New and improved tools and methods for enhanced biosynthesis of natural products in microorganisms

Zhiqing Wang¹, Patrick C. Cirino¹

1. Department of Chemical and Biomolecular Engineering, University of Houston, Houston, TX 77204, USA

Abstract

Engineering efficient biosynthesis of natural products in microorganisms requires optimizing gene expression levels to balance metabolite flux distributions and to minimize accumulation of toxic intermediates. Such metabolic optimization is challenged with identifying the right gene targets, and then determining and achieving appropriate gene expression levels. After decades of having a relatively limited set of gene regulation tools available, metabolic engineers are recently enjoying an ever-growing repertoire of more precise and tunable gene expression platforms. Here we review recent applications of natural and designed transcriptional and translational regulatory machinery for engineering biosynthesis of natural products in microorganisms. Customized trans-acting RNAs (sgRNA, asRNA and sRNA), along with appropriate accessory proteins, are allowing for unparalleled tuning of gene expression. Meanwhile metabolite-responsive transcription factors and riboswitches have been implemented in strain screening and evolution, and in dynamic gene regulation. Further refinements and expansions on these platform technologies will circumvent many long-term obstacles in natural products biosynthesis.

Introduction

Metabolic engineers enjoy an expanding genetic tool set, characterized biosynthetic pathways, and genetically “friendly” host organisms, enabling an expanding biochemical product[1]. In addition to the common difficulties associated with functionally expressing enzymes and metabolic pathways involved in the synthesis of pharmaceutically-relevant natural products in a genetically amenable host, the greater biosynthetic complexity associated with these compounds curbs progress in their overproduction via metabolic engineering, compared to many simpler compounds (and pathways) more compatible with the host’s native metabolism. Often, increased production is not simply correlated with increased pathway gene expression. Accordingly, grand challenges in metabolic engineering that are of particular relevance in natural products biosynthesis include identifying appropriate genes to target for altered expression, balancing flux distribution of key metabolites toward biosynthesis and cell maintenance, and minimizing accumulation of toxic intermediates or stresses associated with pathway gene expression.

Advances in synthetic biology tools that allow for more precise, flexible and dynamic control over expression of targeted genes have facilitated progress in microbial production of natural products. This review focuses on the recent use of transcriptional and translational regulatory machinery for improving natural products biosynthesis in microbial hosts. Attention is paid to demonstrated applications of these tools in the context of metabolic engineering, rather than tool development. It is also worth noting that while the pharmaceutical or other commercial relevance of some natural products described here may be debatable, their biosynthetic pathways are nonetheless relevant as they share key enzymes, regulatory elements, and/or precursor metabolites with many other natural product biosynthetic pathways. We broadly categorize the regulators as inducible *trans*-acting RNA platforms (not responsive to metabolites within the biosynthetic pathways), and metabolite-responsive regulators. These distinct approaches offer different advantages. For example, non-metabolite-responsive gene regulators can serve as inducible and reversible gene silencers to improve biosynthesis by reducing (or eliminating) host

pathways that compete for resources. Meanwhile regulators that act in response to metabolites within a biosynthetic pathway can dynamically regulate pathway genes, for example to prevent toxic accumulation of intermediates. Metabolite-responsive regulators can also be used as genetically-encoded molecular reporters or mediators of a selective growth advantage, which in turn enable identification of new gene targets and pathway engineering.

Inducible *trans*-Acting RNA Platforms

CRISPR-based Regulators

The RNA-guided DNA targeting and cleavage associated with the CRISPR (clustered regularly interspaced short palindromic repeats) bacterial adaptive immunity systems[2,3] have been repurposed as regulators of gene expression[4,5]. Cas9 of the Type II CRISPR system is the most well-characterized CRISPR protein[3]. In this system, a synthetic *trans*-acting RNA called “single guide RNA” (sgRNA) contains a hairpin “handle” that interacts with Cas9, a base-pairing RNA segment (~20 bases) that directs the Cas9-sgRNA complex to the desired complementary DNA segment, and a transcription terminator (refer to Figure 1). Insertion of two silencing mutations in the Cas9 nuclease domains yielded a catalytically inactive variant called dCas9[4]. Coexpression of dCas9 and sgRNAs inhibits transcription specifically at DNA sequences complementary to the sgRNA base-pairing segments, without cleaving the DNA[4]. This transcriptional-level gene silencing tool was named CRISPRi (CRISPR interference) (Figure 1a). Key features of CRISPRi, as demonstrated in *E. coli*, include i) ability to tune the extent of gene silencing (e.g. by changing the sgRNA sequence or the number of sgRNAs targeting a gene); ii) ability to reverse gene silencing (e.g. by turning off sgRNA expression); and iii) ability to silence multiple genes simultaneously (using multiple sgRNAs). dCAS9 has also been fused to transcriptional activators,

resulting in sgRNA-guided gene activation (Figure 1b,c) [6–8]. Below we summarize recent applications of CRISPRi technologies to engineer natural product biosynthesis in bacteria and yeast.

Acetyl-CoA and malonyl-CoA are valuable precursor metabolites used in both anabolic metabolism (e.g. fatty acid biosynthesis) and natural product biosynthesis (e.g. flavonoid and polyketide synthesis). Flux of these metabolites must be balanced between chemical synthesis and cellular demands. To improve intracellular availability of acetyl- and malonyl-CoA for biosynthesis of the plant flavonoid naringenin in *E. coli* engineered to express this pathway, Cress *et al.* targeted for CRISPRi-based repression the transcription factor (TF) FadR, which activates fatty acid biosynthesis and represses β -oxidation[9]. In the same study, a computational metabolic optimization method (OptForce[10]) was used to identify four additional gene targets (*fumC*, *sucC*, *sucD* and *scpC*) that were repressed using CRISPRi to further enhance naringenin production in *E. coli* (up to 2.5-fold). Similarly, in a study conducted by Wu *et al.*, thirty different *E. coli* genes known to be involved in acetyl- and malonyl-CoA metabolism were individually repressed by CRISPRi to study their effects on the intracellular concentrations of these -CoAs[11]. Sixteen individual genes (when silenced) were found to increase the intracellular concentrations of acetyl-CoA and/or malonyl-CoA by more than 2-fold. Importantly, most of these genes are involved in central metabolic pathways, and simple chromosomal deletions of many of these could be lethal or very compromising (depending on the growth medium). In order to optimally balance between cell growth and malonyl-CoA availability, they then designed sgRNAs for high-, medium-, and low-silencing efficacy targeting twelve candidate genes, and compared the silencing effects on cell growth and malonyl-CoA concentrations. Finally, the top target genes were selected for combinatorial gene silencing, resulting in 7.4-fold improved naringenin biosynthesis compared to the control *E. coli* strain.

CRISPRi was similarly used to direct succinyl-CoA flux to a representative metabolic pathway (biosynthesis of 4-hydroxybutyrate (4-HB), which in turn is incorporated into polyhydroxyalkanoates) by adjusting the expression levels of genes involved in succinyl-CoA utilization in *E. coli*[12]. sgRNAs of

different silencing efficacy were designed for each of five endogenous genes, and the resulting transcript levels were measured. The best combination of five sgRNAs increased 4HB content within their biosynthesized P(3HB-co-4HB) copolymer by ~13-fold.

Koffas and coworkers used CRISPRi to construct synthetic, orthogonal repressor-promoter pairs[13]. A 4-nucleotide region in the T7 promoter was randomized and screened for varied promoter strengths. The authors then designed CRISPR spacer sequences for eleven of these mutated promoters and characterized the resulting dCas9 repression efficacy and repressor-promoter orthogonality. As a proof-of-concept, three orthogonal synthetic repressor-promoter pairs were used to tune carbon flux through the branched, multi-gene violacein pathway expressed in *E. coli*. Branch points in this pathway result in production of four possible compounds, and different combinations of repressors regulating pathway genes resulted in different product distributions.

Zalatan *et al.* developed the dCas9 system to act as a gene-specific transcriptional activator in yeast by fusing sgRNAs to RNA hairpin sequences that recruit specific RNA-binding proteins, and then fusing the corresponding RNA-binding proteins to a transcriptional activator domain (e.g. VP64) (Figure 1c)[8]. The sgRNA fusions were termed scaffold RNAs (scRNA). They demonstrated no crosstalk between different scRNA-binding protein pairs, multivalent recruitment of activation domains, and stronger gene activation compared to using an unmodified sgRNA with the dCas9-VP64 fusion protein. Transcriptional repression was similarly demonstrated by fusing an RNA-binding protein to a gene silencing domain (Figure 1d). This system was used to design an RNA program such that different combinations of three scRNAs (regulating three different genes) directed flux to different products of the above-mentioned bacterial violacein biosynthetic pathway, expressed in *Saccharomyces cerevisiae*. Regulated expression of dCas9 by a galactose-inducible promoter provided an added level of control over scRNA-encoded circuits.

Actinomycetes are among the most important producers of pharmaceutically-relevant secondary metabolites, but technical challenges in genetically modifying these organisms have limited progress in engineering their metabolism. Lee and coworkers recently developed CRISPR-Cas9 and the corresponding CRISPRi systems that work in the model actinomycete *Streptomyces coelicolor* (commonly used for polyketide antibiotics production), as demonstrated by manipulating this strain's actinorhodin biosynthetic gene cluster[14]. This powerful and broadly applicable technology should enable rapid progress in actinomycete metabolic engineering.

Antisense RNA

Antisense RNA (asRNA) is a small *trans*-acting RNA fragment that binds to a complementary 5' untranslated region (5'-UTR) and/or open reading frame (ORF) of transcribed mRNA and suppresses gene expression by blocking translation (refer to Figure 1e) [15]. asRNA tools have gained popularity for use in *E. coli*. A library of asRNAs that target 79 essential genes in *E. coli* has been reported [16], as have inducible asRNA vectors targeting 71 genes involved in *E. coli* central metabolism [17]. These constructs can be readily applied for metabolic engineering purposes.

To enhance naringenin production in *E. coli*, Wu *et al.* designed multiple asRNAs targeting *fabB* and *fabF*, two genes involved in malonyl-CoA consumption for fatty acid biosynthesis[18]. They studied how different asRNAs targeting different regions of *fabB* and *fabF* mRNA affected intracellular malonyl-CoA levels. An inducible promoter allowed for controlled expression of the asRNAs at a suitable culture density / growth phase. The best combination of asRNAs increased naringenin production by 4.3-fold, compared to the control strain without asRNAs. Similarly, Yan and coworkers used asRNA in *E. coli* to downregulate *fabD* (encoding malonyl-CoA-ACP transacylase; initiating fatty acid biosynthesis), resulting in a 4.5-fold increased intracellular malonyl-CoA concentration[19]. A heterologous biosynthetic pathway that produces 4-hydroxycoumarin (a malonyl-CoA-derived precursor to warfarin)

was expressed in *E. coli*, and inclusion of the *fabD* asRNA increased 4-hydroxycoumarin production by 2.5-fold, without impairing cell growth. Downregulation of *fabD* using this same asRNA also increased biosynthesis of the phenylpropanoid resveratrol by 1.70-fold, and naringenin by 1.53-fold (both via heterologous biosynthesis pathways).

Scalcinati *et al.* used asRNA-based repression in *S. cerevisiae* to increase production of α -santalene, a compound commonly used in perfumery and aromatherapy industries [20]. The authors produced α -santalene by introducing a plant-derived santalene synthase (SanSyn) that catalyzes the conversion of farnesyl pyrophosphate (FPP) to α -santalene. FPP is an essential metabolite used in the biosynthesis of many physiologically important compounds, including the steroid precursor squalene. The authors downregulated squalene synthase (SQS, encoded by *ERG9* gene) to enhance availability of FPP for α -santalene production. Notably, *ERG9* is an essential gene in *S. cerevisiae*[21]. A glucose-responsive promoter, HXT2 (induced when glucose concentrations are low[22]), was used to express an asRNA fragment complementary to the 5'-UTR of *ERG9* mRNA, as well as part of the *ERG9* ORF. As a result, *ERG9* expression was moderately repressed during exponential phase (high glucose level) and maximal repression of *ERG9* was achieved during stationary phase (low glucose level). This strategy successfully increased α -santalene production in *S. cerevisiae* by 3.4-fold. Note that in this example gene silencing (via asRNA expression) is metabolically induced. While simple in concept, this potentially powerful approach adds a layer of complexity to the gene circuitry, but is expected to be widely adopted.

Small Regulatory RNA

trans-Acting small regulatory RNAs (sRNAs) are comprised of a targeting antisense RNA that binds complementary mRNA, and a scaffold that recruits host factor Hfq, which can subsequently induce degradation of the mRNA via RNase E (refer to Figure 1f) [23]. Lee and coworkers developed a system to design temperature-inducible synthetic sRNAs that comprise three parts: the *E.coli* MicC sRNA

scaffold, a 24-nt antisense segment, and a terminator[24]. This group designed synthetic sRNAs to improve production of tyrosine and cadaverine in *E. coli*[25]. They showed that synthetic sRNA binding energy correlates well with repression efficiency, and so gene expression can be fine-tuned by rational design of sRNA. The authors tested sRNA-based combinatorial knockdown of four genes involved in tyrosine metabolism, using 14 different *E. coli* strains, resulting in dramatically enhanced tyrosine production. Next, in an *E. coli* strain engineered to convert lysine to cadaverine, sRNAs were designed for eight genes potentially involved in diversion of cadaverine biosynthesis. The best-performing sRNA was *anti-murE* (sRNA to repress *murE*), which increased cadaverine production by >150%. In a similar study, 128 different synthetic sRNAs were screened for increased production of 1,3-diaminopropane by engineered *E. coli* [26]. Five genes were able to increase production by up to 35%, and two of these are essential genes (which could not be deleted).

Meng *et al.* used metabolic flux distribution analysis to identify 23 genes involved in the pentose phosphate and nucleotide metabolism pathways as targets for sRNA suppression, to improve production of the polyketide 6-deoxyerythronolide B (6dEB, a precursor to the antibiotic erythromycin) in *E. coli*[27]. The effects of downregulating each individual gene were assessed, leading to selection of three genes from the pentose phosphate pathway and six genes from nucleotide metabolism to co-repress in combinations (one from each pathway). The best combination led to nearly 3-fold increased production in 6dEB, and the highest yield yet reported for polyketide heterologous biosynthesis in *E. coli*.

Finally, Liu *et al.* applied the sRNA strategy to increase *N*-acetylglucosamine (GlcNAc) production in *Bacillus subtilis*[28]. GlcNAc synthesis, glycolysis, and peptidoglycan synthesis pathways compete for the same precursors. In order to direct flux to GlcNAc synthesis, two sRNAs targeting *pfk* (encoding 6-phosphofructokinase) and *glmM* (encoding phosphoglucosamine mutase), were designed to limit glycolysis and peptidoglycan synthesis, respectively. To achieve tighter repression, *E. coli* Hfq

protein was coexpressed. Simultaneous repression of *pfk* and *glmM* with sRNA increased GlcNAc yield 4.3-fold.

Metabolite-responsive Gene Regulators

Transcriptional Regulatory Proteins

Allosteric transcriptional regulatory proteins (or TFs) that respond to small molecule inducers (or effectors) are ubiquitous in nature and have been readily adapted as powerful tools for targeted control over gene expression and as endogenous molecular biosensors. Both of these applications of regulatory proteins have enabled advances in metabolic engineering (previous reviews include [29,30]). Recently, Chou and Keasling developed a feedback-regulated evolution of phenotype (FREP) strategy[31] in which mutator gene *mutD5* (whose expression reduces mismatch repair[32]) is under the control of a TF that represses transcription when in complex with the target metabolite. Low intracellular concentrations of the target metabolite (effector of repression) triggers relatively high mutation rates and expression of a reporter gene. Mutation-facilitated evolution of effector biosynthesis correlates with decreased mutation rate and reporter gene expression. FREP was used to increase lycopene production in *E. coli* by increasing availability of the precursor isopentenyl diphosphate (IPP). Here, a synthetic TF containing an IPP recognition domain fused to a DNA binding domain was designed. Using their feedback-regulated evolution construct lycopene production was increased several fold after 24-hour continuous evolution and after a longer term (432 hour) evolution with intermittent isolation and passage of only the highest producers. FREP was similarly demonstrated in the evolution of tyrosine production by *E. coli* (using a tyrosine-dependent transcriptional repressor).

Church and coworkers characterized the operational range and “escape rate” of a variety of regulatory proteins as sensors of chemical inducers (including several flavonoids) in the context of genetic selections[33]. They then implemented a dual selection strategy whereby expression of the TolC membrane protein (conferring resistance to sodium dodecyl sulfate under positive selection, and sensitivity to colicin E1 during negative selection) is under control of the TtgR repressor, inducible by the flavonoid naringenin. Several rounds of multiplex automated genome engineering (MAGE)[34] targeting genes involved in naringenin biosynthesis, each followed by “toggled” (positive/negative) selection afforded an *E. coli* strain that produced 36 times more naringenin than the parent strain. The same strategy was implemented to improve glucaric acid production in *E. coli*, this time with TolC expression under the control of glucaric acid-activated CdaR regulator.

Siedler *et al.* identified FdeR from *Herbaspirillum seropedicae* SmR1 as a transcriptional activator in response to naringenin, and repressor QdoR from *B. subtilis*, inducible by the flavonoids quercetin and kaempferol[35]. Using these TFs to control fluorescent protein expression, biosensors for their respective effectors were constructed and characterized in *E. coli*. The QdoR-biosensor showed a tight correlation between cellular fluorescence and *in vivo* conversion of dihydrokaempferol to kaempferol, demonstrating the value of such biosensors for gene identification and pathway engineering.

Keasling and coworkers designed and characterized TF/promoter sensor systems that respond to short/medium-chain alcohols and dicarboxylic acids in *E. coli*[36]. In this work the TF BmoR from *Thauera butanivorans* was shown to activate gene expression in *E. coli* in response to 1-butanol, and a BmoR-based reporter system was used to screen for enhanced production of alcohols derived from 2-keto acid biosynthesis in *E. coli*. Similarly, several recent studies have adapted previously characterized TFs that regulate gene expression in response to specific amino acids as sensors for screening *Corynebacterium glutamicum* mutants that overproduce these amino acids sensors[37–40].

With the understanding that some intermediate metabolites within a biosynthesis pathway are toxic (especially when accumulated), Keasling and coworkers sought to dynamically regulate metabolic pathways using promoters responsive to such toxic intermediates [41]. Accumulation of FPP in isoprenoid biosynthetic pathways expressed in *E. coli* leads to toxicity that impairs productivity. To identify FPP-responsive promoters in *E. coli*, transcriptomes were compared between a strain engineered to accumulate FPP (using a biosynthetic pathway that begins with mevalonate) and another that produces FPP but also converts it to the non-toxic isoprenoid product amorphaadiene (via enzyme ADS). Two resulting promoters were characterized: gene expression from P_{gadE} is downregulated in response to FPP, while gene expression is upregulated from P_{rstA} in response to FPP. To dynamically regulate amorphaadiene biosynthesis, P_{gadE} was used to regulate the FPP biosynthesis pathway, while P_{rstA} regulated expression of ADS. The resulting amorphaadiene titer was 2-fold higher than that from the pathway controlled by two IPTG-inducible promoters. The same approach was used in the same work to identify other promoters that are instead responsive to HMG-CoA, another toxic intermediate involved in many isoprenoid biosynthetic pathways. Gene expression from promoter P_{fapB} was found to increase in response to HMG-CoA.

FapR from *B. subtilis* is a repressor protein that binds the *fapO* operator sequence, and releases the operator upon binding malonyl-CoA[42,43]. Koffas and coworkers incorporated *fapO* into a hybrid T7 promoter, resulting in a malonyl-CoA-sensing device that could control and optimize malonyl-CoA-dependent biosynthetic pathways in *E. coli*[44]. This group then discovered that the FapR protein could activate gene expression from the *E. coli* promoter P_{GAP} in the absence of malonyl-CoA, that malonyl-CoA inhibits this activation, and that the dynamic range (in response to malonyl-CoA) can be tuned by incorporating *fapO* sites within the P_{GAP} promoter. A fatty acid synthesis pathway (FAS) representing a malonyl-CoA metabolic sink was next placed under control of the hybrid T7-*fapO* promoter, and the genes encoding acetyl-CoA carboxylase (ACC), whose expression results in overproduction of malonyl-CoA, were placed downstream of P_{GAP} (depicted in Figure 2a). Upon constitutive FapR expression, the

resulting genetic circuit provided dynamic pathway control that improved fatty acid production relative to the “uncontrolled” strains. Zhang and coworkers similarly constructed a malonyl-CoA-sensing negative feedback controller (Figure 2b), this time by placing ACC expression under control of a LacI-repressible promoter, and LacI expression under control of a promoter repressed by FapR (at *fapO*) in the absence of malonyl-CoA[45]. The enhanced fatty acid productivity using this feedback circuit was attributed to alleviated toxicity associated with ACC overexpression.

Gene mining and screening for putative TFs can be a powerful approach to identify novel ligand-responsive regulators and promoters (e.g. [41,46,47]). Alternately, one may seek altered/customized effector specificity (and sensitivity) of a characterized TF, using protein engineering methods. Many such examples exist, including recent studies that have yielded sensors for octanoyl-homoserine lactone[48,49], vanillin[50], 4-de-dimethylamino-anhydrotetracycline[51], alkanes with different chain lengths[52], 2,4-dinitrotoluene[53], fucose, gentiobiose, lactitol, and sucralose[54]. Engineering TFs as biosensors has also been recently reviewed[30].

Demonstrated applications of such customized regulatory proteins as sensors in natural products biosynthesis are scarcer. Our group has focused on altering ligand recognition by the *E. coli* regulatory protein AraC[55–57]. An AraC-based mevalonate sensor was used to screen a heterologous mevalonate pathway operon gene library for enhanced mevalonate production by *E. coli*[56]. Monitoring mevalonate depletion using this same biosensor could prove useful for engineering mevalonate-dependent isoprenoid biosynthetic pathways. An AraC-based biosensor that is responsive to triacetic acid lactone (TAL) was similarly constructed and used to screen a library of plant-derived 2-pyrone synthase (a polyketide synthase that naturally produces TAL) variants in *E. coli*, resulting in ~20-fold increased TAL production in *E. coli*[57]. Continuing with this theme, Chen *et al.* isolated an AraC variant that responds to ectoine, which has cosmeceutical and pharmaceutical values. The ectoine reporter was used to screen a library of mutants from the heterologous gene cluster responsible for ectoine biosynthesis from L-aspartate- β -

semialdehyde, resulting in an *E. coli* clone showing 3.3-fold increased ectoine production compared to the parent strain[58]. AraC-based biosensors that respond to a variety of other small molecules such as vanillin and salicylic acid have more recently been developed (manuscripts in preparation).

Riboswitches

Riboswitches are a class of RNA-based regulatory elements employed by eukaryotes and prokaryotes [59]. A typical riboswitch contains an aptamer domain (input domain) that binds a metabolite effector, and a gene-regulation domain (output domain) whose secondary structure dictates transcription/translation activity and is dependent on the corresponding aptamer's structure (and hence ligand binding status) (Figure 3a,b). Aptazymes are a variation of the riboswitch, in which the output domain is a hammerhead ribozyme whose activity depends on aptamer binding status (Figure 4c)[60]. Riboswitch mechanisms and variations have been reviewed[59], and methods to design the aptamer and regulation domains have been described[61,62]. Building synthetic riboswitches is an appealing alternative to engineering ligand recognition in regulatory proteins due to the modularity between domains, and the absence of challenges associated with functional protein expression.

Michener and Smolke used an aptazyme that incorporates an archetypal aptamer for theophylline, to facilitate high-throughput screening for bacterial P450 monooxygenase (BM3) variants with enhanced demethylase activity in the conversion of caffeine to theophylline in *S. cerevisiae* (here, increased intracellular theophylline concentrations correlated with increased GFP expression from the riboswitch) [63]. After multiple rounds of directed evolution, the best BM3 variant increased theophylline production by ~23-fold compared to the parent. Eckdahl *et al.* similarly used caffeine demethylase activity as a metric for desired metabolic output in *E. coli*, using a different riboswitch (developed previously[64]) in which theophylline binding is coupled to expression of TetA, conferring tetracycline resistance[65].

The natural aptazyme *glmS* found in many bacteria triggers self-cleavage upon binding glucosamine 6-phosphate (GlcN6P), a precursor of N-acetyl glucosamine (GlcNAc). Lee and Oh inserted this aptazyme into the 3'-UTR of FCY1 (encoding cytosine deaminase, which converts fluorocytosine to the toxic growth inhibitor fluorouracil) in *S. cerevisiae*[66]. Using this aptazyme, the authors screened a random mutagenesis library of an endogenous fructose-6-phosphate amidotransferase (GFA1) for enhanced production of GlcN6P, based on the size of colonies growing in the presence of fluorocytosine. The isolated clones went through another round of screening by comparing their growth rates in the presence of fluorocytosine. Using the *S. cerevisiae* strain harboring their improved GFA1 variant and employing the same aptazyme, the authors then screened a library of *E. coli* haloacid dehalogenase enzymes for ability to convert GlcNAc6P to GlcNAc, this time screening for reduced cell growth on fluorocytosine (i.e. depletion of GlcNAc6P). A resulting clone showed ~4-fold increased GlcNAc production compared to control strains.

Lysine-responsive riboswitches have been used to engineer lysine-overproducing strains. Yang *et al.* constructed a lysine-OFF riboswitch (represses gene expression upon binding of lysine) using the lysine-sensitive 5'-UTR sequence of the *E. coli lysC* gene[67]. This riboswitch was added upstream of *tetA* to facilitate high-throughput selection for lysine overproducers in the presence of NiCl₂ (here, expression of TetA renders the cells sensitive to NiCl₂). A promoter library was constructed for *ppc*, encoding phosphoenolpyruvate carboxylase, aiming to optimize the expression level of this enzyme for improved lysine production. Four rounds of enrichment yielded *E. coli* strains with dramatically increased lysine production compared to the parent strain. Riboswitches for other amino acids should be similarly useful for engineering other overproduction strains.

Applying the aforementioned lysine-OFF riboswitch, Wang *et al.* evolved a chimeric aspartate kinase (AK) for improved lysine production in *E. coli*[68]. AK catalyzes the first step of aspartate-derived amino acid biosynthesis, including lysine, but is regulated by end-metabolites via feedback inhibition. BT

is a chimeric AK resistant to lysine-feedback inhibition[69], but having insufficient activity to afford substantial lysine production in *E. coli*. The authors constructed a random mutagenesis library of BT and used the lysine-OFF switch to select for BT variants that conferred increased lysine production. *E. coli* expressing an evolved BT variant produced lysine titers similar to commercially-used lysine-producing *E. coli*.

Conclusions

Enhanced control over gene expression in microorganisms, along with increasing identification and understanding of which genes to target, are enabling continued improvements in the biosynthesis of many natural products. Whereas more traditional metabolic engineering efforts have primarily relied on a small number of inducible promoters and host gene knockouts, ncRNA-based systems that allow for temporal and tunable control over gene silencing are a powerful alternative, e.g. when targeting many genes simultaneously, or essential genes. Use of *trans*-acting ncRNA should prove to be time saving and portable across different hosts, and especially useful when in the stage of identifying candidate genes to improve biosynthesis. New TFs that respond to different effector molecules are continuously being identified and characterized, and many have already proven useful in metabolic engineering efforts. Meanwhile a recent surge in the design of TF-based biosensors having customized effector specificity and sensitivity suggests a forthcoming surge in their use as biosensors/reporters, and as dynamic pathway controllers. Notably, the different approaches described here may be combined, for example by placing sgRNA/dCas9 or asRNA under control of a metabolite-responsive regulator, thereby allowing for dynamic regulation of pathways containing that metabolite. Such next-generation applications should become mainstream as users grow comfortable with these new tools.

Acknowledgements: Financial support was provided by the National Science Foundation (awards CBET 1511425 and CBET 1159398).

Figure 1. RNA-guided gene regulators. (a) dCas9 complexed with a sgRNA acts as a transcriptional repressor. (b) When dCas9 is fused with RNA polymerase (RNAP) ω subunit, the complex binds to target DNA, facilitates RNAP assembly and activates gene expression. (c) A scaffold RNA can be fused 3' of the sgRNA to recruit RNA binding proteins. When the RNA binding protein is fused with a transcription activator (TA), binding of dCas9 complex to DNA activates gene expression. (d) When the RNA binding protein is fused with a transcription repressor (TR), dCas9 complex represses gene expression (Figure 1a-d adapted from reference 8). (e) Binding of asRNA to complementary mRNA blocks the ribosome binding site (RBS) and stops translation initiation. (f) The antisense segment (red) of a sRNA recognizes the targeted sequence of the mRNA, and the recruiting module (black) recruits Hfq protein, which subsequently triggers mRNA degradation by Rnase E.

Figure 2. Two examples of dynamic regulation. (a) Both production and consumption of malonyl-CoA are regulated by the malonyl-CoA-responsive TF FapR. When the malonyl-CoA level is high, expression of ACC is low and the desired malonyl-CoA-dependent production pathway is high, but when malonyl-CoA levels are low the production pathway is repressed so malonyl-CoA is kept available for cell growth (reference 44). (b) Synthesis of malonyl-CoA is controlled by feedback regulation conferred by FapR and LacI. Malonyl-CoA is only synthesized when its intracellular level is low. IPTG is used to initially induce ACC for malonyl-CoA production. The production pathway is under an inducible promoter, allowing temporal control (reference 45).

Figure 3. Common mechanisms of riboswitches (adapted from reference 59). (a) Regulation at the translational level. Binding ligand results in changes in mRNA secondary structure and renders the ribosome binding site (RBS) inaccessible to ribosome. (b) Regulation at the transcriptional level. In the

absence of ligand, the intrinsic terminator stops transcription, while binding ligand eliminates the terminator structure. (c) The mechanism of an aptazyme. Binding ligand activates the hammerhead ribozyme and triggers self-cleavage, which subsequently leads to mRNA degradation.

References

1. Sun X, Shen X, Jain R, Lin Y, Wang J, Sun J, Wang J, Yan Y, Yuan Q: **Synthesis of chemicals by metabolic engineering of microbes.** *Chem. Soc. Rev.* 2015, **44**:3760–85.
2. Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, Moineau S, Mojica FJM, Wolf YI, Yakunin AF, et al.: **Evolution and classification of the CRISPR-Cas systems.** *Nat. Rev. Microbiol.* 2011, **9**:467–77.
3. Barrangou R: **Diversity of CRISPR-Cas immune systems and molecular machines.** *Genome Biol.* 2015, **16**:247.
4. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA: **Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.** *Cell* 2013, **152**:1173–83.
5. Luo ML, Mullis AS, Leenay RT, Beisel CL: **Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression.** *Nucleic Acids Res.* 2015, **43**:674–81.
6. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, et al.: **CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes.** *Cell* 2013, **154**:442–51.

7. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM: **CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering.** *Nat. Biotechnol.* 2013, **31**:833–8.
8. Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, La Russa M, Tsai JC, Weissman JS, Dueber JE, Qi LS, et al.: **Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds.** *Cell* 2014, **160**:339–350.
9. Cress BF, Toparlak ÖD, Guleria S, Lebovich M, Stieglitz JT, Englaender JA, Jones JA, Linhardt RJ, Koffas MAG: **CRISPathBrick: Modular Combinatorial Assembly of Type II-A CRISPR Arrays for dCas9-Mediated Multiplex Transcriptional Repression in *E. coli*.** *ACS Synth. Biol.* 2015, **4**:987–1000.
- * The authors developed a system for the design and modular cloning of CRISPRi constructs for simultaneous repression of multiple genes in *E. coli*. Multiple examples are presented, and considerations are made for repressing genes within operons.
10. Ranganathan S, Suthers PF, Maranas CD: **OptForce: an optimization procedure for identifying all genetic manipulations leading to targeted overproductions.** *PLoS Comput. Biol.* 2010, **6**:e1000744.
11. Wu J, Du G, Chen J, Zhou J: **Enhancing flavonoid production by systematically tuning the central metabolic pathways based on a CRISPR interference system in *Escherichia coli*.** *Sci. Rep.* 2015, **5**:13477.
- ** CRISPRi was used to achieve fine-tuning of central metabolic pathways to efficiently channel carbon flux towards malonyl-CoA, without impairing cell growth and allowing for significantly

improved production of a natural product. Emphasis is placed on the value of this approach for attenuating (and not eliminating) the expression levels of multiple essential genes.

12. Lv L, Ren Y-L, Chen J-C, Wu Q, Chen G-Q: **Application of CRISPRi for prokaryotic metabolic engineering involving multiple genes, a case study: Controllable P(3HB-co-4HB) biosynthesis.** *Metab. Eng.* 2015, **29**:160–8.
13. Cress BF, Jones JA, Kim DC, Leitz QD, Englaender JA, Collins SM, Linhardt RJ, Koffas MAG: **Rapid generation of CRISPR/dCas9-regulated, orthogonally repressible hybrid T7-lac promoters for modular, tuneable control of metabolic pathway fluxes in *Escherichia coli*.** *Nucleic Acids Res.* 2016, doi:10.1093/nar/gkw231.
14. Tong Y, Charusanti P, Zhang L, Weber T, Lee SY: **CRISPR-Cas9 Based Engineering of Actinomycetal Genomes.** *ACS Synth. Biol.* 2015, **4**:1020–9.
15. Thomason MK, Storz G: **Bacterial Antisense RNAs: How Many Are There, and What Are They Doing?** *Annu. Rev. Genet.* 2010, **44**:167–188.
16. Meng J, Kanzaki G, Meas D, Lam CK, Crummer H, Tain J, Xu HH: **A genome-wide inducible phenotypic screen identifies antisense RNA constructs silencing *Escherichia coli* essential genes.** *FEMS Microbiol. Lett.* 2012, **329**:45–53.
17. Nakashima N, Ohno S, Yoshikawa K, Shimizu H, Tamura T: **A vector library for silencing central carbon metabolism genes with antisense RNAs in *Escherichia coli*.** *Appl. Environ. Microbiol.* 2014, **80**:564–73.
18. Wu J, Yu O, Du G, Zhou J, Chen J: **Fine-tuning of the fatty acid pathway by synthetic**

- antisense RNA for enhanced (2S)-naringenin production from L-tyrosine in *Escherichia coli*.** *Appl. Environ. Microbiol.* 2014, **80**:7283–7292.
19. Yang Y, Lin Y, Li L, Linhardt RJ, Yan Y: **Regulating malonyl-CoA metabolism via synthetic antisense RNAs for enhanced biosynthesis of natural products.** *Metab. Eng.* 2015, **29**:217–26.
- * This study demonstrates that when downregulating multiple essential genes with asRNA, repression strengths can be tuned by changing asRNA sequence/structure to address the issue of cell growth impairment.
20. Scalcinati G, Knuf C, Partow S, Chen Y, Maury J, Schalk M, Daviet L, Nielsen J, Siewers V: **Dynamic control of gene expression in *Saccharomyces cerevisiae* engineered for the production of plant sesquiterpene α -santalene in a fed-batch mode.** *Metab. Eng.* 2012, **14**:91–103.
21. Yuan J, Ching C-B: **Dynamic control of ERG9 expression for improved amorpho-4,11-diene production in *Saccharomyces cerevisiae*.** *Microb. Cell Fact.* 2015, **14**:38.
22. Ozcan S, Johnston M: **Three different regulatory mechanisms enable yeast hexose transporter (HXT) genes to be induced by different levels of glucose.** *Mol. Cell. Biol.* 1995, **15**:1564–72.
23. Aiba H: **Mechanism of RNA silencing by Hfq-binding small RNAs.** *Curr. Opin. Microbiol.* 2007, **10**:134–9.
24. Yoo SM, Na D, Lee SY: **Design and use of synthetic regulatory small RNAs to control gene expression in *Escherichia coli*.** *Nat. Protoc.* 2013, **8**:1694–707.
25. Na D, Yoo SM, Chung H, Park H, Park JH, Lee SY: **Metabolic engineering of *Escherichia coli***

using synthetic small regulatory RNAs. *Nat. Biotechnol.* 2013, **31**:170–4.

- ** The authors constructed 122 synthetic sRNAs designed to suppress 130 genes involved in the cadaverine metabolic network. This revealed gene targets for increasing cadaverine production, including two essential genes that could not have been identified by screening gene deletions.
26. Chae TU, Kim WJ, Choi S, Park SJ, Lee SY: **Metabolic engineering of *Escherichia coli* for the production of 1,3-diaminopropane, a three carbon diamine.** *Sci. Rep.* 2015, **5**:13040–13053.
27. Meng H-L, Xiong Z-Q, Song S-J, Wang J, Wang Y: **Construction of polyketide overproducing *E. coli* strains via synthetic antisense RNAs based on *in silico* fluxome analysis and comparative transcriptome analysis.** *Biotechnol. J.* 2015, doi:10.1002/biot.201500351.
28. Liu Y, Zhu Y, Li J, Shin H, Chen RR, Du G, Liu L, Chen J: **Modular pathway engineering of *Bacillus subtilis* for improved N-acetylglucosamine production.** *Metab. Eng.* 2014, **23**:42–52.
29. Gredell JA, Frei CS, Cirino PC: **Protein and RNA engineering to customize microbial molecular reporting.** *Biotechnol. J.* 2012, **7**:477–99.
30. Mahr R, Frunzke J: **Transcription factor-based biosensors in biotechnology: current state and future prospects.** *Appl. Microbiol. Biotechnol.* 2016, **100**:79–90.
31. Chou HH, Keasling JD: **Programming adaptive control to evolve increased metabolite production.** *Nat. Commun.* 2013, **4**:2595.
- ** A powerful method for continuous evolution of metabolite biosynthesis is presented. Evolution of increased producers is monitored by the expression of a reporter gene in a real-time fashion. Notably, several synthetic TFs were readily designed and used as sensors in bacteria and yeast.

32. Schaaper RM, Radman M: **The extreme mutator effect of *Escherichia coli* mutD5 results from saturation of mismatch repair by excessive DNA replication errors.** *EMBO J.* 1989, **8**:3511–6.
33. Raman S, Rogers JK, Taylor ND, Church GM: **Evolution-guided optimization of biosynthetic pathways.** *Proc. Natl. Acad. Sci.* 2014, **111**:201409523.
- ** The authors characterize in detail the biosensor capacity of multiple TFs and riboswitches, and describe broadly applicable methods to couple metabolite concentrations to cell fitness. Iterative genome engineering and toggling between positive and negative selection provides a mechanism for the evolution of higher producing strains.
34. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM: **Programming cells by multiplex genome engineering and accelerated evolution.** *Nature* 2009, **460**:894–8.
35. Siedler S, Stahlhut SG, Malla S, Maury J, Neves AR: **Novel biosensors based on flavonoid-responsive transcriptional regulators introduced into *Escherichia coli*.** *Metab. Eng.* 2014, **21**:2–8.
36. Dietrich JA, Shis DL, Alikhani A, Keasling JD: **Transcription factor-based screens and synthetic selections for microbial small-molecule biosynthesis.** *ACS Synth. Biol.* 2013, **2**:47–58.
37. Binder S, Schendzielorz G, Stäbler N, Krumbach K, Hoffmann K, Bott M, Eggeling L: **A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level.** *Genome Biol.* 2012, **13**:R40.
38. Schendzielorz G, Dippong M, Grünberger A, Kohlheyer D, Yoshida A, Binder S, Nishiyama C, Nishiyama M, Bott M, Eggeling L: **Taking control over control: use of product sensing in**

- single cells to remove flux control at key enzymes in biosynthesis pathways.** *ACS Synth. Biol.* 2014, **3**:21–9.
- * Production of many metabolites may be limited by end-product feedback inhibition. Here, metabolite-responsive TF-based high-throughput screening was used to engineer enzyme variants that override this control mechanism to maximize production.
39. Mustafi N, Grünberger A, Kohlheyer D, Bott M, Frunzke J: **The development and application of a single-cell biosensor for the detection of l-methionine and branched-chain amino acids.** *Metab. Eng.* 2012, **14**:449–457.
40. Mahr R, Gätgens C, Gätgens J, Polen T, Kalinowski J, Frunzke J: **Biosensor-driven adaptive laboratory evolution of l-valine production in *Corynebacterium glutamicum*.** *Metab. Eng.* 2015, **32**:184–194.
41. Dahl RH, Zhang F, Alonso-Gutierrez J, Baidoo E, Batth TS, Redding-Johanson AM, Petzold CJ, Mukhopadhyay A, Lee TS, Adams PD, et al.: **Engineering dynamic pathway regulation using stress-response promoters.** *Nat. Biotechnol.* 2013, **31**:1039–46.
- * A generally applicable approach is presented whereby transcriptome analysis is used to identify promoters that are up- and down-regulated by stress-inducing metabolic intermediates. The identified promoters can be used in dynamic regulations to optimize biosynthesis.
42. Schujman GE, Paoletti L, Grossman AD, de Mendoza D: **FapR, a bacterial transcription factor involved in global regulation of membrane lipid biosynthesis.** *Dev. Cell* 2003, **4**:663–72.
43. Schujman GE, Guerin M, Buschiazio A, Schaeffer F, Llarrull LI, Reh G, Vila AJ, Alzari PM, de

- Mendoza D: **Structural basis of lipid biosynthesis regulation in Gram-positive bacteria.** *EMBO J.* 2006, **25**:4074–83.
44. Xu P, Li L, Zhang F, Stephanopoulos G, Koffas M: **Improving fatty acids production by engineering dynamic pathway regulation and metabolic control.** *Proc. Natl. Acad. Sci. U. S. A.* 2014, **111**:11299–304.
- ** An example where a TF enables dynamic regulation of a biosynthetic pathway. The engineered circuit was able to dynamically regulate the expression of enzymes involved in the synthesis and consumption of malonyl-CoA, resulting in an oscillatory malonyl-CoA pattern and balanced metabolism between product formation and cell growth.
45. Liu D, Xiao Y, Evans BS, Zhang F: **Negative Feedback Regulation of Fatty Acid Production Based on a Malonyl-CoA Sensor–Actuator.** *ACS Synth. Biol.* 2015, **4**:132–140.
46. Stanton BC, Nielsen AAK, Tamsir A, Clancy K, Peterson T, Voigt CA: **Genomic mining of prokaryotic repressors for orthogonal logic gates.** *Nat. Chem. Biol.* 2014, **10**:99–105.
47. Uchiyama T, Abe T, Ikemura T, Watanabe K: **Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes.** *Nat. Biotechnol.* 2004, **23**:88–93.
48. Collins CH, Arnold FH, Leadbetter JR: **Directed evolution of *Vibrio fischeri* LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones.** *Mol. Microbiol.* 2004, **55**:712–723.
49. Collins CH, Leadbetter JR, Arnold FH: **Dual selection enhances the signaling specificity of a**

- variant of the quorum-sensing transcriptional activator LuxR. *Nat. Biotechnol.* 2006, **24**:708–12.
50. de Los Santos ELC, Meyerowitz JT, Mayo SL, Murray RM: **Engineering Transcriptional Regulator Effector Specificity Using Computational Design and *in vitro* Rapid Prototyping: Developing a Vanillin Sensor.** *ACS Synth. Biol.* 2015, doi:10.1021/acssynbio.5b00090.
51. Henssler E-M, Scholz O, Lochner S, Gmeiner P, Hillen W: **Structure-based design of Tet repressor to optimize a new inducer specificity.** *Biochemistry* 2004, **43**:9512–8.
52. Reed B, Blazeck J, Alper H: **Evolution of an alkane-inducible biosensor for increased responsiveness to short-chain alkanes.** *J. Biotechnol.* 2012, **158**:75–9.
53. Lönneborg R, Varga E, Brzezinski P: **Directed evolution of the transcriptional regulator DntR: isolation of mutants with improved DNT-response.** *PLoS One* 2012, **7**:e29994.
54. Taylor ND, Garruss AS, Moretti R, Chan S, Arbing M, Cascio D, Rogers JK, Isaacs FJ, Kosuri S, Baker D, et al.: **Engineering an allosteric transcription factor to respond to new ligands.** *Nat. Methods* 2016, **In press**.
55. Tang S-Y, Fazelinia H, Cirino PC: **AraC regulatory protein mutants with altered effector specificity.** *J. Am. Chem. Soc.* 2008, **130**:5267–71.
56. Tang S-Y, Cirino PC: **Design and application of a mevalonate-responsive regulatory protein.** *Angew. Chem. Int. Ed. Engl.* 2011, **50**:1084–6.
57. Tang S, Qian S, Akinterinwa O: **Screening for enhanced triacetic acid lactone production by recombinant *Escherichia coli* expressing a designed triacetic acid lactone reporter.** *J. Am.*

Chem. Soc. 2013, **135**:10099–10103.

- * A unique example in which the effector specificity of a TF (AraC) was customized for the purpose of developing a molecular reporter for high-throughput screening of an expressed enzyme library for improved production of a minimal polyketide.
- 58. Chen W, Zhang S, Jiang P, Yao J, He Y, Chen L, Gui X, Dong Z, Tang S-Y: **Design of an ectoine-responsive AraC mutant and its application in metabolic engineering of ectoine biosynthesis.** *Metab. Eng.* 2015, **30**:149–55.
- 59. Groher F, Suess B: **Synthetic riboswitches - A tool comes of age.** *Biochim. Biophys. Acta* 2014, **1839**:964–973.
- 60. Berens C, Groher F, Suess B: **RNA aptamers as genetic control devices: The potential of riboswitches as synthetic elements for regulating gene expression.** *Biotechnol. J.* 2015, **10**:246–57.
- 61. Ogawa A: **Rational design of artificial ON-riboswitches.** *Methods Mol. Biol.* 2014, **1111**:165–81.
- 62. Schneider C, Suess B: **Identification of RNA aptamers with riboswitching properties.** *Methods* 2015, **97**:44–50.
- 63. Michener JK, Smolke CD: **High-throughput enzyme evolution in *Saccharomyces cerevisiae* using a synthetic RNA switch.** *Metab. Eng.* 2012, **14**:306–316.
- 64. Desai SK, Gallivan JP: **Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation.** *J. Am. Chem. Soc.* 2004, **126**:13247–54.

65. Eckdahl TT, Campbell AM, Heyer LJ, Poet JL, Blauch DN, Snyder NL, Atchley DT, Baker EJ, Brown M, Brunner EC, et al.: **Programmed evolution for optimization of orthogonal metabolic output in bacteria.** *PLoS One* 2015, **10**:1–27.
66. Lee SW, Oh MK: **A synthetic suicide riboswitch for the high-throughput screening of metabolite production in *Saccharomyces cerevisiae*.** *Metab. Eng.* 2015, **28**:143–150.
67. Yang J, Seo SW, Jang S, Shin S-I, Lim CH, Roh T-Y, Jung GY: **Synthetic RNA devices to expedite the evolution of metabolite-producing microbes.** *Nat. Commun.* 2013, **4**:1413.
- * This study demonstrates the use of riboswitches as metabolite sensors coupled to high-throughput screening/selection. In conjunction with combinatorial mutagenesis, enhanced producing strains were enriched to 75% of the total population after four rounds of growth competition cycles.
68. Wang J, Gao D, Yu X, Li W, Qi Q: **Evolution of a chimeric aspartate kinase for L-lysine production using a synthetic RNA device.** *Appl. Microbiol. Biotechnol.* 2015, **99**:8527–36.
69. Kato C, Kurihara T, Kobashi N, Yamane H, Nishiyama M: **Conversion of feedback regulation in aspartate kinase by domain exchange.** *Biochem. Biophys. Res. Commun.* 2004, **316**:802–8.

Figure 1

Figure 2

Figure 3

RNA-guided Repressors

Metabolite Sensors