ACS SyntheticBiology

Review

pubs.acs.org/synthbio

Synthetic Biology and Metabolic Engineering

Gregory Stephanopoulos*

Department of Chemical Engineering, Massachusetts Institute of Technology, Building 56 Room 469C, 77 Massachusetts Ave, Cambridge, Massachusetts 02139, United States

ABSTRACT: Metabolic engineering emerged 20 years ago as the discipline occupied with the directed modification of metabolic pathways for the microbial synthesis of various products. As such, it deals with the engineering (design, construction, *and* optimization) of native as well as non-natural routes of product synthesis, aided in this task by the availability of synthetic DNA, the core enabling technology of synthetic biology. The two fields, however, only partially overlap in their interest in pathway engineering. While fabrication of biobricks, synthetic cells, genetic circuits, and nonlinear cell dynamics, along with pathway engineering, have occupied researchers in the field of synthetic biology, the sum total of these areas does not constitute a coherent definition of synthetic biology with a distinct intellectual foundation and well-defined areas of application. This paper reviews the origins of the two fields and advances two distinct paradigms for each of them: that of unit operations for metabolic engineering and electronic circuits for synthetic biology. In this context, metabolic engineering is about engineering cell factories for the biological manufacturing of chemical and pharmaceutical products, whereas the main focus of synthetic biology is fundamental biological research facilitated by the use of synthetic DNA and genetic circuits.

KEYWORDS: *metabolic engineering, synthetic biology*

A couple of years ago I was invited to contribute a chapter to a book on synthetic biology. I happily accepted and asked for the table of contents at which point I discovered that, at least in the planning stage, the envisioned book overlapped at approximately 80% with the textbook on *Metabolic Engineering* that I had co-authored in 1998 (with Jens Nielsen and Aristos Aristidou). This was not the first time that synthetic biology was being *de facto* defined in an almost identical manner as metabolic engineering. It is indeed quite natural to define a field by *what it does*, as long as this is new and distinct from other fields and encompasses a well-defined body of rich intellectual content. However, while there is an explicit and universally accepted definition of metabolic engineering, to my knowledge, to date there has been no definition of synthetic biology. It is not my intention to define here synthetic biology. This paper is an attempt to describe the origins of the two disciplines, their intellectual foundations and areas of synergism, and challenges they face in view of the increasing scientific and public awareness of their existence and technological capabilities.

METABOLIC ENGINEERING

Metabolic engineering emerged at the beginning of the decade of the 1990s following a period of intense inquiry into the technological manifestations of genetic engineering and applied molecular biology. This culminated with two seminal papers[1,2] that essentially initiated the field of metabolic engineering. A conference followed soon thereafter (1996), along with a journal (*Metabolic Engineering*) and the first book in the field, *Metabolic Engineering: Principles and Methodologies* (Academic Press, 1998). Following a period of stimulating discussion the field was defined as the *directed modulation of metabolic pathways using methods of recombinant technology for the purpose of overproducing fuels and chemical and pharmaceutical products.*[1] Metabolic engineering, making extensive use of molecular methods of gene modulation, seemed very similar to genetic engineering, so a natural question that arose initially was how the two fields differed. After considerable deliberation, a distinct focus emerged for metabolic engineering, namely, investigation of the properties of *integrated* metabolic pathways and genetic regulatory networks, as opposed to individual genes and enzymes, which was the subject of most molecular biological research at that (pre-systems biology) time. In this sense, metabolic engineering preceded systems biology by championing the need for a *systemic view of metabolic pathways* and approaches for their optimal functioning.

In the following years it became very clear that metabolic engineering was a lot more than simply stitching genes together to build a basic functioning pathway. One can successfully express the totality of pathway genes to produce a few milligrams of product, but a cost-effective process cannot be realized until all three of titer, rate (or productivity), and yield (TRY) have been optimized. So, while a pathway can be built in a few months, it can take much longer to improve it to the point that it can support a commercial process. This brings into focus a basic asymmetry of the current publication and intellectual property (IP) ownership system that rewards the first publication of a pathway while ignoring the vast amount of effort and innovation that are required to improve such pathways so that they reach the TRY figures of merit required for a commercial operation. At the same time, it underlines the basic elements of metabolic engineering as the field aiming at pathway design, construction, and optimization. These elements encompass a lot more than genetic engineering and molecular biology and include components from graph theory,

Special Issue: Synthetic Biology for Strain Development

Received: September 23, 2012
Published: November 1, 2012

   dx.doi.org/10.1021/sb300094q | *ACS Synth. Biol.* 2012, 1, 514–525

Table 1

Intellectual Foundation of Metabolic Engineering
1. Enumeration and design of all pathways for converting a specified feedstock A to a target product Z. Methods and criteria for pathway ranking.
2. Thermodynamic analysis (feasibility) of promising candidate pathways.
3. Determination of pathway fluxes. Use of isotopic tracers for pathway validation and metabolic flux analysis.
4. Genome-scale models. Applications to identification of gene modulation targets and determination of optimal gene expression profiles.
5. Kinetic analysis of pathways. Distribution of kinetic control (MCA). Identification and elimination of kinetic bottlenecks.
6. Inverse metabolic engineering. Pathway optimization via rational and combinatorial methods.
7. Analysis of kinetics of synthetic genetic networks and gene circuits.

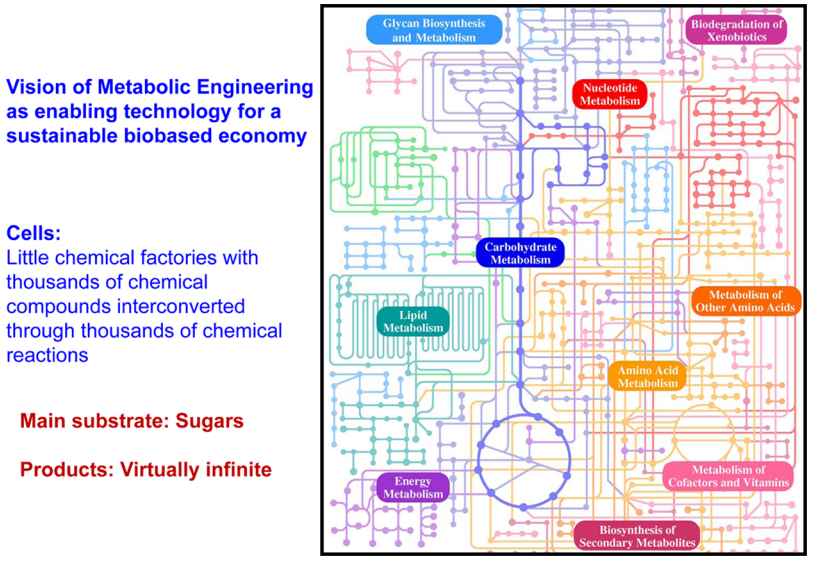

Figure 1. Vision of metabolic engineering as enabling technology of a sustainable biobased economy. Cells: Little chemical factories with thousands of chemical compounds interconverted through thousands of chemical reactions. The main substrate is sugar, and the number of products that can be made is virtually infinite.

chemical reaction engineering, biochemistry, and optimization, as summarized in Table 1.

It is important to recognize that, while cell and pathway performance was the ultimate goal of metabolic engineering, the *means* by which this goal was pursued also included synthetic genetic constructs (networks or circuits) that regulated the performance of the metabolic network. Such circuits constitute today a main activity of synthetic biology, which however had first been advanced in the context of metabolic pathway optimization and metabolic engineering.[3]

It should be noted that the listing of Table 1 is a rather telegraphic description of the topics of interest to metabolic engineering. Thus, enumeration of pathways connecting a designated substrate with a target product has occupied researchers for the past 30 years.[4−6] It requires complex concepts of combinatorics and graph theory. The number of possible pathways has exploded following sequencing of numerous genomes and increasing success in expressing heterologous pathways in various hosts: while before one was searching for all possible pathways within the genome of the host organism, nowadays the search has expanded to include the genomes of essentially all organisms. As such, methods for *ranking* candidate pathways are critical in reducing the pathways to be examined to a reasonable number. Thermodynamics can provide some criteria for pathway ranking[5] but other methods have been explored as well (discussed extensively in a recent review[7]).

Fluxes and flux determination is a central component of metabolic engineering along with the special requirements for generating reliable flux estimates, such as accurate tracer enrichment measurements and assessment of acceptable confidence intervals.[8−10] Fluxes are most informative when viewed as differences from a base state because it is then that they can be useful in identifying rate-controlling enzymes. Distribution of kinetic control, as initially pioneered by metabolic control analysis (MCA),[11,12] is indispensable in understanding the response of pathways following single and multiple enzyme modulation and can provide guidance into identifying target enzymes whose modulation will bring about the greatest change in the pathway flux. MCA concepts can have profound impact in designing and interpreting experiments of industrial as well as medical relevance (designing target therapies). In this regard, elementary mode analysis[13,14] constitutes a rational approach to identifying the basic elementary pathways carrying the overall flux toward a designated product, as well as the fraction of total flux carried by each one of the elementary pathways (or modes).[15] This information allows the systematic identification of specific enzymes that carry the majority of the product flux, which thus become targets for modulation.

To bypass the need for time- and resource-demanding experiments with stable isotope tracers required for the determination of the *actual* metabolic fluxes, other methods were developed for the determination of fluxes supporting *maximum growth.* As long as one does not forget this caveat, such maximum growth-supporting fluxes can be determined for numerous genomes using genome-scale models and linear optimization approaches applied on flux balance models.[16]

dx.doi.org/10.1021/sb300094q | *ACS Synth. Biol.* 2012, 1, 514−525

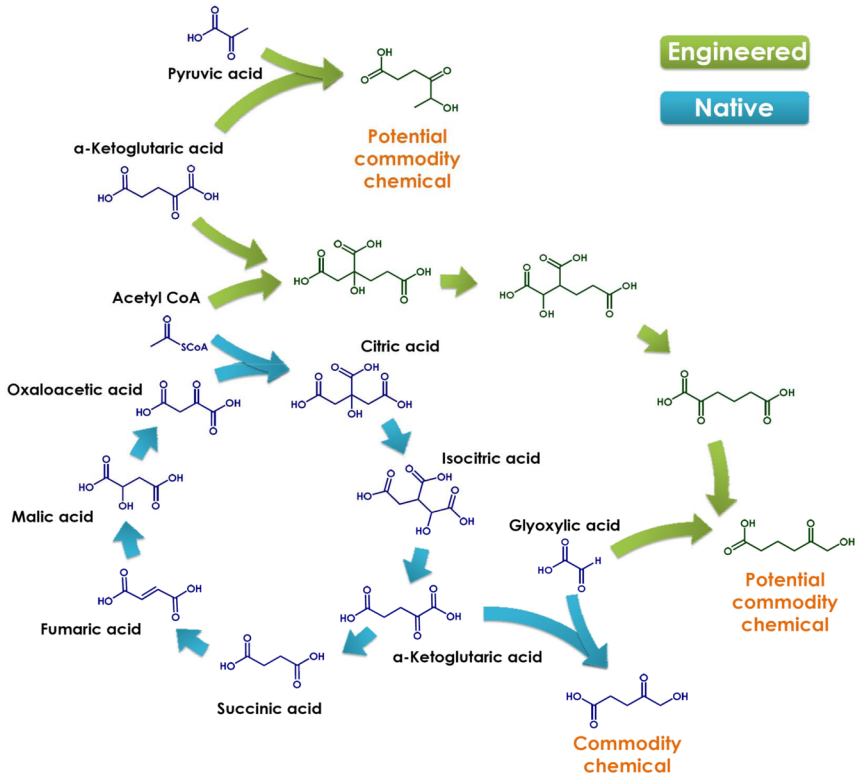

Figure 2. Schematic of native and non-natural pathways. Native pathways (blue) comprise only native reactions, whereas non-natural pathways import reactions from other organisms (green) through heterologous gene expression.

Optimal profiles of such fluxes can then be determined such as to optimize the rate or yield of product formation. Numerous algorithms have been published[17,18] identifying gene knock-out and overexpression targets using these calculated fluxes that support maximum growth. Although not always easy to implement experimentally, optimal flux profiles enhance the researcher's insights regarding pathway functioning and rational methods to improve pathway performance. In certain cases, production and growth are coupled so that growth optimization (and gene knock-out targets promoting maximum growth) also yield maximum product.[19] It should be also noted that, while metabolic engineering was initially based on rational pathway design methods, it soon expanded its portfolio to also include combinatorial methods in recognition of the inadequacy of available kinetic and regulatory models to support optimal pathway design on a global scale. Application of combinatorial methods on a rationally designed pathway usually improves performance by 2- to 4-fold over the rationally designed pathway. Of particular importance to combinatorial methods and associated approaches of inverse metabolic engineering[19] are high-throughout screens allowing selection of improved mutants and identification of the particular genetic element(s) responsible for enhanced performance.[20,21]

As can be seen from the above, metabolic engineering has evolved into a discipline of very rich intellectual content that goes far beyond the control of pathway gene expression. With respect to applications, after a period of small hesitant steps with pathways characterized by well understood kinetics and regulation, researchers are now emboldened to undertake the modulation of pathways for the production of high volume commodity chemicals besides the initial high-priced therapeutics and chemicals. Examples include biopolymers, fuels (ethanol, isobutanol, *n*-butanol, hydrocarbons, oils, and lipids), chemicals (succinic acid, butanediol, acrylic acid, lactic acid, isoprene), and numerous specialty chemicals. The driving force in these developments is concern about sustainability and the associated increasing interest in the production of products from renewable resources, namely, sugars derived directly from sugar cane or corn but also from cellulosic biomass at some point in the near future. Responding to the above market forces, technology advances, primarily through metabolic engineering, have provided the enabling technological platform required for realizing the above vision of a biobased economy.

Figure 1 depicts a schematic of this vision supported by technologies of metabolic engineering. The key concept here is that sugar (or other substrate) conversion to various products is most efficiently done by microorganisms. Bioprocesses that use engineered organisms for production are characterized by high product selectivity (relatively to chemical processes), which allows the design of smaller scale plants with lower capital cost and focused market scope. In many cases, bioproducts are cost-competitive with their chemical counterparts without considering any premium due to their renewable origin and environmental friendly nature.

A note on the application of metabolic engineering to mammalian cell cultures is in order. These systems were developed in the 1980s as technologies for the production of monoclonal antibodies and biopharmaceuticals. Culturing mammalian cells was the preferred technology due to the inability of bacterial and yeast systems to carry out the post-translational modifications (glycosylation) required for a functional protein product. Metabolic engineering methods of stoichiometric balancing and controlled nutrient feed[22] contributed to increasing monoclonal antibody titers from a few milligrams per liter to gram/liter levels. Also, genetic approaches to extend cell viability increased dramatically cell

dx.doi.org/10.1021/sb300094q | *ACS Synth. Biol.* 2012, 1, 514–525

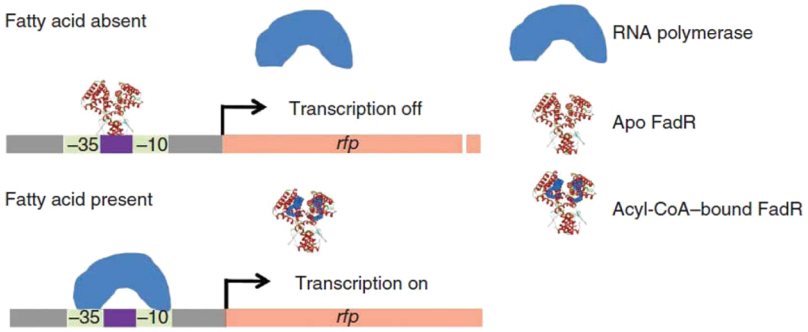

Figure 3. Schematic of sensor-actuator combination enabled by genetic circuits. The schematic depicts the design of a FA/acyl-CoA biosensor.[50] In the absence of fatty acid, FadR binds to the FadR-recognition site of the promoter, prevents the RNA polymerase from binding to the promoter, and represses the transcription of the gene *rfp*. When fatty acid is present, fatty acid is activated to acyl-CoA, which antagonizes the DNA binding activity of FadR. This allows RNA polymerase to bind to the promoter and turn on *rfp* transcription.

culture productivity.[23,24] More in line with the traditional concept of metabolic engineering, the glycosylation pathways were similarly engineered both in mammalian cells and recently in yeast, to generate cell systems with improved product glycosylation profiles, or entirely engineered yeast cells accomplishing glycosylation identical to that of mammalian cells, a most remarkable achievement of metabolic engineering with far reaching implications.[25,26]

Looking into the future we can expect metabolic engineering applications to increase dramatically. While one can think of various schemes to classify the envisioned pathway engineering applications (see, for example, the original classification by Cameron[27]), one classification is to divide them into two basic categories: those utilizing native pathways (for example all amino acid pathways) and those relying on non-natural pathways constructed with the use of heterologous genes (for example, pathways starting from amino acids and introducing reactions conducting new chemistry on the amino acid scaffold). Figure 2 depicts the two pathway types: products that can be synthesized by the native, blue, reactions belong to the first category, whereas products requiring the expression of heterologous, green, enzymes would belong to the second category. Such products can be non-natural, and examples include propanediol[28] and styrene,[29] among many others. Their synthesis is greatly facilitated by the availability of genes from numerous sources that are properly codon-optimized for optimal expression in the selected microbial host. It is this aspect of synthetic biology that has impacted the most the field of metabolic engineering.

SYNTHETIC BIOLOGY

The origins of synthetic biology are not well established. Some researchers would argue that the key initial event was the realization of high-throughput chemical synthesis of DNA. This was soon followed by the emergence of companies offering synthetic DNA as product. While some of these ventures aspired to a higher level of applications encompassing pathway design, construction, and expression in microorganisms, only those that focused on the cost-effective supply of synthetic DNA survived the market competition. Others will argue that the main events that initiated synthetic biology were the construction of the first genetic counter[30] and also the construction of the first genetic toggle switch in *Escherichia coli*,[31] which demonstrated that an artificial genetic element could initiate a different process or control an existing one inside a microbe. Yet, others will support the expression of a complete pathway with synthetic nucleic acid elements as the key event that started synthetic biology.

The above differing points of view are in sharp contrast to the origin of metabolic engineering and have profound implications regarding the definition of synthetic biology:

1. If the high-throughput chemical synthesis of DNA is the starting point of synthetic biology, then its nature is that of a new synthetic technology facilitating research by virtue of easy and inexpensive availability of a key reagent in pathway construction and modulation. Clearly, tremendous advances have been made in this field continuously expanding the length and type of DNA that can be chemically synthesized to the point that it has now reached the scale of small genomes.[32] This gives rise to the concept of *synthetic cells* and their impact in dealing with various pressing problems of energy and the environment, a topic we will revisit later.

2. If one accepts the construction and demonstration of a genetic counter or toggle switch as key defining event, then the focus shifts to a different, more biological, domain whereby easy availability of synthetic DNA sequences opens new doors for the investigation of fundamental biological questions. Here the counter/toggle switch were soon followed by numerous genetic control circuits, including various types of gates implementing different configurations of gene expression control. Such controls have now been applied to a large number of situations generally conforming to the following prototype: a genetic sensing element for the measurement of some metabolite or other analyte is combined with an actuator initiating control action when the level of the analyte exceeds a threshold. Figure 3 depicts the main idea in the context of a fatty acid biosensor, which is not all that novel from a control theory point of view or in light of prior work on the design and use of cross regulation systems for regulating cloned gene expression and protein production in bacteria.[33] Yet, it is very innovative when one considers that the whole device is coded by a synthetic piece of DNA allowing a cell to be viewed as a device that senses its environment and make a decision about its function, physiology, or mutation rate.[34] These genetic control elements can be applied to numerous situations and have led to ingenious devices for medical therapies.[35] Additionally, similar concepts and devices have been applied to the generation of interesting dynamics in metabolic and other systems.[36] It is expected that these concepts will allow the construction of innovative genetic control modules that will help elucidate basic biological mechanisms such as circadian rhythms, quorum sensing, and complex signaling pathways. Clearly, the

dx.doi.org/10.1021/sb300094q | *ACS Synth. Biol.* 2012, 1, 514–525

intellectual inquiries supported by these tools direct synthetic biology toward fundamental biological research as well as biomedical sensor-actuator types of applications.

3. Although there are not many cases that one would characterize *distinct* applications of synthetic biology to pathway design and construction, this area has received probably the most attention, perhaps because of its industrial relevance. From an intellectual point of view, however, this area has very little, if anything, to add to the body of work defining metabolic engineering. One can say that most all pathway examples of synthetic biology are better classified as examples of metabolic engineering as they comprise little more than chemical synthesis (by a vendor) of DNA and transformation of cells with the corresponding vectors.

BIOBRICKS AND SYNTHETIC CELLS

The ability to chemically synthesize at high throughput prescribed DNA sequences had two immediate consequences: First, it fueled the construction, collection and dissemination of individual genetic elements (referred to as biobricks) useful in the assembly of more complex genetic systems such as pathways, gates, and switches. Second, it gave rise to the vision of artificial chromosomes and totally synthetic cells as natural outcome of the synthesis of ever increasing strings of DNA.

The concept of modular construction of complex genetic systems expands on similar approaches taken by various vendors who have been offering vectors, adaptors, promoters, primers, etc. for several years. Yet, the concept of biobrick collection and associated foundation, student competitions, and other activities (IGEM) have been positive steps in promoting the image of biology and introducing unconventional methods in biological lab education. As such, they have met with remarkable success and become the best ambassadors of synthetic biology among undergraduate students and educators. Time will tell how well these structures meet the test of time, but overall, the concept of biobricks has been effective in popularizing molecular biology and expanding its image among students by including a critical hands-on element that makes it more applied and relevant to societal needs.

While the above contribution is undeniable, assertions that biobricks are the critical missing link in constructing metabolic pathways, cell compartments, or totally synthetic cells are unfounded and certainly overstated. These statements assert that, as a blueprint of cellular function is soon to become available, availability of biobricks will limit the realization of efficient metabolic pathways or artificial cells. I would argue that a blueprint of cellular function is nowhere in sight and our understanding of biology has been and will continue to be for a long time the limiting component in any attempt to reconstruct or emulate biological systems, in whole or in part. The undeniable progress of recent years in enhancing our understanding of basic biology by no means suggests that we are about to have a blueprint for designing artificial cells. One can sympathize with such statements amidst the excitement about the emergence of a new technology; however, they cannot be taken seriously in designing research programs or funding policies. Large-scale research programs aiming at "productizing" the construction of such biological components *in a vacuum* are misguided, especially when they are justified by the expectation that the availability of these components will accelerate by several orders of magnitude the design and optimization of metabolic pathways. These programs promote the idea that biobricks are the limiting component in constructing cell factories for commercial applications. I suggest that this view reveals little appreciation of metabolic engineering or experience with microbe engineering. Parts availability never was an issue in the construction of a successful pathway; biology most often has slowed progress in these endeavors.

Biobricks, such as modules controlling gene expression, can aid, in principle, the combinatorial optimization of a metabolic pathway by allowing the construction of numerous combinations of gene expression of various pathway modules (see the multivariable, modular pathway optimization recently published.[37] Combinatorial pathway optimization is critically dependent on equally high-throughput methods for *assessing* pathway function. Also, past experience suggests that, while constructing a number of pathway variants can be useful in optimizing a pathway, the requisite number that allows one to probe the relevant physiological space is rather small. This can change with ever increasing numbers of pathway modules and combinatorial controls; however, a proof of concept with a small system is needed before embarking on large-scale programs aiming at the synthesis of millions of potentially unnecessary biobricks. This calls for a critical reassessment of such programs, to ensure that the right priorities and context are set in their design and implementation.

In considering the possibility of chemically synthesizing a whole cell one should ask whether (a) this is possible, and (b) what would be the reason(s) for attempting such an undertaking in the first place. Considering the continuing advances in the length, fidelity, and speed by which chromosomal-size lengths of DNA can be chemically synthesized, one can envision the day when a totally artificial cell will be constructed. The recently announced creation of an entirely synthetic genome starting from a digitized sequence and followed by successful expression and creation of an organism with only synthetically generated DNA[38] supports this optimism. This will be a defining landmark for synthetic biology and momentous development in biological research. Such an achievement will connect for the first time the *design* of the cell with its *function*. Iterations on design principles will lead to improvements of certain cell properties and our understanding of overall cellular function and properties. Therefore, there is no denying that the main impetus for pursuing this goal should be advancement of fundamental biological research. Put differently, we are far more likely to understand the properties of a system, however complex, that we managed to design and construct compared to one about which there is still doubt as to the underlying mechanisms of its basic functions.

While I presume that there is little disagreement on the above, pursuing synthetic cells as better biocatalysts for solving major environmental and energy problems generates skepticism and a good deal of confusion regarding the motives of this research. First, there is absolutely no basis for the assertion that synthetic cells will have better properties for the cost-effective production of biofuels or fixation of carbon dioxide, to mention two technological areas mentioned in this context. The claims that synthetic cells will better satisfy the TRY figures of merit required for a commercial process (a) ignore improvements in cellular function brought about by evolution and (b) assume that we either now or in the near future will possess the blue print for not only designing a functional cell but also optimizing it to the level of robust performance required for commercial operation. In other words, it is not clear, nor have arguments been put forth to support this assertion, why a totally synthetic cell would be better than a well-engineered one using the

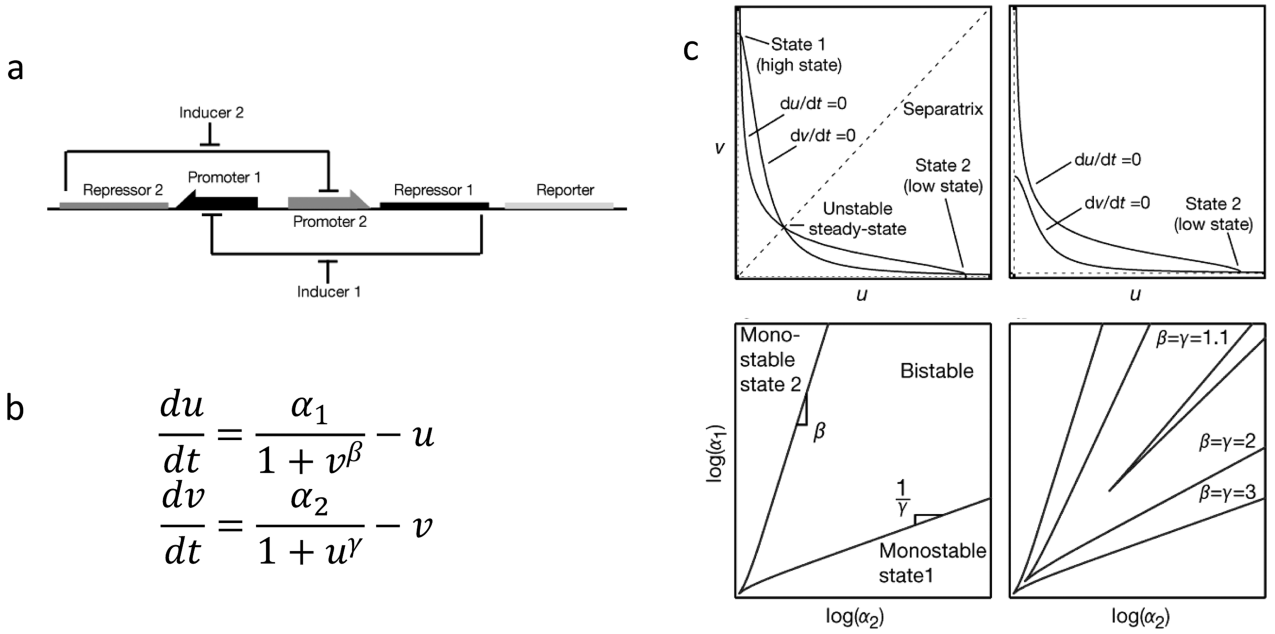

Figure 4. A bistable genetic "toggle switch" in *E. coli*, using a combination of two promoters and two repressors.[31] (a) Genetic construct. Repressor 1 inhibits transcription from Promoter 1 and is induced by Inducer 1. Repressor 2 inhibits transcription from Promoter 2 and is induced by Inducer 2. (b) Equations describing the system dynamics: u and v are the concentrations of the repressors 1 and 2, respectively; α represents the effective rate of synthesis of the repressor (which is a function of RNA polymerase biding, open-complex formation, and other genetic events); and β and γ represent the cooperativity of binding to the promoters. (c) Operating diagrams for the stability of the system as functions of the repressor rate of synthesis and cooperativity.

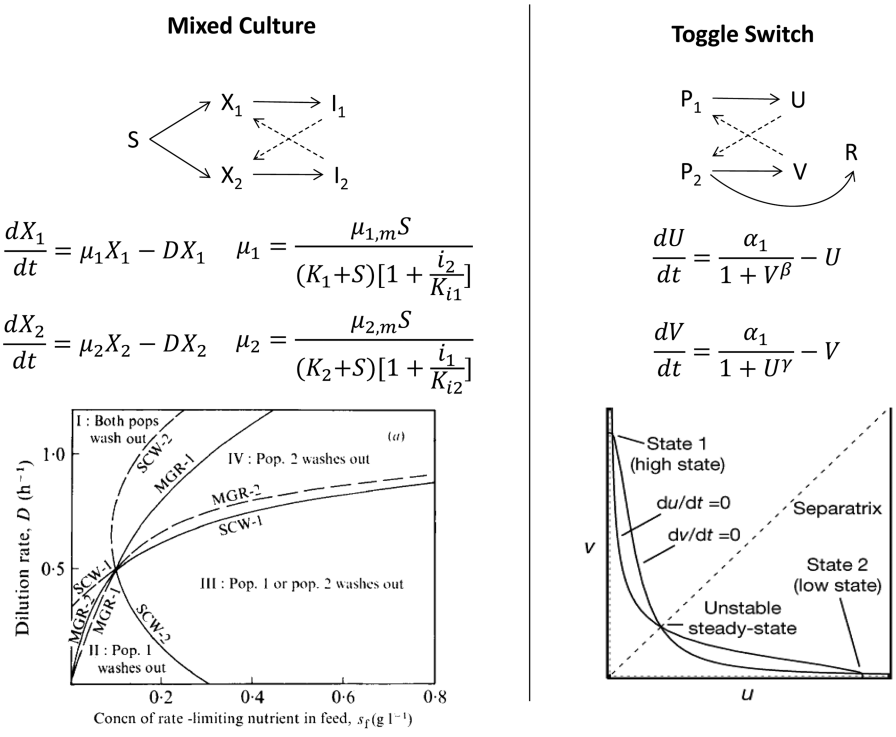

Figure 5. Chemostat dynamics for a mixed culture with inhibitors that do not act on the populations that produce them. The growth of each population (X_i) is expressed by the product of the specific growth rate (μ_i) and the cell density of the population, X_i. Specific growth rates are given by the indicated models, which are functions of the limiting substrate (S) and the inhibitor concentration (I). Substrate consumption for growth is assumed to be proportional to the rate of growth with the proportionality constant given by the growth yield (Y = grams cells formed per gram of substrate consumed). The analogy with the interaction of the toggle switch should be noted. The operating diagram shown in the figure was obtained for the following values of the model parameters: μ_{m1} = 2.0 (h^{-1}), μ_{m2} = 1.0 (h^{-1}), K_1 = 0.3 (g/L), K_2 = 0.1 (g/L), $K_{i1}/(a_2Y_2)$ = 0.3 (g/L), $K_{i2}/(a_1Y_1)$ = 0.6 (g/L).

dx.doi.org/10.1021/sb300094q | *ACS Synth. Biol.* 2012, 1, 514–525

principles discussed earlier. All in all, one cannot seriously argue that synthetic cells will be the panacea for mankind's problems in any foreseeable future.

A final note of caution is in order regarding the use of the terms synthetic cells and synthetic biology in industrial applications for the production of consumer goods. These terms evoke strong reactions reminiscent of the early day GMO controversy. The announcement of the semiartificial cell rekindled these concerns despite the fact that there was actually nothing unusual with the methods used to construct these cells. Nevertheless, these nascent concerns can surface at any time and jeopardize otherwise perfectly fine GMO biocatalysts, so such terms should be used with discretion in describing commercial applications of synthetic biology.

MARVELING NONLINEAR SYSTEM DYNAMICS

Gene expression can be monitored by GFP fluorescence. Monitoring gene expression dynamics can reveal fundamental information about the mechanisms of transcription and translation and also produce fascinating fluorescence oscillations at the culture but especially at the single cell level.[39,40] Synthetic biology allows the design of special gene expression systems involving feedback interactions that give rise to nonlinear kinetic expressions. Such systems can exhibit rich dynamic behavior ranging from asymptotic stable steady states, to stable focal points, to stable oscillations. This dynamic behavior contains information about the kinetics of protein−protein and protein−nucleic acid interactions. There is a certain fascination with the dynamics of such constructs stemming from the fact that they originate from a prescribed string of nucleic acid sequence. However, neither the concepts underlying this dynamic behavior nor the methods of bifurcation analysis are new as they have been applied extensively long ago in the context of chemical reaction system dynamics.[41]

Consider, as example, the repressor−inducer system of Figure 4. The dynamics of this system are captured by the pair of differential equations for the concentrations of the two repressors u and v as shown in the figure. The first term in each equation describes the cooperative repression of constitutively transcribed promoter, and the second term $(-u)$ is the rate of its degradation. At certain promoter strengths (α), the system is bistable, meaning that it has two stable steady states and can be interconverted between the two by a perturbation that allows it to cross the intermediary separatrix passing through the unstable steady state. Here, the stable steady states are high and low expression of the reporter gene (in this system GFP). By contrast, a monostable system has only one accessible steady state, in this case either the low or high expression state.

Consider now a system of two mixed microbial cultures growing in a continuous flow system (chemostat). The dynamics of this system are described by two differential equations for the time rate of change of the concentrations of the two species in the chemostat, as shown in Figure 5.[42] Similar to the previous system, the first term in each equation is the rate of growth of the species, and the second term is the dilution due to the flow through the reactor. Here, the key parameter is the specific growth rate of each of the growing species, which is usually a function of the growth-limiting substrate. Specific growth rates, however, can also depend on metabolites produced by other organisms growing in the same culture. This dependence can be a positive (growth-enhancing) or a negative one, such as growth inhibition. The point is that as a result of such interactions, specific growth rates can be nonlinear functions of metabolite concentrations, giving rise to equally fascinating dynamic behavior of the mixed culture as a whole and the concentrations of the constituent populations. Interactions studied are numerous, including indirect interactions such as competition, inhibition, commensalism, ammensalism, and mutualism and direct interactions such as predator−prey relationships, as well as combinations among them. They all produced fascinating dynamic behavior, which was confirmed experimentally for many of these systems. Figure 5 depicts the dynamics of a mixed culture where the specific growth rates depend, in addition to the growth-limiting substrate S, also on the concentrations of inhibitors produced by the opposite population. For the parameters shown in the legend it can be seen that the simple system of Figure 5 can have a single growth steady state with either population 1 or population 2 growing (regions IV and II, respectively) or multiple steady states (region III) where population 1 or population 2 grows depending on the initial conditions of the system. This behavior is identical to the toggle switch system shown in Figure 4. It is noted that numerous configurations for the types of interactions between the two populations exist, yielding a very rich diversity of dynamic behavior, as summarized in the cited reference and many other similar papers on mixed culture dynamics.

Very similar behavior was also observed with chemically reacting systems operating in continuous flow reactors (open systems). The dynamics of these systems too is captured by differential equations of the general type described in the above two examples, namely, a dilution term and a generation term containing the rate of reaction that can be a nonlinear function of the concentrations of reactants and products. It is noted that the dimensionality of the system is determined by the number of independent reactions considered, not by the number of reactants and products participating in the reactions. Steady states (fixed points) are obtained by setting the derivatives equal to zero, just as it is done in the operating diagrams of Figures 4 and 5. It is noted that the nature of reaction rates determines the types of dynamics of the system. For linear kinetics, only simple dynamics are possible. However, for the case of nonlinear kinetics, rich dynamic behavior is expected and has been experimentally observed.[43−45] Numerous configurations of chemically reacting systems were studied in the 1960s and 1970s where earlier theories of nonlinear dynamics, methods of bifurcation analysis, and some new ideas were explored in their analysis; see ref 46 for a comprehensive review of the state of the art of oscillating chemical reactions in 1967. The rationale for studying these systems was the expectation of elucidating the kinetics of the underlying chemical reactions. This promise was only partially fulfilled, but fascinating dynamics, often involving colorful systems (such as the Belousov−Zhabotinsky reaction[47]), were observed in the course of the experimentation.

One should note the striking similarities between gene expression dynamics and population interactions presently studied in the domain of synthetic biology and the earlier research of chemical reaction and mixed culture systems. There are two points to consider in bringing out the analogies between these two classes of systems. First, there is extensive literature on this subject that has been largely ignored by synthetic biology researchers. Gene expression, like the large majority of cellular processes, is defined mostly by chemical reactions. As such, all such processes will exhibit similar dynamic behavior, largely determined by the kinetics of

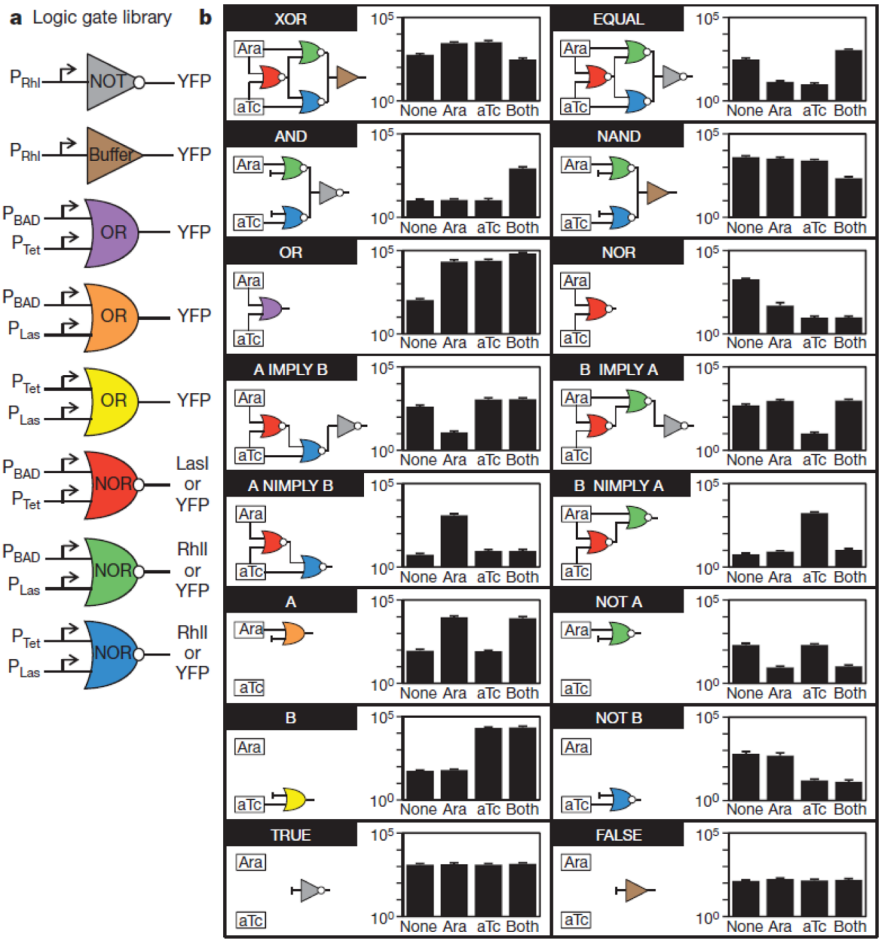

Figure 6. Metabolic network as collection of circuits.[53]

chemical interactions. This is clearly also observed in current research. Second, as mentioned, the main justification advanced for carrying out these studies then was the quest for otherwise unavailable information about reaction rates and chemical kinetics. This promise was only partially fulfilled. One of the reasons was that different reaction mechanisms can give similar dynamics and that a necessary step in identifying reaction mechanism is to first understand the type of bifurcation that a system undergoes as it enters the observed dynamic behavior. These lessons should serve as guidance in the design of similar research in synthetic circuits and synthetic biology today.

TWO PARADIGMS: UNIT OPERATIONS VERSUS ELECTRONIC CIRCUITS

The relationship between metabolic engineering and synthetic biology comes into sharp contrast when one considers the underlying fundamental paradigms of the two fields. These paradigms have not been externalized as such to date, so some justification of asserting them is in order. Thus, metabolic engineering is concerned with systems or networks of chemical reactions forming metabolic pathways in cells. The goal is to process continuously and in as seamless manner as possible material from one reaction (i.e., enzyme) to another until the final product is formed. Accumulation of metabolites suggests imbalances in the rates of various processing steps, which is undesirable as it slows down the overall system. As such, cells indeed behave as little chemical plants converting feedstocks (i.e, substrates) to products, while they maintain themselves by replenishing the key catalysts (i.e., enzymes). The main difference with a real chemical plant is the lack of any physical separation between the units carrying out the various steps and the lack of any need for separating reactants from products and recycling unreacted compounds to the main reactor. This is due to the exquisite specificity of enzymes that minimizes byproduct formation and yields processes with very high overall specificity.

The picture described above is the concept of *unit operations* of chemical plants. A chemical plant consists of units such as reactors, separation columns, and mixing and holding tanks. Auxiliary units also exist for energy, steam and water generation, and processing, reminiscent of currency metabolites for energy and redox transfer in cells. The concepts underlying modeling and optimization of such units and the overall plant are similar to those applied for the modeling and optimization of metabolic networks. A real plant also comprises control elements that maintain operation at a desirable steady state and take corrective action when deviations from this steady state are detected. These controls are implemented in a metabolic network at the enzyme level via natural allostery but also by engineering proteins to bring about more desirable properties of feedback regulation (combination of metabolic and protein engineering[48]). Another example is the adoption of methods and concepts from the design of heat exchange networks to the design and optimization of regulatory networks in metabolic pathways.[49] To date such methods have been satisfactory in maintaining metabolic networks at robust steady states. One can envision, however, situations where these more local,

dx.doi.org/10.1021/sb300094q | *ACS Synth. Biol.* 2012, 1, 514–525

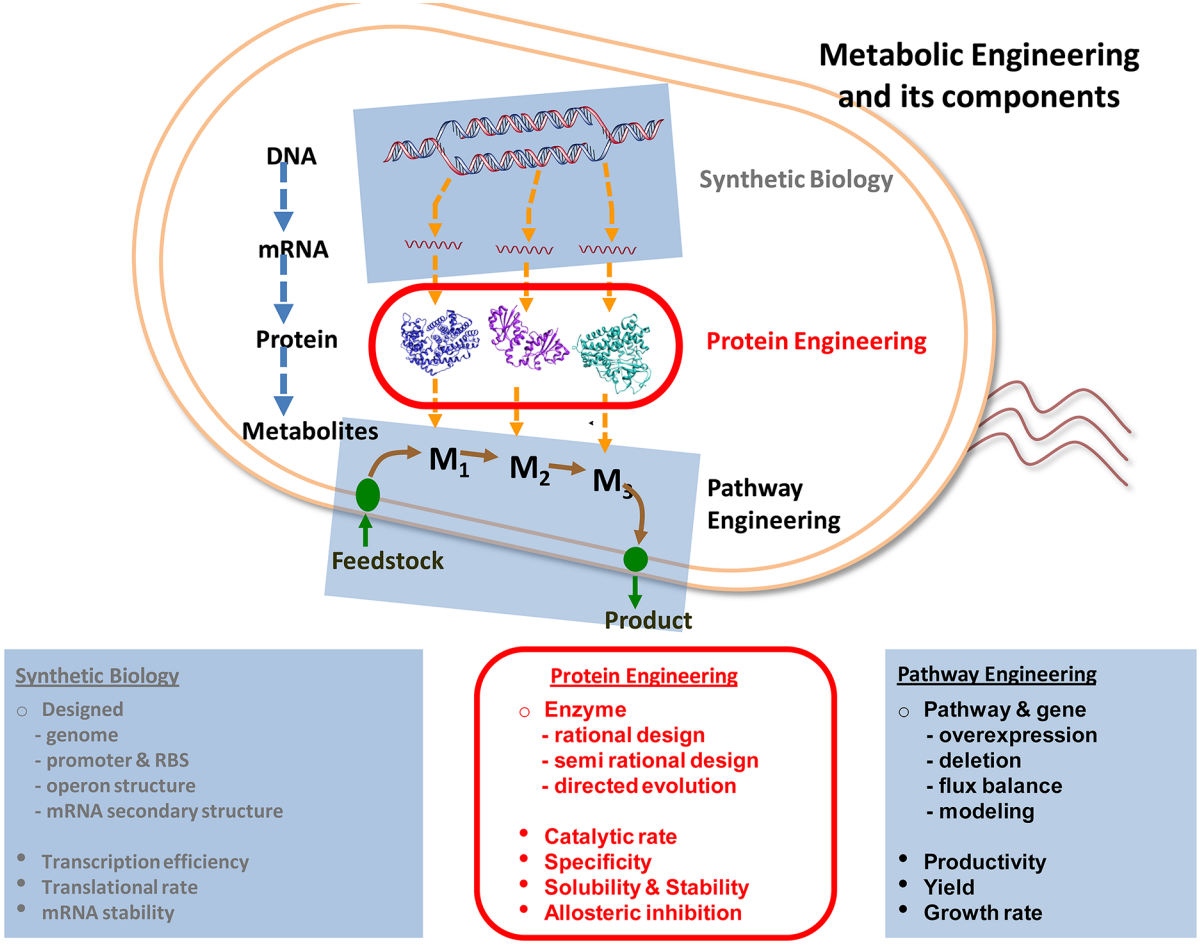

Figure 7. Metabolic engineering as a synthesis of synthetic biology and protein and pathway engineering (from P. Ajikumar and C. Pirie).

enzymatic controls will be inadequate due, mainly, to as yet not well understood distal network interactions. Here is where more sophisticated control structures advanced by synthetic biology can play an important role.[50]

The unit operations view of metabolic networks and cell factories is sharply contrasted by the electronic circuit view of metabolism embedded in the mind-frame of synthetic biology. The origin of this analogy is not clear and the justification even less obvious. Circuits are devices with on−off type of function transducing step inputs to step outputs and generally unrelated to chemical kinetics and reactions. They are designed for transmission and storage of information encoded in digital form. On the other hand, chemical reactions, which underlie any metabolic or cellular network, are distinctly analog in their function, which is manifestation of chemical reaction kinetics. Furthermore, it is not clear what purpose is served by the digital framework of metabolic network modeling. Flux distribution among various pathways is controlled by the regulation of enzymes at key junctions of the metabolic network, not activation of on−off switches in the network as is the case with digital electronic networks. Similarly, rates through the network are determined by the structure of the network and the relative activity of enzymes. If the goal is to analyze the properties of such networks and optimize their performance as measured by the production of a metabolite, this is done far more efficiently and reliably in a chemical unit operations framework rather than a digital one defined by gates, switches, and circuits.

A digital framework for the study of cellular function is better understood in the context of gene expression. Genes have been commonly assumed to be on or off, a view that is more in line with the circuit paradigm of synthetic biology. This approach can establish the connectivity among genes from global gene transcription data, a useful outcome overall. However, one should bear in mind that this connectivity will be (a) phenomenological, (b) hard to differentiate whether it is direct or indirect, and (c) devoid of any deeper mechanistic understanding that probes into the actual chemical reactions that mediate and regulate gene transcription.

Be that as it may, one cannot ignore the ubiquitous presence of circuits and the digital paradigm in publications of synthetic biology. Figure 6 is such an example. Here, the circuit representation was used to construct and describe a metabolic network consisting of combinations of promoters and repressors responsive to exogenous feeding of various compounds. Promoters and repressors were arranged in such a way that allows the cell to perform a logical operation (NOT, OR, and NOR), based on the presence of two compounds, and return in a digital response, in this case florescence. By using mutiple colonies performing different logical operations all 16 possible two-component logic gates were constructed (Figure 6). Although this work was done in spatially separated cells, it would be possible in the future to implement increasingly complex functions, and therefore behaviors, into individual cells, allowing for very specific responses to external and internal stimuli. These are useful concepts but not entirely

dx.doi.org/10.1021/sb300094q | ACS Synth. Biol. 2012, 1, 514−525

novel as very similar concepts and representations were used in prior work of biochemical systems theory[51] and systems biology.[40]

Molecular switches of the type contemplated by current research in synthetic biology will have important applications in metabolic engineering. A recent example is the use of a dynamic sensor-regulator system (DSRS) to improve biofuel production in *E. coli*.[50] As shown in Figure 3, the production rate of key pathway enzymes, which is used to control pathway flux, is modulated by the availability of key precursors and intermediates. This leads to a cellular production system capable of balancing cellular metabolism to reduce cell stress and increase overall yield.

On the other hand, the circuit-centric view of metabolism is not a particularly useful approach to understanding such complex systems and improving them for commercial applications. The same holds for understanding molecular mechanisms of gene transcription and translation and other important biological processes. Elucidation of these processes at a fundamental level will require the study of the chemical reactions that mediate these mechanisms. To this end, a circuit representation will have little to offer.

EPILOGUE

There are many synergies between metabolic engineering and synthetic biology, and the two fields need one another. Metabolic engineering is about designing, engineering, and optimizing pathways for the production of a variety of products, such as fuels, materials, and chemicals, including specialty, pharmaceutical, and commodity. To this end, it has established rich intellectual content and an effective portfolio of tools and methodologies. GM organisms of future applications will be constructed using these tools, as well as synthetic DNA provided by synthetic biology for building non-natural pathways for the production of current and novel products. Besides synthetic DNA, synthetic biology can also contribute advanced molecular switches for controlling the state of the metabolism in robust microbes suitable for commercial processes. Protein engineering can improve enzyme activity and specificity, and finally, pathway engineering can balance cofactors and currency metabolites, improve yields, and direct product synthesis in the most effective way. The above components, synthetic biology, protein and pathway engineering, are integrated, as indicated in Figure 7, in the overall scheme of metabolic engineering, which true to its integrating and systemic nature also is concerned with the overall physiological state and well being of the organism.

Synthetic biology, first and foremost, needs to define itself. This definition should include the intellectual foundations of the field, its tools, and its goals. Additionally, it should be *distinct* from other areas so that it does not replicate existing fields with a different name. A definition may be based on what the field *is* or what the field *does*. With respect to the latter synthetic biology *does* synthetic DNA, genetic switches for controlling metabolism, and molecular sensor-actuator combinations for controlling disease and other biomedical applications.[52,50] From an intellectual and more fundamental perspective, a good part of synthetic biology research has studied nonlinear dynamics of simple, prescribed biological processes. Besides the rich dynamic behavior of such systems, this work promises to unveil fundamental biological mechanisms. By and large, however, this work is not novel, and similarities to earlier work on chemical systems should be noted and properly acknowledged.

Synthetic biology has generated a lot of excitement among scientists and concerns among regulators. The sources of the excitement are (a) hands-on experiments among students making use of collections of synthetic genetic modules for the construction of diverse genetic elements and engineering microbes for various applications; (b) the vision of a totally synthetic cell, a noble goal for advancing basic biological research but one that cannot be justified by the prospect of developing biological processes to solve pressing problems in energy and the environment; and (c) the eagerness to advance programs for productizing the manufacturing of genetic modules, an initiative of uncertain utility that should be proven in small scale before undertaking large scale implementation. Regulators, on the other hand, are not comfortable with the prospect of reopening issues of recombinant cells that were assumed to be settled and are now re-emerging in connection with discussions about synthetic cells. A clear definition of synthetic biology will aid in settling this issue.

All in all, metabolic engineering is about engineering, whereas synthetic biology is about biology. As mentioned, the former will benefit from the tools of the latter in the synthesis and control of non-natural pathways. Synthetic biology too will benefit from the methods of metabolic engineering in the areas of pathway design, analysis, and optimization. The greatest benefit will be the adoption by synthetic biology of the chemi-centric, unit operations-based paradigm of metabolic engineering that recognizes chemistry as a fundamental science of most all biological processes.

AUTHOR INFORMATION

Corresponding Author

*E-mail: gregstep@mit.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

I wish to acknowledge the immense help of my students Ben Woolston and Steven Edgar for discussion and reference and figure preparation, as well as Professors Vassily Hatzimanikatis (EPFL), Jens Nielsen (Chalmers), Tillman Gerngross (Dartmouth), and Mike Jewett (Northwestern University) for critical reading of the manuscript and many useful suggestions.

REFERENCES

(1) Bailey, J. E. (1991) Toward a science of metabolic engineering. *Science 252*, 1668−1675.

(2) Stephanopoulos, G., and Vallino, J. (1991) Network rigidity and metabolic engineering in metabolite overproduction. *Science 252*, 1675−1681.

(3) Heinemann, M., and Panke, S. (2006) Synthetic biology–putting engineering into biology. *Bioinformatics 22*, 2790−2799.

(4) Hatzimanikatis, V., et al. (2005) Exploring the diversity of complex metabolic networks. *Bioinformatics 21*, 1603−1609.

(5) Mavrovouniotis, M., and Stephanopoulos, G. (1992) Synthesis of biochemical production routes. *Comput. Chem. Eng. 16*, 605−619.

(6) Yousofshahi, M., Lee, K., and Hassoun, S. (2011) Probabilistic pathway construction. *Metab. Eng. 13*, 435−444.

(7) Medema, M. H., van Raaphorst, R., Takano, E., and Breitling, R. (2012) Computational tools for the synthetic design of biochemical pathways. *Nat. Rev. Microbiol. 10*, 191−202.

(8) Stephanopoulos, G. (1999) Metabolic fluxes and metabolic engineering. *Metab. Eng. 1*, 1−11.

(9) Antoniewicz, M. R., Kelleher, J. K., and Stephanopoulos, G. (2007) Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. *Metab. Eng. 9*, 68−86.
(10) Antoniewicz, M. R., Kelleher, J. K., and Stephanopoulos, G. (2006) Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metab. Eng. 8*, 324−337.
(11) Burns, J., and Cornish-Bowden, A. (1985) Control analysis of metabolic systems. *Trends Biochem. Sci. 10*, 16.
(12) Fell, D. A. (1992) Metabolic control analysis: a survey of its theoretical and experimental development. *Biochem. J. 286*, 313−330.
(13) Schuster, S., Dandekar, T., and Fell, D. A. (1999) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends Biotechnol. 17*, 53−60.
(14) Trinh, C., Wlaschin, A., and Srienc, F. (2009) Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. *Appl. Microbiol. Biotechnol. 81*, 813−826.
(15) Stephanopoulos, G., Aristidou, A., and Nielsen, J. (1998) *Metabolic Engineering: Principles and Practices*, Academic Press Inc., San Diego.
(16) 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.
(17) 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.
(18) Feist, A. M., Henry, C. S., Reed, J. L., Krummenacker, M., Joyce, A. R., Karp, P. D., Broadbelt, L. J., Hatzimanikatis, V., and Palsson, B. O. (2007) A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol. 3*, 121.
(19) Alper, H., Jin, Y.-S., Moxley, J. F., and Stephanopoulos, G. (2005) Identifying gene targets for the metabolic engineering of lycopene biosynthesis in Escherichia coli. *Metab. Eng. 7*, 155−164.
(20) Santos, C. N. S., and Stephanopoulos, G. (2008) Melanin-based high-throughput screen for L-tyrosine production in Escherichia coli. *Appl. Environ. Microbiol. 74*, 1190−1197.
(21) Santos, C. (2012) Rational, combinatorial, and genomic approaches for engineering L-tyrosine production in Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A. 109*, 13538−13543.
(22) Xie, L., and Wang, D. (1996) High cell density and high monoclonal antibody production through medium design and rational control in a bioreactor. *Biotechnol. Bioeng. 51*, 725−729.
(23) Minas, W., and Bailey, J. E. (1995) Co-overexpression of prlF increases cell viability and enzyme yields in recombinant Escherichia coli expressing Bacillus stearothermophilus alpha-amylase. *Biotechnol. Prog. 11*, 403−411.
(24) Majors, B. S., Betenbaugh, M. J., Pederson, N. E., and Chiang, G. G. (2008) Enhancement of transient gene expression and culture viability using Chinese hamster ovary cells overexpressing Bcl-x(L). *Biotechnol. Bioeng. 101*, 567−578.
(25) Shelikoff, M. (1996) A modeling framework for the study of protein glycosylation. *Biotechnol. Bioeng. 50*, 73−90.
(26) Hamilton, S. R., Bobrowicz, P., Bobrowicz, B., Davidson, R. C., Li, H., Mitchell, T., Nett, J. H., Rausch, S., Stadheim, T. A., Wischnewski, H., Wildt, S., and Gengross, T. U. (2003) Production of complex human glycoproteins in yeast. *Science 301*, 1244−1246.
(27) Cameron, D. C. (1993) Cellular and metabolic engineering. *Appl. Biochem. Biotechnol. 38*, 105−140.
(28) Cameron, D. C., Altaras, N. E., Hoffman, M. L., and Shaw, A. J. (1998) Metabolic engineering of propanediol pathways. *Biotechnol. Prog. 14*, 116−125.
(29) McKenna, R., and Nielsen, D. R. (2011) Styrene biosynthesis from glucose by engineered E. coli. *Metab. Eng. 13*, 544−554.
(30) Elowitz, M. B., and Leibler, S. (2000) A synthetic oscillatory network of transcriptional regulators. *Nature 403*, 335−338.
(31) Gardner, T. S., Cantor, C. R., and Collins, J. J. (2000) Construction of a genetic toggle switch in Escherichia coli. *Nature 403*, 339−342.
(32) Smith, H. O., Iii, C. A. H., Pfannkoch, C., and Venter, J. C. (2003) Generating a synthetic genome by whole genome assembly. *Proc. Natl. Acad. Sci. U.S.A. 100*, 15440−15445.
(33) Chen, W., Kallio, P. T., and Bailey, J. E. (1995) Process characterization of a novel cross-regulation system for cloned protein production in Escherichia coli. *Biotechnol. Prog. 11*, 397−402.
(34) Good, M. C., Zalatan, J. G., and Lim, W. A. (2011) Scaffold proteins: hubs for controlling the flow of cellular information. *Science 332*, 680−686.
(35) Weber, W., and Fussenegger, M. (2009) The impact of synthetic biology on drug discovery. *Drug Discovery Today 14*, 956−963.
(36) Fung, E., Wong, W., Suen, J., Bulter, T., and Lee, S. (2005) A synthetic gene − metabolic oscillator. *Nature 435*, 118−122.
(37) Ajikumar, P. K., Xiao, W.-H., Tyo, K. E. J., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T. H., Pfeifer, B., and Stephanopoulos, G. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science 330*, 70−74.
(38) Gibson, D. G., Glass, J. I., Lartique, C., Noskov, V. N., Chuang, R. Y., Algire, M.A., Benders, G. A., Montaque, M.G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z. Q., Segall-Shapiro, T. H., Calvey, C. H., Parmar, P. P., Hutchison, C. A., 3rd, Smith, H. O., and Venter, J. C. (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science 329*, 52−56.
(39) Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S., and Elowitz, M. B. (2005) Gene regulation at the single-cell level. *Science 307*, 1962−1965.
(40) Geva-Zatorsky, N., Rosenfield, N., Itzkovitz, S., Milo, R., Sigal, A., Dekel, E., Yamitzky, T., Liron, Y., Polak, P., Lahav, G., and Alon, U. (2006) Oscillations and variability in the p53 system. *Mol. Syst. Biol. 2* (2006), 0033.
(41) Kevrekidis, I., Schmidt, L., and Aris, R. (1984) On the dynamics of periodically forced chemical reactors. *Chem. Eng. Commun. 30*, 323−330.
(42) De Freitas, M. J., and Fredrickson, a. G. (1978) Inhibition as a factor in the maintenance of the diversity of microbial ecosystems. *J. Gen. Microbiol. 106*, 307−320.
(43) McKarnin, M., Schmidt, L., and Aris, R. (1988) Response of nonlinear oscillators to forced oscillations: Three chemical reaction case studies. *Chem. Eng. Sci. 43*, 2833−2844.
(44) Middya, U., Graham, M. D., Luss, D., and Sheintuch, M. (1993) Pattern selection in controlled reaction−diffusion systems. *J. Chem. Phys. 98*, 2823.
(45) Gavalas, G. (1968) *Nonlinear Differential Equations of Chemically Reacting Systems*, Springer-Verlag, Berlin.
(46) Higgins, J. (1967) The theory of oscillating reactions. *Ind. Eng. Chem. 59*, 18−62.
(47) Murray, J. D. (1974) On a model for the temporal oscillations in the Belousov-Zhabotinsky reaction. *J. Chem. Phys. 61*, 3610.
(48) Leonard, E., Kumaran, P., Thayer, K., Xiao, W., and Mo, J. D. (2010) Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. *Proc. Natl. Acad. Sci. U.S.A. 107*, 13654−13659.
(49) Hatzimanikatis, V., Floudas, C. a, and Bailey, J. E. (1996) Optimization of regulatory architectures in metabolic reaction networks. *Biotechnol. Bioeng. 52*, 485−500.
(50) Zhang, F., Carothers, J. M., and Keasling, J. D. (2012) Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol. 30*, 354−359.
(51) Irvine, H. (1987) Biochemical systems theory and metabolic control theory: 1. Fundamental similarities and differences. *Math. Biosci. 86*, 127−145.
(52) Dong, H., Tao, W., Zhang, Y., and Li, Y. (2012) Development of an anhydrotetracycline-inducible gene expression system for solvent-producing Clostridium acetobutylicum: A useful tool for strain engineering. *Metab. Eng. 14*, 59−67.

(53) Tamsir, A., Tabor, J. J., and Voigt, C. A. (2011) Robust multicellular computing using genetically encoded NOR gates and chemical "wires". *Nature 469*, 212−215.