

REVIEW

Open Access



Directed evolution of hydrocarbon-producing enzymes

Jochem R. Nielsen¹, Joseph Kennerley^{1,2} and Wei E. Huang^{1*}

Abstract

Enzymes capable of catalysing the production of hydrocarbons hold promise for sustainable fuel synthesis. However, the native activities of these enzymes are often insufficient for their exploitation in industrial bioprocesses. Enzyme engineering approaches including directed evolution (DE) can be used to improve the properties of enzymes to meet desirable standards for their industrial application. In this review, we summarise DE methods for engineering hydrocarbon-producing enzymes, including both screening- and selection procedures. The efficacy of DE depends on several factors, including sensitive and accurate detection of enzyme activity, the throughput of screening or selection steps, and the scale of diversity generation. Although DE is a well-established approach, its application in engineering hydrocarbon-producing enzymes has not been widely demonstrated. This can be attributed to the physiochemical properties of the target molecules, such as aliphatic hydrocarbons, which can be insoluble, gaseous, and chemically inert. Detection of these molecules *in vivo* presents several unique challenges, as does dynamically coupling their abundance to cell fitness. We conclude with a discussion on future directions and potential advancements in this field.

Keywords Directed evolution, Hydrocarbons, Drop-in fuel, Enzyme engineering, Screening and selection, Growth-coupling

Introduction

Earth's visibly changing climate demands global attention and necessitates reduced greenhouse gas (GHG) emissions across diverse sectors. 20–25% of GHG emissions originate from the unsustainable combustion of hydrocarbon fuels in personal and commercial transport sectors [21]. While electric vehicles significantly reduce CO₂ emission and contribute to sustainable development, hydrocarbon fuels produced from carbon neutral processes still hold an advantage due to their high energy density. For example, hydrocarbon fuels remain the primary energy source for the aviation sector [20]. Over the

past decades, transportation fuels in some economies have been replaced to some degree with bioethanol and biodiesel; for example, Brazil, where biofuels represent approximately 25% of transport fuel consumption [22]. Although these positively contribute to reducing net emissions, the chemical composition of these fuels is not identical to fossil hydrocarbons, necessitating technical developments in engine technology, such as flex-fuel engines, to facilitate widespread adoption. Otherwise, such fuels must be blended with conventional fuels before use in most internal combustion engines (ICEs), reducing their scalability and resulting in the so-called “blend wall”. In addition, concerns about arable land requirements for biofuel production limit their attractiveness and scalability as fossil fuel alternatives.

The biological production of aliphatic hydrocarbons presents a promising solution to some of the longstanding issues in the biofuel sector. Firstly, the production of chemically identical hydrocarbons (‘drop-in’ fuels)

*Correspondence:

Wei E. Huang
wei.huang@eng.ox.ac.uk

¹ Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK

² Oxford SimCell Ltd, Begbroke Science Park, Begbroke OX5 1PF, UK



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

alleviates the blend wall issue, allowing either neat usage or significantly increased blend percentages. Secondly, microbial cell factories can, either natively or through genetic engineering, consume low-cost substrates including various waste streams, CO₂, or substrates derived from CO₂ such as acetate and formate. This can shift bio-based processes from the current mostly crop-derived substrates towards truly sustainable production [78], alleviating our reliance on a finite lithospheric carbon reservoir.

Natural microbial bioproduction of accessible hydrocarbons at industrially relevant titres, rates, and yields (TRY) is rare [40]. Further, most widely studied microbes produce no immediately useful fuel molecules, including gasses such as butane and propane (main components in liquified petroleum gas, LPG) or liquid C₈–C₁₆ n-alkanes (components of petrol and kerosene fuels) [110, 111]. This necessitates metabolic engineering to establish a biosynthesis pathway for the desired product, which ideally will require minimal subsequent processing for compatibility with existing fuel blends. Some product formation can often be achieved this way, but the resulting TRY metrics are usually far below those required for commercial exploitation. Fortunately, the modern synthetic biologist has a variety of tools available to them for targeted alteration of cell metabolism, enabling the application of engineering principles to cells to improve production. Although some standard techniques can be used to improve TRY by redirecting carbon flux towards product and balancing cofactor availability, sometimes the (terminal) enzymes involved directly in product synthesis remain limiting [4, 57].

In these cases, enzyme engineering approaches can be used to alter enzyme solubility, activity, (thermo) stability, or substrate specificity, potentially leading to improvements in production [65, 67]. Strategies including rational and semi-rational design can be used to target and alter specific regions or residues to achieve a desired change to an enzyme's properties. Rational design benefits from a comprehensive understanding of the enzyme's structure and catalytic mechanism, traditionally achieved by determining a crystal structure of the enzyme in question [45], [102]. A wide variety of enzyme classes have been successfully engineered using small rationally designed 'smart' libraries using structural information as input [14]. However, even with structural information success is not assured, rational enzyme engineering can lead to disappointing results when designed alterations have unexpected effects, such as a radical loss in stability or activity due to unforeseen interactions between residues. Besides this, the determination of an enzyme's crystal structure is laborious, expensive and, in some cases, cannot be realised at all.

The development of *in silico* protein structure prediction and protein design tools have vastly reduced our reliance on experimental protein structure determination, often representing the bottleneck to this approach. Considerable advancements in this field include the development of Rosetta- and, more recently, AlphaFold algorithms [39, 72]. AlphaFold, a machine-learning based protein structure prediction model, is capable of achieving unprecedented speeds and accuracies in predicting protein folding, enabling researchers to predict structures for almost all known proteins [89]. Although the volume of structural data generated by AlphaFold has become vast, gathering information concerning the relationship between a predicted enzyme structure and its biological performance—referred to as labelled data—remains a challenge and hinders machine-learning approaches [90, 98]. To address this challenge, advances are also being made in the prediction and improvement of protein functionality using unlabelled data [51]. In most cases however, to truly elucidate this relationship either a function must be experimentally validated and assigned to an enzyme structure, or a desired functionality must be identified and isolated from a pool of diverse enzyme structures, the latter of which is achieved through screening or selection.

Semi-rational enzyme engineering methods permit researchers a more naïve starting point in the engineering pipeline compared to full rational design. These techniques often rely on sequence-level data (as opposed to structural data) to help select residues to mutagenise, increasing the chance that a mutation has a functional effect and drastically reducing library sizes compared to methods using random mutagenesis. For example, phylogenetic datasets analysed through multiple sequence alignments (MSAs) can be interrogated to identify evolutionary 'hotspots' within sequences, uncover co-evolving residues, and generate ancestral protein sequences, the latter of which often display higher thermostability and altered substrate specificities [32, 61]. Combinations of (predicted) structural information and evolutionary sequence-level have also been successfully used to engineer a variety of enzyme properties [87].

An alternative approach that can function without any prior knowledge of an enzyme's structure or catalytic mechanism, and which can interrogate a much wider solution space for high-performing mutants, is random mutagenesis combined with directed evolution (DE), which uses iterative rounds of mutagenesis and screening or selection to identify enzymes with desirable traits from a large pool of variants [5]. To use this approach, the desired activity of the enzyme must be measurable (screenable), allowing high-performing variants to be screened and isolated. Alternatively, the target activity

must contribute to fitness (selectable) within a population of cells carrying enzyme variants, enabling those cells with desirable variants to become the population's dominant genotype [66]. This requires the development of a robust method to screen or select for desirable enzymatic activity, which in itself is the most challenging aspect of DE.

Screening and selection methods designed for DE can be used to identify desirable variants from the relatively small libraries often generated by (semi)-rational enzyme engineering methods. However, when random mutagenesis is used for diversity generation, then not only must the DE method be robust, but it must also allow for a high throughput. Random mutations have a vastly lower chance of being beneficial as opposed to neutral or detrimental, concomitantly requiring a vastly higher throughput to identify beneficial mutations in a population of variants [12]. While high throughput is required for any DE approach utilising random mutagenesis, establishing a DE method for hydrocarbon production faces some unique challenges compared to other classes of enzyme products, which we discuss in this review.

The application of directed evolution methods to engineer biocatalysts has been well described elsewhere [6, 13, 71, 105], as have advancements specifically in the field of in vivo directed evolution [26, 68, 82, 92]. In this review, we summarise the development and use of

DE approaches specifically for hydrocarbon-producing pathways and enzymes. Finally, we provide perspectives and suggestions for the further development of the field, hopefully accelerating the use of DE methods to establish commercially viable sustainable fuel production.

Hydrocarbon biosynthesis

Various hydrocarbon-producing enzymes have been identified in nature. Already since the 1960's waxes, alkanes and alkenes have been detected in insect, plant and cyanobacterial extracts. These hydrocarbons have disparate functions, ranging from protection against desiccation in insects and plants to promoting cell wall flexibility and salt tolerance in cyanobacteria [10, 35, 37, 46, 106]. Research efforts over the last two decades have led to the initial identification, characterisation, and application of several of the metabolic pathways responsible for hydrocarbon biosynthesis (Fig. 1). For the purposes of this review, we focus on bacterial hydrocarbon synthesis enzymes since many of these have been heterologously expressed and their catalytic activity characterised. In addition, we use the term 'hydrocarbons' to refer to aliphatic hydrocarbons (alkanes/alkenes), since these compounds are abundant in natural gas, gasoline, kerosene, and diesel, hence they have significant potential as drop-in fuels.

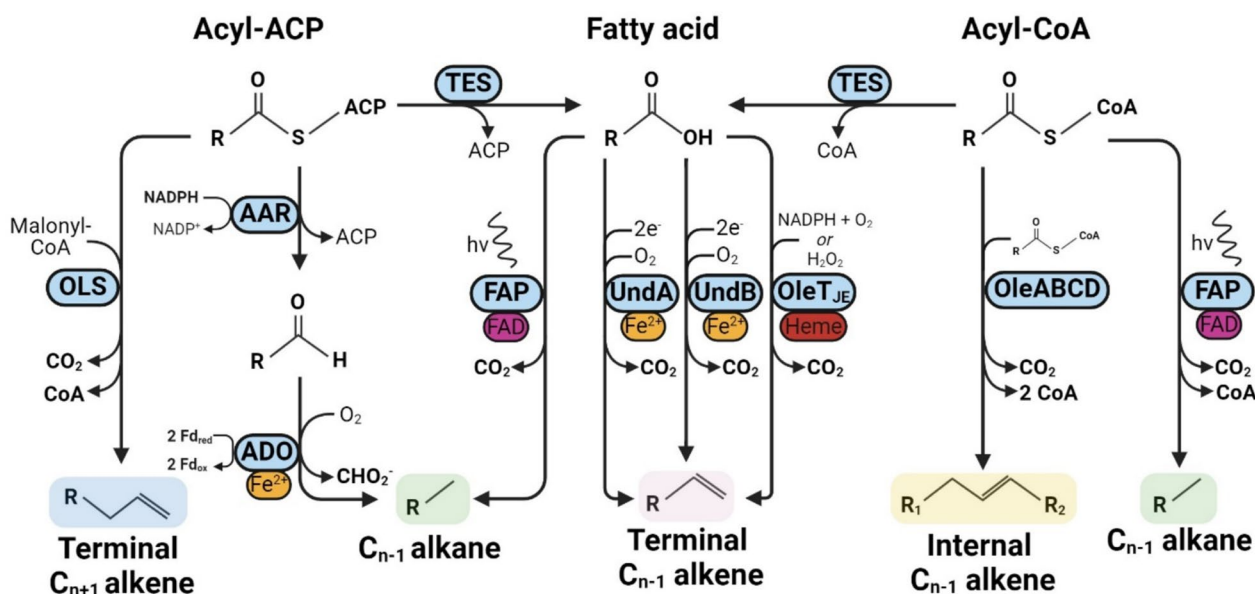


Fig. 1 Enzymatic pathways for hydrocarbon biosynthesis. All currently known biosynthetic pathways for alkene and alkane biosynthesis, abbreviations and native substrate specificities in terms of carbon chain length are indicated in parenthesis where relevant. Thioesterase (TES), olefin synthase (OLS, C₂₀), acyl-ACP reductase (AAR), aldehyde deformylating oxygenase (ADO, C₄-C₁₈), fatty acid photodecarboxylase (FAP, C₁₂-C₁₈), α -olefin synthases UndA (C₁₀-C₁₄)/UndB (C₆-C₁₇)/OleT_{JE} (C₄-C₂₀) and head-to-head fatty acid condensation enzymes OleABCD (C₂₅-C₃₂). Cofactors, coenzymes, prosthetic groups or metal ions required for catalysis are indicated including nicotinamide adenine dinucleotide phosphate (NADP⁺/NADPH), acyl carrier protein (ACP), coenzyme A (CoA), heme, Fe²⁺ and FAD (flavin adenine dinucleotide)

The constituent molecules of a drop-in compatible biofuel should be chemically similar to those found in a petroleum counterpart; primarily paraffins (n-alkanes, generally with carbon length C_4 – C_{12} for gasoline, C_9 – C_{25} for diesel, and C_5 – C_{15} for jet fuel), olefins (n-alkenes), isoparaffins (branched alkanes), naphthenes (cycloalkanes), and aromatics (mono- or polycyclic unsaturated hydrocarbons). n-Alkanes are a major component of transportation fuels and biologically derived alkanes can directly displace fossil-derived alkanes without impacting fuel quality or requiring changes to engine infrastructure, hence why biosynthetic pathways for producing n-alkanes are attractive candidates for optimisation and upscaling.

Aromatic fuel hydrocarbons—such as benzene, toluene, and xylene (BTX)—cannot yet be reliably produced directly using microbes, though biosynthesis of toluene has been observed and some key metabolic enzymes identified [9]. Derivatives of these aromatics, including phenol [96] and vanillin [63], have been biosynthesised, and directed evolution has been applied to biosensors for some derivatives [28], but there have been no significant attempts so far to use directed evolution to improve biological production of aromatic fuel hydrocarbons.

Terminal alkene biosynthesis

The first isolated enzyme responsible for terminal alkene biosynthesis was found to be a cytochrome P450 from the bacterium *Jeotgalicoccus* sp., dubbed OleT_{JE}, which catalyses the decarboxylation of fatty acids, producing alkenes [74]. Its substrate specificity was found to be extremely broad, enabling the production of short-, medium- and long chain C_{n-1} olefins (C_4 – C_{20}) from the respective fatty acids with concomitant release of CO_2 [24]. The OleT_{JE} enzyme also displayed flexibility in terms of its electron source, which can either be H_2O_2 without the need for oxygen or via a reduced electron carrier in the presence of oxygen [54].

Another terminal alkene biosynthetic pathway was discovered in *Synechococcus* sp. PCC7002, which was initially observed to produce C_{19} alkenes through an unknown mechanism [97]. Elucidation and characterisation of the pathway revealed that a multidomain protein named olefin synthase (*ols*), with homology to type I polyketide synthases, was responsible for C_{19} terminal alkene synthesis through an elongation and subsequent decarboxylation mechanism [59]. This cyanobacterium did not accumulate free fatty acids over C_{18} in length, which excluded the possibility that direct decarboxylation of free fatty acids was responsible for C_{19} alkene biosynthesis. Instead, the synthesis of C_{19} alkenes was found to occur through additional acyltransferase activity which

elongated a C_{18} fatty acid with malonyl-CoA, followed by a final decarboxylation step yielding a C_{19} product.

A reverse genetics approach was used to identify the enzymes responsible for 1-undecene biosynthesis in *Pseudomonas*. This revealed a mononuclear iron oxidase, UndA, which utilised O_2 and a reducing agent to decarboxylate fatty acids to corresponding terminal alkenes [76]. Further study later revealed it to constitute a new class of diiron decarboxylase [56]. The UndA enzyme could convert C_{10} – C_{14} fatty acids to their corresponding C_{n-1} 1-alkenes but was unable to utilise substrates with chain lengths $\leq C_8$ and $\geq C_{16}$, indicating a narrow substrate specificity range [76]. Upon further study another 1-undecene producing enzyme was found by the same group, dubbed UndB, which also catalysed the oxidative decarboxylation of fatty acids. This enzyme was found to have a broader substrate specificity than UndA, ranging from C_6 – C_{17} fatty acids, although with a preference for C_{10} – C_{14} [75]. In contrast to UndA, the UndB enzyme is membrane-bound, which has hindered biochemical analysis of its catalytic mode of action. However, it displays homology to other diiron desaturases, making it likely that UndB also contains a diiron core. The biosynthesis of 1-alkenes in yeast was achieved using UndB [110, 111], and it was found capable of catalysing decarboxylation of dicarboxylic acids [99].

Compared to OleT_{JE} and UndA, in vivo expression of the membrane-bound UndB enzyme results in higher production of 1-alkenes both in *E. coli* and *S. cerevisiae* [52]. Most engineering efforts have focussed on redox partner optimisation for the monooxygenase function of OleT_{JE} [24, 29], [54]. Only one recently published preprint describes the use of DE to evolve UndB [94]. The authors described using in silico guided residue selection and alanine scanning to identify UndB variants with improved substrate specificity towards C_9 – C_{14} dicarboxylic acids, achieving an approximate fivefold improvement in conversion of C_{10} dicarboxylic acid to 1,7-octadiene. Apart from this pre-print, no other work has reported using DE for the engineering of OleT_{JE}, Ols, UndA or UndB.

Internal alkene biosynthesis

The biosynthesis of internal alkenes occurs via the head-to-head condensation of fatty acids. The *oleABCD* genes from *Micrococcus luteus* were found to be responsible for biosynthesis of $C_{27:3}$ and $C_{29:3}$ alkenes, via consumption of two fatty acyl-CoA and the release of one CO_2 [8]. Soon after, more *ole*-type operons were discovered in a variety of bacteria. Some of these were characterised, revealing that a wide array of olefin products can be generated through this pathway, ranging from C_{24} – C_{31} in length and with up to 9 unsaturated bonds [86].

Alkane biosynthesis

Saturated alkanes are biosynthesised in plants—primarily as very-long-chain (VLC) alkanes which comprise cuticular waxes—most notably by ECERIFERUM 1 and 3 (CER1/CER3) in *Arabidopsis thaliana* [1, 11, 15]. In insects, an NADPH-dependent cytochrome P450 from the CYP4G family was shown to be a highly conserved enzyme involved in alkane biosynthesis [104]. Such enzymes, native to eukaryotes, are commonly expressed in yeast for functional characterisation [11], though specific applications for biofuel production have not been extensively explored. Eukaryotic enzymes can be difficult to express in prokaryotes, limiting their practical use in model bacteria like *E. coli*, though CER1 has been expressed in *Acinetobacter baylyi* as part of a two-stage process to convert CO₂ to acetate and then acetate to alkanes [48]. The carbon chain length specificity of plant and insect alkane biosynthesis enzymes (C₂₀–C₃₄) exceeds the range required for drop-in kerosene- or petroleum fuels, although some products may be suitable as diesel replacements. Attempts to use these enzymes for scalable bio-based production of these long-chain alkanes will have to contend with limited substrate solubility. The vast majority of research on alkane biosynthesis circumvents this challenge by focussing on using and engineering cyanobacterial alkane biosynthesis enzymes.

In 2010, a landmark study by Schirmer et al. reported the identification and characterisation of enzymes responsible for alkane biosynthesis in the cyanobacterium *Synechococcus elongatus* PCC7942 [77]. The proposed pathway consisted of an Acyl-ACP Reductase (AAR) and a terminal enzyme initially designated an aldehyde decarbonylase, owing to its supposed side-product being carbon monoxide. This was corrected in 2011, when the side-product of its reaction was found to be formate rather than CO, resulting in a renaming of the enzyme to Aldehyde Deformylating Oxygenase (ADO) [50, 95]. The AAR enzyme produces free fatty aldehydes using NADPH as an electron donor and fatty acyl-ACP as a preferred substrate, although it can also utilise acyl-CoA. The resulting fatty aldehyde product is then deformylated by ADO to produce a C_{n-1} alkane and formate as a by-product. The terminal ADO enzyme requires input of four electrons to perform the deformylation reaction, which are preferentially delivered by a ferredoxin (Fd)/ferredoxin reductase (FNR) system [109]. An FNR/Fd pair is therefore hypothesised to be the endogenous reducing system for ADO in cyanobacteria. Native ADO displays a wide substrate specificity ranging from C₄–C₁₈ aldehydes, and several studies have successfully improved the production of short-chain alkanes through rational enzyme- and/or

metabolic engineering [3, 19, 41, 43, 81]. For example, structure-based rational engineering of ADO led to an approximate doubling of propane titre [43]. Besides chain length specificity, other engineering attempts have yielded ADO variants with a 2- to 4-fold increased k_{cat}/K_M [93] and increased thermostability [80]. Having been a target of metabolic- and enzyme engineers for over a decade, maximum titres of the heterologously expressed AAR/ADO pathway have currently reached the g/L level in *E. coli* [16, 31]. However, production rates and yields still remain commercially infeasible: an earlier techno-economic study reported that biobased alkane production must achieve a theoretical yield of at least 90% to compare with ethanol in terms of cost efficiency [17]. Since most studies reported ADO to be significantly rate limiting it has become an attractive enzyme engineering target. Even so, only a handful of studies have used DE approaches to engineer ADO, which are described in later sections.

A more recently discovered alkane- and alkene-forming enzyme catalyses the light-driven decarboxylation of fatty acids or fatty acyl-CoAs [49, 83, 84]. This activity was exhibited by the *Chlorella variabilis* fatty acid photodecarboxylase (cvFAP), which displays a substrate specificity of C₁₂–C₁₈ fatty acids with a preference for C₁₆ and C₁₇ and requires blue light (~400–500 nm) for its activity. Rational engineering of the enzyme resulted in improved specificity towards short-chain product formation between 29 and 552-fold, allowing the light-driven generation of alkanes as short as ethane [103]. Only recently has the cvFAP enzyme first been engineered via DE, resulting in a comprehensive study where the enzyme was repurposed to enable highly selective decarboxylative radical cyclisation, achieving a 184-fold improvement in yield of cyclised product [38]. Despite the novel screening pipeline devised and used successfully in this work, product identification and quantification were performed via analytical techniques such as GC, GC–MS or HPLC. Consequently, the throughput of such an approach is still constrained and necessitates relatively small variant library sizes.

Although DE has not been widely used for engineering alkane biosynthesis enzymes, other work has instead focussed on alkane degradation. For example, iterative rounds of DE and screening were successfully used to engineer a P450_{BM3} enzyme towards a high specificity for propane monooxygenase activity, allowing the selective conversion of this gaseous n-alkane. This successful enzyme engineering campaign demonstrated that DE is a viable strategy for targeting the properties of alkane-interacting enzymes [30].

Challenges facing directed evolution of hydrocarbon-producing enzymes

Establishing a DE protocol varies in difficulty, depending mostly on the physical and chemical properties of the target product, and the extent to which the target enzyme has been characterised in terms of its activity and expression in the host cell. The detection and accurate quantification of a target product for DE is subject to several requirements (Table 1). Hydrocarbons satisfy few or none of these requirements, depending on their specific chemical composition. The most notable difficulty originates from the hydrophobicity of the substrates and/or products, and their ensuing poor solubility in physiological conditions. For example, if a certain hydrocarbon detection system relies on the compound being in solution, a poorly soluble product may not be detected, resulting in a false negative reading. Similarly, if the hydrocarbon product is gaseous, it may simply evaporate and no longer be available for detection in solution. A core requirement for directed evolution is the strong linkage between a detectable phenotype (e.g. enzyme-catalysed product formation) and the genotype of the responsible variant (its isolated sequence). The physicochemical properties of hydrocarbons can disrupt this linkage through evaporation, aggregation, and diffusion in and out of cells/vesicles. Hence, establishing a robust DE protocol for hydrocarbon biosynthesis is particularly challenging.

With these challenges in mind, attempts have been made to establish screening and selection protocols for hydrocarbon synthesis. The majority of these focus on detection and reporting of hydrocarbon production through a biosensor regulating expression of a reporter gene, although approaches involving selection instead

of screening are starting to be adopted. We summarise the development of both these approaches for the DE of hydrocarbon biosynthesis enzymes in the following sections.

Hydrocarbon screening methods

Chemical screening methods

The use of chemicals that result in a colorimetric or fluorometric change upon activity of a target enzyme represents a straightforward and intuitive screening method for directed evolution (Fig. 2B). One such approach was used in the DE of a P450 decarboxylase of the CYP152 family (CYP-Sm46Δ29), where OleT_{JE} was included as a reference enzyme, using a colorimetric screen for H₂O₂ consumption [101]. The OleT_{JE} enzyme can use H₂O₂ as a sole electron and oxygen source to catalyse fatty acid decarboxylation, and the authors exploited this principle by expressing CYP-Sm46Δ29 and OleT_{JE} mutant libraries in catalase-deficient *E. coli* to identify H₂O₂-consuming variants. This approach revealed a CYP-Sm46Δ29 variant with a 3.2-fold improved k_{cat}/K_M for lauric acid decarboxylation, although engineering of the OleT_{JE} enzyme itself did not yield any significant improvements.

Transcription factor-based reporter systems

Many DE approaches are based on measuring the abundance of a product of interest via a reporter system. By pinpointing genes that play a role in either catabolism or anabolism of the desired product, genetic elements involved in transcriptional responses to its presence can be identified and characterised. These genetic elements can then become part of a toolkit used in synthetic constructs to drive reporter gene expression [107, 108]. For example, such a construct could consist of a

Table 1 Properties of hydrocarbons in the context of directed evolution

Property	Ideal enzyme product for DE*	Hydrocarbons
Solubility	High Product is ideally highly soluble to enable interactions with proteins and/or other chemicals	Low Often insoluble in aqueous environments, can be gaseous or solid depending on chain length, tends to localise to the cell membrane or aggregate due to hydrophobicity
Compartmentalisation	High No extracellular leakage of the product. In case of cell lysis, released product cannot be taken up by living cells	Low Short chain hydrocarbons are volatile, medium- and long chain hydrocarbons can passively diffuse through cell membranes. Can result in poor linkage between enzyme performance and sequence
Interaction between product and host metabolism	Depending For screening purposes, interaction with host metabolism should be low. For selection, coupling product formation to host metabolism can be a useful strategy	Low For most model organisms used as screening or selection platforms, hydrocarbons are not a native metabolite
Toxicity	Low Ideally, the target product should not be toxic	Medium Short and medium-chain hydrocarbons can be toxic. Longer chain hydrocarbons exhibit low toxicity

* Assuming a DE system is used where intracellular enzyme product concentration correlates with product detection response (screening) or growth (rate) (selection)

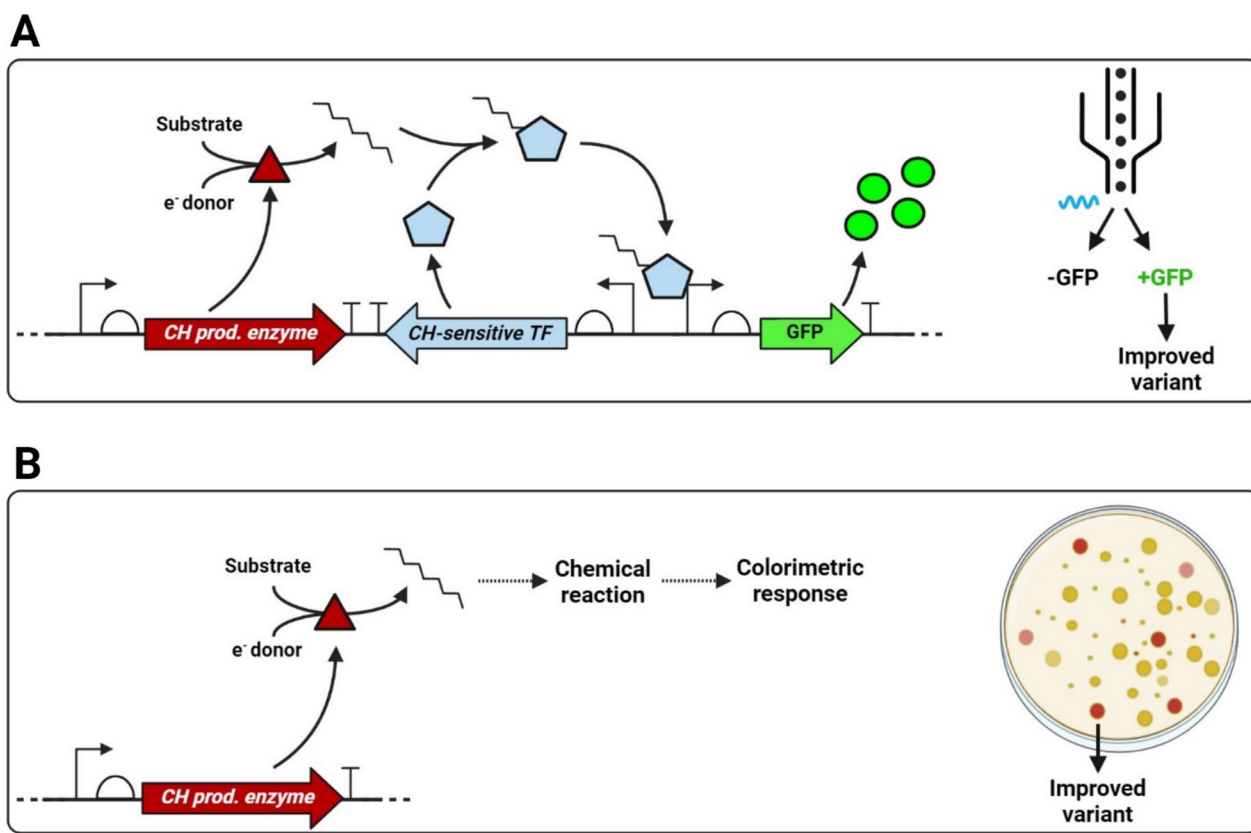


Fig. 2 Screening methods for hydrocarbon biosynthesis enzymes. **A** Schematic representation of a synthetic genetic construct consisting of a hydrocarbon production gene (CH prod. enzyme), hydrocarbon-detecting transcription factor (CH-sensitive TF) and a TF-responsive element, in this case a GFP reporter gene. The expression of GFP is linked to hydrocarbon presence, and GFP-expressing cells can be sorted by a FACS procedure resulting in enrichment for improved hydrocarbon biosynthesis enzymes. **B** A chemical reaction with either the product (hydrocarbon) or substrate(s) can be coupled to a colorimetric response to identify isolates with (improved) activity

hydrocarbon-sensitive transcription factor (TF) regulating the expression of a green fluorescent protein (GFP) (Fig. 2A).

Hydrocarbon-degrading microbes have been an abundant source of useful hydrocarbon-sensitive TFs with various affinities for a wide range of carbon chain lengths. For example, *Pseudomonas oleovorans* carries the alkS/ P_{alkB} system and *Acinetobacter baylyi* carries the alkR/ P_{alkM} system [69, 88], [107, 108]. Both systems exhibit comparable control architecture—activating expression of a native alkane monooxygenase catalysing degradation of environmental alkanes—and have been extensively characterised and repurposed for hydrocarbon detection. These systems have been used to detect hydrocarbon contamination in water using fluorescent or luminescent outputs [85], [107, 108], and to study cellular uptake of alkanes [33]. As pathways for biosynthesis of hydrocarbons are characterised and find use in bioproduction, so too have natural hydrocarbon biosensors been engineered to improve their sensitivity and specificity, in

pursuit of improved hydrocarbon quantification in vivo (see Table 2).

A recent study reported the first example of alkane-sensitive biosensors being successfully used in DE, in this case to engineer ADO [18]. In this work, the alkane-responsive TF AlkS was engineered to increase its sensitivity to longer-chain alkanes. Four rounds of random mutagenesis were performed using bidirectional fluorescence activated cell sorting (FACS) screening, where cells displaying low fluorescence in the absence of alkanes were isolated first, followed by the addition of alkanes and isolation of the top 1% of fluorescent cells. After optimising the biosensor, the authors randomly mutagenised the AAR and ADO genes over four rounds and identified isolates displaying a 13-fold improvement in alkane titre. This work highlights the importance of optimising the hydrocarbon-detecting TF before attempting a DE approach, and the successful engineering of AAR-ADO reported in this study demonstrates the potential of biosensor-based DE strategies.

Table 2 Transcription factor-based reporter systems for aliphatic hydrocarbon detection

Detector—reporter	Specificity	Dynamic range	Study findings	Reference
$P_{alkS}/alkR-P_{alkM}-GFP$	$C_{13} - C_{17}$	1–3 mg alkanes per g CDW	Improved sensor performance, combined two alkane response elements	[100]
$P_{alkB}/alkS^*-P_{alkS^-}-eGFP$	$C_{10} - C_{17}$	50–500 mg/L	Engineered AlkS for improved response range, DE of AAR and ADO	[18]
$P_{\Delta}alkS^*-P_{alkS^-}-GFP$	$C_5 - C_{10}$	Not indicated	Engineered AlkS for short-chain alkane sensitivity	[70]
$P_{alkRM}/alkR-P_{alkM^-}-GFP$ and P_{luxAB}	$C_{12} - C_{18}$	Detection of 40 mM alkanes, no range given	Monitoring aldehyde and alkane flux simultaneously	[47, 48]

Advantages and limitations of screening methods

A clear advantage of chemical screening systems is their simplicity. In addition, since they do not rely on biological parts for their functionality as is the case with TF-based screens, these methods are relatively robust, but also inflexible. A considerable disadvantage is their low throughput, as variants must usually be physically separated to discern differences in response. This limits many of these approaches to agar plate-based methods with a maximum throughput of $\sim 10^3$ – 10^4 variants per day, which is useful only for screening smaller libraries (e.g. smart libraries or site saturation mutagenesis libraries). In general, however, we expect that DE methods applied to randomly mutagenised libraries will move away from agar plate-based screening and towards higher-throughput methods.

An advantage of TF-based detection is that a response is generated based on direct interaction with the target molecule. It is also possible to swap out one biosensor system for another, for example if specificity towards a certain carbon chain length is preferred. Theoretically, hydrocarbon-responsive TFs can be engineered towards specificity for a wide range of chain lengths. It would therefore be possible to construct a library of TF variants with different specificities to a range of reaction products, enabling the swapping of detector modules to suit the needs of the DE campaign. Methods using TF-based detection could also be used to track the consumption of relevant substrates to infer biosynthesis rates of the product of interest. However, chain length specificity of the TF will be of similar importance as with product-detecting TFs, as discussed earlier. A TF-based system that detects the use of other molecules involved in the relevant reaction (such as electron donors) could also be envisaged, although interpretation of the sensor's response may become complex due to the ubiquitous and cyclical nature of intracellular redox reactions. That being said, TF-based systems that are specific and sensitive independently of the substrate or product's chain length would be highly

useful and make them similar to chemical detection methods.

However, this approach also has several limitations. Firstly, although interaction with the target hydrocarbon is direct, any chosen TF will not elicit an equal transcriptional response to all hydrocarbons. The output of the DE campaign is therefore defined by the specificity and sensitivity of the chosen TF. To improve the chances of a successful outcome, TFs will likely need to be engineered for maximum affinity toward the desired hydrocarbon molecule prior to their use in DE studies, as exemplified in recent work [18]. In addition, as discussed earlier, most hydrocarbons can freely diffuse into and out of the cell. This presents an issue for hydrocarbon detection-based screening methods, especially when using longer cultivation times, as non-producing cells can take up exogenous hydrocarbons, leading to increased false positive rates (so-called 'cheater' populations). Overcoming this requires either a) spatial separation of variants, vastly reducing potential throughput; b) rapid detection of improved variants before significant product leakage occurs or c) the use of counterselection methods to identify and remove cheating cell populations.

The use of biosensor-based approaches that are not coupled to growth relies on the availability of specialist equipment to perform high-throughput screening, such as FACS for single-cell quantification of fluorescence and sorting. If specialist equipment is not available, screening throughput has to be lowered several orders of magnitude to allow manual microtiter or agar plate screening. Increasing the throughput of screening methods therefore also increases procedure complexity and costs, lowering accessibility. Since detrimental mutations are far more common than desirable ones, enzyme screening campaigns inevitably result in the observation of mostly low-performing isolates. Whilst this could be valuable for gathering information about the enzyme's sequence-structure-function relationships, especially to generate labelled data for AI-based

rational mutagenesis approaches, it still means that most resources are spent on observing non- or poorly functioning variants.

Growth-coupled selection methods

Another DE strategy involves coupling the activity of the target enzyme to cell fitness. Host cells expressing enzyme variants with increased activity gain a growth advantage and outcompete their neighbours, mimicking the principles of natural selection to enrich high-performing variants within a diverse population. Selection methods are often based on the target enzyme activity rescuing a natural or engineered metabolic deficiency. When growth is dependent on the activity of an enzyme the enzyme is deemed to be growth-coupled (for a comprehensive overview of growth coupling principles (see [2]). Growth-coupled approaches offer a low-cost and procedurally simple method for enzyme engineering, as they consolidate both screening and selection steps into a single process operation. As growth (rate) is the output metric, populations of variants can be enriched for those conferring growth benefits using selective serial passages of the population, or by selective cultivation in turbidostats. The throughput of such approaches is usually only limited by the transformation efficiency of the host, which can be in the range of 10^8 – 10^{10} CFU/ μg DNA for *E. coli* [64], or by the target's mutation rate in the case of in vivo mutagenesis [34]. Since all successfully transformed hosts are simultaneously assessed for their fitness in selective conditions, the throughput of growth-coupled selection is amongst the highest of all existing enzyme engineering approaches, whilst only requiring minimal resources. One caveat to this is that the success rate of identifying a beneficial variant from a pool of competing variants can be limited by clonal interference [23].

Directed evolution by selection is only as good as the imposed selection pressure and its stringency [5]. It is therefore imperative that the selection pressure for such an approach is chosen carefully and extensively characterised. When growth is coupled to product synthesis, it is important to establish the relationship between output growth (rate) and the target product biosynthesis rate to establish a dynamic range of the selection platform. This information will be useful in avoiding either a maximisation of growth rate with unevolved enzyme activity, leaving no selective pressure for further improvement of the enzyme, or a situation where enzyme activity is too low to rescue growth to begin with. To avoid the first issue, concentrations of the growth-coupled enzyme can be reduced by swapping constitutive promoters to lower-strength variants, reducing induction levels of inducible promoters, or

by its fusion to a protein degradation tag. However, in most cases growth-coupled schemes are bottlenecked by the initial step of coupling growth solely to enzyme activity.

Hydrocarbon-producing enzymes vary in their activity, but most have catalytic activity far below those of strongly growth-coupled central metabolic enzymes [7]. For example, ADO is infamous for its sluggish activity, displaying turnover numbers of $\sim 1 \text{ min}^{-1}$ in vitro [57]. The light-driven fatty acid photodecarboxylase *cvFAP* was shown to be more efficient but still below the median turnover number of central and secondary metabolic enzymes, achieving a k_{cat} of $\sim 0.9 \text{ s}^{-1}$ [83]. Ensuring these relatively inefficient enzymes can rescue growth therefore requires their coupling to sensitive growth-rescue mechanisms, such as essential but low-flux metabolic pathways. Growth can also be made co-dependent instead of fully reliant on target enzyme activity, by supplying small amounts of the target product to help first establish growth. As the enzyme evolves towards higher activity this supplementation can be reduced until cell growth is fully dependent on the target activity. This approach of 'weaning' the cell off a rescue carbon source requires a comprehensive characterisation of the relevant dynamic range.

Hydrocarbons are a challenging class of molecules to target for DE approaches. Engineering cells to be dependent on hydrocarbon-producing enzymes therefore requires creatively designed selection methods, examples of which we have compiled below.

Transcription factor-based growth rescue

Hydrocarbon-sensitive transcription factors (TFs) have been widely used in the context of enzyme screening, as described earlier. These hydrocarbon-sensitive TFs can be viewed as modular hydrocarbon detection units. They do not necessarily have to be linked to expression of a reporter molecule, such as GFP, but can instead be employed to drive expression of genes essential for growth. This way, the functionality of these TF systems can be transformed from that of a screening method to a selection method (Fig. 3A).

These growth-essential enzymes may either be of direct metabolic or structural relevance, involved in processing of important central metabolites or structural components (e.g. TF-based restoration of an enzyme producing cell wall components), or they may rescue growth by alleviating an external stressor, for example by detoxifying some molecule added to the growth media. The gene circuit architecture is common to both approaches, but the rescue mechanism is different. Notably, the latter approach allows for easily adjustable stringency of selection by varying the concentration of the toxic additive.

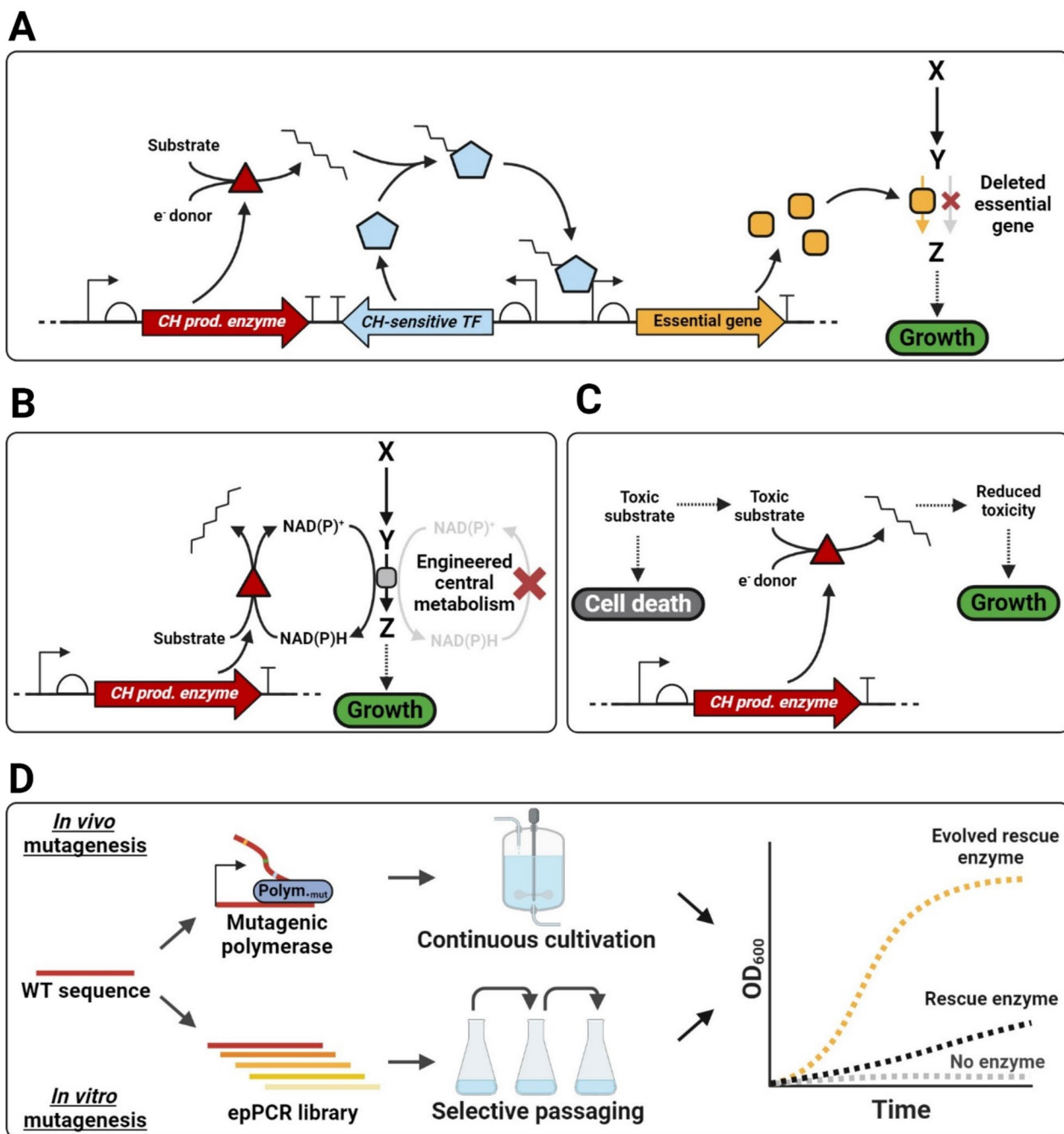


Fig. 3 Selection methods for hydrocarbon-producing enzymes. **A** Schematic representation of a synthetic genetic construct consisting of a hydrocarbon production gene (CH prod. enzyme), hydrocarbon-detecting transcription factor (CH-sensitive TF) and a TF-responsive element. The responsive element consists of an essential gene to enable hydrocarbon-dependent essential gene expression. **B** Schematic of the principle of cofactor-based growth coupling. The microbial selection platform can be engineered to exhibit a deficiency in the oxidation or reduction of redox cofactors. These microbes are unable to grow unless a rescue reaction is introduced restoring redox cofactor cycling, in this example through the NAD(P)H-dependent reduction of substrate to hydrocarbon. **C** A toxic concentration of substrate can be used arrest growth or induce cell death unless the substrate is converted to the target hydrocarbon compound, resulting in growth and selection for enzyme variants with higher activity. **D** Evolution using these systems can be performed *in vivo* using continuous mutagenesis tools and continuous cultivation (top). Alternatively, more traditional methods such as error-prone PCR (epPCR) can be used with selection by serial passaging (bottom). Target enzyme activity is directly coupled to growth, so evolved enzymes confer a higher growth rate to the selection strain

This was exemplified in a prior study detailing the use of TF-based detection of small molecules to establish a growth-coupled system [25]. The authors coupled a TF sensitive to 1-butanol to the expression of tetracycline transporter gene *tetA*. The resulting strain could grow in tetracycline-containing medium only when supplemented with 1-butanol, or when 1-butanol was biosynthesised in vivo. This system was then used to engineer the RBS sequences of a mixed alcohol biosynthesis pathway. Although this work did not target aliphatic hydrocarbon synthesis, the underlying principle of the work demonstrates its viability.

Another approach requiring more complex gene circuitry was used to achieve a similar system with tuneable stringency. A convergently oriented inducible promoter system was constructed to modulate expression of a TF-regulated rescue gene by direct transcriptional interference, where the rescue gene was responsible for synthesis of cell wall components [42]. This system was used to select for alkane/alkene biosynthesis in *Acinetobacter baylyi*, using its native alkane-sensitive $\text{alkR}/P_{\text{alkM}}$ transcriptional regulator and promoter system [42]. The *A. baylyi* strain was first made auxotrophic for the essential amino acid lysine by knocking out 4-hydroxy-tetrahydropicolinate reductase (*dapB*). This gene was then re-integrated into the chromosome but under control of the alkane-sensitive $\text{alkR}/P_{\text{alkM}}$ system, in place of the native alkane monooxygenase *alkM*, which was removed to prevent degradation of hydrocarbon product.

Growth of this strain could be rescued either by direct supplementation of lysine in the medium, or by addition of mineral oil, inducing expression of the essential gene and thereby complementing growth. Using this system, a mixed population of *A. baylyi* strains transformed with low-, medium-, and high-producing alkane biosynthesis plasmids were shown to be enriched for the high alkane production plasmids over time. When assessed separately, strains with high-producing alkane biosynthesis plasmids displayed increased growth rates and decreased lag phases compared to those transformed with low-producing alkane production plasmids.

This study demonstrated that TF-based hydrocarbon detection could also be used to establish a selection system for hydrocarbon production, as opposed to just a screening method. Due to the growth-based selection principle only a selective cultivation of the strain would be required to achieve population enrichment of positive clones, considerably simplifying the process compared to a screening method. This approach does still suffer from some of the same limitations as TF-based screening approaches, including potential cross-feeding of product in the population and the dependence on TF specificity and genetic circuit integrity.

Cofactor-based growth coupling

A relatively new concept in the field of growth-coupled enzyme engineering is to use redox cofactor auxotrophy as a selection method [62]. In this approach, strains unable to grow due to an imposed deficiency in the reduction or oxidation of a certain cofactor pair (e.g. NAD^+/NADH or $\text{NADP}^+/\text{NADPH}$) can be used to select for enzymes that alleviate this redox imbalance. This method offers a way to couple growth to redox cofactor-requiring enzymes or pathways, without having to use the target product as the selective molecule directly (Fig. 3B). Since redox cofactors are the selected molecules in this approach, issues caused by low solubility and poor compartmentalisation of hydrocarbon products can be avoided.

Aliphatic hydrocarbons can be biosynthesised through carbon chain-elongating steps in metabolism, such as fatty acid biosynthesis or reverse beta oxidation. These pathways involve several reductive reactions requiring significant input of NAD(P)H , making them viable targets for cofactor-based growth-coupling using strains deficient in the cycling of NAD(P)H back to their oxidised counterparts NADP^+ or NAD^+ . Several hydrocarbon biosynthesis enzymes also require cofactors, such as AAR (NADPH , directly), ADO (NAD(P)H , ferredoxin-mediated) and OleT_{JE} (NAD(P)H , cytochrome P450 reductase). Redox cofactor auxotrophs have already been used to engineer a 3-hydroxy-acyl-CoA reductase [55], carboxylic acid reductase [44], P450 cytochrome [58] and a nitroreductase [79], to name several. The breadth of enzyme activity targetable by this DE approach makes it a promising technology for engineering hydrocarbon-producing enzymes.

Although this approach has not been used to engineer hydrocarbon production so far, we expect cofactor-based growth coupling to become a more widely used tool in the DE repertoire due to its unique product-independent property and improved resilience against inactivating mutations in the growth-coupling gene circuit.

Substrate toxicity alleviation

Instead of selecting for the biosynthesis of a target product, the consumption of a substrate can also be linked to cell fitness if the substrate is toxic and the product less so. Consumption of the toxic substrate by the target enzyme alleviates its toxic effects, either reducing the growth lag phase of the host cell, improving its growth rate, or a combination of the two (Fig. 3C).

Although this approach has a limited scope, since the target enzyme must have a toxic substrate and a less toxic product, it provides a direct way to link enzyme activity to growth, making the gene circuit less prone to inactivation by the host. One study has successfully

employed this approach to engineer aldehyde deforming oxygenase (ADO). The aldehyde substrate for ADO is toxic to *E. coli* depending on concentration and carbon chain length. The authors exploited this property by adding inhibitory concentrations of aldehydes to *E. coli* cultures, where they found that ADO-expressing cells exhibited improved growth [53]. After a round of random mutagenesis, the authors identified an *E. coli* protease degradation target on the C-terminus of the protein, which when removed or protected by a His-tag significantly improved the concentration of soluble ADO protein in vivo. The authors used an engineered strain deficient in several native aldehyde reductases to ensure aldehyde consumption by the host was limited. This approach lends itself well to enzymes that catalyse the conversion of toxic substrates that are poorly, or not at all, degraded to non-toxic products by the host.

A second relevant study, although not focussing on aliphatic hydrocarbon production, demonstrated that the detoxification of medium-chain fatty acids can be used as a selective principle [36]. In this work, the authors engineered carboxylic acid reductases for improved activity against medium-chain fatty acids (MCFA), to generate the concomitant alcohols. One aspect of the work echoes the previous ADO-targeting study exploiting aldehyde toxicity, namely that endogenous pathways capable of detoxifying the substrate were removed. For example, deletions of a MCFA exporter and an aldehyde dehydrogenase were performed to ensure accumulation of alcohol product. Considering several hydrocarbon-forming enzymes use fatty acids directly as a substrate, this approach warrants further exploration.

For many hydrocarbon-producing enzymes, their substrate specificity in terms of carbon chain length is a major focal point for enzyme engineering. Substrate toxicity alleviation provides an attractive method to customise selective pressure towards specificity for a certain carbon chain length.

Advantages and limitations of hydrocarbon selection methods

The use of selection methods for engineering hydrocarbon enzymes offers a way to perform ultra-high throughput DE with minimal requirements for complex equipment and delicate, often robotics-assisted, procedures. These methods are therefore highly attractive for applying selection to large combinatorial variant libraries, performing continuous mutagenesis, and evolving whole-cell systems towards improved product synthesis.

While the advantages are considerable, growth-coupled selection methods are challenging to establish. Some level of metabolic engineering is almost certainly required, and extensive characterisation of the resulting metabolic

deficiencies must be performed before a DE campaign can be initiated. In addition, the catalytic rate of hydrocarbon biosynthesis enzymes is lower on average than essential enzymes in central metabolism [7]. Establishing growth-coupling solely to target enzyme activity therefore requires careful planning regarding which essential metabolic node to target, since this defines the dynamic range within which the selective pressure will operate. Establishing selection based on essential gene expression may therefore be a useful alternative to the direct metabolic coupling of enzyme activity, since expression strength can also be modulated through other means such as RBS swapping or using multi-copy constructs.

Besides establishing a selection method, ensuring the required selective pressures remain in place during DE procedures can be challenging as growth is often severely inhibited. Fundamentally, even with the most sophisticated growth-coupling methods, the researcher's pursuit of "productivity" is imperfectly aligned with the host microbe's pursuit of "survival", the latter being achievable either by developing along prescribed evolutionary routes toward hydrocarbon production or by simply circumventing the genetic circuits put in place to encourage such evolution. Unlike lower-throughput DE methods leveraging external screening and selection machinery, growth-coupling DE methods rely more on constructed genetic circuits which are open to manipulation by the host organism. Preserving regulatory gene modules governing rescue gene expression against mutation, especially in the face of stringent selection pressure over many generations, is a major obstacle to development of this DE approach. Switchable expression by a hydrocarbon-sensitive TF can easily mutate to become constitutive, restoring rescue genes and returning the host to a high-growth phenotype independent of hydrocarbon productivity. Escape mutants disrupting the DE process will assuredly appear, requiring either an extremely stringent selection method incorporating contingencies for repair of compromised genetic circuits, or a counterselection method to combat cheating populations [27].

Future perspectives

Directed evolution methods for hydrocarbon-producing enzymes are challenging to establish, often requiring preliminary genetic and protein engineering to establish high-throughput screening and selection procedures. Much ground has been covered in recent years, but several research areas remain unapproached.

The sensitivity, specificity and robustness of hydrocarbon biosensors has steadily improved over the past decade, with recent work demonstrating its potential for engineering the AAR-ADO pathway [18]. Despite these improvements, the majority of studies still apply them

solely for fluorescent reporter gene expression, even though this approach requires specialised equipment for repeat rounds of external screening and sorting, precluding its use for continuous evolution. We envision these biosensors being used in growth-coupled designs, where the detection of hydrocarbons is coupled to the expression of gene (circuits) linked to growth. Although this would require more extensive initial strain engineering and characterisation, these growth-coupled selection methods can interrogate a vastly larger solution space of variants, while also offering reduced operational complexity and equipment requirements, potentially accelerating the discovery of improved enzyme/pathway variants.

The use of substrate toxicity as a selection method for aliphatic hydrocarbon production has only been used to engineer ADO [53], but other potential targets for this approach are fatty acid decarboxylating enzymes such as OleT_{JE}, UndA, UndB and *cvFAP*. Short- and medium-chain fatty acids are toxic to *E. coli* and *S. cerevisiae* above certain concentrations depending on their chain length [36, 73], meaning that enzymes catalysing their conversion to less toxic alkanes or alkenes could potentially be engineered using this approach. Efficient production of gaseous hydrocarbons such as propane and butane is also highly sought after, but performing DE using the gasses themselves as the target enzyme product is challenging due to a lack of compartmentalisation (assuming that their intracellular concentration determines the DE system's response). Alleviating short-chain fatty acid toxicity as the selective principle may enable easier engineering of these enzymes. When using *E. coli* as a host for enzyme selection, exogenous fatty acids can also be easily excluded from cell metabolism by deletion of *fadD*, avoiding their activation to acyl-CoA and subsequent degradation.

Studies focussed on engineering hydrocarbon biosynthesis have mostly reported either host-level metabolic engineering (e.g. carbon flux improvements, redox balancing) or rational mutagenesis and screening specifically targeting only terminal hydrocarbon synthesis enzymes. However, when hydrocarbons are synthesised from a sole carbon source (i.e. the substrate of the terminal enzyme is not supplemented), some of the DE methods set out in this review could be used to evolve entire metabolic pathways and other cellular processes towards improved hydrocarbon synthesis. Besides improving bioprocess performance using common substrates such as glucose, growth-coupled designs enable the simultaneous engineering of product biosynthesis and substrate catabolism. This can be exploited to evolve cell factories towards optimal assimilation and conversion of CO₂-derived carbon sources. In vivo continuous targeted hypermutation

[60] or whole-genome mutagenesis tools [91] represent suitable technologies for performing this type of evolutionary engineering towards improved product synthesis, though exempting from mutation the crucial gene circuitry responsible for growth-coupling remains an obstacle. Consequently, in vivo mutagenesis coupled with a high-throughput DE approach has not been performed in the context of hydrocarbon biosynthesis.

This review focussed on novel methods to engineer (terminal) enzymes involved in hydrocarbon biosynthesis, which we consider a crucial bottleneck for maturation of the field. Indeed, full exploitation of the screening and selection methods described in this review may offer considerable benefits over currently used strategies. It is acknowledged, however, that in the context of hydrocarbon biosynthesis, there may be physiochemical limitations hindering the success of enzyme engineering campaigns. For example, the often poor solubilities exhibited by substrates and products of hydrocarbon biosynthetic pathways may fundamentally limit catalytic performance, at least when the enzyme(s) of choice is/are utilised in vivo. Strategies including the use of cell-free systems with optimised media for hydrocarbon biosynthesis could remove some of these limitations. Alternatively, bespoke reactor design could ensure the reaction environment is better suited for hydrocarbon biosynthesis through the use of pressurisation or biphasic production processes. These and other strategies could help ensure that the enzyme fitness landscape being traversed during directed evolution remains free of obstacles.

Conclusion

Living systems are undoubtedly the most sophisticated manufacturing platforms on earth; able to adapt, self-repair, and transform diverse substrates into complex structures under mild conditions with unmatched molecular precision. But the sheer physiological complexity of living systems means biological engineers are typically operating with incomplete knowledge. An engineering approach able to rapidly explore a broad solution space, while avoiding the need for extensive knowledge of all factors at play, represents a powerful tool for optimising such a system. Directed evolution is such a tool; a massively parallel black-box approach to the traditional engineering design-build-test-learn cycle, attempting to emulate and expedite a natural strategy that can be observed in our living world.

Biosynthesis of drop-in compatible hydrocarbon fuels is a promising field, but has struggled to establish processes with high titres, rates, and yields. Even though several hydrocarbon-producing enzymes and their related pathways have been identified and extensively characterised over the past decade and a half, the

limited activities of these enzymes make them unsuited to commercial use. A considerable bottleneck to the improvement of hydrocarbon-producing enzymes and, more holistically, industrially useful hydrocarbon-producing strains, is a lack of appropriate screening or selection methods to identify improved variants in a high-throughput manner. While directed evolution seems well-suited to tackle these problems, designing directed evolution strategies for hydrocarbon synthesis is challenging and requires creative implementation of multiple facets of synthetic biology. In this review, we have summarised the development of directed evolution methods for hydrocarbon production and discussed their advantages and limitations.

The most widely reported directed evolution systems utilise a screening method based on a hydrocarbon-sensitive genetic construct driving expression of a fluorescent reporter protein. Although these systems initially displayed low selectivity and sensitivity, considerable improvements have been made over the years. These efforts have resulted in screening systems that can now detect alkanes with broader chain length specificity, while also exhibiting improved dynamic range and lower leakiness. As a result, recent studies have reported the application of these tools for the directed evolution of ADO, achieving promising improvements in activity.

A relatively unexplored method for engineering hydrocarbon production is directed evolution by selection. Many of the limitations of screening approaches can be circumvented using selection methods, but these may require more extensive genetic engineering and characterisation to establish. The selection methods summarised in this work include selection by substrate toxicity alleviation, cofactor-based growth coupling, and hydrocarbon-sensitive transcription factors driving expression of essential genes. Besides their use as enzyme engineering platforms, strains displaying a growth dependence on hydrocarbon biosynthesis may be valuable assets in the context of bioprocess stability and antibiotic-free biomanufacturing. Despite these attractive features, the body of work encompassing growth-based selection for hydrocarbon biosynthesis is minimal and remains an area with high potential.

In summary, the field of drop-in hydrocarbon biosynthesis is steadily moving from discovery and characterisation of useful genetic parts, through to metabolic and protein engineering efforts to improve the useful function of said parts, towards leveraging the principles of directed evolution to accelerate the traditional engineering design-build-test-learn cycle. By fully utilising the continuously expanding toolbox of synthetic biology, screening and selection of hydrocarbon-producing

enzymes and organisms will be made increasingly accessible. Hopefully, this will accelerate the development of novel enzyme variants and new, industrially relevant microbial strains that can help drive sustainable hydrocarbon synthesis towards commercial viability.

Acknowledgements

This work was supported by the BBSRC (grant BB/M011224/1) and the Oxford Interdisciplinary Bioscience Doctoral Training Partnership. W.E.H. thanks EPSRC (EP/M002403/1 and EP/N009746/1) for financial support. Figures in this article were produced using Biorender.com.

Author contribution

JRN and JK wrote the main manuscript text and prepared the Tables and Figures. WEH supervised the work. All authors reviewed the manuscript.

Funding

This research was funded by Biotechnology and Biological Sciences Research Council (Grant No. BB/M011224/1, EPSRC, EP/M002403/1)

Data availability

No datasets were generated or analysed during the current study.

Declarations

Competing interests

The authors declare no competing interests.

Received: 25 January 2025 Accepted: 24 July 2025

Published online: 12 August 2025

References

1. Aarts MGM, Keijzer CJ, Stiekema WJ, Pereira A. Molecular characterization of the CER1 gene of Arabidopsis involved in epicuticular wax biosynthesis and pollen fertility. *Plant Cell*. 1995;7(12):2115. <https://doi.org/10.2307/3870155>.
2. Alter TB, Ebert BE. Determination of growth-coupling strategies and their underlying principles. *BMC Bioinform*. 2019;20(1):1–17. <https://doi.org/10.1186/S12859-019-2946-7>.
3. Amer M, Toogood H, Scrutton NS. Engineering nature for gaseous hydrocarbon production. *Microb Cell Fact*. 2020;19(1):1–14. <https://doi.org/10.1186/S12934-020-01470-6>.
4. Andre C, Kim SW, Yu XH, Shanklin J. Fusing catalase to an alkane-producing enzyme maintains enzymatic activity by converting the inhibitory byproduct H₂O₂ to the cosubstrate O₂. *Proc Natl Acad Sci*. 2013;110(8):3191–6. <https://doi.org/10.1073/pnas.1218769110>.
5. Arnold FH. Design by directed evolution. *Acc Chem Res*. 1998;31(3):125–31.
6. Arnold FH. Directed evolution: bringing new chemistry to life. *Angew Chem (International Ed in English)*. 2017;57(16):4143. <https://doi.org/10.1002/anie.201708408>.
7. Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS, Milo R. The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry*. 2011;50(21):4402–10. <https://doi.org/10.1021/BI2002289>.
8. Beller HR, Goh EB, Keasling JD. Genes involved in long-chain alkene biosynthesis in *Micrococcus luteus*. *Appl Environ Microbiol*. 2010;76(4):1212–23. <https://doi.org/10.1128/AEM.02312-09>.
9. Beller HR, Rodrigues AV, Zargar K, Wu YW, Saini AK, Saville RM, Pereira JH, Adams PD, Tringe SG, Petzold CJ, Keasling JD. Discovery of enzymes for toluene synthesis from anoxic microbial communities. *Nat Chem Biol*. 2018;14(5):451–7. <https://doi.org/10.1038/s41589-018-0017-4>.
10. Berla BM, Saha R, Maranas CD, Pakrasi HB. Cyanobacterial alkanes modulate photosynthetic cyclic electron flow to assist growth under cold stress. *Sci Rep*. 2015;5(1):1–12. <https://doi.org/10.1038/srep14894>.

11. Bernard A, Domergue F, Pascal S, Jetter R, Renne C, Faure J-D, Haslam RP, Napier JA, Lessire R, Joubès J. 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*. 2012;24(7):3106–18. <https://doi.org/10.1105/TPC.112.099796>.
12. Bloom JD, Arnold FH. In the light of directed evolution: pathways of adaptive protein evolution. *Proc Natl Acad Sci*. 2009;106(supplement-1):9995–10000. <https://doi.org/10.1073/pnas.0901522106>.
13. Bommaris AS, Paye MF. Stabilizing biocatalysts. *Chem Soc Rev*. 2013;42(15):6534–65. <https://doi.org/10.1039/C3CS60137D>.
14. Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, Robins K. Engineering the third wave of biocatalysis. *Nature*. 2012;485(7397):185–94. <https://doi.org/10.1038/nature11117>.
15. Bourdenx B, Bernard A, Domergue F, Pascal S, Léger A, Roby D, Pervent M, Vile D, Haslam RP, Napier JA, Lessire R, Joubès J. Overexpression of Arabidopsis ECERIFERUM1 promotes wax very-long-chain alkane biosynthesis and influences plant response to biotic and abiotic stresses. *Plant Physiol*. 2011;156(11):29–45. <https://doi.org/10.1104/PP.111.172320>.
16. Cao YX, Xiao WH, Zhang JL, Xie ZX, Ding MZ, Yuan YJ. Heterologous biosynthesis and manipulation of alkanes in *Escherichia coli*. *Metab Eng*. 2016;38:19–28. <https://doi.org/10.1016/j.ymben.2016.06.002>.
17. Caspeta L, Nielsen J. Economic and environmental impacts of microbial biodiesel. *Nat Biotechnol*. 2013;31(9):789–93. <https://doi.org/10.1038/nbt.2683>.
18. Chen D, Xu S, Li S, Tao S, Li L, Chen S, Wu L. Directly evolved AlkS-based biosensor platform for monitoring and high-throughput screening of alkane production. *ACS Synth Biol*. 2023;12(3):832–41. <https://doi.org/10.1021/ACSSYNBIO.2C00620>.
19. Choi YJ, Lee SY. Microbial production of short-chain alkanes. *Nature*. 2013;502(7472):571–4. <https://doi.org/10.1038/nature12536>.
20. IPCC. Climate change 2022: mitigation of climate change. 2022. <https://www.ipcc.ch/report/ar6/wg3/>
21. CO2 Emissions in 2022—analysis—IEA. 2023. <https://www.iea.org/reports/co2-emissions-in-2022>
22. da Costa AO, dos Santos LBO, Henriques RM, Araujo RB. Implementation of bioenergy in Brazil - 2021 update. 2021. https://www.ieabioenergy.com/wp-content/uploads/2021/11/CountryReport2021_Brazil_final.pdf
23. De Visser JAGM, Rozen DE. Clonal interference and the periodic selection of new beneficial mutations in *Escherichia coli*. *Genetics*. 2006;172(4):2093. <https://doi.org/10.1534/GENETICS.105.052373>.
24. Dennig A, Kuhn M, Tassoti S, Thiessenhusen A, Gilch S, Bülter T, Haas T, Hall M, Faber K. Oxidative decarboxylation of short-chain fatty acids to 1-alkenes. *Angew Chem Int Ed*. 2015;54(30):8819–22. <https://doi.org/10.1002/ANIE.201502925>.
25. Dietrich JA, Shis DL, Alikhani A, Keasling JD. Transcription factor-based screens and synthetic selections for microbial small-molecule biosynthesis. *ACS Synth Biol*. 2013;2(1):47–58. <https://doi.org/10.1021/SB30091D>.
26. d'Oelsnitz S, Ellington A. Continuous directed evolution for strain and protein engineering. *Curr Opin Biotechnol*. 2018;53:158–63. <https://doi.org/10.1016/j.copbio.2017.12.020>.
27. d'Oelsnitz S, Kim W, Burkholder NT, Javanmardi K, Thyer R, Zhang Y, Alper HS, Ellington AD. Using fungible biosensors to evolve improved alkaloid biosyntheses. *Nat Chem Biol*. 2022;18(9):981–9. <https://doi.org/10.1038/s41589-022-01072-w>.
28. Du H, Liang Y, Li J, Yuan X, Tao F, Dong C, Shen Z, Sui G, Wang P. Directed evolution of 4-hydroxyphenylpyruvate biosensors based on a dual selection system. *Int J Mol Sci*. 2024;25(3):1533. <https://doi.org/10.3390/ijms25031533>.
29. Fang B, Xu H, Liu Y, Qi F, Zhang W, Chen H, Wang C, Wang Y, Yang W, Li S. Mutagenesis and redox partners analysis of the P450 fatty acid decarboxylase OleTJE. *Sci Rep*. 2017;7(1):1–7. <https://doi.org/10.1038/srep44258>.
30. Fasan R, Mehareenna YT, Snow CD, Poulos TL, Arnold FH. Evolutionary history of a specialized P450 propane monooxygenase. *J Mol Biol*. 2008;383(5):1069–80. <https://doi.org/10.1016/j.jmb.2008.06.060>.
31. Fatma Z, Hartman H, Poolman MG, Fell DA, Srivastava S, Shakeel T, Yazdani SS. Model-assisted metabolic engineering of *Escherichia coli* for long chain alkane and alcohol production. *Metab Eng*. 2018;46:1–12. <https://doi.org/10.1016/j.ymben.2018.01.002>.
32. Fram B, Su Y, Truebridge I, Riesselman AJ, Ingraham JB, Passera A, Napier E, Thadani NN, Lim S, Roberts K, Kaur G. Simultaneous enhancement of multiple functional properties using evolution-informed protein design. *Nat Commun*. 2024;15(1):5141. <https://doi.org/10.1038/s41467-024-49119-x>.
33. Grant C, Deszcz D, Wei YC, Martínez-Torres RJ, Morris P, Folliard T, Sreenivasan R, Ward J, Dalby P, Woodley JM, Baganz F. Identification and use of an alkane transporter plug-in for applications in biocatalysis and whole-cell biosensing of alkanes. *Sci Rep*. 2014;4(1):1–9. <https://doi.org/10.1038/srep05844>.
34. Greener A, Callahan M, Jerpseth B. An efficient random mutagenesis technique using an *E. coli* mutator strain. *Mol Biotechnol*. 1997;7(2):189–95. <https://doi.org/10.1007/BF02761755>.
35. Holze H, Schrader L, Buellbach J. Advances in deciphering the genetic basis of insect cuticular hydrocarbon biosynthesis and variation. *Heredity*. 2020;126(2):219–34. <https://doi.org/10.1038/s41437-020-00380-y>.
36. Hu Y, Zhu Z, Gradischnig D, Winkler M, Nielsen J, Siewers V. Engineering carboxylic acid reductase for selective synthesis of medium-chain fatty alcohols in yeast. *Proc Natl Acad Sci USA*. 2020;117(37):22974–83. <https://doi.org/10.1073/PNAS.2010521117>.
37. Jetter R, Kunst L. Plant surface lipid biosynthetic pathways and their utility for metabolic engineering of waxes and hydrocarbon biofuels. *Plant J*. 2008;54(4):670–83. <https://doi.org/10.1111/J.1365-313X.2008.03467.X>.
38. Ju S, Li D, Mai BK, Liu X, Vallota-Eastman A, Wu J, Valentine DL, Liu P, Yang Y. Stereodivergent photobiocatalytic radical cyclization through the repurposing and directed evolution of fatty acid photodecarboxylases. *Nat Chem*. 2024;2024:1–9. <https://doi.org/10.1038/s41557-024-01494-0>.
39. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596:583–9. <https://doi.org/10.1038/s41586-021-03819-2>.
40. Kang M-K, Nielsen J. Biobased production of alkanes and alkenes through metabolic engineering of microorganisms. *J Industrial Microbiol Biotechnol*. 2016;44(4):613–22. <https://doi.org/10.1007/S10295-016-1814-Y>.
41. Kallio P, Pásztor A, Thiel K, Akhtar MK, Jones PR. An engineered pathway for the biosynthesis of renewable propane. *Nat Commun*. 2014;5(1):4731. <https://doi.org/10.1038/ncomms5731>.
42. Kennerley J. An evolving synthetic biological platform for alkane production. University of Oxford; 2022. https://doi.org/10.5287/ora-z6r2z_aypk.
43. Khara B, Menon N, Levy C, Mansell D, Das D, Marsh ENG, Leys D, Scrutton NS. Production of propane and other short-chain alkanes by structure-based engineering of ligand specificity in aldehyde-deformylating oxygenase. *ChemBioChem*. 2013;14(10):1204–8. <https://doi.org/10.1002/cbic.201300307>.
44. Kramer L, Le X, Rodriguez M, Wilson MA, Guo J, Niu W. Engineering carboxylic acid reductase (CAR) through a whole-cell growth-coupled NADPH recycling strategy. *ACS Synth Biol*. 2020;9(7):1632–7. <https://doi.org/10.1021/ACSSYNBIO.0C00290>.
45. Krishna R, Wang J, Ahern W, Sturmfels P, Venkatesh P, Kalvet I, Lee GR, Morey-Burrows FS, Anishchenko I, Humphreys IR, McHugh R, Vafeados D, Li X, Sutherland GA, Hitchcock A, Neil Hunter C, Kang A, Brackenbrough E, Bera AK, et al. Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science*. 2024. <https://doi.org/10.1126/SCIENCE.ADL2528>.
46. Lea-Smith DJ, Ortiz-Suarez ML, Lenn T, Nürnberg DJ, Baers LL, Davey MP, Parolini L, Huber RG, Cotton CAR, Mastroianni G, Bombelli P, Ungerer P, Stevens TJ, Smith AG, Bond PJ, Mullineaux CW, Howe CJ. Hydrocarbons are essential for optimal cell size, division, and growth of cyanobacteria. *Plant Physiol*. 2016;172(3):1928–40. <https://doi.org/10.1104/PP.16.01205>.
47. Lehtinen T, Santala V, Santala S. Twin-layer biosensor for real-time monitoring of alkane metabolism. *FEMS Microbiol Lett*. 2017;364(6):53. <https://doi.org/10.1093/femsle/fnx053>.
48. Lehtinen T, Virtanen H, Santala S, Santala V. Production of alkanes from CO2 by engineered bacteria. *Biotechnol Biofuels*. 2018;11(1):228. <https://doi.org/10.1186/s13068-018-1229-2>.

49. Li J, Ma Y, Liu N, Eser BE, Guo Z, Jensen PR, Stephanopoulos G. Synthesis of high-titer alka(e)nes in *Yarrowia lipolytica* is enabled by a discovered mechanism. *Nat Commun.* 2020;11(1):1–13. <https://doi.org/10.1038/s41467-020-19995-0>.
50. Li N, Chang WC, Warui DM, Booker SJ, Krebs C, Bollinger JM. Evidence for only oxygenative cleavage of aldehydes to alk(a)e)nes and formate by cyanobacterial aldehyde decarboxylases. *Biochemistry.* 2012;51(40):7908–16. <https://doi.org/10.1021/bi300912n>.
51. Liu C, Wu J, Chen Y, Liu Y, Zheng Y, Liu L, Zhao J. Advances in zero-shot prediction-guided enzyme engineering using machine learning. *ChemCatChem.* 2025;17(5): e202401542. <https://doi.org/10.1002/cctc.202401542>.
52. Liu K, Li S. Biosynthesis of fatty acid-derived hydrocarbons: perspectives on enzymology and enzyme engineering. *Curr Opin Biotechnol.* 2020;62:7–14. <https://doi.org/10.1016/j.COPBIO.2019.07.005>.
53. Liu Y, Chen J, Khusnutdinova AN, Correia K, Diep P, Batyrova KA, Nemr K, Flick R, Stogios P, Yakunin AF, Mahadevan R. A novel C-terminal degron identified in bacterial aldehyde decarboxylases using directed evolution. *Biotechnol Biofuels.* 2020;13(1):114. <https://doi.org/10.1186/s13068-020-01753-5>.
54. Liu Y, Wang C, Yan J, Zhang W, Guan W, Lu X, Li S. Hydrogen peroxide-independent production of α -alkenes by OleT JE P450 fatty acid decarboxylase. *Biotechnol Biofuels.* 2014;7(1):1–12. <https://doi.org/10.1186/1754-6834-7-28>.
55. Machado HB, Dekishima Y, Luo H, Lan EI, Liao JC. A selection platform for carbon chain elongation using the CoA-dependent pathway to produce linear higher alcohols. *Metab Eng.* 2012;14(5):504–11. <https://doi.org/10.1016/j.YMBEN.2012.07.002>.
56. Manley OM, Fan R, Guo Y, Makris TM. Oxidative decarboxylase UndA utilizes a dinuclear iron cofactor. *J Am Chem Soc.* 2019;141(22):8684–8. <https://doi.org/10.1021/JACS.9B02545>.
57. Marsh ENG, Waugh MW. Aldehyde decarboxylases: enigmatic enzymes of hydrocarbon biosynthesis. In *ACS catalysis*. American Chemical Society; 2013. vol 3, issue 11, p 2515–2521 <https://doi.org/10.1021/cs400637t>
58. Maxel S, King E, Zhang Y, Luo R, Li H. Leveraging oxidative stress to regulate redox balance-based, in vivo growth selections for oxygenase engineering. *Cite This: ACS Synth Biol.* 2020;2020:3133. <https://doi.org/10.1021/acssynbio.0c00380>.
59. Mendez-Perez D, Begemann MB, Pfeleger BF. Modular synthase-encoding gene involved in α -olefin biosynthesis in *Synechococcus* sp. strain pCC 7002. *Appl Environ Microbiol.* 2011;77(12):4264–7. <https://doi.org/10.1128/AEM.00467-11>.
60. Molina RS, Rix G, Mengiste AA, Álvarez B, Seo D, Chen H, Hurtado JE, Zhang Q, García-García JD, Heins ZJ, Almhjell PJ, Arnold FH, Khalil AS, Hanson AD, Dueber JE, Schaffer DV, Chen F, Kim S, Fernández LÁ, et al. In vivo hypermutation and continuous evolution. *Nat Rev Methods Prim.* 2022;2(1):1–22. <https://doi.org/10.1038/s43586-022-00119-5>.
61. Musil M, Khan RT, Beier A, Stourac J, Konegger H, Damborsky J, Bednar D. FireProtASR: a web server for fully automated ancestral sequence reconstruction. *Brief Bioinform.* 2021;22(4):bbaa337. <https://doi.org/10.1093/bib/bbaa337>.
62. Nielsen JR, Weusthuis RA, Huang WE. Growth-coupled enzyme engineering through manipulation of redox cofactor regeneration. *Biotechnol Adv.* 2023. <https://doi.org/10.1016/j.BIOTECHADV.2023.108102>.
63. Ni J, Tao F, Du H, et al. Mimicking a natural pathway for *de novo* biosynthesis: natural vanillin production from accessible carbon sources. *Sci Rep.* 2015;5:13670. <https://doi.org/10.1038/srep13670>.
64. Nishimura A, Morita M, Nishimura Y, Sugino Y. A rapid and highly efficient method for preparation of competent *Escherichia coli* cells. *Nucleic Acids Res.* 1990;18(20):6169. <https://doi.org/10.1093/nar/18.20.6169>.
65. Orsi E, Schada von Borzyskowski L, Noack S, Nickel PI, Lindner SN. Automated in vivo enzyme engineering accelerates biocatalyst optimization. *Nat Commun.* 2024;15(1):1–14. <https://doi.org/10.1038/s41467-024-46574-4>.
66. Packer MS, Liu DR. Methods for the directed evolution of proteins. *Nat Rev Genet.* 2015;16(7):379–94. <https://doi.org/10.1038/nrg3927>.
67. Paul C, Hanefeld U, Hollmann F, Qu G, Yuan B, Sun Z. Enzyme engineering for biocatalysis. *Mol Catal.* 2024;555: 113874. <https://doi.org/10.1016/J.MCAT.2024.113874>.
68. Raman S, Rogers JK, Taylor ND, Church GM. Evolution-guided optimization of biosynthetic pathways. *Proc Natl Acad Sci USA.* 2014;111(50):17803–8. <https://doi.org/10.1073/PNAS.1409523111>.
69. Ratajczak A, Geißdörfer W, Hillen W. Expression of alkane hydroxylase from *Acinetobacter* sp. strain ADP1 is induced by a broad range of n-alkanes and requires the transcriptional activator AlkR. *J Bacteriol.* 1998;180(22):5822–7. <https://doi.org/10.1128/JB.180.22.5822-5827.1998>.
70. Reed B, Blazeck J, Alper H. Evolution of an alkane-inducible biosensor for increased responsiveness to short-chain alkanes. *J Biotechnol.* 2012;158(3):75–9. <https://doi.org/10.1016/J.JBIOTECH.2012.01.028>.
71. Reetz MT, Carballeira JD. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat Protoc.* 2007;2(4):891–903. <https://doi.org/10.1038/nprot.2007.72>.
72. Rohl CA, Strauss CE, Misura KM, Baker D. Protein structure prediction using Rosetta. In: *Methods in enzymology*. 2004. vol. 383, p. 66–93. [https://doi.org/10.1016/S0076-6879\(04\)83004-0](https://doi.org/10.1016/S0076-6879(04)83004-0)
73. Royce LA, Liu P, Stebbins MJ, Hanson BC, Jarboe LR. The damaging effects of short chain fatty acids on *Escherichia coli* membranes. *Appl Microbiol Biotechnol.* 2013;97(18):8317. <https://doi.org/10.1007/S00253-013-5113-5>.
74. Rude MA, Baron TS, Brubaker S, Alibhai M, Del Cardayre SB, Schirmer A. Terminal olefin (1-alkene) biosynthesis by a novel P450 fatty acid decarboxylase from *Jeotgaliococcus* species. *Appl Environ Microbiol.* 2011;77(5):1718–27. <https://doi.org/10.1128/AEM.02580-10>.
75. Rui Z, Harris NC, Zhu X, Huang W, Zhang W. Discovery of a family of desaturase-like enzymes for 1-alkene biosynthesis. *ACS Catal.* 2015;5(12):7091–4. <https://doi.org/10.1021/ACSCATAL.5B01842>.
76. Rui Z, Li X, Zhu X, Liu J, Domigan B, Barr I, Cate JHD, Zhang W. Microbial biosynthesis of medium-chain 1-alkenes by a nonheme iron oxidase. *Proc Natl Acad Sci.* 2014;111(51):18237–42. <https://doi.org/10.1073/PNAS.1419701112>.
77. Schirmer A, Rude MA, Li X, Popova E, Del Cardayre SB. Microbial biosynthesis of alkanes. *Science.* 2010;329(5991):559–62. <https://doi.org/10.1126/science.1187936>.
78. Scown CD, Keasling JD. Sustainable manufacturing with synthetic biology. *Nat Biotechnol.* 2022;40(3):304–7. <https://doi.org/10.1038/s41587-022-01248-8>.
79. Sellés Vidal L, Murray JW, Heap JT. Versatile selective evolutionary pressure using synthetic defect in universal metabolism. *Nat Commun.* 2021;12(1):1–15. <https://doi.org/10.1038/s41467-021-27266-9>.
80. Shakeel T, Gupta M, Fatma Z, Kumar R, Kumar R, Singh R, Sharma M, Jade D, Gupta D, Fatma T, Yazdani SS. A consensus-guided approach yields a heat-stable alkane-producing enzyme and identifies residues promoting thermostability. *J Biol Chem.* 2018;293(24):9148–61. <https://doi.org/10.1074/jbc.RA117.000639>.
81. Sheppard MJ, Kunjapur AM, Prather KL. Modular and selective biosynthesis of gasoline-range alkanes. *Metab Eng.* 2016;33:28–40. <https://doi.org/10.1016/j.jymben.2015.10.010>.
82. Simon AJ, d’Oelsnitz S, Ellington AD. Synthetic evolution. *Nat Biotechnol.* 2019;37(7):730–43. <https://doi.org/10.1038/S41587-019-0157-4>.
83. Sorigué D, Légeret B, Cuiné S, Blangy S, Moulin S, Billon E, Richaud P, Brugière S, Couté Y, Nurizzo D, Müller P, Brettel K, Pignol D, Arnoux P, Li-Beisson Y, Peltier G, Beisson F. An algal photoenzyme converts fatty acids to hydrocarbons. *Science.* 2017;357(6354):903–7. <https://doi.org/10.1126/science.aan6349>.
84. Sorigué D, Légeret B, Cuiné S, Morales P, Mirabella B, Guédeney G, Li-Beisson Y, Jetter R, Peltier G, Beisson F. Microalgae synthesize hydrocarbons from long-chain fatty acids via a light-dependent pathway. *Plant Physiol.* 2016;171(4):2393–405. <https://doi.org/10.1104/PP.16.00462>.
85. Sticher P, Jaspers MCM, Stemmler K, Harms H, Zehnder AJB, Van Der Meer JR. Development and characterization of a whole-cell bioluminescent sensor for bioavailable middle-chain alkanes in contaminated groundwater samples. *Appl Environ Microbiol.* 1997;63(10):4053–60. <https://doi.org/10.1128/AEM.63.10.4053-4060.1997>.
86. Sukovich DJ, Seffernick JL, Richman JE, Gralnick JA, Wackett LP. Widespread head-to-head hydrocarbon biosynthesis in bacteria and role of OleA. *Appl Environ Microbiol.* 2010;76(12):3850–62. <https://doi.org/10.1128/AEM.00436-10>.
87. Sumida KH, Núñez-Franco R, Kalvet I, Pellock SJ, Wicky BI, Milles LF, Dauparas J, Wang J, Kipnis Y, Jameson N, Kang A. Improving protein

- expression, stability, and function with ProteinMPNN. *J Am Chem Soc.* 2024;146(3):2054–61. <https://doi.org/10.1021/jacs.3c10941>.
88. van Beilen JB, Wubbolts MG, Witholt B. Genetics of alkane oxidation by *Pseudomonas oleovorans*. *Biodegradation.* 1994;5(3–4):161–74. <https://doi.org/10.1007/BF00696457>.
 89. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A, Židek A. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 2022;50(D1):D439–44.
 90. Vidal LS, Isalan M, Heap JT, Ledesma-Amaro R. A primer to directed evolution: current methodologies and future directions. *RSC Chemical Biology.* 2023;4(4):271–91. <https://doi.org/10.1039/d2cb00231k>.
 91. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM. Programming cells by multiplex genome engineering and accelerated evolution. *Nature.* 2009;460(7257):894–8. <https://doi.org/10.1038/nature08187>.
 92. Wang Y, Xue P, Cao M, Yu T, Lane ST, Zhao H. Directed evolution: methodologies and applications. *Chem Rev.* 2021;121(20):12384–444. <https://doi.org/10.1021/ACS.CHEMREV.1C00260>.
 93. Wang Q, Bao L, Jia C, Li M, Li JJ, Lu X. Identification of residues important for the activity of aldehyde-deformylating oxygenase through investigation into the structure-activity relationship. *BMC Biotechnol.* 2017;17:1–9. <https://doi.org/10.1186/s12896-017-0351-8>.
 94. Ward T, Zou Z, Wu S, Gerngross D-Bsse D, Zürich E, Lozhkin B, Chen D. A microbial cell factory combining an artificial metathase with a fatty acid decarboxylase for new-to-nature synthesis of cycloalkenes. 2023. <https://doi.org/10.21203/RS.3.RS-3706252/V1>
 95. Warui DM, Li N, Nørgaard H, Krebs C, Bollinger JM, Booker SJ. Detection of formate, rather than carbon monoxide, as the stoichiometric coproduct in conversion of fatty aldehydes to alkanes by a cyanobacterial aldehyde decarboxylase. *J Am Chem Soc.* 2011;133(10):3316–9. <https://doi.org/10.1021/ja111607x>.
 96. Wiercx NJ, Ballerstedt H, de Bont JA, Wery J. Engineering of solvent-tolerant *Pseudomonas putida* S12 for bioproduction of phenol from glucose. *Appl Environ Microbiol.* 2005;71(12):8221–7. <https://doi.org/10.1128/AEM.71.12.8221-8227.2005>.
 97. Winters K, Parker PL, Van Baalen C. Hydrocarbons of blue-green algae: geochemical significance. *Science.* 1969;163(3866):467–8. <https://doi.org/10.1126/science.163.3866.467>.
 98. Wittmann BJ, Johnston KE, Wu Z, Arnold FH. Advances in machine learning for directed evolution. *Curr Opin Struct Biol.* 2021;69:11–8. <https://doi.org/10.1016/j.sbi.2021.01.008>.
 99. Wu S, Zhou Y, Gerngross D, Jeschek M, Ward TR. Chemo-enzymatic cascades to produce cycloalkenes from bio-based resources. *Nat Commun.* 2019;10(1):1–10. <https://doi.org/10.1038/s41467-019-13071-y>.
 100. Wu W, Zhang L, Yao L, Tan X, Liu X, Lu X. Genetically assembled fluorescent biosensor for in situ detection of bio-synthesized alkanes. *Nat Publ Group.* 2015;5:10907. <https://doi.org/10.1038/srep10907>.
 101. Xu H, Liang W, Ning L, Jiang Y, Yang W, Wang C, Qi F, Ma L, Du L, Fourage L, Zhou YJ, Li S. Directed evolution of P450 fatty acid decarboxylases via high-throughput screening towards improved catalytic activity. *ChemCatChem.* 2020;12(1):80–4. <https://doi.org/10.1002/CCTC.201901347>.
 102. Xu J, Li T, Huang WE, Zhou NY. Semi-rational design of nitroarene dioxygenase for catalytic ability toward 2,4-dichloronitrobenzene. *Appl Environ Microbiol.* 2024. <https://doi.org/10.1128/AEM.01436-23>.
 103. Xu W, Chen Y, Li D, Wang Z, Xu J, Wu Q. Rational design of fatty acid photodecarboxylase enables the efficient decarboxylation of medium- and short-chain fatty acids for the production of gasoline bio-alkanes. *Mol Catal.* 2022;524: 112261. <https://doi.org/10.1016/J.MCAT.2022.112261>.
 104. Qiu Y, Tittiger C, Wicker-Thomas C, Le Goff G, Young S, Wajnberg E, Fricaux T, Taquet N, Blomquist GJ, Feyereisen R. An insect-specific P450 oxidative decarboxylase for cuticular hydrocarbon biosynthesis. *Proc Natl Acad Sci USA.* 2012;109(37):14858–63. <https://doi.org/10.1073/PNAS.1208650109>.
 105. Yi D, Bayer T, Badenhorst CP, Wu S, Doerr M, Höhne M, Bornscheuer UT. Recent trends in biocatalysis. *Chem Soc Rev.* 2021;50(14):8003–49. <https://doi.org/10.1039/DOCS01575J>.
 106. Yamamori T, Kageyama H, Tanaka Y, Takabe T. Requirement of alkanes for salt tolerance of Cyanobacteria: characterization of alkane synthesis genes from salt-sensitive *Synechococcus elongatus* PCC7942 and salt-tolerant *Aphanothece halophytica*. *Lett Appl Microbiol.* 2018;67(3):299–305. <https://doi.org/10.1111/LAM.13038>.
 107. Zhang D, He Y, Wang Y, Wang H, Wu L, Aries E, Huang WE. Whole-cell bacterial bioreporter for actively searching and sensing of alkanes and oil spills. *Microb Biotechnol.* 2012;5(1):87–97. <https://doi.org/10.1111/J.1751-7915.2011.00301.X>.
 108. Zhang D, Zhao Y, He Y, Wang Y, Zhao Y, Zheng Y, Wei X, Zhang L, Li Y, Jin T, Wu L, Wang H, Davison PA, Xu J, Huang WE. Characterization and modeling of transcriptional cross-regulation in *Acinetobacter baylyi* ADP1. *ACS Synth Biol.* 2012;1(7):274–83. <https://doi.org/10.1021/SB3000244>.
 109. Zhang J, Lu X, Li JJ. Conversion of fatty aldehydes into alk (a/e)nes by in vitro reconstituted cyanobacterial aldehyde-deformylating oxygenase with the cognate electron transfer system. *Biotechnol Biofuels.* 2013;6(1):86. <https://doi.org/10.1186/1754-6834-6-86>.
 110. Zhou YJ, Hu Y, Zhu Z, Siewers V, Nielsen J. Engineering 1-alkene biosynthesis and secretion by dynamic regulation in *Acinetobacter baylyi*. 2018;7(2):584–90. <https://doi.org/10.1021/ACSSYNBIO.7B00338>.
 111. Zhou YJ, Kerkhoven EJ, Nielsen J. Barriers and opportunities in bio-based production of hydrocarbons. *Nat Energy.* 2018;3(11):925–35. <https://doi.org/10.1038/s41560-018-0197-x>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.