Synthetic biologists may soon design and build engineered biological systems.

The Promise of Synthetic Biology

Jay Keasling is director of the Berkeley Center for Synthetic Biology, University of California, Berkeley, and Lawrence Berkeley National Laboratory.

Jay Keasling

It has been estimated that for every successful drug compound, 5,000 to 10,000 compounds must be introduced into the drug-discovery pipeline. On average, it takes $802 million and 10 to 15 years to develop a successful drug. Given this very low success rate and the incredibly high costs, drug companies must introduce as many drug candidates into their pipelines as possible.

Natural products have been important sources of drug leads; as much as 60 percent of successful drugs are of natural origin (Cragg et al., 1997), and some of the most potent natural products have been used as anticancer, antibacterial, and antifungal drugs. However, most natural products evolved for purposes other than the treatment of human disease. Thus, even though they can sometimes function as human therapeutics, their pharmacological properties may not be optimal. Furthermore, many are produced in miniscule amounts in their native hosts, thus making them expensive to harvest.

Organic chemistry methodologies are widely used to synthesize pharmaceuticals (of natural origin or not) and functionalize pharmaceutically relevant natural products. With appropriate protection and deprotection steps, chiral centers and functionalities can be introduced into molecules with precision. With the advent of combinatorial chemical synthesis, researchers have been able to construct entire families of molecules substituted at several positions with several different substituents, thus allowing drug companies to fill drug-discovery pipelines with variations of promising leads.

Despite the creation of complicated molecules made possible by advances in organic synthesis methodologies, the performance of these molecules is hardly comparable to the ease, specificity, and "green-ness" of enzymes. Indeed, many organic synthesis routes now incorporate one or more enzymes to perform transformations that are particularly difficult using nonenzymatic routes. Furthermore, enzymes are now being used for the *in vitro*, combinatorial functionalization of complex molecules. The next logical step in the synthesis of chemotherapeutics is the use of enzymes for combinatorial synthesis inside the cell, which would allow the production of drug candidates from inexpensive starting materials and avoid the need for purification of enzymes, which may be necessary for *in vitro* synthesis.

Biological Engineering for the Synthesis of Drugs

Rich, versatile biological systems are ideally suited to solving some of the world's most significant challenges, such as converting cheap, renewable resources into energy-rich molecules; producing high-quality, inexpensive drugs to fight disease; detecting and destroying chemical or biological agents; and remediating polluted sites. Over the years, significant strides have been made in engineering microorganisms to produce ethanol, bulk chemicals, and valuable drugs from inexpensive starting materials; to detect and degrade nerve agents as well as less toxic organic pollutants; and to accumulate metals and reduce radionuclides.

However, meeting these biological engineering challenges requires long development times, largely because of a lack of useful tools that would enable engineers to easily and predictably reprogram existing systems, let alone build new enzymes, signal transduction pathways, genetic circuits, and, eventually, whole cells. The ready availability of these tools would drastically alter the biotechnology industry, leading to less expensive pharmaceuticals, renewable energy, and biological solutions to problems that do not currently offer sufficient monetary returns to justify the high cost of biological research.

Most of the biological engineering tools currently available to scientists and engineers have not changed significantly since genetic engineering began in the 1970s. Biologists still use natural, gene-expression control systems (promoters with cognate repressors/activators). The ability to place a single heterologous gene under the control of one of these native promoters and produce large quantities of a protein of interest is the basis for the modern biotechnology industry.

Although redesigned biological control systems have been generally effective for their intended purposes (controlling rather roughly the expression of a single gene or a few genes), not surprisingly they are often inadequate for more complicated engineering tasks (e.g., controlling very large, heterologous, metabolic pathways or signal transduction systems). In addition, these borrowed "biological parts" retain many of the features that were beneficial in their native forms but make them difficult to use for purposes other than the ones for which they evolved. Well characterized standard biological parts, and larger devices made from such parts, would make biological engineering more predictable and enable the construction and integration of larger systems than is currently possible.

No standards have been defined for building larger biological devices.

In almost every other field of engineering, standards have been developed for building large integrated systems by assembling components from various manufacturers. However, biologists and engineers have not yet defined standards for the parts that might allow them to build larger biological devices. The design and construction of new devices (e.g., genetic-control systems) would benefit greatly from standards governing how various parts (e.g., regulatory proteins, promoters, ribosome binding sites) should interact and be assembled. Setting standards would also encourage manufacturing firms to develop parts.

Biological engineering has been held back because many of the most effective biological parts (promoters, genes, plasmids, etc.) have been patented and are available only to companies that can afford the royalty payments. This has not only increased the cost of drug development, but also hampered the development of new biological solutions to problems that may not have significant monetary payoffs (basically, anything other than drug development). Open-source biological parts, devices, and eventually whole cells would reduce the cost of engineering biological systems, make biological engineering more predictable, and encourage the development of novel biological solutions to some of

our most challenging problems. The development of open-source biological technology would improve awareness of, and minimize possible future biological risks, in the same way that open-source software tends to promote a constructive and responsive community of users and developers.

Synthetic Biology

Synthetic biology is the design and construction of new biological entities, such as enzymes, genetic circuits, and cells, or the redesign of existing biological systems. The goal of synthetic biology, which builds on advances in molecular, cellular, and systems biology, is to transform biology in the same way that synthesis transformed chemistry and integrated circuit design transformed computing. The element that distinguishes synthetic biology from traditional molecular and cellular biology is the focus on (1) the design and construction of core components (parts of enzymes, genetic circuits, metabolic pathways, etc.) that can be modeled, understood, and engineered to meet specific performance criteria, and (2) the assembly of these smaller parts and devices into larger integrated systems to solve specific problems. Just as engineers now design integrated circuits based on the known physical properties of materials and then fabricate functioning circuits and entire processors (with relatively high reliability), synthetic biologists will soon design and build engineered biological systems.

Unlike many other areas of engineering, however, biology is nonlinear and less predictable, and much less is known about parts and how they interact. Hence, the overwhelming physical details of natural biology (gene sequences, protein properties, biological systems) must be organized and recast via a set of design rules that hide information and manage complexity, thereby enabling the engineering of multicomponent integrated biological systems. Only when this is accomplished will designs of significant scale be possible.

Synthetic biology arose from four different intellectual premises. The first is the scientific idea that a practical test of understanding is the ability to reconstitute a functional system from its basic parts. Using synthetic biology, scientists are testing models of how biology works by building systems based on models and measuring differences between expectations and observations. Second, some consider biology an extension of chemistry, and thus synthetic biology can be considered an extension of synthetic chemistry. Attempts to manipulate living systems at the molecular level will likely lead to a better understanding, and new types, of biological components and systems. Third, natural living systems evolved to ensure their continued existence; they are not optimized for human understanding and intention. By thoughtfully redesigning natural living systems, it is possible simultaneously to test our current understanding and potentially implement engineered systems that are easier to interact with and study. Fourth, biology can be used as a technology, and biotechnology, broadly redefined, includes the engineering of integrated biological systems for the purposes of processing information, producing energy, manufacturing chemicals, and fabricating materials.

Although the emergence of the discipline of synthetic biology was motivated by these agendas, progress has only been practical since the recent advent of two foundational technologies, DNA sequencing, which has increased our understanding of the components and organization of natural biological systems, and synthesis, which has enabled us to begin to test the designs of (1) new, synthetic biological parts (Allert et al., 2004; Basu et al., 2004; Becskei and Serrano, 2000; Cane et al., 2002; Datsenko and Wanner, 2000; De Luca and Laflamme, 2001; Dwyer and Hellinga, 2004; Gardner and Collins, 2000; Gardner et al., 2000; Geerlings et al., 2001; Gerasimenko et al., 2002; Godfrin-Estevenon et al., 2002; Guet et al., 2002; Kobayashi et al., 2004; McDaniel et al., 1997) and (2) new biological systems (Bignell and Thomas, 2001; Blake and Isaacs, 2004; Hughes and Shanks, 2002; Iijima et al., 2004; Irmler et al., 2000; Judd et al., 2000; Kumar et al., 2004; Le Borgne et al., 2001; Martin et al., 2001, 2002, 2003; Okamoto et al., 2004). Each of these examples demonstrates the incredible potential of synthetic biology, as well as the foundational scientific and engineering challenges that must be met for the engineering of biology to become routine.

References

Allert, M., S.S. Rizk, L.L. Looger, and H.W. Hellinga. 2004. Computational design of receptors for an organophosphate surrogate of the nerve agent soman. Proceedings of the National Academy of Sciences 101(21): 7907–7912.

Basu, S., R. Mehreja, S. Thiberge, M.-T. Chen, and R. Weiss. 2004. Spatiotemporal control of gene expression with pulse-generating networks. Proceedings of the National Academy of Sciences 101(17): 6355–6360.

Becskei, A., and L. Serrano. 2000. Engineering stability in gene networks by autoregulation. Nature 405(6786): 590–593.

Bignell, C., and C.M. Thomas. 2001. The bacterial ParA-ParB partitioning proteins. Journal of Biotechnology 91(1): 1–34.

Blake, W.J., and F.J. Isaacs. 2004. Synthetic biology evolves. Trends in Biotechnology 22(7): 321–324.

Cane, D.E., F. Kudo, K. Kinoshita, and C. Khosla. 2002. Precursor-directed biosynthesis: biochemical basis of the remarkable selectivity of the erythromycin polyketide synthase toward unsaturated triketides. Chemistry and Biology 9(1): 131–142.

Cragg, G.M., D.J. Newman, and K.M. Snader. 1997. Natural products in drug discovery and development. Journal of Natural Products 60(1): 52–60.

Datsenko, K.A., and B.L. Wanner. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of Sciences 97(12): 6640–6645.

De Luca, V., and P. Laflamme. 2001. The expanding universe of alkaloid biosynthesis. Current Opinion in Plant Biology 4(3): 225–233.

Dwyer, M.A., and H.W. Hellinga. 2004. Periplasmic binding proteins: a versatile superfamily for protein engineering. Current Opinion in Structural Biology 14(4): 495–504.

Gardner, T.S., C.R. Cantor, and J.J. Collins. 2000. Construction of a genetic toggle switch in *Escherichia coli*. Nature 403(6767): 339–342.

Gardner, T.S., and J.J. Collins. 2000. Neutralizing noise in gene networks. Nature 405(6786): 520–521.

Geerlings, A., F.J. Redondo, A. Contin, J. Memelink, R. van der Heijden, and R. Verpoorte. 2001. Biotransformation of tryptamine and secologanin into plant terpenoid indole alkaloids by transgenic yeast. Applied Microbiology Biotechnology 56(3-4): 420–424.

Gerasimenko, I., Y. Sheludko, X. Ma, and J. Stockigt. 2002. Heterologous expression of a Rauvolfia cDNA encoding strictosidine glucosidase, a biosynthetic key to over 2000 monoterpenoid indole alkaloids. European Journal of Biochemistry 269(8): 2204–2213.

Godfrin-Estevenon, A.M., F. Pasta, and D. Lane. 2002. The parAB gene products of *Pseudomonas putida* exhibit partition activity in both *P. putida* and *Escherichia coli*. Molecular Microbiology 43(1): 39–49.

Guet, C.C., M.B. Elowitz, W. Hsing, and S. Leibler. 2002. Combinatorial synthesis of genetic networks. Science 296(5572): 1466–1470.

Hughes, E.H., and J.V. Shanks. 2002. Metabolic engineering of plants for alkaloid production. Metabolic Engineering 4(1): 41–48.

Iijima, Y., D.R. Gang, E. Fridman, E. Lewinsohn, and E. Pichersky. 2004. Characterization of geraniol synthase from the peltate glands of sweet basil. Plant Physiology 134(1): 370–379.

Irmler, S., G. Schroder, B. St. Pierre, N.P. Crouch, M. Hotze, J. Schmidt, D. Strack, U. Matern, and J. Schroder. 2000. Indole alkaloid biosynthesis in Catharanthus roseus: new enzyme activities and identification of cytochrome P450 CYP72A1 as secologanin synthase. The Plant Journal 24(6): 797–804.

Judd, E.M., M.T. Laub, and H.H. McAdams. 2000. Toggles and oscillators: new genetic circuit designs. Bioessays 22(6): 507–509.

Kobayashi, H., M. Kaern, M. Araki, K. Chung, T.S. Gardner, C.R. Cantor, and J.J. Collins. 2004. Programmable cells: interfacing natural and engineered gene networks. Proceedings of the National Academy of Sciences 101(22): 8414–8419.

Kumar, P., C. Khosla, and Y. Tang. 2004. Manipulation and analysis of polyketide synthases. Methods in Enzymology 388: 269–293.

Le Borgne, S., B. Palmeros, F. Bolivar, and G. Gosset. 2001. Improvement of the pBRINT-Ts plasmid family to obtain marker-free chromosomal insertion of cloned DNA in *E. coli*. Biotechniques 30(2): 252–254, 256.

Martin, V.J.J., C.D. Smolke, and J.D. Keasling. 2002. Redesigning cells for production of complex organic molecules. ASM News 68: 336–343.

Martin, V.J.J., D.J. Pitera, S.T. Withers, J.D. Newman, and J.D. Keasling. 2003. Engineering the mevalonate pathway in *Escherichia coli* for production of terpenoids. Nature Biotechnology 21(7): 796–802.

Martin, V.J.J., Y. Yoshikuni, and J.D. Keasling. 2001. The in vivo synthesis of plant sesquiterpenes in *Escherichia coli*. Biotechnology and Bioengineering 75(5): 497–503.

McDaniel, R., C.M. Kao, S.J. Hwang, and C. Khosla. 1997. Engineered intermodular and intramodular polyketide synthase fusions. Chemistry and Biology 4(9): 667–674.

Okamoto, A., K. Tanaka, and I. Saito. 2004. DNA logic gates. Journal of the American Chemical Society 126(30): 9458–9463.