Chapter Title:

Alkane biosynthesis in bacteria

Authors:

Steven Brown, Josh Loh, Stephen J. Aves and Thomas P. Howard

Author affiliations:

Steven Brown

Synthace Ltd., The London Bioscience Innovation Centre, 2 Royal College Street, London, NW1 0NH, UK

Email: s.brown@synthace.com

Josh Loh

School of Natural and Environmental Sciences, Faculty of Science and Engineering, Newcastle University, NE1 7RU, UK

Email: J.Loh2@newcastle.ac.uk

Stephen J. Aves

Biosciences, College of Life and Environmental Sciences, Geoffrey Pope Building, University of Exeter, EX4 4QD, UK

Email: S.J.Aves@exeter.ac.uk

Thomas P. Howard (corresponding author)

School of Natural and Environmental Sciences, Faculty of Science and Engineering, Newcastle University, NE1 7RU, UK,

The Centre for Synthetic Biology and the Bioeconomy, Newcastle University, NE1 7RU, UK

Email: Thomas.Howard@newcastle.ac.uk

Tel: +44 (0)191 208 48 54

Abstract:

Biofuels are a commercial reality with ethanol comprising approximately 10% of the US retail fuel market, and biodiesels contribute a little under 5% to the EU retail fuel market. These biofuels molecules are derived from the fermentation of sugars by yeast (ethanol), and from the chemical modification of animal fats and plant oils (biodiesel). However, these biofuel molecules are chemically distinct from the petroleum fuels that they are blended with. Petroleum-based fuels are predominantly composed of alkane and alkene hydrocarbons. These differences impact on fuel properties and infrastructure compatibility resulting in a 'blend wall' that - without significant infrastructure re-alignment and associated costs - limits the use of biofuels. For this reason, there is great interest in biosynthetic routes for alkane and alkene production. Here we will review the known biological routes to alka/ene biosynthesis with a focus on bacterial alkane and alkene biosynthetic pathways. Specifically, we will review pathways for which the underlying genetic components have been identified. We will also investigate the development of engineered metabolic pathways that permit the production of alkanes and alkenes that are not

naturally synthesised in bacteria (heterologous production), but are suitable for industrial commercial application. Finally, we will highlight some of the challenges facing this research area as it moves from proof-of-principle studies towards industrialisation.

1. Introduction

Biogenic production of alkane and alkene hydrocarbons has received much attention in recent years. This is largely driven by the growth of the biofuels sector in response to climate change legislation and fuel security concerns. Current retail biofuels fall into two categories: alcohols (primarily ethanol) that are blended with petrol, and both fatty acid ethyl esthers (FAEEs) and fatty acid methyl esthers (FAMEs) that are blended with diesel to form fatty-acid derived biodiesel. Ethanol is produced predominantly by yeast fermentation and is mixed with petrol (gasoline). Biodiesels, derived from animal fats or plant oils, are mixed with diesel. These first-generation biofuels, most notably ethanol, have demonstrated a rapid increase in their penetration of the fuel market. In the US for example, ethanol content in petrol rose from little over 1% in 2001 to nearly 10% within a decade. Six years later, however, and ethanol remains at approximately 10% of the market. The failure to penetrate the fuel market further is in part due to the challenges that these biofuel molecules present to the fuel sector. They are not wholly compatible with our petroleum-focussed infrastructure. Ethanol, for example is hygroscopic, meaning that it has a high tendency to absorb water from the air. The presence of water makes it corrosive to transport infrastructure and reduces the heat of fuel combustion. Biodiesel is also moderately hygroscopic, and, if not blended correctly, can form waxes within fuels at cool temperatures. Both of these biofuels therefore have limitations imposed on the ratio of biofuel to petroleum - the so called ‘blend wall’. This is typically in the region of 10-15%. It is to circumvent this blend wall that there has been a renewed interest in alkane biogenesis.

Alkane and alkene hydrocarbon molecules are chemically and structurally identical to the molecules found within petroleum-based fuels (**Fig. 1**). As such there is no theoretical or practical limitation on their inclusion within current transport infrastructure. Within the last decade, the underlying genetic components of alkane biogenesis have been discovered in diverse species from across all kingdoms. This knowledge is in turn permitting a greater understanding of the biochemical processes involved and spurring the search for further discoveries. Coupled to this ‘discovery-focussed’ effort, there are significant advances in engineering natural biosynthetic pathways to produce alkanes and alkenes in microbial hosts, such as *Escherichia coli* and *Saccharomyces cerevisiae*. The aim of this research is to improve yields and to direct carbon flux towards infrastructure-compatible molecules - so called ‘drop-in’ biofuels. This chapter will review current knowledge of the underlying genetic and biochemical basis for alka/ene biogenesis in bacteria, and will describe efforts to manipulate these biological systems to develop metabolic pathways that can produce alkane and alkene hydrocarbons tailored to the specific requirements of our transport infrastructure. The biosynthesis of other hydrocarbon molecules, for example isoprenoids, will not be considered here, nor will we review the extensive literature manipulating and enhancing fatty acid biosynthesis. For reviews on these topics see Schrader and Bohlmann (2015) and Mehrer et al. (2016) respectively.

2. Natural alkane biosynthetic pathways

Biogenic sources of alkane have been known for some considerable time (Albro and Dittmer 1969a, b), and alkane biosynthesis can be observed across the biological domains. Examples include alkanes as components in waxes of plants (Bernard et al. 2012) and animals (Cheesbrough and Kolattukudy 1988), in storage reserves and

membrane components of microalgae (Dennis and Kolattukudy 1992; Grossi et al. 2000; Sorigue et al. 2016) and in pheromones, defence compounds and cuticular waxes of insects (Howard 1982; Howard et al. 1982; Reed et al. 1994; Tillman et al. 1999; Qiu et al. 2012). Alkanes are also known to be synthesised in fungi (Griffin et al. 2010; Gianoulis et al. 2012; Spakowicz and Strobel 2015) and in many bacterial species (McInnes et al. 1980; Beller et al. 2010; Schirmer et al. 2010; Sukovich et al. 2010b; Rude et al. 2011) though their physiological roles are less well understood. In eukaryotic systems, the genetic components have been identified in plant cuticular wax biosynthesis and in fruit fly wax biosynthesis. In plants, very long chain (VLC) alkanes are an integral component of the waxy cuticle, comprising up to 70% of the cuticle in the model species *Arabidopsis thaliana*. VLC alkanes have a carbon chain length (C_n) of between C_{20} and C_{36}. In *Arabidopsis*, the proteins ECERIFERUM1 (CER1), ECERIFERUM3 (CER3) and the endoplasmic reticulum-associated cytochrome b5 isoform CYTB5s interact to enable the conversion of fatty acids to VLC alkanes (Bourdenx et al. 2011; Bernard et al. 2012). This proposed conversion is via a fatty acyl-CoA intermediate, with the subsequent release of CO. Clear evidence for the role of CER1/CER3 and CYTB5 has been demonstrated by the reconstitution of this pathway in the yeast, *S. cerevisiae* and the resulting production of VLC alkanes (Bernard et al. 2012). Within the fruit fly *Drosophila melanogaster*, a P450 enzyme of the CYP4G family oxidatively produces alkanes from fatty aldehydes (Qiu et al. 2012). The enzyme reaction allows the production of the alkane heptadecane from the fatty aldehyde octadecanal, with the release of CO_2. The biochemical and genetic mechanisms underlying the production of alkanes and alkenes in other eukaryotic systems remain to be elucidated. In this section, we will review the remainder of the alkane and alkene biosynthetic pathways for which the genetic components have been identified (**Fig. 2**). All of these biogenic routes are found within bacteria. They include the cyanobacterial biosynthesis of alkanes and alkenes (section 2.1) catalysed by an acyl-[acyl carrier protein] (ACP) reductase and an aldehyde deformylase oxygenase, the production of VLC alkanes (section 2.2) via the *oleABCD* gene cluster, and the production of mid-chain length terminal alkenes (2.3) that can be accomplished via three independent mechanisms.

2.1 Cyanobacterial biosynthesis of alkanes and alkenes

The discovery and elucidation of a two-step alkane biosynthetic pathway in cyanobacteria stimulated a renewed interest in alkane biosynthesis (Schirmer et al. 2010). The cyanobacterial alkane biosynthetic pathway comprises a two-step reaction involving firstly, the reduction of intermediates of fatty acid biosynthesis (specifically the growing fatty acyl-ACP chain) to a fatty aldehyde, and secondly, the deformylation of the fatty aldehyde to an C_{n-1} alkane or alkene (Das et al. 2011; Warui et al. 2011). The genes responsible for this reaction were discovered through a subtractive genomics approach (Schirmer et al. 2010). Cyanobacteria benefit from reproducible reports of alkane biosynthesis (Winters et al. 1969) and the availability of whole genome sequences. An assessment of 11 cyanobacterial species for their ability to make alkanes in culture identified ten that could synthesise alkanes and one that could not. Subtractive analysis of the available genomes identified 17 genes common to all species that produced alkanes but absent in the one species that did not. While it was possible to assign biological functions to ten of these genes, seven candidates remained without clear biological roles. Of these seven, two of these genes were hypothesised as likely candidates and as a result they were cloned from *Synechococcus elongatus* PCC7942. These were orf1594 and orf1593. They were expressed in *E. coli* and, when combined, the result was the biosynthesis of heptadecane (C_{17}) and pentadecane (C_{15}) (**Fig. 1A**). Further investigations of homologs of orf1594

and orf1593 in other cyanobacteria identified this to be the common genetic basis for alkane biosynthesis across all cyanobacteria investigated within this study (Schirmer et al. 2010).

The two genes, orf1594 and orf1593, encode an acyl-ACP reductase (AAR) and an aldehyde deformylating oxygenase (ADO) respectively (**Fig. 2A**) (Schirmer et al. 2010; Das et al. 2011; Eser et al. 2011). Both AAR and ADO are soluble enzymes and the reactions proceed in the cytosol, in contrast to the *Arabidopsis* CER1/CER3/CYTB5s and *Drosophila* CYP4G proteins which are membrane associated. The ADO enzyme is responsible for generating alkanes from fatty aldehydes, though it can also generate C_{n-1} fatty aldehydes and alcohols (Aukema et al. 2013). Initially, it was suggested that the decarbonylation of fatty aldehyde produces carbon monoxide as a by-product (Schirmer et al. 2010). Further investigations with *Nostoc punctiforme* ADO heterologously expressed in *E. coli* have indicated that the by-product was in fact formate (Warui et al. 2011). The cyanobacterial ADO is a member of the non-heme dinuclear iron oxygenase family of enzymes. This family of enzymes use molecular oxygen and a reducing system, typically ferredoxin, ferredoxin oxidoreductase and NADPH, to reduce the enzyme at the start of each reaction. Initially, studies of ADO from *Prochlorococcus marinus* MIT9313, heterologously expressed and purified from *E. coli*, indicated that the reaction mechanism also produces formate and alkane under anaerobic conditions (Das et al. 2011). Other experiments on *N. punctiforme*, *Synechococcus* sp RS9917 and *Synechococcus* sp PCC6803 also exhibited alkane biosynthesis in anaerobic conditions (Eser et al. 2011). It was therefore thought that the enzyme may convert fatty aldehydes to alkane in the absence of molecular oxygen. If ADO operates in the absence of oxygen it was hypothesised that water provides the oxygen required for the formation of formate (Eser et al. 2011). Radiolabelling studies with ^{18}O-water suggested that a hydrolytic event was solely responsible for the conversion of fatty aldehyde to its corresponding *n*-1alkane and formate (Eser et al. 2011). A reducing system is still required even under anaerobic conditions and reaction rates were further improved using a non-physiological reducing system, 5-methylphenazinium methylsulphate (PMS):NADPH as opposed to the ferredoxin system (Eser et al. 2011). In direct contradiction to these investigations, an equivalent study with ADO from *N. punctiforme* suggested an absolute dependence on molecular oxygen (Li et al. 2011) Later studies have confirmed this view, with labelled $^{18}O_2$ and ^{18}O-water demonstrating a strict dependence of ADO on molecular oxygen with *P. marinus* and *N. punctiforme* orthologs (Li et al. 2012). The discrepancy with the early work was attributed to oxygen contamination in a supposedly anoxic condition during enzymatic assays. To reflect the formation of formate and use of oxygen, the original nomenclature for the enzyme aldehyde decarbonylase has been changed to aldehyde deformylating oxygenase.

Catalytic turnover of ADO is incredibly slow *in vitro* with single turnover rates measured in hours (Das et al., 2011; Eser et al., 2011; Warui et al., 2011). Solubility of the substrate may play a part in the slow turnover rates, as even at low concentrations long substrates such as octadecanal forms micelles. Low rates are typical for enzymes that catalyse reactions involving insoluble substrates. Shorter chain length substrates, such as heptanal, are also suitable substrates (Eser et al., 2011). There are also improvements in rates when the reaction is reconstituted in the presence of the endogenous cyanobacterial reducing system (Zhang et al., 2013). Reaction rates in these circumstances can be measured at 0.4 min^{-1} against octadecanal. The low activity and stability of the ADO presents a significant challenge for its deployment in metabolic engineering, though it has undoubtedly proven a popular choice of enzyme for alkane biosynthesis.

2.2 Very-long chain (VLC) alkene biosynthesis

The presence of VLC alkenes has been reported in many Gram-positive and Gram-negative bacteria (Sukovich et al. 2010a), including in *Sarcina lutea* (Albro and Dittmer 1969a, b), *Micrococcus luteus* (Beller et al. 2010), *Arthrobacter aurescens* (Frias et al. 2009), *Stenotrophomonas maltophilia* (Bonnett et al. 2011) and *Shewanella* species (Sugihara et al. 2010; Sukovich et al. 2010b). These alkenes typically fall into the range C_{25} to C_{33} and can be highly unsaturated molecules. The highly unsaturated alkene identified in *Shewanella* is shown in **Fig. 1B**. The biosynthesis of these long chain hydrocarbons is proposed to be important in maintaining membrane integrity, particularly in marine environments and cold conditions (Nichols et al. 1995; Sukovich et al. 2010b). This function is not restricted to bacterial membranes, with similar physiological function proposed in some unicellular marine eukaryotes (Grossi et al. 2000). The genes responsible have been termed *ole* genes (the name derived from olefin synthesis) and biosynthesis of alkenes requires the presence of three or four genes, *oleA*, *oleB*, *oleC* and *oleD* (Sukovich et al. 2010a; Christenson et al. 2017a). A combination of biochemical deduction and genome analysis was responsible for the identification of three *ole* genes in *M. luteus* (Beller et al. 2010). In this instance, research benefited from well-documented alkene biosynthesis in the closely related *S. lutea* species (Albro and Dittmer 1969a, b). The pathway in *S. lutea* was hypothesised to involve an initial decarboxylation of a fatty acyl-CoA substrate, followed by a 'head-to-head' condensation reaction. The availability of the genome from the closely related *M. luteus* allowed searches for homologs of enzymes involved in condensing reactions in fatty acid biosynthesis to identify three candidate genes (Beller et al. 2010). Of these, one corresponded closely to the β-ketoacyl-ACP synthase II (KASII) gene *fabF* and another corresponded to the β-ketoacyl-ACP synthase III (KASIII) gene *fabH*. These two genes were also present in a six-gene cluster encoding several other enzymes of fatty acid biosynthesis, strongly suggesting their role was in the synthesis of fatty acids rather than alkenes. The third gene, Mlut_13230, was also homologous to *fabH* but the sequence similarity diverged to a greater extent than the other candidate gene. Heterologous expression of these candidate genes in *E. coli* did not result in the production of alkenes. Careful examination of the metabolite profile of *E. coli* expressing Mlut_13230 however did reveal the appearance of monoketones. This is consistent with the proposed mechanism of alkene biosynthesis. As a result, further analysis of the *M. luteus* genome revealed that Mlut_13230 is present as part of a three-gene cluster, comprising Mlut_13230, Mlut_13240 and Mlut_13250. It was shown that expression of all three of the genes in this gene cluster in *E. coli* results in the biosynthesis of VLC alkenes (Beller et al. 2010).

In two similar studies, VLC alkenes were identified in *Shewanella oneidensis* (Sukovich et al. 2010b) and *Shewanella* sp. strain osh08 (Sugihara et al. 2010). The principal alkene produced was a highly unsaturated molecule, containing nine C=C double bonds. NMR and GC/MS analysis revealed it to be 3,6,9,12,15,19,22,25,28-hentriacontanonaene, shown in **Fig. 1A**. Genetic evidence indicates that the production of this alkene in *Shewanella* requires the biosynthesis of poly-unsaturated fatty acids (PUFAs). Strains in which PUFA biosynthesis is inhibited through the removal of key PUFA biosynthetic genes are unable to synthesise alkenes (Sugihara et al. 2010; Sukovich et al. 2010b). Unlike in *M. luteus*, however, *Shewanella* alkene biosynthesis requires four genetic components. One of the proteins identified in *S. oneidensis*, gi 24373309, has a 31% homology at the amino acid level to Mlut_13230. Another, identified in *S. oneidensis* as gi 24373312, shows 31% homology to Mlut_13250, while the two remaining proteins both have homology to Mlut_13240 from *M. luteus*. In fact, on closer examination it is evident that Mlut_13240 is a fusion protein of these proteins (Beller et al. 2010; Sukovich et al. 2010b). Genomic analysis indicates that the occurrence of this type of 'head-to-head'

condensation is widespread in bacteria and different genomic arrangements have been described (Sukovich et al. 2010a). It is now recognised that the *oleABCD* gene cluster encodes the following enzymes: *oleA* encodes a thiolase, *oleB* encodes an alpha/beta hydrolase, *oleC* encodes an AMP-dependent ligase/synthase and *oleD* encodes short chain dehydrogenase/reductase. None of the *ole* genes have yet been identified in the Archaea or Eukaryota (Sukovich et al. 2010a).

The biosynthesis of these VLC alkenes commences with the reaction catalysed by the OleA thiolase (**Fig. 2B**). In this reaction, the non-decarboxylative Claisen condensation of two fatty acyl-CoA molecules, such as tetradecanoyl-CoA (C_{14}), results in the production of a C_{27} β-keto acid, 2-myristoylmyristic acid (Frias et al. 2011). The next step, catalysed by OleD, an NADPH-dependent 2-alkyl-3-ketoalkanoic acid reductase results in the production of a hydroxyl alkanoic acids (Bonnett et al. 2011). Finally, OleC has been shown generate a thermally labile β -lactone which can spontaneously, and non-biologically decarboxylate to an alkene. However, it is currently believed that this final reaction is catalysed *in vivo* by OleB - a β-lactone decarboxylase (Kancharla et al. 2016; Christenson et al. 2017b). Recent evidence indicates that OleB, OleC and OleD enzymes assemble into a large, multiprotein complex (Christenson et al. 2017a). It is proposed that such a configuration retains the highly reactive beta-lactone intermediate produced by the OleC-catalysed reaction from unwanted reactions - such multi-protein complexes are observed across the biological domains (for examples see Singleton et al. (2014). Genetic tools available in *S. oneidensis* permit the removal of the endogenous *oleA* gene and its replacement with homologs from other species. When *oleA* is replaced with the equivalent gene from *S. maltophilia* there is an alteration in the alkene profile. In the wild type strain, 3,6,9,12,15,19,22,25,28-hentriacontanonaene is the sole product, whereas in the modified strain an abundance of more saturated monoketones is produced (Sukovich et al. 2010b). Likewise, swapping in *oleA* genes from a range of diverse bacteria alters the hydrocarbon profile in such a way that it resembles the hydrocarbon profile from the donor *oleA*, rather than the host *S. oneidensis* (Sukovich et al. 2010a).

2.3 Biosynthesis of mid-chain length alkenes

The third example of hydrocarbon biogenesis in bacteria is the production of mid-chain length terminal alkenes and there are three distinct genetic examples of how this may be achieved. These include their production via cytochrome P450-catalysed reactions in *Jeotgalicoccus* sp. (Rude et al. 2011), via a non-heme iron oxidase in *Pseudomonas aeruginosa* (Rui et al. 2014) and via a modular polyketide synthase (PKS) in *Synechococcus* sp. (Mendez-Perez et al. 2011; Mendez-Perez et al. 2014). These reactions produce mid-chain length 1-alkenes far shorter than those described in section 2.2, with chain lengths typically in the region of C_{10} (for the non-heme iron oxidase) and C_{18} to C_{20} (for the P450 and PKS reactions). Mid-chain length terminal alkenes are of interest not just for their fuel properties but also, because they can be readily derivatised, they are important precursor molecules for commodity chemicals such as plastics, lubricants and detergents.

A reverse genetics approach was used to isolate the genes responsible for medium chain length (C_{18} to C_{20}) alkene biosynthesis in *Jeotgalicoccus* sp. (Rude et al. 2011). *Jeotgalicoccus* species are low GC Gram positive firmicutes. Analysis of micrococci reported to produce alkenes revealed several that produced very long chain alkenes (as discussed in section 2.2) and one that produced medium chain length alkenes. This ability to produce medium chain length alkenes rather than VLC alkenes was identified in a *Jeotgalicoccus* species, and subsequent investigation revealed this was a common feature in many *Jeotgalicoccus* species. Both feeding assays

- in which media were supplemented with fatty acids - and the ability of crude cell lysates to convert fatty acids to the corresponding C_{n-1} alkenes, suggested direct activity on fatty acids (Rude et al. 2011). Importantly, the ability to assay alkene production in this manner permits protein purification and fractionation experiments. As a result, fractions of partially purified proteins with the ability to convert fatty acids to alkenes were identified that contained two candidate proteins. One of these was identified as orf880, and heterologous expression of orf880 in *E. coli* resulted in the production of pentadecene and 1,10-heptadecadiene. The gene was termed *oleT* and it encodes an enzyme that is a cytochrome P450 fatty acid peroxygenase (Rude et al. 2011), specifically a member of the cyp152 subfamily of P450s. They are found in a diverse range of bacteria, including many that catalyse the hydroxylation of fatty acids. Members of this family are active using H_2O_2, rather than O_2, NADPH or other redox partners. $OleT_{JE}$ has been shown to catalyse the formation of C_{n-1} alkenes through the H_2O_2-dependent decarboxylation of C_{12}, C_{14}, C_{16}, C_{18} and C_{20} saturated fatty acids (Belcher et al. 2014). Other P450s known with similar activity include a P450 from *Rhodotorula minuta* that converts isovalerate to isobutene (Fukuda et al. 1994).

Another pathway for the biosynthesis of mid-chain length alkenes is found in *P. aeruginosa* (Rui et al. 2014). Discovery of the gene responsible for the biosynthesis of the 11-carbon semi-volatile 1-undecane began with feeding labelled putative substrates [12-^{13}C]dodecanoic acid and [1-^{13}C] dodecanoic acid to *P. aeruginosa* cultures. The result was the production of [12-^{13}C]undecene and [U-^{12}C11]undecene, a result that confirmed both the substrate (fatty acid), and that the terminal carboxylic acid moiety is removed in the reaction (Rui et al. 2014). A high-throughput screen of approximately 6000 fosmid-containing *E. coli* cultures was employed and identified a single gene for 1-undecene production, termed *undA*. UndA is a small protein of 261 amino acids. Biochemical analysis reveals that Fe^{2+} is essential for its function *in vitro*. Moreover, whilst the dominant function *in vivo* appears to be the production of 1-undecene, it is also capable of acting on fatty acids in the range of C_{10} to C_{14}. The reaction is proposed to proceed via the sequential binding of the carboxylate moiety of the fatty acid to ferrous iron within UndA, followed by binding of O_2 to form a Fe(III) superoxide complex. Electron transfer results in the production of 1-undecene, CO_2 and H_2O and the reduction of an unstable Fe(IV)=O species. Importantly, unlike AAR/ADO and the *oleABCD* pathway, both $OleT_{JE}$ and UndA work directly on fatty acid substrates to produce the corresponding C_{n-1} alkene in a single enzyme catalysed reaction.

Finally, 1-alkene production by a modular synthase-encoding gene from *Synechococcus* sp. PCC 7002 has been reported (Mendez-Perez et al. 2011; Mendez-Perez et al. 2014). *Synechococcus* sp. 7002 has been demonstrated to produce $C_{19:1}$ and $C_{19:2}$ alkenes, but interestingly, the largest fatty acid detected from *Synechococcus* sp. 7002 is a C_{18} fatty acid. This suggests that the mechanism is not analogous to the various C_{n-1} mechanisms for generating mid-chain length terminal alkenes, or the cyanobacterial ADO mechanism. Rather, an elongation-decarboxylation mechanism is proposed (Mendez-Perez et al. 2011; Mendez-Perez et al. 2014). The *ols* gene (olefin synthesis) was identified as a putative candidate, and Δ*ols* strains did not contain alkenes. The Ols system comprises a single, multi-domain protein, and is found in only a small number of cyanobacteria (Xie et al. 2017). The biochemical steps are yet to be fully elucidated.

3. Synthetic alkane biosynthetic pathways

The identification of genetic components responsible for alkane biosynthesis across the biosphere is the crucial first step in the production of petroleum-replica hydrocarbons by microbial fermentation. However, the demands

the fuel industry places on the chemical and structural diversity for alkanes are not the same as the evolutionary pressure that has led to the biochemistry capable of biosynthesising alkanes (Rude and Schirmer 2009; Jimenez-Diaz et al. 2017). Natural biosynthetic routes predominantly produce very long chain alkanes (e.g. *A. thaliana, M. luteus*), terminal alkenes (*oleT*) or a narrow range of linear alkanes (dominated by C_{17} and C_{15} in cyanobacteria). For alkane biosynthesis to be of commercial relevance, as well as biological interest, novel pathways must be developed that permit the production of the range and diversity of chemicals blended in fuel.

3.1 Altering starting substrates

The alkane profile of organisms is greatly influenced by the first step in the alkane biosynthetic pathway. For example, the cyanobacterial enzyme AAR reveals a preference for C_{18} over C_{16} fatty acyl-ACPs (Schirmer et al. 2010; Howard et al. 2013), while the nature of the OleA present in *Shewanella* greatly affects the resulting alkane profile (Sukovich et al. 2010a). The ability to manipulate the output of alkane biosynthetic pathways will therefore be impacted by the ability to control which acyl chains enter the biosynthetic pathway. As a result, many groups have successfully identified novel routes for providing fatty aldehydes for alkane biosynthesis.

In one study, the fatty acid reductase (FAR) complex from *Photorhabdus luminescens*, used to provide fatty aldehyde for bacterial luminescence, replaced the activity of AAR in an engineered *E. coli* (Howard et al. 2013). The *P. luminescens* FAR complex (encoded by *luxCDE*) exploits fatty acids, fatty acyl-CoAs and fatty acyl-ACP molecules to generate fatty aldehydes (Meighen 1998), thereby widening the pool of substrates available to feed into alkane biosynthesis. When coupled with the activity of ADO, there was a change in the abundance and composition of the alkane profile that more closely reflected the composition of the fatty acid profile than did the reconstituted AAR/ADO pathway (Howard et al. 2013). In two further examples, the carboxylic acid reductase (CAR), a large, single chain polypeptide from *Mycobacterium marinum* (Akhtar et al. 2013) and a fatty acyl-CoA reductase from *Acinetobacter* sp. M-1 (Yan et al. 2016) were coupled with ADO to enable the used of free fatty acid pools, rather than fatty acyl-ACP, as starting substrate. By using fatty acids rather than fatty acyl-ACP, the nature of the alkanes produced can be altered, and any future metabolic engineering strategies can benefit from the extensive literature on improving fatty acid yields in bacteria.

Another study exploited the plant alkane biosynthetic system, rather than ADO but also permitted access to the fatty acid pool in *E. coli* (Choi and Lee 2013). In this instance, the authors increased the natural ability of *E. coli* cells to produce fatty acyl-CoAs from fatty acids, via fatty acyl-CoA synthetase (FadD), and coupled this activity with a fatty acyl-CoA reductase (ACR) activity from *Clostridium acetobutylicum* to reduce fatty acyl-CoAs to fatty aldehydes. The final conversion of these aldehydes to alkanes was completed following addition of an *E. coli* codon-optimised *Arabidopsis CER1* gene. This report is important in two respects. Firstly, previous attempts to heterologously express an active CER1 protein had been unsuccessful (Bourdenx et al. 2011; Bernard et al. 2012). Results indicated that CER1/CER3 are mandatory for activity and that they rely on an interaction with their partner CYTB5s. By providing modifications upstream in *E. coli* that provide the required fatty aldehyde, the authors have been able to demonstrate that the difficulties are, at least in part, metabolic in nature; without CER3 - a putative acyl-CoA reductase - there is no substrate for CER1 to act on. Secondly, this analysis supports the assertion that CER1 catalyses the second stage of alkane biosynthesis and is therefore informative as to the metabolic routes for alkane biosynthesis *in planta*.

3.2 Manipulating chain length

The chain length of alkanes and alkenes is one of the most crucial contributors to the properties of the fuel itself. Chain lengths in the region of C_5 to C_{10} are the dominant chain length distribution found in petrol (gasoline) fuel blends, whereas C_{12} to C_{18} chain length are found in jet fuels and diesel. The ability to alter the chain length is therefore of great interest and has been demonstrated by several groups independently.

The first group to demonstrate the ability to manipulate the alkane output was a team of undergraduate students from the University of Washington, competing in the Internationally Genetically Engineered Machines (iGEM) competition in 2011 (Harger et al. 2013). They achieved this through a combination of genetic and media manipulations. To do this, the two-step cyanobacterial AAR/ADO pathway was expressed in *E. coli*. An additional modification was made through the addition of the gene encoding KASIII from *Bacillus subtilis* (*fabHB*). The KASIII enzyme from *B. subtilis*, unlike its *E. coli* counterpart, accepts a wider range of primer molecules into fatty acid biosynthesis. In addition to this genetic manipulation they also grew their *E. coli* in the presence of the three-carbon molecule propanoate. The presence of propanoate in *E. coli* growth media increases the pool of propanoyl-CoA molecules which are available for incorporation into fatty acid biosynthesis via the introduced KASIII. This ensured that - at least for some cycles - a three-carbon compound was incorporated into the fatty acid elongation cycle as well as the normal two carbon acetyl-CoA. The resultant fatty acids are therefore one carbon longer than the typical fatty acids. Following the C_{n-1} rule for cyanobacterial AAR/ADO alkane biosynthesis, the combined effect of these genetic and media manipulations is the production of even- as well as odd-chain length alkanes.

Entirely genetically encoded manipulations of alkane output have also been successful. For example, very short chain alkanes (e.g. propane) have been achieved in *E. coli* (Choi and Lee 2013; Kallio et al. 2014; Sheppard et al. 2016; Zhang et al. 2016). These manipulations are invariably made possible because of the switch from an acyl-ACP dependent system, to one that can exploit a manipulated fatty acid pool. In each instance, alkane production in *E. coli* whether via the FadD/ACR/CER1 manipulations of Choi and Lee (2013), or via CAR (Akhtar et al. 2013; Kallio et al. 2014), FAR (Howard et al. 2013), or ACR (Yan et al. 2016), was redirected to shorter chain alkanes by the inclusion of various thioesterases.

3.3 Production of branched alkanes

Branched chain alkanes (**Fig. 1C**) are crucial for fuel performance, preventing stacking (gelling) of fuel at cold temperatures or altitude. To date there are no elucidated natural pathways that produce mid-chain length branched alkanes suitable for retail fuels. To demonstrate that alkane biosynthesis can be directed towards such branched molecules *E. coli* expressing the FAR/ADO pathway was analysed for the ability to incorporate exogenous branched fatty acids into alkanes. The results indicated that branched molecules could be used by this pathway (Howard et al. 2013). Establishing these capabilities at the genetic level in *E. coli* is complicated by the fact that *E. coli* cannot naturally synthesise branched fatty acids (Choi et al. 2000; Smirnova and Reynolds 2001) though other bacteria, notably *B. subtilis*, do produce branched fatty acids (Oku and Kaneda 1998). The difficulty in establishing branched fatty acid biosynthesis in *E. coli* is two-fold. *In vitro* assembly of *E. coli* fatty acid biosynthesis indicates firstly that biosynthesis of branched fatty acids relies on the addition of appropriate branched primer molecules, and secondly that it relies on the addition of an alternative KASIII enzyme (Choi et al. 2000). To make branched fatty acids in *E. coli* therefore both the KASIII gene from *B. subtilis* needs to be

heterologously expressed, and suitable activity to generate the branched primer molecules, not naturally present in *E. coli*, is also required. For this to occur a further four genes are required that code for the multi-enzyme branched chain keto-dehydrogenase complex. Introducing all five genes leads to the production of branched fatty acids in *E. coli*, and the further addition of the FAR/ADO pathway results in the appearance of branched alkanes (Howard et al. 2013). The biosynthesis of short chain branched alkanes has also been established (Sheppard et al. 2016). Taken together with the results discussed in 3.2, it is apparent that it is entirely possible to synthesise a range of linear and branched alkanes and alkenes of the chain lengths appropriate for petroleum-replacement biofuels that can be blended depending upon the required fuel properties.

4. Research needs

Further advances in the metabolic engineering of microbes for alkane biosynthesis has naturally shifted towards increasing yields and efficiencies of the biosynthetic processes. These may be considered in four broad areas: choice of organism and growth conditions; pathway engineering; enzyme engineering and the removal of the alkane product. In addition, further discovery of the genetics underlying alkane biosynthesis across all living systems will improve our understanding of the evolutionary pressures and biochemical diversity of alkane biosynthesis, with a resulting impact inspiring novel engineering strategies.

4.1 Choice of chassis

Whilst it is possible to relatively rapidly develop a proof of principle strain that produces a product of interest, it is far more challenging to develop a strain that meets commercial targets and fits within a biorefinery concept (Runguphan and Keasling 2014). The choice of cell factory is critical in the assessment of the industrial production of chemicals and recently there has been a consolidation and focus on a few industrial cell factories (Rumbold et al. 2009; Vickers et al. 2010; Hong and Nielsen 2012). The proven ability of cell factories such as *E. coli*, *Corynebacterium glutamicum*, *Aspergillus niger*, *Pichia stipitis* and *S. cerevisiae*, to perform robustly within industrial processes highlights their potential as industrial cell factories that may replace the current production of chemicals from crude oil in a sustainable way. The large scale industrial production of cellulosic ethanol using the yeast *S. cerevisiae* ("Project Liberty": a joint venture between POET and DSM in the USA) shows the economic feasibility of using this cell factory for production of a relatively cheap, high volume commodity and is a success story related to the production of biofuels.

The AAR/ADO pathway has been shown to operate in a wide range of microbes beyond *E. coli*, including non-AAR/ADO cyanobacterial species (Yoshino et al. 2015) and the chemoautotrophic bacterium *Cupriavidus necator* (Crepin et al. 2016). The latter is already grown as a commercial concern for production of bioplastics. Alkane biosynthesis has also been demonstrated in eukaryotic microbes such as *S. cerevisiae* (Bernard et al. 2012; Buijs et al. 2015; Kang et al. 2017) and in the filamentous fungus *Aspergillus carbonarius* (Sinha et al. 2017). It is also important to note however that highly significant gains were made in the production of semi-synthetic artemisinin not only from metabolic engineering strategies, but through optimisation of fermentation and extraction conditions (Westfall et al. 2012; Paddon et al. 2013). It is therefore a key challenge to metabolic engineers to consider not only the optimisation of the genetic components (choice of chassis and pathway engineering for example) but to include the optimisation of environmental (i.e. fermentation) conditions at the same time (Mukhopadhyay et al. 2008).

4.2 Pathway engineering

Exploiting fatty acid biosynthesis for fuel production is advantageous, as many species have high carbon flux into these energy rich chemicals. Large increases in yields have been achieved in fatty acid and fatty acid-derived chemicals (reviewed in Mehrer et al. (2016)) and similar methods can be applied to increasing alkane biosynthesis efficiencies. There are early examples of this being applied successfully to microbial production of alkanes. Greater alkane titres have been achieved by manipulating the supply of substrates to the pathway and removal of competitive reactions (Cao et al. 2016; Song et al. 2016). In addition to these specific genetic manipulations, advances in computer guided metabolic engineering strategies will greatly assist these efforts (Patel et al. 2016).

4.3 Enzyme engineering

One means of addressing the need to improve system performance is through improvements to the performance of the enzymes responsible for catalysing the conversion from substrate (typically fatty acids) through to alkanes and alkenes. This is particularly important given the slow catalytic turnover of many of the enzymes. There are already examples of successful catalytic manipulations of both the ADO and OleT systems through their fusion to alternative reducing systems, or proteins capable of removing inhibiting compounds. As a peroxygenase, $OleT_{JE}$ uses H_2O_2 as its redox partner. However, excess reactive oxygen species can cause cellular damage and result in apoptosis. The cell therefore elicits many responses leading to rapid detoxification of reactive oxygen species and the removal of a driver of $OleT_{JE}$ alkene biosynthesis. To circumvent this studies have shown that $OleT_{JE}$ can perform H_2O_2-independent catalysis *in vitro* using either a flavodoxin/flavodoxin reductase system or a P450 RhFRED reductase domain from *Rhodococcus* sp. (Liu et al. 2014). Furthermore, an OleT-RhFRED fusion has been shown to perform the same enzymatic activity, but supported by NADPH and oxygen, in an engineered fatty acid-overproducing strain of *E. coli* (Liu et al. 2014). An alternative strategy has been to fuse $OleT_{JE}$ with an alditol oxidase (AldO) from *Streptomyces coelicolor* (Matthews et al. 2017). The addition of AldO fused to $OleT_{JE}$ enables local generation of H_2O_2 from polyols (e.g. glycerol, sorbitol and xylitol) and an increased conversion of tetradecanoic acid to alkenes compared to direct addition of H_2O_2. For ADO, H_2O_2 has been shown to reversibly inhibit its catalytic activity. In order to circumvent this, the creation of a fusion protein consisting of ADO fused to a catalase capable of removing local H_2O_2 led to a dramatic five-fold improvement in catalytic turnover *in vivo* (Andre et al. 2013). In each instance these protein engineering efforts demonstrate that catalytic activity can be maintained and indeed improved when the proteins are assembled into larger structures providing optimism that different strategies for improving alkane production through protein manipulations can be achieved.

4.4 Removal of product

As with any metabolic engineering strategy, as pathways improve and titres of target molecules increase, it is likely that toxicity of the product will become a problem. This is an under explored area that will need to be developed to maximise outputs of petroleum-replica hydrocarbons. Manipulating efflux pumps to increase yield of target chemicals has been successful (Dunlop et al. 2011; Lennen et al. 2012; Kato et al. 2015). One of the key challenges identified in this research however, is the need to increase efflux pump efficiency and specificity, rather than simply increasing the number of pumps. The latter scenario can result in physiological damage to membrane integrity with deleterious effects on cell survival.

4.5 Gene Discovery

Alkane biosynthesis is known in many living organisms yet only a small number of the biochemical pathways and the underlying genetic components are so far known. Elucidation of new pathways, for example in fungi (Gianoulis et al. 2012), mammalian waxes (Cheesbrough and Kolattukudy 1988), algae (Dennis and Kolattukudy 1992), or bacteria, will enhance our understanding of the biochemistry and enzymology involved in alkane biosynthesis. This knowledge will broaden the choices available to bioengineers for developing suitable bioprocesses and facilitate greater forward engineering of pathways for commercial alkane biosynthesis.

5. Acknowledgments

J.L. would like to thank the BBSRC for PhD funding (BB/M017036/2).

6. References

Akhtar MK, Turner NJ, Jones PR (2013) Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities. Proc Natl Acad Sci U S A 110 (1):87

Albro PW, Dittmer JC (1969a) The biochemistry of long-chain, nonisoprenoid hydrocarbons. I. Characterization of the hydrocarbons of Sarcina lutea and the isolation of possible intermediates of biosynthesis. Biochemistry 8 (1):394

Albro PW, Dittmer JC (1969b) The biochemistry of long-chain, nonisoprenoid hydrocarbons. IV. Characteristics of synthesis by a cell-free preparation of Sarcina lutea. Biochemistry 8 (8):3317

Andre C, Kim SW, Yu X-H, Shanklin J (2013) Fusing catalase to an alkane-producing enzyme maintains enzymatic activity by converting the inhibitory byproduct H2O2 to the cosubstrate O2. Proc Natl Acad Sci U S A 110 (8):3191

Aukema KG, Makris TM, Stoian SA, Richman JE et al. (2013) Cyanobacterial aldehyde deformylase oxygenation of aldehydes yields n-1 aldehydes and alcohols in addition to alkanes. ACS Catal 3 (10):2228

Belcher J, McLean KJ, Matthews S, Woodward LS et al. (2014) Structure and biochemical properties of the alkene producing cytochrome P450 OleTJE (CYP152L1) from the Jeotgalicoccus sp. 8456 bacterium. J Biol Chem 289 (10):6535

Beller HR, Goh EB, Keasling JD (2010) Genes involved in long-chain alkene biosynthesis in Micrococcus luteus. Appl Environ Microbiol 76 (4):1212

Bernard A, Domergue F, Pascal S, Jetter R et al. (2012) Reconstitution of plant alkane biosynthesis in yeast demonstrates that Arabidopsis ECERIFERUM1 and ECERIFERUM3 are core components of a very-long-chain alkane synthesis complex. Plant Cell 24 (7):3106

Bonnett SA, Papireddy K, Higgins S, del Cardayre S et al. (2011) Functional characterization of an NADPH dependent 2-alkyl-3-ketoalkanoic acid reductase involved in olefin biosynthesis in Stenotrophomonas maltophilia. Biochemistry 50 (44):9633

Bourdenx B, Bernard A, Domergue F, Pascal S et al. (2011) Overexpression of Arabidopsis ECERIFERUM1 promotes wax very-long-chain alkane biosynthesis and influences plant response to biotic and abiotic stresses. Plant Physiol 156 (1):29

Buijs NA, Zhou YJ, Siewers V, Nielsen J (2015) Long-chain alkane production by the yeast Saccharomyces cerevisiae. Biotechnol Bioeng 112 (6):1275

Cao YX, Xiao WH, Zhang JL, Xie ZX et al. (2016) Heterologous biosynthesis and manipulation of alkanes in Escherichia coli. Metab Eng 38:19

Cheesbrough TM, Kolattukudy PE (1988) Microsomal preparation from an animal tissue catalyzes release of carbon monoxide from a fatty aldehyde to generate an alkane. J Biol Chem 263 (6):2738

Choi K-H, Heath RJ, Rock CO (2000) β-ketoacyl-acyl carrier protein synthase III (FabH) Is a determining factor in branched-chain fatty acid biosynthesis. J Bacteriol 182 (2):365

Choi YJ, Lee SY (2013) Microbial production of short-chain alkanes. Nature 502 (7472):571

Christenson JK, Jensen MR, Goblirsch BR, Mohamed F et al. (2017a) Active multienzyme assemblies for long-chain olefinic hydrocarbon biosynthesis. J Bacteriol 199 (9)

Christenson JK, Richman JE, Jensen MR, Neufeld JY et al. (2017b) beta-Lactone synthetase found in the olefin biosynthesis pathway. Biochemistry 56 (2):348

Crepin L, Lombard E, Guillouet SE (2016) Metabolic engineering of Cupriavidus necator for heterotrophic and autotrophic alka(e)ne production. Metab Eng 37:92

Das D, Eser BE, Han J, Sciore A et al. (2011) Oxygen-independent decarbonylation of aldehydes by cyanobacterial aldehyde decarbonylase: a new reaction of diiron enzymes. Angew Chem Int Ed Engl 50 (31):7148

Dennis M, Kolattukudy PE (1992) A cobalt-porphyrin enzyme converts a fatty aldehyde to a hydrocarbon and CO. Proc Natl Acad Sci USA 89 (12):5306

Dunlop MJ, Dossani ZY, Szmidt HL, Chu HC et al. (2011) Engineering microbial biofuel tolerance and export using efflux pumps. Mol Syst Biol 7:487

Eser BE, Das D, Han J, Jones PR et al. (2011) Oxygen-independent alkane formation by non-heme iron-dependent cyanobacterial aldehyde decarbonylase: investigation of kinetics and requirement for an external electron donor. Biochemistry 50 (49):10743

Frias JA, Richman JE, Wackett LP (2009) C29 olefinic hydrocarbons biosynthesized by Arthrobacter species. Appl Environ Microbiol 75 (6):1774

Frias JA, Richman JE, Erickson JS, Wackett LP (2011) Purification and characterization of OleA from Xanthomonas campestris and demonstration of a non-decarboxylative Claisen condensation reaction. J Biol Chem 286 (13):10930

Fukuda H, Fujii T, Sukita E, Tazaki M et al. (1994) Reconstitution of the isobutene-forming reaction catalyzed by cytochrome P450 and P450 reductase from Rhodotorula minuta: decarboxylation with the formation of isobutene. Biochem Biophys Res Commun 201 (2):516

Gianoulis TA, Griffin MA, Spakowicz DJ, Dunican BF et al. (2012) Genomic analysis of the hydrocarbon-producing, cellulolytic, endophytic fungus Ascocoryne sarcoides. PLoS Genet 8 (3):e1002558

Griffin MA, Spakowicz DJ, Gianoulis TA, Strobel SA (2010) Volatile organic compound production by organisms in the genus Ascocoryne and a re-evaluation of myco-diesel production by NRRL 50072. Microbiology 156 (Pt 12):3814

Grossi V, Raphel D, Aubert C, Rontani J-F (2000) The effect of growth temperature on the long-chain alkenes composition in the marine coccolithophorid Emiliania huxleyi. Phytochemistry 54 (4):393

Harger M, Zheng L, Moon A, Ager C et al. (2013) Expanding the product profile of a microbial alkane biosynthetic pathway. ACS Synth Biol 2 (1):59

Hong KK, Nielsen J (2012) Metabolic engineering of Saccharomyces cerevisiae: a key cell factory platform for future biorefineries. Cell Mol Life Sci 69 (16):2671

Howard RW (1982) Chemical ecology and biochemistry of insect hydrocarbons. Annu Rev Entomol 27:149

Howard RW, McDaniel CA, Nelson DR, Blomquist GJ et al. (1982) Cuticular hydrocarbons of *Reticulitermes virginicus* (Banks) (Isoptera, Rhinotermitidae) and their role as potential species-recognition and caste-recognition cues. Journal of Chemical Ecology 8 (9):1227

Howard TP, Middelhaufe S, Moore K, Edner C et al. (2013) Synthesis of customized petroleum-replica fuel molecules by targeted modification of free fatty acid pools in *Escherichia coli*. Proc Natl Acad Sci U S A 110 (19):7636

Jimenez-Diaz L, Caballero A, Perez-Hernandez N, Segura A (2017) Microbial alkane production for jet fuel industry: motivation, state of the art and perspectives. Microb Biotechnol 10 (1):103

Kallio P, Pasztor A, Thiel K, Akhtar MK et al. (2014) An engineered pathway for the biosynthesis of renewable propane. Nat Commun 5:4731

Kancharla P, Bonnett SA, Reynolds KA (2016) Stenotrophomonas maltophilia OleC-Catalyzed ATP-Dependent Formation of Long-Chain Z-Olefins from 2-Alkyl-3-hydroxyalkanoic Acids. Chembiochem 17 (15):1426

Kang MK, Zhou YJ, Buijs NA, Nielsen J (2017) Functional screening of aldehyde decarbonylases for long-chain alkane production by Saccharomyces cerevisiae. Microb Cell Fact 16 (1):74

Kato A, Takatani N, Use K, Uesaka K et al. (2015) Identification of a Cyanobacterial RND-Type Efflux System Involved in Export of Free Fatty Acids. Plant Cell Physiol 56 (12):2467

Lennen RM, Politz MG, Kruziki MA, Pfleger BF (2012) Identification of transport proteins involved in free fatty acid efflux in *Escherichia coli*. Journal of Bacteriology, vol 195.

Li N, Norgaard H, Warui DM, Booker SJ et al. (2011) Conversion of fatty aldehydes to alka(e)nes and formate by a cyanobacterial aldehyde decarbonylase: cryptic redox by an unusual dimetal oxygenase. J Am Chem Soc 133 (16):6158

Li N, Chang WC, Warui DM, Booker SJ et al. (2012) Evidence for only oxygenative cleavage of aldehydes to alk(a/e)nes and formate by cyanobacterial aldehyde decarbonylases. Biochemistry 51 (40):7908

Liu Y, Wang C, Yan J, Zhang W et al. (2014) Hydrogen peroxide-independent production of a-alkenes by OleTJE P450 fatty acid decarboxylase. Biotechnology for Biofuels 7 (1):1

Matthews S, Tee KL, Rattray NJ, McLean KJ et al. (2017) Production of alkenes and novel secondary products by P450 OleTJE using novel H2 O2 -generating fusion protein systems. FEBS Lett 591 (5):737

McInnes AG, Walter JA, Wright JL (1980) Biosynthesis of hydrocarbons by algae: Decarboxylation of stearic acid to N-heptadecane in Anacystis nidulans determined by 13 C-and 2 H-labeling and 13 C nuclear magnetic resonance. Lipids 15 (9):609

Mehrer CR, Hernández Lozada NJ, Lai R-Y, Pfleger BF (2016) Production of fatty acids and derivatives by metabolic engineering of bacteria. In: Consequences of Microbial Interactions with Hydrocarbons, Oils, and Lipids: Production of Fuels and Chemicals. pp 1

Meighen EA (1998) Enzymes and genes from the lux operons of bioluminescent bacteria. Annual Review of Microbiology 42:151

Mendez-Perez D, Begemann MB, Pfleger BF (2011) Modular synthase-encoding gene involved in alpha-olefin biosynthesis in Synechococcus sp. strain PCC 7002. Appl Environ Microbiol 77 (12):4264

Mendez-Perez D, Herman NA, Pfleger BF (2014) A desaturase gene involved in the formation of 1,14-nonadecadiene in Synechococcus sp. strain PCC 7002. Appl Environ Microbiol 80 (19):6073

Mukhopadhyay A, Redding AM, Rutherford BJ, Keasling JD (2008) Importance of systems biology in engineering microbes for biofuel production. Curr Opin Biotechnol 19 (3):228

Nichols D, Nichols PD, McKMeekin TA (1995) A new n-C31:9 polyene hydrocarbon from Antarctic bacteria. FEMS Microbiol Lett 125 (2-3):281

Oku H, Kaneda T (1998) Biosynthesis of branched-chain fatty acids in Bacillus subtilis. J Biol Chem 263 (34):18386

Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K et al. (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. Nature 496 (7446):1

Patel KG, Welch M, Gustafsson C (2016) Leveraging Gene Synthesis, Advanced Cloning Techniques, and Machine Learning for Metabolic Pathway Engineering. In: Van Dien S (ed) Metabolic Engineering for Bioprocess Commercialization. Springer International Publishing, pp 53

Qiu Y, Tittiger C, Wicker-Thomas C, Le Goff G et al. (2012) An insect-specific P450 oxidative decarbonylase for cuticular hydrocarbon biosynthesis. Proc Natl Acad Sci U S A 109 (37):14858

Reed JR, Vanderwel D, Choi S, Pomonis JG et al. (1994) Unusual mechanism of hydrocarbon formation in the housefly: cytochrome P450 converts aldehyde to the sex pheromone component (Z)-9-tricosene and CO2. Proc Natl Acad Sci U S A 91 (21):10000

Rude MA, Schirmer A (2009) New microbial fuels: a biotech perspective. Curr Opin Microbiol 12 (3):274

Rude MA, Baron TS, Brubaker S, Alibhai M et al. (2011) Terminal olefin (1-alkene) biosynthesis by a novel P450 fatty acid decarboxylase from Jeotgalicoccus species. Applied and Environmental Microbiology 77 (5):1718

Rui Z, Li X, Zhu X, Liu J et al. (2014) Microbial biosynthesis of medium-chain 1-alkenes by a nonheme iron oxidase. Proc Natl Acad Sci U S A 111 (51):18237

Rumbold K, van Buijsen HJ, Overkamp KM, van Groenestijn JW et al. (2009) Microbial production host selection for converting second-generation feedstocks into bioproducts. Microb Cell Fact 8:64

Runguphan W, Keasling JD (2014) Metabolic engineering of Saccharomyces cerevisiae for production of fatty acid-derived biofuels and chemicals. Metab Eng 21:103

Schirmer A, Rude MA, Li X, Popova E et al. (2010) Microbial biosynthesis of alkanes. Science 329 (5991):559

Schrader J, Bohlmann J (2015) Biotechnology of Isoprenoids, vol 148. Advances in Biochemical Engineering/Biotechnology. Springer International Publishing,

Sheppard MJ, Kunjapur AM, Prather KL (2016) Modular and selective biosynthesis of gasoline-range alkanes. Metab Eng 33:28

Singleton C, Howard TP, Smirnoff N (2014) Synthetic metabolons for metabolic engineering. J Exp Bot 65 (8):1947

Sinha M, Weyda I, Sorensen A, Bruno KS et al. (2017) Alkane biosynthesis by Aspergillus carbonarius ITEM 5010 through heterologous expression of Synechococcus elongatus acyl-ACP/CoA reductase and aldehyde deformylating oxygenase genes. AMB Express 7 (1):18

Smirnova N, Reynolds KA (2001) Branched-chain fatty acid biosynthesis in *Escherichia coli*. Journal of Industrial Microbiology and Biotechnology 27 (4):246

Song X, Yu H, Zhu K (2016) Improving alkane synthesis in Escherichia coli via metabolic engineering. Applied microbiology and biotechnology 100 (2):757

Sorigue D, Legeret B, Cuine S, Morales P et al. (2016) Microalgae synthesize hydrocarbons from long-chain fatty acids via a light-dependent pathway. Plant Physiol 171 (4):2393

Spakowicz DJ, Strobel SA (2015) Biosynthesis of hydrocarbons and volatile organic compounds by fungi: bioengineering potential. Applied microbiology and biotechnology 99 (12):4943

Sugihara S, Hori R, Nakanowatari H, Takada Y et al. (2010) Possible biosynthetic pathways for all cis-3,6,9,12,15,19,22, 25,28-hentriacontanonaene in bacteria. Lipids 45 (2):167

Sukovich DJ, Seffernick JL, Richman JE, Gralnick JA et al. (2010a) Widespread head-to-head hydrocarbon biosynthesis in bacteria and role of OleA. Appl Environ Microbiol 76 (12):3850

Sukovich DJ, Seffernick JL, Richman JE, Hunt KA et al. (2010b) Structure, function, and insights into the biosynthesis of a head-to-head hydrocarbon in Shewanella oneidensis strain MR-1. Appl Environ Microbiol 76 (12):3842

Tillman JA, Seybold SJ, Jurenka RA, Blomquist GJ (1999) Insect pheromones - an overview of biosynthesis and endocrine regulation. Insect Biochem Molec 29 (6):481

Vickers CE, Blank LM, Kromer JO (2010) Chassis cells for industrial biochemical production. Nat Chem Biol 6 (12):875

Warui DM, Li N, Norgaard H, Krebs C et al. (2011) Detection of formate, rather than carbon monoxide, as the stoichiometric co-product in conversion of fatty aldehydes to alkanes by a cyanobacterial aldehyde decarbonylase. Journal of the American Chemical Society 133 (10):3316

Westfall PJ, Pitera DJ, Lenihan JR, Eng D et al. (2012) Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. Proc Natl Acad Sci U S A 109 (3):111

Winters K, Parker PL, Van Baalen C (1969) Hydrocarbons of blue-green algae: geochemical signfficance. Science 163 (3866):467

Xie M, Wang W, Zhang W, Chen L et al. (2017) Versatility of hydrocarbon production in cyanobacteria. Applied microbiology and biotechnology 101 (3):905

Yan H, Wang Z, Wang F, Tan TW et al. (2016) Biosynthesis of chain-specific alkanes by metabolic engineering in Escherichia coli. Engineering in Life Sciences 16 (1):53

Yoshino T, Liang Y, Arai D, Maeda Y et al. (2015) Alkane production by the marine cyanobacterium Synechococcus sp. NKBG15041c possessing the alpha-olefin biosynthesis pathway. Applied microbiology and biotechnology 99 (3):1521

Zhang L, Liang Y, Wu W, Tan X et al. (2016) Microbial synthesis of propane by engineering valine pathway and aldehyde-deformylating oxygenase. Biotechnol Biofuels 9:80

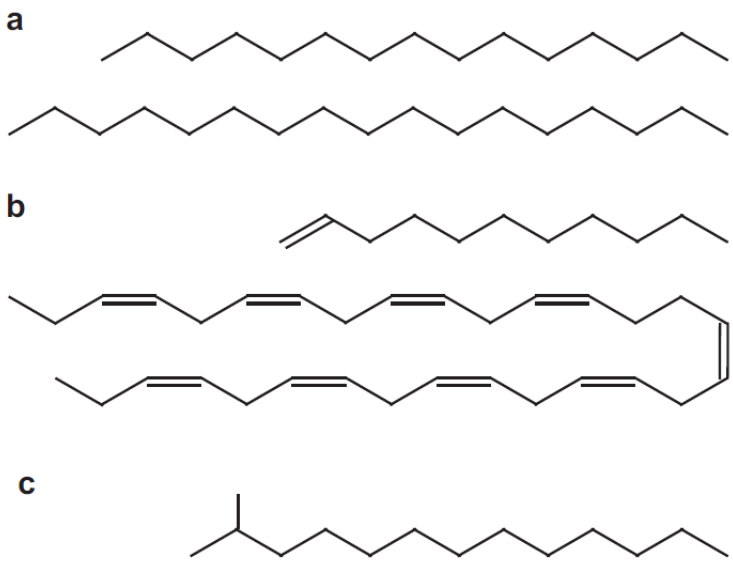

Fig. 1 Examples of alkane and alkene molecules that can be synthesized by microbes. (a) Pentadecane (*top*) and heptadecane (*bottom*) are produced by cyanobacteria (Schirmer et al. 2010). (b) Alkenes include 1-undecene (*top*) produced by *Pseudomonas* sp. (Rui et al. 2014) and the highly unsaturated very-long-chain alkene, 3,6,9,12,15,19,22,25,28-hentriacontanonaene (*bottom*) produced by *Shewanella* sp. (Sukovich et al. 2010). (c) Branched alkanes bearing a methyl group (e.g., methyl tridecene illustrated) can be produced by the creation of synthetic metabolic pathways (Howard et al. 2013)

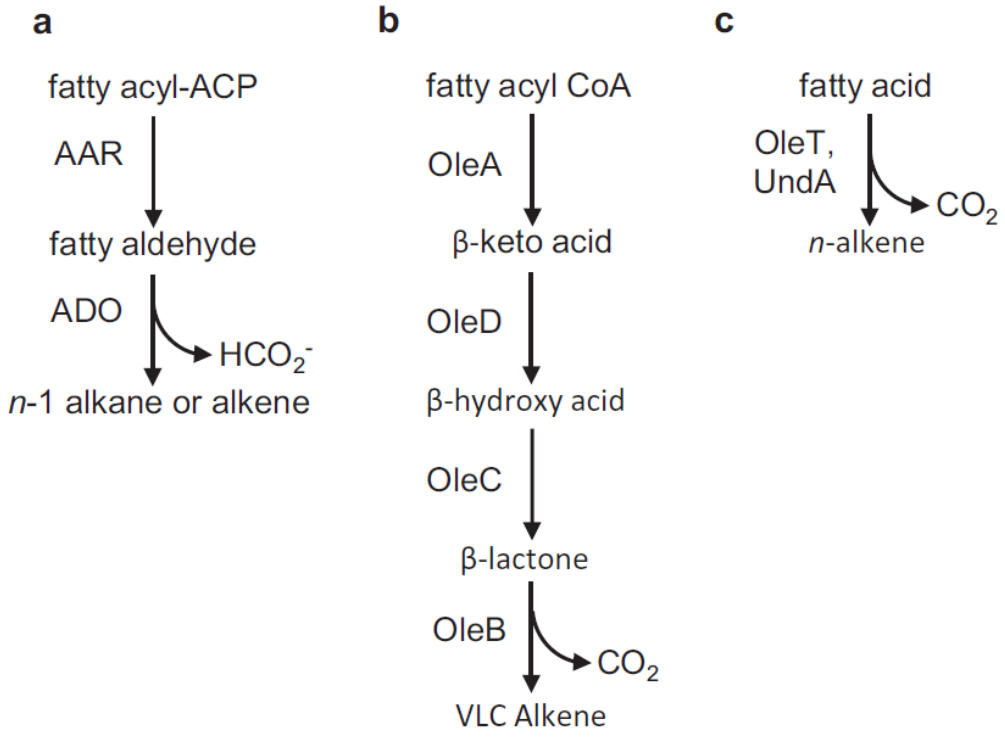

Fig. 2 Overview of bacterial alkane and alkene biogenesis. (a) Cyanobacterial conversion of fatty acyl-ACP molecules to C_{n-1} alkanes and alkenes via acyl-ACP reductase (AAR) and aldehyde- deformylating oxygenase (ADO) (Schirmer et al. 2010). (b) Biosynthesis of VLC alkenes by the *oleABCD* gene cluster enzymes OleA, OleD, OleC and OleB (Beller et al. 2010; Sukovich et al. 2010; Sugihara et al. 2010; Christenson et al. 2017b). (c) Biosynthesis of mid-chain length alkenes directly from fatty acids accomplished using either OleT (Rude et al. 2011) or UndA (Rui et al. 2014)

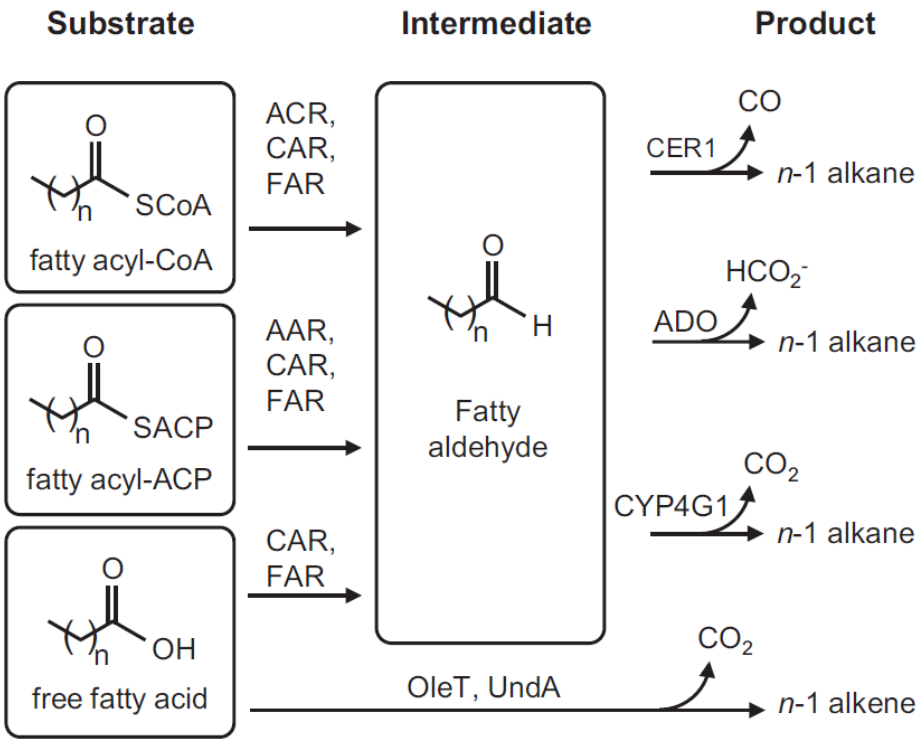

Fig. 3 Various substrates have been used as a starting point for alkane and alkene biosynthesis in genetically engineered pathways. These include fatty acyl-CoAs, fatty acyl-ACPs, and fatty acids. The enzymes that can exploit these substrates are a fatty acyl-CoA reductase (ACR) from *Acinetobacter* sp. (Yan et al. 2016), the carboxylic acid reductase (CAR) from *Mycobacterium marinum* (Akhtar et al. 2013), the fatty acid reductase (FAR) complex from *Photorhabdus luminescens* (Howard et al. 2013), OleT from *Jeotgalicoccus* sp. (Rude et al. 2011), and UndA from *Pseudomonas* sp. (Rui et al. 2014). With the exception of OleT and UndA, the other reactions generate a fatty aldehyde that can then be converted to an *n*-1 alkane or alkene using an aldehyde reductase from *Arabidopsis thaliana* (CER1) (Choi and Lee 2013), cyanobacterial aldehyde- deformylating oxygenase (ADO) (e.g., Akhtar et al. 2013; Howard et al. 2013), or CYP4G1 from *Drosophila melanogaster* (Qiu et al. 2012). Different permutations of these various steps have been successfully assembled, primarily in *E. coli* or *S. cerevisiae*