtacactct cagagcctaa tcttaaagag
      841 ttcgtggaat cagccgactt catcttgatg cttggagtta aactcacaga ctcttcaaca
      901 ggagccttca ctcatcattt aaatgaaaat aaaatgattt cactgaatat agatgaagga
      961 aaaatattta acgaaagaat ccaaaatttt gattttgaat ccctcatctc ctctctctta
     1021 gacctaagcg aaatagaata caaaggaaaa tatatcgata aaaagcaaga agactttgtt
     1081 ccatcaaatg cgcttttatc acaagaccgc ctatggcaag cagttgaaaa cctaactcaa
     1141 agcaatgaaa caatcgttgc tgaacaaggg acatcattct ttggcgcttc atcaattttc
     1201 ttaaaatcaa agagtcattt tattggtcaa cccttatggg gatcaattgg atatacattc
```

1261 ccagcagcat taggaagcca aattgcagat aaagaaagca gacacctttt atttattggt
1321 gatggttcac ttcaacttac agtgcaagaa ttaggattag caatcagaga aaaaattaat
1381 ccaatttgct ttattatcaa taatgatggt tatacagtcg aaagagaaat tcatggacca
1441 aatcaaagct acaatgatat tccaatgtgg aattactcaa aattaccaga atcgtttgga
1501 gcaacagaag atcgagtagt ctcaaaaatc gttagaactg aaaatgaatt tgtgtctgtc
1561 atgaaagaag ctcaagcaga tccaaataga atgtactgga ttgagttaat tttggcaaaa
1621 gaaggtgcac caaaagtact gaaaaaaatg ggcaaactat ttgctgaaca aaataaatca
1681 taatccacga gttaaggaga gggggttcca atgtcttatc ctgagaaatt tgaaggtatc
1741 gctattcaat cacacgaaga ttggaaaaac ccaaagaaga caaagtatga cccaaaacca
1801 ttttacgatc atgacattga cattaagatc gaagcatgtg gtgtctgcgg tagtgatatt
1861 cattgtgcag ctggtcattg gggcaatatg aagatgccgc tagtcgttgg tcatgaaatc
1921 gttggtaaag ttgtcaagct agggcccaag tcaaacagtg ggttgaaagt cggtcaacgt
1981 gttggtgtag gtgctcaagt cttttcatgc ttggaatgtg accgttgtaa gaatgataat
2041 gaaccatact gcaccaagtt tgttaccaca tacagtcagc cttatgaaga cggctatgtg
2101 tcgcagggtg gctatgcaaa ctacgtcaga gttcatgaac attttgtggt gcctatccca
2161 gagaatattc catcacattt ggctgctcca ctattatgtg gtggtttgac tgtgtactct
2221 ccattggttc gtaacggttg cggtccaggt aaaaaagttg gtatagttgg tcttggtggt
2281 atcggcagta tgggtacatt gatttccaaa gccatggggg cagagacgta tgttatttct
2341 cgttcttcga gaaaaagaga agatgcaatg aagatgggcg ccgatcacta cattgctaca
2401 ttagaagaag gtgattgggg tgaaaagtac tttgacacct tcgacctgat tgtagtctgt
2461 gcttcctccc ttaccgacat tgacttcaac attatgccaa aggctatgaa ggttggtggt
2521 agaattgtct caatctctat accagaacaa cacgaaatgt tatcgctaaa gccatatggc
2581 ttaaaggctg tctccatttc ttacagtgct ttaggttcca tcaaagaatt gaaccaactc
2641 ttgaaattag tctctgaaaa agatatcaaa atttgggtgg aaacattacc tgttggtgaa
2701 gccggcgtcc atgaagcctt cgaaaggatg gaaaagggtg acgttagata tagatttacc
2761 ttagtcggct acgacaaaga attttcagac tagggatccc tcgaggaagc ttgggcccga
2821 acaaaaactc atctcagaag aggatctgaa tagcgccgtc gaccatcatc atcatcatca
2881 ttgagtttaa acggtctcca gcttggctgt tttggcggat gagagaagat tttcagcctg
2941 atacagatta aatcagaacg cagaagcggt ctgataaaac agaatttgcc tggcggcagt
3001 agcgcggtgg tcccacctga ccccatgccg aactcagaag tgaaacgccg tagcgccgat
3061 ggtagtgtgg ggtctcccca tgcgagagta gggaactgcc aggcatcaaa taaaacgaaa
3121 ggctcagtcg aaagactggg cctttcgttt tatctgttgt ttgtcggtga actaattatc
3181 tagactgcag ttgatcgggc acgtaagagg ttccaacttt caccataatg aaataagatc
3241 actaccgggc gtatttttg agttatcgag attttcagga gctaaggaag ctaaaatgga
3301 gaaaaaaatc actggatata ccaccgttga tatatcccaa tggcatcgta aagaacattt
3361 tgaggcattt cagtcagttg ctcaatgtac ctataaccag accgttcagc tggatattac
3421 ggccttttta aagaccgtaa agaaaaataa gcacaagttt tatccggcct ttattcacat
3481 tcttgcccgc ctgatgaatg ctcatccgga atttcgtatg gcaatgaaag acggtgagct
3541 ggtgatatgg gatagtgttc acccttgtta caccgttttc catgagcaaa ctgaaacgtt
3601 ttcatcgctc tggagtgaat accacgacga tttccggcag tttctacaca tatattcgca
3661 agatgtggcg tgttacggtg aaaacctggc ctatttccct aaagggttta ttgagaatat
3721 gtttttcgtc tcagccaatc cctgggtgag tttcaccagt tttgatttaa acgtggccaa
3781 tatggacaac ttcttcgccc ccgttttcac catgggcaaa tattatacgc aaggcgacaa
3841 ggtgctgatg ccgctggcga ttcaggttca tcatgccgtt tgtgatggct tccatgtcgg
3901 cagaatgctt aatgaattac aacagtactg cgatgagtgg cagggcgggg cgtaatttga
3961 tatcgagctc gcttggactc ctgttgatag atccagtaat gacctcagaa ctccatctgg

4021 atttgttcag aacgctcggt tgccgccggg cgtttttat tggtgagaat ccaagcctcg
4081 gtgagaatcc aagcctcgat caacgtctca ttttcgccaa aagttggccc agggcttccc
4141 ggtatcaaca gggacaccag gatttattta ttctgcgaag tgatcttccg tcacaggtat
4201 ttattcggcg caaagtgcgt cgggtgatgc tgccaactta ctgatttagt gtatgatggt
4261 gtttttgagg tgctccagtg gcttctgttt ctatcagctg tccctcctgt tcagctactg
4321 acggggtggt gcgtaacggc aaaagcaccg ccggacatca gcgctagcgg agtgtatact
4381 ggcttactat gttggcactg atgagggtgt cagtgaagtg cttcatgtgg caggagaaaa
4441 aaggctgcac cggtgcgtca gcagaatatg tgatacagga tatattccgc ttcctcgctc
4501 actgactcgc tacgctcggt cgttcgactg cggcgagcgg aaatggctta cgaacggggc
4561 ggagatttcc tggaagatgc caggaagata cttaacaggg aagtgagagg gccgcggcaa
4621 agccgttttt ccataggctc cgccccccctg acaagcatca cgaaatctga cgctcaaatc
4681 agtggtggcg aaacccgaca ggactataaa gataccaggc gtttcccccct ggcggctccc
4741 tcgtgcgctc tcctgttcct gcctttcggt ttaccggtgt cattccgctg ttatggccgc
4801 gtttgtctca ttccacgcct gacactcagt tccgggtagg cagttcgctc caagctggac
4861 tgtatgcacg aacccccccgt tcagtccgac cgctgcgcct tatccggtaa ctatcgtctt
4921 gagtccaacc cggaaagaca tgcaaaagca ccactggcag cagccactgg taattgattt
4981 agaggagtta gtcttgaagt catgcgccgg ttaaggctaa actgaaagga caagttttgg
5041 tgactgcgct cctccaagcc agttacctcg gttcaaagag ttggtagctc agagaacctt
5101 cgaaaaaccg ccctgcaagg cggttttttc gttttcagag caagagatta cgcgcagacc
5161 aaaacgatct caagaagatc atcttattaa tcagataaaa tatttctaga tttcagtgca
5221 atttatctct tcaaatgtag cacctgaagt cagccccata cgatataagt tgtaattctc
5281 atgtttgaca gcttatcatc gataagcttc cgatggcgcg ccgagaggct ttacacttta
5341 tgcttccggc tgaattcgcg gccgcttcta gagttcgcgc gcgaaggcga agcggcatgc
5401 atttacgttg acaccatcga atggtgcaaa acctttcgcg gtatggcatg atagcgcccg
5461 gaagagagtc aattcagggt ggtgaatgtg aaaccagtaa cgttatacga tgtcgcagag
5521 tatgccggtg tctcttatca gaccgtttcc cgcgtggtga accaggccag ccacgtttct
5581 gcgaaaacgc gggaaaaagt ggaagcggcg atggcggagc tgaattacat tcccaaccgc
5641 gtggcacaac aactggcggg caaacagtcg ttgctgattg gcgttgccac ctccagtctg
5701 gccctgcacg cgccgtcgca aattgtcgcg gcgattaaat ctcgcgccga tcaactgggt
5761 gccagcgtgg tggtgtcgat ggtagaacga agcggcgtcg aagcctgtaa agcggcggtg
5821 cacaatcttc tcgcgcaacg cgtcagtggg ctgatcatta actatccgct ggatgaccag
5881 gatgccattg ctgtggaagc tgcctgcact aatgttccgg cgttatttct tgatgtctct
5941 gaccagacac ccatcaacag tattattttc tcccatgaag acggtacgcg actgggcgtg
6001 gagcatctgg tcgcattggg tcaccagcaa atcgcgctgt tagcgggccc attaagttct
6061 gtctcggcgc gtctgcgtct ggctggctgg cataaatatc tcactcgcaa tcaaattcag
6121 ccgatagcgg aacgggaagg cgactggagt gccatgtccg gttttcaaca aaccatgcaa
6181 atgctgaatg agggcatcgt tcccactgcg atgctggttg ccaacgatca gatggcgctg
6241 ggcgcaatgc gcgccattac cgagtccggg ctgcgcgttg gtgcggatat ctcggtagtg
6301 ggatacgacg ataccgaaga cagctcatgt tatatcccgc cgtcaaccac catcaaacag
6361 gattttcgcc tgctggggca aaccagcgtg gaccgcttgc tgcaactctc tcagggccag
6421 gcggtgaagg gcaatcagct gttgcccgtc tcactggtga aaagaaaaac caccctggcg
6481 cccaatacgc aaaccgcctc tccccgcgcg ttggccgatt cattaatgca gctggcacga
6541 caggtttccc gactggaaag cgggcagtga gcgcaacgca attaatgtga gttagcgcga
6601 attgatctgg tttgacagct tatcatcgac tgcacggtgc accaatgctt ctggcgtcag
6661 gcagccatcg gaagctgtgg tatggctgtg caggtcgtaa atcactgcat aattcgtgtc
6721 gctcaaggcg cactcccgtt ctggataatg ttttttgcgc cgacatcata acggttctgg

6781 caaatattct gaaatgagct gttgacaatt aatcatccgg ctcgtataat gtgtggaatt
6841 gtgagcggat aacaatt
//

A2.10. pBMO#61

```
LOCUS       (Bmo#61)\pTRC:PD       6917 bp    DNA     circular     11-JAN-2011
SOURCE
  ORGANISM
COMMENT     This file is created by Vector NTI
            http://www.invitrogen.com/
COMMENT     VNTDATE|574702854|
COMMENT     VNTDBDATE|578857667|
COMMENT     LSOWNER|
COMMENT     VNTNAME|(Bmo#61) pTRC:PDC,ADH6|
COMMENT     VNTAUTHORNAME|Jeffrey Dietrich|
COMMENT     VNTAUTHOREML|jadietrich@berkeley.edu|
FEATURES             Location/Qualifiers
     misc_feature    complement(4363..4368)
                     /ApEinfo_label=rbs
                     /ApEinfo_fwdcolor=#fcc466
                     /ApEinfo_revcolor=#fcc466
                     /vntifkey="21"
                     /label=rbs
     misc_feature    complement(5548..6630)
                     /ApEinfo_label=LacIq
                     /ApEinfo_fwdcolor=#910b6e
                     /ApEinfo_revcolor=#910b6e
                     /vntifkey="21"
                     /label=LacIq
     misc_feature    4421..5265
                     /ApEinfo_label=p15A, OripACYC
                     /ApEinfo_fwdcolor=#feffb1
                     /ApEinfo_revcolor=#feffb1
                     /vntifkey="21"
                     /label=p15A,\OripACYC
     misc_binding    6895..6917
                     /ApEinfo_label=LacO
                     /ApEinfo_fwdcolor=cornflower blue
                     /ApEinfo_revcolor=cornflower blue
                     /vntifkey="20"
                     /label=LacO
     misc_feature    3245..3250
                     /ApEinfo_label=PstI
                     /ApEinfo_fwdcolor=#0080ff
                     /ApEinfo_revcolor=#0080ff
                     /vntifkey="21"
                     /label=PstI
     misc_feature    2866..3233
                     /ApEinfo_label=TrrnB
                     /ApEinfo_fwdcolor=#9191ff
```

/ApEinfo_revcolor=#9191ff
/vntifkey="21"
/label=TrrnB
misc_feature 6687..6897
/ApEinfo_label=pTRC
/ApEinfo_fwdcolor=#2b2bff
/ApEinfo_revcolor=#2b2bff
/vntifkey="21"
/label=pTRC
misc_feature 3356..4012
/ApEinfo_label=CmR (EcoRI-KO)
/ApEinfo_fwdcolor=#fff54c
/ApEinfo_revcolor=#fff54c
/vntifkey="21"
/label=CmR\(EcoRI-KO)
misc_feature 1771..2853
/vntifkey="21"
/label=ADH6
RBS 1744..1770
/vntifkey="32"
/label=RBS_50K
/note="synthetic RBS (50K)_RBS calculator"
insertion_seq 6858..2924
/vntifkey="14"
/label=DNA2.0
/note="DNA2.0 order"
primer 2866..2925
/vntifkey="27"
/label=F132/F133
primer 6858..6917
/vntifkey="27"
/label=F130/F131
primer 1911..1936
/vntifkey="27"
/label=S36F/S37R
primer 2804..2829
/vntifkey="27"
/label=S38F/S39R
primer 2836..2895
/vntifkey="27"
/label=F138F/F139R
primer 6910..6
/vntifkey="27"
/label=DC133
primer 1744..1795
/vntifkey="27"

/label=DC134R
primer complement(6880..6)
/vntifkey="27"
/label=DC145R
primer 1768..1827
/vntifkey="27"
/label=DC146F
misc_feature 37..1743
/vntifkey="21"
/label=PDC
primer 1714..1773
/vntifkey="27"
/label=DC152/DC153
primer complement(6894..36)
/vntifkey="27"
/label=DC155
primer 1744..1803
/vntifkey="27"
/label=DC156
primer 7..66
/vntifkey="27"
/label=DC154
BASE COUNT 1721 a 1659 c 1778 g 1758 t 1 others
ORIGIN
1 agatctccaa tatataataa aatatggagg aatgcgatga gttatactgt cggtacctat
61 ttagcggagc ggcttgtcca gattggtctc aagcatcact tcgcagtcgc gggcgactac
121 aacctcgtcc ttcttgacaa cctgcttttg aacaaaaaca tggagcaggt ttattgctgt
181 aacgaactga actgcggttt cagtgcagaa ggttatgctc gtgccaaagg cgcagcagca
241 gccgtcgtta cctacagcgt cggtgcgctt tccgcatttg atgctatcgg tggcgcctat
301 gcagaaaacc ttccggttat cctgatctcc ggtgctccga acaacaatga ccacgctgct
361 ggtcacgtgt tgcatcacgc tcttggcaaa accgactatc actatcagtt ggaaatggcc
421 aagaacatca cggccgccgc tgaagcgatt tataccccgg aagaagctcc ggctaaaatc
481 gatcacgtga ttaaaactgc tcttcgtgag aagaagccgg tttatctcga aatcgcttgc
541 aacattgctt ccatgccctg cgccgctcct ggaccggcaa gcgcattgtt caatgacgaa
601 gccagcgacg aagcttcttt gaatgcagcg gttgaagaaa ccctgaaatt catcgccnac
661 cgcgacaaag ttgccgtcct cgtcggcagc aagctgcgcg cagctggtgc tgaagaagct
721 gctgtcaaat ttgctgatgc tcttggtggc gcagttgcta ccatggctgc tgcaaaaagc
781 ttcttcccag aagaaaaccc gcattacatc ggtacctcat ggggtgaagt cagctatccg
841 ggcgttgaaa agacgatgaa agaagccgat gcggttatcg ctctggctcc tgtctttaac
901 gactactcca ccactggttg gacggatatt cctgatccta agaaactggt tctcgctgaa
961 ccgcgttctg tcgtcgttaa cggcattcgc ttccccagcg tccatctgaa agactatctg
1021 acccgtttgg ctcagaaagt ttccaagaaa accggtgctt tggacttctt caaatccctc
1081 aatgcaggtg aactgaagaa agccgctccg gctgatccga gtgctccgtt ggtcaacgca
1141 gaaatcgccc gtcaggtcga agctcttctg accccgaaca cgacggttat tgctgaaacc
1201 ggtgactctt ggttcaatgc tcagcgcatg aagctcccga acggtgctcg cgttgaatat
1261 gaaatgcagt ggggtcacat tggttggtcc gttcctgccg ccttcggtta tgccgtcggt

1321 gctccggaac gtcgcaacat cctcatggtt ggtgatggtt ccttccagct gacggctcag
1381 gaagtcgctc agatggttcg cctgaaactg ccggttatca tcttcttgat caataactat
1441 ggttacacca tcgaagttat gatccatgat ggtccgtaca acaacatcaa gaactgggat
1501 tatgccggtc tgatggaagt gttcaacggt aacggtggtt atgacagcgg tgctggtaaa
1561 ggcctgaagg ctaaaaccgg tggcgaactg gcagaagcta tcaaggttgc tctggcaaac
1621 accgacggcc caaccctgat cgaatgcttc atcggtcgtg aagactgcac tgaagaattg
1681 gtcaaatggg gtaagcgcgt tgctgccgcc aacagccgta agcctgttaa caagctcctc
1741 tattccacga gttaaggaga gggggttcca atgtcttatc ctgagaaatt tgaaggtatc
1801 gctattcaat cacacgaaga ttggaaaaac ccaaagaaga caaagtatga cccaaaacca
1861 ttttacgatc atgacattga cattaagatc gaagcatgtg gtgtctgcgg tagtgatatt
1921 cattgtgcag ctggtcattg gggcaatatg aagatgccgc tagtcgttgg tcatgaaatc
1981 gttggtaaag ttgtcaagct agggcccaag tcaaacagtg ggttgaaagt cggtcaacgt
2041 gttggtgtag gtgctcaagt cttttcatgc ttggaatgtg accgttgtaa gaatgataat
2101 gaaccatact gcaccaagtt tgttaccaca tacagtcagc cttatgaaga cggctatgtg
2161 tcgcagggtg gctatgcaaa ctacgtcaga gttcatgaac attttgtggt gcctatccca
2221 gagaatattc catcacattt ggctgctcca ctattatgtg gtggtttgac tgtgtactct
2281 ccattggttc gtaacggttg cggtccaggt aaaaaagttg gtatagttgg tcttggtggt
2341 atcggcagta tgggtacatt gatttccaaa gccatggggg cagagacgta tgttatttct
2401 cgttcttcga gaaaaagaga agatgcaatg aagatgggcg ccgatcacta cattgctaca
2461 ttagaagaag gtgattgggg tgaaaagtac tttgacacct tcgacctgat tgtagtctgt
2521 gcttcctccc ttaccgacat tgacttcaac attatgccaa aggctatgaa ggttggtggt
2581 agaattgtct caatctctat accagaacaa cacgaaatgt tatcgctaaa gccatatggc
2641 ttaaaggctg tctccatttc ttacagtgct ttaggttcca tcaaagaatt gaaccaactc
2701 ttgaaattag tctctgaaaa agatatcaaa atttgggtgg aaacattacc tgttggtgaa
2761 gccggcgtcc atgaagcctt cgaaaggatg gaaaagggtg acgttagata tagatttacc
2821 ttagtcggct acgacaaaga attttcagac tagggatccc tcgaggaagc ttgggcccga
2881 acaaaaactc atctcagaag aggatctgaa tagcgccgtc gaccatcatc atcatcatca
2941 ttgagtttaa acggtctcca gcttggctgt tttggcggat gagagaagat tttcagcctg
3001 atacagatta aatcagaacg cagaagcggt ctgataaaac agaatttgcc tggcggcagt
3061 agcgcggtgg tcccacctga ccccatgccg aactcagaag tgaaacgccg tagcgccgat
3121 ggtagtgtgg ggtctcccca tgcgagagta gggaactgcc aggcatcaaa taaaacgaaa
3181 ggctcagtcg aaagactggg cctttcgttt tatctgttgt ttgtcggtga actaattatc
3241 tagactgcag ttgatcgggc acgtaagagg ttccaacttt caccataatg aaataagatc
3301 actaccgggc gtatttttg agttatcgag attttcagga gctaaggaag ctaaaatgga
3361 gaaaaaaatc actggatata ccaccgttga tatatcccaa tggcatcgta aagaacattt
3421 tgaggcattt cagtcagttg ctcaatgtac ctataaccag accgttcagc tggatattac
3481 ggccttttta aagaccgtaa agaaaaataa gcacaagttt tatccggcct ttattcacat
3541 tcttgcccgc ctgatgaatg ctcatccgga atttcgtatg gcaatgaaag acggtgagct
3601 ggtgatatgg gatagtgttc acccttgtta caccgttttc catgagcaaa ctgaaacgtt
3661 ttcatcgctc tggagtgaat accacgacga tttccggcag tttctacaca tatattcgca
3721 agatgtggcg tgttacggtg aaaacctggc ctatttccct aaagggttta ttgagaatat
3781 gtttttcgtc tcagccaatc cctgggtgag tttcaccagt tttgatttaa acgtggccaa
3841 tatggacaac ttcttcgccc ccgttttcac catgggcaaa tattatacgc aaggcgacaa
3901 ggtgctgatg ccgctggcga ttcaggttca tcatgccgtt tgtgatggct tccatgtcgg
3961 cagaatgctt aatgaattac aacagtactg cgatgagtgg cagggcgggg cgtaatttga
4021 tatcgagctc gcttggactc ctgttgatag atccagtaat gacctcagaa ctccatctgg

4081 atttgttcag aacgctcggt tgccgccggg cgtttttat tggtgagaat ccaagcctcg
4141 gtgagaatcc aagcctcgat caacgtctca ttttcgccaa aagttggccc agggcttccc
4201 ggtatcaaca gggacaccag gatttattta ttctgcgaag tgatcttccg tcacaggtat
4261 ttattcggcg caaagtgcgt cgggtgatgc tgccaactta ctgatttagt gtatgatggt
4321 gtttttgagg tgctccagtg gcttctgttt ctatcagctg tccctcctgt tcagctactg
4381 acggggtggt gcgtaacggc aaaagcaccg ccggacatca gcgctagcgg agtgtatact
4441 ggcttactat gttggcactg atgagggtgt cagtgaagtg cttcatgtgg caggagaaaa
4501 aaggctgcac cggtgcgtca gcagaatatg tgatacagga tatattccgc ttcctcgctc
4561 actgactcgc tacgctcggt cgttcgactg cggcgagcgg aaatggctta cgaacggggc
4621 ggagatttcc tggaagatgc caggaagata cttaacaggg aagtgagagg gccgcggcaa
4681 agccgttttt ccataggctc cgccccccctg acaagcatca cgaaatctga cgctcaaatc
4741 agtggtggcg aaacccgaca ggactataaa gataccaggc gtttcccccct ggcggctccc
4801 tcgtgcgctc tcctgttcct gcctttcggt ttaccggtgt cattccgctg ttatggccgc
4861 gtttgtctca ttccacgcct gacactcagt tccgggtagg cagttcgctc caagctggac
4921 tgtatgcacg aacccccccgt tcagtccgac cgctgcgcct tatccggtaa ctatcgtctt
4981 gagtccaacc cggaaagaca tgcaaaagca ccactggcag cagccactgg taattgattt
5041 agaggagtta gtcttgaagt catgcgccgg ttaaggctaa actgaaagga caagttttgg
5101 tgactgcgct cctccaagcc agttacctcg gttcaaagag ttggtagctc agagaacctt
5161 cgaaaaaccg ccctgcaagg cggttttttc gttttcagag caagagatta cgcgcagacc
5221 aaaacgatct caagaagatc atcttattaa tcagataaaa tatttctaga tttcagtgca
5281 atttatctct tcaaatgtag cacctgaagt cagccccata cgatataagt tgtaattctc
5341 atgtttgaca gcttatcatc gataagcttc cgatggcgcg ccgagaggct ttacacttta
5401 tgcttccggc tgaattcgcg gccgcttcta gagttcgcgc gcgaaggcga agcggcatgc
5461 atttacgttg acaccatcga atggtgcaaa acctttcgcg gtatggcatg atagcgcccg
5521 gaagagagtc aattcagggt ggtgaatgtg aaaccagtaa cgttatacga tgtcgcagag
5581 tatgccggtg tctcttatca gaccgtttcc cgcgtggtga accaggccag ccacgtttct
5641 gcgaaaacgc gggaaaaagt ggaagcggcg atggcggagc tgaattacat tcccaaccgc
5701 gtggcacaac aactggcggg caaacagtcg ttgctgattg gcgttgccac ctccagtctg
5761 gccctgcacg cgccgtcgca aattgtcgcg gcgattaaat ctcgcgccga tcaactgggt
5821 gccagcgtgg tggtgtcgat ggtagaacga agcggcgtcg aagcctgtaa agcggcggtg
5881 cacaatcttc tcgcgcaacg cgtcagtggg ctgatcatta actatccgct ggatgaccag
5941 gatgccattg ctgtggaagc tgcctgcact aatgttccgg cgttatttct tgatgtctct
6001 gaccagacac ccatcaacag tattattttc tcccatgaag acggtacgcg actgggcgtg
6061 gagcatctgg tcgcattggg tcaccagcaa atcgcgctgt tagcgggccc attaagttct
6121 gtctcggcgc gtctgcgtct ggctggctgg cataaatatc tcactcgcaa tcaaattcag
6181 ccgatagcgg aacgggaagg cgactggagt gccatgtccg gttttcaaca aaccatgcaa
6241 atgctgaatg agggcatcgt tcccactgcg atgctggttg ccaacgatca gatggcgctg
6301 ggcgcaatgc gcgccattac cgagtccggg ctgcgcgttg gtgcggatat ctcggtagtg
6361 ggatacgacg ataccgaaga cagctcatgt tatatcccgc cgtcaaccac catcaaacag
6421 gattttcgcc tgctggggca aaccagcgtg gaccgcttgc tgcaactctc tcagggccag
6481 gcggtgaagg gcaatcagct gttgcccgtc tcactggtga aaagaaaaac caccctggcg
6541 cccaatacgc aaaccgcctc tccccgcgcg ttggccgatt cattaatgca gctggcacga
6601 caggtttccc gactggaaag cgggcagtga gcgcaacgca attaatgtga gttagcgcga
6661 attgatctgg tttgacagct tatcatcgac tgcacggtgc accaatgctt ctggcgtcag
6721 gcagccatcg gaagctgtgg tatggctgtg caggtcgtaa atcactgcat aattcgtgtc
6781 gctcaaggcg cactcccgtt ctggataatg ttttttgcgc cgacatcata acggttctgg

6841 caaatattct gaaatgagct gttgacaatt aatcatccgg ctcgtataat gtgtggaatt
6901 gtgagcggat aacaatt
//