Manufacturing Molecules Through Metabolic Engineering

Jay D. Keasling

Science **330**, 1355 (2010);
DOI: 10.1126/science.1193990

REVIEW

Manufacturing Molecules Through Metabolic Engineering

Jay D. Keasling[1,2,3]

Metabolic engineering has the potential to produce from simple, readily available, inexpensive starting materials a large number of chemicals that are currently derived from nonrenewable resources or limited natural resources. Microbial production of natural products has been achieved by transferring product-specific enzymes or entire metabolic pathways from rare or genetically intractable organisms to those that can be readily engineered, and production of unnatural specialty chemicals, bulk chemicals, and fuels has been enabled by combining enzymes or pathways from different hosts into a single microorganism and by engineering enzymes to have new function. Whereas existing production routes use well-known, safe, industrial microorganisms, future production schemes may include designer cells that are tailor-made for the desired chemical and production process. In any future, metabolic engineering will soon rival and potentially eclipse synthetic organic chemistry.

The term "metabolic engineering" was coined in the late 1980s–early 1990s (*1*). Since that time, the range of chemicals that can be produced has expanded substantially, in part due to notable advances in fields adjacent to metabolic engineering: DNA sequencing efforts have revealed new metabolic reactions and variants of enzymes from many different organisms; extensive databases of gene expression, metabolic reactions, and enzyme structures allow one to query for desired reactions and design or evolve novel enzymes for reactions that do not exist; new genetic tools enable more precise control over metabolic pathways; new analytical tools enable the metabolic engineer to track RNA, protein, and metabolites in a cell to identify pathway bottlenecks; and detailed models of biology aid in the design of enzymes and metabolic pathways. Yet even with these substantial developments, microbial catalysts are not as malleable as those in synthetic organic chemistry, and metabolic engineers must weigh many trade-offs in the development of microbial catalysts: (i) cost and availability of starting materials (e.g., carbon substrates); (ii) metabolic route and corresponding genes encoding the enzymes in the pathway to produce the desired product; (iii) most appropriate microbial host; (iv) robust and responsive genetic control system for the desired pathways and chosen host; (v) methods for debugging and debottlenecking the constructed pathway; and (vi) ways to maximize yields, titers, and productivities (Fig. 1). Unfortunately, these design decisions cannot be made independently of each other: Genes cannot be expressed, nor will the resulting enzymes function, in every host; products or metabolic intermediates may be toxic to one host but not another host; different hosts have different levels of sophistication of genetic tools available; and processing conditions (e.g., growth, production, product separation and purification) are not compatible with all hosts. Even with these many challenges, metabolic engineering has been successful for many applications, and with continued developments more applications will be possible.

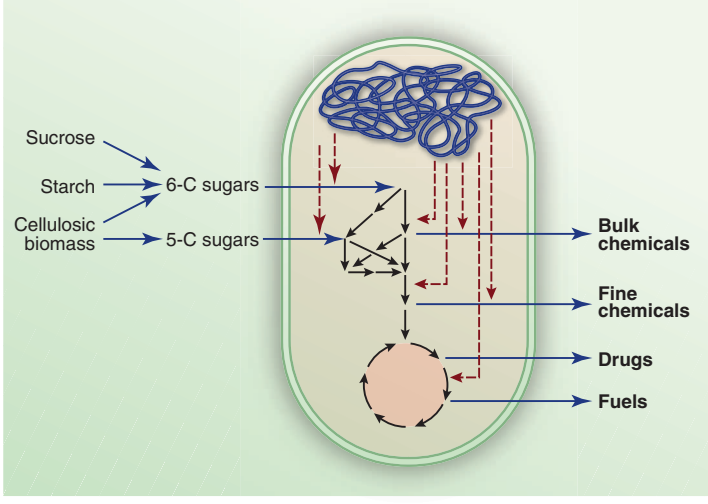

Fig. 1. Conversion of sugars to chemicals by means of microbial catalysts.

Starting Materials, Products, and Metabolic Routes

One area where metabolic engineering has a sizable advantage over synthetic organic chemistry is in the production of natural products, particularly active pharmaceutical ingredients (APIs), some of which are too complex to be chemically synthesized and yet have a value that justifies the cost of developing a genetically engineered microorganism. The cost of starting materials is generally a small fraction of their cost, and relatively little starting material is necessary so availability is not an issue. Most APIs fall into three classes of natural products, and many of the biosynthetic pathways for their precursors have been reconstituted in heterologous hosts.

Alkaloids are nitrogen-containing, low molecular weight compounds found primarily in and derived from plants and widely used as drugs. Two recent studies conclude that the large group of benzyl isoquinoline alkaloids (BIAs) will one day be producible in *Escherichia coli* and *Saccharomyces cerevisiae* (*2*). Unfortunately, the BIAs are only one of four major alkaloid groups, all of which are produced through different pathways. As the metabolic pathways for other alkaloids are discovered in their natural producers, many more of these valuable molecules could be produced microbially.

Polyketides and nonribosomal peptides (NRPs) have found broad use as APIs, veterinary agents, and agrochemicals. Naturally occurring polyketides and NRPs are produced by a number of bacteria and fungi using large, modular enzymes. Their titers and yields in the native producers have been improved through traditional strain engineering and advanced metabolic engineering. More recently, some of the most valuable molecules have been produced with engineered industrial hosts (*3*). Recombination of various synthase modules allows one to produce a nearly infinite range of chemicals (*4*, *5*), opening up the possibility that they may one day be used to produce fine and bulk chemicals.

Isoprenoids have found use as fragrances and essential oils, nutraceuticals, and pharmaceuticals. Many isoprenoids have been produced microbially, including carotenoids and various plant-derived terpenes (*6*–*8*), taking advantage of terpene synthases to form the most complicated part of the molecules and hydroxylases to introduce hydroxyl group that can be subsequently functionalized chemically or biologically (*7*, *9*). Isoprenoids are one of the few classes of natural products where there are alternative precursor production pathways. An example of using metabolic engineering and synthetic chemistry together to produce an API is the semisynthesis of the antimalarial drug artemisinin with *S. cerevisiae* engineered to produce artemisinic acid, the most complex part of the molecule, and synthetic chemistry to produce artemisinin from the microbially sourced

[1]Joint BioEnergy Institute, 5885 Hollis Street, Emeryville, CA 94608, USA. [2]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [3]Synthetic Biology Engineering Research Center, Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA. E-mail: keasling@berkeley.edu

artemisinic acid (*6*, *7*, *9*). Beyond producing natural products, laboratory evolution or rational engineering of terpene cyclases, terpene hydroxylases, and a host of other terpene-functionalizing enzymes (*8*, *10*–*12*) and combinatorial expression of these evolved enzymes in a heterologous host will enable the production of unnatural terpenes, some of which might be more effective than the natural product for the treatment of human disease.

Although individual metabolic pathways have been developed to produce natural products derived from a single pathway, there is an opportunity to synthesize multisubstituent APIs (e.g., Taxol) or other molecules from the products of multiple biosynthetic pathways. This will require simultaneous expression of multiple precursor pathways in a single microorganism, as well as "ligases" that can assemble multiple substituents together into a single molecule. The benefit would be the synthesis of complicated molecules that might not otherwise be produced.

Although not as valuable as pharmaceuticals, many fine chemicals have been produced economically from natural and engineered microorganisms, including amino acids, organic acids, vitamins, flavors, fragrances, and nutraceuticals. For fine chemicals, profit margins are generally much lower than for APIs and may be affected by substrate availability and cost. Some of these molecules are sufficiently complicated that they cannot be produced economically by any route other than biological production, whereas others have chemical routes. For some important products (fragrances, flavors, amino acids), heterologous hosts have been engineered to enhance their production. Yet we have barely begun to investigate what will be possible to produce.

In contrast, bulk chemicals such as solvents and polymer precursors are rarely produced from microorganisms, because they can be produced inexpensively from petroleum by chemical catalysis. Due to fluctuations in petroleum prices and recognition of dwindling reserves, trade imbalances, and political considerations, it is now possible to consider production of these inexpensive chemicals from low-cost starting materials such as starch, sucrose, or cellulosic biomass (e.g., agricultural and forest waste, dedicated energy crops, etc.) with a microbial catalyst. For example, 1,3-propanediol (1,3-PDE), a useful intermediate in the synthesis of polyurethanes and polyesters, is now being produced from glucose by *E. coli* engineered with genes from *Klebsiella pneumoniae* and *S. cerevisiae* (*13*). There is an opportunity to produce many other bulk chemicals (e.g., polymer precursors) by using metabolically engineered cells, but the key will be to produce the exact molecule needed for existing products rather than something "similar but green" that will require extensive product testing before it can be used.

By far the highest-volume (and lowest-margin) application for engineered metabolism is the production of transportation fuels. For many of the same reasons that it is desirable to produce petroleum-derived chemicals using biological systems, it is desirable to produce transportation fuels from readily available, inexpensive, renewable sources of carbon. There is a long history of using microorganisms to produce alcohols, primarily ethanol and butanol. Although much of the work on these alcohols was done by traditional strain mutagenesis and selection, more recent work focused on engineering yeasts and bacteria to produce ethanol or butanol from a variety of sugars while eliminating routes to side products and improving the tolerance of the host to the alcohol (*14*). Larger, branched-chain alcohols can be produced by way of the Ehrlich pathway. By incorporating broad substrate-range 2-keto acid decarboxylases and alcohol dehydrogenases, several microbes have now been engineered to produce these fuels (*15*, *16*). These alcohols are generally considered better fuels than ethanol and butanol and can also be used to produce a variety of commodity chemicals.

Recent advances in metabolic pathway and protein engineering have made it possible to engineer microorganisms to produce hydrocarbons with properties similar or identical to those of petroleum-derived fuels and thus compatible with our existing transportation infrastructure. Linear hydrocarbons (alkanes, alkenes, and esters) typical of diesel and jet fuel can be produced by way of the fatty acid biosynthetic pathway (*17*–*19*). For diesel in cold weather and jet fuel at high altitudes, branches in the chain are beneficial—regularly branched and cyclic hydrocarbons of different sizes with diverse structural and chemical properties can be produced via the isoprenoid biosynthetic pathway (*20*, *21*). Both the fatty acid–derived and the isoprenoid-derived fuels diffuse (or are pumped) out of the engineered cells and phase separate in the fermentation, making purification simple and reducing fuel cost.

Although the pathways described above produce a wide range of fuel-like molecules, there are many other molecules that one might want to produce, such as short, highly branched hydrocarbons (e.g., 2,2,4-trimethyl pentane or isooctane) that would be excellent substitutes for petroleum-derived gasoline. Additionally, most petroleum fuels are mixtures of large numbers of components that together create the many important properties of the fuels. It should be possible to engineer single microbes or microbial consortia to produce a mixture of fuels from one of the biosynthetic pathways or from multiple biosynthetic pathways. Indeed, some enzymes produce mixtures of products from a single precursor—maybe these enzymes could be tuned to produce a fuel mixture ideal for a particular engine type or climate.

To make these new fuels economically viable, we must tap into inexpensive carbon sources (namely, sugars from cellulosic biomass). Given the variety of sugars in cellulosic biomass, the fuel producer must be able to consume both five- and six-carbon sugars. Because many yeasts do not consume five-carbon sugars, recent developments in engineering yeast to catabolize these sugars will make production of these fuels more economically viable (*22*). Engineering fuel-producing microorganisms to secrete cellulases and hemicellulases to depolymerize these sugar polymers into sugars before uptake and conversion into fuels has the potential to substantially reduce the cost of producing the fuel.

Hosts and Expression Systems

From the applications cited above, it should be evident that the product, starting materials, and production process all affect host choice. Some of the most important qualities one must consider when choosing a host are whether the desired metabolic pathway exists or can be reconstituted in that host; if the host can survive (and thrive) under the desired process conditions (e.g., ambient versus extremes of temperature, pH, ionic strength, etc.); if the host is genetically stable (both with the introduced pathway and not susceptible to phage attack); and if good genetic tools are available to

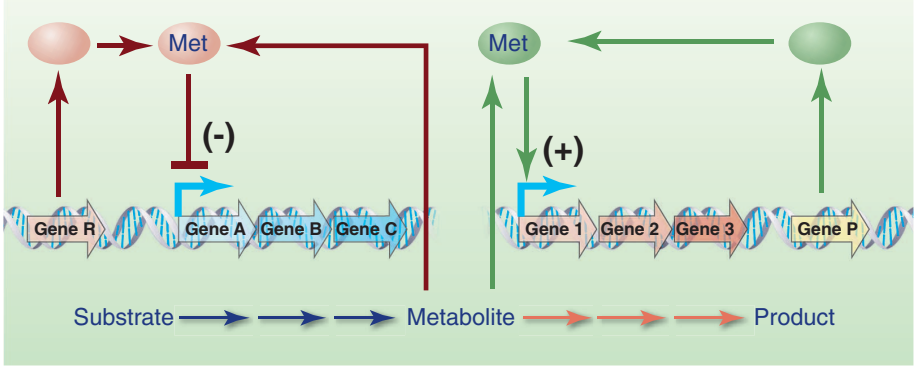

Fig. 2. Use of synthetic regulators to modulate metabolic pathways that have a toxic intermediate. Regulatory proteins or RNAs bind the toxic metabolite and down-regulate the biosynthetic pathway and up-regulate the consumption pathway.

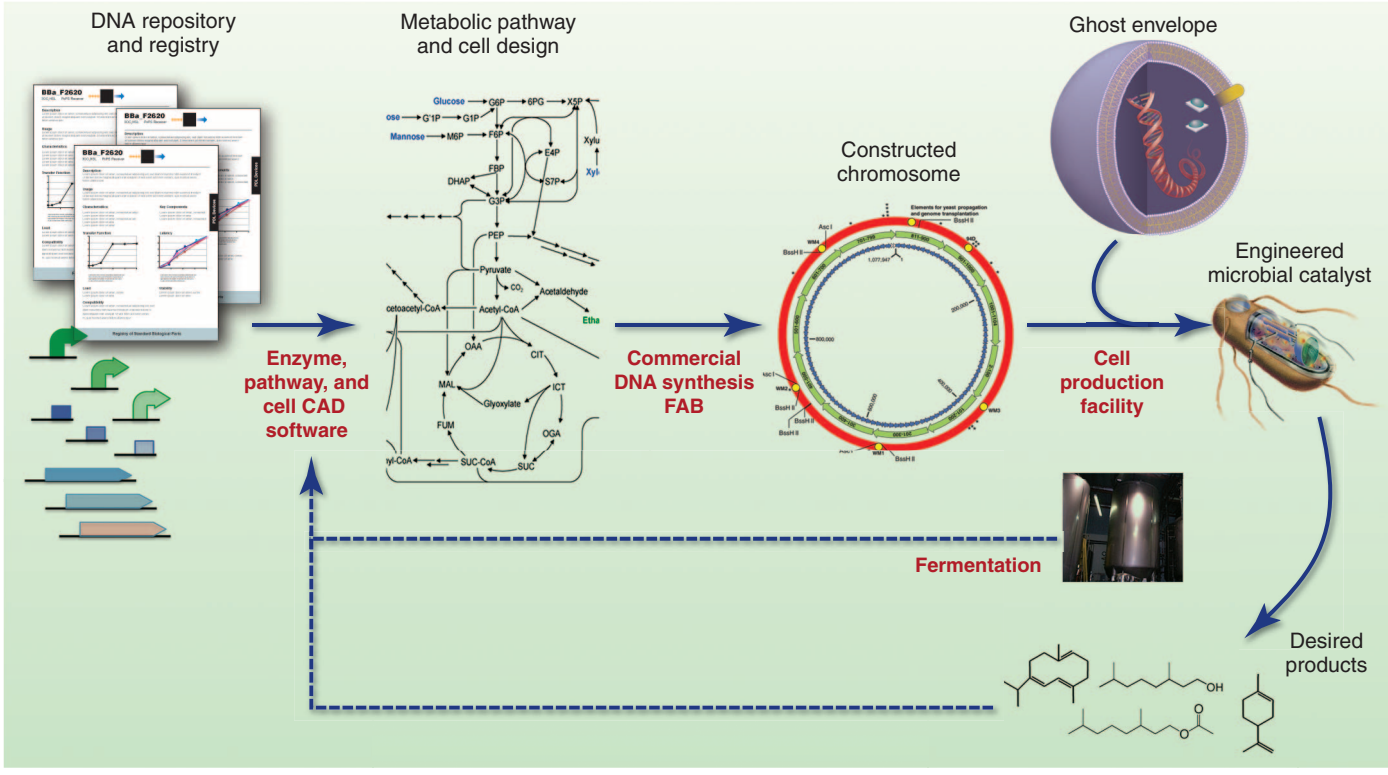

Fig. 3. The future of engineered biocatalysts. Pathways, enzymes, and genetic controls are designed from characteristics of parts (enzymes, promoters, etc.) by means of pathway and enzyme CAD software. The chromosomes encoding those elements are synthesized at a FAB and incorporated into a ghost envelope to obtain the new catalyst. The design of the engineered catalyst is influenced by the desired product and the production process.

manipulate the host. Widely used, heterologous hosts include *E. coli*, *S. cerevisiae*, *Bacillus subtilis*, and *Streptomyces coelicolor*, to name a few. Although *E. coli* and *S. cerevisiae* excel in the genetic tools available, *E. coli* has the disadvantage of being susceptible to phage attack. And while *B. subtilis* and *S. coelicolor* have the ability to easily express polyketide synthases, they have fewer genetic tools available than either *S. cerevisiae* or *E. coli*. Although minimal, bacterial hosts may have scientific interest (*23*), minimal hosts that require addition of many nutrients or that cannot cope with stresses in processing will probably not find a niche in industrial chemical or fuel production where cost is critical. Thus, it is essential to have genetic tools for existing industrial hosts that can grow on simple, inexpensive carbon sources and salts or on an inexpensive, undefined medium with minimal additions (*24*, *25*).

The key issue necessitating good genetic tools is the introduction of foreign genes encoding the metabolic pathway and control over their expression to maximize yields and titers. The genes encoding the transformational enzymes in metabolically engineered cells do not need to be highly expressed, but must be produced in catalytic amounts sufficient to adequately transform the metabolic intermediates into the desired products at a sufficient rate. Expression of the desired genes at too high a level will rob the cell of metabolites that might otherwise be used to produce the desired molecule of interest, particularly important for production of low-margin chemicals, while underexpressed genes will create pathway bottlenecks. Furthermore, because intermediates of a foreign metabolic pathway can be toxic to the heterologous host (*6*), which results in decreased production of the desired final compound, it is essential that the relative levels of the enzymes be coordinated.

Central to any genetic manipulation is the vector used to carry and/or harbor the transforming DNA in the host. Important features of the cloning vector include segregational stability, minimal and consistent copy number in all cells of a culture, and the ability to replicate and express large sequences of DNA. There is growing recognition that one or only a few copies of a gene are needed, particularly for metabolic engineering applications. With the ability to vary promoter (*26*) and ribosome binding strength (*27*), as well as the stabilities of the mRNA (*28*) and the resulting protein produced from it, there are many factors other than copy number that can be manipulated to alter enzyme production.

Promoters play an essential role in controlling biosynthetic pathways. Inducible promoters are one of the easiest and most effective ways to regulate gene expression, but it is essential that the promoter be induced consistently in all cells of a culture (*29*). Constitutive promoters (*26*) and promoters that respond to a change in growth condition or to an important intermediary metabolite (*30*) allow for inexpensive, inducer-free gene expression, which is particularly important where cost is an issue (Fig. 2). Although there are many inducible promoters for bacteria, the small number of inducible promoters for yeast and other potential industrial hosts makes regulation of metabolic pathways in those organisms more challenging than in bacteria.

Because production of complicated molecules often requires several enzymes, it is desirable to coordinate expression of the genes encoding these enzymes to prevent accumulation of toxic intermediates and bottlenecks in biosynthetic pathways. There are many ways to coordinate expression of multiple genes, such as using a nonnative RNA polymerase or transcription factor to induce multiple promoters (*31*); grouping multiple, related genes into operons; varying the ribosome binding strength for the enzymes encoded in the operon (*27*); or controlling segmental mRNA stability of each coding region to regulate the amount of each enzyme produced (*32*). One of the limitations to expressing multiple genes in yeast is the lack of internal ribosomal entry sequences (IRESs) that are available for higher eukaryotes. The development of yeast IRESs would allow one to express genes encoding metabolic

pathways without the need for a promoter for each gene.

Debottlenecking, Debugging, and Process Optimizing

Even with a kit full of tools, building a biosynthetic pathway is made difficult without accurate blueprints. In almost all areas of engineering, there are models and simulation tools that allow one to predict which components to assemble to obtain a larger system with a desired function or characteristic. Similar biological design tools are in their infancy. However, metabolic models that incorporate cell composition and gene regulation have become relatively predictive and may one day be used to design metabolic pathways and predict the level of gene expression needed to achieve a particular flux through a reaction or pathway (*33*).

Regardless of how sophisticated the design tools and how good the blueprint, there will always be "bugs" in the engineered system. Analogous to software debuggers that allow one to find and fix errors in computer code, the development of similar tools for biological debugging would reduce development times for optimizing engineered cells. For the development of microbial chemical factories, functional genomics can serve in the role of debugging routines (*34*), because imbalances in a metabolic pathway often elicit a stress response in central metabolism (due to protein overproduction or accumulation of toxic intermediates or end products) (*6*, *35*). Information from one or more of these techniques can be used to diagnose the problem and modify expression of genes in the metabolic pathway or in the host to improve titer and/or productivity.

Many desirable chemicals will be toxic to the producer, particularly at the high titers needed for industrial-scale production. Taking advantage of the cell's native stress response pathways can be an effective way to alleviate part or all of the toxicity (*36*). Even better, transporters could be used to pump the desired product outside the cell, reducing intracellular toxicity and purifying the product from the thousands of contaminating intracellular metabolites (*37*).

Designer Cells for Designer Chemicals

One can envision a future when a microorganism is tailor-made for production of a specific chemical from a specific starting material, much like chemical engineers build refineries and other chemical factories from unit operations (Fig. 3). The chemical and physical characteristics of the product and starting materials would be considered in the design of the organism to minimize both production and purification costs (e.g., operating the engineered cell at the boiling point of volatile, toxic products to drive production and reduce product toxicity). The cell envelope would be designed to be resistant to the specific desired chemical, and the cell wall would be designed to make the organism tolerant to industrial processing conditions. Specific, engineered transporters would be incorporated into the membrane to pump the desired product out of the cell and keep it out and to import the desired starting material. The biosynthetic pathway would be constructed from a parts registry containing all known enzymes by means of retrosynthesis software (*38*), and done so to maximize yield and minimize the time required to grow the organism and produce the desired chemical from the desired starting material. In the event that an enzyme does not exist for a particular reaction or set of reactions, one would use computer-aided design (CAD) software to design the desired enzyme (*39*).

Once the cell has been designed in the computer, the genetic control system would be designed to control expression of all the genes at the correct time and at the appropriate levels. Redundancies in the genetic control system would be engineered to ensure that design parameters are maintained regardless of the transient changes the cell encounters during the production process. Simulations and scenario planning would test various designs, including genetic control system failure. Safety for the plant operators and the environment would be an essential design criterion. When the genetic controls were fully designed and tested, the chromosome(s) would be designed and constructed. Gene location, modularity, and ease of construction are but a few of the important considerations in designing the chromosome. The chromosome would be ordered from a commercial DNA manufacturer. Depending on the state of the technology at the time, the chromosome would arrive in pieces and be assembled in the constructed envelope or would be completely assembled at the factory and sent to another location to be introduced into the ghost cell. One can even envision a day when cell manufacturing is done by different companies, each specializing in certain aspects of the synthesis—one company constructs the chromosome, one company builds the membrane and cell wall (the "bag"), one company fills the bag with the basic molecules needed to boot up the cell.

Until this future arrives, manufacturing of molecules will be done with well-known, safe, industrial microorganisms that have tractable genetic systems. Continued development of tools for existing, safe, industrial hosts, cloning and expressing genes encoding precursor production pathways, and the creation of novel enzymes that catalyze unnatural reactions will be necessary to expand the range of products that can be produced from biological systems. When more of these tools are available, metabolic engineering should be just as powerful as synthetic chemistry, and together the two disciplines can greatly expand the number of products available from renewable resources.

References and Notes

1. J. E. Bailey, *Science* **252**, 1668 (1991).
2. K. M. Hawkins, C. D. Smolke, *Nat. Chem. Biol.* **4**, 564 (2008).
3. B. A. Pfeifer, S. J. Admiraal, H. Gramajo, D. E. Cane, C. Khosla, *Science* **291**, 1790 (2001).
4. H. G. Menzella *et al.*, *Nat. Biotechnol.* **23**, 1171 (2005).
5. V. Siewers, R. San-Bento, J. Nielsen, *Biotechnol. Bioeng.* **106**, 841 (2010).
6. V. J. J. Martin, D. J. Pitera, S. T. Withers, J. D. Newman, J. D. Keasling, *Nat. Biotechnol.* **21**, 796 (2003).
7. D. K. Ro *et al.*, *Nature* **440**, 940 (2006).
8. E. Leonard *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13654 (2010).
9. M. C. Chang, R. A. Eachus, W. Trieu, D. K. Ro, J. D. Keasling, *Nat. Chem. Biol.* **3**, 274 (2007).
10. Y. Yoshikuni, T. E. Ferrin, J. D. Keasling, *Nature* **440**, 1078 (2006).
11. C. Schmidt-Dannert, D. Umeno, F. H. Arnold, *Nat. Biotechnol.* **18**, 750 (2000).
12. J. A. Dietrich *et al.*, *ACS Chem. Biol.* **4**, 261 (2009).
13. C. E. Nakamura, G. M. Whited, *Curr. Opin. Biotechnol.* **14**, 454 (2003).
14. E. J. Steen *et al.*, *Microb. Cell Fact.* **7**, 36 (2008).
15. G. K. Donaldson, A. C. Eliot, D. Flint, L. A. Maggio-Hall, V. Nagarajan, U.S. Patent 20070092957 (2007).
16. S. Atsumi, T. Hanai, J. C. Liao, *Nature* **451**, 86 (2008).
17. E. J. Steen *et al.*, *Nature* **463**, 559 (2010).
18. H. R. Beller, E.-B. Goh, J. D. Keasling, *Appl. Environ. Microbiol.* **76**, 1212 (2010).
19. A. Schirmer, M. A. Rude, X. Li, E. Popova, S. B. del Cardayre, *Science* **329**, 559 (2010).
20. S. T. Withers, S. S. Gottlieb, B. Lieu, J. D. Newman, J. D. Keasling, *Appl. Environ. Microbiol.* **73**, 6277 (2007).
21. N. S. Renninger, D. J. McPhee, World Patent 200804555 (2008).
22. H. W. Wisselink, M. J. Toirkens, Q. Wu, J. T. Pronk, A. J. van Maris, *Appl. Environ. Microbiol.* **75**, 907 (2009).
23. D. G. Gibson *et al.*, *Science* **329**, 52 (2010).
24. H. H. Wang *et al.*, *Nature* **460**, 894 (2009).
25. G. Pósfai *et al.*, *Science* **312**, 1044 (2006).
26. P. R. Jensen, K. Hammer, *Appl. Environ. Microbiol.* **64**, 82 (1998).
27. H. M. Salis, E. A. Mirsky, C. A. Voigt, *Nat. Biotechnol.* **27**, 946 (2009).
28. C. D. Smolke, T. A. Carrier, J. D. Keasling, *Appl. Environ. Microbiol.* **66**, 5399 (2000).
29. A. Khlebnikov, K. A. Datsenko, T. Skaug, B. L. Wanner, J. D. Keasling, *Microbiology* **147**, 3241 (2001).
30. W. R. Farmer, J. C. Liao, *Nat. Biotechnol.* **18**, 533 (2000).
31. H. Alper, G. Stephanopoulos, *Metab. Eng.* **9**, 258 (2007).
32. B. F. Pfleger, D. J. Pitera, C. D. Smolke, J. D. Keasling, *Nat. Biotechnol.* **24**, 1027 (2006).
33. J. S. Edwards, R. U. Ibarra, B. O. Palsson, *Nat. Biotechnol.* **19**, 125 (2001).
34. J. H. Park, K. H. Lee, T. Y. Kim, S. Y. Lee, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7797 (2007).
35. L. Kizer, D. J. Pitera, B. F. Pfleger, J. D. Keasling, *Appl. Environ. Microbiol.* **74**, 3229 (2008).
36. H. Alper, J. Moxley, E. Nevoigt, G. R. Fink, G. Stephanopoulos, *Science* **314**, 1565 (2006).
37. M. J. Dunlop, J. D. Keasling, A. Mukhopadhyay, *Syst. Synth. Biol.* **4**, 95 (2010).
38. K. L. Prather, C. H. Martin, *Curr. Opin. Biotechnol.* **19**, 468 (2008).
39. J. B. Siegel *et al.*, *Science* **329**, 309 (2010).

40. This work was supported in part by the Synthetic Biology Engineering Research Center, which is funded by National Science Foundation Award No. 0540879 and by the Joint BioEnergy Institute, which is funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231. Competing financial interests: The author is a founder of and owns equity in Amyris and LS9.

10.1126/science.1193990