Synthetic biology for microbial production of lipid-based biofuels
Leo d’Espaux¹, Daniel Mendez-Perez¹, Rachel Li¹,² and Jay D Keasling¹,²,³,⁴

The risks of maintaining current CO₂ emission trends have led to interest in producing biofuels using engineered microbes. Microbial biofuels reduce emissions because CO₂ produced by fuel combustion is offset by CO₂ captured by growing biomass, which is later used as feedstock for biofuel fermentation. Hydrocarbons found in petroleum fuels share striking similarity with biological lipids. Here we review synthetic metabolic pathways based on fatty acid and isoprenoid metabolism to produce alkanes and other molecules suitable as biofuels. We further discuss engineering strategies to optimize engineered biosynthetic routes, as well as the potential of synthetic biology for sustainable manufacturing.

Addresses
¹ Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, CA 94608, United States
² Department of Plant and Microbial Biology, University of California, Berkeley, CA 94270, United States
³ QB3 Institute, Department of Bioengineering, University of California, Berkeley, CA 94270, United States
⁴ QB3 Institute, Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94270, United States

Corresponding author: Keasling, Jay D (jdkeasling@lbl.gov)

Introduction
Over the last century, human use of fossil fuels has raised atmospheric CO₂ to levels 40% higher than at any other time in the 800 000-year record [1,2]. The rising CO₂ can be linked to global climate change, including more frequent extreme weather events and rising mean global temperatures. The already-observed temperature increase of 0.7 °C is projected to reach 2–8 °C by the end of the century, with potentially catastrophic consequences for our biosphere [1,3].

Of all anthropogenic CO₂ emissions, one quarter arise from the combustion of liquid transportation fuels [4]. Curbing these emissions will require a multi-faceted approach, including improved standards for vehicle fuel economy and emissions, alternative-powered vehicles, and biofuels. Biofuels reduce emissions because CO₂ produced by fuel combustion is offset by CO₂ captured by growing biomass, which is in turn used to produce more fuel (Figure 1). ‘Drop-in’ biofuels that can be used with existing vehicles — especially trucks and planes that are impractical to power using current fuel cell technology — are particularly desirable [5]. Several governments including those of the United States, China, and the European Union have instituted mandates for biofuels to constitute an increasing percentage of total transportation fuel usage in the coming years [6].

Liquid transportation fuels
Today, only 2% of all transportation fuel is bio-based. By far the most prevalent biofuel is ethanol produced by microbial fermentation of sugars, blended into gasoline as a volume booster and oxygenate [7]. However, ethanol can only be used as 10% of the blend due to its low energy density (and other factors). In the United States, most gasoline is already blended at this 10% ethanol limit.

Besides ethanol, the other predominant biofuel today is fatty acid alkyl ester (FAAE, marketed as biodiesel). FAAE is produced by thermochemical esterification of plant oils with alcohol — typically methanol or ethanol — and used as a diesel substitute. FAAE is chemically different from petroleum diesel, and from gasoline. Gasoline is primarily composed of linear and ringed C4–C9 hydrocarbons, diesel and jet fuel C8–21. FAAE, on the other hand, is composed of methyl-esters or ethyl-esters of linear C16–C22 alkyl chains (Figure 2). The increased oxygen content of FAAE leads to more complete fuel combustion, decreasing particulate and CO emissions [8]. However, oil crops used as feedstocks for FAAE production have low yields and divert agricultural resources from food crops.

The limitations of first-generation biofuels have generated interest in genetically engineering microorganisms to perform the bioconversion of an abundant and inexpensive feedstock into a biofuel. Here we discuss how various enzymes can be combined to biologically produce molecules suitable as transportation fuels, focusing on lipid-based replacements for diesel and jet fuels produced in the two most well-known microbial hosts, Escherichia coli
Carbon cycle for a microbial biofuel. Biofuels reduce emissions because CO₂ produced by fuel combustion is offset by CO₂ captured by growing biomass, which is in turn used to produce more fuel. With synthetic biology, it may be possible to produce fuel from various sources of carbon and energy. It may also be possible to produce fuels, or other molecules, with improved properties using the diverse bioconversions observed in living organisms.

and *Saccharomyces cerevisiae*. We also provide an overview of the tools and methodologies of synthetic biology for creating and optimizing biological designs, and outlooks on its potential for future biomanufacturing.

**Hydrocarbons and lipids**

The hydrocarbons we use as fuels today share striking similarity with some of the lipids most organisms use to store energy (Figure 2). Lipids are naturally energy-dense, and many exhibit other properties desirable in a biofuel. In fact, early demonstrations of the internal combustion (diesel) engine used peanut oil as fuel [9]. The triacylglycerides (TAGs) predominant in oils are produced through the fatty acid biosynthetic pathway. This nearly universally conserved pathway produces other major cellular components (such as phospholipids) as well as a great variety of other molecules, several which are structurally similar to fuels.

**Fatty acid metabolism**

Fatty acid metabolism begins with the carboxylation of acetyl-CoA to malonyl-CoA by acetyl-CoA carboxylase (ACC). Fatty acid synthase (FAS) then condenses one acetyl-CoA starter unit and several malonyl-CoA extender units iteratively to produce a linear acyl chain typically 12–22 carbons long, depending on the organism. Fatty acids are released as acyl thioesters, bound to either coenzyme A (CoA, in type I FAS) or to acyl carrier protein (ACP, in type II FAS). Most fungi and mammals employ type I FAS, most bacteria type II. Type II FAS is organized as discrete multifunctional polypeptides. By contrast type I FAS is organized as multimeric complexes of one or more polypeptides, each containing multiple enzyme activities.

After synthesis, acyl thioesters are mainly routed biologically toward membrane phospholipids (containing two acyl chains bound to a glycerol backbone) or energy storage TAGs (containing three). Most organisms can degrade TAGs or other fatty acids through β-oxidation. Some organisms perform additional bioconversions, such as consuming alkanes [10], or producing wax esters, polyhydroxalkanoates, fatty alcohols, or other compounds through variations of this versatile pathway.

**Engineered microbial production of fatty-acid derived biofuels**

Genes encoding enzymes that perform desired chemical conversions can be introduced into an easy-to-culture and genetically tractable microbial host, allowing the engineered strain to convert a simple feedstock — e.g., glucose — into a target molecule. A number of enzymes have been identified that catalyze the conversion of fatty acids or their intermediates into different products with good fuel properties.
Hydrocarbons and lipids. Compounds present in liquid transportation fuels are chemically similar to compounds produced through fatty acid biosynthesis. (a) Iso-octane is a major component of gasoline, hexadecane of diesel fuel, and FAAE of biodiesel. FAE is typically produced using methanol and producing fatty acid methyl ester (FAME). (b) TAGs are produced in many organisms as energy storage molecules, and phospholipids as the main structural components of cellular membranes. Fatty acid biosynthesis produces acyl thioester intermediates that give rise to phospholipids, TAGs, and other molecules. The acyl chains in these thioesters can be bound to acyl carrier protein (ACP) or CoA, depending on the organism. Generally bacteria employ acyl-ACPs and eukaryotes acyl-CoAs. (c) Lesser-known microorganisms contain intriguing lipids. Annamox bacteria produce fatty acids containing linearly concatenated cyclobutane rings termed ‘ladderanes’. Archaeal membranes are composed of phospholipids containing isoprene-chains linked through ether (rather than ester) linkages.

FAAE — molecularly identical to oil-crop biodiesel — has been produced by heterologously expressing a wax ester synthase (WS) catalyzing the esterification of an acyl thioester with ethanol [11,12]. Several WS enzymes have been shown to catalyze this reaction [13].

Alkanes and alkenes — the major constituents of petroleum diesel — have been produced through various bioengineered routes, such as the reduction of acyl-thioesters (or free fatty acids) into fatty aldehydes followed by decarbonylation [14–16]. Other routes include the decarboxylation of free fatty acids directly into α-alkenes by a bacterial cytochrome P450 [16], or polyketide synthase (PKS)-mediated extension-decarboxylation [17**].

Similarly, different pathways can be assembled to produce molecules not currently used as fuels, but with likely suitable properties, including fatty alcohols [12,18], methyl ketones [19,20], ω-hydroxy and dicarboxylic acids [21], and other fatty acid-derived products (Figure 3). It should be noted that different tailoring enzymes have different preferences for substrate chain length and terminal moiety, for example, acyl-ACP, acyl-CoA, or free fatty acids. These can mirror the nature of their host’s central FAS pathway — type I produces acyl-CoA, type II acyl-ACP. However, it is possible to introduce a type I FAS pathway into a host that natively employs type II FAS [22**], or vice versa [23]. Additionally, the FAS pathway can be modified to incorporate branched amino acids and produce branched-chain fatty acids [24]. Chain branching lowers freezing point, which is important for fuel performance in high altitudes or cold weather.

Isoprenoid metabolism
Branding is a defining feature of isoprenoids, another main class of lipids that contains molecules similar to those found in fuels. Also known as terpenoids, these molecules are defined by being formed from 5-carbon isoprene building blocks into thousands of molecules, encompassing 60% of known natural products [25]. Isoprenoids perform diverse functions in all kingdoms of life. In plants, for example, isoprenoids are important...
Biosynthetic routes for the production of natural and synthetic fuels from glucose. Fatty acid biosynthesis (pink) naturally produces phospholipids for membrane composition, and TAGs for energy storage. Isoprenoid biosynthesis (green) naturally produces sterols and other compounds. These pathways can be coopted using heterologous genes to produce a number of biofuel molecules. From acyl thiosteres: 1, esterification with ethanol by wax synthase [11] to produce FAEE; 2, reduction followed by decarboxylation [14–16], or PKS-mediated extension-decarboxylation [32] to produce alkenes; 3, reduction either directly [22**] or through fatty aldehydes intermediates [19] to produce fatty acids; 4, other routes to other products [12,19]; 5, heterologous FAS pathways [22**]; 6, monoterpen synthases can modify C10 geranyl-PP to produce pinene, limonene, or other monoterpenes [28,33]; 7, sesquiterpene synthases can modify C15 farnesyl-PP to form farnesene, bisabolene, or other sesquiterpenes [30,31]. Unsaturated lipids can be chemically hydrogenated for biofuel production (e.g., farnesene to farnesane). ACC, acetyl-CoA carboxylase; FAS, fatty acid synthase. The isoprenoid pathway shown is the mevalonate (MEV) pathway. Bacteria employ an alternative route, the DXP pathway, not shown for simplicity.

secondary metabolites. In archaea, they are part of primary metabolism and comprise the hydrophobic chains of cell membranes (Figure 2).

Isoprenoids are synthesized through one of two pathways: the 1-deoxy-D-xylulose 5-phosphate (DXP) pathway (native to most bacteria) or the mevalonate (MVA) pathway (native to most eukaryotes and archaea). The MVA pathway begins with three acetyl-CoA molecules, which combine to form mevalonate through six enzyme-catalyzed steps. The DXP pathway begins with pyruvate and glyceraldehyde-3-phosphate, which form DXP through seven steps. Both pathways end with the formation of the five-carbon building blocks isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). IPP and DMAPP are condensed and modified in various ways by various kinds of terpene synthases (TPSs) to form thousands of products. C10 and C15 isoprenoids, called respectively monoterpenes and sesquiterpenes, have appropriate carbon numbers for a liquid biofuel.
Engineered microbial production of isoprenoid derived biofuels

Monoterpenes (C10) are produced by the condensation of IPP and DMAPP into geranyl pyrophosphate (GPP) catalyzed by GPP synthase, and further modified by any number of monoterpane synthases. Synthetic pathways to monoterpenes such as α-pinene [26], sabine [27], limonene [28], and geraniol [29], have been successfully constructed in microbial hosts.

Sesquiterpenes (C15) are produced by the condensation of two IPP molecules and one DMAPP into farnesyl pyrophosphate (FPP) by FPP synthase, and subsequent modification by sesquiterpene synthases. Combining pathways to FPP with various sesquiterpene synthases in microbial hosts has produced bisabolene [30] and α-farnesene [31].

Metabolic engineering strategies for lipid-based biofuels

Having chosen a set of enzyme activities as a biosynthetic route to a biofuel, achieving good production levels is often challenging and time-consuming. Traditional metabolic engineering often employs a ‘pull-push-block’ approach. ‘Pulling’ on a pathway by overexpressing terminal enzymes or providing an irreversible sink — such as partitioning into a separate phase [12] — can create a thermodynamic driving force for product formation. ‘ Blocking’ consumption of products or intermediates can be achieved by deleting genes catalyzing undesirable reactions [12,34,35]. ‘Pushing’ involves overcoming bottlenecks that may form along the pathway. The problems, and solutions, are often specific to the pathway and host organism.

In fatty acid biosynthesis, for example, a common bottleneck is the carboxylation of acetyl-CoA into malonyl-CoA by ACC [36]. In Escherichia coli, ACC activity is inhibited by the product acyl-ACP. This inhibition can be minimized by converting acyl-ACP to other products that do not inhibit ACC — for example, by expressing a thioesterase [12]. In S. cerevisiae, ACC activity is inhibited post-translationally by phosphorylation in response to signals — such as glucose depletion. Mutating phosphorylation sites in ACC has resulted in increased titers of fatty acid products [37]. Another layer of regulation in yeast is transcriptional inhibition of ACC in response to elevated fatty acid levels. Replacing the native ACC promoter with one that is constitutively active has led to improved fatty product titers [34].

The availability of the central metabolite acetyl-CoA is important not only for fatty acid-derived products, but also isoprenoids produced through the mevalonate pathway, and many other targets. In S. cerevisiae, acetyl-CoA is provided to different subcellular pools through various biosynthetic routes [38]. The route to cytoplasmic acetyl-CoA is energetically draining and suppressed in the high-glucose conditions typical of laboratory and industrial cultivation (for reasons that are still debated [39,40]). Strategies to overcome this bottleneck have included expressing heterologous pathways for the provision of this central building block [41,42,43,44].

A number of other strategies (reviewed elsewhere [45–48]) have been used to improve production levels of many biofuel molecules (Table 1). Although many improvements in production levels have been realized, there is much room for further optimization to approach maximum theoretical yields.

Synthetic biology tools and methodologies

Synthetic biology today encompasses an increasing number of tools and methodologies to facilitate strain

---

**Table 1**

<table>
<thead>
<tr>
<th>Biofuel</th>
<th>Max theoretical yield (g/g glucose)</th>
<th>Host</th>
<th>Titer (g/L)</th>
<th>Yield (g/g glucose)</th>
<th>Percent of max theoretical yield</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFA (C16)</td>
<td>0.37</td>
<td>E. coli</td>
<td>5.2</td>
<td>0.26</td>
<td>70%</td>
<td>[49]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S. cerevisiae</td>
<td>2.2</td>
<td>0.11</td>
<td>30%</td>
<td>[35]</td>
</tr>
<tr>
<td>FAEE (C18)</td>
<td>0.36</td>
<td>E. coli</td>
<td>1.5</td>
<td>0.075</td>
<td>21%</td>
<td>[50]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S. cerevisiae</td>
<td>0.034</td>
<td>0.0017</td>
<td>&lt;1%</td>
<td>[51]</td>
</tr>
<tr>
<td>Fatty alcohol (C16)</td>
<td>0.34</td>
<td>E. coli</td>
<td>3.8b</td>
<td>0.13</td>
<td>38%</td>
<td>[22]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S. cerevisiae</td>
<td>0.33</td>
<td>0.017</td>
<td>5%</td>
<td>[18]</td>
</tr>
<tr>
<td>Alkanes (C15)</td>
<td>0.30</td>
<td>E. coli</td>
<td>0.08b</td>
<td>0.0027</td>
<td>&lt;1%</td>
<td>[14]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S. cerevisiae</td>
<td>0.0037</td>
<td>0.00019</td>
<td>&lt;1%</td>
<td>[52]</td>
</tr>
<tr>
<td>Bisabolene (C15)</td>
<td>0.27</td>
<td>E. coli</td>
<td>1.1c</td>
<td>0.11</td>
<td>41%</td>
<td>[53]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S. cerevisiae</td>
<td>1.0d</td>
<td>0.050</td>
<td>19%</td>
<td>[30]</td>
</tr>
</tbody>
</table>

Titers and yields of current laboratory-scale demonstrations of selected microbial biofuels produced in E. coli and S. cerevisiae. In general, experiments were performed using 2% glucose in shake flask fermentations.

a Maximum theoretical yields are calculated using an in silico optimization algorithm employing a whole genome-scale reconstruction [54,55].

b This example used 3% glucose rather than 2%.

c This example used 1% glucose rather than 2%.

d This example used 1.8% galactose and 0.2% glucose.
construction and optimization. Synthesizing, sequencing, and introducing DNA sequences into living cells [56**] is cheaper and easier than ever. Codon-optimization, directed evolution [57], screening enzyme libraries, and incorporating non-natural amino acids [58] all provide ways of improving or generating novel enzymatic activities.

Choosing the context in which to express pathway genes — levels, localization, and timing — can drastically affect production levels and growth rates of engineered strains. Heterologously expressed enzymes can alter levels of metabolites and cofactors and disrupt metabolism, cause stress responses, or other as yet poorly understood effects. More fully characterizing and modeling biological systems will reveal principles and design rules for synthetic biology [53*,59**]. These can be implemented using an increasing number of tools to program gene expression: ribosome binding sites (RBSs) [60], promoters [61], trans-acting activators [62], sensors and switches [50,63], enzyme fusions [64], scaffolds [65], localization tags [66], and other genetic ‘parts’. Employing these in ways approaching the complex natural orchestration of metabolism [67] will be necessary for more sophisticated and better performing biological designs.

Although we have constrained our discussion to S. cerevisiae and E. coli, several other microorganisms are already used as hosts for engineered bioproduction [68]. A different host organism may have a significantly different base metabolism, availability of substrates and cofactors, or compatibility with heterologous genes, which can lead to better pathways. Exploring and developing genetic tools for non-traditional hosts will open possibilities for novel metabolic pathways.

Conclusions
With more powerful synthetic biology tools, the concept of ‘host’ gives way to a fundamentally different synthetic organism. It is worth considering how the design objectives differ between a natural organism that evolved to maximize fitness in a given ecological niche and one synthetically constructed to convert a feedstock into a product. The propensity of organisms to grow — rather than to produce a target molecule — can lead to lower product yields. The propensity to mutate — perhaps necessary to adapt to changing environmental conditions — may not be desirable in constant fermentation conditions, and in fact yield to strain instability. These fundamental biological forces have significant impacts on the viability of microbial bioproduction.

For biofuels, strain performance is paramount. A gallon of petroleum gasoline sells for less than a gallon of water [69]. Techno-economic analysis shows that microbial biofuels provide for significant reductions in CO₂ emissions over using petroleum fuels [70]. However, these savings, and the economic viability of such bioprocesses, depend largely on biochemical pathway yields and feedstock costs. Bioprocesses that utilize as feedstock cellulosic biomass — agricultural or wood industry by-products, grasses growing on marginal land, among others — offer maximal CO₂ offsets, and do not compete with food production [71]. In the future it may be possible to engineer strains that grow directly on cellulosic biomass, or other abundant and inexpensive substrates, such as methane or CO₂. Or it may be possible to produce molecules with better performance, or as yet unimagined uses. As synthetic biology matures, this young technology holds vast potential to supplant the fossil economy with a sustainable and versatile biomanufacturing platform.

Acknowledgements
We thank Nicholas Clements, Victor Chubukov, and Maren Wehrs for suggestions while preparing this manuscript. This work was funded by the Joint BioEnergy Institute (JBEI), which is funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under Contract DE-AC02-05CH11231, and by the National Science Foundation under award MCB-1341894.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


15. Choi YJ, Lee SY: Microbial production of short-chain alkanes. [Internet]. Nature 2013, 502:571-574. This paper shows production of short chain hydrocarbons for the first time using a fatty acyl-CoA reductase from Clostridium acetobutylicum and a fatty aldehyde decarboxylase from Arabidopsis thaliana.


22. Haushalter RW, Groff D, Deutsch S, The L, Chavkin Ta. [Brunner SF, Katz L, Keasling JD]: Development of an orthogonal fatty acid biosynthesis system in E. coli for oleochemical production. Metab Eng 2015, 30:1-6. Type I fatty acid synthases were expressed for the first time in parallel with native E. coli fatty acid synthase for the production of fatty alcohols and methyl ketones.


Synthetic biology for microbial production of biofuels d’Espaux et al. 65


61. A whole-genome scale computational model including flux balance analysis and kinetic data is used to predict beneficial strain improvements.


