The zero-sum game of pathway optimization: emerging paradigms for tuning gene expression

<table>
<thead>
<tr>
<th>Journal</th>
<th>Biotechnology Journal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>biot.201100086.R1</td>
</tr>
<tr>
<td>Wiley - Manuscript type</td>
<td>Review</td>
</tr>
<tr>
<td>Date Submitted by the Author</td>
<td>n/a</td>
</tr>
<tr>
<td>Complete List of Authors</td>
<td>Solomon, Kevin; Synthetic Biology Engineering Research Center (SynBERC); MIT, Dept. of Chemical Engineering Prather, Kristala; Synthetic Biology Engineering Research Center (SynBERC); MIT, Dept. of Chemical Engineering</td>
</tr>
<tr>
<td>Primary Keywords</td>
<td>Metabolic Engineering</td>
</tr>
<tr>
<td>Secondary Keywords</td>
<td>Synthetic Biology</td>
</tr>
<tr>
<td>Keywords</td>
<td>flux optimization, tuning of gene expression</td>
</tr>
</tbody>
</table>
The zero-sum game of pathway optimization: emerging paradigms for tuning gene expression

Kevin V. Solomon, Kristala L. J. Prather*

Department of Chemical Engineering
Synthetic Biology Engineering Research Center (SynBERC),
Massachusetts Institute of Technology
Cambridge, MA, 02139, USA

* Corresponding author:
Department of Chemical Engineering
77 Massachusetts Avenue
Room 66-454
Cambridge, MA 02139
Phone: 617.253.1950
Fax: 617.258.5042
Email: kljp@mit.edu

Keywords: flux optimization, metabolic engineering, synthetic biology, tuning of gene expression

Running title: Tuning expression for pathway optimization
Abstract

With increasing price volatility and growing awareness of the lack of sustainability of traditional chemical synthesis, microbial chemical production has been tapped as a promising renewable alternative for the generation of diverse, stereospecific compounds. Nonetheless, many attempts to generate them are not yet economically viable. Due to the zero sum nature of microbial resources, traditional strategies of pathway optimization are attaining minimal returns. This result is in part a consequence of the gross changes in host physiology resulting from such efforts and underscores the need for more precise and subtle forms of gene modulation. In this review, we describe alternative strategies and emerging paradigms to address this problem and highlight potential solutions from the emerging field of synthetic biology.
**Introduction**

Microbial production systems display a remarkable flexibility in the diversity and enantioselectivity of the compounds that they can generate. These compounds have historically been natural products such as ethanol, amino acids, acetone and antibiotics. However, with the introduction of ever more sophisticated tools, a range of natural and unnatural products have been made in engineered hosts including compounds such as hydroxyacids[1-3], isoprenoids[4, 5], polyketides[6, 7], and biopolymers[8, 9]. While several of these processes have been successfully commercialized [10-12], many remain economically infeasible and are the subject of intense optimization efforts.

In optimizing microbial pathways, the objectives are to maximize product flux, yield and selectivity. Traditionally, this problem has been approached by an analysis of the metabolic pathway that leads to removing branch points that lower product yield and selectivity (gene inactivation) and increasing the flux of intermediates through the pathway (gene overexpression). The power of such methods has improved tremendously with the advent of computational tools such as Flux Balance Analysis (FBA) [13, 14] and bilevel optimization [15-17] to identify flux bottlenecks, yet, they are still fundamentally constrained by the interconnectedness and finite nature of microbial resources (Figure 1).

Gene inactivations may necessitate media supplementation, impair cellular function and are sometimes infeasible for non-linear production pathways. Overexpression of pathway genes, on the other hand, comes at the expense of endogenous ones due to consumption of common precursors and titration of cellular machinery such as polymerases and ribosomes and may lead to growth inhibition, reduced expression and even cell death [18-20]. In certain hosts the heat shock response is stimulated by protein
overexpression [21, 22] further limiting the degree of overexpression possible. Moreover, successfully overexpressing or knocking out genes does not guarantee improved productivity. Decoupling the native regulation of flux within the pathway in these ways may lead to the accumulation of intermediates that can inhibit pathway enzymes [3, 23] or are bacteriostatic [1, 24, 25]. These challenges are not insurmountable, but they do underscore the need for more tools in pathway optimization. This review will highlight novel approaches to pathway optimization and describe emerging paradigms for flux manipulation.

**Downregulation of related pathways**

Modulation of gene expression, such as downregulation of undesired branch points, has been identified as a fruitful avenue for increased pathway productivity [16, 17]. In contrast to gene inactivation, downregulation offers the ability to redirect metabolite flux into production pathways while maintaining sufficient flux for endogenous processes. Moreover, in cases of drastic differences in catalytic efficiency of competing enzymes, it may prove more efficient than overexpression of pathway enzymes. Downregulation may be implemented in many different ways. One promising method, amenable to implementation in a wide variety of hosts and pathways, is the use of antisense RNA (asRNA) mediated inhibition of translation [26-30].

One such example of asRNA use in pathway optimization is found in the engineering of *Clostridium acetobutylicum*. Predating the rise of petrochemical sources, *C. acetobutylicum* was an industrially relevant source of solvents such as acetone and butanol [33] which it naturally ferments as part of its lifecycle [31, 32]. Recent volatility
in the price of chemical feedstocks and increasing concern regarding the sustainability of traditional chemical synthetic routes have led to renewed interest in the species with a focus on controlling the distribution of products [27, 29, 34, 35]. The Papoutsakis group used an asRNA approach to downregulate the CoA transferase which catalyzes the formation of acetone (ctfA1) to shift these strains to a primarily alcohologenic mode of production (ethanol and butanol), obtaining the highest ethanol titers reported at the time in C. acetobutylicum [29, 34]. Similar success has been reported for the engineering of glutamate synthesis from Corynebacterium glutamicum. C. glutamicum is a natural overproducer of amino acids and an industrial source of several of these including glutamate [36] which is produced from the transamination of α-ketoglutarate, a citric acid cycle intermediate. Utilizing an asRNA approach, Kim and coworkers [28] increased the cell specific productivity of glutamate by inhibiting activity of 2-oxoglutarate dehydrogenase thereby allowing sufficient flux of α-ketoglutarate through the citric acid cycle for energy production while diverting additional precursors to increase glutamate synthesis. Finally, asRNA technology has been utilized in the synthesis of cobalamin (Vitamin B₁₂) in Bacillus megaterium to improve titers and yields by 20% [30].

The use of downregulation extends beyond the realm of small molecule synthesis and has similar applications in recombinant protein production where acetate has been established to have an inhibitory effect on specific protein expression and bacterial growth [37-40]. Controlling acetate production by inactivation of phosphotransacetylase (pta) or acetate kinase (ackA) genes in E. coli, which shunt excess acetyl-CoA to acetate, has a deleterious effect on the cellular redox state [40], carbon flux [41], and ultimately growth [41]. Diverse solutions such as process-based schemes [37] and metabolic
engineering of the host to shunt the excess acetyl-CoA to acetoin [39] have been
developed to address the issue. Nonetheless, these solutions are not scalable to all
methods of culture and inhibit ATP synthesis by acetate secretion. Thus, Kim and Cha
[42] chose an antisense based scheme to minimize detrimental physiological effects.
Through minor antisense inhibition of *ackA* and *pta*, Kim and Cha were able to reduce
acetate formation by more than 20% while simultaneously observing a 60% improvement
in the production of green fluorescent protein with negligible impact on cellular growth.

These examples of asRNA inhibition are not the only examples of pathway
downregulation. Alternative strategies such as those utilizing the effect of codon bias on
translational efficiency in *C. glutamicum* [43, 44], repressible promoters in *S. cerevisiae*
[45-47] and titrating inducible promoters in *E. coli* [48] have been used with great
success to increase product yields and/or titers. Moreover, the last decade has seen
intense efforts to regulate genes at the transcriptional and post translational levels
culminating in several novel methods such as regulated suppression of amber mutations
[49], inducible protein degradation [50], engineered allostery [51] and riboregulators [52, 53]. Despite the fact that many of these emerging technologies have yet to mature and
attain widespread adoption, particularly in an industrial context, the growing interest in
asRNA points to its relative ease of implementation. While unexplored in these studies,
another advantage of downregulation is the possibility of dynamic control of gene
expression.

**Dynamic Expression Profiles**
When maximizing product titers and yields for industrial scale fermentation, carbon flux is shifted from the normal balance of metabolic intermediates and shunted into the desired product. This shift is frequently at odds with the goals of the cell, i.e. maintaining metabolic flux levels and maximizing biomass. Thus, genetic alterations that alter metabolic flux will incur a redistribution of metabolites to compensate for the change with some inhibition of growth. Gadkar et al. [54] studied this issue in silico as it applied to glycerol and ethanol production. In their work, they pursued a bilevel optimization strategy analogous to that of OptKnock [15] where product titers are maximized subject to growth maximization and other physical constraints to determine gene candidates for upregulation or deletion. However, unlike OptKnock, they also optimized the timing of these genetic changes. For glycerol production, simulations of a biphasic approach to gene expression resulted in a 30% improvement in titers over a static strategy. Similarly, ethanol titers were improved by 40% over a static strategy and 90% over wildtype behavior. These cases and more were further studied by Anesiadis and coworkers [55] with the simulated behavior of genetic elements from synthetic biology, as opposed to instantaneous switching in expression, and came to a similar conclusion: dynamic control of gene expression may be implemented to increase pathway productivity.

One of the first experimental demonstrations of this paradigm was elegantly performed in 2000. In trying to produce lycopene in *E. coli*, Farmer and Liao [56] sought to overexpress 2 key rate limiting enzymes: phosphoenolpyruvate synthase (Pps), which controls the pool of a glycolytic intermediate needed for lycopene biosynthesis, and isopentenyl diphosphate isomerase (Idi), which pulls glycolytic intermediates into the
lycopene biosynthetic pathway. However, overexpressing them statically from a tac promoter hindered growth, yields and titers. Thus, they engineered a gene circuit/metabolite control system in which expression of pps and idi was directly tied to the availability of acetyl phosphate, a proxy for glycolytic flux and cellular health. Using this approach, they were able to overexpress these enzymes to higher levels than that seen using a static approach while maintaining cellular viability and ultimately improve titers by 50%, productivity three-fold and carbon yields by more than an order of magnitude. Similar control systems have also been developed to drive protein expression through the use of quorum sensing in *E. coli* [57, 58]. Such systems allow for coordinated delayed induction across multiple cellular populations in addition to transmitting the metabolic load state of the host [59] thereby mitigating potential challenges associated with protein overexpression. Moreover, they are modular and readily amenable to integration in complex circuits [57] where Boolean logic and sensor functions can be implemented for tight pathway regulation in combination with other strategies for cumulative effects. A hypothetical example of this is presented in Figure 2 where sensing and logic (AND) operations are used to drive expression of pathway genes and product only when high cell densities and carbon flux are achieved.

**Emerging paradigms**

With an eye towards the creation of sophisticated gene circuits and networks for both pathway regulation and biosynthesis, the emerging discipline of synthetic biology has established a paradigm of developing reusable modules or “parts” and “devices” to control gene expression [60-62]. Towards this end, libraries of sensors, control
elements, promoters [63-65], and ribosome binding sites (RBS) [66] among others have been developed. Many of these libraries are curated within the Registry of Standard Biological Parts (http://partsregistry.org) and are freely available to the community. Through these libraries of parts, network components may be individually selected, tuned and regulated to achieve the necessary phenotype.

The rise of part libraries has also been accompanied by the development of computer aided design (BioCAD) tools to facilitate the design of ever more complex circuits [67-71]. However, they are dependent on the availability of datasheets [72] or other experimental characterization to describe them which are typically context dependent and not readily generalizable to all scenarios. Moreover, the current lack of generic insulators for these parts results in feedback from downstream parts, or retroactivity [73], which can further perturb performance from expectation. Nonetheless, there has been some success with the engineering of systems from these libraries using both theoretical and experimental approaches. For example, using an equilibrium statistical thermodynamic model, Salis and coworkers [74] were able evaluate the effects of the 5’ UTR on translation culminating in the design of novel RBSs able to achieve expression levels spanning 5 orders of magnitude. Their software tool, RBSCalculator (https://salis.psu.edu/software/), also allows for relative expression tuning of a given sequence. Empirical and combinatorial approaches to the tuning of gene expression from library components have also proven successful in optimizing yields of lycopene and mevalonate production pathways [63, 75]. Finally, a combination of both theoretical and experimental characterization has been used to design and develop tuned systems with little post hoc adjustment [76].
More recently, new part classes such as engineered enzyme complexes have been developed. As discussed previously, manipulating flux gives rise to a myriad of challenges such as the physiological consequences of flux imbalance and titration of cellular machinery. These undesired effects can be attenuated with gene modulation and gene circuits with some tradeoff in selectivity due to reduced pathway intermediates.

Inspired by natural solutions to minimize this tradeoff [77, 78], Deuber et al. [79] engineered an enzyme scaffold scheme to recruit multiple pathway enzymes in a single complex. By colocalizing enzymes in this way, diffusional limitations are effectively nullified and toxic metabolites can be maintained at locally high, but globally low, concentrations to maximize pathway flux with minimal disruption to endogenous processes. Furthermore, this scaffolding is scalable and generalizable to many enzymes and pathways in specified stoichiometries allowing for efficient spatial organization of pathway genes [79-81]. Using this system, mevalonate titers were improved by 70-fold when compared to scaffold free control. Moreover, with rate limiting enzymes being expressed at nominally low levels, the scaffolded constructs were able to achieve these yields without the growth inhibition seen in an unscaffolded design [79]. Alternative strategies such as direct protein fusions have also proved successful in improving pathway productivity [81].

**Conclusions and Perspectives**

Microbial production systems have enormous potential to synthesize many valuable chemical compounds in a sustainable manner. However, optimizing these systems for economic feasibility remains a challenge, in part, due to the zero sum nature
of intracellular metabolites (Figure 1). Redirecting these metabolites into pathways of interest necessitates a loss of flux elsewhere and titration of cellular machinery away from endogenous processes resulting in negative physiological consequences. These concerns may be attenuated to some degree by microbial consortia. Such mixed populations of cells are able to achieve more complex tasks, are more robust to environmental changes and are able to be organized by function [82]. More importantly, this functional specialization allows for the distribution of the metabolic burden across populations resulting in overall healthier cultures and potentially more efficient pathways. This advantage is offset, however, by the recalcitrant nature of genetic manipulations of all but a few species and the potential for competition between populations making a single organism, single population solution the most tractable solution for the immediate future.

Despite the limitations of finite cellular resources, the use of tools which precisely modulate expression levels has led to much improvement in pathway function by mitigating the effects of the pathway on host physiology. Synthetic biology has further contributed genetic parts and tools that allow for more precise application of regulation through mechanisms such as basic computation [66, 83, 84], sensing [56, 85] and timed expression [57, 58] with demonstrated improvements in productivity. Moreover, these modifications may all be combined for cumulative pathway improvement (Figure 2).

With increased understanding of the consequences of metabolic perturbations and evermore sophisticated regulation of expression, yields of microbial production systems may soon be economically competitive with traditional synthesis culminating in the realization of widespread microbial production.
Acknowledgments

This work was supported by the Synthetic Biology Engineering Research Center (SynBERC) funded by the National Science Foundation (Grant Number EEC-0540879), and a NSF CAREER Award (Grant Number CBET-0954986). Supplemental assistance was also provided by the Natural Sciences and Engineering Research Council (NSERC) of Canada through a fellowship to KVS. The authors would also like to thank Diana M. Bower for assistance with final proofing of the manuscript.

The authors have declared no conflict of interest.
References


[2] Lee, S.-H., Park, S., Lee, S., Hong, S., Biosynthesis of enantiopure (S)-3-hydroxybutyric acid in

[3] Tseng, H.-C., Harwell, C., Martin, C., Prather, K., Biosynthesis of chiral 3-hydroxyvalerate from single


produced in metabolically engineered Escherichia coli results in a strong fiber. *PNAS* 2010, 107, 14059-
14063.

[9] Wang, F., Lee, S. Y., Production of poly(3-hydroxybutyrate) by fed-batch culture of filamentation-


[27] Tummala, S. B., Welker, N. E., Papoutsakis, E. T., Design of antisense RNA constructs for
1923-1934.

[28] Kim, J., Hirasawa, T., Sato, Y., Nagahisa, K., et al., Effect of odhA overexpression and odhA
antisense RNA expression on Tween-40-triggered glutamate production by Corynebacterium glutamicum.

[29] Sillers, R., Al-Hinai, M. A., Papoutsakis, E. T., Aldehyde–alcohol dehydrogenase and/or thiolase
overexpression coupled with CoA transferase downregulation lead to higher alcohol titers and selectivity in


[31] Grupe, H., Gottschalk, G., Physiological events in Clostridium acetobutylicum during the shift from
acidogenesis to solventogenesis in continuous culture and presentation of a model for shift induction. *Appl


12-15.

[34] Tummala, S. B., Junne, S. G., Papoutsakis, E. T., Antisense RNA downregulation of coenzyme A
transferase combined with alcohol-aldehyde dehydrogenase overexpression leads to predominantly


6, 261-272.


Figure Captions

Figure 1
The zero sum challenges with traditional pathway optimization strategies. In the original pathway (A), only one media supplement (yellow circle) is needed to generate product (blue circle) and essential metabolites (aqua circle). However, gene inactivation (B) necessitates additional supplementation to generate the essential metabolite while overexpression (C) increases the pool of desired intermediate at the expense of expression and flux through the other enzymatic steps potentially limiting growth. The sum of these effects on the host’s health controls the degree of success on overall pathway production. Metabolite flux is proportional to the line thickness while metabolite pool size is represented by the circle area. Dashed lines indicate an absence of metabolite flux/pools when compared to wildtype.

Figure 2
A hypothetical example of a complex regulatory circuit utilizing multiple modules or parts. The LuxR/LuxI quorum sensing system (luxI not shown), mediated by N-acyl homoserine lactone (AHL), is used to drive the expression of glnAp2, an acetyl phosphate (ACP) sensor [86]. Sufficient carbon flux through central metabolism will lead to accumulation of ACP. The presence of GlnAp2 and ACP serve as inputs to an AND gate (binding of ACP to GlnAp2) whose output is expression of pathway genes and, ultimately, synthesis of product (triangles). Product is only produced when biomass and carbon flux is high, i.e. from a healthy culture.
Figure 1

A

B

C
Fig 1. The zero sum challenges with traditional pathway optimization strategies. In the original pathway (A), only one media supplement (yellow circle) is needed to generate product (blue circle) and essential metabolites (aqua circle). However, gene inactivation (B) necessitates additional supplementation to generate the essential metabolite while overexpression (C) increases the pool of desired intermediate at the expense of expression and flux through the other enzymatic steps potentially limiting growth. The sum of these effects on the host’s health controls the degree of success on overall pathway production. Metabolite flux is proportional to the line thickness while metabolite pool size is represented by the circle area. Dashed lines indicate an absence of metabolite flux/pools when compared to wildtype.

76x42mm (300 x 300 DPI)
Figure 2. A hypothetical example of a complex regulatory circuit utilizing multiple modules or parts. The LuxR/LuxI quorum sensing system (luxI not shown), mediated by N-acyl homoserine lactone (AHL), is used to drive the expression of glnAp2, an acetyl phosphate (ACP) sensor [86]. Sufficient carbon flux through central metabolism will lead to accumulation of ACP. The presence of GlnAp2 and ACP serve as inputs to an AND gate (binding of ACP to GlnAp2) whose output is expression of pathway genes and, ultimately, synthesis of product (triangles). Product is only produced when biomass and carbon flux is high, i.e. from a healthy culture.