Metabolic engineering: enabling technology for biofuels production

Mitchell Tai and Gregory N. Stephanopoulos

The biofuels industry is rapidly growing with the goal of providing a sustainable fuel alternative to petroleum products. However, two challenges persist in limiting current biofuels technologies as a viable alternative: (1) the cost of feedstock and (2) the quality of the fuel. Metabolic engineering shows the potential to overcome these two challenges, enabling the bioconversion of cheaper feedstocks into improved fuels in a sustainable and environmentally friendly manner. Metabolic engineering is the systematic attempt to understand, design, and engineer cellular metabolic networks using a wide range of interdisciplinary tools and strategies. It is rapidly emerging as the enabling technology behind the development of the next generation of biofuels. Several examples of current research demonstrate how metabolic engineering has begun to contribute both novel and creative solutions toward bringing promising biofuel technologies to fruition: production of higher chain alcohols, fermentation of lignocellulosic material, and production of fatty acid derivatives. © 2012 John Wiley & Sons, Ltd.

How to cite this article: WIREs Energy Environ 2012. doi: 10.1002/wene.5

ENGINEERING THE FUTURE OF BIOFUELS

The past few years have introduced a flurry of interest over renewable energy sources. Biofuels have attracted attention as renewable alternatives to liquid transportation fuels. There are numerous potential advantages over fossil fuels: sustainable supply, diversification of energy sources, energy independence and security, rural development, and reduction in greenhouse emissions.1 However, achieving adequate scale requires a tremendous effort in research and development beyond what has thus far been achieved. The field of metabolic engineering is well suited to develop the future technologies that will give us widespread, cost-effective, and sustainable transportation fuels.

Metabolic engineering is the improvement of cellular activities by manipulation of metabolic networks through the use of recombinant deoxyribonucleic acid technology.2 Interdisciplinary advances in metabolic engineering have yielded powerful strategies and methods to understand and manipulate whole metabolic pathways with confidence.3,4 To date, numerous efforts have successfully engineered and optimized metabolic networks to produce high-value targets for use in the pharmaceutical and fine chemicals industries.5 However, attention is now being turned toward commodity-scale processes, which require both cost-efficiency and robustness.1

Currently, the most prevalent biofuels are ethanol produced from corn or sugarcane and biodiesel produced from vegetable oils. Under current production processes, however, neither biofuel is economically competitive or well integrable into existing petroleum-based technologies and infrastructure.6 Two developmental challenges underpin these shortcomings: (1) the need for a better feedstock and (2) the need for a better fuel. However, these challenges also represent key opportunities to develop the next generation of biofuel technologies. A central element in these technologies will be the use of metabolic engineering to develop the biological platforms that produce these biofuels.

Engineering for Improved Feedstocks

For the past few years, production of ethanol from corn and biodiesel from vegetable oils has been increasing rapidly. Last year, the United States
production capacity of corn ethanol exceeded 13 billion gallons per year (bgy), approaching 10% of the national gasoline demand. Meanwhile, global biodiesel production is approaching 5.0 bgy, with a majority coming from Europe.

However, production of these biofuels from plants like corn or rapeseed also competes for arable cropland needed for food. This adds undesirable price sensitivities between biofuels and food and has already shown adverse effects on food prices. Transforming forests or existing cropland can also sometimes have the effect of increasing greenhouse gas emissions, counteracting the carbon emissions benefit of biofuels.

The primary cost for producing biofuels is the cost of the feedstock: 60% in the case of corn ethanol and 80% for soybean biodiesel. Even with gains in process yield, current crop-based feedstocks will still limit the overall profitability of biofuels. Currently, upward of half of the production cost of these biofuels needs to be supported by government subsidies.

The next generation of feedstocks will need to have lower land requirements and lower production cost, yet maintain high production capacity to bring biofuels closer to economic viability. Metabolic engineering allows us to bridge the feedstock gap by enabling the utilization of cheaper and more sustainable substrates by introducing catabolic pathways and optimizing metabolic networks for the conversion of feedstock to fuel. Indeed, yield optimization has been a critical aspect of virtually all biochemical engineering processes in recent history. Metabolic engineering of organisms toward this end only serves to continue this tradition, pushing yields beyond what is naturally observed. Furthermore, microbe-based biofuel production also reduces the cropland requirements compared to crop-based methods, decreasing competition with food production.

Engineering for Improved Fuels

Although new feedstocks are explored, a simultaneous search continues for the next generation of fuel types. Current biofuels have some persistent disadvantages that limit their incorporation into existing infrastructure.

Ethanol, although widely produced, has relatively poor fuel characteristics. Ethanol is hygroscopic, capable of absorbing water, which can lead to corrosion. The energy content is also low, containing only 70% of the energy per volume of gasoline. Also, as ethanol is produced by fermentation, the resulting beer is dilute, containing roughly 10% ethanol. Subsequent distillation to separate the ethanol is very energy intensive.

Biodiesel is a better fuel, but also has some disadvantages. It is not well suited for use at low temperatures because of a high cloud point, and still often requires large quantities of petroleum-derived methanol as part of its production. It also has only 89% of the energy content of its analog, petrodiesel.

Current biofuel characteristics limit their integration into existing infrastructure. Because of this, there is a high transition barrier to adoption of biofuels, and both ethanol and biodiesel are often blended only at low concentrations into conventional fuels.

Development of better fuels that have high energy density and can be integrated into existing pipelines and engines will be needed if biofuels are to be more widely adopted and have a reasonable hope to replace fossil fuels. Through metabolic engineering of production pathways, alternative products can be made that have characteristics closer to their petroleum equivalents, easing the barrier for adoption. These alternatives range from slight modifications to existing metabolites, to new pathways that create naturally unique compounds. These naturally rare products will also require extensive pathway engineering and optimization to achieve effective production capacities—one of the central strengths of metabolic engineering.

TOOLS OF METABOLIC ENGINEERING

To understand how metabolic engineering plays a role in biofuels development and how it takes an interdisciplinary approach to problem solving, it is important to first understand its main strategies and tools. The strategies of metabolic engineering can be compartmentalized into three steps: (1) understanding, (2) designing, and (3) engineering the metabolic network. Each of these steps uses tools and technologies adopted from a range of disciplines. An overview of the strategies of metabolic engineering can be found in Figure 1.

Understanding the Metabolic Network

The first step in metabolic engineering is to understand the complex network of enzymatic reactions that compose a cell’s metabolism. In addition to the enzymology of participating enzymes, this requires information on the structure and behavior of the pathways that connect these enzymes. Knowledge of the pathway chemistry and stoichiometry allows us to
FIGURE 1  Strategies of metabolic engineering revolve around the understanding, design, and engineering of metabolic networks and pathways to produce desired molecular products from biological platforms. These strategies employ techniques and technologies from a range of disciplines, from omics technology to synthetic biology.

calculate theoretical yields, which are often used as benchmarks for pathway engineering efficacy. Comprehensive systems-level data about these complex networks is acquired through omics technologies and bioinformatics. Omics technologies involve using genomic, transcriptomic, proteomic, or metabolomic data to quantify the system behavior of the cell along various functional axes (e.g., growth, tolerance, productivity). Bioinformatics is the method of extracting biological meaning by identifying significant patterns, motifs, and connections within these large, complex data sets. These techniques enable us to develop a systems-level perspective on cellular activity and an understanding of important contributing networks.

As an example, metabolic flux analysis derived from metabolomic data allows us to observe the flow of material through cellular metabolic pathways. Like a material balance, these fluxes describe the distribution of material throughout the cell’s metabolic network and can help identify branch points and competing pathways relevant to our desired product. Fluxes also help to determine the degree of engagement of various enzymes in the pathway, allowing us to identify rate-limiting steps and control points.

Because any biological manipulation will rarely ever produce only an isolated response, it is important to observe the system-level response of our engineering efforts. Using bioinformatics and omics technologies allows us to understand the interactions, connections, and responses between different parts of the system to predict and control the metabolic network.

Designing the Metabolic Network

Once we have sufficient understanding of the organism and its cellular activities, we are then able to develop and design specific strategies to obtain our desired product. Although we can introduce, remove, or otherwise modify pathways, identifying the most effective actions a priori can help save much time and effort. Modern methods to do so are found in the field of computational systems biology.

A main goal of computational systems biology is to reconstruct cellular networks in silico, which can model the behavior of the cell. Starting with a cellular model, one is able to simulate and characterize how possible pathway manipulations will affect the system overall. Evaluation of these changes can help identify the ideal genetic targets that will maximize our objectives.

One such method of evaluation is called elementary mode analysis, which uses a systems engineering approach to decompose metabolic networks into uniquely organized pathways that can be used to evaluate cellular phenotypes, metabolic network regulation, network robustness, and network fragility. As an extension, neural networks can also be used to make sense of exceptionally difficult systems and to subsequently predict future behavior.

Engineering the Metabolic Network

Once targets and pathways are identified, the next task is to implement these changes in vivo. This involves genetic manipulation of the host organism using molecular biology. The term synthetic biology describes the systematic approach to pathway manipulation through standardized biological components for the purpose of increasing their programmability and robustness. Under this framework, genetic elements are modularized to simplify the process of genetic engineering. These elements can then be used to introduce new genes, knockout existing genes, or modifying existing deoxyribonucleic acid sequences. Modules can be built up to produce whole pathways and can also be rearranged to optimize expression.

Numerous nonrational techniques are also available that extend the reach of traditional strain improvement: high throughput screening, directed evolution, gene shuffling, and combinatorial engineering. These techniques increase the efficiency of strain improvement by sampling a much larger phenotypic search space, opening up the possibility to select for
changes in less-intuitive or distal targets. An interesting example is the use of transcriptional engineering to achieve phenotypic diversity.16 This technique involves mutagenesis of transcription factors to produce global changes in the transcription and expression of genes in the cell, which would not be feasible through isolated point mutations.17

Finally, optimizing to maximize flux through the production pathway is often the most difficult engineering task. Tuning the expression of genes many times is necessary as intermediates and cofactors need to be balanced within the pathway. Furthermore, troubleshooting the pathway often involves alleviating any number of potential bottlenecks that may impede flux, such as competing pathways, lack of enzymatic driving force, cofactor imbalances, insufficient enzyme activity, unbalanced enzyme expression, transport issues, enzyme regulation, and toxicity.

Because of its interdisciplinary nature, metabolic engineering will be limited to the available tools and technologies of the fields from which it draws. However, as the state of the art progresses and development improves at the interface of these related disciplines, metabolic engineers will have increased power to evaluate, design, and manipulate cells to produce the desired products that will be in demand in the future. These abilities will most certainly be necessary for the development of the next generation of biofuels.

**METABOLIC ENGINEERING ENABLES BIOFUELS DEVELOPMENT**

In the following section, we will review various researches that attempt to address these biofuel challenges and how metabolic engineering is central to enabling these technologies. A schematic of relevant metabolic pathways for biofuels production can be found in Figure 2.

**Production of Higher Chain Alcohols**

Because ethanol has several undesirable fuel properties, higher chain alcohols have received attention as possible fuels. For example, n-butanol has 20% higher energy content than ethanol. In addition, it is more hydrophobic and thus less susceptible to inducing corrosion.18 Traditionally, butanol is produced through refining petroleum or from acetone-butanol-ethanol fermentation. The ability to robustly produce higher chain alcohols represents a step toward biofuels with characteristics approaching that of gasoline.

Utilizing metabolic engineering, pathways for the production of higher chain alcohols have been introduced into organisms such as *Escherichia coli*.18,19 The use of the model bacteria allows for the study and efficient development of novel pathways. Two pathways have recently emerged to successfully produce of C4 or C5 alcohols: coenzyme A (CoA)-mediated and nonfermentative pathways.

The CoA-mediated pathway involves utilizing the native pathway of the butanol-producing organism, *Clostridium acetobutylicum*. Introduction of five *Clostridium* genes into *E. coli* is sufficient for production of n-butanol from acetyl-CoA. However, initial titers were at best about 1 g/L, which is much lower than the 10 g/L typically produced by native *n*-butanol producers.19

More recently, extensive engineering of the pathway has led to insights into maximizing the production through this pathway. Three major bottlenecks were discovered that could be alleviated by (1) balancing the expression of upstream and downstream enzymes, (2) balancing cofactor utilization and generation, and (3) engineering driving force to increase flux toward the product.20,21 Indeed, these bottlenecks seem to be recurring obstacles in many efforts to engineer high production pathways. After addressing these bottlenecks, titers of 30 g/L could be achieved at about 70% theoretical maximum yield.20

The second pathway utilizes a creative nonfermentative approach, producing C4 and C5 branched chain alcohols from intermediates in the amino acid metabolic network.22 The introduction of a 2-ketoacid decarboxylase and an alcohol dehydrogenase allows the conversion of a variety of 2-keto acid metabolites found in amino acid synthesis pathways into their analogous branched chain alcohols: 1-propanol, isobutanol, *n*-butanol, 2-methyl-1-butanol, 3-methyl-1-butanol, and 2-phenylethanol. Branched chain alcohols have higher octane numbers than their straight-chained counterparts, making them better fuels. This process has the advantage of avoiding CoA-mediated chemistry, while also leveraging the wealth of understanding from decades of research on metabolic engineering for amino acid production. As such, they were able to obtain isobutanol production of over 20 g/L at 86% of the theoretical maximum yield.22 Optimization of cofactor imbalances produced strains that achieved 100% theoretical maximum yield, suggesting some of the same bottlenecks may exist in this alternative pathway.23

**Fermentation of Lignocellulosic Material**

In the search for improved feedstocks, the push toward cellulosic biofuels is a clear choice. Cellulosic biomass eliminates the need to compete with food
crop production as an estimated 1.3+ billion dry tons per year of biomass is potentially available in the United States. Two issues highlight how metabolic engineering can enable industrial utilization of this feedstock: xylose utilization and cellulose utilization.

Because hydrolysis of lignocellulosic biomass results in 20–30% carbohydrates in the form of xylose, utilization of pentose sugars is one of the first steps toward efficiently using cellulosic materials. *Saccharomyces cerevisiae*, the most productive of ethanologenic organisms, cannot ferment xylose; it lacks the ability to convert xylose into xylulose, although xylulose is metabolized within the pentose phosphate pathway (PPP). Transferring the xylose reductase (XR) and xylitol dehydrogenase (XDH) enzymes from *Schefersomyces stipitis* (formerly *Pichia stipitis*) enables the growth of yeast on xylose and production of ethanol.

However, growth and production are considerably slower than on glucose, and significant amounts of xylitol are often produced. Xylitol is the intermediate of the XR/XDH pathway, and most understand this to result from differences in cofactor specificity between reduced nicotinamide adenine dinucleotide phosphate (NADPH)-dependent XR and nicotinamide adenine dinucleotide (NAD)-dependent XDH. The cofactor imbalance has been addressed in two different ways: use of (1) xylose isomerase pathway or (2) protein mutagenesis to switch cofactor specificity.

Additional factors to the limited productivity are the lack of dedicated pentose transporters, low PPP flux, and inability for the cell to identify xylose as a fermentable sugar. However, more recently, progress has been made in these areas through additional metabolic engineering strategies: introducing heterologous xylose transporters, overexpressing...
PPP enzymes, engineering cofactor specificity, and evolutionary engineering (for a comprehensive review see Matsushika et al.).

Cellulose on the contrary is a polysaccharide composed of β(1→4) linked glucose molecules. Enzymatic digestion is most commonly used to break these chains down into free glucose molecules. However, this process is somewhat inefficient, requiring a separate enzymatic unit operation. To improve the efficiency of this process, recently, cellodextrin transporters from Neurospora crassa were introduced into S. cerevisiae allowing for utilization of cellobiose, cellotriose, and cellotetraose. Upon cellodextrin uptake, β-glucosidase breaks down cellodextrin into monomeric glucose, allowing for immediate catabolism. When coupled with enzymatic digestion, this process improves utilization efficiency and allows for the simultaneous saccharification and fermentation of cellulose.

Because of its preference for glucose, S. cerevisiae will natively repress the utilization of alternative substrates as well as down-regulate fermentative pathways after glucose is depleted. This results in inefficient and delayed utilization of alternative substrates like xylose, a phenomenon known as diauxic shift. However, the cellodextrin pathway seems to avoid this by eschewing the generation of extracellular glucose. As a result, strains harboring both cellodextrin and xylose utilization pathways remarkably exhibit the ability to simultaneously coferment glucose and xylose.

Metabolic engineering has shown how organisms can be adapted to utilize originally foreign substrates and feedstocks. It can also be used to circumvent regulatory bottlenecks to achieve more robust feedstock utilization. Although further regulatory and metabolic understanding is still necessary, current progress has shown great promise in enabling the production of biofuels from lignocellulosic materials.

Production of Fatty Acid Derivatives
The use of fatty acids and their derivatives as biofuels has numerous potential benefits. These long-chain hydrocarbon products have high energy density, are hydrophobic, and chemically resemble their fossil fuel analogs. Termed as drop-in biofuels, these products could be quickly integrated into existing infrastructure with little modification.

The most attractive pathway for production is the use of cytosolic thioesterase enzymes to produce free fatty acids, which can then be converted enzymatically or catalytically into a variety of derivative products: alkanes, alkenes, fatty acid esters, and alcohols. First discovered by Cho and Cronan, expression of cytosolic thioesterase ‘tesA converts the lipid synthesis intermediate acyl carrier protein (ACP) into its fatty acid form. Disrupting the fatty acid synthesis pathway in this manner circumvents many of the typical fatty acid products that are often used as triggers for feedback regulation. This engineered pathway remarkably decouples production from growth by relieving feedback inhibition of a tightly regulated pathway, providing driving force for fatty acid synthesis at all stages of the cell cycle. The fatty acids can then be modified into a variety of derivatives: free fatty acids, alkanes, and fatty acid ethyl esters (FAEE). The process has even been extended into microalgae, which promises biofuels production from photosynthesis. Additional engineering modifications to increase precursor availability included overexpression of acetyl-CoA carboxylase and coexpression of plant thioesterase. The best reported efforts produced 4.5 g/L/day of fatty acids through these engineering strategies.

In addition, this pathway has been used to advance the concept of a consolidated bioprocess (CBP). CBP entails having one organism, which can convert feedstock to product with little to no additional processing in between. This saves on operating costs and dramatically simplifies the production process.

Steen et al. report engineering an organism that can produce FAEE from xylan. As xylan is one of the primary constituents of hemicellulose, this process demonstrates the use of a single organism to consolidate initial enzymatic degradation of xylan into xylose, the conversion of xylose into fatty acids, and the transesterification of those fatty acids into ethyl esters. The basis of this work was the discovery and utilization of a wax ester synthase that is able to perform the biodiesel transesterification reaction enzymatically. Expression of xylanase and introduction of the ethanologenic pathway allowed for the in vivo transesterification of fatty acids to produce FAEE that can be used directly as biodiesel.

Robust control of the length and saturation of products will be a key achievement in the development of fatty acid derivatives production as a biofuel technology. Also, because fatty acid synthesis is tightly regulated and critical to many cellular functions, some creative approaches to metabolic engineering will be needed to increase yields and productivity without adversely affecting cell function. Nonetheless, a biofuel that chemically resembles petroleum fuels and can easily be integrated into existing technology has the potential to quickly become a relevant component of the biofuel industry.
CONCLUSION

A systems-level interdisciplinary approach is necessary for effective strategies to tackle today’s global energy and environmental problems. The tools and strategies of metabolic engineering are well suited for addressing the persistent challenges facing a successful transition away from petroleum transportation fuels. As such, metabolic engineering will be instrumental in developing the next generation of cost-effective and robust transportation fuels, which will come from cheaper, more sustainable feedstocks and have better fuel characteristics.

It remains to be seen what processes and technologies will successfully establish sustainable alternatives to petroleum transportation fuels. However, because of obstacles in current feedstocks and fuels, it is necessary to continue research in technologies that can overcome existing limitations. Metabolic engineering is uniquely poised to develop and implement the next-generation biofuels using a systems-level approach from multiple disciplines.

Furthermore, metabolic engineering allows us to explore unconventional strategies that are naturally uncommon: nonfermentative production of branched higher alcohols, cellodextrin transport for xylose and glucose cofermentation, and derivatization of fatty acid products. Radical pathway manipulation is also on the horizon: engineering of nitrogen flux for the conversion of protein into biofuels could further improve overall yields while recycling reduced nitrogen; design of artificial nonphotosynthetic carbon fixation pathways could open a new means of production processes without light (and thus land) as a limiting factor. Metabolic engineering enables both promising and exciting opportunities for alternative energy, which have the potential for great societal ramifications. It will also be central in the long road ahead to develop these opportunities into robust, efficient industrial-scale technologies.

REFERENCES


