

# Genes & Development

## Designing biological systems

David A. Drubin, Jeffrey C. Way and Pamela A. Silver

*Genes & Dev.* 2007 21: 242-254

Access the most recent version at doi:[10.1101/gad.1507207](https://doi.org/10.1101/gad.1507207)

---

### References

This article cites 99 articles, 38 of which can be accessed free at:  
<http://www.genesdev.org/cgi/content/full/21/3/242#References>

Article cited in:

<http://www.genesdev.org/cgi/content/full/21/3/242#otherarticles>

### Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

---

### Topic collections

Articles on similar topics can be found in the following collections

[Systems Biology](#) (3 articles)

---

### Notes

---

To subscribe to *Genes and Development* go to:  
<http://www.genesdev.org/subscriptions/>

---

## REVIEW

# Designing biological systems

David A. Drubin,<sup>1</sup> Jeffrey C. Way,<sup>2</sup> and Pamela A. Silver<sup>1,3</sup>

<sup>1</sup>Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA; <sup>2</sup>EMD Lexigen Research Center, Billerica, Massachusetts 01821, USA

**The design of artificial biological systems and the understanding of their natural counterparts are key objectives of the emerging discipline of synthetic biology. Toward both ends, research in synthetic biology has primarily focused on the construction of simple devices, such as transcription-based oscillators and switches. Construction of such devices should provide us with insight on the design of natural systems, indicating whether our understanding is complete or whether there are still gaps in our knowledge. Construction of simple biological systems may also lay the groundwork for the construction of more complex systems that have practical utility. To realize its full potential, biological systems design borrows from the allied fields of protein design and metabolic engineering. In this review, we describe the scientific accomplishments in this field, as well as its forays into biological part standardization and education of future biological designers.**

Synthetic biology refers to a variety of experimental approaches that either seek to modify or mimic biological systems. This nascent discipline has been extensively described (e.g., Arkin 2001; Benner 2003; Brent 2004; Benner and Sismour 2005; Endy 2005; McDaniel and Weiss 2005; Sismour and Benner 2005; Sprinzak and Elowitz 2005; Andrianantoandro et al. 2006; Arkin and Fletcher 2006; Breithaupt 2006; Chin 2006a,b; Heinemann and Panke 2006; Tucker and Zilinskas 2006), often in a prospective manner, describing what synthetic biology ought to become and could accomplish. In this review, we describe what has been done thus far in synthetic biology and how the field is actively being moved forward, as well as survey various prospective views that articulate different visions of synthetic biology's future.

One goal of synthetic biology is the engineering of biological systems to fulfill a particular purpose. This purpose may be a specialized function or a means of gaining biological understanding. In reaching this goal, synthetic biologists pursue two distinct paths or a combination thereof: (1) the ectopic use of natural cellular components and systems, or (2) the generation of unnatural

chemical systems that mimic living systems, especially in regard to replication and Darwinian selection. This review will primarily focus on accomplishments of the first strategy; however, readers are referred to several fine reviews on the generation of artificial, life-emulating systems (Szostak et al. 2001; Benner 2003; Benner and Sismour 2005).

According to the synthetic biology paradigm, a synthetically programmed cell should be composed of many subsystems that operate successfully as a result of extensive characterization and educated design. System construction is the product of iterative cycles of computer modeling, biological assembly, and testing. In this way, much from the field of systems biology can be applied to synthetic biology, as it is the understanding of the functional system as a whole that is required to design an operable synthetic system. The systems/synthetic relationship is reciprocal, as what is learned from building working systems can be applied to better understanding natural systems and components.

As synthetic biology is an amalgamation of biology and engineering, in a discussion of a synthetic biology *raison d'être*, it is imperative to consider the perspectives of both disciplines. From an engineering point of view, biology is a tool or a system of parts that can be used to construct devices that accomplish a variety of tasks. The engineering influence seeks out the simplicity in biological systems, and brings standardization and modular design principles to biology.

From a biological point of view, synthetic biology is first and foremost a way of thoroughly testing concepts and systems. Synthesis complements observation and analysis (Benner and Sismour 2005), and allows us to both learn the dimensions of systems we may not have previously observed by analysis alone, and to test our models by reconstruction.

## Harnessing nature's toolbox

The birthing throes of synthetic biology echo from a combination of technologies and revelations that emerged in the 1960s and 1970s. These include the understanding of gene circuitry and the protein-based regulation of gene transcription as well as the tools provided by recombinant DNA technology (e.g., Jacob and Monod 1961; Szybalski and Skalka 1978). These foundational concepts and methods made synthetic biology possible,

[*Keywords:* Biological engineering, genetic engineering, metabolic engineering, protein design, synthetic biology]

<sup>3</sup>Corresponding author.

EMAIL [pamela\\_silver@hms.harvard.edu](mailto:pamela_silver@hms.harvard.edu); FAX (617) 432-5012.

Article is online at <http://www.genesdev.org/cgi/doi/10.1101/gad.1507207>.

in principle. In practice, recombinant DNA allowed the production of engineered proteins, but the engineering of whole organisms has had to wait for four developments: inexpensive DNA synthesis, rapid and inexpensive DNA sequencing, a large collection of components in the form of well-characterized genetic modules from years of study and from genome sequencing efforts, and a self-conscious importation of engineering approaches. The importance of DNA synthesis and sequencing in synthetic biology has been reviewed by others (Carlson 2003; Knight 2005), so we will focus on the availability of modular genetic parts. For synthetic biology, the modules of biological systems, from DNA to RNA, protein, and pathway levels, constitute a natural set of building blocks for engineering.

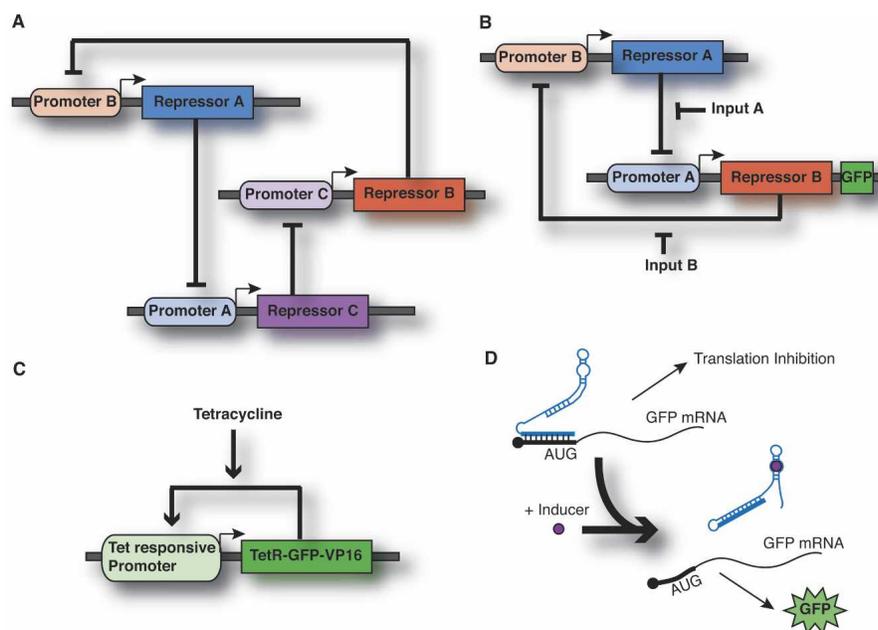
### Synthetic biology endeavors

By employing the simple modules provided by nature, many synthetic biologists aim to both build useful devices and to study artificial systems that may replicate features of natural systems. What follows is a survey of work within the synthetic biology community.

### Synthetic oscillators and switches

The first synthetic gene circuits were engineered in bacteria, and more recently have been established in yeast and mammalian systems (Hasty et al. 2002; Kaern et al. 2003; Sprinzak and Elowitz 2005). These early synthetic studies characterized design strategies inspired by both natural biological systems (Wall et al. 2003, 2004) and human-designed nonbiological systems. The studies illustrate the important role of mathematical modeling in designing gene circuits.

Oscillations of protein levels are vital to coordinating cellular events such as the cell cycle or circadian rhythms. In the case of the ring oscillator, or “repressilator,” developed by Elowitz and colleagues (Elowitz and Leibler 2000), a transcriptional control system was engineered to mimic the behavior of these oscillating systems. The circuit design involves a loop of three repressors (the lac, lambda, and tetracycline repressors), each repressing the transcription of the next in the cycle (Fig. 1A). An important part of the design process was the development of a rough quantitative model of the system in order to determine the vital parameters and parameter spaces in which the system is oscillatory. Gene



**Figure 1.** Synthetic oscillators and switches. (A) A circular network of transcriptional repressors, each repressing the expression of the next in the series, is the basis of the genetic ring oscillator, or “repressilator.” Oscillatory output was observed via GFP expression regulated by the tetracycline repressor, one of the repressors in the circuit (Elowitz and Leibler 2000). (B) A genetic toggle switch that uses a mutual repression motif to achieve bistability. It comprises two genes, each encoding a transcriptional repressor that represses expression of the other gene. Transient exposure to either input A or B will shift the stable steady state of the system to either expression of repressor B or repressor A, respectively. Output in the circuit shown here will be in GFP expression when the steady state favors repressor B, and no GFP when repressor A is expressed (Gardner et al. 2000). (C) A bistable positive feedback loop expressed in yeast in which a tetracycline-dependent activator turns on its own expression (Beckskei et al. 2001). (D) The RNA-based “antiswitch” relies on ligand-binding regions of RNA that, when bound to ligand, induce changes in RNA structure. In the antiswitch designs, the RNA is engineered such that when bound to an inducer ligand, the antiswitch either exposes or hides a region of RNA homologous to a region of a target mRNA in the 5' UTR, encompassing the translational start site. In this way, when the antisense sequence is exposed, translation of the target mRNA is repressed. Both an on and off antiswitch (the “on” version is shown here) have been designed that differ in the response to the presence of ligand (Bayer and Smolke 2005).

Drubin et al.

and protein modules, such as efficiently repressed promoters and protein degradation tags, were then assembled in a manner consistent with the model. Elowitz and Leibler (2000) describe only a single construct, suggesting that the approximate calculations used to design the repressilator were good enough to get a system that worked the first time (at some level). The output of the final repressilator design, while oscillatory as desired, showed much more variability in the oscillation period than do natural oscillators. Such noise in gene networks has implications in the function of natural systems and design of predictable synthetic gene networks (e.g., Becskei and Serrano 2000; Elowitz et al. 2002; Raser and O'Shea 2004, 2005; Hooshangi et al. 2005; Pedraza and van Oudenaarden 2005; Rosenfeld et al. 2005; Guido et al. 2006; Mettetal et al. 2006). One characteristic of natural oscillators such as the circadian rhythm and the heart beat is that they involve many more gene products than the repressilator, which is a rather minimal system (Chen et al. 1996; Anantharam et al. 2003; Shirasu et al. 2003). The heart, for example, is an oscillator that is not coupled to an external cycle, and its rhythm is thought to involve a very large number of proteins, especially when ion channel splice variants are considered. By classical molecular genetics it is very difficult to tease out the function of heart channels and other proteins in the limitation of period variability per se, as opposed to modulating changes in period in response to stress, maintaining function in the face of changes in blood ions and energy sources, and maintaining robustness in the face of heart damage. Synthetic biology offers an opportunity to analyze biological oscillation without the complications of dealing with a system that evolved in response to a variety of constraints.

Gardner et al. (2000) constructed a genetic toggle switch, which uses a network topology that confers bistability and memory to a system. There are generally two means of achieving bistability in a network: a positive feedback loop or mutual repression (i.e., double negative feedback) (Ferrell 2002; Hasty et al. 2002). In the case of Gardner's genetic toggle switch, a mutual repression system was employed using two genes that each code for a transcriptional repressor of the other gene (Fig. 1B). Various combinations of the lac repressor, tetracycline repressor, and the temperature-sensitive lambda cI857 repressor were used. The systems were designed such that an external stimulus inhibits the activity of one repressor over the other, and pushes the system to one stable state, where the one gene is repressed and the other fully transcribed. This system also demonstrates hysteresis, such that once the switch is flipped toward one steady state it remains there, even in the absence of the original stimulus, until a specific perturbation shifts the system to the other steady state. Like the repressilator, modeling of this pathway was crucial in determining the important parameters for robust bistability.

An important value of the toggle switch system is its utility as a pathway module to create networks with novel functions, and thus more complex programmable cells. This was demonstrated when Kobayashi et al.

(2004) used a *lacI*/lambda repressor toggle switch as a memory subsystem within a larger system that sensed DNA damage and recorded the result with this switch. This system used a toggle switch with a wild-type lambda repressor and lac repressor, and used the natural ability of this repressor to be cleaved upon DNA damage and induction of the SOS response. Kobayashi et al. (2004) then constructed a strain in which the *traA* gene, which activates biofilm formation by *Escherichia coli*, is repressed by lambda repressor, so that when the switch is flipped to the lac-on/lambda-off mode by DNA damage, biofilm formation results. Finally, Kobayashi et al. (2004) created a sensor circuit in which a sensor for acyl-homoserine lactone (AHL) activates a *lacI* gene, so that a *lacI*/lambda repressor toggle switch in the same cell can be flipped into the *lacI* mode. AHL is the quorum-sensing signal that is produced constitutively by *E. coli* to represent cell density. Cells with this artificial construct show density-dependent flipping of the toggle switch, which is registered through expression of a reporter. These results indicate that it is possible to construct artificial regulatory circuits that result in permanent phenotypic changes, and illustrate how individual modules may be combined to build progressively more complex circuits.

A discussion of switches inevitably recalls the naturally occurring genetic toggle switch found in phage lambda (Ptashne 2004). The components of this switch have been characterized in extensive, quantitative detail, which has allowed this system to become a staple for mathematical modeling (McAdams and Arkin 1997). Atsumi and Little (2004) constructed variant lambda phages in which the *cro* gene was replaced by *lacI*, and various alleles of lac operators were introduced at lambda P<sub>L</sub>, P<sub>RM</sub>, and P<sub>R</sub>. This design strategy takes advantage of the fact that *lacO* is normally located downstream from the transcription start site, so that this operator can be inserted without interfering with the function of the lambda operators. Atsumi and Little (2004) found a large number of phages that could grow lytically, some of which could lysogenize and a few of which had properties fairly similar to wild-type lambda.

Becskei et al. (2001) constructed a simple yeast-based switch in which the well-known tetracycline-dependent activator (Gossen and Bujard 1992; Gossen et al. 1995) turns on its own synthesis, an example of a positive feedback loop (Fig. 1C). In this system, cells initially in an "off" state could be flipped onto an "on" state in the presence of tetracycline. In response to a graded increase in tetracycline, a bimodal distribution of cells is created that are either in the on or off state. Thus, at the population level, the distribution of cells in the on and off states can be controlled by varying levels of inducer. Even in the presence of tetracycline, the off state is metastable, and will flip into the on state stochastically. This yeast-based positive feedback loop may be considered a noise-based switch (Hasty et al. 2000). The bistability resulting from this eukaryotic positive feedback loop is achieved via the cooperativity inherent in the tetracycline-dependent activator and eukaryotic transcription

itself (Polach and Widom 1996; Vashee et al. 1998; Wang et al. 1999).

Using the same general design concepts as described in the bacterial toggle switch, Kramer and colleagues (Kramer et al. 2004; Kramer and Fussenegger 2005) designed a mammalian (CHO cell)-based hysteretic toggle switch using streptogramin and macrolide-dependent transcription factors, each fused to a KRAB repression domain. The system switches between two steady states in response to the specific inducers. With the gene regulatory tools available in eukaryotic circuit designs, there is great potential for the generation and study of more complex systems, integrating such mechanisms as regulated nuclear transport.

Synthetic switches in both prokaryotes and eukaryotes can also be mediated via RNA devices, in the form of engineered riboregulators (Isaacs et al. 2004; Bayer and Smolke 2005), which has been recently reviewed in great detail (Isaacs et al. 2006). Briefly, RNA-based gene regulation is evident naturally in several different forms, and generally represses gene expression through either targeted mRNA degradation or inhibition of translation in *cis* and *trans* fashions. Engineered modular riboregulators can also activate gene expression (Isaacs et al. 2004). In a variation of these themes, riboswitches function by the binding of a particular metabolite to the 5' untranslated region (UTR) of a target mRNA that induces the RNA to fold in such a way as to block the ribosome-binding site and translation. Recently, Bayer and Smolke (2005) have used these ligand-binding regions from the 5' UTRs of riboswitch RNA and coupled them with an antisense module that targets a specific mRNA to prevent translation, in this case of a green fluorescent protein (GFP) reporter, in yeast. In this "antiswitch" design, the binding of the specific ligand to the binding region of the RNA affects the proper availability of the antisense region for its target transcript (Fig. 1D). This can either prevent antisense activity or activate it depending on the switch design. In one form, the presence of the ligand makes the antisense region of the antiswitch available for binding of the target mRNA, and thus represses expression. In another form, the ligand induces sequestration of the antisense region and activates expression of the target mRNA. In both forms, the system demonstrated a fast response. Key consequences of this system have been characterizations of important parameters in RNA circuit design, as by introducing specific mutations in portions of the RNA sequences, the switching behavior of the system can be altered.

Work thus far on synthetic circuits in single-celled organisms has generated devices that are rather reminiscent of the naturally occurring circuits in these same organisms. It remains to be seen whether more complex devices can be built in bacteria; whether current attempts are limited by the organisms, the creativity of the designers, and/or the fact that bio-circuit design is in an early stage. In their book *The Plausibility of Life*, Kirschner and Gerhart (2005) argue for the idea of "evolvability"; put anthropomorphically, some evolutionary precursors gave rise to diverse types of organisms be-

cause they had design features that were easier for evolution to work with. For example, the regulated specification of cell type seen in insects and mammals may have allowed the evolution of a wide variety of body types, while the invariant cell lineages seen in nematodes may have limited us to a large number of nematode species that all look like nematodes.

In the context of synthetic biology, it remains to be seen whether natural evolvability corresponds to engineerability. For example, it may prove difficult to design bacteria that show complex pattern formation because the cell wall may make it impossible for cells to communicate by cell contact, and communication solely by diffusible small molecules may be too difficult to use for anything interesting. Similarly, in spite of the evidence for some level of action at a distance on DNA, the mechanisms of bacterial gene regulation may mean that there is an inherent limit to the complexity of synthetic biological machines we can build in prokaryotes. Addressing these questions will be one challenge for future synthetic biologists.

#### *Artificial cell-cell interaction systems*

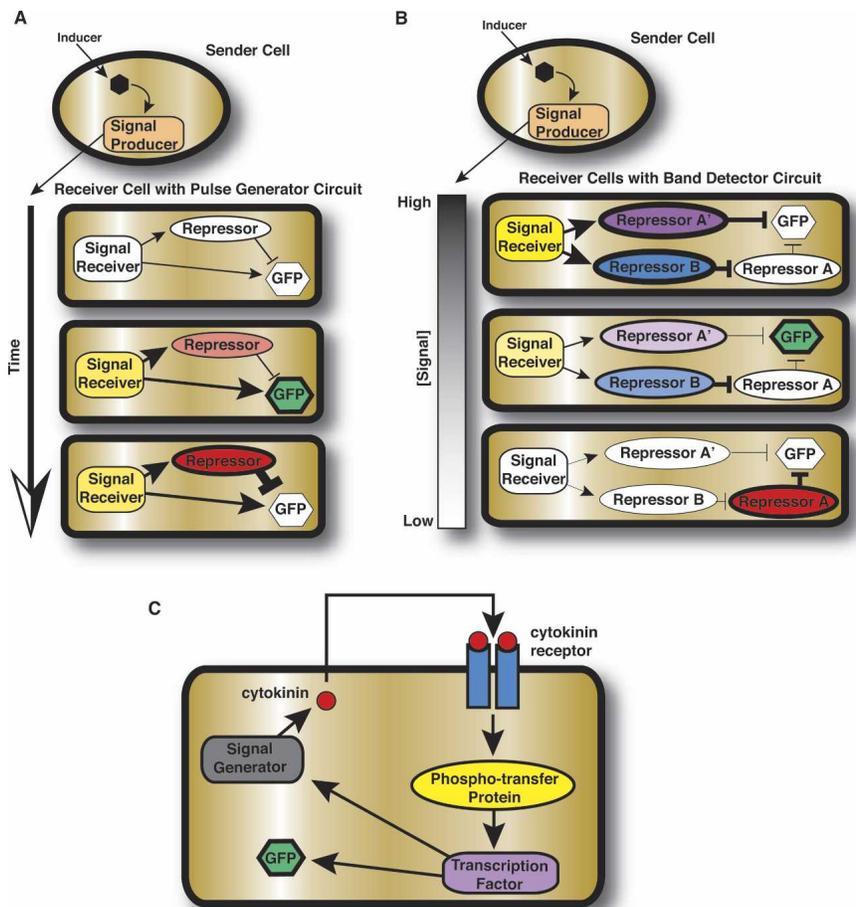
Synthetic circuits have also been engineered in the context of multicellular systems. Basu et al. (2004, 2005) designed two synthetic systems from well-characterized natural modules that are comprised of two different cell types in a bacterial population. In both systems, one type of bacteria is a sender cell that produces a signal, while the other type of cell is a signal receiver. The receiver cells dictate the behavioral response of the system to the sender cell inducer signal, which in both cases is the freely diffusing molecule AHL.

The first communication circuit was designed as a pulse generator, which responds to the spatiotemporal characteristics of the inducer signal (Basu et al. 2004). In response to a long-lasting increase of inducer produced by the sender cells, the receiver cells respond with a pulse of GFP. The receiver cell circuit was designed such that both GFP (pulse reporter) and the lambda repressor are expressed in response to AHL concentration via the AHL-activated LuxR transcription factor (Fig. 2A). Transcription of GFP is also repressible by lambda repressor. As AHL levels rise, there is a pulse of GFP expression, the result of the initial rise in GFP expression followed by its transcriptional repression (coupled with fast GFP degradation) by the concomitant rise in lambda repressor. This pulse-generating circuit was modeled and engineered to produce various characteristic responses in such aspects as duration, kinetics, and intensity depending on the versions of parts used, such as different operator-binding sites. The end result of this system was the demonstration of receiver cells with a varying response to sender cells based on the rate of increase of inducer concentration and their distance from the sender cell.

In the second multicellular system the receiver cell was designed as a band detector of sender cell AHL production (Basu et al. 2005). Specifically, the receiver cells

Drubin et al.

**Figure 2.** Artificial cell–cell communication networks. (A) In the multicell bacterial pulse generator network, the sender cell synthesizes the signaling molecule AHL in response to the inducer tetracycline. A signal is released and builds up locally over time. Initially, in the receiver cells, there is no GFP expression. The presence of AHL then activates the LuxR signal receiver, which then activates transcription of the GFP and lambda repressor genes. GFP protein is then expressed. The GFP gene is also repressible by lambda repressor. Over time, lambda repressor accumulates above the threshold required to repress GFP transcription, and with a fast degradation of GFP protein, the GFP signal fades. The result is a radiating wave that moves outward from the signal (Basu et al. 2004). (B) The receiver cell band detector circuit is also built around LuxR as the AHL signal receiver. In the presence of low concentration of AHL, LuxR is not active, LacI (repressor A, which represses GFP expression), is expressed, and GFP is off. In an intermediate amount of AHL, the receiver cell LuxR activates transcription of a second copy of the LacI gene (repressor A') and lambda repressor (repressor B). LacI does not accumulate enough in this condition to repress the GFP gene, and lambda repressor represses expression of the first LacI gene (repressor A), thus allowing GFP expression. When AHL concentration is high, the second LacI (repressor A') accumulates enough to repress GFP expression (Basu et al. 2005). (C) In the yeast-based quorum sensor device, the sender and receiver cells are combined into one cell. A cell generates the *A. thaliana* cytokinin signal (a diffusible small molecule), which is secreted and activates the heterologous cytokinin receptor on itself and other identical yeast cells. The activated receptor triggers a modified yeast osmosensing–signaling cascade via the YPD1 phospho-transfer protein, which activates the SKN7 transcription factor. SKN7 then activates GFP transcription as well as transcription of the cytokinin generator, which confers positive feedback regulation in the network. As cell density increases, the signal is amplified as well as GFP reporter expression (Chen and Weiss 2005).



respond with GFP expression to specific concentrations ranges of inducer, which on solid media is dictated by the distance of the receiver cell from the sender. The receiver cell circuit is again governed by the LuxR response to AHL, which then modulates an elegant circuit design of three repressors (lambda repressor, and two versions of LacI) directing GFP expression in a certain range of AHL concentration (Fig. 2B). The engineering of this system is a further example of the use of modeling to elicit desired behavior by guiding design and parameter tuning. By modulating the physical properties (e.g., degradation rates) and amounts of the circuit components, the concentration ranges within which GFP is expressed can be specified. With further modeling and experimentation, various complex patterns of GFP expression can be observed on a field of receiver cells that are dependent on the positions, amounts, and density of sender cells.

Cell–cell communication has also been explored in eukaryotes with a novel synthetic circuit in yeast by Chen and Weiss (2005). To create the systems, components

from the *Arabidopsis thaliana* cytokinin hormone pathway (for cytokinin synthesis and the cytokinin receptor) were connected to the endogenous downstream elements of the SLN1–YPD1–SSK1/SKN7 osmosensor signaling system in yeast. The yeast pathway was then linked to a GFP reporter via an SKN7-responsive synthetic promoter. In one system, the sender cell secretes the cytokinin signal and activates the heterologous signaling pathway in the receiver cells. A synthetic quorum-sensing network was achieved when the sender and receiver circuits were integrated into the same cell, with the cytokinin production under positive feedback regulation (Fig. 2C).

These experiments on artificial cell interaction systems represent the first synthetic-biological examination of the role of diffusion gradients in development, an area with a long, rich history of experimentation and model-building (e.g., Day and Lawrence 2000; Gregor et al. 2005). These initial demonstration experiments illustrate that the basic phenomenon of cells responding dif-

ferentially to a diffusion gradient can indeed be recapitulated by an artificial system with rather few components. As mentioned above, it will be interesting to see how many spatially organized cell types can be generated with diffusion gradients alone, and whether direct cell-cell interactions and oriented asymmetric cell divisions are also necessary to generate the complexity characteristic of multicellular animals.

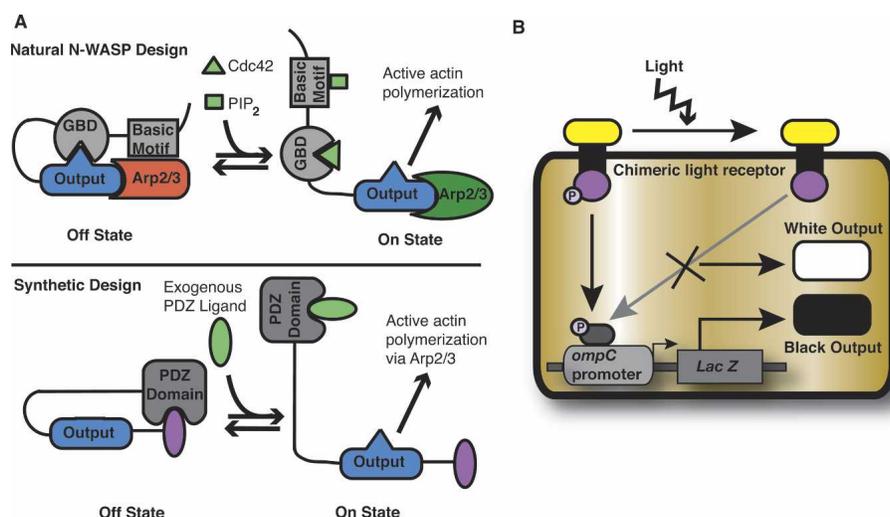
### Engineering signal transduction

Many cellular functions and interactions with the environment are mediated by multiple, interlinking signal transduction cascades that are made up of complex circuits of proteins that can be enzymatic and/or structural in nature. These proteins are comprised of many different modular domains, which confer specific functions and pathway connectivity and dictate the inputs and outputs of a signaling network (Bhattacharyya et al. 2006; Seet et al. 2006). Gating of a signaling protein or pathway may be regulated by direct modification (phosphorylation, for example) or through the binding of a specific ligand. Understanding and manipulating such regulatory mechanisms will increase the complexity and flexibility of synthetic network designs (Dueber et al. 2004; Pawson and Lindling 2005).

Dueber et al. (2003) describe the modular reprogramming of an allosteric protein signaling switch in yeast. Specifically, the hybrid protein was constructed around

the N-WASP-regulated actin polymerization output (Rohatgi et al. 1999; Prehoda et al. 2000). However, instead of the normal inputs that induce allosteric activation of the N-WASP output domain, the hybrid protein was engineered to have a different autoinhibitory input domain (e.g., a PDZ domain) that responded to different inducers (e.g., a PDZ domain ligand; Fig. 3A). In this way, the N-WASP output was coupled to heterologous inputs, creating a novel signaling circuit. Additionally, they created a library of different gating behaviors using both the PDZ- and SH3-binding domains with a variety of domain ligand affinities as the basis of the autoinhibitory interactions. As in other synthetic biology endeavors, the generation process of the chimeric signaling protein enabled the designers to observe and understand how certain parameters affected the behavior of the switch.

Signaling inputs and outputs can also be rewired via scaffold and adapter proteins, which link signaling proteins in a cascade together. The yeast scaffold proteins Ste5 and Pbs2, which normally mediate  $\alpha$ -factor and osmolarity responses respectively, were fused and engineered to channel an  $\alpha$ -factor input to an osmo-response output (Park et al. 2003). Chimeric adapter proteins have also been constructed, in which the phosphotyrosine recognition domains Grb2 and ShcA were fused to a Fadd death effector domain, effectively channeling mitogenic or transforming receptor tyrosine kinase signals to the triggering of cell death (Howard et al. 2003). The future of fast responding genetic circuits lies in understanding



**Figure 3.** Engineered signal transduction. (A) An allosteric protein switch based on the natural N-WASP allosteric switch. Naturally, N-WASP can activate the Arp2/3 complex, which mediates actin polymerization, via its output domain. In the off state, the N-WASP output domain is repressed by the GTPase-binding domain (GBD) and basic domain. In the presence of Cdc42 and PIP<sub>2</sub>, the autoinhibition is relieved, and the output domain functions to activate Arp2/3. In the synthetic switch, the mediators of the intramolecular autoinhibition are swapped out, thus changing the input of the switch. In this case, a PDZ-binding domain is used with a C-terminal PDZ ligand. In the off state, the binding of the PDZ domain to its ligand conformationally blocks the output domain activation function. When an exogenous PDZ ligand is added, the intramolecular PDZ interaction is disrupted, and the output domain is free to stimulate actin polymerization via Arp2/3 (Dueber et al. 2003). (B) Light sensation in *E. coli* based on a synthetic sensor-kinase system (the “bacterial camera”) (Levskaya et al. 2005). The chimeric phytochrome Cph1 and EnvZ histidine kinase serves as the light sensor. In the absence of light, the kinase activity is active, phosphorylating the OmpR transcription factor and inducing expression of LacZ, resulting in a black output when exposed to S-gal (a substrate that becomes black when cleaved by  $\beta$ -galactosidase). When exposed to light, the kinase and transcription factor become inactive, and LacZ is no longer expressed, leaving the cells light in the presence of S-gal.

Drubin et al.

and manipulating signaling proteins and their corresponding enzymatic cascades. Not only will this improve the kinetics of synthetic pathways, but will allow the wiring of novel, more complex synthetic cell signaling circuits.

Another form of engineering signaling pathways is the establishment of ectopic signaling responses. Some early examples of this include yeast that were conferred a signaling response to various exogenous stimuli, such as  $\beta$ -adrenergic receptor agonists (King et al. 1990) and somatostatin (Price et al. 1995), via the ectopic expression of specific G-protein-coupled receptors (GPCRs) linked to signal through the endogenous yeast pheromone response pathway. Recently, a novel bacterial network was engineered to “see” light (Fig. 3B; Levskaya et al. 2005). In this case, *E. coli* were designed to function as a bacterial film when plated as a lawn and exposed to light. The light-sensing function was conferred upon the *E. coli* by the creation of a protein chimera, consisting of an extracellular light-sensing phytochrome module (Cph1) from *Synechocystis*, coupled with the intracellular histidine kinase signaling domain from the natural *E. coli* EnvZ, part of the EnvZ–OmpR two-component system. A phytochrome is also a two-component signaling system that has an extracellular sensor domain and an intracellular response-regulator. In addition, it was necessary to engineer the bacteria to metabolically produce the light-responding part of the photoreceptor, phycocyanobilin, through the introduction of specific biosynthetic genes. The EnvZ–OmpR part of the hybrid system confers the signal from light to transcriptional regulation at a target promoter regulating a reporter gene, in this case *lacZ*. When there is light, the reporter gene is repressed, and the cell is unable to process the S-gal (3,4-Cyclohexenoesuletin  $\beta$ -D-galactopyranoside) substrate to a black precipitate. Thus, exposure of a bacterial lawn to a particular pattern of light results in a clear, white patterned region within an unexposed background of black. Ultimately, this work demonstrated a novel function imparted to an organism and used in an innovative way by using both engineered signaling pathways and metabolic engineering. In this vein, bacteria were also recently engineered with heterologous sensors in order to invade the environment of a tumor (Anderson et al. 2006).

### Metabolic pathway engineering

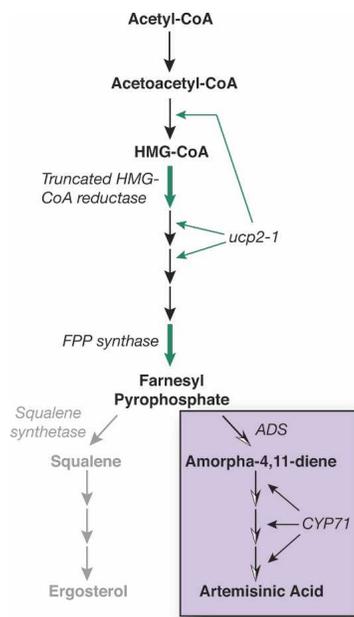
The concepts of synthetic biology have also been applied to metabolic engineering (Khosla and Keasling 2003; Chang and Keasling 2006). Using engineered pathways and modular biosynthesis cascades in tractable organisms such as *E. coli* and *Saccharomyces cerevisiae*, normal cellular metabolism may be altered to produce non-natural metabolites or to shift the balance of synthesis toward a key metabolite of interest. This metabolite may be the final output sought for or a key intermediate that can be easily processed further to the molecule of choice. An important aspect of metabolic engineering is integrating the new pathway into a cell, taking into account

the native metabolites and pathway operations necessary for maintenance of essential cell functions.

A key venture, representative of this synthetic biology arena, has been made by the Keasling laboratory (Roth and Acton 1989; Martin et al. 2003; Ro et al. 2006), which is pursuing the cost-effective synthesis of an immediate precursor to the antimalarial drug artemisinin, artemisinic acid. Faced with the prevalence of multi-drug-resistant strains of *Plasmodium falciparum*, effective drugs are needed at an affordable cost for Third-World patients. Artemisinin is an effective treatment that is being used in cocktail therapies to treat strains resistant to standard, cheaper treatments. Artemisinin is naturally produced by the *Artemisia annua* (sweet wormwood), but is difficult and expensive to obtain (Schmid and Hofheinz 1983; Enserink 2005). An adult course of artemisinin-based treatment through the World Health Organization costs \$2.40 US, compared with just \$0.20 US for a course of the standard chloroquine treatment. A key factor in artemisinin’s cost is its extensive preparation time, up to at least 8 mo from planting to processing, and raw material short supply in the face of increasing demand.

To bring the costs of artemisinin down, Ro et al. (2006) have described the engineering of yeast to produce artemisinic acid. The crux of this endeavor lies in the natural yeast mevalonate pathway, which normally leads to the biosynthesis of sterols from acetyl-CoA arising from the metabolism of simple sugars. In the engineered system (Fig. 4), the mevalonate pathway is tuned to produce an increased amount of the intermediate farnesyl pyrophosphate (FPP). FPP is primarily converted to squalene and then sterols, but can be harnessed by the ectopic expression of enzymes from *A. annua* for the production of artemisinic acid. However, just simply expressing a few key enzymes is not enough for efficient production of artemisinic acid in yeast. The natural mevalonate pathway was optimized with respect to FPP production, as monitored via the synthesis of the artemisinic acid precursor amorphodiene from FPP by *A. annua* amorphodiene synthase. The output was systematically optimized by overexpression of a truncated, soluble form of HMG-CoA reductase (which catalyzes the first committed step in FPP production), reducing the expression of *ERG9* (which converts FPP to squalene) by expressing it from the methionine-repressible  $P_{MET3}$  promoter, and including a gain-of-function mutation in *UCP-2*, which regulates sterol production; these changes increase FPP production and also decrease the shuttling of FPP to natural sterol production. Coupling such an optimized amorphodiene producer with the expression of a specific cytochrome P450 monooxygenase from *A. annua*, the completed yeast strain successfully produces artemisinic acid at higher specific productivity than naturally obtained from *A. annua*. With further work on yield optimization and scaled-up harvesting, they predict artemisinin-based therapies at significantly reduced production time and cost.

The production of artemisinic acid in yeast is just one example of the many successful ventures in metabolic



**Figure 4.** Engineering artemisinin acid production in yeast. Three types of modifications were introduced: (1) enhanced expression of enzymes that lead to production of an intermediate from general metabolism (FPP) using the general regulatory mutation *ucp2-1* and specific overexpression of two enzymes, (2) down-regulation of the initial step for ergosterol synthesis by reduced transcription of squalene synthetase, and (3) introduction of genes for amorphaadiene synthase (ADS), a specific cytochrome P450 (CYP71AV1), and its redox partner (NADPH: cytochrome p450 oxidoreductase) from *A. annua*, recreating the pathway for artemisinin acid production (Ro et al. 2006).

engineering. Others include the engineered production of polyketides, a diverse family of natural metabolites that includes many useful pharmaceuticals (e.g., erythromycin), in heterologous hosts (Pfeifer and Khosla 2001), and the creation of Golden Rice, via the introduction of the provitamin A biosynthetic pathway into rice endosperm, to address the vitamin A deficiency in populations that rely on rice as a primary staple (Ye et al. 2000; Beyer et al. 2002). Current pursuits also include the microbial production of many other currently expensive natural drugs, such as taxol (Dejong et al. 2006) for cancer and prostratin for HIV (Hezareh 2005).

#### Novel biosensors through protein engineering

In the examples described thus far, synthetic biologists have used protein modules as provided by nature. When the modules provided by nature are not adequate for a given purpose, protein engineering techniques may complement the existing array of biological parts. The continued improvement in computer-driven protein design, including design of entire protein domains with novel folds (Kuhlman et al. 2003) and protein interaction surfaces (Kortemme et al. 2004) should allow design of proteins with novel and useful activities.

Looger et al. (2003) have presented an important study in this regard, demonstrating the use of a structure-based computational design method for the engineering of ligand-binding specificities. As a platform for the redesigned ligand-binding proteins, five members of the periplasmic-binding protein (PBP) superfamily from *E. coli* were used (Dwyer and Hellinga 2004), each naturally binding glucose, ribose, arabinose, glutamine, and histidine, respectively. The computational design method involves the use of high-resolution three-dimensional structures to find an amino acid sequence that would create a binding surface for the ligand of interest. Based on amino acid residues in contact with the natural ligand, the method employs a detailed algorithm to determine a rank-ordered list of optimal amino acid changes to the protein platform. Ultimately, starting with the listed members of the PBP family, binding sites to trinitrotoluene (TNT), L-lactate, and serotonin were engineered; all new target ligands are chemically unrelated from the native ligands. This group then used these redesigned receptor molecules to connect the binding of their new target ligand to their natural signaling cascade in *E. coli*. The engineered receptors responded specifically to their new targets, and not their original ligands, and activated a reporter gene linked to the natural signaling cascade. In the case of the TNT sensor, this technique has provided a novel biological detector of explosives, and the cellular detection of L-lactate and serotonin have applications in medicine. Ribose-binding protein has also been conferred enzymatic activity by the same group through a computational design method (Dwyer et al. 2004). The expression of natural mammalian GPCRs in yeast, such as the somatostatin receptor construct of Price et al. (1995) are early examples of biosensor construction, but these were limited to the specificities of naturally occurring GPCRs. The recent work from the Hellinga group described here illustrates how computer-driven protein engineering can enhance the human design of biological systems.

#### Minimal cells and synthetic genomes

One of the goals of synthetic biology is to gain a better understanding of life and how the systems comprising a cell functionally integrate. A major strategy to address this is to define the minimal components in a genome that are sufficient for life. A cell may become easier to study and predictably engineer if one knows the essential, well-defined elements.

There are currently two complementary strategies for studying the simplest organism: the top-down approach, where simple organisms are simplified further through the removal of nonessential genetic elements; and the bottom-up approach, where the synthesis of artificial cells is attempted, component by component (Deamer 2005; Forster and Church 2006). Work toward the latter strategy has seen recent progress with the demonstration of a functional lipid enclosed bioreactor, where the expression of GFP by an enclosed cell-free expression system derived from *E. coli* and wheat germ was maintained

Drubin et al.

for 4 d as a result of external nutrient uptake (Noireaux and Libchaber 2004). In addition, gene transcription and protein synthesis have been coupled within lipid vesicles (Ishikawa et al. 2004). The top-down approach is being pursued in several simple organisms such as *E. coli* (Kolisnychenko et al. 2002; Posfai et al. 2006) and *Mycoplasma* (Glass et al. 2006). Once a minimal genetic complement is established for a system, it is envisioned that a whole chromosome with this minimal genome could be synthesized and inserted into a cell, or generated by a small number of sequential replacement steps. In addition to finding the minimal set of genes, work has also been performed to show that reorganization of a T7 viral genome, in this case via separation of natural genes that overlap each other (thus increasing the overall size of the genome), results in a functional virus (Chan et al. 2005).

One specific, directed application of a synthetically generated minimal genome is in the creation of a cellular vessel that would avail the expanded use of nonnatural amino acids (outside the natural 20 amino acids) in proteins *in vivo*. The incorporation of nonnatural amino acids in cellular proteins is a strategy of synthetic biology that seeks to generate novel proteins and protein functions via the combinations of an expanded library of amino acid side chain chemistries (Wang et al. 2006). Such a strategy is also important in the study of natural protein folding and function. Currently, >30 nonnatural amino acids have been artificially inserted in proteins *in vivo* cotranslational insertion. These nonnatural amino acids have been of many varieties, including, among others, those conferring fluorescence and redox activities.

The current strategies for *in vivo* incorporation of nonnatural amino acids are limited; only very few different nonnatural amino acids can be incorporated into the same protein, or different proteins in the same cell. With the ability to synthesize a specific genome sequence, alternative codons for the 20 natural amino acids could be eliminated in the entire genome. In this way, just 20 codons would specify each of the 20 natural amino acids, providing additional codons to be used to code for nonnatural amino acids. Provided an organism with a genome of this design would function, the incorporation of a variety of nonnatural amino acids in the same cell and even in the same protein without nonspecific effects on cellular function could be realized.

Although the size of even a small bacterial genome (on the order of 300–500,000 nucleotides) may currently be technologically challenging for *de novo* generation, there have been key strides made in genome synthesis, as will be discussed below. Already the genomes of infamous pathogens, such as poliovirus (Cello et al. 2002) and the 1918 pandemic influenza virus (Tumpey et al. 2005), have been synthesized and used to create infectious virus. In the case of influenza, obtaining live virus has allowed the study of the potential threat of future pandemic viruses and insight into what types of countermeasures we might employ should a new outbreak arise. With the ability to generate whole genomes from

scratch, multiple gene additions and modifications will also become more tractable, allowing the creation of even more complex synthetic systems and functionally specific cells. The major obstacle to the ready, widespread realization of the benefits resulting from gene and genome synthesis lies in the cost.

## Technical advances

### DNA synthesis

The efficient and cheap synthesis of large pieces of DNA is a driving force for the progression of some synthetic biology endeavors (Carlson 2003). Once synthesis costs drop below a certain level, the time and labor involved in constructing a multiple component circuit via plasmid-based cloning and insertion into cells will become more expensive than commercial synthesis. Limiting factors such as DNA length, synthesis fidelity, time, and cost are being brought to levels where the synthesis of genes for engineered systems is increasingly becoming a more realistic alternative to standard gene manipulation techniques.

The first report of the synthesis of extended stretches of DNA was in 1995, with the PCR-based synthesis of a 2.7 kb plasmid from a pool of short, overlapping synthesized oligonucleotides (Stemmer et al. 1995). The first synthesized genome, poliovirus, proved in principle that genome synthesis could be accomplished (Cello et al. 2002). However, the technique is much too laborious and costly for application toward larger sequences. The successful synthesis of  $\phi$ X174 bacteriophage reduced the time to generate a similarly sized genome and offered the prospect of building yet larger genomes (Smith et al. 2003). However, the technique used selection for infectivity (coupled with sequencing) to identify the phage genomes that were functional and without significant mutations. Smith et al. (2003) point out that without the selectivity, the subgenome cassettes used to build the genome would harbor about two mutations per kilobase just due to errors in the oligo synthesis, and thus unfit for whole chromosome and genome synthesis. Sequencing and repair by site-directed mutagenesis is again proposed. However, this adds time and cost. Ultimately, the weakness of these techniques is that they begin with the use of oligos synthesized by current methods, which are expensive and have too high a rate of errors for use on the genomic scale.

Another gene synthesis technology based on the same principles of oligo assembly as the previous accomplishments has recently been demonstrated. In this case, however, the high cost and error rate of oligo synthesis and laborious multistep assembly, sequencing, and site-directed error correction are addressed. The technique couples large-scale parallel custom oligo synthesis on microfluidic microarrays with mismatch error correction and a single-step polymerase assembly multiplexing reaction to build the gene or gene combinations of interest (Tian et al. 2004). Ultimately, an ~14.6-kb assembly was generated with two errors, a rate significantly less

than previous methods. Further reduction in error rate is proposed by the addition of natural DNA mismatch repair proteins into the procedure.

Technological innovations such as these are quickly driving down the costs of accurate gene synthesis. While more accessible than in the past, the current cost is still prohibitive for regular, everyday use. With the rapid development and integration of less labor-intensive ways of obtaining fast synthesis with low errors, the era of gene synthesis is soon approaching where biological hypotheses will be tested via the design of genes and genomes made to order.

### *Standard biological parts*

Like early 20th Century electronic circuit construction, modern genetic circuit construction effectively “hand-wires” DNA components, such that assembly is slow, and the results of similar but nonidentical genetic circuit designs may not be consistent. The lack of standardization in both gene assembly methods and the modules themselves is a current bottleneck for the efficient advancement of synthetic biology endeavors. To address these issues, one current synthetic biology ideology includes the following (Endy 2005): (1) use of well-characterized parts, (2) standardization of assembly methods, and (3) abstraction.

Well-characterized “parts” are generally the parts that molecular biologists know and love: the lac promoter, the CMV promoter, lacZ, GFP, and so on. One goal of synthetic biology is that the quantitative characteristics of such parts should be well documented. This part of the synthetic biology program has been articulated based on an analogy with the microchip industry, where the physical specifications of chips are extremely well characterized. For example, it is important to know the temperature range in which a given chip will function, so that the specs for appliances using the chips can be defined. This level of characterization obviously has rarely been achieved for commonly used biological parts, and remains mostly anecdotal.

One vision of standard assembly is currently articulated in the “biobricks” concept (Knight 2003; Phillips and Silver 2006). According to this idea, a given part will have a set of restriction sites that will allow standard cloning and construction of a series of parts to form a synthetic gene unit such as a hybrid promoter and/or a fusion protein. If all of the synthetic biology community uses this or some other standard assembly convention, then all parts can be shared and combined without compatibility issues. The reliance on standard assembly is, albeit currently necessary, only an intermediate step. As the feasibility of synthesizing longer pieces of DNA develops, scientists and bioengineers will be able to design and order combinations of standard, well-characterized parts directly.

Finally, abstraction is anticipated when large-scale constructions will be done. Again, by analogy to the microchip industry, chip designers do not need to know the details of how the components work, such as the atomic

details of semiconductor materials, and programmers do not need to know anything at all about how the chips work. By habit, molecular biologists also work at some level of abstraction; most practitioners have no idea about the three-dimensional structure of the proteins that they use. What is anticipated is that some fairly complex parts that consist of many components can be abstracted to an input–output curve, in which the internal workings can be ignored.

The culmination of all these ideals is currently exemplified in the Registry of Standard Biological Parts (<http://parts.mit.edu>). The mission of the Registry is to catalog a library of fully characterized biological modules, from individual genes or parts of genes to complete gene circuits; the idea being that if one is looking for a particular functionality in a designed biological system, one can search the catalog and obtain the component or circuit that fulfills the need of the design. This detailed database will help spare arduous “reinvention” and optimization time resulting from unknown parameters and undefined problems.

It remains to be seen how the parts/standardization/abstraction approach will play out in practice. Several potential problems can be envisioned. First, who will characterize the parts? The motivation to do this is not part of the value system of current, practicing molecular biologists, for whom qualitative characterization is generally considered adequate. Second, what are the parameters that should be characterized? These may be different for every type of part: For promoters, one might want the maximal transcription rate, the maximally repressed transcription rate, the response curve relative to concentrations of a regulatory protein, etc.; for a reporter gene, one might want the excitation and emission wavelengths. But what about the translation efficiency based on codon usage, folding times, or degradation rates of the mRNA and protein? In the case of genes that are to be integrated into the chromosome, stochastic effects on gene expression may (or may not) be largely compared with the characteristics that are measured. Third, the repeated use of characterized parts (e.g., promoters) may lead to a risk of genetic recombination between them.

### **Educational initiatives**

To establish itself as a discipline, the synthetic biology community has sought ways to introduce biologists and engineers to synthetic biology. There are already some curricula in place that can be used as models for the development of others. For example, there are synthetic biology courses offered at University of California at Berkeley, Massachusetts Institute of Technology, and Boston University. There is also an annual summer competition organized by MIT—the international Genetically Engineered Machine (iGEM) competition (see [http://parts2.mit.edu/wiki/index.php/Main\\_Page](http://parts2.mit.edu/wiki/index.php/Main_Page))—in which students from around the world join together as teams in an attempt to engineer novel biological devices. The competitions, which began in 2004 with five teams, grew to 37 teams in 2006, and included entries from the

Drubin et al.

United States, Europe, Asia, and South America. In these competitions, teams of students design and attempt to build interesting biological devices, the products of which are contributed as devices and parts to the Registry of Standard Biological Parts for use and further development and characterization. Hopefully, with such programs in place, they will inspire and serve as examples for the development of teaching materials and expansion of institutions providing education in synthetic biology.

### Future enterprises

Previous reviews (cited above) have set forth an ambitious program for synthetic biology. We expect that DNA synthesis costs will continue to fall and computers will become more powerful, but DNA costs are already rather low and the computing power of a typical molecular biology lab is probably greater than that of the entire world in 1970. Ultimately, the rate-limiting factor for the future development of synthetic biology may actually be human creativity. Future "technology development" may, in fact, consist of novel ways of thinking about life that allow us to build things that are truly new.

On the academic front, synthetic biology addresses the questions "How do we know that we completely understand a biological system? How can we be certain that we aren't missing something?" By traditional approaches, one can mutate and perturb a system and obtain the desired results, but this will not allow one to definitively answer these questions. The answer proposed by synthetic biology is that successfully rebuilding a biological system, using different parts with the same quantitative properties, is necessary. In this sense, the synthetic biologists follow the words of Richard Feynman ([http://en.wikipedia.org/wiki/Richard\\_Feynman](http://en.wikipedia.org/wiki/Richard_Feynman)): "What I cannot create, I do not understand."

### Acknowledgments

We thank Jake Wintermute for reviewing the manuscript, and the Synthetic Biology Working Group for input on content. This work was supported by funding from the National Institutes of Health to P.A.S. and an NIH Ruth L. Kirschstein National Research Service Award to D.A.D.

### References

- Anantharam, A., Markowitz, S.M., and Abbott, G.W. 2003. Pharmacogenetic considerations in diseases of cardiac ion channels. *J. Pharmacol. Exp. Ther.* **307**: 831–838.
- Anderson, J.C., Clarke, E.J., Arkin, A.P., and Voigt, C.A. 2006. Environmentally controlled invasion of cancer cells by engineered bacteria. *J. Mol. Biol.* **355**: 619–627.
- Andrianantoandro, E., Basu, S., Karig, D.K., and Weiss, R. 2006. Synthetic biology: New engineering rules for an emerging discipline. *Mol. Syst. Biol.* **2**: 2006.0028.
- Arkin, A.P. 2001. Synthetic cell biology. *Curr. Opin. Biotechnol.* **12**: 638–644.
- Arkin, A.P. and Fletcher, D.A. 2006. Fast, cheap and somewhat in control. *Genome Biol.* **7**: 114.
- Atsumi, S. and Little, J.W. 2004. Regulatory circuit design and evolution using phage  $\lambda$ . *Genes & Dev.* **18**: 2086–2094.
- Basu, S., Mehreja, R., Thiberge, S., Chen, M.T., and Weiss, R. 2004. Spatiotemporal control of gene expression with pulse-generating networks. *Proc. Natl. Acad. Sci.* **101**: 6355–6360.
- Basu, S., Gerchman, Y., Collins, C.H., Arnold, F.H., and Weiss, R. 2005. A synthetic multicellular system for programmed pattern formation. *Nature* **434**: 1130–1134.
- Bayer, T.S. and Smolke, C.D. 2005. Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nat. Biotechnol.* **23**: 337–343.
- Becskei, A. and Serrano, L. 2000. Engineering stability in gene networks by autoregulation. *Nature* **405**: 590–593.
- Becskei, A., Seraphin, B., and Serrano, L. 2001. Positive feedback in eukaryotic gene networks: Cell differentiation by graded to binary response conversion. *EMBO J.* **20**: 2528–2535.
- Benner, S.A. 2003. Synthetic biology: Act natural. *Nature* **421**: 118.
- Benner, S.A. and Sismour, A.M. 2005. Synthetic biology. *Nat. Rev. Genet.* **6**: 533–543.
- Beyer, P., Al-Babili, S., Ye, X., Lucca, P., Schaub, P., Welsch, R., and Potrykus, I. 2002. Golden Rice: Introducing the  $\beta$ -carotene biosynthesis pathway into rice endosperm by genetic engineering to defeat vitamin A deficiency. *J. Nutr.* **132**: 506S–510S.
- Bhattacharyya, R.P., Remenyi, A., Yeh, B.J., and Lim, W.A. 2006. Domains, motifs, and scaffolds: The role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu. Rev. Biochem.* **75**: 655–680.
- Breithaupt, H. 2006. The engineer's approach to biology. *EMBO Rep.* **7**: 21–23.
- Brent, R. 2004. A partnership between biology and engineering. *Nat. Biotechnol.* **22**: 1211–1214.
- Carlson, R. 2003. The pace and proliferation of biological technologies. *Biosecur. Bioterror.* **1**: 203–214.
- Cello, J., Paul, A.V., and Wimmer, E. 2002. Chemical synthesis of poliovirus cDNA: Generation of infectious virus in the absence of natural template. *Science* **297**: 1016–1018.
- Chan, L.Y., Kosuri, S., and Endy, D. 2005. Refactoring bacteriophage T7. *Mol. Syst. Biol.* **1**: 2005.0018.
- Chang, M.C. and Keasling, J.D. 2006. Production of isoprenoid pharmaceuticals by engineered microbes. *Nat. Chem. Biol.* **2**: 674–681.
- Chen, M.T. and Weiss, R. 2005. Artificial cell–cell communication in yeast *Saccharomyces cerevisiae* using signaling elements from *Arabidopsis thaliana*. *Nat. Biotechnol.* **23**: 1551–1555.
- Chen, J.N., Haffter, P., Odenthal, J., Vogelsang, E., Brand, M., van Eeden, F.J., Furutani-Seiki, M., Granato, M., Hamerschmidt, M., Heisenberg, C.P., et al. 1996. Mutations affecting the cardiovascular system and other internal organs in zebrafish. *Development* **123**: 293–302.
- Chin, J.W. 2006a. Modular approaches to expanding the functions of living matter. *Nat. Chem. Biol.* **2**: 304–311.
- Chin, J.W. 2006b. Programming and engineering biological networks. *Curr. Opin. Struct. Biol.* **16**: 551–556.
- Day, S.J. and Lawrence, P.A. 2000. Measuring dimensions: The regulation of size and shape. *Development* **127**: 2977–2987.
- Deamer, D. 2005. A giant step towards artificial life? *Trends Biotechnol.* **23**: 336–338.
- Dejong, J.M., Liu, Y., Bollon, A.P., Long, R.M., Jennewein, S., Williams, D., and Croteau, R.B. 2006. Genetic engineering of taxol biosynthetic genes in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **93**: 212–224.
- Dueber, J.E., Yeh, B.J., Chak, K., and Lim, W.A. 2003. Reprogramming control of an allosteric signaling switch through

- modular recombination. *Science* **301**: 1904–1908.
- Dueber, J.E., Yeh, B.J., Bhattacharyya, R.P., and Lim, W.A. 2004. Rewiring cell signaling: The logic and plasticity of eukaryotic protein circuitry. *Curr. Opin. Struct. Biol.* **14**: 690–699.
- Dwyer, M.A. and Hellinga, H.W. 2004. Periplasmic binding proteins: A versatile superfamily for protein engineering. *Curr. Opin. Struct. Biol.* **14**: 495–504.
- Dwyer, M.A., Looger, L.L., and Hellinga, H.W. 2004. Computational design of a biologically active enzyme. *Science* **304**: 1967–1971.
- Elowitz, M.B. and Leibler, S. 2000. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**: 335–338.
- Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. 2002. Stochastic gene expression in a single cell. *Science* **297**: 1183–1186.
- Endy, D. 2005. Foundations for engineering biology. *Nature* **438**: 449–453.
- Enserink, M. 2005. Infectious diseases. Source of new hope against malaria is in short supply. *Science* **307**: 33.
- Ferrell Jr., J.E. 2002. Self-perpetuating states in signal transduction: Positive feedback, double-negative feedback and bistability. *Curr. Opin. Cell Biol.* **14**: 140–148.
- Forster, A.C. and Church, G.M. 2006. Towards synthesis of a minimal cell. *Mol. Syst. Biol.* **2**: 45.
- Gardner, T.S., Cantor, C.R., and Collins, J.J. 2000. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**: 339–342.
- Glass, J.I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M.R., Maruf, M., Hutchison III, C.A., Smith, H.O., and Venter, J.C. 2006. Essential genes of a minimal bacterium. *Proc. Natl. Acad. Sci.* **103**: 425–430.
- Gossen, M. and Bujard, H. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci.* **89**: 5547–5551.
- Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W., and Bujard, H. 1995. Transcriptional activation by tetracyclines in mammalian cells. *Science* **268**: 1766–1769.
- Gregor, T., Bialek, W., Deruyter van Steveninck, R.R., Tank, D.W., and Wieschaus, E.F. 2005. Diffusion and scaling during early embryonic pattern formation. *Proc. Natl. Acad. Sci.* **102**: 18403–18407.
- Guido, N.J., Wang, X., Adalsteinsson, D., McMillen, D., Hasty, J., Cantor, C.R., Elston, T.C., and Collins, J.J. 2006. A bottom-up approach to gene regulation. *Nature* **439**: 856–860.
- Hasty, J., Pradines, J., Dolnik, M., and Collins, J.J. 2000. Noise-based switches and amplifiers for gene expression. *Proc. Natl. Acad. Sci.* **97**: 2075–2080.
- Hasty, J., McMillen, D., and Collins, J.J. 2002. Engineered gene circuits. *Nature* **420**: 224–230.
- Heinemann, M. and Panke, S. 2006. Synthetic biology—Putting engineering into biology. *Bioinformatics* **7**: 512.
- Hezareh, M. 2005. Prostratin as a new therapeutic agent targeting HIV viral reservoirs. *Drug News Perspect.* **18**: 496–500.
- Hooshangi, S., Thiberge, S., and Weiss, R. 2005. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc. Natl. Acad. Sci.* **102**: 3581–3586.
- Howard, P.L., Chia, M.C., Del Rizzo, S., Liu, F.F., and Pawson, T. 2003. Redirecting tyrosine kinase signaling to an apoptotic caspase pathway through chimeric adaptor proteins. *Proc. Natl. Acad. Sci.* **100**: 11267–11272.
- Isaacs, F.J., Dwyer, D.J., Ding, C., Pervouchine, D.D., Cantor, C.R., and Collins, J.J. 2004. Engineered riboregulators enable post-transcriptional control of gene expression. *Nat. Biotechnol.* **22**: 841–847.
- Isaacs, F.J., Dwyer, D.J., and Collins, J.J. 2006. RNA synthetic biology. *Nat. Biotechnol.* **24**: 545–554.
- Ishikawa, K., Sato, K., Shima, Y., Urabe, I., and Yomo, T. 2004. Expression of a cascading genetic network within liposomes. *FEBS Lett.* **576**: 387–390.
- Jacob, F. and Monod, J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**: 318–356.
- Kaern, M., Blake, W.J., and Collins, J.J. 2003. The engineering of gene regulatory networks. *Annu. Rev. Biomed. Eng.* **5**: 179–206.
- Khosla, C. and Keasling, J.D. 2003. Metabolic engineering for drug discovery and development. *Nat. Rev. Drug Discov.* **2**: 1019–1025.
- King, K., Dohlman, H.G., Thorner, J., Caron, M.G., and Lefkowitz, R.J. 1990. Control of yeast mating signal transduction by a mammalian  $\beta$  2-adrenergic receptor and Gs  $\alpha$  subunit. *Science* **250**: 121–123.
- Kirschner, M. and Gerhart, J. 2005. *The plausibility of life: Resolving Darwin's dilemma*. Yale University Press, New Haven, CT.
- Knight, T. 2003. Idempotent vector design for standard assembly of biobricks. <http://hdl.handle.net/1721.1/21168>. In *DSPACE*. Massachusetts Institute of Technology Artificial Intelligence Laboratory, MIT Synthetic Biology Working Group.
- Knight, T.F. 2005. Engineering novel life. *Mol. Syst. Biol.* **1**: 2005.0020.
- Kobayashi, H., Kaern, M., Araki, M., Chung, K., Gardner, T.S., Cantor, C.R., and Collins, J.J. 2004. Programmable cells: Interfacing natural and engineered gene networks. *Proc. Natl. Acad. Sci.* **101**: 8414–8419.
- Kolisnychenko, V., Plunkett III, G., Herring, C.D., Feher, T., Posfai, J., Blattner, F.R., and Posfai, G. 2002. Engineering a reduced *Escherichia coli* genome. *Genome Res.* **12**: 640–647.
- Kortemme, T., Joachimiak, L.A., Bullock, A.N., Schuler, A.D., Stoddard, B.L., and Baker, D. 2004. Computational redesign of protein–protein interaction specificity. *Nat. Struct. Mol. Biol.* **11**: 371–379.
- Kramer, B.P. and Fussenegger, M. 2005. Hysteresis in a synthetic mammalian gene network. *Proc. Natl. Acad. Sci.* **102**: 9517–9522.
- Kramer, B.P., Viretta, A.U., Daoud-El-Baba, M., Aubel, D., Weber, W., and Fussenegger, M. 2004. An engineered epigenetic transgene switch in mammalian cells. *Nat. Biotechnol.* **22**: 867–870.
- Kuhlman, B., Dantas, G., Ireton, G.C., Varani, G., Stoddard, B.L., and Baker, D. 2003. Design of a novel globular protein fold with atomic-level accuracy. *Science* **302**: 1364–1368.
- Levskaia, A., Chevalier, A.A., Tabor, J.J., Simpson, Z.B., Lavery, L.A., Levy, M., Davidson, E.A., Scouras, A., Ellington, A.D., Marcotte, E.M., et al. 2005. Synthetic biology: Engineering *Escherichia coli* to see light. *Nature* **438**: 441–442.
- Looger, L.L., Dwyer, M.A., Smith, J.J., and Hellinga, H.W. 2003. Computational design of receptor and sensor proteins with novel functions. *Nature* **423**: 185–190.
- Martin, V.J., Pitera, D.J., Withers, S.T., Newman, J.D., and Keasling, J.D. 2003. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat. Biotechnol.* **21**: 796–802.
- McAdams, H.H. and Arkin, A. 1997. Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci.* **94**: 814–819.
- McDaniel, R. and Weiss, R. 2005. Advances in synthetic biology: On the path from prototypes to applications. *Curr. Opin. Biotechnol.* **16**: 476–483.
- Mettetal, J.T., Muzzey, D., Pedraza, J.M., Ozbudak, E.M., and van Oudenaarden, A. 2006. Predicting stochastic gene expression dynamics in single cells. *Proc. Natl. Acad. Sci.* **103**: 7304–7309.

Drubin et al.

- Noireaux, V. and Libchaber, A. 2004. A vesicle bioreactor as a step toward an artificial cell assembly. *Proc. Natl. Acad. Sci.* **101**: 17669–17674.
- Park, S.H., Zarrinpar, A., and Lim, W.A. 2003. Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. *Science* **299**: 1061–1064.
- Pawson, T. and Linding, R. 2005. Synthetic modular systems—Reverse engineering of signal transduction. *FEBS Lett.* **579**: 1808–1814.
- Pedraza, J.M. and van Oudenaarden, A. 2005. Noise propagation in gene networks. *Science* **307**: 1965–1969.
- Pfeifer, B.A. and Khosla, C. 2001. Biosynthesis of polyketides in heterologous hosts. *Microbiol. Mol. Biol. Rev.* **65**: 106–118.
- Phillips, I. and Silver, P.A. 2006. A new biobrick assembly strategy designed for facile protein engineering. <http://hdl.handle.net/1721.1/32535>. In *DSPACE*. Massachusetts Institute of Technology Artificial Intelligence Laboratory, MIT Synthetic Biology Working Group.
- Polach, K.J. and Widom, J. 1996. A model for the cooperative binding of eukaryotic regulatory proteins to nucleosomal target sites. *J. Mol. Biol.* **258**: 800–812.
- Posfai, G., Plunkett III, G., Feher, T., Frisch, D., Keil, G.M., Umenhoffer, K., Kolisnychenko, V., Stahl, B., Sharma, S.S., de Arruda, M., et al. 2006. Emergent properties of reduced-genome *Escherichia coli*. *Science* **312**: 1044–1046.
- Prehoda, K.E., Scott, J.A., Mullins, R.D., and Lim, W.A. 2000. Integration of multiple signals through cooperative regulation of the N-WASP–Arp2/3 complex. *Science* **290**: 801–806.
- Price, L.A., Kajkowski, E.M., Hadcock, J.R., Ozenberger, B.A., and Pausch, M.H. 1995. Functional coupling of a mammalian somatostatin receptor to the yeast pheromone response pathway. *Mol. Cell. Biol.* **15**: 6188–6195.
- Ptashne, M. 2004. *A genetic switch: Phage lambda revisited*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Raser, J.M. and O’Shea, E.K. 2004. Control of stochasticity in eukaryotic gene expression. *Science* **304**: 1811–1814.
- Raser, J.M. and O’Shea, E.K. 2005. Noise in gene expression: Origins, consequences, and control. *Science* **309**: 2010–2013.
- Ro, D.K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J., et al. 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**: 940–943.
- Rohatgi, R., Ma, L., Miki, H., Lopez, M., Kirchhausen, T., Takenawa, T., and Kirschner, M.W. 1999. The interaction between N-WASP and the Arp2/3 complex links Cdc42-dependent signals to actin assembly. *Cell* **97**: 221–231.
- 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.
- Roth, R.J. and Acton, N. 1989. A simple conversion of artemisinic acid into artemisinin. *J. Nat. Prod.* **52**: 1183–1185.
- Schmid, G. and Hofheinz, W. 1983. Total synthesis of qinghaosu. *J. Am. Chem. Soc.* **105**: 624–625.
- Seet, B.T., Dikic, I., Zhou, M.M., and Pawson, T. 2006. Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Biol.* **7**: 473–483.
- Shirasu, N., Shimohigashi, Y., Tominaga, Y., and Shimohigashi, M. 2003. Molecular cogs of the insect circadian clock. *Zool. Sci.* **20**: 947–955.
- Sismour, A.M. and Benner, S.A. 2005. Synthetic biology. *Expert Opin. Biol. Ther.* **5**: 1409–1414.
- Smith, H.O., Hutchison III, C.A., Pfannkoch, C., and Venter, J.C. 2003. Generating a synthetic genome by whole genome assembly:  $\phi$ X174 bacteriophage from synthetic oligonucleotides. *Proc. Natl. Acad. Sci.* **100**: 15440–15445.
- Sprinzak, D. and Elowitz, M.B. 2005. Reconstruction of genetic circuits. *Nature* **438**: 443–448.
- Stemmer, W.P., Cramer, A., Ha, K.D., Brennan, T.M., and Heyneker, H.L. 1995. Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* **164**: 49–53.
- Szostak, J.W., Bartel, D.P., and Luisi, P.L. 2001. Synthesizing life. *Nature* **409**: 387–390.
- Szybalski, W. and Skalka, A. 1978. Nobel prizes and restriction enzymes. *Gene* **4**: 181–182.
- Tian, J., Gong, H., Sheng, N., Zhou, X., Gulari, E., Gao, X., and Church, G. 2004. Accurate multiplex gene synthesis from programmable DNA microchips. *Nature* **432**: 1050–1054.
- Tucker, J.B. and Zilinskas, R.A. 2006. The promise and perils of synthetic biology. *New Atlantis* **12**: 25–45.
- Tumpey, T.M., Basler, C.F., Aguilar, P.V., Zeng, H., Solorzano, A., Swayne, D.E., Cox, N.J., Katz, J.M., Taubenberger, J.K., Palese, P., et al. 2005. Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science* **310**: 77–80.
- Vashee, S., Melcher, K., Ding, W.V., Johnston, S.A., and Kodadek, T. 1998. Evidence for two modes of cooperative DNA binding in vivo that do not involve direct protein–protein interactions. *Curr. Biol.* **8**: 452–458.
- Wall, M.E., Hlavacek, W.S., and Savageau, M.A. 2003. Design principles for regulator gene expression in a repressible gene circuit. *J. Mol. Biol.* **332**: 861–876.
- Wall, M.E., Hlavacek, W.S., and Savageau, M.A. 2004. Design of gene circuits: Lessons from bacteria. *Nat. Rev. Genet.* **5**: 34–42.
- Wang, J., Ellwood, K., Lehman, A., Carey, M.F., and She, Z.S. 1999. A mathematical model for synergistic eukaryotic gene activation. *J. Mol. Biol.* **286**: 315–325.
- Wang, L., Xie, J., and Schultz, P.G. 2006. Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct.* **35**: 225–249.
- Ye, X., Al-Babili, S., Klott, A., Zhang, J., Lucca, P., Beyer, P., and Potrykus, I. 2000. Engineering the provitamin A ( $\beta$ -carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* **287**: 303–305.