Synthetic Biology Approaches to Hydrocarbon Biosensors: A Review

Claudia F. Moratti1,2, Colin Scott2 and Nicholas V. Coleman1*

1School of Life and Environmental Science, Faculty of Science, University of Sydney, Sydney, NSW, Australia, 2CSIRO Synthetic Biology Future Science Platform, Canberra, ACT, Australia

Monooxygenases are a class of enzymes that facilitate the bacterial degradation of alkanes and alkenes. The regulatory components associated with monooxygenases are nature’s own hydrocarbon sensors, and once functionally characterised, these components can be used to create rapid, inexpensive and sensitive biosensors for use in applications such as bioremediation and metabolic engineering. Many bacterial monooxygenases have been identified, yet the regulation of only a few of these have been investigated in detail. A wealth of genetic and functional diversity of regulatory enzymes and promoter elements still remains unexplored and unexploited, both in published genome sequences and in yet-to-be-cultured bacteria. In this review we examine in detail the current state of research on monooxygenase gene regulation, and on the development of transcription-factor-based microbial biosensors for detection of alkanes and alkenes. A new framework for the systematic characterisation of the underlying genetic components and for further development of biosensors is presented, and we identify focus areas that should be targeted to enable progression of more biosensor candidates to commercialisation and deployment in industry and in the environment.

Keywords: biosensor, hydrocarbon, alkane, alkene, monooxygenase, transcription factor, regulation, bacteria

1 INTRODUCTION

Synthetic biology repurposes the native functions of organisms to engineer creative solutions to problems. The metabolic diversity of bacteria is among their most useful properties for synthetic biology; this feature is the product of millions of years of evolutionary adaptation to diverse niches, nutrients, and stresses. Hydrocarbon-metabolising bacteria are especially interesting for synthetic biology applications, including bioremediation (Das and Chandran, 2011), biocatalysis (Coscolín et al., 2019), and biosensors (Sticher et al., 1997).

Nearly eighty different bacterial genera have been identified as degraders of at least one petrochemical, although only a fraction are genetically or biochemically well-characterised. Monooxygenases are the key enzymes responsible for bacterial degradation of alkanes and alkenes in such bacteria, and the regulatory systems of monooxygenases are therefore of significant interest. There is a need to better characterise these sensing systems to better understand the metabolic diversity of bacteria and to capitalise on these biological switches.

Several eukaryotic species are also capable of hydrocarbon degradation, including many genera of fungi and yeast, and one alga (Prince, 2010). Similarly to bacterial systems, monooxygenases are responsible for the eukaryotic metabolism of aliphatic alkanes. To date only the cytochrome P450 class of monooxygenases have been identified to perform this role (Das and Chandran, 2011; Prenafeta-Boldu et al., 2019) and only under aerobic conditions, unlike in bacteria where hydrocarbon metabolism is facilitated by several classes of monooxygenases and can also occur
anaerobically. In fungal systems, other types of hydrocarbons including alkenes and polycyclic aromatics are typically only partially degraded and it is rare for hydrocarbons to act as a sole carbon source (Prince, 2010; Prenafeta-Boldu et al., 2019). Eukaryotic hydrocarbon degradation has already been extensively reviewed (Cerniglia and Crow, 1981; Prince, 2010; Das and Chandran, 2011; Beier et al., 2014; Prenafeta-Boldu et al., 2019) and so is not included in the scope of this review. Moreover, the variety of monooxygenases involved in bacterial hydrocarbon degradation is interesting and warrants a narrow focus, especially when considering these systems for biotechnology applications.

Hydrocarbon-sensing systems consist of regulatory proteins (transcription factors) that bind to an inducer and then interact with operator sequences near the promoter of the gene being controlled, resulting in a change in expression levels of that gene (Figure 1). In some cases, the sensing system is divided into two proteins, one which binds the inducer, and another which interacts with the operator; the induction signal in these cases is passed from the first protein to the second. A promoter is a sequence of DNA upstream of a gene that recruits RNA polymerase for transcription (Browning and Busby, 2016). Operators act as binding sites for specific transcription factors and have features including direct or inverted repeats. An inducer is a compound that interacts with the regulatory protein in a way which changes the binding of the protein to the operator.

A regulatory protein is considered a repressor if it binds to the operator in the absence of an inducer, preventing transcription. An inducer will relieve repression by binding to the protein in a way that prevents the protein from remaining bound to the operator. In contrast, activators are regulatory proteins that only bind to the operator once they have complexed with an inducer; in these cases, transcription will be switched on after binding of the activator protein. Transcription-factor-based biosensors can be easily constructed from these regulatory components by replacing the metabolic genes normally controlled by the system with a readily detectable output such as fluorescence. The broad use of bacterial two-component systems in biosensors has been recently reviewed (Lazar and Tabor, 2021).

There has been much research into the bacterial metabolism and detection of aromatic hydrocarbons (Chauhan et al., 2008; Plotnikova et al., 2016; Reineke et al., 2020). Less is known about metabolism and sensing of aliphatic hydrocarbons, in particular the shorter chain gaseous compounds. Development of biosensors for detection of \( n \)-alkanes and \( n \)-alkenes has mostly been aimed at monitoring the clean-up of oil spills in seawater (Harayama et al., 2004; Tecon and Van der Meer, 2008; Tecon et al., 2010; Das and Chandran, 2011; Reed et al., 2012; Zhang et al., 2012a; Sevilla et al., 2017; Varjani, 2017), with a consequent focus on biosensor systems detecting octane (Jiang et al., 2020) and bacterial hosts suited to marine environments (Sevilla et al., 2015). The limited scope of research on aliphatic biosensors is unfortunate given their potentially many applications.

Aliphatic hydrocarbon biosensors are potentially useful in the chemical industry for real-time monitoring of reactants or products (Minak-Bernero et al., 2004), in the fresh produce industry for the detection of ethene (Janssen et al., 2014), and in the natural gas industry for both safety and maintenance applications (e.g., leak detection) (Nandimandalam et al., 2018). There has also been interest in the biofuels industry in biosensors for detecting the intracellular concentrations of biosynthesized alkanes (Wu et al., 2015). Finally, there are potential medical applications for alkane biosensors in analysis of breath samples for lung cancer diagnosis (Tan et al., 2016).

This review will provide a framework for the development of transcription-factor-based biosensors to help guide future development efforts. The framework will then be used to assess the current research landscape in the case of aliphatic hydrocarbons, including review of the bacteria and catabolic enzymes associated with these substrates, the types of regulatory genes involved, and consideration of the limiting factors in the research development pipeline. Methods commonly used to fulfill the requirements of each step of the framework will also be summarised throughout.

2 A FRAMEWORK FOR BACTERIAL BIOSENSOR DEVELOPMENT

A framework has been developed here to show the key stages in the development of transcription-factor-based biosensors; this includes clear research goals that must be achieved to progress
from concept to deployment. When applied to the biosensor literature to date, this framework allows identification of the limiting steps, which is important to focus funding and research in the areas which will provide the most reward for effort.

The first step in the development of any biosensor is the identification of a biological part that could be repurposed for a sensing application; this might be an enzyme, antibody or nucleic acid (van der Meer, 2010). Research then progresses to characterisation, where the function of the candidate sensor is confirmed. This is usually done via the deletion and/or complementation of the gene and/or promoter element in the presence of the predicted inducer compound(s). Characterisation may be done in stages depending on how the sensing system was first discovered. For example, if it was found associated with specific catabolic genes, the function of these is often first confirmed before the associated sensing elements are characterised.

The third checkpoint in biosensor development is proof-of-concept in a controlled laboratory environment. At this stage, various key parameters such as sensitivity, selectivity, shelf life and response time would be determined. Next is the development phase, where the biosensor circuitry and context are adapted with the intention of optimising those key parameters. Finally, the device can proceed to be engineered for market deployment. This stage includes activities such as: seeking regulatory approval, benchmarking analysis, production scale-up, process optimisation, cellular immobilisation, and company formation.

Applying the above framework to the specific case of aliphatic hydrocarbon biosensors typically involves two rounds of identification and characterisation. First, a monooxygenase-encoding gene is identified, and the corresponding enzyme function is characterised. Secondly, the transcription factor responsible for monooxygenase expression is identified and the regulatory mechanism of the system is characterised. There will be a strong focus on the characterisation strategies for these regulatory mechanisms in this review because it is at this stage where greater research efforts are most needed, as will be made apparent.

3 IDENTIFICATION AND CHARACTERISATION OF MONOOXYGENASES

Identification of a biological part may be deliberate or coincidental, rigorous or speculative. Bioinformatic detection could be targeted at the transcription factor itself, or at the associated catabolic genes. The identification stage typically
### TABLE 1 | Organisation and properties of bacterial monooxygenases involved in aliphatic hydrocarbon degradation.

| Monoxygenase family | Description* | Substrate(s) | Inducer(s) | Example |
|---------------------|--------------|--------------|------------|---------|
| **SDIMO**: Soluble di-iron monooxygenases | Group 1 | Contains 4 protein subunits (αβγδ)-C-R, encoded by 6 genes | benzene, toluene, xylenes, styrene, naphthalene, indole, C7-C10 alkenes, C7 chloroalkanes, chloroform, isoprene | Propene monooxygenase XamoABCDEF from Xanthobacter Py2, Zhou et al. (1999) |
| | Group 3, includes sMMOs: soluble methane monooxygenases | Contains 4 protein subunits (αβγδ)-C-X-R, encoded by 4 genes | C7-C10 alkanes and haloalkanes, C7-C10 ethers, cyclohexane, benzene, toluene, styrene, pyridine, methanol | Low Cu/biomass ratio (sMMO only), n-butanol (BMO) Hanson and Hanson (1996), Kurth et al. (2008) |
| | | | | Soluble methane monooxygenase MmoxXBCDC from Methylcoccus capsulatus Bath, Stanyhlende et al. (1990), Rosenzweig et al. (1993) |
| CuMMO: Copper membrane monooxygenases | Includes pMMOs: Particulate methane monooxygenases | Contains 3 protein subunits (αβγδ)-C-R, encoded by 3 genes | C1-C5 alkanes and haloalkanes, C2-C4 alkenes and haloalkanes, C2-C4 ethers, chloroethylene, methyl tert-butyl ether, N-nitosodimethylamine | High copper/biomass ratio (pMMO only) Hanson and Hanson (1996), propane, butane, tert-butyl alcohol, C4-C6 ethers Johnson et al. (2004) |
| | | | Propane, tetrahydrofuran, dioxane | Particulate methane monooxygenase PmABCD from Mycobacterium TY-6 Kotani et al. (2006) |
| alkB: integral-membrane non-heme di-iron monooxygenase | alkB, alkM | Contains 3 protein subunits (αβγδ)-F-R, encoded by 3 genes | C2-C13 alkanes, C2-C20 alkenes, propylene, 1-butenes | Akane hydroxylase AkBFGHJKL from Pseudomonas putida Gpo1 van Beilen et al. (1994), AkMva from Acinetobacter sp. M1, Tani et al. (2000), Throne-Holst et al. (2007) |
| | | | | Cytochrome P450 alkane hydroxylase from Alcanivorax dieselii Ei et al. (2011) |
| CYP: soluble heme-dependent cytochrome P450s | CYP153 | Diverse in structure, requires 3 components | C2-C16 alkanes, C7-C10 alkenes, limonene, cyclohexene, styrene, medium and long-chain fatty acids | Cytochrome P450 4A10 alkane hydroxylase from Alcanivorax dieselii Ei et al. (2011) |
| | | | | | |
| | Class VII CYP P450 | Contains 1 gene consisting with 2 domains; a heme domain and a reductase domain | C14-C16, C24, C26 | Medium- and long-chain alkanes Minoridi et al. (2015) |
| FMO: Soluble flavin-binding monooxygenase | AlaM | Contains 1 gene | C14-C16 alkenes, pristane, phytane | Akma from Acinetobacter sp. DSM 17874, Throne-Holst et al. (2007) |

(Continued on following page)
TABLE 1 | (Continued) Organisation and properties of bacterial monoxygenases involved in aliphatic hydrocarbon degradation.

| Monoxygenase family | Descriptiona | Substrate(s) | Inducer(s) | Example |
|---------------------|--------------|--------------|------------|---------|
| Dioxygenase         | Contains 1 gene, requires Cu²⁺ but not NAD(P)⁺ Maeng et al. (1998) | C₁₀-C₂₀ alkanes, C₂₀-C₃₀ alkenes, arylnaphthalene and tridecylbenzene Maeng et al. (1998) | C₁₀-C₂₀ alkanes Maeng et al. (1998), Sakai et al. (1998) | Acinetobacter sp. M1 Sakai et al. (1998) |
| LadA                | Contains 1 gene consisting of 3 domains; a monoxygenase domain, and 2 NAD(P)⁺ oxidation domains Feng et al. (2007), Tourova et al. (2016) | C₂₂-C₃₆ alkanes Feng et al. (2007), Wang and Shao. (2013), Wang and Shao. (2014), Tourova et al. (2016) | C₂₀-C₃₀ alkanes Liu et al. (2011) | LadhA from Geobacillus thermodenitrificans NG80-2 Feng et al. (2007) |

*a,b,γ, oxygenase subunit; R, reductase subunit; C, coupling protein; F, ferredoxin; X, protein of unknown function.

This is tentatively inferred by the homology of the group 5 and group 6 SDIMOs to the group 4 alkene MOs.

involves the study of pure cultures of bacteria, database mining and/or metagenomics.

The identification of the monoxygenase responsible for aliphatic hydrocarbon oxidation is a natural first step in biosensor development for this class of compounds. The rationale for this is two-fold. Firstly, monoxygenase genes make better targets for bioprospecting (Holmes and Coleman, 2008) than the associated regulatory genes, which tend to be more diverse and thus more difficult to detect by PCR or sequence analysis. Secondly, in many cases, the inducers of aliphatic hydrocarbon degradation genes are not the alkanes or alkenes, but rather the downstream metabolites (alcohols, aldehydes, epoxides) (Kurth et al., 2008) and it is the monoxygenase enzymes that play the crucial role of converting the hydrocarbon of interest into the inducer molecule.

Monoxygenases are responsible for the oxidation of methyl or methylene groups in n-alkanes, which is the first step in the catabolism of n-alkanes in all aerobic bacteria. The alcohols produced by this catalysis can then be oxidised into aldehydes or ketones, and then ultimately into fatty acids (Van Beilen et al., 2003; Kotani et al., 2006). Monoxygenases are of equal importance for alkene oxidation, with the resultant epoxides processed via a series of coenzyme M or glutathione-mediated reactions (Vlieg et al., 2000; Ensign, 2001; Mattes et al., 2010). It is not uncommon for one bacterial isolate to possess multiple monoxygenases of different types (Rojo, 2009), which may include both alkane and alkene-oxidising enzymes.

3.1 Methods for Identifying the Presence of Monoxygenases in Bacterial Species

Monoxygenase genes are most commonly identified in cultures obtained from enrichment and isolation on aliphatic hydrocarbons as the sole carbon source (Holmes and Coleman, 2008). Alternatively, these genes can be retrieved by metagenomic analysis, preferably using DNA from environments enriched in hydrocarbons (Musumeci et al., 2017; Gacesa et al., 2018), or by trawling pre-existing sequence data, which is now abundant due to the decreased costs of DNA sequencing. Monoxygenase homologues fall into distinct classes (Table 1), with known conserved sequence regions, and thus they can easily be identified purely based on sequence analysis in genomes or metagenomes, and functional approaches are not required at the gene discovery stage. The main classes of monoxygenases and their relevant properties are summarised in Table 1.

There is great interest in detecting and recovering new monoxygenases, due to their interesting catalytic properties and also due to their linkage to useful regulators for biosensor construction. Nested PCR with degenerate primers enabled recovery of novel soluble di-iron monoxygenases (SDIMO) from soils, sediments, and enrichment cultures, and was also useful for identifying interesting isolates which contained multiple SDIMO genes (Coleman et al., 2006). Other PCR approaches have been invaluable for screening isolate collections for AlkB type monoxygenases (Smits et al., 1999; Van Beilen et al., 2003). Metaproteomics approaches have been used to identify novel alkene monoxygenases in enrichments from vinyl-chloride contaminated groundwater (Chuang et al., 2010), and to identify archaeal ammonia monoxygenases (a copper-containing membrane monoxygenase, CuMMO) in marine samples (Morris et al., 2010). Novel CuMMOs have also been isolated from an oilsands tailing pond using stable isotope probing and qPCR methods (Rochman et al., 2020).

3.2 Functional Characterisation of Monoxygenases

Once a novel target gene has been identified, it is most important for biosensor development to confirm that the monoxygenase genes are indeed inducible by hydrocarbons (or metabolites thereof) (Vogne et al., 2010). Determining other parameters such as substrate range and kinetics are also very important for related applications like bioremediation. Techniques used to confirm the function of novel monoxygenases may include biochemical assays (resting cells, cell extracts etc.), omics approaches (e.g., transcriptomics, proteomics), or genetic methods (knockouts, knockdowns, heterologous expression).

Alkane-degrading bacterial species are diverse and some are well-characterised, with Actinobacteria such as Corynebacterium, Mycobacterium, Nocardia and Rhodococcus dominating when gaseous substrates are used (Shennan, 2006) and
Proteobacteria such as *Pseudomonas*, *Acinetobacter*, and *Alcanivorax* more typical when liquid substrates are used (Nie et al., 2014a). The best-characterised alkane monooxygenase is AlkB from *P. putida* GP01, isolated on hexane, which can oxidise C5-C9 n-alkanes (Baptist et al., 1963; van Beilen et al., 1994). AlkB requires accessory proteins (AlkG, AlkT) to deliver electrons from NADH to enable the activation of molecular oxygen. Many other classes of monooxygenases can also attack alkanes, including iron, flavin, and copper-requiring enzymes (Table 1, also see reference Moreno and Rojo, 2017). The gene arrangements of a selection of characterised alkane monooxygenases can be seen in Figure 2.

There are fewer examples of alkene-oxidising enzymes (Table 1), with the two best-studied systems being the propene monooxygenases of *Rhodococcus rhodochrous* (Gallagher et al., 1997; Gallagher et al., 1998; Smith et al., 1999) and *Xanthobacter* Py2 (Ginkel, 1987; Ginkel et al., 1987; Small and Ensing, 1997; Champreda et al., 2004). Despite having the same primary substrate, these two enzymes are diverse in both sequence and structure; the former is encoded by four genes encoding three enzyme subunits (Smith et al., 1999) while the latter is a six gene, four component system (Small and Ensing, 1997; McCarrl et al., 2018). Major advances since the alkene monooxygenases were last reviewed (Ensign, 2001; Shennan, 2006) include the identification, characterisation, and heterologous expression of the genes encoding the ethene monooxygenases (EtnABCD) found in *Nocardioides* and *Mycoabacterium* spp. (Coleman and Spain, 2003b; Mattes et al., 2005; Coleman et al., 2011a; McCarrl et al., 2018) and the investigation of these enzymes as biocatalysts for epoxide synthesis (Owens et al., 2009; Cheung et al., 2013).

The number of homologs in databases and the number of publications relevant to each representative monooxygenase are shown in Table 2. These numbers reflect the sheer volume of candidate enzymes that have been identified to date; the research at this early stage of the biosensor development framework is abundant. The relationship between monooxygenase genes and hosts is complex and it is likely that the corresponding genes have been subject to extensive lateral gene transfer (Coleman and Spain, 2003a; Das et al., 2015; Minerdi et al., 2015; Liang et al., 2016a; Khadka et al., 2018). The SDIMO and CuMMO type monooxygenases have quite restricted distribution compared to the AlkB and CYP153 enzymes; the latter are common in the genomes of environmental bacteria (Nie et al., 2014b).

### 4 IDENTIFICATION AND CHARACTERISATION OF REGULATORY SYSTEMS

#### 4.1 Initial Identification of Regulatory System Components

Important questions to be answered for newly discovered regulatory systems include: What are the inducers? Is it a one-component or two-component system? Are there additional layers of control apart from the hydrocarbon (or hydrocarbon metabolite) inducer? The investigation typically begins with bioinformatics, *via* identification of regulator gene(s), and the cognate promoter and operator sequences. Regulator genes are usually identified based on their proximity to the monooxygenase genes, and *via* sequence similarities to known transcription factors. Promoters can be identified on the basis of consensus sequences (e.g., −35 and −10 motifs) and their location upstream of the monooxygenase genes. Operators typically have an inverted repeat structure and will be located proximal to the promoter, either upstream (activators) or downstream (repressors) (Browning and Busby, 2016).

Pull-down assays offer an alternative method of identifying regulator proteins (Ji et al., 2019), in which a DNA containing the promoter sequence is immobilized on beads, cell extracts are washed over the beads, then proteins bound to the promoter sequence can be identified by mass spectrometry. This method allows regulatory proteins to be identified independent of bioinformatic predictions, but it does require that the promoter sequence is known, and it may be complicated by competitive binding of multiple proteins to the promoter; this may cause regulators that bind with lower

---

**Table 2** Classification and quantification of monooxygenase homologues and related publications in public databases.

| Representative enzyme subunit used for BLAST analysis | Monooxygenase homologues | Related publications |
|--------------------------------------------------------|--------------------------|---------------------|
| XamA, Xanthobacter Py2 | O87082 Aj006979.1 | 951 |
| MnoX, Methylcoccus capsulatus Bath | P22869 M00050.3 | 810 |
| EtnC, M. chuburense NES4 | D2K2E0 G1U74752.1 | 37 |
| PrmA, Gordonia TYS | Q768T5 AB11920 | 1,551 |
| PrmA, Mycobacterium TY-6 | Q08KF2 AB250938 | 959 |
| PmcO1, Methylcoccus capsulatus Bath | Q603F1 AE017282 | 1,103 |
| AlkB, P. putida GP01 | P12691 AJ245436 | 5,010 |
| CYP153, Alcanivorax dieselolaei | D0Q1H3 GQ980250 | 5,046 |
| CYP116B5, A. radiore sistens S13 | G9BWN9 HQ685986 | 2,787 |
| AimA, Acinetobacter sp. | AO2421HB23 MH857335 | 5,060 |
| LdA, Burkholderia sp. | A0A095EJX9 CP007785 | 4,798 |
| AkMa, Acinetobacter sp. M1 | GQPQK2 AB049410 | 5,012 |

| Scopus keyword(s) | No. of publications |
|-------------------|---------------------|
| “Propene monooxygenase” | 5 |
| “Soluble methane monooxygenase” | 426 |
| “Ethene monooxygenase” | 8 |
| “Propene monooxygenase” | 38 |
| “Particulate methane monooxygenase” | 500 |
| “AlkB” | 1855 |
| “CYP153” | 67 |
| “CYP116B5” | 2 |
| “AimA” AND “monooxygenase” | 14 |
| “LdA” AND “monooxygenase” | 14 |
| “AlmA” | 5 |

*Homologues were defined as BLAST matches with >40% amino acid identity.*
affinity to be overlooked. Other protein-DNA binding assays useful for identifying promoter/regulator pairs are electrophoretic mobility shift assays (EMSA) and DNase footprinting; in the former case, regulators are detected by their retardation of the gel mobility of a DNA containing the promoter sequence, while in the latter case, they are found by their ability to protect the promoter sequence from DNase digestion (Read et al., 2009).

Vogne et al. (2010) describe five criteria that should be addressed in the characterisation of catabolic regulatory factors, summarised below. In the case of alkane regulatory systems, these criteria have only been met in a handful of cases.

(1) There must be evidence for involvement of the regulator in gene expression.
(2) The genes being controlled must be identified.
(3) The promoter elements associated with the regulatory protein must be identified.
(4) The expression of the regulator gene itself must be investigated.
(5) The inducer compound and its relationship with the regulator must be understood.

### 4.2 Experimental Methods for Characterisation of Regulatory Systems

Two general approaches can be used to confirm the function of regulatory system elements; either untargeted methods e.g., proteomics and transcriptomics, or targeted methods e.g., heterologous expression, knock-ins, knockouts, knockdowns, pull-downs, gel shift assays, or footprinting. These approaches are described below.

Transcriptomics gives information about the expression patterns of different genes under different conditions. If exposure to a suspected inducer leads to higher expression levels of a particular gene, this provides strong preliminary evidence that the enzyme encoded by that gene is part of a metabolic pathway controlled by that inducer; searching upstream from the induced gene then allows discovery of the likely promoter. Alternatively, promoters can be found via transcriptomics via pinpointing intergenic regions which are not themselves transcribed. Finally, transcriptomics data can provide clues about the control of the regulatory genes themselves, such as whether these are constitutively expressed, or part of positive or negative feedback loops.

Heterologous expression can be used to confirm the role of regulatory proteins, e.g., by cloning the regulator gene and its cognate promoter and operator elements into a plasmid and adding a reporter gene downstream of the promoter. Exposure to the correct inducer will result in expression of the reporter gene if the combination of regulatory protein, promoter and operator sequences is correct. While a good starting point, there are limitations to this method, as follows; there may be multiple regulatory proteins required to evoke the desired response, the regulatory proteins might be hard to express in a heterologous host due to codon usage issues or due to strain-specific genes required for regulator protein function e.g., chaperones, the structure and type of plasmid used may unexpectedly impact the outcomes due to effects arising from copy number, gene orientation, or transcription read-through.

Gene knock-in methods are another useful targeted approach to testing hypotheses about regulatory systems. This involves integrating a reporter gene downstream of the promoter in the genome of the native organism, either in front of the metabolic genes, or replacing them. Exposure to the correct inducer should yield expression of the reporter gene. Knock-in methods have several advantages over heterologous expression in plasmids, most notably that they sidestep problems arising from codon usage, plasmid copy number, and altered genomic context. The major disadvantage of gene knock-ins is that they may be technically more difficult to generate, depending on the host organism.

Gene knockouts can also be used to interrogate the components of hydrocarbon regulatory systems. Knocking out the putative regulatory gene can be done via homologous recombination (usually replacing the regulator with a resistance gene) or via CRISPR-Cas-based methods; the latter are preferable due to the increased specificity and higher frequency of deletion mutants obtainable, but the choice here may be limited by the genetic tools available in the host species of interest. The impact of the knockout on the host organism’s phenotype can be readily tested, e.g., its ability to oxidise hydrocarbons. For repressors, a knockout should yield a constitutive hydrocarbon-oxidising phenotype, while for activators, a decreased or abolished ability to oxidize hydrocarbons would be expected.

Gene knockdowns enable the regulatory gene to be turned down or off temporarily via technologies such as interfering RNA (RNAi) (Hannon, 2002; Kim and Rossi, 2018) or dead Cas9 (dCas9) proteins (Dong et al., 2018). The overall strategy here is similar to the knock-outs described above, with the impact tested either via measuring hydrocarbon oxidation in the resultant recombinants or via a reporter gene if this has been integrated in place of the metabolic genes. An advantage of the knock-down approach is that it can be used to test inactivation of regulators of essential genes, since the bacteria can be grown first, then the knock-down activated, e.g., testing methane monoxygenase regulators in obligate methanotrophs.

### 5 Regulatory Systems for Alkane Monoxygenases

The most well-studied alkane monoxygenase regulatory systems are those associated with the AlkB and CYP153 monoxygenases, but there is also information available on the regulators of AlkB2, AlkM, AlmA, AlkW, BmoXYBZDC, PrmABC and SmoABCD. Taken together, these represent nine regulatory systems across seven species that have been experimentally characterised. If putative regulatory proteins identified by sequence analysis only are included, this count increases to 19 systems (Tables 3, 4). Details of the elements of these systems and their functions are described in the sections below. The positioning of a selection of regulatory genes relative to the relevant monoxygenase gene cluster can be seen in Figure 2.
It is important to note that the well-characterised alkane regulatory systems represent only a small fraction of the total diversity, and comparisons between these indicate the limitations in extrapolating conclusions from one system to another. The bias in research towards AlkB and CYP153 has left large knowledge gaps for other systems, and a further limitation is that more work has been focused on the monooxygenases rather than their regulatory systems—this is a major bottleneck in the

| Promoter name | Species | MO gene | Promoter sequence* | References |
|---------------|---------|---------|--------------------|------------|
| PalkB         | P. putida Gpo1 | alkBFGHJKL | TAAATATTATCCAGATAACTACCACACGAGATTTGAAGCA GGCCTAAGGTGAATGCTACATTTACCATTCTCA | Yuste et al. (1998), Ganosa et al. (2000) |
| PalkB         | P. putida P1 | alkBFGHJKL | TAAATATTATCCAGATAACTACCACACGAGATTTGAAGCA GGCCTAAGGTGAATGCTACATTTACCATTCTCA | van Beilen et al. (2001) |
| PalkB1        | A. borkumensis AP1 | alkSB1GHU | AAACCTACCCCTTTGGAGTGTGTTTGGGCTCTATTTCT AATTTGCTACATTTGAAC ATTTGCTACATTTGAAC | van Beilen et al. (2004) |
| PalkB2        | A. borkumensis AP1 | alkB2 | ATCTACTTAACTTTGTTTGGACAATCTCATGAACTATGG GTAGTTCACCTGGTGA | van Beilen et al. (2004) |
| Palkk         | A. borkumensis SK2 | CYP153 P450-1 | GGGCCCTTCTTGCTGTGATTTTTGGCGATAATCATCCTTTG GGAATGCCCAGTAATGCTTG | Sevilla et al. (2017) |
| PalkB         | B. cepacia RR10 | alkB | GGGCTTTGGAAGCGCGCGACCGGTCATATACGG CAGAAC | Marin et al. (2001) |
| PalkW1        | Dietzia sp. DQ12-45-1b | alkW1 | CTGGACAAAAATGAGATCGTGCTGAGACACCGCAAGG TTTTTGCTACATTTGCAACCAAT | Liang et al. (2016b) |
| PalkB1        | P. aeruginosa RR1 | alkB1 | TAGTGATCGTGCACTGCGACCCGGCTCGGTGTGGTGT TGGCTGTTGT | Marin et al. (2003) |
| PalkB2        | P. aeruginosa RR1 | alkB2 | TGGGATGGGGCGAGCGGCGGCAAAGGACGACAATTG TAGATGCACACATC | Marin et al. (2003) |
| Palkk         | Dietzia sp. DQ12-45-1b | CYP153 | CTGCCGCAAATGAGCCATCGGTGTGCTAATGAGCCAATGCTGGAAGCTAC | Liang et al. (2016a) |
| PalkM         | Acinetobacter ADP1 | alkM | ATGCCAAAGAGCCTTTGAGAGTTTTTGGAGACATACA ACTGTTTAAATATATATAAGATGAGGATATGAGGACAGTAAT | Rataczak et al. (1998a) |
| Palks2        | P. aeruginosa SJTD-1 | alkB2 | AAGCGCGCGACAGAACGAGGATGTTGGGACGGG CAGGGCGCGCAGGGCGAGGACAATTTGCATCAAGAC | Ji et al. (2019) |
| Ppmm          | Rhodococcus sp. BCP1 | pmABC | GGGCCCTGGGTCGACAGGGATATTTGTCCGGAATGT GACCTGCGTCATACAGACGAGTAACGGGGTGCG GCACGGGTTGTGTACGTGACAC | Cappelletti et al. (2015) |
| Psmo          | Rhodococcus sp. BCP1 | smABC | CCTACCCGGTCTAATTCGCTGAGGTAGTGTCA ACCCGGTTGGAATGATGTTGGATCTCAA | Cappelletti et al. (2015) |
| Palkk         | Rhodococcus sp. BCP1 | alkB | AAGCGTGGTTACAAATTACGATATTGGTTTCACCCTGTA GACCCAAATGTCCGTCAATCTAACATGAC | Cappelletti et al. (2011) |

*The −35 and −10 motifs are underlined, transcription start points are in bold, and inverted repeats are shown with arrows. All of these elements were identified in prior studies except in the case of Palkk of A. borkumensis AP1, where we have tentatively identified the −35, −10, and start point as part of this study.
### TABLE 4 | Summary of identified and/or characterised monooxygenase regulatory systems from literature.

| Regulator family | Nature of regulator | Regulator gene | Cognate promoter | Bacterial strain | Monooxygenase | Inducers | Evidence for regulator function | References |
|------------------|---------------------|----------------|------------------|------------------|---------------|----------|---------------------------------|------------|
| LuxR/MalT        | Activator           | AlkS           | $P_{\text{alkB}}$ | *P_{\text{alkB}}$ | P. putida GP01| alkBF/GHJKL/alkST | C$_{6}$-C$_{10}$ n-alkanes | Heterologous expression in E.coli; | Stich et al. (1997), Panke et al. (1999), Canosa et al. (2003), van Beilen et al. (2001) |
|                  |                     |                | $P_{\text{alkB}}$ | P. putida P1    | alkBF/GHJKL/ | C$_{6}$-C$_{12}$ n-alkanes | Inferred from results with GP01 | Schneiker et al. (2006) |
|                  |                     |                | $P_{\text{alkB}1}$ | A. borkumensis SK2 | alkSB1GHJ | C$_{6}$-C$_{12}$ n-alkanes | Sequence analysis only | van Beilen et al. (2004) |
|                  |                     |                | $P_{\text{alkB}1}$ | A. borkumensis AP1 | alkSB1GHJ | C$_{6}$-C$_{12}$ n-alkanes | S1 nuclease protection assay; lacZ | Schneiker et al. (2006), Sevilla et al. (2017) |
| AraC/XyS         | Activator           | cypR           | $P_{\text{bmx}}$ | A. borkumensis SK2 | CYP153 | C$_{6}$-C$_{18}$ n-alkanes | transcriptional fusion | Schneiker et al. (2006), Sevilla et al. (2017) |
|                  |                     |                | $P_{\text{bmx}}$ | Dietzia sp. DQ12-45-b1 | CYP115 | C$_{5}$-C$_{14}$ n-alkanes | Promoter-lacZ fusion assays; gene-inactivation; RACE analysis | Lian et al. (2016a) |
|                  |                     |                | $P_{\text{alkM}}$ | Acinetobacter sp. ADP1 | alkM | C$_{7}$-C$_{18}$ n-alkanes | Gene-inactivation; lacZ chromosomal fusions | Ratajczak et al. (1998a, 1998b) |
|                  |                     |                | $P_{\text{alkMa}}$ | Acinetobacter sp. M1 | alkMa | C$_{7}$-C$_{22}$ n-alkanes | DNA footprinting assay; EMSA | Liu et al. (2011) |
|                  |                     |                | $P_{\text{alkMb}}$ | Acinetobacter sp. M1 | alkMb | C$_{12}$-C$_{26}$ n-alkanes | Sequence analysis only | Liu et al. (2011) |
|                  |                     | $P_{\text{Orf1}}$ | — | A. dieselolei B5 | alkB2 | C$_{8}$-C$_{16}$ n-alkanes | Sequence analysis only | Li et al. (2000) |
|                  |                     | $P_{\text{Orf3}}$ | — | A. dieselolei B5 | CYP153 | C$_{5}$-C$_{14}$ n-alkanes | Sequence analysis only | Lian et al. (2011) |
| None             | Repressor           | almR           | — | A. dieselolei B5 | almA | C$_{22}$-C$_{30}$ n-alkanes | Gene-inactivation experiments | Liu et al. (2011), Wang and Shao (2014) |
| TetR             | Repressor           | alkX           | $P_{\text{alkX1}}$ | Dietzia sp. DQ12-45-b1 | alkW1X | C$_{7}$-C$_{24}$ fatty acids | Sequence analysis only | Wang and Shao (2012) |
|                  |                     |                | $P_{\text{alkX10}}$ | A. hongdengensis A-11-3 | alkB1 | C$_{12}$-C$_{24}$ n-alkanes | Sequence analysis only | Wang and Shao (2012) |
| GntR             | Repressor           | gntR           | $P_{\text{alkB2}}$ | A. borkumensis SK2 | alkB2 | C$_{8}$-C$_{16}$ n-alkanes | Sequence analysis only | van Beilen et al. (2004), Schneiker et al. (2006) |
|                  |                     |                | $P_{\text{alkB2}}$ | A. hongdengensis A-11-3 | alkB2 | C$_{12}$-C$_{24}$ n-alkanes | Sequence analysis only | Wang and Shao (2012) |
|                  |                     |                | $P_{\text{PA1526}}$ | P. aeruginosa RRI/PA O 1 | alkB2 | C$_{12}$-C$_{30}$ n-alkanes | Sequence analysis only | Marin et al. (2003), Smits et al. (2003) |
| LysR             | Repressor           | OrgA           | $P_{\text{alkB2}}$ | P. aeruginosa SJTD-1 | alkB2 | C$_{12}$-C$_{30}$ n-alkanes | Gene-inactivation; EMSA; DNase I footprinting; promoter-GFP plasmid assays | Ji et al. (2019) |
|                  |                      |                | $P_{\text{bmaR}}$ | T. butanivorans | bmoXYB2DC | C$_{2}$-C$_{5}$ n-alkanes | Gene-inactivation; promoter-lacZ fusion assays | Kurth et al. (2008) |
|                  |                      |                | $P_{\text{bmaR}}$ | Rhodococcus sp. BCP1 | pmABC | C$_{3}$-C$_{4}$ n-alkanes | Sequence analysis only | Cappelletti et al. (2015) |
|                  |                      |                | $P_{\text{bmaR}}$ | Rhodococcus sp. RH41 | — | — | Sequence analysis only | Cappelletti et al. (2015) |
|                  |                      |                | $P_{\text{bmaR}}$ | Rhodococcus sp. BCP1 | pmABC | C$_{3}$-C$_{4}$ n-alkanes | Sequence analysis only | Cappelletti et al. (2015) |
|                  |                      |                | $P_{\text{bmaR}}$ | M. smegmatis MC2 155 | pmABC | — | Sequence analysis only | Cappelletti et al. (2015) |
|                  |                      |                | $P_{\text{bmaR}}$ | Rhodococcus sp. strain BCP1 | smaABC | C$_{1}$-C$_{7}$ n-alkanes | Sequence analysis only | Coleman et al. (2011a), Cappelletti et al. (2015) |

* Asterisks here indicate that the promoter has not been named in previous reports. Promoter names assigned here are based on previous naming conventions.
development of alkane biosensors. A better understanding of the regulation of alkane oxidation systems is needed, via following the characterisation criteria (Vogne et al., 2010) outlined in the previous section.

5.1 Promoters and Operator Sequences

Identifying promoter and operator elements is an essential part of understanding the mechanism of action of transcription factors. Sigma factors are key players in transcriptional initiation and in bacteria σ^70 and σ^54 are the two dominant subclasses. Each recognizes and binds to specific promoter elements that allows for the recruitment and correct coordination of RNA polymerase. Promoters can be identified by locating either the σ^70 promoter elements at the −10 and −35 positions relative to the transcriptional start site (Paget and Helmann, 2003), or the σ^54 promoter elements at the −12 and −24 positions (Buck and Cannon, 1992; Francke et al., 2011). The known promoters and operator sequences associated with alkane monooxygenases are summarised in Table 3. The degree of characterisation of these elements varies across species and there is much work still to be done to understand the details of promoter-protein relationships in the hydrocarbon metabolism regulators.

The operator elements associated with monooxygenase promoters consist of repeat sequences adjacent to the −35 site. Nearly all the monooxygenase promoter sequences identified in literature show such repeats (Table 3). Sometimes these sequences are imperfect direct repeats, like the repeat sequence identified in the PalkB promoter drops to negligible levels when the operator is removed (van Beilen et al., 2004), implying that the putative AlkS binding site upstream of alkB2 is not functional, despite its strong homology to functional operator sequences.

5.2 Understanding Expression of Regulatory Genes, and the Inducer-Protein Relationship

Most alkane monooxygenase regulators are activator proteins (Moreno and Rojo, 2019), induced by the alkane directly, e.g. AlkS in P. putida Gpo1 (Kok et al., 1989), or induced by a downstream metabolite, e.g. BmoR in T. butanivorans (Kurth et al., 2008). The latter systems depend on the monooxygenase having a non-zero level of expression in the “switched off” state, so the inducer can be made from the alkane. Most alkane regulatory systems display self-regulation, such as AlkS in P. putida Gpo1. In the absence of alkanes, alkS is expressed from P_{alkS1} and expression levels are kept low by a self-repressive effect of AlkS on P_{alkS1}. Upon the addition of alkanes, AlkS activates the adjacent P_{alkS2} promoter, driving high levels of AlkS expression in a positive feedback loop, and also repressing expression from P_{alkS1} (Canosa et al., 2000). In Dietzia sp. DQ12-45-b1, fatty acids reduce AlkX repression on the alkW1 promoter in a positive feedback loop (Liang et al., 2016b). Meanwhile, in A. borkumensis AP1, the regulatory gene alkS is expressed constitutively, independent of the presence of alkanes (van Beilen et al., 2004).

Species that contain multiple monooxygenases complicate the understanding of induction processes due to possible overlaps between the inducer range and substrate range of the monooxygenases. For example, in A. dieseleolai B5, the CYP153 is expressed in the presence of C_{18}-C_{36} n-alkanes, the almA monooxygenase is induced by C_{22}-C_{36} n-alkanes, both alkB1 and alkB2 monooxygenases are induced by C_{12}-C_{26} n-alkanes, and alkB1 and almA expression can also be upregulated by the branched alkanes pristane and phytane (Liu et al., 2011). An overlapping substrate range for two AlkB-type monooxygenases is also seen in A. borkumensis AP1, where C_{10}, C_{12}, C_{14} and C_{16} n-alkanes induce both alkB1 and alkB2 (van Beilen et al., 2004). The correlations of inducers to regulators can be teased apart via the generation of deletion mutants, but this is not always straightforward, e.g., in situations where one monooxygenase can generate the metabolite inducer for another.

5.3 Alkene/Alkane Specificity

The AlkM monooxygenase of A. baylyi ADP1 is an alkane-induced system activated by the AlkR regulator (Ratajczak et al., 1998b). However, AlkR also responds strongly to the alkene octadecene, which is second only to octadecane in its strength as an inducer (Zhang et al., 2012b). The shorter alkene dodecene is also a very effective inducer for the ADP1 AlkR regulator. These findings emphasise the fact that the size of the
inducer molecule is more important than the presence of a double bond in determining whether it will act as an effective inducer. It is likely that other “alkane-inducible” regulators also respond to similarly-sized alkenes, regardless of the substrate range of the cognate monoxygenase. This is part of the relationship between inducer compounds and regulatory proteins that should be explored further.

5.4 Structures of Alkane Regulators

Transcriptional regulators have two protein domains—a DNA-binding domain, and a sensing domain. DNA-binding domains are readily identifiable because they contain conserved motifs, such as helix-turn-helix and zinc finger domains (Harrison, 1991), while sensing domains are more diverse and hard to identify based solely on bioinformatics. Diverse families of transcription factors have been recruited as alkane sensors, including proteins from the LuxR, AraC, TetR, GntR, LysR and Fis families (Table 4). Each of these families have unique organisation and features, e.g., LuxR and AraC family proteins have the DNA-binding motif at the C-terminus, while GntR and TetR family proteins have the DNA-binding domain at the N-terminus (Gallegos et al., 1997; Smitriev et al., 1999; Santos et al., 2012; Cuthbertson and Nodwell, 2013). There is also great diversity within families, e.g., the AlkR and CypR regulators of Acinetobacter sp. ADP1 and Dietzia sp. DQ12-45-b1 have sequence motifs common to AraC/XylS family proteins, but the overall sequence identity between these two proteins is only 26% (Liang et al., 2016a).

The AraC/XylS regulators also have different inducer ranges even within the same species. This is particularly noticeable when comparing the alkR, alkRa and alkRb candidates from Acinetobacter sp. ADP1 and M1. There are three distinct inducer ranges for these monoxygenases, suggesting no such pattern exists at the regulatory family level. In some cases, however, there are similarities in the inducer range of each family of regulators. The inducers for the characterised LuxR-MalT family proteins have the DNA-binding motif at the C-terminus, while GntR and TetR family proteins have the DNA-binding domain at the N-terminus (Gallegos et al., 1997; Schrijver et al., 1999; Santos et al., 2012; Cuthbertson and Nodwell, 2013). There is also great diversity within families, e.g., the AlkR and CypR regulators of Acinetobacter sp. ADP1 and Dietzia sp. DQ12-45-b1 have sequence motifs common to AraC/XylS family proteins, but the overall sequence identity between these two proteins is only 26% (Liang et al., 2016a).

5.5 Differences in Regulation of AlkB-Like Monoxygenases

Given the diversity in sequences and bacterial hosts of AlkB monoxygenases (Nie et al., 2014a), it is not surprising that the regulation of these systems is also diverse. In most cases, the 7-alkane is the inducer, however in at least one case (Rhodococcus sp. BCPI), an alcohol can also act as inducer (Cappelletti et al., 2015). Comparison of the regulation of the ADP1 and GP01 AlkM and AlkB enzymes is informative (note that the ADP1 AlkM monoxygenase is a homolog of AlkB, at 41% amino acid identity). Regulation of AlkM in ADP1 is simple, and occurs solely via the AlkR protein, which is constitutively expressed at low levels in the cell, and activates alkM expression in the presence of long-chain alkanes (Ratajczak et al., 1998a). In contrast, the situation in P. putida GP01 is more complex (Yuste et al., 1998), with the alkBFGHJKL cluster controlled by the AlkS regulator via a positive feedback loop, as described in a previous section. The regulator AlkS is partly responsible for limitations on the hydrocarbon substrate range of P. putida GP01, since the range of inducers that it recognises is narrower than the oxidation range of AlkB.

Differences in the regulation of AlkB-type monoxygenases have implications for future biosensor development. Ratajczak et al. (1998b) observed that while both medium- and long-chain 7-alkanes (C7-C19) induce alkM transcription in Acinetobacter sp. ADP1, the organism can only grow on larger alkanes (＞C12), and this induction pattern was confirmed in an ADP1 AlkR-based biosensor (Zhang et al., 2012a). Interestingly, the opposite pattern of inducers vs. growth substrates is seen in GP01. Shingler (2010) refers to these situations as “regulatory bottlenecks” for catabolic performance, and this should be kept in mind when investigating new bacterial isolates for development of hydrocarbon biosensors, i.e., the range of compounds a strain can utilise may not match the inducer range of its regulatory proteins. These phenomena are not unique to AlkB regulation; e.g., in Rhodococcus sp. BCPI, transcription of smoABCD is induced by methane despite BCP1 being unable to use methane as a carbon source (Cappelletti et al., 2015).

Comparison of the AlkS regulators in P. putida GP01 and A. borkumensis API emphasises the diversity in bacterial hydrocarbon-sensing systems. There is 32% sequence identity between these two AlkS proteins, both contain a helix-turn-helix DNA binding domain, and both recognise a binding site upstream of the promoter which contains a 20 bp inverted repeat; this motif is common to LuxR family regulators (Santos et al., 2012). The AlkS protein from P. putida GP01 can cross-activate expression of the alkBI gene in A. borkumensis (van Beilen et al., 2004), suggesting that the helix-turn-helix site
in the proteins and the cognate operator element in both species are functionally comparable. Despite these similarities, there are also many differences in the function of these two AlkS regulators, such as the constitutive vs. positive feedback mode of control of the regulator gene (see above section), and the presence in *P. putida* GP01 of an additional inhibitory global regulation network via Hfq and Crc (Moreno and Rojo, 2017). Understanding the expression patterns of the regulatory proteins themselves is a key part of characterising the regulatory system as a whole.

### 5.6 Limitations of Sequence Analysis

Bioinformatic analysis of open reading frames in proximity to monooxygenase genes can be used to identify putative regulatory genes. For example, a gene upstream of the *alkB2* monooxygenase in *A. borkumensis* SK2 encodes a GntR homolog, so a logical hypothesis might be that this GntR homolog is the regulator of the monooxygenase. More weight is added to this hypothesis when further analysis reveals that *A. hongdengensis* A-11-3, *P. aeruginosa* RR1, and *P. aeruginosa* SJTD-1 also contain GntR-like proteins in the same position relative to their *alkB2* genes (Wang and Shao, 2012; Liu et al., 2014). However, further work in strain SJTD-1 surprisingly revealed that CrgA, a LysR-type regulatory protein, was actually the regulator of the *alkB2* gene, despite the fact that this regulator is >300 kb away from the monooxygenase gene in the genome (Ji et al., 2019). This result was validated using knock-out, EMSA, footprinting, and promoter-probe assays, but it is not involved in the monooxygenase regulation.

### 5.7 Added Complexity: Further Layers of Regulation of Monooxygenase Expression

The transcription factors that respond to alkanes or their metabolites are not the only players in the regulation of alkane oxidation genes. Aside from methanotrophs and a few other obligate hydrocarbonoclastic bacteria, most hydrocarbon-assimilating species isolated to date are heterotrophic generalists, and do not preferentially utilise n-alkanes. Therefore, confirming the absence of other more easily-utilised carbon sources is just as important for these bacteria as sensing the presence of the hydrocarbon. This is managed by catabolite repression (Moreno and Rojo, 2019). There is evidence that chemotaxis towards alkanes and alkane uptake are also strongly linked to the initial sensing and subsequent metabolism of alkanes in *A. dieselolei*, further reflecting that the expression of regulatory proteins for hydrocarbon detection can be influenced by, and linked to, other cellular functions (Moreno and Rojo, 2019).

Catabolite repression has been well-studied in *P. putida* GP01, where the alkane degradation pathways are repressed by succinate, lactate, pyruvate or rich complex media like LB (Yuste et al., 1998; Dinamarca et al., 2003). Similarly, the *prmA* and *smoA* monooxygenases of *Rhodococcus* sp. BCP1 are repressed by succinate, glucose or LB medium in the presence of alkanes (Cappelletti et al., 2015), although AlkB from the same strain is not affected in the same way (Cappelletti et al., 2011). In *B. cepacia* RR10, glucose, arabinose, lactose and fructose repress *alkB* expression in the presence of the inducer tetradecanol (Marín et al., 2001). Interestingly, catabolite repression is not seen when *alkB* from *P. putida* GP01 is heterologously expressed in *E. coli*, implying differences in these mechanisms between species (Staijen et al., 1999). This aligns with the fact that the preferred carbon sources of the two differ—glucose for *E. coli*, organic acids for *Pseudomonas*. It is likely that similar global catabolite repression systems exist in other facultative hydrocarbon oxidisers that are less well-characterised.

Product repression is another layer of control of expression of monooxygenases, and this can be exerted either directly or indirectly. In the case of the butane monooxygenase of *T. butanivorans*, fatty acids generated by the butane degradation pathway (typically butyrate) directly bind to and repress the monooxygenase (Doughty et al., 2006). Fatty acids also repress CYP153 in *Dietzia* sp. DQ12-45-1b (Liang et al., 2016a) and *alkB* in *B. cepacia* RR10 (Marín et al., 2001). The ability of fatty acids to repress alkane oxidation pathways has most likely evolved to...
preventing potential toxicity from the accumulation of such products (Moreno and Rojo, 2019).

Growth phase can also impact the regulation of monooxygenase-mediated pathways. Expression of the \( \text{alkB} \) monooxygenase in \( P. \) putida Gp01 was significantly decreased when cells entered stationary phase compared to exponential phase (Yuste et al., 1998). This trend is also seen in \( A. \) borkumensis strains AP1 and SK2, where expression of \( \text{alkB} \) and \( \text{alkB2} \) decreased in stationary phase (Schneiker et al., 2006). The opposite is seen with \( \text{alkB} \) expression in \( B. \) cepacia RR10 where detectable transcripts were highest at early stationary phase (Marin et al., 2001). Meanwhile, in \( P. \) aeruginosa PA01, \( \text{alkB1} \) is strongly expressed in late exponential phase while \( \text{alkB2} \) is expressed more in early exponential phase (Marin et al., 2003). The molecular mechanisms behind these growth-phase dependent differences in regulation in these different species are unclear.

### 6 REGULATORY SYSTEMS FOR ALKENE MONOOXYGENASES

The enantioselectivity of epoxidation by alkene monooxygenases makes them valuable in the production of pharmaceutical precursors and other fine chemicals (Owens et al., 2009; Cheung et al., 2013). For this reason, research thus far has focused on the monooxygenase itself, leaving large gaps in our understanding of the regulation of these systems. An important distinction between the alkene and alkane monooxygenases is the fact that while alkane monooxygenases show activity on both alkanes and alkenes, and their regulatory proteins often accept both kinds of substrates if the carbon chain length is in the right range, the reverse is not true, and alkanes are generally not good substrates or inducers for alkene monooxygenases (Zhou et al., 1999). Because far fewer alkene-oxidising systems have been studied, the sections below are organised differently to the corresponding material on alkane regulation above, and we have taken a case-by-case approach, rather than attempting to draw general conclusions across all the systems.

**Figure 3** shows the gene configuration of the two cases discussed below—*Xanthobacter autotrophicus* Py2 (XamoABCDEF/Xamo) and *Nocardoides* sp. JS614 (EtnABCD/EtnMO). The shared subunits in these monooxygenases are 23–28% identical to each other, and it is clear the gene arrangement between the two is quite different. The sequence identity of each subunit is also compared to the alkB cluster from *Rhodococcus rhodochrous* B276, an archetypal alkene monooxygenase, for reference. As shown in the figure, the sequence identities between the B276 and JS614 subunits are much higher, between 41 and 60%. This emphasizes the similarities in both organization and sequence in Actinobacterial clusters. The sequence identities between B276 and Py2 subunits were only 23–32%. The regulation of the *R. rhodochrous* monooxygenase has not been characterized and so won’t be examined in detail in this review.

#### 6.1 Regulation of Alkene Monooxygenase in *Xanthobacter autotrophicus* Py2

The alkene monooxygenase of *Xanthobacter* Py2 (XamoABCDEF/Xamo) is a 6-component enzyme that belongs to SDIMO group 1 (**Figure 3; Table 1**). Strain Py2 was isolated on propene and Xamo was originally identified as a propene-oxidising enzyme (Ensign, 1996), but it also attacks aromatic compounds, and these can support the growth of Py2 (Zhou et al., 1999). One challenge in understanding how Xamo is regulated is uncertainty surrounding the nature of the inducer—it could be the alkene or an epoxide metabolite (Ensign, 1996); resolving this issue is tricky since the wild-type host will rapidly convert alkenes to epoxides, and to date, no effective heterologous expression systems have been reported either for the monooxygenase or the regulators. Early work on this system by Ensign (1996) showed that Xamo induction was possible even in cells grown on glucose, although Small and Ensign later reported that catabolite repression of Xamo occurred during growth on other carbon sources (Small and Ensign, 1997).

A promoter upstream of the XamoA subunit has been identified via sequence analysis, which is 60% identical to the sigma54-dependent promoter consensus sequence (Zhou et al., 1999). This is likely to be the promoter involved in the regulation of the monooxygenase expression.

Analysis of cosmid libraries provided evidence that the control of the Py2 propene monooxygenase might be part of a larger system. Complementation of mutants of Py2 with cosmid clones restored growth on epoxypropane, but interestingly, this activity became constitutive. This indicated that the 22 kb section of DNA in the cosmid contained the metabolic genes but not the necessary regulatory genes (Swaving et al., 1995), and further implied that the system was under at least one layer of negative regulation. The fact that the Py2 Xamo can also oxidise benzene, toluene and phenol suggests that this monooxygenase could be under long range control typical of aromatic hydrocarbon monooxygenases (Zhou et al., 1996).

A later shotgun proteomics study identified a candidate transcriptional regulator for the Py2 monooxygenase (Xaut_4864), that is encoded by a gene located between the Xamo gene cluster and the epoxide carboxylase gene cluster (Broberg and Clark, 2010). Xaut_4864 has a DNA binding domain similar to the MerR family of regulators but is missing 16 of 32 conserved residues typically found in MerR proteins, including the critical Cys82, and thus Xaut_4864 is thought to behave differently to MerR proteins in that it most likely does not need a metal ion to function. Xaut_4864 was identified as a protein expressed only in propylene-grown cells and thus is likely to be an activator of Xamo expression that also induces its own expression.

#### 6.2 Regulation of Alkene Monooxygenase in *Nocardoides* sp. JS614

The alkene monooxygenase of *Nocardoides* JS614 (EtnABCD/EtnMO) is quite different from that of Py2; this is a 4-component enzyme, part of SDIMO group 4 (**Figure 3; Table 1**), and enables...
growth of the host on ethene and vinyl chloride. Despite being the first ethene-oxidising bacterium to be fully genome-sequenced (Coleman et al., 2011b), the regulatory mechanisms surrounding alkene degradation in *Nocardioides* sp. JS614 are still unclear. Early work on this organism revealed an unusual starvation response (Mattes et al., 2005; Chuang and Mattes, 2007), in which alkene-starved cultures or cultures pre-grown on acetate exhibited long lag periods before growth on alkenes recommenced. A key finding was that addition of ethene oxide (epoxyethane) eliminated the lag periods, suggesting that the epoxide was the inducer, and that the starvation response was due to the bacteria being initially unable to generate this inducer from the alkene substrate.

Peptide mass fingerprinting identified seven proteins in JS614 cells expressed in the presence of vinyl chloride, epoxymethane and ethene (Chuang and Mattes, 2007); these were all identified as components of the monooxygenase or putative downstream metabolic enzymes (e.g., dehydrogenases and transferases) but no regulatory proteins were detected. Inspection of the genome sequence of JS614 reveals that a two-component sensing system is encoded immediately adjacent to the alkene catabolic genes (see Figure 3 in Mattes et al., 2010); this consists of a CdaR family

### TABLE 5 | Putative regulatory genes associated with monooxygenase clusters in *Mycobacterium chubuense* NBB4.

| Gene cluster | MO class | Substrates | Putative regulatory gene(s) | Regulatory protein family | References |
|--------------|----------|------------|-----------------------------|--------------------------|------------|
| smoXYBC1Z    | SDIMO Group 3 | C₂-C₆ alkanes and alkenes | MYCCH_RS28740 | AcoR; CadC1 HTH domain | Martin et al. (2014) |
|              |          |            | MYCCH_RS28735 | SigC-type transcription factor | Coleman et al. (2011a) |
| pmoABCD      | SDIMO Group 4 | Propene<sup>a</sup> | MYCCH_RS26755 | CdaR; GAF domain | Coleman et al. (2011a) |
|              | CuMMO    | C₂-C₆ alkanes and alkenes | MYCCH_RS28775 | CadR; His HTH domain and PEP-CTERM-box | Coleman et al. (2012) |
| smoABCD      | SDIMO Group 6 | Propane<sup>a</sup> | MYCCH_RS26425 | LuxR; REC and HTH domains | Coleman et al. (2011a) |
|              |          |            | MYCCH_RS26430 | Histidine kinase; GAF domain | Coleman et al. (2011a) |
| etnABCD      | SDIMO Group 4 | C₂-C₆ alkanes and chlorinated alkenes | EtnR1; MYCCH_RS29055 | CdaR; PucR HTH domain | Coleman et al. (2011a), Moratti et al. (2016) |
|              |          |            | EtnR2; MYCCH_RS29050 | | |
| CYP153; fdx-cyp-för | | | MYCCH_RS28400 | AraC | Coleman et al. (2011a) |
|              |          | | MYCCH_RS28420 | TetR | Coleman et al. (2011a) |
|              |          | | MYCCH_RS06610 | TetR | Coleman et al. (2011a) |

<sup>a</sup>These substrates predicted based on studies of homologous monooxygenases.
transcription activator and a DmcR family sensor kinase. Homologs of these regulators are also found adjacent to the alkene metabolic genes in other ethene-oxidising bacteria, but there is only preliminary experimental evidence to date to confirm their functions (see below Section 7).

Another clue about regulation of alkene oxidation in JS614 comes from Taylor et al. (2010), who found that the addition of ethene oxide to cultures allowed expansion of the growth substrate range to include propane and butene, which are normally cometabolised, but cannot support growth. These findings reinforce the hypothesis developed from earlier work that the epoxide is the inducer of the monooxygenase in JS614, and also highlight again the important distinction between inducers of regulatory proteins and substrates of catabolic enzymes. Understanding this distinction is critical to the successful development of biosensors and emphasises the importance of characterisation of regulatory proteins themselves, rather than trying to ascertain inducers based solely on the substrates of the monooxygenase.

7 **MYCOLICIBACTERIUM CHUBUENSE**

**NBB4: A CASE STUDY OF DIVERSE ALKANE AND ALKENE MONOOXYGENASES AND REGULATORS**

*Mycocibacterium chubuense* NBB4 (Coleman et al., 2006) is a hydrocarbon degrader originally isolated on ethene that can grow on many alkanes and alkenes, and is capable of co-metabolism of vinyl chloride and 1,2-dichloroethane (Coleman et al., 2011a). These activities are attributable to the diverse monooxygenases in NBB4 cells, including SDIMOs, AlkB, p450, and CuMMO (Table 5). Coordination of the activities of all these monooxygenases must involve complex regulation, which is thus far not understood. However, the availability of a genome sequence has allowed identification of the likely regulatory genes associated with each monooxygenase gene cluster (Figure 4; Table 5), and generation of hypotheses about their functions, as described below.

7.1 Alkane Monoxygenase Regulation: smoXYB1C1Z, hmoCAB, smoABCD

The alkane monooxygenases in *M. chubuense* NBB4 are a curious and diverse group. The smoXYB1C1Z gene cluster encodes a group 3 SDIMO that is active on C2-C4 alkanes and alkenes (Martin et al., 2014). There are two putative regulator genes near the smo genes; one of these is predicted to encode an AcoR-like regulatory protein that is distantly related to the regulators of other group 3 SDIMOs (BmoR and MmoR; 10% amino acid identity) (Coleman et al., 2011a) and has 30% identity to the closest characterised protein (acetoin dehydrogenase activator AcoR from *B. subtilis* 168 (Ali et al., 2001). The other regulator near the smo genes encodes a SigC-like sigma factor; there is no parallel for this in the other group 3 SDIMOs. There is thus far no experimental evidence to confirm that either of these regulatory proteins control smoXYB1C1Z expression, but their positioning relative to the smo genes indicates that this is likely.

The hmoCAB genes encoding a CuMMO are located immediately upstream of the smoXYB1C1Z genes in the NBB4 genome (Coleman et al., 2012). This organisation and the common substrates shared by both these monooxygenases (gaseous alkanes and alkenes) suggests that they may be co-regulated. Previous hypotheses about the relationship between the Smo and Hmo enzymes include the possibility that they might be a high- and low-affinity pair or that they are expressed in environments with different cofactor metal availabilities (Martin et al., 2014). The hmoCAB cluster is preceded by another acoR-like regulator gene, which suggests that the two clusters may have their own independent regulatory systems. It is possible that there are multiple layers of regulation over these two monooxygenases; this would be consistent with a hypothesis that their expression is controlled by both substrate and cofactor availability.

The smoABCD cluster of NBB4 encodes a group 6 SDIMO which is thus far uncharacterised except for bioinformatics (Coleman et al., 2011a). The smoABCD genes are adjacent to a gene encoding a LuxR-type regulator (van Kessel et al., 2013), which is oriented “head-to-head” with a gene encoding a sensor kinase, with a 29 bp overlap between the 3′ ends of the two genes. It is common for LuxR regulators to belong to two-component transduction systems with a sensor kinase, so it is plausible that this pair of proteins act together in this way. The closest homolog of this NBB4 regulator with a known function is the cold-shock regulator DesR in *B. subtilis* 168, at 26% amino acid identity (Cybulski et al., 2004).

Analysis of the regulation of the smoABCD monooxygenase in *Rhodococcus sp.* BCP1 gives clues to the regulation of smoABCD in NBB4; these two SDIMOs are very similar, sharing 85–94% amino acid identity in the enzyme subunits. In strain BCP1, the promoter region contains the same core inverted repeat, potential binding sites for this pair of proteins control smoXYB1C1Z expression, but their positioning relative to the smo genes indicates that this is likely.

The CYP153 gene cluster in NBB4 is flanked by genes encoding two regulator proteins—a TetR-like protein and an AraC-like protein (Table 5). The araC-like gene and the CYP cluster are divergently transcribed, with just 115 bp separating them; this organisation is consistent with the AraC homolog being the regulator responsible for controlling CYP153 in NBB4, and with promoters of these two genes being located in this 115 bp region. A TetR-like protein is encoded by a gene adjacent to the AlkB-like monooxygenase gene cluster in NBB4. This protein is 31 bp downstream of the alkB and putative rub genes, and oriented in the same direction as both alkB and rub genes.  

---

**Table 5**

| Gene Cluster | Description |
|--------------|-------------|
| smoXYB1C1Z   | Alkane monooxygenase |
| hmoCAB       | Alkene monooxygenase |
| smoABCD      | Alkene monooxygenase |

---
There are many similarities in the configuration of CYP153 and alkB in NBB4 and in Dietzia sp. strain DQ12-45-1b. The alkB homolog in Dietzia (alkW1) is under the control of a TetR-family repressor; this kind of regulation of alkB appears to be unique to the Actinobacteria (Liang et al., 2016b). The regulator alkX is 19 bp downstream of the rubredoxin gene in Dietzia compared to 31 bp in NBB4. The intergenic DNA region upstream of alkB in both species is 53% identical. The ~35–10 sites of the alkW1 promoter each differ by just one base from the equivalent alkB promoter in NBB4, and the operator sequences in NBB4 have strong homology and similar spacing to those seen in Dietzia (Liang et al., 2016a). The configuration of the araC family gene in the CYP153 cluster of Dietzia is also very similar to the situation in NBB4, i.e., oriented divergently and separated from the ferredoxin gene by 118 bp (Liang et al., 2016b).

If the functional parallels between these monooxygenases in NBB4 and Dietzia reflect the sequence similarities, one could infer that the CYP153 and alkB genes in NBB4 work as part of a team to efficiently tackle alkanes across a wide range of chain lengths. Nie et al. described how the inducer and substrate range of CYP153 and alkB in Dietzia are complementary, i.e., CYP153 hydroxylates n-alkanes < C10, while alkW1 acts on > C10 substrates (Nie et al., 2014b). It is worth noting that NBB4 has two further CYP153 genes (not discussed in detail here), which adds further complications to understanding the relationship between enzymes, inducers and substrates (Coleman et al., 2011a), and highlights the fact that much further work is needed to understand the regulation of alkB and CYP153 in NBB4 and similar bacteria.

7.3 Alkene Monooxygenase Regulation: pmoABCD, etnABCD

The substrate ranges of the NBB4 ethene and propene monooxygenases (EtnABCD and PmoABCD, respectively) are similar; both enzymes show activity on C₂−C₈ alkenes, with stronger activity on gaseous alkenes (C₂−C₄) (McCarl et al., 2018). The fact that strain NBB4 possesses two distinct SDIMO enzyme systems with very close overlap in substrate ranges is unusual and poses questions about why this genotype has evolved and how these genes are regulated. This is an excellent example of why monooxygenase regulatory systems warrant further investigation.

The pmoABCD gene cluster in Mycolicibacterium NBB4 has a CdaR family regulator encoded by a gene immediately upstream of the monooxygenase, and divergently oriented from these. This family of regulators was originally studied for their role in sugar diacid regulation in E. coli, and are typically activators containing a helix-turn-helix domain at the C-terminus (Monterrubio et al., 2000). A similar gene organisation is seen for pmoABCD of Mycobacterium M156 (Coleman et al., 2011a), although neither regulatory system has been characterised beyond DNA sequencing and bioinformatic analysis.

Regulation of etnABCD is likely to be done by a pair of regulators EtnR1/EtnR2, encoded by genes upstream of etnABCD. The EtnR1/EtnR2 pair display the typical features of a bacterial two-component regulatory system in which a DNA-binding protein (transcriptional regulator) is phosphorylated by a histidine protein kinase (sensor protein). EtnR1 is a CdaR family protein that contains a DNA-binding helix-turn-helix domain. EtnR2 is distantly related to the DcmR dichloromethane-sensing regulator from Methylobacterium DM4 (16% aa identity, 27% aa similarity), and contains a MADS domain, which is involved in sensing hydrocarbon derivatives in both methanogens and methylo trophs (Anantharaman and Aravind, 2005).

Preliminary research has confirmed an interaction between EtnR1 and a DNA segment containing etnP, its putative cognate promoter, via EMSA (Moratti et al., 2016). Transcriptomics data (unpublished) also shows a significant upregulation (approx. 8-fold) of both etnR1 and etnR2 after exposure of NBB4 cells to ethene; this is consistent with these genes being positive regulators that activate their own expression in a feedback loop, similar to the situation discussed above with AlkS in P. putida GPo1. Importantly, the transcriptomics data from NBB4 shows no reads from the short intergenic region thought to contain the etnP promoter sequence, consistent with this region driving expression but not itself being expressed.

The etnR1 and etnR2 regulators are highly conserved (73–78% aa identity) across many ethene-oxidising isolates, including Mycolicibacterium JS623 (etnR1: WP_015305844, etnR2: WP_015305843), Mycolicibacterium tusciae JS617 (etnR1: WP_006247394, etnR2: WP_006247393), and Mycolicibacterium rhodesiae JS60 (etnR1: WP_014211282, etnR2: WP_014211281). More divergent homologs of these genes (44–59% aa identity) can also be found in many other Actinobacteria, including but not limited to Streptomyces thermoautotrophicus H1 (etnR1: WP_066887198, etnR2: WP_079045917) and Amycolatopsis SYSUP005 (etnR1: WP_101434350, etnR2: WP_158242445). Nearly all these bacteria also contain etnABCD homologs near the regulator genes, although most have not been tested for alkene oxidation. Unlike the regulation of alkB described above, this suggests that there is overall consistency in the regulation of alkene oxidation across all gram-positive bacteria.

8 CONSTRUCTING A HYDROCARBON BIOSENSOR: GENERAL CONSIDERATIONS

There are two general approaches for constructing whole-cell transcription-factor based biosensors. The first approach involves the assembly and expression of the regulatory gene(s), promoters and reporter gene in a plasmid in an appropriate heterologous host strain; this approach allows for more control over expression of the different elements because promoters, ribosome binding sites and other features can be easily and individually modified. The second approach involves the integration of the reporter gene into the genome of the native host, either immediately upstream of the metabolic genes or replacing these; in this situation, there is a gain in stability of the system but the trade-off is that there is less flexibility and construction is more difficult. There are advantages and disadvantages of both approaches, depending on the intended applications of the system (Carpenter et al., 2018).
For environmental applications e.g. in monitoring bioremediation, precise detection at a single cell level is not necessary, and detecting average hydrocarbon concentrations in a population of biosensor cells is acceptable; note that given current legal and biosafety considerations it is more likely that these analyses would be done in vitro rather than in situ. In a metabolic engineering context, detection at a single-cell level may be desirable e.g. for screening clones in a directed evolution library (Dietrich et al., 2012), and in such applications, a resistance gene may be used instead of a reporter gene to select for the target phenotype (Dietrich, 2011; Carpenter et al., 2018). Challenges to using biosensors at a single-cell level include variations in expression level in individual cells, the potential impact on cell viability, and the impact of the surrounding population on single cells (Tecon and van der Meer, 2006; Carpenter et al., 2018).

The type of replication origin and the host species are key considerations for plasmid-based biosensor construction. Different origins of replication will give different copy numbers, which will impact the signal strength produced by the biosensor (Carpenter et al., 2018). Plasmid-based systems enable the use of hosts that are non-pathogenic, fast to grow and easy to transform, but possible disadvantages include a lack of robustness outside the laboratory, problems with the inducible promoter not being recognised correctly, and potential codon usage issues (Carpenter et al., 2018). Codon optimisation or harmonisation methods can be applied to the regulatory protein and/or to the reporter gene to overcome codon usage problems.

Chromosomally-integrated biosensors will be single-copy systems, and the output signal strength thus correspondingly lower than for plasmid-based systems. Chromosomally-integrated biosensors in non-standard hosts can be more useful for environmental applications, or under harsher physicochemical regimes where traditional host strains would not survive (Jiang et al., 2020). Chromosomal biosensors do not require any selection pressure to maintain, which offers a practical advantage in environmental applications (de las Heras and de Lorenzo, 2010). A major disadvantage of chromosomally-integrated biosensors is the difficulty of construction, since non-standard microbial species have more limited genetic tools and methods available. Homologous recombination has typically been used to make the precise insertions required, but more modern methods such as CRISPR will increasingly replace this (Arroyo-Olarte et al., 2021).

Several problems with chromosomal biosensors relate to interference of different kinds; this may arise from other inducers, other enzyme systems, other global regulators, or other cells in the environment (Su et al., 2011; Jiang et al., 2020). Wild-type microbes that have been engineered to contain a hydrocarbon biosensor circuit may might have other enzyme systems that can metabolise the inducer; this will lead to a depressed response or false negative signals. Less well-studied hosts may contain global control networks that interact with the promoter of interest in unknown ways. For environmental applications or mixed culture conditions, major unknown factors arise from the possible interactions between the

### TABLE 6 | Compilation of existing hydrocarbon biosensors.

| Sensing components | Reporter | Host | Type | Inducers and detection limits | Best inducer | Intended application | References |
|---------------------|----------|------|------|-------------------------------|--------------|----------------------|------------|
| alkS/PalK from P. putida GPo1 | luxAB | E. coli DH5a | Plasmid | C8-C10 alkanes, 4–100 nM | C8 alkane | Monitoring bioremediation | Sticher et al. (1997) |
| gfp | E. coli DH5a | Plasmid | C9 alkane, 10 nM to 1 μM | C9 alkane | Monitoring bioremediation | Jaspers et al. (2001) |
| gfp | A. borkumensis | Plasmid | C4 alkane, petrol | N/A | Monitoring bioremediation | Sevilla et al. (2015) |
| gfp | E. coli DH108i | Plasmid | C4–C6 alkanes | N/A | Monitoring bioremediation | Reed (2012) |
| sf-gfp | E. coli DH5a and E. coli HB101 | Plasmid | C9, C10, C11 alkanes | N/A | Intracellular alkane detection; characterisation of AlkL | Grant et al. (2014) |
| alkS/alkB from P. putida GPo1 | luxAB | E. coli TOP10 | Plasmid | C4 alkane, alcohols, aldehydes 10–200 μM | C4 alkane | Intracellular alcohol and aldehyde detection in metabolic engineering | Minak-Bernero et al. (2004) |
| alkS/alkB from A. borkumensis | luxAB | E. coli DH5a | Plasmid | C4 alkane, crude oil, 5 nM | C4 alkane | Monitoring bioremediation | Kumari et al. (2011) |
| alkS/alkB from A. borkumensis | luxAB | A. baylyi ADP WH2 | Chromosomal | C7–C16 alkanes and alkenes, 100 μM | C8 alkane | Monitoring bioremediation | Zhang et al. (2011), Zhang et al. (2012a), Li et al. (2013) |
| luxAB and gfp | A. baylyi ADP1 | Chromosomal | C12 alkane | Intracellular alkane detection in metabolic engineering | Minak-Bernero et al. (2004) |
| gfp | E. coli BL21 DE3 | Plasmid | C16, C17 alkanes | N/A | Intracellular alkane detection | Sonta et al. (2012), Lehtinen et al. (2017) |
| bmOR/PmoO from T. butanivorans sp. nov | tetA-gfp fusion | E. coli DH1 ΔadhE | Plasmid | C2,C4 alcohols, 0.01–100 mM | C2 alcohol | Intracellular alcohol detection in metabolic engineering | Dietrich et al. (2012) |
biosensor and other microbes; these may create or consume inducers, or enhance/inhibit growth or metabolism in other ways.

9 EXISTING ALIPHATIC ALKANE BIOSENSOR PROTOTYPES

There have been eleven unique hydrocarbon biosensors developed to date (Table 6) using components from four alkane-degrading bacteria (P. putida, A. borkumensis, A. baylyi, and T. butanivorans). These have been designed either for monitoring the bioremediation of oil spills (detection of extracellular alkanes) or to increase our understanding of hydrocarbon metabolism and/or engineer strains for hydrocarbon synthesis (detection of intracellular alkanes, alcohols or aldehydes). The bmo/R/P bmo biosensor derived from T. butanivorans is notable in this collection because it is the only example to date of a SDIMO-associated regulator employed as a biosensor (Dietrich et al., 2012), and because it provides an important reminder that these systems are in some cases induced by metabolites (in this case 1-butanol) rather than alkanes; this is not necessarily a disadvantage, and reflects the broader potential of monooxygenase regulatory systems in biotechnology.

9.1 Plasmid-Based Alkane Biosensors

More plasmid-based alkane biosensors have been developed than chromosomally-integrated ones, and the majority of these are based on AlkS/P alk from P. putida GP01 (Table 6). The output signals from these include GFP or LuxAB (detection), or TetA (selection) (Dietrich et al., 2012). Some have the input and output components on separate plasmids (Sticher et al., 1997; Reed et al., 2012), while others use single-plasmid systems (Jaspers et al., 2001; Minak-Bernero et al., 2004; Kumari et al., 2011; Sevilla et al., 2015) (See Figure 1 of Reed et al. (2012) and Figure 1 of Jaspers et al. (2001) for representative plasmid-based biosensor schematics). One novel biosensor uses the monooxygenase genes alkfi/alkbFG to convert alkanes to aldehydes, which then support luciferase activity (Minak-Bernero et al., 2004). Another example of note uses AlkR/P alk from A. baylyi ADP1 to measure pentadecane and heptadecane biosynthesis from the ado and aar genes of Synechococcus integrated into an E. coli host, via a GFP output (Wu et al., 2015). Most plasmid-based biosensors are maintained in lab strains of E. coli such as DH5a (Sticher et al., 1997; Jaspers et al., 2001; Kumari et al., 2011), DH10B (Reed et al., 2012) or TOP10 (Minak-Bernero et al., 2004). An exception is one system using A. borkumensis as a host; this had a slower response time, but higher sensitivity (Sevilla et al., 2015).

9.2 Chromosomal Biosensors

Only two chromosomally-integrated alkane biosensors have been developed to date, both in A. baylyi ADP. The first of these (named ADPWH alk) has luxCDABE integrated upstream of alkM1, controlled by AlkR/P alkM1 (Zhang et al., 2012b), and has been used for biosensing in an oil-contaminated sample (Li et al., 2013) and further engineered for increased functionality by immobilisation on magnetic nanoparticles (Zhang et al., 2011) (see Figure 1A in Zhang et al., 2012a for a representative chromosomal biosensor schematic). The other chromosomally-integrated biosensor allows simultaneous detection of intracellular alkanes and aldehydes. In this sensor, a cassette containing a P alkM gfp fusion generates a fluorescent output for alkane detection while a luxAB cassette reports on aldehyde concentrations (long-chain aldehydes are the substrate for the luciferase). The system allows reporting on both alkane oxidation to aldehydes and alkane synthesis from aldehydes, since it also contains IPTG-inducible alkane biosynthesis genes (aar and ado) (Lehtinen et al., 2017).

9.3 Benefits and Limitations

The low bioavailability of hydrocarbons is a major challenge for alkane biosensors, which all under-report hydrocarbon concentration by around 20% (Sticher et al., 1997; Kumari et al., 2011; Zhang et al., 2012a; Li et al., 2013). Kumari et al. (2011) found that a longer incubation time was required to increase the bioavailability, and hence detection, of longer chain alkanes (>C11), while Sticher et al. (1997) attributed the underestimation to the presence of unknown inhibitor compounds. Li et al. (2013) argued that, despite these shortcomings, the benefits of biosensors still made them valuable, and in the case of their system, the short detection time (0.5–4 vs. 48 h for GC/MS) and the small sample size (1 ml vs. 500 ml required for GC/MS) were major advantages. Biosensors can be surprisingly robust; e.g., Zhang et al. did not observe any instability or loss of function in their ADPWH alk biosensor, which still worked well after storage for a month in water at 4°C (Zhang et al., 2012a).

10 DEVELOPMENT AND OPTIMISATION OF HYDROCARBON BIOSENSORS

An increased understanding of regulatory components opens the door to protein and DNA engineering to optimise the specificity, sensitivity, dynamic range, detection range and response time of these systems for use as biosensors (Ding et al., 2021). The flexibility of regulatory proteins in detecting multiple inducers can be seen as a double-edged sword (Diplock et al., 2010); this may be useful in nature for a bacterium to respond to multiple possible carbon sources, but may not be ideal for biosensing of specific analytes. Factors targeted for optimisation include host strain, promoter sequence, replication origin, ribosome binding site, protein-promoter binding sites, and the sequence of the regulatory protein; the latter may also involve addition of degradation tags to reduce the metabolic burden on the cell (Ding et al., 2021).

Techniques such as site-directed mutagenesis, random PCR mutagenesis and DNA shuffling can be used to generate libraries of regulator variants with improved functions (van der Meer and Belkin, 2010). Regulatory proteins of the AraC/XylS and TetR families have successfully been mutated to alter binding specificity (Galvão et al., 2006); e.g., the R41G mutation in XylS increases the response to 2-ethylbenzoate, while reducing the response to 2-methylbenzoate (Galvão et al., 2006). The
crystal structure of a protein can reveal optimal sites for mutation; this approach was successfully used in the case of DntR, a salicylate-induced activator in *P. putida*, to identify the binding pocket (Lönneborg et al., 2007).

An impressive example of optimisation involved the directed evolution of *AlkS* in an *E. coli*-based *AlkS/PalkB* biosensor (Reed et al., 2012). Two rounds of error-prone PCR resulted in the identification of several mutations that conferred improved response to short-chain alkanes (C5-C9), with the best mutant showing a five-fold increase in fluorescence output in response to hexane compared to wild type *AlkS*. The Q410K and S470T mutations were present in the two best mutants; the latter increases the bulkiness in the putative alkane binding pocket, discouraging binding of larger alkanes (Reed et al., 2012). This study provides clear evidence for the usefulness of directed evolution approaches to alter the substrate range and specificity of alkane-sensing systems.

A novel approach to biosensor development was taken by Zhang et al. (2011), who functionalised the *A. baylyi* ADPWH_alk biosensor using magnetic nanoparticles, to allow for remote manipulation of the reporter cells; this allows biosensor cells to be collected for re-use after deployment in a complex environment by application of a magnetic field. The magnetic nanoparticles had no adverse impact on cellular function or alkane detection and could be attached with an efficacy of 99.96%. The ADPWH_alk biosensor has a shorter response time (30 min) compared to a previously-developed biosensor using the same AlkR/P<sub>alkB</sub> components (10 h) (Ratajczak et al., 1998a); this is potentially attributable to three fortuitous point mutations located near the AlkR binding site (Zhang et al., 2012b).

The 1-butanol biosensor constructed using BmoR/P<sub>BMO</sub> and GFP has also been optimised (Dietrich, 2011). In this case, lower temperature (25°C) and lower levels of BmoR expression resulted in a more robust biosensor. A synthetic ribosome binding site for *gfp* expression was also beneficial, resulting in higher fold-induction and better dynamic range. Other parameters that were optimised included: induction time (early exponential phase was best), inducer concentration (alcohol toxicity was observed above 40 mM 1-butanol), host strain (*ΔadhE* strain lowered background fluorescence) and plasmid origin of replication (a low copy replicon gave no fluorescence).

Calibration of biosensors to traditional detection techniques is a crucial step towards real-world application, as it proves the integrity and reliability of the device. This was done with the *A. baylyi* ADPWH_alk biosensor, by comparing it to GC/MS measurements (Li et al., 2013). Two contaminated soil samples (>5,000 mg petroleum/kg soil) were analysed along with two clean soil samples from an adjacent site. It was found that the biosensor reported ~20% lower levels of oil compared to the GC/MS, which could be attributed to the low bioavailability of alkanes, as discussed above (Sticher et al., 1997). Inducer bioavailability has been a consistent problem for biosensor development, but a counter-argument can also be made that the bioavailable fraction is more relevant than the total hydrocarbons for determining ecotoxicity (Tecon and Van der Meer., 2008). Enhancing the uptake of alkanes may increase their apparent bioavailability; e.g., co-expression of the AlkL transporter in an *E. coli* strain expressing the AlkB monooxygenase resulted in a 100-fold increase in oxidation of large alkanes (>C₁₂) (Grant et al., 2014).

A recent biosensor optimisation study examined the impact of different host organisms on the performance of the biosensor. A dual plasmid biosensor containing *AlkS* and *P<sub>alkB</sub>* with a GFP reporter was transformed into several alkane-assimilating marine bacteria, and also into *E. coli* (Sevilla et al., 2015). Although *E. coli* had the fastest detection rate, *A. borkumensis* was the best candidate overall, giving the most sensitive detection of octane (detection limit of 0.5 µM), and also effectively detecting C₇-C₉ alkanes at 0.012% v/v concentration in a saltwater sample. This is an excellent example of how a hydrocarbon biosensor can be

---

**FIGURE 5 |** Status of development of octane (A), medium-chain alkane (B), and alkene biosensors (C), with priority research areas yielding maximum impacts indicated by the dark outlined arrows.
11 BRINGING HYDROCARBON BIOSENSORS TO MARKET

Despite the above-described successes with optimisation, no hydrocarbon biosensors are commercially available at the time of writing. The deployment phase of a whole-cell microbial biosensor is challenging, particularly for devices that have environmental applications. The use of microbial biosensors in the environment and the associated challenges with their deployment have been extensively reviewed (D’Souza, 2001; van der Meer et al., 2004; Harms et al., 2006; de las Heras and de Lorenzo, 2010; Coleman et al., 2011a; de las Heras and de Lorenzo, 2012; Plotnikova et al., 2016; Shemer and Belkin, 2019; Hicks et al., 2020; Jiang et al., 2020); here we will focus on a few key engineering considerations relevant to the specific case of hydrocarbon biosensors. The potential of these systems has been long acknowledged, and yet the same barriers seem to stand in the way, decades after they were first recognized. These barriers include legal considerations surrounding the release of genetically manipulated bacteria, functional issues about the investment of time and resources required to yield sufficiently optimised systems for field application (Sadana and Sadana, 2011), and functional limitations of these biosensors, such as bioavailability concerns (van der Meer, 2016).

Two linked challenges for biosensor commercialisation are culture scale-up and immobilisation. Large numbers of cells need to be grown in pure culture and kept viable and at high activity; continuous culture is a good option here, but this requires more complex equipment and maintenance and has a higher risk of contamination (Bjerketorp et al., 2006). Cells then need to be preserved, immobilised and/or contained such that they can be deployed safely and effectively without compromising their functionality; these methods may include freeze- or vacuum-drying, immobilisation and/or encapsulation. Cells can be immobilised on solid surfaces like optical fibres or microchips, or encapsulated in soft materials like hydrogels, sol-gel, carrageenan, alginate, polyacrylamide, oxysilane or polyvinyl alcohol (Liu et al., 2007). The choice of encapsulation/immobilisation methods impacts many aspects of biosensor function, such as the rates of gas and solute diffusion, biosensor response time, cell viability, and shelf-life.

Striking a balance between functionality and biosafety is an ongoing challenge for all whole cell-based biosensors that have intended applications outside of the laboratory. The inclusion of toxin/anti-toxin systems, non-canonical amino acids, kill switches, engineered auxotrophy, or conditional origins of replication are examples of methods to prevent unintentional gene transfer from the biosensor to the environment (Wright et al., 2015; Hicks et al., 2020). Given the high cost and labour inputs into their production, the potential to reuse or recycle these biosensors should also be considered (e.g., see above-described example using magnetic nanoparticles (Zhang et al., 2011). If the biosensor is intended for single use, biodegradation options should be assessed, e.g., by choosing a biodegradable immobilisation surface. Testing the function of a biosensor in its intended application environment is very important since other chemical compounds or microbes in the target sample could inhibit or confound its response (van der Meer and Belkin, 2010); elucidating these interactions should be an early focus of the deployment phase of biosensor development.

12 FUTURE RESEARCH PRIORITIES AND CONCLUSIONS

One useful analogy for the landscape of aliphatic hydrocarbon biosensor development is the research funnel. Identifying the bottlenecks in this funnel is the key to efficiently directing research for different kinds of biosensors. At the top of the funnel, there is an abundance of putative monooxygenase and other hydrocarbon catabolic genes in databases that have been tentatively identified using bioinformatics. There is then a steep decline at the next level, representing experimentally characterised catabolic systems, with only a few dozen monooxygenases reaching this milestone. The identification and characterisation of the regulatory systems at the next two lower levels filter the candidates even further. Only a handful of biosensors make it to proof-of-concept stage, with optimisation and development attempts made on an even smaller subset of these, and thus far no candidates have been deployed as commercial products.

The large number of putative monooxygenase genes already existing in databases reflects the constantly decreasing cost of DNA sequencing, and the availability and accessibility of bioinformatic software; this part of the funnel is not the best focus for efforts to develop biosensors. The characterisation of monooxygenases is also not a major limiting factor since many representatives of different monooxygenases have now been at least partially characterised. In contrast, there are strong arguments for focusing research efforts on identifying and characterising regulatory genes and promoters. Extrapolating from previous well-studied systems is of limited usefulness, as previous work suggests different species may have unique regulatory mechanisms, even for similar catabolic genes (Moreno and Rojo, 2019). More effort is needed to overcome a bias in the literature towards C5-C18 alkane sensing systems and AlkB or CYP enzymes; this has led to a neglect of the systems responding to smaller alkanes and alkenes, especially those associated with SDIMOs and CuMMOs. Investing in research on thorough characterisation of regulatory systems will give increased understanding of how these systems function in nature, leading to better biosensors, and also helping to remove legal barriers to implementation in the field.

Very few hydrocarbon regulatory systems have met all five proposed criteria for complete characterisation (van der Meer and Belkin, 2010), and these knowledge gaps will limit the development of biosensors. The complexities of these systems should not be underestimated, and much more research is required to appreciate their intricacies. Untangling the regulation of different hydrocarbon catabolic genes that exist...
within a single strain is an especially useful avenue to pursue (Coleman et al., 2011a); this will give insights into the behaviour of biosensor circuits which must function in the presence of other hydrocarbon catabolic genes and regulators, and will inform the construction of more complex systems that integrate multiple sensing systems into a single cell.

Continued research attempts at the proof-of-concept stage of hydrocarbon biosensors are also warranted. This may involve directed evolution or site-specific modifications of regulatory components to yield biosensors with improved qualities, or proceeding with wild-type sensor systems, many of which already have good sensitivity and specificity (van der Meer and Belkin, 2010). Challenges at the proof-of-concept stage are often due to issues with robustness, shelf-life, and applicability to different real-world environments (Hicks et al., 2020) e.g., how to safely immobilise the cells while maintaining their function. Investigation into the localisation of wild type regulatory proteins could also be valuable, as it can influence the response time of the biosensor (Ding et al., 2021).

Three specific recommendations for research can be drawn from this review (Figure 5). Firstly, there are already two excellent octane biosensors, the plasmid-based $\text{AlkB}$/$\text{PalkB}$/GFP biosensor in E. coli DH10B (Reed et al., 2012) and the chromosomal $\text{alkR}/\text{PalkM}/\text{luxAB}$ sensor in A. baylyi (Zhang et al., 2011; Zhang et al., 2012a; Li et al., 2013); these are ready to progress to the final stage of development and deployment, which should focus on testing their robustness, sensitivity and selectivity in various real-world environments, and on finding the best methods of immobilising or encapsulating the cells. Secondly, several medium-chain alkane biosensor systems that work well in the lab (Zhang et al., 2012b; Sevilla et al., 2015) should now be progressed to the optimisation stages, e.g. to make a suite of sensors, each with selectivity for different single analytes. Finally, the molecular details of alkene-sensing systems need to be much better characterised, since our understanding of these is still rudimentary.

New developments in synthetic biology have greatly expanded the possibilities for hydrocarbon biosensors. The combination of synthetic biology methods with the wealth of novel sequences that continue to appear in genetic databases promises an exciting future for this research field. However, to maximise these possibilities our efforts must be effectively targeted at the appropriate development stages for each biosensor. Making successful commercial biosensors for deployment in real environmental or industrial contexts will require genuinely interdisciplinary efforts including microbiologists, molecular biologists, biochemists, structural biologists, engineers, materials scientists, and mathematical modellers. The microbes have provided the raw materials, but now we must provide the ingenuity and the effort to complete these tasks.

**AUTHOR CONTRIBUTIONS**

CM, NC, and CS contributed to the design of this review and defined the scope. CM carried out the main data collection and analysis. CM and NC drafted the manuscript. NC and CS provided critical revision of the draft.

**FUNDING**

This research is funded by a Research Training Program scholarship provided by the Australian Government, with additional financial, technical, and supervisory support from the CSIRO Future Science Platform in Synthetic Biology.

**REFERENCES**

Alonso, H., and Roujeinikova, A. (2012). Characterization and Two-Dimensional Crystallization of Membrane Component AlkB of the Medium-Chain Alkane Hydroxylase System from Pseudomonas Putida GPo1. *Appl. Environ. Microbiol.* 78 (22), 7946–7953. doi:10.1128/AEM.02053-12

Anantharaman, V., and Aravind, L. (2005). MEADS and Pocr Are Novel Domains with a Predicted Role in Sensing Simple Hydrocarbon Derivatives in Prokaryotic Signal Transduction Systems. *Bioinformatics* 21 (12), 2805–2811. doi:10.1093/bioinformatics/bti418

Arenghi, F. L. G., Barbieri, P., Bertoni, G., and de Lorenzo, V. (2001). New Insights into the Activation of O-xylene Biodegradation in Pseudomonas Stutzeri OX1 by Pathway Substrates. *EMBO Rep.* 2 (5), 409–414. doi:10.1093/EMBOREPORTS/KVE092

Arenghi, F. L. G., Pinti, M., Galli, E., and Barbieri, P. (1999). Identification of the Pseudomonas Stutzeri OX1 Toluene- O-Xylene Monoxygenase Regulatory Gene (touR) and of its Cognate Promoter. *Appl. Environ. Microbiol.* 65 (9), 4057–4063. doi:10.1128/AEM.65.9.4057-4063.1999

Arroyo-Olarte, R. D., Bravo Rodriguez, R., Morales-Ríos, E., and Morales-Ríos, E. (2021). Genome Editing in Bacteria: CRISPR-Cas and beyond. *Microorganisms* 9 (4), 844. doi:10.3390/MICROORGANISMS9040844

Banerjee, R., Jones, J. C., and Lipscomb, J. D. (2019). Soluble Methane Monoxygenase. *Annual Review of Biochemistry* 88, 409–431. doi:10.1146/annurev-biochem-013118-111529

Baptist, J. N., Gholson, R. K., and Coon, M. J. (1963). Hydrocarbon Oxidation by a Bacterial Enzyme System. *Biochim. Biophys. Acta* 69, 40–47. doi:10.1006/0066-3002(63)91223-x

Bjerketorp, J., Håkansson, S., Belkin, S., and Jansson, J. K. (2006). Advances in Preservation Methods: Keeping Biosensor Microorganisms Alive and Active. *Curr. Opin. Biotechnol.* 17 (1), 43–49. doi:10.1016/j.copbio.2005.12.005

Boase-Beier, J., Fawcett, A., Wilson, P., and Schauer, F. (2014). Introduction. *AMB Express* 4 (1), 1–10. doi:10.1186/S13568-014-0075-2

Broberg, C. A., and Clark, D. D. (2010). Shotgun Proteomics of Xanthobacter Autotrophicus Py2 Reveals Proteins Specific to Growth on Propylene. *Appl. Microbiol.* 192 (11), 945–957. doi:10.1128/AEM.00203-010.0623-3

Browning, D. F., and Busby, S. J. W. (2016). Local and Global Regulation of Transcription Initiation in Bacteria. *Nat. Rev. Microbiol.* 14, 636–650. doi:10.1038/nrmicro.2016.103

Buck, M., and Cannon, W. (1992). Specific Binding of the Transcription Factor Sigma-54 to Promoter DNA. *Nature* 358 (6385), 422. doi:10.1038/358422A0

Burrows, K. J., Cornish, A., Scott, D., and Higgins, I. J. (1984). Substrate Specificities of the Soluble and Particulate Methane Mono-Oxygenases of Methylosinus Trichosporum OB3b. *Microbiology* 130 (12), 3327–3333. doi:10.1099/00221287-130-12-3327

Canosa, I., Sanchez-Romero, J. M., Yuste, L., and Rojo, F. (2000). A Positive Feedback Mechanism Controls Expression of AlkB, the Transcriptional Regulator of the Pseudomonas Oleovorans Alkane Degradation Pathway. *Mol. Microbiol.* 35 (4), 791–799. doi:10.1046/j.1365-2958.2000.01751.x

Cappelletti, M., Fedi, S., Fraschini, D., Ohtake, H., Turner, R. J., and Zannoni, D. (2011). Analyses of Both the AlkB Gene Transcriptional Start Site and alkB
Promoter-Inducing Properties of Rhodococcus Sp. Strain BCP1 Grown on N-Alkanes. Appl. Environ. Microbiol. 77 (5), 1619–1627. doi:10.1128/AEM.00871-10

Cappelletti, M., Presentato, A., Milazzo, G., Turner, R. J., Fedi, S., Frascari, D., et al. (2015). Growth of Rhodococcus Sp. Strain BCP1 on Gaseous N-Alkanes: New Metabolic Insights and Transcriptional Analysis of Two Soluble Di-iron Monooxygenase Genes. Front. Microbiol. 6, 393. doi:10.3389/fmicb.2015.00393

Carpenter, A., Paulsen, I., and Williams, T. (2018). Blueprints for Biosensors: Design, Limitations, and Applications. Genes 9, 375–403. doi:10.3390/genes9080375

Cerniglia, C. E., and Crow, S. A. (1981). Metabolism of Aromatic Hydrocarbons by Yeasts. Arch. Microbiol. 129 (1), 9–13. doi:10.1007/BF00417170

Champreda, V., Zhou, N.-Y., and Leak, D. J. (2004). Heterologous Expression of Alkene Monooxygenase Components from Xanthobacter autotrophicus Py2 and Reconstitution of the Active Complex. FEMS Microbiol. Lett. 239 (2), 309–318. doi:10.1016/j.femsle.2004.09.002

Chauhan, A., Fazlurrahman, J. G. O., Oakeshott, J. G., and Jain, R. K. (2008). Bacterial Metabolism of Polycyclic Aromatic Hydrocarbons: Strategies for Bioremediation. Indian J. Microbiol. 48 (1), 95–113. doi:10.1007/s12088-008-0010-9

Cheung, S., McCarl, V., Holmes, A. J., Coleman, N. V., Wilson, N. L., Barry, K., Brettin, T. S., Bruce, D. C., Copeland, A., et al. (2003b). Epoxyalkane:Coenzyme M Transferase in Mycobacterium Strains. J. Bacteriol. 185 (16), 4772–4778. doi:10.1128/JB.185.16.4772-4778.2003

Chung, A. S., JinJin, Y. O. L., Schmidt, L. S., Li, Y., Fogel, S., Smoler, D., et al. (2013). Continuous Cultures. J. Bacteriol. 195 (10), 3038–3046. doi:10.1128/JB.05109-11

Chuah, M. V., et al. (2015). Regulation of Plasmid-encoded Isoprene Biosynthesis of a Terminal-Alkane Hydroxylating Diiron Monooxygenase. Appl. Microbiol. Biotechnol. 98 (3), 1289–1299. doi:10.1007/s00253-014-5726-z

Christiaens, C., Vernay, S., Vaccari, I., Faucon, J., Schurr, U., van der Donk, W. A., and Karperien, M. (2010). CRISPR-Cas Gene Activators for Transcriptional Reprogramming in Bacteria. ACS Synth. Biol. 1, 1–6. doi:10.1021/sb900309x

Chu, Y., Davis, P., Yao, W., and Whited, G. M., et al. (2015). Response to Vinyl Chloride, Ethene, and Epoxyethane in Nocardioides Sp. from a Vinyl Chloride-Contaminated Site. Microbiol. Rep. 3 (3), 297–307. doi:10.1111/micro.12025.x

Chubuense strain NBB4, a Versatile Hydrocarbon Degrader. Front. Microbiol. 3 (3), 440–475. doi:10.3389/FMICR.2012.00018-13
Shennan, J. L. (2006). Utilisation of C2-C4gaseous Hydrocarbons and Isopropane by Microorganisms. J. Chem. Technol. Biotechnol. 81 (3), 237–256. doi:10.1002/ jctb.1388

Shingler, V. (2010). Experimental Evolution of Novel Regulatory Activities in Response to Hydrocarbons and Related Chemicals. Handbook Hydrocarbon Lipid Microbiol., 1235–1245. doi:10.1007/978-3-540-77587-4_86

Siva-Jimenez, H., Garcia-Fontana, C., Cadirci, B. H., Ramos-González, M. I., Ramos, J. L., and Krell, T. (2012). Study of the TmoS/TmoT Two-Component System: towards the Functional Characterization of the Family of TodS/TodT like Systems. Microb. Biotechnol. 5 (4), 489–500. doi:10.1111/j.1751-7915.2011.00322.X

Small, F. J., and Ensigh, S. A. (1997). Alkene Monoxygenase from Xanthobacter Strain Py2. J. Biol. Chem. 272 (40), 24913–24920. doi:10.1074/jbc.272.40.24913

Smith, T. J., LloydGallagherMurrell, J. S., Gallagher, S. C., Fosdike, W. L. J., Murrell, J. C., and Dalton, H. (1999). Heterologous Expression of Alkene Monoxygenase from Rhodococcus Rhodochrous B-276. Eur. J. Biochem. 260 (2), 446–452. doi:10.1046/j.1432-1379.1999.00179.x

Smits, T. H. M., Balada, S. B., Witholt, B., and Van Beilen, J. B. (2002). Functional Analysis of Alkane Hydroxylases from Gram-Negative and Gram-Positive Bacteria. J. Bacteriol. 184 (6), 1733–1742. doi:10.1128/JB.184.6.1733-1742.2002

Smits, T. H. M., Rothlisberger, M., Witholt, B., and Van Beilen, J. B. (1999). Molecular Screening for Alkane Hydroxylase Genes in Gram-Negative and Gram-Positive Strains. Microb. Biotechnol. 1 (4), 307–317. doi:10.1111/1751-7915.00037.x

Smits, T. H. M., Witholt, B., and Van Beilen, J. B. (2003). Functional Characterization of Genes Involved in Alkane Oxidation by Pseudomonas Aeruginosa. Antonie van Leeuwenhoek. 84 (3), 193–200. doi:10.1023/A: 102600622765

Stajen, I. E., Marcionelli, R., and Witholt, B. (1999). Promoter Is under Carbon Catabolite Repression Control in Pseudomonas Oleovorans but Not in Escherichia coli Alk + Reombinants. J. Bacteriol. 181 (5), 1610–1616. doi:10.1128/jb.181.5.1610-1616.1999

Stainthorpe, A. C., Lees, V., Salmond, G. P. C., Dalton, H., and Murrell, J. C. (1990). The Methane Monoxygenase Gene Cluster of Methylococcus Capsulatus (Bath). Gene 91, 27–34. doi:10.1016/0378-1119(90)90158-n

Sticher, P., Jaspers, M. C., Stemmler, K., Harms, K., Zehnder, A. J., and van der MeerVan der Meer, J. R. (1997). Development and Characterization of a Whole-Cell Bioluminescent Sensor for Bioavailable Middle-Chain Alkanes in Contaminated Groundwater Samples. Appl. Environ. Microb. 63 (10), 4035–4060. doi:10.1128/aem.63.10.4035-4060.1997

Stirling, D. I., Colby, J., and Dalon, H. (1979). A Comparison of the Substrate and Electron-Donor Specificities of the Methane Mono-Oxgenases from Three Strains of Methane-Oxidizing Bacteria. Biochem. J. 177 (1), 361–364. doi:10.1042/bj1770361

Stolyar, S., Franke, M., and Lidstrom, M. E. (2001). Expression of Individual Copies of Methylococcus Capsulatus Bath Particulate Methane Monooxygenase Genes. J. Bacteriol. 183 (5), 1810–1812. doi:10.1128/JB.183.5.1810-1812.2001

Su, L., Jia, W., Hou, C., and Lei, Y. (2011). Microbial Biosensors: A Review. Biosens. Bioelectron. 26 (5), 1788–1799. doi:10.1016/j.bios.2010.09.005

Swaging, J., Weijers, C. A., van Ooyen, A. J., and De Bont, J. A. (1995). Complementation of Xanthobacter Py2 Mutants Defective in Epoxyalkane Degradation, and Expression and Nucleotide Sequence of the Complementing DNA Fragment. Microbiology (Reading) 141 (Pt 2), 477–484. doi:10.1099/13500872-141-2-477

Tan, J.-L., Yong, Z.-X., and Lian, C.-K. (2016). Using a Chemiresistor-Based Alkane Sensor to Distinguish Exhaled Breaths of Lung Cancer Patients from Subjects with No Lung Cancer. J. Thorac. Dis. 8 (10), 2772–2783. doi:10.21037/ jtd.2016.10.30

Tani, A., Ishige, T., Sakai, Y., and Kato, N. (2000). Gene Structures and Regulation of the Alkane Hydroxylase Complex in Acinetobacter Sp. J. Bacteriol. 182 (10), 2772–2781. doi:10.1128/JB.182.10.2772

Taylor, A. E., Arp, D. J., Bottomley, P. J., and Semprini, L. (2010). Extending the Alkene Substrate Range of Vinyl Chloride Utilizing Nocardioides Sp. Strain JS614 with Ethene Oxide. Appl. Microb. Biotechnol. 87, 2293–2302. doi:10.1007/s00253-010-2719-8

Tecon, R., Beggah, S., and van der MeerVan der Meer, J. R. (2010). “Bacterial Bioreporter Assays to Measure Hydrocarbons,” in Handbook of Hydrocarbon
