Synthetic Biology: Designing and Building Biological Systems
Desmond S. Lun, Broad Institute

Introduction to Synthetic Biology

Synthetic biology is extreme genetic engineering: instead of tweaking a single gene or a single promoter, synthetic biologists are thinking about designing a whole system bottom up. Synthetic biology raises ethical concerns, for valid reasons, which will be discussed later.

Analogous to computer engineering, synthetic biology is also hierarchical. Whereas in computer engineering a physical layer of transistors is coupled to make logic gates, which are coupled to make modules, which are joined to make the computers, which are joined to make networks, in synthetic biology proteins and genes are coupled to enable biochemical reactions, which are coupled to pathways, which are put together to make cells, which make up a tissue, and tissues make up an organism. Now breathe after reading the previous sentence. There is one difference: computers were originally built by people, whereas the bio systems were designed by evolution and we have discovered them already working.

Examples of Synthetic Biology

We will next inspect the recent literature to understand what synthetic biology is all about.

This is a pioneering synthetic biology project in which an oscillator was introduced to E. Coli. Elowitz and Leibler introduced three transcriptional repressor systems that are not part of any natural biological clock to build an oscillating network, as each gene suppresses the next. The network periodically induces the synthesis of green fluorescent protein as a readout of its state in individual cells. The oscillating fluorescence is pretty cool but not very useful. This artificial clock displays noisy behavior, possibly because of stochastic fluctuations of its components. Building such networks can improve our understanding of naturally occurring networks.


This is another example for a cool yet non-useful work. The authors redesign the genome of bacteriophage T7, a phage that infects E. coli. They take an existing organism and rewrite it so that it works the same way but in a way more appealing to us, trying to improve it bottom up, without the help of evolution. The authors divided the T7 genome to segments and rewrote the first two segments. Any overlapping genes were separated, and restriction sites were introduced for later excision of fragments of interest. The resulting chimeric genome encodes a viable phage that appears to maintain key features of the original while being simpler to model and easier to manipulate, as seen from comparing the wild type to the engineered T7 transfection into E. coli (see plot on right). However, somewhat smaller plaques are formed in this same transfection experiment (figure B on the right).

While the previous two examples engineered DNA, this work engineers proteins. Dueber et al. design a synthetic switch gated by heterologous ligands using N-WASP, an actin regulatory protein. Normally, N-WASP contains an output region ("VCA" domain) that in isolation is constitutively active and stimulates actin polymerization by binding and activating the actin-related protein complex (ARP). Two modular domains, a basic (B) motif and a GTPase-binding domain (GBD) repress activity through autoinhibitory interactions. Two activating stimuli, PIP2 and Cdc42, bind the B and GBD motifs, respectively, and disrupt autoinhibition. Because the two inputs act cooperatively, N-WASP approximates an AND gate. In this pretty boring paper the authors design a strategy for a synthetic single-input switch using N-WASP’s output domain and a PDZ domain-ligand pair as heterologous autoinhibitory module. Likewise they constructed many other dual-input switches by changing the separation of the two domains using this synthetic protein system.


This is another paper that is not useful but provides an example of a complex function engineered by man. Levskaya et al. designed a bacterial system that is switched between different states by red light. The system consists of a synthetic sensor kinase that allows a lawn of bacteria to function as a biological film, such that the projection of a pattern of red light on to the E. coli produces a picture. Cool – use coliform that grows in one’s gut to produce pictures.

5. Ro et al., Production of the antimalarial drug precursor artemisiminic acid in engineered yeast, Nature 440, 940-943 (13 April 2006)

This is the first of several examples of useful application of synthetic biology. Artemisinin is an anti-malaria drug produced from the sweet wormwood plant but is in short supply and unaffordable to most malaria sufferers. Ro et al. engineer S.
cerevisiae to produce high titres of artemisinic acid using an engineered mevalonate pathway, amorphadiene synthase, and a novel cytochrome P450 monooxygenase (CYP71AV1) from sweet wormwood that performs a three-step oxidation of amorpha-4,11-diene to artemisinic acid. The synthesized artemisinic acid is transported out and retained on the outside of the engineered yeast, meaning that a simple and inexpensive purification process can be used to obtain the desired product. This is cool and the idea of recombinant therapy has been successfully implemented in medicine for more than a decade, for example recombinant insulin is grown in E. coli to treat diabetics, and recombinant factor VIII is widely used to treat hemophiliacs. This example is considered synthetic biology and not traditional recombinant DNA technology because of the introduction of a whole new pathway to yeast, instead of a single gene.

6. Ethanol, schmethanol The Economist (27 September 2007)
Microbes can also be engineered to produce biofuels. There is a large effort to get bacteria to cheaply and efficiently produce ethanol from biomass such as cellulose. Unfortunately, this isn’t ideal because ethanol is not the best fuel source. The alternative is to find a way to get bacteria to produce hydrocarbons instead, i.e. petroleum or so-called “biocrude”. If successful, this would produce very pure fuels that recycle the carbon already present in the environment, rather than digging into sequestered carbon sources and having to refine them.

**Design and fabrication are the key components of synthetic biology**

Synthetic biology is an engineering problem more than a science problem. As with any engineering discipline, there are two components required for synthetic biology: design and fabrication.

![Synthetic Biology Diagram](source: Heinemann and Panke, Bioinformatics, 2006)
**Systems Fabrication**

Two fundamental tools required for synthetic biology are DNA sequencing and synthesis. Similar to Moore’s law for the number of transistors per chip, the number of bases able to be sequenced and synthesized is increasing exponentially. The associated costs of reading and writing DNA are also decreasing exponentially. Advances in synthesis lag slightly behind sequencing, but there are many companies worldwide that sell custom DNA. One can imagine a time in the near future when these technologies are good enough and cheap enough to no longer be a limiting factor in biological engineering projects.

The advances in these technologies only provide biologists with the raw materials for synthetic biology. Viable, robust methods and tools for synthetic biology are still needed for the field to advance.

Standardization is necessary for any engineering discipline because simplifies construction of systems. In synthetic biology, BioBricks is a first attempt at standardizing biological parts for use in building more complex biological devices and systems (the Registry of Standard Biological Parts, http://parts.mit.edu). These BioBrick parts are all flanked by the same set of four restriction sites. By carefully selecting these restriction sites, the construction scheme illustrated here allows a simple way to combine parts together in a fairly reliable fashion. Note that although the ends created by S and X are compatible, when they are ligated they do not recreate either restriction site. Using this construction method, one could imagine a catalog of parts from which an engineer could pick components and assemble to create their own biological devices to perform certain, possibly novel, functions.

Source: http://parts.mit.edu/registry/index.php/Assembly:Standard assembly
Providing engineering chassis
Synthetic biology is attractive because it provides researchers with a new and easier way to work with molecular biology. However, although there are now tools like the Registry being developed, there is still the problem of relying on existing organisms to implement these new biological devices. Because of this, there is no guarantee that when a new set of genes are inserted into *E. coli*, they will behave in the predicted way. Thus synthetic biology needs a platform on which all these new parts can be built.

Two of the groups trying to solve this problem are Frederick Blattner at the University of Wisconsin-Madison and George Church at Harvard University. Blattner’s approach is to take an existing *E. coli* strain and try to strip it down until all that remains is what is necessary for the cell to live. Church’s group is trying to build a “minimal cell” from the ground up, by only including genes that are necessary for survival. If these or other groups are successful, then they will have provided a chassis for BioBrick parts or other custom genes to be inserted with predictable behavior.

Systems Design
With the fabrication side of synthetic biology steadily progressing, there can be a shift in focus towards systems design. That is, engineers need to decide how to actually use these new tools to build useful things. One place to start is to take design principles from other fields, such as communication, and apply them to biology.

The *E. coli* metabolism can be thought of as a very large, complex network where each node is a metabolite or some biochemical intermediate, and each link between nodes is a biochemical reaction. These networks are very complex. *E. coli* has 1039 metabolites that participate in 2381 reactions. Modeling this entire network is impractical, but small subnetworks can be analyzed. Once a subnetwork is understood, one can begin to think about what changes can be made to the network. For example, links between nodes can be removed, or new links can be added (genes that regulate the reactions). Also, promoters can be tuned to adjust the expression levels of certain genes without completely knocking them out. This sort of analysis can be done with flux-balance metabolic modeling.

In the figure, we assume that v8 is Biomass – this is what your organism wants to maximize. You can then model your desired output as v9, assume that you can remove and add reactions (connections between nodes). You can then solve this based model to predict the output of the system. Maximization of your synthetic objective then
reduces to a bi-level optimization problem. Where you have control of the possible connections and the organism has control of how much flow enters each node. You have to work towards optimizing your desired output while keeping in mind that the organism will respond to try to optimize its desired output. This is analogous to a game theory problem where you have one objective, and your opponent (the cell) has another.

One can convert this bi-level optimization problem to a single level optimization problem. This is done by translating the organism's optimization problem into a set of constraints which when fulfilled maximize v8. This can then be fed into a Mixed integer linear program, which is a standard solver for these types of problems. This single constraint problem however is not linearly solvable. This becomes a problem when you consider the magnitude of the metabolic networks in biological systems. Currently we can only solve moderately sized problems (~100 nodes). This is done in the Broad Institute using heuristic methods, written in python and GLPK, these programs search over local domain of the problem to try finding the local optimal solutions, like gradient descent. They have currently applied this program to find solutions for biofuel production in E.coli. The problem with this, as with other heuristic methods, is there is no guarantee for performance: the algorithm is non-deterministic.

Ethical Concerns – Biosecurity and Biosafety

There is a large amount of concern with the potential risk involved in synthetic biology. This concern is mostly justified. Along with the large potential of synthetic biology comes a large amount of risk. We need to be careful when dealing with tough ethical situations as they can impair scientific progress. This becomes a bigger concern as the price of DNA synthesis continues to decrease. For example, viral DNA is very short yet has the ability to inflict a lot of damage. While most people will use synthetic biology responsibly, people can use this technology for nefarious purposes. As a proof of concept, a group of researchers synthesized the polio virus purely from knowledge of sequence. They started by synthesizing cDNA from the sequence information, they then moved this DNA to a coated virus, replacing its DNA. To test the efficacy of this virus, they infected mice. As expected, the mice died from this reconstructed virus (Cello, et al, 2002). Another viral reconstruction project is the reconstruction of the 1918 Spanish Flu with the purpose of making a vaccine for the present day flu (Tumpey et. al., 2005). Given that the reconstruction of such deadly pathogens is already a reality with current technologies, we must be careful about the technology we have and consider the potential for abuse of each new technology. We must also begin to consider methods of regulating these technologies, so that we can take advantage of their potential while at the same time safeguard ourselves from their abuse.
Further Reading


Elowitz and Leibler, A synthetic oscillatory network of transcriptional regulators, Nature 403, 335-338 (20 January 2000)


Ro et al., Production of the antimalarial drug precursor artemisinic acid in engineered yeast, Nature 440, 940-943 (13 April 2006)

Ethanol, schmethanol The Economist (27 September 2007)