acs chemical biology

REVIEW

Synthetic Biology for Synthetic Chemistry

Jay D. Keasling*

Departments of Chemical Engineering and Bioengineering, University of California, Berkeley, Berkeley, California 94720, and Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

Downloaded via 155.190.61.27 on July 17, 2025 at 23:24:21 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ABSTRACT The richness and versatility of biological systems make them ideally suited to solve some of the world's most significant challenges, such as converting cheap, renewable resources into energy-rich molecules; producing high-quality, inexpensive drugs to fight disease; and remediating polluted sites. Over the years, significant strides have been made in engineering microorganisms to produce fuels, bulk chemicals, and valuable drugs from inexpensive starting materials; to detect and degrade nerve agents as well as less toxic organic pollutants; and to accumulate metals and reduce radionuclides. The components needed to engineer the chemistry inside a microbial cell are significantly different from those commonly used to overproduce pharmaceutical proteins. Synthetic biology has had and will continue to have a significant impact on the development of these components to engineer cellular metabolism and microbial chassis to host the chemistry. The ready availability of more well-characterized gene expression components and hosts for chemical synthesis, standards for the connection of these components to make larger functioning devices, computer-aided design software, and debugging tools for biological designs will decrease both the time and the support needed to construct these designs. Some of the most important tools for engineering bacterial metabolism and their use for production of the antimalarial drug artemisinin are reviewed.

*Corresponding author,
keasling@berkeley.edu.

Received for review November 27, 2007
and accepted December 27, 2007.
Published online January 18, 2008
10.1021/cb7002434 CCC: $40.75

© 2008 American Chemical Society

Although synthetic organic chemistry has made tremendous advances since its inception with the production of urea (*1*), there are many molecules, from simple fuels and chemical intermediates to complex natural products and their derivatives, that might be best produced inside a cell. Enzymes can catalyze in a single step what might require many steps using synthetic chemistry methodologies (*2*). By coupling multiple enzymes to create a metabolic pathway (in a single cell), one eliminates the need for purification of the chemical intermediates prior to the next reaction step. If intracellular metabolites can be used for the production of the desired chemical, then it can be produced from simple, inexpensive, starting materials, like glucose. The manipulation of intracellular metabolic reactions for the production of a desired chemical is generally referred to as metabolic engineering (*3*–*7*).

Engineering metabolic chemistry generally begins with the introduction of one or more genes encoding enzymes that transform readily available, intracellular intermediates into a desired chemical. Unlike production of pharmaceutical proteins in which the genes are highly expressed to maximize production of the target protein, the genes encoding the transformational enzymes do not need to be highly expressed; rather, the enzymes need to be produced in catalytic amounts only sufficient to adequately transform the metabolic intermediates into the desired products at a sufficient rate. Expression of the desired genes at too high a level will rob the cell of metabolites (nucleotides for the excess mRNA (mRNA), amino acids for the excess protein, *etc.*) that might be otherwise used to produce the desired molecule of interest (*8*). Furthermore, because intermediates of a foreign metabolic pathway can be toxic to a heterologous host (*9*), which can result in decreased production of the desired final compound, it is essential that the relative levels of the enzymes be coordinated in such a way that no intermediate in the path-

IPR2024-01419
Illumina Inc., v. Molecular Loop Biosciences, Inc.
Molecular Loop's Exhibit 2019
EX2019, Page 1

way accumulates to toxic levels. Thus, the metabolic engineer needs a toolset that will allow him or her to accurately control the native and heterologous reactions inside the cell.

Although a number of gene expression devices (promoter systems, vectors, *etc.*) have been created for controlling gene expression, much of the work in this area has been directed at the production of pharmaceutical proteins. Producing large quantities of a single desired protein is a very different endeavor than balancing the expression of multiple genes in such a way as to maximize production of a desired metabolite. Indeed, many of the biological engineering tools currently available to scientists and engineers have not changed significantly since the dawn of genetic engineering in the 1970s: biological engineers largely use natural, gene expression control systems (promoters with their cognate repressors/activators). The ability to place a single heterologous gene under the control of one of these native promoters and produce large quantities of a protein of interest is the basis for the modern biotechnology industry. While these redesigned biological control systems have been generally effective for their intended purpose (controlling rather roughly the expression of a single gene or a few genes), not surprisingly they are often inadequate for more complicated engineering tasks: for example, control of very large, heterologous, metabolic pathways or signal transduction systems. Further, these borrowed "biological parts" retain many of the features that were beneficial in their native form but which make them difficult to use for purposes other than that for which they evolved. Well-characterized standard biological components of the sort that are now available in many other engineering disciplines would make metabolic engineering more predictable and allow construction and integration of larger systems than is currently possible.

In almost every other field of engineering, standards have been developed to allow one to easily assemble components from various manufacturers to build a large integrated system. Biologists and engineers have not yet defined the standards for the various parts that might allow them to build larger biological devices. The design and construction of new devices (*e.g.*, genetic control systems or metabolic pathways) would benefit greatly from a set of standards that would govern how the various parts (*e.g.*, regulatory proteins, promoter, ribosome binding site, enzymes) should interact and be assembled. Setting a standard will, in turn, encourage manufacturing firms to develop biological components as well.

Biological engineering is further hindered by the fact that many of the most effective biological components (promoters, genes, plasmids, *etc.*) have been patented and are available only to those companies that can afford the royalty payments, which increase the cost of biological engineering endeavors and hamper the development of new biological solutions to problems where the eventual monetary payoff is not significant. Open-source biological components and eventually whole cells will decrease the cost of engineering metabolic pathways, make metabolic engineering more predictable with less guesswork, and encourage the development of novel biological solutions to some of our most challenging problems.

Synthetic Biology. Synthetic biology is the design and construction of new biological components, such as enzymes, genetic circuits, and cells, or the redesign of existing biological systems. Synthetic biology builds on the advances in molecular, cell, and systems biology and seeks to transform biology in the same way that synthesis transformed chemistry and integrated circuit design transformed computing. The element that distinguishes synthetic biology from traditional molecular and cellular biology is the focus on the design and construction of core components (parts of enzymes, genetic circuits, metabolic pathways, *etc.*) that can be modeled, understood, and tuned to meet specific performance criteria, and the assembly of these smaller components into larger integrated systems that solve specific problems. Just as engineers now design integrated circuits based on the known physical properties of materials and then fabricate functioning circuits and entire processors (with relatively high reliability), synthetic biologists will soon design and undertake large-scale biological engineering projects. Unlike many other areas of engineering, biology is incredibly nonlinear and less predictable, and there is less knowledge of the components and how they interact. Hence, the overwhelming physical details of natural biology (gene sequences, protein properties, biological systems) must be organized and recast *via* a set of design rules that hide information and manage complexity, thereby enabling the engineering of many-component integrated biological systems. It is only when this is accomplished that designs of significant scale (like constructing com-

EX2019, Page 2

plex metabolic pathways) will be possible in the time periods characteristic of many other engineering fields.

Biological Components and Chassis for Chemical Production. If we wanted to make metabolic engineering easier and faster, what standard biological components and other tools would we need? To answer this question, it is most useful to consider the essential characteristics of a microbial chemical factory. First, the microbial chemical factory should be genetically stable for long periods of cultivation, at least from scale-up to the large-scale production phases; that is, the vector used to carry the genes into and harbor the genes within the cell should be genetically stable, and the host should have minimal ability (or need) to inactivate these heterologous genes. Second, the microbial chassis should be able to grow with minimal supplements to the growth medium to minimize production costs and should be robust to industrial processing conditions (*e.g.*, from small, laboratory-scale reactors to extremely large bioreactors). As the desired chemical or an intermediate in its biosynthesis pathway might be toxic to the microbial factory, the third requirement is the ability to switch on the entire biosynthetic pathway at the correct time in the production process while having the biosynthetic pathway completely repressed during the nonproductive (growth) stages. Fourth, as most chemicals will require multiple enzymatic transformations, it is important to be able to coordinate the expression of many genes simultaneously. Fifth, a computer-aided-design (CAD) system and accurate descriptions of biological components would make the design of the entire systems from off-the-shelf components much easier. Finally, biological debugging routines would make it faster and easier to identify and correct problems in engineered microorganisms.

Chassis. A stable, well-behaved chassis (host cell) is essential for large-scale, and even laboratory-scale, production of the desired molecule. Although minimal, bacterial chassis may have significant scientific interest, industrial production of chemicals, particularly fuels and commodity chemicals, where cost of goods is most critical, requires the use of a chassis capable of synthesizing all biological macromolecules necessary for growth from a simple, inexpensive carbon source and salts of N, P, S, and the like. Hence, minimal hosts that require addition of many complex nutrients or that cannot cope with potential stresses in processing will probably not find a niche in industrial chemical/fuel production.

Robust, well-characterized organisms capable of growing on minimal, inexpensive carbon sources have been and will continue to be of enormous practical use. However, most robust hosts have the capability to inactivate or rid themselves altogether of the foreign genes to minimize the metabolic burden of the foreign DNA (*10–13*). The instability of foreign DNA in a cell is due to a combination of transposons, insertion sequence elements, defective phages, integrases, and site-specific recombinases (*10, 13, 14*). To create a strain of *Escherichia coli* that will stably maintain foreign DNA over timescales relevant to production, Blattner and colleagues (*10, 15, 16*) removed large sections of the *E. coli* chromosome, some of which contained mobile elements and cryptic virulence genes, and in the process created a stable, robust chassis capable of maintaining heterologous genes over long periods of time. The growth rate of the minimized organism was comparable to that of the wild-type hosts, and protein production was excellent. As such, these and other strains should make excellent chassis for chemical production.

Vectors. Central to any genetic manipulation is the vector used to carry and/or harbor the transforming DNA in the host. In general, the chosen genetic vector should be segregationally stable, have as low a copy number as possible to minimize burden on the host cell, and have consistent copy number in all cells of the culture to ensure product consistency. One of the most important features of the expression vector is segregational stability to ensure that all cells in the culture have the plasmid and produce the desired product. In nature, this feature has been achieved in several, unit-copy plasmids (*e.g.*, F and P1 plasmids of *E. coli*) through a partition element that segregates plasmids into each half of the mother cell before division to form two daughter cells (*17*) and through a mechanism that kills daughter cells that do not receive a plasmid at division (*18, 19*). As the latter mechanism leads to a loss of cells (and therefore a waste of metabolic energy in

KEYWORDS

Artemisinin: A sesquiterpene lactone endoperoxide extracted from *Artemisia annua* L that is highly effective against multi-drug-resistant *Plasmodium* spp.

Metabolic engineering: The manipulation of intracellular metabolic reactions for the production or degradation of a target chemical.

Synthetic biology: The design and construction of new biological components, such as enzymes, genetic circuits, and cells, or the redesign of existing biological systems.

Metabolic burden: The consumption of macromolecule precursors and energy due to the expression of foreign genes in a cell.

All-or-none induction or expression: The induction of an inducible promoter to full strength in only a fraction of the cells of a population, the fraction dependent on the concentration of the inducer in the medium.

EX2019, Page 3

synthesizing that cell), the former mechanism is preferable. Few commercially available high-copy-number plasmids have loci encoding these functions and so are relatively unstable.

It is also important that the expression vector have minimal and consistent copy number in all cells of a culture. As cells with a large number of plasmids grow more slowly than cells with a small number of plasmids, because of the additional metabolic burden, cells with a low plasmid copy number will eventually dominate the culture. Furthermore, cells with a large number of plasmids might express the genes that reside on that plasmid to a greater extent than cells with a small number of plasmids. If the genes residing on the plasmid are responsible for synthesizing a metabolic product, the heterogeneous population could produce a heterogeneous product. Consistent copy number in all cells of a culture can be achieved by timing the replication of the plasmid with the cell cycle, to ensure that the copy number has doubled before division (*20–24*), and by evenly partitioning plasmids into future daughter cells preceding division (*25*). Only the low-copy-number plasmids of *E. coli* (and, of course, the chromosome) have replication timed with the cell cycle (*20–24*).

As metabolic engineering often requires the transfer of large sequences of DNA carrying the multiple genes that might be required for the reconstitution of a heterologous pathway in a host, the expression vector must be able to carry large sequences of DNA. The copy numbers of some ColE1-type plasmids have been reported to be size dependent; a plasmid carrying a large insert will have a lower copy number than a plasmid carrying a small insert (*26*). By contrast, the naturally occurring low-copy-number plasmids of *E. coli* and other bacteria are extremely large (100–300 kb) (*27*) and can even replicate the entire chromosome when their origin of replication

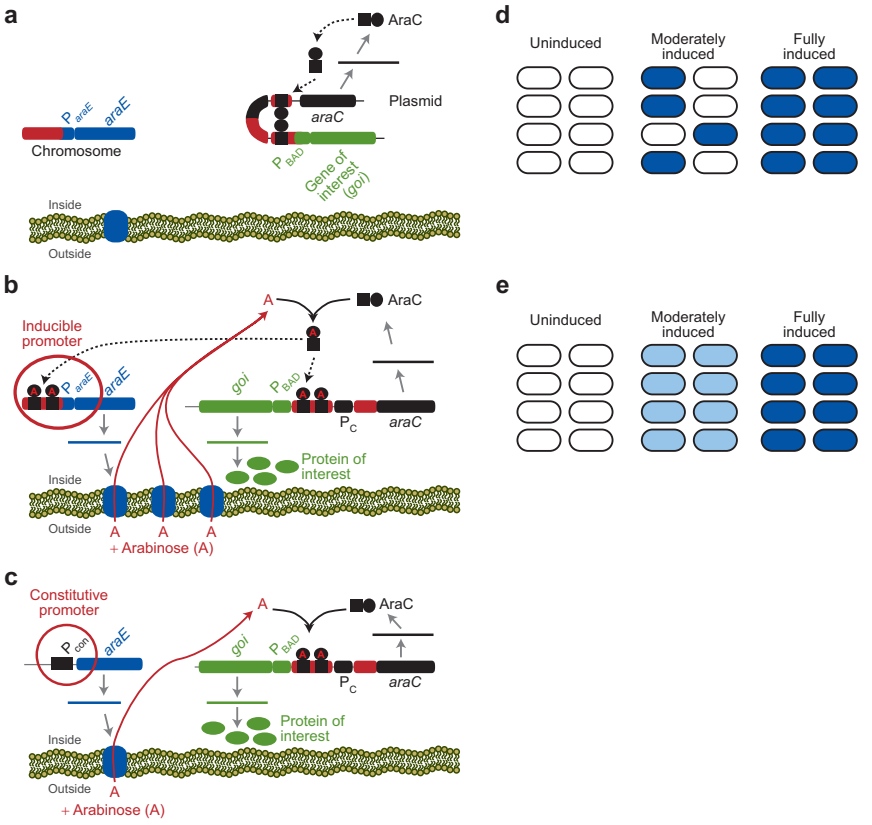

Figure 1. Achieving graded induction from the arabinose-inducible *araBAD* promoter (P_{BAD}). a) In the absence of arabinose, P_{BAD} is repressed by the action of the regulator protein AraC. b) When arabinose is added to a culture, the arabinose is transported into the cell by one of the arabinose transporters (*e.g.*, AraE) and combines with the AraC protein to derepress P_{BAD} as well as promoters controlling expression of the arabinose transporter genes. In a culture of cells containing the arabinose transporters under their native control, the fraction of the cells in the culture that is induced varies with the concentration of arabinose present in the medium (see panel d). In those cells, expression of the arabinose transporter genes is induced, which produces more transporters in the membrane and more arabinose transported into the cell, resulting in maximal expression of the gene of interest (*goi*). c) In a culture of cells containing the arabinose transporters under control of constitutive promoters, all cells have the same number of arabinose transporters (the number of transporters does not change with arabinose concentration in the medium nor does it vary substantially between induced and uninduced cells) and graded induction from the promoter is possible (see panel e). d) Variation in gene expression in cells subject to all-or-none induction of gene expression (represented in panels a and b). Cells are represented by ovals, and the level of gene expression is represented by the shading. In all-or-none induction, the fraction of cells fully induced varies with the concentration of inducer in the medium. e) Variation in gene expression in cells with homogeneous induction (represented in panel c). All cells in a culture are induced to the same level, which varies with the inducer concentration in the medium.

EX2019, Page 4

is substituted for the chromosomal origin of replication (*28*). These low-copy-number plasmids should be adequate to harbor practically any insert.

Of course the most stable expression vector is the chromosome itself. Over the past few years, several methods have been developed to integrate genes into the chromosome (*29–34*). Central to any integration strategy is the location on the chromosome into which the gene(s) of interest are inserted, as this can dramatically impact the expression level of the heterologous gene and any native gene that is interrupted in the process (*35*).

Promoters. As control of gene expression in prokaryotes is dominated by transcription and transcript processing, promoters play an essential role in controlling biosynthetic pathways. Inducible promoters are one of the easiest and most effective ways to turn on gene expression. Several inducible expression systems are now available for use in bacteria. In addition to the versatile and extensively used *lac* and arabinose expression systems, inducible promoters based on the regulatory genes for the upper and meta-cleavage pathways of toluene degradation from the TOL plasmid (*36, 37*), for naphthalene degradation from the NAH plasmid (*38*), and more recently for metabolism of propionate (*39–43*) are now available.

One of the most desirable features of the promoter is that all cells in the culture should be induced uniformly when stimulated, and the level of induction should vary directly with inducer concentration (Figure 1). Unfortunately, some of the most useful, inducible promoters (*e.g.*, lactose and arabinose-inducible promoters) exhibit all-or-none expression, in which the fraction of cells induced in the culture varies with the inducer concentration; one subpopulation of cells is fully induced, while the other subpopulation is uninduced, with the subpopulation size varying with the inducer concentration (Figure 1, panel d). This all-or-none phenomenon could have deleterious effects on the host, as those cells that express the genes would be subject to more metabolic burden (and grow more slowly) than those cells not expressing the genes. If the cells are producing a product whose composition depends on the level of inducible gene expression (*e.g.*, a polymer composed of two different kinds of monomers that impact the properties of the polymer (*44, 45*)), this all-or-none phenomenon could lead to subpopulations of cells with very different product compositions. As the all-or-none phenomenon appears to be a result of transport of the inducer into the cell, there are at least three ways to ensure (or engineer) consistent, regulatable control in all cells of a culture: uncouple transporter expression from inducer control, which has been done for the arabinose expression system (Figure 1, panel c) (*46–49*); synthesize an inducer that can readily diffuse across the cell membrane without a transporter, like IPTG; or use a natural inducer whose transport is not autoregulated, such as propionate (*50–54*).

Induction can also be achieved using promoters that are induced by starvation of cells for a required nutrient. These can be very useful for the production of a chemical during nongrowth periods. The phosphate-starvation promoter of *E. coli* has been widely used for the expression of heterologous genes (*55–59*). Promoters activated by other starvations or the stationary phase are also available (*60–63*). However, it is not currently possible (and might never be possible) to have graded expression from starvation promoters (as is possible with the addition of an inducer). The starvation promoter is generally induced to its maximal expression level (and may or may not relax to a lower level of expression) when the nutrient of which the cell is being starved is nearly depleted in the medium and imminent starvation is sensed. Even though it may not be possible to regulate expression using the limiting nutrient, it may be possible to engineer the strength of starvation promoters (as described below) to tune the level of heterologous gene expression during the starvation phases.

Constitutive promoters can play a special role in the expression of genes involved in the production of low-value chemicals or drugs, where use of an inducer would be too expensive. By varying the nucleotide sequence in the −10 or −35 region of a promoter, in the 16-bp spacer between these regions, or both, it is possible to generate a family of constitutive promoters of different strengths (*64–66*). In theory, this technique should be useful for modifying the strengths of inducible and starvation promoters so long as the sequence

KEYWORDS

Operon: A group of genes under the control of a single promoter or multiple promoters.

Tunable intergenic region (TIGR): Untranslated region between two coding regions of a multicistronic mRNA, which contains sequences that control mRNA processing and translation of one or both of the surrounding coding regions.

Internal ribosome entry sequence (IRES): A sequence between two coding regions on a multicistronic mRNA that allows a eukaryotic ribosome to initiate translation of the 3′ coding region.

CAD: Computer-aided design

EX2019, Page 5

where the controlling protein binds is not significantly altered.

Simultaneous Engagement of Multiple Genes. Construction of a metabolic pathway most often involves the introduction of more than one heterologous gene. One way to coordinate expression of multiple genes is to use different inducible promoters for each gene: gene *X* under the control of P_{lac}, gene *Y* under the control of P_{BAD}, and so on. Indeed, there are several examples of such strategies being used (*58, 59*). The disadvantage of using this control strategy is that multiple inducers need to be added to the medium, which increases the cost of production and complexity of gene expression timing. In addition, many inducible promoters have cross-talk that prevents their simultaneous use for independently controlling the expression of multiple genes (*67*). Alternatively, if it is not necessary to control each promoter independently, inducible promoters of different strengths can be generated using the method described above (*65, 66*). Then, each gene required for the complete metabolic pathway can be placed under the control of an inducible promoter of the appropriate strength so that its activity is optimized in the cell. If all of these promoters are induced by the same molecule, then only one inducer needs to be added to the culture, which should be easier than using multiple inducers; assuming that transcription is roughly the same for all genes under control of the same promoter, the genes should be expressed at the same relative levels, no matter what amount of inducer is added. However, it can be more difficult to construct this family of promoters than to use promoters induced by different molecules, as the desired expression level for each gene needs to be known *a priori* to choose a promoter of the appropriate strength.

Another strategy is to use a non-native RNA polymerase or transcription factor to control the expression of more than one gene. For example, the T7 RNA polymerase and T7 promoters of various strengths have been engineered to control and balance the expression of several genes involved in synthesis of polyketides (*68–70*). Similarly, Stephanopoulos and coworkers (*64, 71*) have randomly mutagenized the σ factor in *E. coli* to alter the regulation of many native genes simultaneously to improve metabolite production and tolerance.

Grouping multiple, related genes into operons, as is done naturally in prokaryotes (*72*), is a convenient means for regulating several genes simultaneously without the need for multiple promoters (Figure 2). Internal

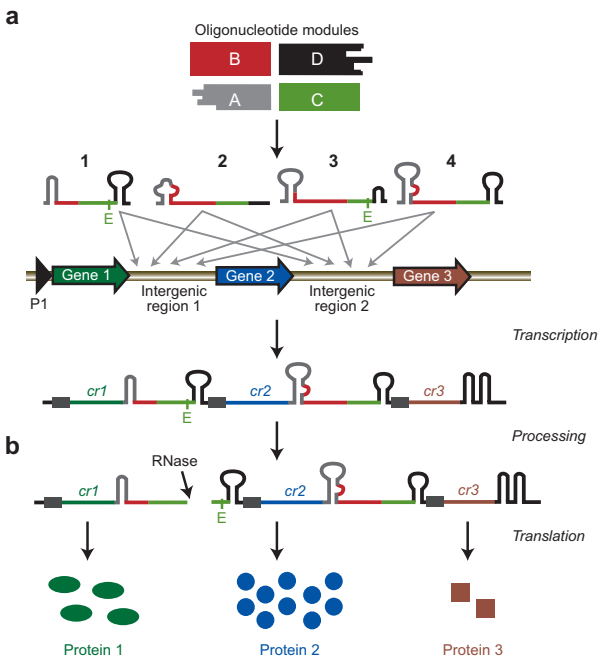

Figure 2. Combinatorial synthesis of intergenic regions from oligonucleotides. a) Intergenic regions were assembled from groups of oligonucleotides (designated A, B, C, D). Some of the oligonucleotides formed hairpins, others did not; some contained RNase sites (designated E), others did not. The oligonucleotides were assembled using PCR, the resulting intergenic regions were then used as megaprimers in a PCR reaction to attach them to the 3′ and 5′ ends of gene 2, and the entire PCR product was inserted between gene 1 and gene 2 in the plasmid. b) The various intergenic regions caused the mRNA coding regions (*cr1*, *cr2*, *cr3*) to be differentially processed and translated, resulting in different levels of the various proteins encoded in the operon.

ribosomal entry sequences from eukaryotic viruses and host stress response pathways perform a similar function and have been harnessed to create operons for heterologous expression of genes in eukaryotes (*73–76*). With a single promoter controlling the transcription of several genes, relative expression of each open reading frame in the operon is controlled by altering post-transcriptional processes such as transcription termination (*77, 78*), mRNA stability (*79, 80*), and translation initiation (*81–83*). Sequences inserted into the intergenic regions of bacterial operons can direct the processing and segmental stability of a transcript containing multiple coding regions (*84, 85*). This type of directed mRNA processing results in differential production of the pro-

EX2019, Page 6

teins encoded in the operon depending on the nature of the intergenic region between the coding regions.

One of the major obstacles to designing and implementing this type of control is the difficulty in decoupling the many interrelated variables involved in post-transcriptional regulation (*86*). We previously demonstrated that it is possible to differentially control the protein levels encoded by two or more genes in an operon using intergenic region sequences (*84, 85*). Recently, we described a system to tune the expression of several genes within operons by generating and screening large libraries of tunable intergenic regions (TIGRs) containing control elements that include mRNA secondary structures, RNase cleavage sites, and ribosome binding site sequestering sequences (*87*). We showed that TIGRs can vary the relative expression of two reporter genes over a 100-fold range and balance expression of three genes in an operon that encodes a heterologous mevalonate biosynthetic pathway, resulting in a 7-fold increase in mevalonate production. This technology should be useful for optimizing the expression of multiple genes in synthetic operons, in both prokaryotes and eukaryotes.

CAD Tools. The design of any microbial chemical factory is made difficult by the lack of simulation tools to predict the types and expression levels of genes necessary to engineer the metabolism of any cell for high-level production of a desired molecule. In almost all other areas of engineering, there are models and simulation tools that allow one to design needed components and to assemble those components into a large, functioning system (*88, 89*). Similar biological design tools are in their infancy.

Because biology is so complicated, most "design" tools are largely analytical and focused on a particular aspect of biology. For example, a number of metabolic engineers have developed models for metabolism to analyze fluxes through central metabolic pathways and to predict the impact of changes in a particular metabolic reaction or pathway on chassis growth and product formation (*90–101*). These models have been very useful for analyzing metabolic fluxes once the changes to metabolism have been made. However, very few models allow one to accurately predict the necessary changes to effect a change in flux. Unfortunately, it is still not possible with any degree of accuracy to predict the functioning of a promoter or any other component in the chassis and the actual levels of expression needed to achieve a particular flux through a reaction or pathway. Integration of flux balance models with models of genetic regulation could greatly facilitate predictive engineering of cellular metabolism.

Debugging Routines. The development of software has been made easier through the development and use of debugging software to quickly identify and correct problems with code (*102, 103*). The development of similar tools for biological debugging would reduce development times for building and optimizing engineered cells. For the development of microbial chemical factories, functional genomics can serve in the role of debugging routines (*104–106*), because the introduction of a metabolic pathway often elicits a stress response (stringent response, unfolded protein response, *etc.*) in the host (*12, 107, 108*). These stresses are reflected in mRNA and proteins expressed at that time and are thus detectable using DNA arrays or proteomics (*12, 107, 108*). Overexpression of a metabolic pathway might rob the cell of central metabolic intermediates resulting in decreased cell growth, a change in the profiles of intracellular and extracellular metabolites, which would be reflected in the metabolite profile, and a change in metabolic fluxes, which would be reflected in the flux profile. Imbalances in a metabolic pathway can result in an accumulation of potentially toxic metabolites that inhibit growth (*9*). Information from one or more of these techniques can be used to then modify expression of genes in the metabolic pathway or in the host to improve titers and/or productivity of the final product.

Synthetic Biology for Artemisinin Biosynthesis. A recent example of the use of several biological components to produce a chemical is the microbial production of the antimalarial drug artemisinin. Malaria is a global health problem that threatens 300–500 million people and kills >1 million people annually (*109*). Disease control is hampered by the occurrence of multi-drug-resistant strains of the malaria parasite *Plasmodium falciparum* (*110, 111*). Synthetic antimalarial drugs and malarial vaccines are currently being developed, but their efficacy against malaria awaits rigorous clinical testing (*112, 113*). Artemisinin, a sesquiterpene lactone endoperoxide extracted from *Artemisia annua* L (family Asteraceae; commonly known as sweet wormwood), is highly effective against multidrug-resistant *Plasmodium* spp., but it is in short supply and unaffordable to most malaria sufferers (*114*). Although total synthesis of artemisinin is difficult and costly (*115*), the semisynthesis of artemisinin or any derivative

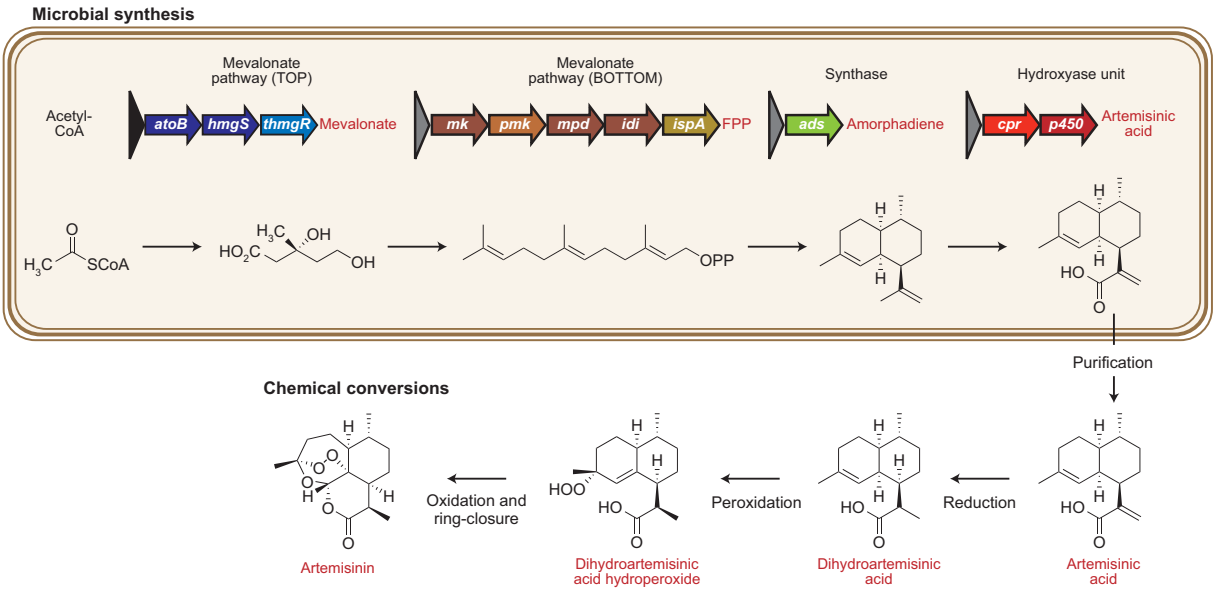

Figure 3. Semisynthesis strategy for producing artemisinin. The genes encoding the mevalonate-based FPP biosynthetic pathway were introduced from *Saccharomyces cerevisiae* (HMG-CoA synthase, *hmgS*; the N-terminally truncated version of HMG-CoA reductase, *thmgR*; mevalonate kinase, *mk*; phosphomevalonate kinase, *pmk*; and mevalonate diphosphate decarboxylase, *mpd*) and *E. coli* (acetoacetyl-CoA synthase, *atoB*; IPP isomerase, *idi*; and FPP synthase, *ispA*) into *E. coli*. The genes encoding the enzymes in the pathway were organized into two operons to allow for easy optimization. The MevT operon contains genes responsible for transforming three acetyl-CoAs into mevalonate, and the MBIS operon contains genes responsible for transforming mevalonate to FPP. In addition, the genes encoding the amorphadiene synthase (*ads*), oxidase (*p450*), and redox partners (*cpr*) were also introduced into the host strain. The microbial strain produces artemisinic acid, which can be transformed to artemisinin using established chemistry.

from microbially sourced artemisinic acid, its immediate precursor, could be cost effective, environmentally friendly, high quality, and reliable (*116, 117*).

We have engineered *E. coli* to produce high titers (up to 100 mg/L) of artemisinic acid using an engineered mevalonate pathway, amorphadiene synthase, and a novel cytochrome P450 monooxygenase (CYP71AV1) from *A. annua* that performs a three-step oxidation of amorpha-4,11-diene to artemisinic acid (Figure 3) (*9, 118–120*). The engineering endeavor made use of many of the tools and methods described above. We chose to bypass the native deoxyxylulose-5-phosphate (DXP) pathway and instead import the heterologous mevalonate pathway from *S. cerevisiae* (*121*) to supplement endogenous isopentenyl pyrophosphate (IPP) biosynthesis (*122*). The mevalonate pathway was divided into two synthetic operons. The MevT operon encodes the enzymes to transform three acetyl-CoA's to mevalonate: acetoacetyl-CoA synthase (from *E. coli*), hydroxymethylglutaryl-CoA (HMG-CoA) synthase (from *S. cerevisiae*), and HMG-CoA reductase (from *S. cerevisiae*). The MBIS operon encodes the enzymes to transform mevalonate to farnesyl pyrophosphate (FPP): mevalonate kinase, phosphomevalonate kinase, and mevalonate diphosphate decarboxylase (all from *S. cerevisiae*), and IPP isomerase and FPP synthase (both from *E. coli*). The operons were split around mevalonate because it could be added to the growth medium to select for cells that harbored a functional MBIS pathway (*9*). Alternatively, mevalonate will diffuse out of cells allowing for it to be used as an indicator for MevT activity (*8*). Both operons were constructed using the native versions of the genes, since *S. cerevisiae* codon usage is so close to that of *E. coli*, consensus *E. coli* ribosome binding sites, and secondary-structure-free spacers between the end of the 5′ gene and the ribosome binding site of the 3′ gene. The operons were placed under the control of *lac* promoters on compatible plasmids (different origins of rep-

EX2019, Page 8

lication, different antibiotic resistance genes) to ensure their maintenance and expression in the host.

Initial work showed that expression of the mevalonate pathway in *E. coli* led to severe growth inhibition that was attributed to toxicity resulting from accumulation of high levels of the prenyl diphosphates (FPP, IPP, and/or DMAPP) (*121*). Coexpression of a codon-optimized terpene synthase (*123*) from the artemisinin pathway, amorpha-4,11-diene (amorphadiene) synthase (ADS) (*124, 125*), alleviated this toxicity and was used to evaluate IPP and DMAPP production as the downstream amorphadiene product. (This same toxicity has been used to identify genes that utilize either IPP/DMAPP or FPP and thereby relieve the toxicity; in this case a phosphatase that hydrolyzes IPP or DMAPP to the corresponding alcohol, isopentenol, which incidentally may be a next-generation biofuel (*126*).) Production of amorphadiene from the imported mevalonate pathway was found to be substantially higher than production obtained with a supplemental DXP pathway (*121*), yielding titers of 280 mg L^{-1} after 24 h or 480 mg L^{-1} in a fed-batch bioreactor (*127*). Closer examination of the mevalonate pathway suggested that availability of mevalonate could limit amorphadiene production (*128*). However, when the MevT expression was increased, small-molecule-dependent growth inhibition of *E. coli* was observed that was attributed to HMG-CoA toxicity (*128*). Expression of an additional copy of *HMGR* alleviated accumulation of HMG-CoA and increased productivity of mevalonate by almost 2-fold by relieving growth inhibition (*128*). The source for the HMG-CoA toxicity has since been identified using DNA arrays and proteomics (Kizer and Keasling, unpublished results). Following up on these results, library-based engineering of the intergenic regions of the polycistronic message for MevT operon was carried out to balance expression levels for the individual genes in the MevT operon, resulting in a 7-fold overall increase in mevalonate titers due to improvements in specific production and cell growth (*129*). These efforts show that carbon can be effectively channeled for production of exogenous isoprenoids in *E. coli* even though the flux through native isoprenoid biosynthesis is relatively low. Over a 4-year period, production of amorphadiene was increased >1 million-fold (Figure 4).

This optimized, amorphadiene-producing *E. coli* was further engineered to produce artemisinic acid, which can be converted directly into the antimalarial drug artemisinin in two chemical steps (Figure 3) (*130*). Addition

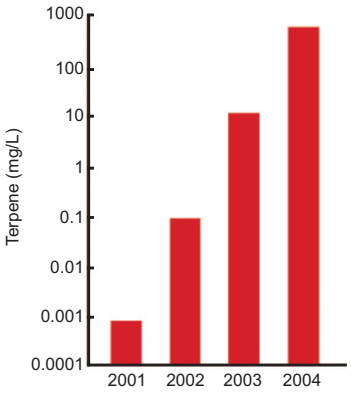

Figure 4. Progress in synthesis of amorphadiene from the period 2001–2004. Production of amorphadiene was increased >1 million-fold over that of the original terpene production strains of *E. coli* (*131*). This progress was made largely using the biological components described herein.

of an amorphadiene oxidase (AMO) and associated redox partners from *A. annua* enabled production of artemisinic acid (*118, 120*). Although the native gene (nAMO) had no detectable *in vivo* or *in vitro* activity, codon-optimization coupled with N-terminal transmembrane domain engineering generated two constructs that were competent to carry out the first oxidation step *in vivo* to generate the alcohol congener of artemisinic acid at low levels (0.18–0.45 mg L^{-1}). Use of the redox partners from *A. annua* increased productivity 12-fold to 5.6 mg L^{-1} of alcohol. Finally, use of the most appropriate promoters and expression vector allowed *in vivo* production of fully oxidized artemisinic acid at levels exceeding 300 mg/L (*118*).

Conclusions and Perspectives. As could be seen in the example above, well-characterized biological components are essential for engineering the metabolism of microorganisms. Continued development of new components and methods is critical to reduce the time required to undertake even more complicated metabolic engineering projects. With these components in hand, it should be possible to significantly reduce development times and costs of microbial chemical factories.

Acknowledgment: The research described in this review was funded through various grants from the National Science Foundation (NSF), including the NSF-funded Synthetic Biology Engineering Research Center; the National Institutes of Health; the UC Discovery Grant Program; the Bill and Melinda Gates Foundation; and the U.S. Department of Energy.

REFERENCES

1. Kauffman, G., and Chooljian, S. (2001) Friedrich Wohler (1800–1882), on the bicentennial of his birth, *Chem. Educ. 6*, 121–133.
2. Koeller, K. M., and Wong, C. H. (2001) Enzymes for chemical synthesis, *Nature 409*, 232–240.
3. Raab, R. M., Tyo, K., and Stephanopoulos, G. (2005) Metabolic engineering, *Adv. Biochem. Eng. Biotechnol. 100*, 1–17.
4. Bailey, J. E. (1991) Towards a science of metabolic engineering, *Science 252*, 1668–1675.
5. Stephanopoulos, G. N., Aristidou, A. A., and Nielsen, J. (1998) *Metabolic Engineering*, Academic Press, San Diego, CA.
6. Stephanopoulos, G. (1999) Metabolic fluxes and metabolic engineering, *Metab. Eng. 1*, 1–11.
7. Khosla, C., and Keasling, J. D. (2003) Metabolic engineering for drug discovery and development, *Nat. Rev. Drug Discovery 2*, 1019–1025.
8. Pfleger, B. F., Pitera, D. J., Newman, J. D., Martin, V. J., and Keasling, J. D. (2007) Microbial sensors for small molecules: development of a mevalonate biosensor, *Metab. Eng. 9*, 30–38.
9. Martin, V. J. J., Pitera, D. J., Withers, S. T., Newman, J. D., and Keasling, J. D. (2003) Engineering the mevalonate pathway in Escherichia coli for production of terpenoids, *Nat. Biotechnol. 21*, 796–802.
10. Posfai, G., Plunkett, G., 3rd, Feher, T., Frisch, D., Keil, G. M., Umenhoffer, K., Kolisnychenko, V., Stahl, B., Sharma, S. S., de Arruda, M., Burland, V., Harcum, S. W., and Blattner, F. R. (2006) Emergent properties of reduced-genome *Escherichia coli*, *Science 312*, 1044–1046.
11. Sanden, A., Prytz, I., Tubulekas, I., Forberg, C., Le, H., Hektor, A. P. N., Pragai, Z., Harwood, C., and Ward, A. (2003) Limiting factors in *Escherichia coli* fed-batch production of recombinant proteins, *Biotechnol. Bioeng. 81*, 158–166.
12. Gill, R. T., Valdes, J. J., and Bentley, W. E. (2000) A comparative study of global stress gene regulation in response to overexpression of recombinant proteins in *Escherichia coli*, *Metab. Eng. 2*, 178–189.
13. Kolisnychenko, V., Plunkett, G., Herring, C., Feher, T., Posfai, J., Blattner, F., and Posfai, G. (2002) Engineering a reduced *Escherichia coli* genome, *Genome Res. 12*, 640–647.
14. Frost, L. S., Leplae, R., Summers, A. O., and Toussaint, A. (2005) *Nat. Rev. Microbiol. 3*, 722.
15. Sharma, S. S., Blattner, F. R., and Harcum, S. W. (2007) Recombinant protein production in an *Escherichia coli* reduced genome strain, *Metab. Eng. 9*, 133–141.
16. Sharma, S. S., Campbell, J. W., Frisch, D., Blattner, F. R., and Harcum, S. W. (2007) Expression of two recombinant chloramphenicol acetyltransferase variants in highly reduced genome *Escherichia coli* strains, *Biotechnol. Bioeng. 98*, 1056–1070.
17. Firshein, W., and Kim, P. (1997) Plasmid replication and partition in *Escherichia coli*: Is the cell membrane the ? *Mol. Microbiol. 23*, 1–10.
18. Gerdes, K., Rasmussen, P. B., and Molin, S. (1986) Unique type of plasmid maintenance function: Postsegregational killing of plasmid-free cells, *Proc. Natl. Acad. Sci., U.S.A. 83*, 3116–3120.
19. Nordstrom, K., and Austin, S. J. (1989) Mechanisms that contribute to the stable segregation of plasmids, *Annu. Rev. Genet. 23*, 37–69.
20. Keasling, J. D., and Cooper, S. (1994) Plasmid replication during the bacterial division cycle, *Methods Mol. Genet. 3*, 380–388.
21. Keasling, J. D., Palsson, B. O., and Cooper, S. (1991) Cell-cycle-specific F plasmid replication: regulation by cell size control of initiation, *J. Bacteriol. 173*, 2673–2680.
22. Keasling, J. D., Palsson, B. O., and Cooper, S. (1992) Replication of the R6K plasmid during the *Escherichia coli* cell cycle, *J. Bacteriol. 174*, 1060–1062.
23. Keasling, J. D., Palsson, B. O., and Cooper, S. (1992) Replication of the P1 prophage is cell-cycle-specific, *J. Bacteriol. 174*, 4457–4462.
24. Keasling, J. D., Palsson, B. O., and Cooper, S. (1992) Replication of mini-F plasmids during the bacterial division cycle, *Res. Microbiol. 143*, 541–548.
25. Kuo, H., and Keasling, J. D. (1995) Plasmid replication during the bacterial division cycle, *Biotechnol. Bioeng. 52*, 633–647.
26. Scott, J. R., and Rownd, R. H. (1980) Regulation of plasmid replication. In *Mechanistic Studies of DNA Replication and Genetic Recombination* (Alberts, B., Ed.), pp 1–8, Academic Press, New York.
27. Holoway, B., and Low, K. B. (1987) F-prime and R-prime factors. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (Neidhardt, F. C., Ed.), pp 1145–1153, American Society for Microbiology, Washington, DC.
28. Tresguerres, E. F., Nieto, C., Casquero, I., and Canovas, J. L. (1986) Host cell variations resulting from F plasmid-controlled replication of the *Escherichia coli* chromosome, *J. Bacteriol. 165*, 424–427.
29. Court, D. L., Sawitzke, J. A., and Thomason, L. C. (2002) Genetic engineering using homologous recombination, *Annu. Rev. Genet. 36*, 361–388.
30. Ellis, H. M., Yu, D., DiTizio, T., and Court, D. L. (2001) High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides, *Proc. Natl. Acad. Sci. U.S.A. 98*, 6742–6746.
31. Lee, E. C., Yu, D., Martinez de Velasco, J., Tessarollo, L., Swing, D. A., Court, D. L., Jenkins, N. A., and Copeland, N. G. (2001) A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA, *Genomics 73*, 56–65.
32. Swaminathan, S., Ellis, H. M., Waters, L. S., Yu, D., Lee, E. C., Court, D. L., and Sharan, S. K. (2001) Rapid engineering of bacterial artificial chromosomes using oligonucleotides, *Genesis 29*, 14–21.
33. Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products, *Proc. Natl. Acad. Sci. U.S.A. 97*, 6640–6645.
34. de Lorenzo, V., Herrero, M., Jakubzik, U., and Timmis, K. N. (1990) Mini-Tn5 transposon derivatives for insertional mutagenesis, promoter probing, and chromosomal insertion of an upper naphthalene catabolic pathway, *J. Bacteriol. 172*, 6890–6972.
35. Sousa, C., de Lorenzo, V., and Cebolla, A. (1997) Modulation of gene expression through chromosomal positioning in *Escherichia coli*, *Microbiology 143*, 2071–2078.
36. Gallegos, M.-T., Schleif, R., Bairoch, A., Hofmann, K., and Ramos, J. L. (1997) AraC-XylS family of transcriptional regulators, *Microbiol. Mol. Biol. Rev. 61*, 393–410.
37. Gallegos, M.-T., Williams, P. A., and Ramos, J. L. (1997) Transcriptional control of the multiple catabolic pathways encoded on the TOL plasmid pWW53 of *Pseudomonas putida* MT53, *J. Bacteriol. 179*, 5024–5029.
38. Yen, K.-M., and Gunsalus, I. C. (1985) Regulation of naphthalene catabolic genes of plasmid NAH7, *J. Bacteriol. 162*, 1008–1013.
39. Lee, S. K., Newman, J. D., and Keasling, J. D. (2005) Catabolite repression of the propionate catabolic genes in *Escherichia coli* and *Salmonella enterica*: evidence for involvement of the cyclic AMP receptor protein, *J. Bacteriol. 187*, 2793–2800.
40. Lee, S. K., and Keasling, J. D. (2005) A propionate-inducible expression system for enteric bacteria, *Appl. Environ. Microbiol. 71*, 6856–6862.
41. Lee, S. K., and Keasling, J. D. (2006) Propionate-regulated high-yield protein production in *Escherichia coli*, *Biotechnol. Bioeng. 93*, 912–918.
42. Lee, S. K., and Keasling, J. D. (2006) A Salmonella-based propionate-inducible expression system for *Salmonella enterica*, *Gene 37*, 6–11.

EX2019, Page 10

43. Lee, S. K., and Keasling, J. D. (2006) Effect of glucose or glycerol as the sole carbon source on gene expression from the Salmonella prpBCDE promoter in *Escherichia coli*, *Biotechnol. Prog. 22*, 1547–1551.
44. Aldor, I., and Keasling, J. D. (2001) Metabolic engineering of poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate) composition in recombinant *Salmonella enterica* serovar Typhimurium, *Biotechnol. Bioeng. 76*, 108–114.
45. Aldor, I. S., Kim, S.-W., Jones, K. L., and Keasling, J. D. (2002) Metabolic engineering of a novel propionate-independent pathway for the production of poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate) in recombinant *Salmonella enterica* serovar Typhimurium, *Appl. Environ. Microbiol. 68*, 3848–3854.
46. Khlebnikov, A., Datsenko, K. A., Skaug, T., Wanner, B. L., and Keasling, J. D. (2001) Homogeneous expression of the P_{BAD} promoter in *Escherichia coli* by constitutive expression of the low-affinity high-capacity AraE transporter, *Microbiology 147*, 3241–3247.
47. Khlebnikov, A., and Keasling, J. D. (2001) Effect of lacY expression on homogeneity of induction from the Ptac and Ptrc promoters by natural and synthetic inducers, *Biotechnol. Prog. 18*, 672–674.
48. Khlebnikov, A., Risa, O., Skaug, T., Carrier, T. A., and Keasling, J. D. (2000) A regulatable arabinose-inducible gene expression system with consistent control in all cells of a culture, *J. Bacteriol. 182*, 7029–7034.
49. Khlebnikov, A., Skaug, T., and Keasling, J. D. (2002) Modulation of gene expression from the arabinose-inducible araBAD promoter, *J. Ind. Microbiol. Biotechnol. 29*, 34–37.
50. Lee, S. K., and Keasling, J. D. (2005) A propionate-inducible expression system for enteric bacteria, *Appl. Environ. Microbiol. 71*, 6856–6862.
51. Lee, S. K., and Keasling, J. D. (2006) Effect of glucose or glycerol as the sole carbon source on gene expression from the Salmonella prpBCDE promoter in *Escherichia coli*, *Biotechnol. Prog. 22*, 1547–1551.
52. Lee, S. K., and Keasling, J. D. (2006) A Salmonella-based, propionate-inducible, expression system for *Salmonella enterica*, *Gene 377*, 6–11.
53. Lee, S. K., and Keasling, J. D. (2006) Propionate-regulated high-yield protein production in *Escherichia coli*, *Biotechnol. Bioeng. 93*, 912–918.
54. Lee, S. K., Newman, J. D., and Keasling, J. D. (2005) Catabolite repression of the propionate catabolic genes in *Escherichia coli* and *Salmonella enterica*: evidence for involvement of the cyclic AMP receptor protein, *J. Bacteriol. 187*, 2793–2800.
55. Shin, P. K., and Seo, J.-H. (1990) Analysis of *E. coli phoA-lacZ* fusion gene expression inserted into a multicopy plasmid and host cell's chromosome, *Biotechnol. Bioeng. 36*, 1097–1104.
56. Wanner, B. L. (1996) Phosphorus assimilation and control of the phosphate regulon. In *Escherichia coli and Salmonella. Cellular and Molecular Biology* (Neidhardt, F. C., Curtiss, R., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., Eds.), pp 1357–1381, ASM Press, Washington, DC.
57. Sharfstein, S. T., Van Dien, S. J., and Keasling, J. D. (1996) Modulation of the phosphate-starvation response by genetic manipulation of the polyphosphate pathways, *Biotechnol. Bioeng. 51*, 434–438.
58. Van Dien, S. J., and Keasling, J. D. (1998) Optimization of polyphosphate degradation and phosphate secretion using hybrid metabolic pathways and engineered host strains, *Biotechnol. Bioeng. 59*, 754–761.
59. Van Dien, S. J., hani, S., Yang, C., and Keasling, J. D. (1997) Manipulation of independent synthesis and degradation of polyphosphate in *Escherichia coli* for investigation of phosphate secretion from the cell, *Appl. Environ. Microbiol. 63*, 1689–1695.
60. Connell, N., Han, Z., Moreno, F., and Kolter, R. (1987) An *E. coli* promoter induced by the cessation of growth, *Mol. Microbiol. 1*, 195–201.
61. Williams, M. D., Ouyang, T. X., and Flickinger, M. C. (1994) Starvation-induced expression of SspA and SspB: the effects of a null mutation in sspA on *Escherichia coli* protein synthesis and survival during growth and prolonged starvation, *Mol. Microbiol. 11*, 1029–1043.
62. Atkinson, M. R., Blauwkamp, T. A., Bondarenko, V., Studitsky, V., and Ninfa, A. J. (2002) Activation of the glnA, glnK, and nac promoters as *Escherichia coli* undergoes the transition from nitrogen excess growth to nitrogen starvation, *J. Bacteriol. 184*, 5358–5363.
63. Park, J. W., Jung, Y., Lee, S. J., Jin, D. J., and Lee, Y. (2002) Alteration of stringent response of the *Escherichia coli* rnpB promoter by mutations in the −35 region, *Biochem. Biophys. Res. Commun. 290*, 1183–1187.
64. Alper, H., and Stephanopoulos, G. (2007) Global transcription machinery engineering: a new approach for improving cellular phenotype, *Metab. Eng. 9*, 258–267.
65. Jensen, P. R., and Hammer, K. (1998) The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters, *Appl. Environ. Microbiol. 64*, 82–87.
66. Jensen, P. R., and Hammer, K. (1998) Artificial promoters for metabolic optimization, *Biotechnol. Bioeng. 58*, 191–195.
67. Lee, S. K., Chou, H. H., Pfleger, B. F., Newman, J. D., Yoshikuni, Y., and Keasling, J. D. (2007) Directed evolution of AraC for improved compatibility of arabinose and lactose-inducible promoters, *Appl. Environ. Microbiol. 73*, 5711–5715.
68. Mutka, S. C., Carney, J. R., Liu, Y., and Kennedy, J. (2006) Heterologous production of epothilone C and D in *Escherichia coli*, *Biochemistry 45*, 1321–3130.
69. Kennedy, J., Murli, S., and Kealey, J. T. (2003) 6-Deoxyerythronolide B analogue production in *Escherichia coli* through metabolic pathway engineering, *Biochemistry 42*, 14342–14348.
70. Murli, S., Kennedy, J., Dayem, L. C., Carney, J. R., and Kealey, J. T. (2003) Metabolic engineering of *Escherichia coli* for improved 6-deoxyerythronolide B production, *J. Ind. Microbiol. Biotechnol. 30*, 500–509.
71. Tyo, K. E., Alper, H. S., and Stephanopoulos, G. N. (2007) Expanding the metabolic engineering toolbox: more options to engineer cells, *Trends Biotechnol. 25*, 132–137.
72. Baga, M., Goransson, M., Normark, S., and Uhlin, B. E. (1988) Processed mRNA with differential stability in the regulation of *E. coli* pilin gene expression, *Cell 52*, 197–206.
73. Komar, A. A., and Hatzoglou, M. (2005) Internal ribosome entry sites in cellular mRNAs: mystery of their existence, *J. Biol. Chem. 280*, 23425–23428.
74. Martin, P., Albagli, O., Poggi, M. C., Boulukos, K. E., and Pognonec, P. (2006) Development of a new bicistronic retroviral vector with strong IRES activity, *BMC Biotechnol. 6*, 4.
75. Chappell, S. A., and Mauro, V. P. (2003) The internal ribosome entry site (IRES) contained within the RNA-binding motif protein 3 (Rbm3) mRNA is composed of functionally distinct elements, *J. Biol. Chem. 278*, 33793–33800.
76. Fernandez-Miragall, O., Ramos, R., Ramajo, J., and Martinez-Salas, E. (2006) Evidence of reciprocal tertiary interactions between conserved motifs involved in organizing RNA structure essential for internal initiation of translation, *RNA 12*, 223–234.
77. Nudler, E., and Gottesman, M. E. (2002) Transcription termination and anti-termination in *E. coli*, *Genes Cells 7*, 755–768.
78. Mandal, M., and Breaker, R. R. (2004) Gene regulation by riboswitches, *Nat. Rev. Mol. Cell Biol. 5*, 451–463.
79. Arraiano, C. M., and Maquat, L. E. (2003) Post-transcriptional control of gene expression: effectors of mRNA decay, *Mol. Microbiol. 49*, 267–276.

80. Kushner, S. R. (2004) mRNA decay in prokaryotes and eukaryotes: different approaches to a similar problem, *IUBMB Life 56*, 585–594.
81. de Smit, M. H., and van Duin, J. (1994) Control of translation by mRNA secondary structure in *Escherichia coli*. A quantitative analysis of literature data, *J. Mol. Biol. 244*, 144–510.
82. Isaacs, F. J., Dwyer, D. J., Ding, C., Pervouchine, D. D., Cantor, C. R., and Collins, J. J. (2004) Engineered riboregulators enable post-transcriptional control of gene expression, *Nat. Biotechnol. 22*, 841–847.
83. Majdalani, N., Vanderpool, C. K., and Gottesman, S. (2005) Bacterial small RNA regulators, *Crit. Rev. Biochem. Mol. Biol. 40*, 93–113.
84. Smolke, C. D., Carrier, T. A., and Keasling, J. D. (2000) Coordinated, differential expression of two genes through directed mRNA cleavage and stabilization by secondary structures, *Appl. Environ. Microbiol. 66*, 5399–5405.
85. Smolke, C. D., and Keasling, J. D. (2002) Effect of gene location, mRNA secondary structures, and RNase sites on the expression of two genes in an engineered operon, *Biotechnol. Bioeng. 80*, 762–776.
86. Deana, A., and Belasco, J. G. (2005) Lost in translation: the influence of ribosomes on bacterial mRNA decay, *Genes Dev. 19*, 2526–2533.
87. Pfleger, B. F., Pitera, D. J., Smolke, C. D., and Keasling, J. D. (2006) Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes, *Nat. Biotechnol. 24*, 1027–1032.
88. Jansen, D. (2003) *The Electronic Design Automation Handbook*, Springer-Verlag, New York.
89. Mead, C., and Conway, L. (1979) *Introduction to VLSI Systems*, p 396, Addison-Wesley Longman Publishing Co, Boston, MA.
90. Covert, M. W., Schilling, C. H., Famili, I., Edwards, J. S., Goryanin, I. I., Selkov, E., and Palsson, B. O. (2001) Metabolic modeling of microbial strains in silico, *Trends Biochem. Sci. 26*, 179–186.
91. Edwards, J. S., Ibarra, R. U., and Palsson, B. O. (2001) In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data, *Nat. Biotechnol. 19*, 125–130.
92. Varma, A., and Palsson, B. O. (1994) Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110, *Appl. Environ. Microbiol. 60*, 3724–3731.
93. Pramanik, J., and Keasling, J. D. (1997) A stoichiometric model of *Escherichia coli* metabolism: incorporation of growth-rate dependent biomass composition and mechanistic energy requirements, *Biotechnol. Bioeng. 56*, 398–421.
94. Pramanik, J., and Keasling, J. D. (1998) Effect of carbon source and growth rate on biomass composition and metabolic flux predictions of a stoichiometric model, *Biotechnol. Bioeng. 60*, 230–238.
95. Tang, Y. J., Chakraborty, R., Martin, H. G., Chu, J., Hazen, T. C., and Keasling, J. D. (2007) Flux analysis of central metabolic pathways in Geobacter metallireducens during reduction of soluble Fe(III)-nitrilotriacetic acid, *Appl. Environ. Microbiol. 73*, 3859–3864.
96. Tang, Y. J., Hwang, J. S., Wemmer, D. E., and Keasling, J. D. (2007) *Shewanella oneidensis* MR-1 fluxome under various oxygen conditions, *Appl. Environ. Microbiol. 73*, 718–729.
97. Tang, Y. J., Meadows, A. L., Kirby, J., and Keasling, J. D. (2007) Anaerobic central metabolic pathways in *Shewanella oneidensis* MR-1 reinterpreted in the light of isotopic metabolite labeling, *J. Bacteriol. 189*, 894–901.
98. Burgard, A. P., Pharkya, P., and Maranas, C. D. (2003) Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization, *Biotechnol. Bioeng. 84*, 647–657.
99. Suthers, P. F., Burgard, A. P., Dasika, M. S., Nowroozi, F., Van Dien, S., Keasling, J. D., and Maranas, C. D. (2007) Metabolic flux elucidation for large-scale models using (13)C labeled isotopes, *Metab. Eng. 9*, 387–405.
100. Christensen, B., Gombert, A. K., and Nielsen, J. (2002) Analysis of flux estimates based on (13)C-labelling experiments, *Eur. J. Biochem. 269*, 2795–2800.
101. Christensen, B., and Nielsen, J. (2000) Metabolic network analysis. A powerful tool in metabolic engineering, *Adv. Biochem. Eng. Biotechnol. 66*, 209–231.
102. Orden, A. (1952) Solution of systems of linear inequalities on a digital computer. In *ACM Annual Conference/Annual Meeting, Pittsburgh, PA*, pp 91–95, Association for Computing Machinery, Pittsburgh, PA.
103. Campbell, R. V. D. (1952) Evolution of automatic computation. In *ACM Annual Conference/Annual Meeting, Pittsburgh, PA*, pp 29–32, Association for Computing Machinery, Pittsburgh, PA.
104. Park, J. H., Lee, K. H., Kim, T. Y., and Lee, S. Y. (2007) Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation, *Proc. Natl. Acad. Sci. U.S.A. 104*, 7797–7802.
105. Park, S. J., Lee, S. Y., Cho, J., Kim, T. Y., Lee, J. W., Park, J. H., and Han, M. J. (2005) Global physiological understanding and metabolic engineering of microorganisms based on omics studies, *Appl. Microbiol. Biotechnol. 68*, 567–579.
106. Bro, C., and Nielsen, J. (2004) Impact of 'ome' analyses on inverse metabolic engineering, *Metab. Eng. 6*, 204–211.
107. Oh, M. K., and Liao, J. C. (2000) DNA microarray detection of metabolic responses to protein overproduction in *Escherichia coli*, *Metab. Eng. 2*, 201–209.
108. Oh, M. K., and Liao, J. C. (2000) Gene expression profiling by DNA microarrays and metabolic fluxes in *Escherichia coli*, *Biotechnol. Prog. 16*, 278–286.
109. *World Malaria Report 2005* (2005) World Health Organization, Geneva, Switzerland.
110. Korenromp, E. L., Williams, B. G., Gouws, E., Dye, C., and Snow, R. W. (2003) Measurement of trends in childhood malaria mortality in Africa: an assessment of progress toward targets based on verbal autopsy, *Lancet Infect. Dis. 3*, 349–358.
111. Marsh, K. (1998) Malaria disaster in Africa, *Lancet 352*, 924.
112. Hoffman, S. (2004) Save the children, *Nature 430*, 940–941.
113. Vennerstrom, J. L., Arbe-Barnes, S., Brun, R., Charman, S. A., Chiu, F. C., Chollet, J., Dong, Y., Dorn, A., Hunziker, D., Matile, H., McIntosh, K., Padmanilayam, M., Santo Tomas, J., Scheurer, C., Scorneaux, B., Tang, Y., Urwyler, H., Wittlin, S., and Charman, W. N. (2004) Identification of an antimalarial synthetic trioxolane drug development candidate, *Nature 430*, 900–904.
114. Enserink, M. (2005) Infectious diseases. Source of new hope against malaria is in short supply, *Science 307*, 33.
115. Schmid, G., and Hofheinz, W. (1983) Total synthesis of Qinghaosu, *J. Am. Chem. Soc. 105*, 624–625.
116. Acton, N., and Roth, R. J. (1992) On the conversion of dihydroartemisinic acid into artemisinin, *J. Org. Chem. 57*, 3610–3614.
117. Haynes, R. K., and Vonwiller, S. C. (1994) Cyclic peroxyacetal lactone, lactol and ether compounds. U. S. Patent 5310946.
118. Chang, M. C., Eachus, R. A., Trieu, W., Ro, D. K., and Keasling, J. D. (2007) Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s, *Nat. Chem. Biol. 3*, 274–277.
119. Newman, J. D., Marshall, J., Chang, M., Nowroozi, F., Paradise, E., Pitera, D., Newman, K. L., and Keasling, J. D. (2006) High-level production of amorpha-4,11-diene in a two-phase partitioning bioreactor of metabolically engineered *Escherichia coli*, *Biotechnol. Bioeng. 95*, 684–691.

EX2019, Page 12

120. Ro, D. K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S., Kirby, J., Chang, M. C., Withers, S. T., Shiba, Y., Sarpong, R., and Keasling, J. D. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast, *Nature 440*, 940–943.
121. Martin, V. J., Pitera, D. J., Withers, S. T., Newman, J. D., and Keasling, J. D. (2003) Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids, *Nat. Biotechnol. 21*, 796–802.
122. Kakinuma, K., Dekishima, Y., Matsushima, Y., Eguchi, T., Misawa, N., Takagi, M., Kuzuyama, T., and Seto, H. (2001) New approach to multiply deuterated isoprenoids using triply engineered *Escherichia coli* and its potential as a tool for mechanistic enzymology, *J. Am. Chem. Soc. 123*, 1238–1239.
123. Martin, V. J., Yoshikuni, Y., and Keasling, J. D. (2001) The in vivo synthesis of plant sesquiterpenes by *Escherichia coli*, *Biotechnol. Bioeng. 75*, 497–503.
124. Merke, P., Bengtsson, M., Bouwmeester, H. J., Posthumus, M. A., and Brodelius, P. E. (2000) Molecular cloning, expression, and characterization of amorpha-4,11-diene synthase, a enzyme of artemisnin biosynthesis in *Artemisia annua* L, *Arch. Biochem. Biophys. 381*, 173–180.
125. Chang, Y. J., Song, S. H., Park, S. H., and Kim, S. U. (2000) Amorpha-4,11-diene synthase of *Artemisia annua*: cDNA isolation and bacterial expression of a terpene synthase involved in artemisinin biosynthesis, *Arch. Biochem. Biophys. 383*, 178–184.
126. Withers, S. T., Gottlieb, S. S., Lieu, B., Newman, J. D., and Keasling, J. D. (2007) Identification of isopentenol biosynthetic genes from Bacillus subtilis by a screening method based on isoprenoid precursor toxicity, *Appl. Environ. Microbiol. 73*, 6277–6283.
127. Newman, J. D., Marshall, J., Chang, M., Nowroozi, F., Paradise, E., Pitera, D., Newman, K. L., and Keasling, J. D. (2006) High-level production amorpha-4,11-diene in a two-phase partitioning bioreactor of metabolically engineered *Escherichia coli*, *Biotechnol. Bioeng. 95*, 684–691.
128. Pitera, D. J., Paddon, C., Newman, J. D., and Keasling, J. D. (2007) Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*, *Metab. Eng. 9*, 193–207.
129. Pfleger, B. R., Pitera, D. J., Smolke, C. D., and Keasling, J. D. (2006) Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes, *Nat. Biotechnol. 24*, 1027–1032.
130. Roth, R. J., and Acton, N. (1989) A simple conversion of artemisinic acid into artemisinin, *J. Nat. Prod. 52*, 1183–1185.
131. Martin, V. J. J., Yoshikuni, Y., and Keasling, J. D. (2001) The in vivo synthesis of plant sesquiterpenes in *Escherichia coli*, *Biotechnol. Bioeng. 75*, 497–503.