

# ACCEPTED VERSION

Samuel D. Munday, Natasha K. Maddigan, Rosemary J. Young, Stephen G. Bell  
**Characterisation of two self-sufficient CYP102 family monooxygenases from  
Ktedonobacter racemifer DSM44963 which have new fatty acid alcohol product profiles**  
Biochimica et Biophysica Acta - General Subjects, 2016; 1860(6):1149-1162

© 2016 Elsevier B.V. All rights reserved.

This manuscript version is made available under the CC-BY-NC-ND 4.0 license  
<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Final publication at <http://dx.doi.org/10.1016/j.bbagen.2016.01.023>

## PERMISSIONS

<https://www.elsevier.com/about/our-business/policies/sharing>

Accepted Manuscript

Authors can share their [accepted manuscript](#):

### Immediately

- via their non-commercial personal homepage or blog
- by updating a [preprint](#) in arXiv or RePEc with the [accepted manuscript](#)
- via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
- directly by providing copies to their students or to research collaborators for their personal use
- for private scholarly sharing as part of an invitation-only work group on [commercial sites with which Elsevier has an agreement](#)

### After the embargo period

- via non-commercial hosting platforms such as their institutional repository
- via commercial sites with which Elsevier has an agreement

In all cases [accepted manuscripts](#) should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license – this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our [hosting policy](#)
- not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article

**1 February 2021**

<http://hdl.handle.net/2440/100927>

**Characterisation of two self-sufficient CYP102 family monooxygenases from  
*Ktedonobacter racemifer* DSM44963 which have new fatty acid alcohol product  
profiles**

Samuel D. Munday, Natasha K. Maddigan, Rosemary J. Young and Dr. Stephen G. Bell\*

Department of Chemistry, University of Adelaide, SA 5005, Australia

\* To whom correspondence should be addressed.

Stephen G. Bell ([stephen.bell@adelaide.edu.au](mailto:stephen.bell@adelaide.edu.au)) Tel: +61 8 83134822; Fax: +61 8 83034380

## **Abstract**

### **Background**

Two self-sufficient CYP102 family encoding genes (*Krac\_0936* and *Krac\_9955*) from the bacterium *Ktedonobacter racemifer* DSM44963, which possesses one of the largest bacterial genomes, have been identified.

### **Methods**

Phylogenetic analysis of both the encoded cytochrome P450 enzymes, *Krac0936* and *Krac9955*. Both enzymes were produced and their turnovers with fatty acid substrates assessed *in vitro* and using a whole-cell oxidation system.

### **Results**

*Krac0936* hydroxylated straight chain, saturated fatty acids predominantly at the  $\omega$ -1 and  $\omega$ -2 positions using NADPH as the cofactor. *Krac0936* was less active towards shorter unsaturated fatty acids but longer unsaturated acids were efficiently oxidised. *cis,cis*-9,12-Octadecadienoic and pentadecanoic acids were the most active substrates tested with *Krac0936*. Unusually *Krac9955* showed very low levels of NAD(P)H oxidation activity though coupling of the reducing equivalents to product formation was high. The product distribution of tridecanoic, tetradecanoic and pentadecanoic acid oxidation by *Krac9955* favoured oxidation at the  $\omega$ -4,  $\omega$ -5 and  $\omega$ -6 positions, respectively.

### **Conclusion**

*Krac0936* and *Krac9955* are self-sufficient P450 monooxygenases. *Krac0936* has a preference for pentadecanoic acid over other straight chain fatty acids and showed little or no activity with dodecanoic or octadecanoic acids. *Krac9955* preferably oxidised shorter fatty acids compared to *Krac0936* with tridecanoic having the highest levels of product formation. Unlike *Krac0936* and P450Bm3, *Krac9955* showed lower activities with unsaturated fatty acids.

**General Significance** In this study of two of the CYP enzymes from *K. racemifer* we have shown that this bacterium from the *Chloroflexi* phylum contains genes which encode new proteins with novel activity.

## Highlights

Identification of two self-sufficient cytochrome P450 genes in *K. racemifer* DSM44963

Analysis of the relationship of Krac0936 and Krac9955 with other P450Bm3 like enzymes

Purification and preliminary characterisation of both enzymes

The observation of modified selectivity of fatty acid oxidation by Krac0936 compared to P450Bm3

Krac9955 has an unusual substrate range and product selectivity

## Keywords

Cytochrome P450 enzymes, monooxygenases, fatty acid oxidation, *Ktedonobacter*, *Chloroflexi*

## Abbreviations

BSTFA, N,O-Bis(trimethylsilyl)trifluoroacetamide; CYP, cytochrome P450 enzyme; DEAE, diethylaminoethanol; DMSO, dimethyl sulfoxide; EMM, *E. coli* minimal media; EtOH, ethanol; GC, gas chromatography; HPLC, high performance liquid chromatography; IPTG, Isopropyl  $\beta$ -D-1-thiogalactopyranoside; MS, mass spectrometry; NADH, reduced Nicotinamide adenine dinucleotide; NADPH, reduced Nicotinamide adenine dinucleotide phosphate; PCR, Polymerase chain reaction; TMSCl, Trimethylsilyl chloride; Tris, Tris(hydroxymethyl)aminomethane; 2xYT Yeast Extract Tryptone

## 1. Introduction

The P450 cytochrome (CYP) superfamily of *b*-type heme-dependent monooxygenases, members of which are found in virtually every organism, have a highly conserved tertiary structure and share a characteristic protein fold.[1, 2] P450s catalyse the oxidation of a broad array of organic substrates, allowing them to perform physiological roles as varied as steroid hormone synthesis and the breakdown of xenobiotics.[1-3] As such they are intrinsically linked to the catabolic and anabolic pathways of bacteria. The signature reaction of a P450 is the insertion of a single oxygen atom from molecular dioxygen into a carbon–hydrogen bond to give the corresponding alcohol, the second oxygen atom being reduced to water.[4] These reactions are often carried out with high regio- and stereoselectivity.[5, 6] P450-mediated enzymatic oxidation reactions can therefore provide efficient biocatalytic routes for generating alcohols from alkanes which are not feasible using contemporary organic synthetic methods.[7-10]

CYP102A1 (P450<sub>Bm3</sub>) from *Bacillus megaterium* is a soluble 119.5 kDa enzyme in which a 55 kDa P450 heme domain is fused to a 65 kDa reductase domain.[8] It requires only NADPH and oxygen to function and hydroxylates fatty acids at sub-terminal positions, close to the  $\omega$ -terminus.[11] It acts on range of saturated straight-chain fatty acids, with pentadecanoic acid being the most active with the rates falling away as the chain-length increases or decreases.[8] The product profiles of a selection of fatty acids revealed that oxidation at the  $\omega$ -1,  $\omega$ -2 and  $\omega$ -3 C–H bonds are favoured though there was variation among substrates of differing lengths.[12] Longer chain unsaturated fatty acids such as arachidonic, palmitoleic (*cis*-9-hexadecenoic) and linoleic (*cis,cis*-9,12-octadecadienoic) acid are also rapidly and efficiently oxidised by CYP102A1.[13, 14]

Other self-supporting fatty acid oxidising P450s have been identified and investigated including CYP505 (P450<sub>foxy</sub>), a sub-terminal fatty acid hydroxylase from the fungus *Fusarium oxysporum*. [15] Studies of the closely related enzymes from the same family as CYP102A1, for example, A2 and A3 from *B. subtilis*, [16] A5 from *B. cereus*, [17] A7 from *B. licheniformis* [18]

and CYP102D1 from *Streptomyces avermitilis* MA4680 have also been reported.[19] Minor changes in the fatty acid binding profiles and in the regioselectivity of oxidation are observed across these systems. However the consensus among this family of P450 enzymes is that oxidation occurs at C–H bonds close to the  $\omega$ -terminus. For example CYP102A5 oxidised hexadecanoic acid to yield a 2:1 ratio of the  $\omega$ -2 to  $\omega$ -1 fatty acid alcohols with little or no oxidation at  $\omega$ -3, while CYP102A1 oxidation resulted in a 1:2:1 ratio of the three products.[17] The CYP102D1 enzyme is the only self-sufficient CYP102 enzyme from outside the *Bacillus* family of bacteria to be characterised so far. CYP102D1 is a fatty acid hydroxylase but there are differences in the types of substrates accepted and oxidised including the aromatic substrates indole and anthracene. CYP102D1 was also reported to have optimal activity with octadecanoic acid rather than pentadecanoic acid among straight chain, saturated fatty acids. Octadecanoic acid was oxidised by CYP102D1 at the  $\omega$ -1,  $\omega$ -2 and  $\omega$ -3 positions in a ratio of 36:61:3 compared to CYP102A1 which yielded the same products in the following ratio; 39:47:14.[19] Members of the CYP102 family which are not fused to a reductase domain have also been identified. The best characterised being CYP102B1 from *Streptomyces coelicolor* and CYP102H1 from *Nocardia farcinica*. [20, 21] CYP102B1 was found to oxidise arachidonic acid but not hexadecanoic nor octadecanoic acids while CYP102H1 was reported to oxidise linoleic acid at the  $\omega$ -1 position.

Despite extensive efforts, the precise physiological role of these CYP102 enzymes is not yet known. Branched-chain saturated fatty acids, which play a key role in the regulation of membrane fluidity, also induced CYP102A1 expression, and it has been suggested that these are more likely to be the enzyme's natural substrates, since they account for almost 90% of the fatty acid content of *Bacilli*. [22-24]

The self-sufficient CYP102 systems exhibit the highest turnover frequency ( $> 1000 \text{ min}^{-1}$ ) of any CYP enzymes and this overcomes one of the major hurdles to their use as catalysts.[8, 17, 25] Therefore these systems have been used as platforms for designing monooxygenase biocatalysts via protein engineering. To this end, the effects of various amino

acid residues located in the substrate access channel and active site of enzymes have been investigated by both site-directed and random mutagenesis.[7, 8, 26-29] Furthermore natural variants of CYP102A1 have been found to have slightly differing activities with dodecanoic acid and other non-natural substrates.[30]

*Ktedonobacter racemifer* DSM44963 (or strain SOSP1-21) is an aerobic, non-motile, filamentous, mesophilic, gram-positive heterotroph, capable of growing under microaerophilic conditions.[31] It belongs to the Phylum *Chloroflexi*, or green non-sulfur bacterium, a group of bacteria that is highly isolated in the genealogy and comprises, to date, very few organisms. The *K. racemifer* genome has been decoded and is among the largest bacterial genomes with 11,453 protein-coding genes and therefore provides an opportunity to discover new and interesting protein families.[31] There are two potential self-sufficient CYP102 family member encoding genes in this bacterium (genes *Krac\_0936* and *Krac\_9955*). Given its large genome and the evolutionary niche occupied by *K. racemifer* we set out to ascertain if the role of these CYP102 family members is similar to those of the currently characterised members of this CYP family.

## 2. Experimental

### 2.1 Cloning, expression and purification

The genes encoding full-length Krac0936 and Krac9955 were isolated from the genomic DNA of *K. racemifer* DSM44963 (obtained from DSMZ, Germany) using the polymerase chain reaction (PCR) with the following primers; Krac0936 FP, 5'-**ttaattcatatg**caaatagaagaaaagacactcg-3'. Krac0936 RP, 5'-ttaatta**agctt**ctattagtgccccagacatccaccag-3' Krac9955 FP, 5'-ttaatt**catatg**gtacgatcccagctagacaccatc-3' and Krac9955 RP, 5'-ttaatta**agctt**tattaactggtgccccagatgtcg-3'. The restriction sites NdeI and HindIII are highlighted in bold and the stop and start codons are underlined. The PCR products were cloned into the vector pET26a(+) using the restriction enzyme NdeI and HindIII. The sequences of both genes were confirmed by DNA sequencing (Australian Genome Research Facility, Adelaide) using the T7F and T7T primers of the pET26 vector and the following primers.

CYP102FP2	ctgcagcgagcaaatccagacgac
CYP102FP3	tatgcaaaacaagatacgg
CYP102FP4	gaacagtctgctaaaaaag
CYP102FP5	ggaacagctgaaggaac
CYP102FP6	aattgtcgacagcgccg
CYP102FP7	gaagcacattatacttcgg
CYP102RP	aatctgtatcaggctga

The recombinant plasmids were transformed into *Escherichia coli* strain BL21(DE3) and the transformed cells were cultured in 2xYT medium at 37 °C with 30 µg mL<sup>-1</sup> kanamycin. When the OD<sub>600</sub> of the culture reached 0.6 – 0.8 the temperature was reduced to 20 °C, 0.5 mM IPTG, 3 mL L<sup>-1</sup> of trace elements solution (0.74 g CaCl<sub>2</sub>·H<sub>2</sub>O, 0.18 g ZnSO<sub>4</sub>·7H<sub>2</sub>O, 0.132 g MnSO<sub>4</sub>·4H<sub>2</sub>O, 20.1 g Na<sub>2</sub>EDTA, 16.7 g FeCl<sub>3</sub>·6H<sub>2</sub>O, 0.10 g CuSO<sub>4</sub>·5H<sub>2</sub>O, 0.25 g CoCl<sub>2</sub>·6H<sub>2</sub>O), and 1 mM 5-aminolevulinic acid were added in order to induce the production of the protein.

After further growth for 18 hours at 20 °C, cells were harvested by centrifugation, resuspended in 40 mM potassium phosphate, pH 7.4, 1 mM in dithiothreitol (buffer P) and lysed



by sonication on ice. The crude extracts were then centrifuged at 37000 g for 25 min at 4 °C to remove the cell debris. The P450s were isolated from the supernatant by ammonium sulfate precipitation (30-55% fraction) before desalting using a Sephadex G-25 column pre-equilibrated with buffer P. The red coloured protein was loaded onto a GE-Healthcare DEAE fast-flow Sepharose column (200 x 50 mm) pre-equilibrated with buffer P from which the protein was eluted using a linear gradient of 80-400 mM ammonium sulfate in buffer P. The red P450 fractions were collected and concentrated by ultrafiltration before storage at -20 °C in 50% (v/v) glycerol. Glycerol and salts were removed from proteins immediately prior to experiments using a GE Healthcare 5 mL PD-10 desalting column pre-equilibrated with 50 mM Tris buffer, pH 7.4. The protein concentration of Krac0936 was estimated using the extinction coefficients of CYP102A1.[11]

### *2.2 Substrate binding and NADPH turnover rate determinations*

For substrate binding assays the P450 enzymes were diluted to ~2.0 μM using 50 mM Tris, pH 7.4. The high spin heme content was estimated, to approximately ±5%, by comparison with a set of spectra generated from the sum of the appropriate percentages of the spectra of the substrate-free form (>95% low spin, Soret maximum at 418 nm) and camphor-bound form (>95% high spin, Soret maximum at 392 nm) of wild-type CYP101A1.

NADPH turnovers were run in 1200 μL of 50 mM oxygenated Tris, pH 7.4 at 30 °C, containing 0.2 μM enzyme and 120 μg bovine liver catalase. Assays were held at 30 °C for 1 min prior to addition of 250 μM to 1 mM substrate from a 20, 50 or a 100 mM stock in DMSO or EtOH. NADPH was added from a 20 mg mL<sup>-1</sup> stock to a final concentration of ~320 μM. The NADPH oxidation rate was derived using  $\epsilon_{340} = 6.22 \text{ mM}^{-1} \text{ cm}^{-1}$ .

### *2.3 Whole-cell oxidation turnovers*

Cells were grown and protein produced as described above. The resulting cell pellet was resuspended in double the volume of *E. coli* minimal media containing 2% DMSO; (EMM;

K<sub>2</sub>HPO<sub>4</sub>, 7 g, KH<sub>2</sub>PO<sub>4</sub>, 3 g, Na<sub>3</sub>citrate, 0.5 g, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1 g, MgSO<sub>4</sub>, 0.1 g, 20% glucose, 20 mL and glycerol, 1% v/v in one litre) and added to a baffled flask for increased aeration. The substrate was added to a final concentration of 1 mM before shaking at 150 rpm at 30 °C. After 16 hours the turnover was then centrifuged (15 min, 5000g) and the supernatant isolated. Samples (1 mL) of the turnover were taken for analysis via GC-MS.

#### 2.4 Product analysis

After the *in vitro* NADPH oxidation assays or whole-cell incubations had finished 990 µL of the reaction mixture was mixed with 10 µL of an internal standard solution (*trans*-4-phenyl-3-buten-2-one, 20 mM). For the fatty acids 2 µL of concentrated HCl was added. The mixture was extracted three times with 400 µL of ethyl acetate and the organic extracts were combined and dried over MgSO<sub>4</sub>. Solvent was evaporated under a stream of dinitrogen and the sample dissolved in 200 µL acetonitrile. For fatty acids excess (25 µL) BSTFA/TMSCl (99:1) was added and the mixture left for at least 120 min to produce the trimethylsilyl ester of the carboxylic acid group and trimethylsilyl ether of the alcohol, if formed. The reaction mixtures were used directly for GC-MS analysis. For straight chain, saturated fatty acids and 10-undecenoic acid the oven temperature was held at 120 °C for 3 min and then increased at 7 °C min<sup>-1</sup> up to 220 °C where it was maintained for 7 min (total 24.3 min). For the longer unsaturated fatty acids better separation of the products was obtained using the following method; 120 °C for 3 min then increased to 240 °C at a rate of 5 °C min<sup>-1</sup> where it was held for 5 min (total 32 min). The retention times for the trimethylsilyl (TMS) derivatives are given in the supporting information. Products were identified by matching the GC-MS mass spectra to those expected for ω-n trimethylsilyl derivatised alcohols. Products were calibrated against derivatised product samples of 12-hydroxydodecanoic, 15-hydroxypentadecanoic, 3-hydroxytetradecanoic and 16-hydroxyhexadecanoic acids using the assumption that isomeric mono-oxygenated products would give comparable responses (Table S2). Samples containing a range of concentrations of the chosen product in 50 mM Tris, pH 7.4 were extracted as above. The ratios of the integrated peak

areas of the products and internal standard were determined and plotted against product concentration.

### *2.5 Phylogenetic analysis*

CYP enzyme sequences were used as given in the National Center for Biotechnology Information (NCBI) database or the P450 homepage[32] and the sequence identities and similarities were obtained using a BLAST search (at NCBI). Sequence alignments were performed using ClustalW. The evolutionary history was inferred by using the Maximum Likelihood method based on the Jones-Taylor-Thornton (JTT) matrix-based model.[33] The tree with the highest log likelihood is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbour-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA6.[34]

### 3 Results

#### 3.1 Phylogenetic analysis of *Krac0936* and *Krac9955*

*Ktedonobacter racemifer* DSM44963 contains two genes, *Krac\_0936* and *Krac\_9955*, which encode proteins which are similar to CYP102A1 (50% and 41% sequence identity, respectively). We utilised the Basic Local Alignment Search Tool (BLAST) to search for similar enzymes to those encoded by the *Krac\_0936* and *Krac\_9955* genes from *K. racemifer* DSM44963 (whose amino acids sequences are henceforth called *Krac0936* and *Krac9955*). *Krac0936* (Genbank: EFH80345.1) and *Krac9955* (Genbank: EFH88481.1) share 52% sequence identity with each other across both domains of the protein. *Krac0936* is the only other protein in the NCBI database which shares >50% similarity with *Krac9955* (Table 1). However, *Krac0936* has a greater sequence homology (54-55%) with other proteins such as the CYP102 type enzymes from *Bradyrhizobium diazoefficiens* USDA 110, (formerly *B. japonicum*, CYP102A6),[35] *Rhodopseudomonas palustris* HaA2 (CYP102A13),[36] *Singulisphaera acidiphila*,[37] *Afipia broomeae*[38] and *Variovorax paradoxus*[39] (Table 1a). *Krac9955* shares >40% sequence identity with other CYP102A subfamily members while the identity with CYP102D1 and CYP505 are <40% (Table 1). Generally *Krac0936* has a higher sequence identity with CYP102A family members than *Krac9955* does. It has >40% identity with CYP102D1 but <40% with CYP505.

When comparing just the heme domains of the two systems (defined in this instance as the first 473 amino acids) with other members of the CYP102 family the levels of sequence identity and similarities increased slightly (Table S1) indicating that the similarity of the heme domain is greater than for the whole enzyme. Given that alignment of the heme domain sequences shows that the extremities (residues 1–32 and 447–473) are poorly conserved the sequence similarity would be even higher in areas critical for activity and substrate binding (Fig. S1).[8]

In order to better understand the similarity of the CYP102 enzymes from different bacterial strains we analysed the phylogenetic relationship of the holoproteins and heme domains. Both phylogenetic trees show that Krac0936 and Krac9955 cluster with CYP102A subfamily enzymes, rather than the members of the other CYP102 subfamilies (B through to J). However there are different levels of sequence correlation within the CYP102A subfamily (Fig. 1 and Table 1). The holoproteins, CYP102A1, A2, A3, A5 and A7, which are all from strains of *Bacillus*, [8, 16-18] cluster together along with the enzyme from *S. acidiphila*. Krac0936 and Krac9955 are grouped more closely with other CYP102A holoenzymes including CYP102A6 from *B. diazoefficiens* USDA 110, CYP102A11 from *Erythrobacter* sp. NAP1, [40] and the enzymes from *R. palustris* strains HaA2 and BisB5 (CYP102A13), *A. broomeae*, *V. paradoxus* and *Pseudonocardia asaccharolytica* [41] (Fig. 2a). Within this grouping of CYP102A subfamily enzymes Krac9955 clusters with the enzyme from *P. asaccharolytica*.

When only the heme domains were considered, the CYP102A subfamily splits into three distinct groupings (Fig. 1). Krac9955 and the enzyme from *P. asaccharolytica* pair together but apart from the others. The other two clusters are similar to those observed with the holoproteins except that Krac0936 now aligns more closely with the CYP102A enzymes from *S. acidiphila* and the *Bacillus* species.

The alignment of the amino acid sequences of Krac0936 and Krac9955 with the other CYP102A enzymes shows a great deal of similarity with respect to insertions and deletions (Fig. S1 and Table 2). Krac0936 and Krac9955 have a two amino acid insertion at residues 228 (CYP102A1 numbers given) and two sets of three amino acid insertions at 302 and 385 of the heme domain. More significant changes are observed in the linker region between the two domains (from residue 470 onwards) and there is a five amino acid insertion at residue 788 in the reductase domain. The alignment of the heme domain, which controls substrate binding and oxidation, shows that the F/G and G/H loops, the first half of the G-helix and the domain extremities (residues 1–32 and 447–473) are poorly conserved (Fig. S1). Despite the relatively

low sequence similarity within the CYP102A subfamily members the residues that are likely to be important for core CYP enzyme function are uniform (Table 2). The acid-alcohol pair of the I-helix, which is involved in proton delivery and oxygen activation,[42, 43] the phenylalanine amino acid seven residues towards the N-terminus from the heme iron ligating cysteine,[42] and the glutamate and arginine pair (EXXR), which is proposed to be involved in heme incorporation,[44] are all conserved across the CYP102A subfamily members analysed in this work (Tables 2 and 3).

The residues that make up the active site of the CYP102 family enzymes can be predicted based upon sequence alignment with those found in the structurally characterised CYP102A1.[45] A comparison of the residues in the substrate recognition sites; SRS1 (69–92, in CYP102A1), SRS2 (181–188), SRS3 (200–208), SRS4 (I-helix, residues 253–271), SRS5 (325–335) and SRS6 (434–441) and those which have been targeted in previous mutagenesis studies can also be informative (Table 3).[29, 46] The Arg47 and Tyr51 residues at the entrance of the access channel of CYP102A1 are not conserved across the other CYP102 enzymes, though the arginine residue is present in Krac0936 and in the enzyme from *S. acidiphila*. [47-49] In SRS1, Leu86, Phe87 and Thr88 are conserved across all the enzymes analysed here with the exception of the equivalent amino acid to Thr88 in Krac9955 which is a serine residue (Table 3). However other residues which have been shown to alter the substrate selectivity of CYP102A1 when modified such as Ala74, Val78 and Ala82 are not well conserved across the CYP102 members.[29, 50, 51] The equivalent residues to Ala74, Val78 and Ala82 are proline, asparagine and alanine in Krac0936 and asparagine, asparagine and isoleucine in Krac9955. These residues are also altered in the other enzymes (Table 3). Leu181 in SRS2 is completely conserved, but Leu188 shows considerable variation (glycine and alanine in Krac0936 and Krac9955, respectively) across the CYP102A enzymes.[26, 51] Ala184 is maintained among the CYP102A subfamily members from strains of *Bacillus* but not those from the other bacteria (except *P. asaccharolytica*). Ala184 is modified to serine in Krac0936 and threonine in Krac9955 (Table 3).

The residues of SRS3 show considerable variation and it has been reported that they are not critical for substrate binding in CYP102A1. Structural comparisons have also led to the proposal that SRS3 is not a substrate recognition site in the CYP102 family.[52]

In the I-helix and SRS4 many of the important active site residues are conserved for example Thr260, Ile263 and Ala264. Although variations are found in other enzymes at the equivalent residue to Ile259 it is conserved in Krac0936 and Krac9955 (Table 3). SRS5 which is just after the glutamate and arginine pair (EXXR) contains two prolines which are conserved in all the CYP102A subfamily enzymes analysed here but not CYP102D1. Ala328 and Ala330 which have been targeted in several mutagenesis studies (both site-directed and directed evolution) are conserved in many of the CYP102A enzymes.[27, 53-57] Ala328 is completely conserved, with the exception of CYP102D1, while Ala330 shows more variation in the non-*Bacillus* CYP102 enzymes. This alanine is modified to leucine in Krac0936 and methionine in Krac9955 (Table 3). The key active site residues in SRS6 are conserved across all of the CYP102A subfamily enzymes and the equivalent residues to Leu437, Thr438 and Lys440, are found in all the of the enzymes considered here.[47] Other modifications to the residues of the proposed substrate recognition sites are shown in Table 3 and Table S1a. Overall more of the residues of the SRSs in CYP102A1 are conserved in Krac0936 than in Krac9955.

Kim and co-workers identified an <sup>34</sup>ADELGPIE<sup>41</sup> sequence in the CYP102A1 enzyme (reported as AXEXGPIF in the other CYP102A subfamily members) at the entrance to the access channel which was modified in CYP102D1.[19] This region is somewhat conserved across the CYP102A subfamily enzymes evaluated here but there are some variations in Krac0936 (ARELGPIY) and Krac9955 (AQRHGPIF, Table 2).

The residues in the reductase domain of these self-sufficient enzymes which have been shown to be important in FMN and FAD binding and in controlling the redox properties of these enzymes are conserved (Table S1b and c). These include those corresponding to Ser965, Arg966, Lys972, Tyr974 and Trp1046 of CYP102A1, which are all close to the FAD cofactor and are

important in the specificity of this enzyme for NADPH over NADH.[58] The residues Ser490, Asn491, Try536, Asn537, Trp574 and Thr577 of CYP102A1 which interact with the FMN cofactor are also conserved in Krac0936 and Krac9955 (Table S2).[59]

Overall the conservation of many of these residues in the active site, the substrate access channel and the reductase domain suggest a similar function for the CYP102A enzymes. However, modifications in the sequence of the heme domain are likely to be responsible for the alterations in substrate binding affinity and product distributions which are observed across the family.[8, 15-19, 60] These differences are greater in the non-*Bacillus* members of the CYP102A subfamily including Krac0936 and Krac9955 (*vide infra*). Krac0936 and Krac9955 notably occupy a different phylogenetic space than the CYP102A enzymes from *Bacillus* bacteria, though the heme domain of Krac0936 seems to share more similarity when considered separately compared to the whole protein. Based on the differences in the amino acid sequences outlined above the self-sufficient enzymes from other bacteria, and in particular Krac9955, may interact differently with fatty acids, act on divergent substrates or have distinct functions.

### 3.2 Protein production and fatty acid oxidation using whole-cell oxidation systems

The Krac0936 and Krac9955 enzymes were produced using *E. coli*. When compared to CYP102A1, which was used as a control, the levels of protein production for both were observed to be low (Fig S2). *E. coli* cells producing Krac0936 were faint orange in colour compared to those of Krac9955 which were beige whereas those of CYP102A1 were distinctly orange coloured. Codon analysis of the genes *Krac\_9955* and *Krac\_0936* indicated that there were more clusters of rare codons in *Krac\_0936* than in CYP102A1 (Fig. S3). In addition *Krac\_9955* contained particularly high densities of rare codon clusters (Fig. S3) and this, alongside poor heme incorporation and protein folding, may be potential reasons for the low level of protein production using *E. coli*.

We attempted to purify the proteins using two ion exchange chromatography steps, a method which results in pure protein with CYP102A1.[26] Krac0936 could be purified in low



yamounts using this method. The yield of Krac0936 was increased by incorporating an ammonium sulfate precipitation step before the first ion exchange chromatography step. SDS-page analysis indicated protein of the correct size (Krac0936; 121.4 kDa) was being generated (Fig. S2). After purification, Krac0936 was tested for the characteristic Soret absorbance of a reduced ferrous CO-bound P450. The binding of CO to the ferrous form results in a shift of the Soret peak to 450 nm, indicating the viability of the Krac0936 enzyme (Fig. S4).[61] A shoulder at 420 nm was observed, which may be due to an impurity not removed during purification or could arise from the act of bubbling CO through the protein sample, which, if performed too vigorously, can denature sensitive P450 enzymes. The levels of Krac9955 production using *E. coli* were too low for purification. After ammonium sulfate precipitation and ion exchange chromatography a crude mixture of proteins from the cell lysate containing Krac9955 could be isolated. SDS page analysis revealed a mix of proteins with a small band at the correct mass (Krac9955 122.2 kDa, not shown).

Investigation of the substrate range and product distributions of the Krac0936 and Krac9955 enzymes was initially undertaken by taking advantage of the self-sufficient nature of the CYP102A subfamily members. *E. coli* whole-cell systems producing each enzyme were used to study the oxidation of dodecanoic, tetradecanoic and hexadecanoic acids. Aliquots of the supernatant were extracted and analysed via GC-MS after BSTFA/TMSCl derivatisation. The CYP102A1, Krac0936 and Krac9955 whole-cell oxidation systems all resulted in the generation of hydroxylated fatty acid products from dodecanoic acid indicating that functional holoenzyme is being generated in *E. coli* (Fig. 2, Fig. S5 and S6).

Three main products were observed in the turnovers with CYP102A1 for all three acids (Fig. S5). These products were also observed in the whole-cell turnovers of Krac0936 and Krac9955 with dodecanoic acid and these were identified as the  $\omega$ -1,  $\omega$ -2 and  $\omega$ -3 hydroxylated fatty acids by analysis of the mass spectra of the derivatised products (Scheme 1, Fig. 2 and S7). Analysis of the mass spectrum fragmentation pattern of the three products showed an increase in

the base peak at  $m/z$  117.1, 131.1 and 145.1 AMU, respectively. This is consistent with cleavage next to the OSiMe<sub>3</sub> group, resulting in the fragments CH<sub>3</sub>CHOSiMe<sub>3</sub><sup>+</sup>, CH<sub>3</sub>CH<sub>2</sub>CHOSiMe<sub>3</sub><sup>+</sup> and CH<sub>3</sub>CH<sub>2</sub>CH<sub>2</sub>CHOSiMe<sub>3</sub><sup>+</sup>, respectively. Confirmatory additional peaks were found for the  $\omega$ -1,  $\omega$ -2 and  $\omega$ -3 hydroxylated fatty acids at the expected molecular weight of the doubly derivatised product minus 15, for  $\omega$ -1, minus 29 for  $\omega$ -2 and minus 43 for  $\omega$ -3 (Fig. S7).

The Krac9955 whole-cell oxidation system resulted in diminished levels of metabolite formation compared to Krac0936, presumably due to the low levels of protein production. The oxidation of tetradecanoic acid by Krac9955 was significantly less active when compared to Krac0936 and CYP102A1 (Fig. S5 and S6). The oxidation of hexadecanoic acid by Krac0936 generated low levels of hydroxy-fatty acids ( $\omega$ -1,  $\omega$ -2 and  $\omega$ -3) and analysis of the Krac9955 turnovers assays showed no detectable product.

The ratios of the products formed were calculated by integrating the areas of the peaks. As reported in the literature CYP102A1 hydroxylated the substrates in all three positions with no clear preference and the ratios of hydroxylation at the three sites differed with the chain length of the substrates (Table 4 and Fig. S5).[8] Krac0936 catalysed oxidation of the fatty acids generated the same three hydroxy fatty acids as CYP102A1 but with altered regioselectivity. Each showed minimal levels of oxidation at  $\omega$ -3 (<10%). An almost equal preference for the  $\omega$ -1 and  $\omega$ -2 positions was observed with dodecanoic acid (Fig. 2). However with tetradecanoic acid there was a greater preference for hydroxylation at  $\omega$ -1 (Fig. S6). There was an absence of ion peaks corresponding to the cleavage fragments of CH<sub>2</sub>OSiMe<sub>3</sub><sup>+</sup> (103.1 AMU), which would correspond to hydroxylation at the  $\omega$  position.

Krac9955 oxidation of dodecanoic acid showed a strong preference for hydroxylation at the  $\omega$ -3 position (Fig. 2). Very little oxidation was detected at the  $\omega$ -1 position and in addition we observed low levels of oxidation at the  $\omega$ -4 and  $\omega$ -5 positions ( $\omega$ -4;  $m/z$  159.1 AMU, CH<sub>3</sub>(CH<sub>2</sub>)<sub>3</sub>CHOSiMe<sub>3</sub><sup>+</sup>; and an additional peak at 303.1 AMU; 360 minus 57:  $\omega$ -5;  $m/z$  173.1

AMU,  $\text{CH}_3(\text{CH}_2)_4\text{CHOSiMe}_3^+$ , and an additional peak at 289.1 AMU; 360 minus 71, Table 4, Scheme 2, Fig. S7). The low levels of oxidation in the tetradecanoic acid turnover prevented accurate quantitation of the product distribution but the major products arose from oxidation at the  $\omega$ -4 and  $\omega$ -5 C-H bonds with small peaks due to oxidation at more terminal C-H bonds being observed (Fig. S6 and S7).

All of the whole-cell turnovers including CYP102A1 generated the greatest amount of product with dodecanoic acid. As dodecanoic acid has the highest solubility it would allow the substrate to cross the cell membrane more easily, enabling the enzyme to oxidise it. Therefore the results from the whole-cell oxidation systems cannot be interpreted as a direct measurement of relative enzyme activity and substrate preference. For example with straight chain, saturated fatty acids CYP102A1 is reported to have the highest *in vitro* activity with pentadecanoic acid and has a higher activity with tetradecanoic acid than dodecanoic acid.[8]

### 3.3 *In vitro* analysis of Krac0936 and Krac9955 activity with straight chain saturated fatty acids

The purification of Krac0936 allowed qualitative assessment of the spin state of the heme iron upon the addition of various straight chain fatty acid substrates using UV/Vis analysis. It was found that pentadecanoic acid binding drives the heme iron to the high-spin five-coordinate heme form which absorbs at ~390 nm, 65% HS, with the other fatty acid substrates resulting in lower spin-state shifts (Fig. S8). Tridecanoic acid induced the next highest shift to the high spin form, 50% HS, with those of dodecanoic, tetradecanoic, and hexadecanoic acids being much lower. By way of comparison the spin state shifts observed with CYP102A1 after addition of dodecanoic, tetradecanoic and hexadecanoic acids were reported as 40%, 65% and 95%, respectively.[25]

We also monitored the NAD(P)H oxidation activity of Krac0936 with these saturated fatty acid substrates (Table 6). Addition of NADH did not support monooxygenase activity while NADPH was oxidised by Krac0936 in the presence of pentadecanoic acid. In agreement with spin state shift data pentadecanoic acid resulted in the highest rate of NADPH oxidation (Fig. 3). Tetradecanoic acid brought about the next highest activity, followed by tridecanoic acid and

hexadecanoic acid (Table 6). Heptadecanoic acid showed low levels of NADPH oxidation activity and addition of dodecanoic acid resulted in a rate that was not elevated above that of the leak rate (Table 6 and Fig. 3). It was also found that the enzyme had little or no activity with the equivalent fatty alcohols and methyl esters (C12, C14 and C16, Table 6).

The products of the *in vitro* oxidations were analysed by GC-MS. The turnovers of fatty alcohols and methyl esters did not yield any observable product arising from monooxygenase activity. In contrast to the whole-cell oxidation systems the *in vitro* turnover of dodecanoic acid by Krac0936 generated the lowest amount of product with the coupling of product formation to NADPH reducing equivalents being only 8% (Table 6 and Fig. S9). The formation of  $\omega$ -1,  $\omega$ -2 and  $\omega$ -3 alcohols by Krac0936 were assigned in the same manner as previously described and the product distribution broadly agrees with those obtained from the whole-cell oxidation system (Table 4). The turnovers of the other fatty acid substrates also generated the  $\omega$ -1,  $\omega$ -2 and  $\omega$ -3 products (Scheme 1) with pentadecanoic acid having the highest coupling efficiency, 83%. The amount of product decreased if the chain length was shorter or longer than C15 (Table 6 and Fig. 4a-d). The coupling efficiency of hexadecanoic acid turnovers was high (60%) while that of heptadecanoic acid was lower. No product was observed for longer fatty acids (Table 6). The coupling efficiencies of the turnovers of tetradecanoic and tridecanoic acids were also high (>61%) before a sharp drop off for shorter fatty acids (Table 6).

All of the Krac0936 catalysed fatty acid turnovers generated very little product at the  $\omega$ -3 position with the major products arising from oxidation at  $\omega$ -1 and  $\omega$ -2 (Table 6, Fig. 4a-d, and Fig. S9). Like those of dodecanoic acid the monooxygenase turnovers of pentadecanoic and tridecanoic acids showed minimal selectivity between the  $\omega$ -1 and  $\omega$ -2 C–H bonds. Tetradecanoic acid oxidation by Krac0936 was the most selective for alcohol formation at  $\omega$ -1 with hexadecanoic acid showing a slight preference for oxidation at  $\omega$ -2 (Table 4). Deviation in the product distribution between the *in vitro* and *in vivo* turnovers was observed in the Krac0936

oxidation of hexadecanoic acid and this may arise from the low level of metabolite formation in the whole-cell turnover making accurate quantitation more difficult (Table 4).

The low production levels of Krac9955 in *E. coli* prevented the generation of significant quantities of the enzyme in a pure form impeding detailed *in vitro* characterisation. However we performed activity assays on the crude mixture obtained after ammonium sulfate precipitation and ion-exchange chromatography. The activity was low for all of the fatty acids with no detectable increase in the NADPH oxidation above the leak rate. However significant levels of product were generated for undecanoic (coupling 34%) through to pentadecanoic acids (coupling 73%, Table 6). In agreement with the whole-cell oxidation data, the Krac9955 turnover of hexadecanoic acid resulted in no product formation. Turnover of decanoic acid resulted in very low levels of hydroxylated fatty acids (coupling; 7%). Tridecanoic acid oxidation generated the highest total amount of hydroxylated fatty acid metabolites with the coupling of the reducing equivalents to product formation being 93% (Table 6). Moderate but lower levels of alcohol products were observed with dodecanoic, tetradecanoic and pentadecanoic acids (69-86%, Table 6). No oxidation products arising from monooxygenase activity were observed for fatty alcohols or methyl esters (C12, C14 and C16) and no activity was detected with NADH (data not shown).

With dodecanoic acid the major product was at the  $\omega$ -3 position (72%) in agreement with the data from the whole-cell oxidation system (Fig. 2 and Table 5). The main minor product was at  $\omega$ -2 with low levels of oxidation at the  $\omega$ -1,  $\omega$ -4 and  $\omega$ -5 positions being observed, *vide supra* (Table 5). Increasing or reducing the length of the fatty acid resulted in a systematic change in the product distribution. The major product arising from the undecanoic acid turnovers was 9-hydroxyundecanoic acid ( $\omega$ -2, 78%) with minor products at  $\omega$ -1 and  $\omega$ -3 (Fig. 5a). The major site of oxidation of tridecanoic was at  $\omega$ -4 (56%) with  $\omega$ -3 and  $\omega$ -2 as the main minor products (Fig. 5b and Table 5). Small amounts of the  $\omega$ -1,  $\omega$ -5 and  $\omega$ -6 ( $m/z$  187.1 AMU,  $\text{CH}_3(\text{CH}_2)_5\text{CHOSiMe}_3^+$  and 289.2 AMU; 374 minus 85) alcohols were also observed (Table 5

and Fig. S7). As the chain length of the fatty acid increased in tetradecanoic and pentadecanoic acid the positions of C–H bond oxidation moved to increasing subterminal positions and became less selective. For tetradecanoic acid the major products were  $\omega$ -4 and  $\omega$ -5 (Table 5 and Fig. 5c) and for pentadecanoic acid the fatty acid oxidation predominantly occurred at  $\omega$ -6,  $\omega$ -5 and  $\omega$ -4 (Table 5 and Fig. 5d). The minor products also shifted towards the carboxyl group with oxidation products being observed at  $\omega$ -7 ( $m/z$  201.2 AMU,  $\text{CH}_3(\text{CH}_2)_6\text{CHOSiMe}_3^+$  and 289.2 AMU; 388 minus 99, Fig. S7) for tetradecanoic acid while for pentadecanoic acid low levels of an additional minor product at  $\omega$ -8 were also observed ( $\omega$ -8;  $m/z$  215.2 AMU,  $\text{CH}_3(\text{CH}_2)_7\text{CHOSiMe}_3^+$  and 289.2 AMU; 402 minus 113, Table 5 and Fig. S7). The oxidation of decanoic acid by Krac9955 was inefficiently coupled generating little product compared to undecanoic and dodecanoic acid (Table 6 and Fig. S10). The product distribution was unusual with  $\omega$ -1 and  $\omega$ -3 being the major products and lower levels of  $\omega$ -2 being observed (Table 5). There was no evidence of oxidation at the terminal or the  $\omega$ -4 positions.

### 3.4 Analysis of the activity of Krac0936 and Krac9955 with unsaturated chain fatty acids

All the CYP102A family enzymes from *Bacillus* strains, with the exception of CYP102A7, have been reported to turnover certain unsaturated fatty acids at higher activities than the saturated analogues.[8, 16-18] Like CYP102A7, the fungal enzyme CYP505A1 has lower activity with unsaturated fatty acids whereas the activity of CYP102D1 with saturated and unsaturated fatty acids is comparable.[15, 19, 62] *cis*-9-Octadecanoic (oleic) acid induced an 80% shift to the high-spin form on addition to Krac0936 with *cis,cis*-9,12-Octadecadienoic (linoleic) inducing a lower shift (50%, Fig. S8). *cis*-9-Hexadecenoic (palmitoleic) acid also induced a higher shift than both hexadecanoic and pentadecanoic acids but that caused by *cis*-9-tetradecenoic (myristoleic) acid was low (Fig. S8).

Krac0936 was less active with *cis*-9-tetradecenoic acid than tetradecanoic acid (Table 6). The turnover generated two hydroxylated metabolites (coupling efficiency 19%) with a clear

preference for oxidation at the  $\omega$ -1 position, 93%, over  $\omega$ -2, 7% (Scheme 1 and Fig. 6a). The turnover activity of *cis*-9-hexadecenoic acid by Krac0936 was lower than that of hexadecanoic acid due to reduced NADPH oxidation rates but the coupling efficiency was greater (Table 6). The turnover resulted in a product distribution ratio of 70:18:12 for the  $\omega$ -3: $\omega$ -2: $\omega$ -1 alcohols a significant change compared to the saturated fatty acids (Fig.6b). *cis*-9-Octadecenoic acid was turned over by Krac0936 with a NADPH oxidation rate comparable to that of pentadecanoic acid but the coupling was lower resulting in a reduced product formation rate for the unsaturated fatty acid (Table 6, Scheme 1 and Fig. 3). The products generated were confined to the  $\omega$ -1,  $\omega$ -2 and  $\omega$ -3 positions with only 10% at  $\omega$ -3 and a roughly equal amount of the  $\omega$ -1 and  $\omega$ -2 alcohols (Table 4, Fig. 6c and Fig. S7). *cis,cis*-9,12-Octadecadienoic acid was the most active substrate tested having the highest NADPH activity and a product formation rate comparable to pentadecanoic acid (Table 6). The major metabolite was formed from hydroxylation at  $\omega$ -2 (74%) with oxidation at  $\omega$ -1 being the main minor product (24%) and just 2% selectivity for  $\omega$ -3 (Table 4, Fig. 6d and Fig. S7).

The longer unsaturated, C18, fatty acids were not turned over by Krac9955. *cis*-9-Hexadecenoic and *cis*-9-tetradecenoic acids were oxidised albeit with lower levels of coupling (12-14%) compared to tetradecanoic acid (Table 6). The major product arising from the turnover of *cis*-9-hexadecenoic acid was the epoxide with the allylic alcohol at  $\omega$ -8 ( $m/z = 213.2$  AMU,  $\text{CH}_3(\text{CH}_2)_5\text{CH}=\text{CHCHOSiMe}_3^+$ ) being the only other product detected (Table 5 and Scheme 2). The epoxidation products were identified by GC-MS co-elution with standards of the epoxides generated by reaction of the unsaturated fatty acids with *m*-chloroperbenzoic acid and analysis of the MS fragmentation patterns (Fig. S7). GC-MS analysis revealed that other potential minor products may be present but the MS spectra of these were not consistent with those of likely hydroxylation products and these peaks may arise from impurities (Fig. S9). For the oxidation of *cis*-9-tetradecenoic acid the epoxide was also the major product with allylic oxidation at  $\omega$ -6 ( $m/z$

= 185.25 AMU,  $\text{CH}_3(\text{CH}_2)_3\text{CH}=\text{CHCHOSiMe}_3^+$ ) being the minor product (Table 5, Scheme 2, Fig. 7a and Fig. S7).

Given the preference of Krac9955 for shorter chain fatty acids we also tested 10-undecenoic acid as a substrate. This generated less product than the turnover of undecanoic acid (coupling 12%, Table 6) but the reaction was 100% regioselective for oxidation at the allylic  $\omega$ -2 C-H bonds ( $m/z$  = 129.1 AMU,  $\text{CH}_2=\text{CHCHOSiMe}_3^+$ , Scheme 2, Fig. 7b and Fig. S7).

Overall Krac0936 is selective for the oxidation of fatty acids and not the equivalent alcohols or methyl esters. It has a preference for pentadecanoic acid over the other straight chain fatty acids. Tetradecanoic acid showed moderate activity followed by tridecanoic acid and the hexadecanoic and heptadecanoic acids. While dodecanoic acid was hydroxylated, the activity was very low and octadecanoic was not hydroxylated. Krac0936 showed a preference for hydroxylation of straight chain, saturated fatty acids at the  $\omega$ -1 and  $\omega$ -2 positions with low levels of oxidation at the  $\omega$ -3 position. The preference of the  $\omega$ -1 position is greatest for tetradecanoic acid. Unlike all the previously characterised CYP102 and CYP505 family members, including Krac0936 from the same bacterium, the selectivity of fatty acid oxidation by Krac9955 shifts the product selectivity significantly away from the  $\omega$ -1,  $\omega$ -2 and  $\omega$ -3 sub-terminal positions as the length of the fatty acid side chain increases (Fig. S11). Krac0936 was able to efficiently oxidise longer chain unsaturated fatty acids whereas Krac9955 was not.



## 4 Discussion

Krac0936 and Krac9955 are self-sufficient, fatty acid oxidising, heme monooxygenases from the CYP102 family. They share many similarities to the previously studied CYP102A1, A2, A3, A5, A7 and CYP102D1 but there are differences in their activity with saturated and unsaturated fatty acids. Both, Krac0936 and Krac9955 function efficiently on a narrower range of straight chain saturated fatty acids compared to CYP102A1. The activity of Krac0936 is highest with pentadecanoic acid but does not oxidise C11 or C18 saturated fatty acids while dodecanoic and heptadecanoic acids are poor substrates. The optimal activity of Krac9955 is with tridecanoic acid. It does not oxidise hexadecanoic acid and has minimal activity with decanoic acid.

Krac0936 showed a preference for oxidation at the  $\omega$ -1 and  $\omega$ -2 carbons and very low levels of oxidation at  $\omega$ -3 (<10%). With tetradecanoic acid the selectivity for oxidation at  $\omega$ -1 was 73%. With other self-sufficient enzymes from the *Bacillus* and fungal species the distribution of the products is spread more evenly across these three sub-terminal positions. CYP102A5, from *Bacillus cereus*, showed the highest preference for the  $\omega$ -1 and  $\omega$ -2 positions (Table S3). However, unlike Krac0936, oxidation at  $\omega$ -2 over  $\omega$ -1 is favoured. Krac9955 shows an even more dramatic departure from the product distributions of the previously studied CYP102 family members. The favoured positions of hydroxylation on the fatty acid shifted towards the centre of the molecule as the chain length increased with the preferred site of oxidation being eight carbons away from the carboxylic acid group. These modifications in substrate and product selectivity must be related to changes in the size and shape of the active site as a result of changes in the primary structure between the CYP102 enzymes.

Krac0936 was more active towards longer unsaturated fatty acids in line with activities observed with other CYP102A subfamily enzymes. So while octadecanoic acid is not hydroxylated by Krac0936 the addition of one or two alkene functionalities results in significant increases in the product formation activity. Oxidation at  $\omega$ -1,  $\omega$ -2 and  $\omega$ -3 positions was still

favoured though there was an increased preference for oxidation at  $\omega$ -3 for *cis*-9-hexadecenoic acid. The presence of the second double bond in *cis,cis*-9,12-octadecadienoic acid closer to the  $\omega$  terminus compared to *cis*-9-octadecadienoic acid altered the product distribution to favour oxidation at  $\omega$ -2. No products arising from epoxidation were observed in the turnover mixtures of Krac0936 and *cis*-9-hexadecenoic acid which may reflect the increased preference for this enzyme for the C–H bonds closer to the  $\omega$  terminus compared to CYP102A1.[13] Krac9955 does not oxidise the longer unsaturated fatty acids (C18). 10-Undecenoic acid was regioselectively oxidised by Krac9955 at  $\omega$ -2. However the product formation was reduced compared to undecanoic acid. The level of product generated with unsaturated C14 and C16 fatty acids was low and only products arising from allylic C–H bond oxidation and epoxidation were observed and characterised.

Phylogenetic analysis showed that Krac0936 and Krac9955 cluster with members of the CYP102A subfamily. The CYP102 subfamilies (B-J) show lower sequence homology. It is of note that some of the other subfamilies such as B, C, H and E are not self-sufficient enzymes and would require separate electron transfer proteins. A set of CYP102A enzymes from gram negative bacteria including A13 (both HaA2 and BisB5 from *R. palustris* strains), A6, A11 and the enzymes from *A. broomeae* and *V. paradoxus* cluster together and were clearly separated from the *Bacillus* members of this P450 family. None of these enzymes have been characterised. *K. racemifer* and *Streptomyces* are gram positive strains of bacteria while *Bacillus* species are gram negative. Krac9955 is closely related to a CYP102 enzyme from *P. asaccharolytica* which is also gram positive. The heme domain Krac0936 is grouped together with the enzyme from the gram positive *S. acidiphila* and to a lesser extent with the *Bacillus* CYP102 family members.

It seems that the different CYP102 enzymes may have evolved to undergo different functions within cellular metabolism that are unique to the bacterium in which they are found. This hypothesis is supported by differences in the amino acid sequences of the residues which are

expected to line the active site and access channel of the enzymes. These differences would be expected to result in the observed alterations in the substrate range and product distributions. The critical amino acids involved in the mechanism of oxygen activation and electron transfer seem to be well conserved across the self-sufficient species as are other important residues in the active site such as Phe87 in SRS1 and Leu437 in SRS6 (CYP102A1 amino acid sequence given).

More detailed characterisation of the pure proteins including structural studies and substrate binding and activity assays would enable further insight into differences between the CYP102 enzymes. The different product selectivity and the low NADPH oxidation activity of Krac9955 compared to the other CYP102 family members is particularly intriguing. Despite the relatively efficient whole-cell oxidation of shorter chain fatty acids, the low levels of protein production of Krac0936 and, in particular, Krac9955 would hamper efforts to apply these enzymes as biocatalysts. The GC content of *Ktedonobacter racemifer* and of the *Krac\_0936* and *Krac\_9955* genes is not particularly high (53.8%, 60.0% and 55.3%, respectively compared to the K12 strain of *E. coli* which has GC content of 50.8%). However the presence of groups of rare codons in Krac0936 and in particular Krac9955 compared to CYP102A1 could oppress protein production of these long proteins in *E. coli*. As yet we cannot rule out that heme or flavin incorporation may be suboptimal in these systems.

The observation of the less promiscuous nature of Krac0936 and Krac9955 for saturated fatty acids compared to the other CYP102A1 is interesting. All other CYP102 enzymes studied, with the exception of CYP102A7, show a preference for polyunsaturated fatty acids such as (5Z,8Z,11Z,14Z)-5,8,11,14-eicosatetraenoic acid (arachidonic acid). The CYP102A1 enzyme and others have also been shown to oxidise branched fatty acids which are present in *B. megaterium*. Detailed analysis of the substrate binding and of the product distributions of Krac0936 and Krac9955 with branched and polyunsaturated fatty acids would also be informative. The selectivity, both regio- and stereo-, of the hydroxylation and epoxidation of these substrates would also be of significant interest. The fatty acid spectrum of *K. racemifer* DSM44963 consists

of *iso*-methylpentadecanoic acid (*iso*-C16:0, 20.1%) as the most frequent fatty acid, followed by *anteiso*-methylhexadecanoic acid (*anteiso*-C17:0, 18.5%) and *iso*-methylhexadecanoic (*iso*-C17:0, 15.0%).[31]

## **Acknowledgements**

The authors thank the University of Adelaide for the award of a M. Phil Scholarship (to SDM).

SGB thanks the Australian Research Council for a Future Fellowship (FT FT140100355).

## References

- [1] P.R. Ortiz de Montellano, *Cytochrome P450: Structure, Mechanism and Biochemistry*, 3rd ed., Plenum Publishers, New York, 2005.
- [2] A. Sigel, H. Sigel, R. Sigel, *The Ubiquitous Roles of Cytochrome P450 Proteins*, John Wiley & Sons, Weinheim, 2007.
- [3] F.P. Guengerich, Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity, *Chem. Res. Toxicol.*, 14 (2001) 611-650.
- [4] P.R. Ortiz de Montellano, Hydrocarbon hydroxylation by cytochrome P450 enzymes, *Chem. Rev.*, 110 (2010) 932-948.
- [5] S.G. Bell, L.L. Wong, Iron: Heme Proteins, Mono- & Dioxygenases, in: R.B. King (Ed.) *Encyclopedia of Inorganic Chemistry* 2005.
- [6] V.B. Urlacher, S.G. Bell, L.L. Wong, The Bacterial Cytochrome P450 Monooxygenases: P450cam and P450BM-3, in: R.D. Schmid, V.B. Urlacher (Eds.) *Modern Biooxidation*, Wiley, New York, 2007, pp. 99-122.
- [7] S.G. Bell, N. Hoskins, C.J.C. Whitehouse, L.L. Wong, Design and Engineering of Cytochrome P450 Systems, in: A. Sigel, H. Sigel, R.K.O. Sigel (Eds.) *Met. Ions Life Sci.*, 2007, pp. 437-476.
- [8] C.J. Whitehouse, S.G. Bell, L.L. Wong, P450(BM3) (CYP102A1): connecting the dots, *Chem. Soc. Rev.*, 41 (2012) 1218-1260.
- [9] G.D. Roiban, M.T. Reetz, Expanding the toolbox of organic chemists: directed evolution of P450 monooxygenases as catalysts in regio- and stereoselective oxidative hydroxylation, *Chem. Commun.*, 51 (2015) 2208-2224.
- [10] R. Fasan, Tuning P450 Enzymes as Oxidation Catalysts, *ACS Catalysis*, 2 (2012) 647-666.
- [11] S.S. Boddupalli, R.W. Estabrook, J.A. Peterson, Fatty acid monooxygenation by cytochrome P-450BM-3, *J Biol. Chem.*, 265 (1990) 4233-4239.
- [12] P.P. Ho, A.J. Fulco, Involvement of a single hydroxylase species in the hydroxylation of palmitate at the omega-1, omega-2 and omega-3 positions by a preparation from *Bacillus megaterium*, *Biochim. Biophys. Acta*, 431 (1976) 249-256.
- [13] J.F. Buchanan, A.J. Fulco, Formation of 9,10-epoxypalmitate and 9,10-dihydroxypalmitate from palmitoleic acid by a soluble system from *Bacillus megaterium*, *Biochem. Biophys. Res. Commun.*, 85 (1978) 1254-1260.
- [14] J.H. Capdevila, S. Wei, C. Helvig, J.R. Falck, Y. Belosludtsev, G. Truan, S.E. Graham-Lorence, J.A. Peterson, The highly stereoselective oxidation of polyunsaturated fatty acids by cytochrome P450BM-3, *J. Biol. Chem.*, 271 (1996) 22663-22671.
- [15] N. Nakayama, A. Takemae, H. Shoun, Cytochrome P450foxy, a catalytically self-sufficient fatty acid hydroxylase of the fungus *Fusarium oxysporum*, *J. Biochem.*, 119 (1996) 435-440.
- [16] M.C. Gustafsson, O. Roitel, K.R. Marshall, M.A. Noble, S.K. Chapman, A. Pessegueiro, A.J. Fulco, M.R. Cheesman, C. von Wachenfeldt, A.W. Munro, Expression, purification, and characterization of *Bacillus subtilis* cytochromes P450 CYP102A2 and CYP102A3: flavocytochrome homologues of P450 BM3 from *Bacillus megaterium*, *Biochemistry*, 43 (2004) 5474-5487.
- [17] P.K. Chowdhary, M. Alemseghed, D.C. Haines, Cloning, expression and characterization of a fast self-sufficient P450: CYP102A5 from *Bacillus cereus*, *Arch. Biochem. Biophys.*, 468 (2007) 32-43.
- [18] M. Dietrich, S. Eiben, C. Asta, T.A. Do, J. Pleiss, V.B. Urlacher, Cloning, expression and characterisation of CYP102A7, a self-sufficient P450 monooxygenase from *Bacillus licheniformis*, *Appl. Microbiol. Biotechnol.*, 79 (2008) 931-940.
- [19] K.Y. Choi, E. Jung, D.H. Jung, B.P. Pandey, H. Yun, H.Y. Park, R.J. Kazlauskas, B.G. Kim, Cloning, expression and characterization of CYP102D1, a self-sufficient P450 monooxygenase from *Streptomyces avermitilis*, *FEBS J.*, 279 (2012) 1650-1662.

- [20] D.C. Lamb, L. Lei, B. Zhao, H. Yuan, C.J. Jackson, A.G. Warrilow, T. Skaug, P.J. Dyson, E.S. Dawson, S.L. Kelly, D.L. Hachey, M.R. Waterman, Streptomyces coelicolor A3(2) CYP102 protein, a novel fatty acid hydroxylase encoded as a heme domain without an N-terminal redox partner, *Appl. Environ. Microbiol.*, 76 (2010) 1975-1980.
- [21] Y.-H. Chung, J.-W. Song, K.-Y. Choi, J. Yoon, K.-M. Yang, J.-B. Park, Cloning, expression, and characterization of P450 monooxygenase CYP102H1 from *Nocardia farcinica*, *J. Korean Soc. Appl. Biol. Chem.*, 55 (2012) 259-264.
- [22] M.J. Cryle, R.D. Espinoza, S.J. Smith, N.J. Matovic, J.J. De Voss, Are branched chain fatty acids the natural substrates for P450(BM3)?, *Chem. Commun.*, (2006) 2353-2355.
- [23] N. English, C.N. Palmer, W.L. Alworth, L. Kang, V. Hughes, C.R. Wolf, Fatty acid signals in *Bacillus megaterium* are attenuated by cytochrome P-450-mediated hydroxylation, *Biochem. J.*, 327 ( Pt 2) (1997) 363-368.
- [24] T. Kaneda, Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance, *Microbiol. Rev.*, 55 (1991) 288-302.
- [25] B. Rowlatt, J.A. Yorke, A.J. Strong, C.J. Whitehouse, S.G. Bell, L.L. Wong, Chain length-dependent cooperativity in fatty acid binding and oxidation by cytochrome P450BM3 (CYP102A1), *Protein Cell*, 2 (2011) 656-671.
- [26] C.J. Whitehouse, S.G. Bell, H.G. Tufton, R.J. Kenny, L.C. Ogilvie, L.L. Wong, Evolved CYP102A1 (P450(BM3)) variants oxidise a range of non-natural substrates and offer new selectivity options, *Chem. Commun.*, (2008) 966-968.
- [27] A. Seifert, S. Vomund, K. Grohmann, S. Kriening, V.B. Urlacher, S. Laschat, J. Pleiss, Rational Design of a Minimal and Highly Enriched CYP102A1 Mutant Library with Improved Regio-, Stereo- and Chemoselectivity, *Chembiochem*, 10 (2009) 853-861.
- [28] J.A. McIntosh, P.S. Coelho, C.C. Farwell, Z.J. Wang, J.C. Lewis, T.R. Brown, F.H. Arnold, Enantioselective intramolecular C-H amination catalyzed by engineered cytochrome P450 enzymes in vitro and in vivo, *Angew. Chem. Int. Ed. Engl.*, 52 (2013) 9309-9312.
- [29] S. Kille, F.E. Zilly, J.P. Acevedo, M.T. Reetz, Regio- and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution, *Nat. Chem.*, 3 (2011) 738-743.
- [30] J.Y. Kang, S.Y. Kim, D. Kim, D.H. Kim, S.M. Shin, P.S. H., K.H. Kim, H.C. Jung, J.G. Pan, Y. Joung, Y.T. Chi, H. Chae, T. Ahn, C.H. Yun, Characterization of diverse natural variants of CYP102A1 found within a species of *Bacillus megaterium*, *AMB Express*, 1 (2011) 1.
- [31] Y.J. Chang, M. Land, L. Hauser, O. Chertkov, T.G. Del Rio, M. Nolan, A. Copeland, H. Tice, J.F. Cheng, S. Lucas, C. Han, L. Goodwin, S. Pitluck, N. Ivanova, G. Ovchinnikova, A. Pati, A. Chen, K. Palaniappan, K. Mavromatis, K. Liolios, T. Brettin, A. Fiebig, M. Rohde, B. Abt, M. Goker, J.C. Detter, T. Woyke, J. Bristow, J.A. Eisen, V. Markowitz, P. Hugenholtz, N.C. Kyrpides, H.P. Klenk, A. Lapidus, Non-contiguous finished genome sequence and contextual data of the filamentous soil bacterium *Ktedonobacter racemifer* type strain (SOSP1-21), *Stand. Genomic Sci.*, 5 (2011) 97-111.
- [32] D.R. Nelson, The cytochrome p450 homepage, *Hum. Genomics*, 4 (2009) 59-65.
- [33] D.T. Jones, W.R. Taylor, J.M. Thornton, The rapid generation of mutation data matrices from protein sequences, *Comput. Appl. Biosci.*, 8 (1992) 275-282.
- [34] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: Molecular Evolutionary Genetics Analysis version 6.0, *Mol. Biol. Evol.*, 30 (2013) 2725-2729.
- [35] T. Kaneko, Y. Nakamura, S. Sato, K. Minamisawa, T. Uchiumi, S. Sasamoto, A. Watanabe, K. Idesawa, M. Iriguchi, K. Kawashima, M. Kohara, M. Matsumoto, S. Shimpo, H. Tsuruoka, T. Wada, M. Yamada, S. Tabata, Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110, *DNA Res.*, 9 (2002) 189-197.
- [36] Y. Oda, F.W. Larimer, P.S. Chain, S. Malfatti, M.V. Shin, L.M. Vergez, L. Hauser, M.L. Land, S. Braatsch, J.T. Beatty, D.A. Pelletier, A.L. Schaefer, C.S. Harwood, Multiple genome

sequences reveal adaptations of a phototrophic bacterium to sediment microenvironments, *Proc. Natl. Acad. Sci. U S A*, 105 (2008) 18543-18548.

[37] M. Guo, X. Han, T. Jin, L. Zhou, J. Yang, Z. Li, J. Chen, B. Geng, Y. Zou, D. Wan, D. Li, W. Dai, H. Wang, Y. Chen, P. Ni, C. Fang, R. Yang, Genome sequences of three species in the family Planctomycetaceae, *J. Bacteriol.*, 194 (2012) 3740-3741.

[38] D.J. Brenner, D.G. Hollis, C.W. Moss, C.K. English, G.S. Hall, J. Vincent, J. Radosevic, K.A. Birkness, W.F. Bibb, F.D. Quinn, et al., Proposal of *Afipia* gen. nov., with *Afipia felis* sp. nov. (formerly the cat scratch disease bacillus), *Afipia clevelandensis* sp. nov. (formerly the Cleveland Clinic Foundation strain), *Afipia broomeae* sp. nov., and three unnamed genospecies, *J. Clin. Microbiol.*, 29 (1991) 2450-2460.

[39] J.I. Han, H.K. Choi, S.W. Lee, P.M. Orwin, J. Kim, S.L. Laroe, T.G. Kim, J. O'Neil, J.R. Leadbetter, S.Y. Lee, C.G. Hur, J.C. Spain, G. Ovchinnikova, L. Goodwin, C. Han, Complete genome sequence of the metabolically versatile plant growth-promoting endophyte *Variovorax paradoxus* S110, *J. Bacteriol.*, 193 (2011) 1183-1190.

[40] M. Koblizek, J. Janouskovec, M. Obornik, J.H. Johnson, S. Ferriera, P.G. Falkowski, Genome sequence of the marine photoheterotrophic bacterium *Erythrobacter* sp. strain NAP1, *J. Bacteriol.*, 193 (2011) 5881-5882.

[41] K. Reichert, A. Lipski, S. Pradella, E. Stackebrandt, K. Altendorf, *Pseudonocardia asaccharolytica* sp. nov. and *Pseudonocardia sulfidoxydans* sp. nov., two new dimethyl disulfide-degrading actinomycetes and emended description of the genus *Pseudonocardia*, *Int. J. Syst. Bacteriol.*, 48 Pt 2 (1998) 441-449.

[42] J.P. Clark, C.S. Miles, C.G. Mowat, M.D. Walkinshaw, G.A. Reid, S.N. Daff, S.K. Chapman, The role of Thr268 and Phe393 in cytochrome P450 BM3, *J. Inorg. Biochem.*, 100 (2006) 1075-1090.

[43] T.L. Poulos, Heme enzyme structure and function, *Chem. Rev.*, 114 (2014) 3919-3962.

[44] S. Rupasinghe, M.A. Schuler, N. Kagawa, H. Yuan, L. Lei, B. Zhao, S.L. Kelly, M.R. Waterman, D.C. Lamb, The cytochrome P450 gene family CYP157 does not contain EXXR in the K-helix reducing the absolute conserved P450 residues to a single cysteine, *FEBS Lett.*, 580 (2006) 6338-6342.

[45] H. Li, T.L. Poulos, The structure of the cytochrome p450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid, *Nat. Struct. Biol.*, 4 (1997) 140-146.

[46] R. Fasan, Y.T. Meharena, C.D. Snow, T.L. Poulos, F.H. Arnold, Evolutionary history of a specialized p450 propane monooxygenase, *J. Mol. Biol.*, 383 (2008) 1069-1080.

[47] A.B. Carmichael, L.L. Wong, Protein engineering of *Bacillus megaterium* CYP102. The oxidation of polycyclic aromatic hydrocarbons, *Eur. J. Biochem.*, 268 (2001) 3117-3125.

[48] C.J. Whitehouse, S.G. Bell, W. Yang, J.A. Yorke, C.F. Blanford, A.J. Strong, E.J. Morse, M. Bartlam, Z. Rao, L.L. Wong, A Highly Active Single-Mutation Variant of P450(BM3) (CYP102A1), *Chembiochem*, 10 (2009) 1654-1656.

[49] C.J. Whitehouse, W. Yang, J.A. Yorke, H.G. Tufton, L.C. Ogilvie, S.G. Bell, W. Zhou, M. Bartlam, Z. Rao, L.L. Wong, Structure, electronic properties and catalytic behaviour of an activity-enhancing CYP102A1 (P450(BM3)) variant, *Dalton Trans*, 40 (2011) 10383-10396.

[50] W.C. Huang, A.C. Westlake, J.D. Marechal, M.G. Joyce, P.C. Moody, G.C. Roberts, Filling a hole in cytochrome P450 BM3 improves substrate binding and catalytic efficiency, *J. Mol. Biol.*, 373 (2007) 633-651.

[51] Q.S. Li, J. Ogawa, R.D. Schmid, S. Shimizu, Engineering cytochrome P450 BM-3 for oxidation of polycyclic aromatic hydrocarbons, *Appl. Environ. Microbiol.*, 67 (2001) 5735-5739.

[52] C.A. Hasemann, R.G. Kurumbail, S.S. Boddupalli, J.A. Peterson, J. Deisenhofer, Structure and function of cytochromes P450: a comparative analysis of three crystal structures, *Structure*, 3 (1995) 41-62.

[53] C.K. Chen, R.E. Berry, T. Shokhireva, M.B. Murataliev, H. Zhang, F.A. Walker, Scanning chimeragenesis: the approach used to change the substrate selectivity of fatty acid



monooxygenase CYP102A1 to that of terpene omega-hydroxylase CYP4C7, *J. Biol. Inorg. Chem.*, 15 (2010) 159-174.

[54] M.W. Peters, P. Meinhold, A. Glieder, F.H. Arnold, Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3, *J. Am. Chem. Soc.*, 125 (2003) 13442-13450.

[55] C.J. Whitehouse, S.G. Bell, L.L. Wong, Desaturation of Alkylbenzenes by Cytochrome P450(BM3) (CYP102A1), *Chem. Eur. J.*, 14 (2008) 10905-10908.

[56] C.J. Whitehouse, N.H. Rees, S.G. Bell, L.L. Wong, Dearomatisation of o-xylene by P450BM3 (CYP102A1), *Chem. Eur. J.*, 17 (2011) 6862-6868.

[57] C.J. Whitehouse, W. Yang, J.A. Yorke, B.C. Rowlatt, A.J. Strong, C.F. Blanford, S.G. Bell, M. Bartlam, L.L. Wong, Z. Rao, Structural Basis for the Properties of Two Single-Site Proline Mutants of CYP102A1 (P450(BM3)), *Chembiochem*, 11 (2010) 2549-2556.

[58] M.G. Joyce, I.S. Ekanem, O. Roitel, A.J. Dunford, R. Neeli, H.M. Girvan, G.J. Baker, R.A. Curtis, A.W. Munro, D. Leys, The crystal structure of the FAD/NADPH-binding domain of flavocytochrome P450 BM3, *FEBS J.*, 279 (2012) 1694-1706.

[59] I.F. Sevrioukova, H. Li, H. Zhang, J.A. Peterson, T.L. Poulos, Structure of a cytochrome P450-redox partner electron-transfer complex, *Proc. Natl. Acad. Sci. U S A*, 96 (1999) 1863-1868.

[60] M. Budde, S.C. Maurer, R.D. Schmid, V.B. Urlacher, Cloning, expression and characterisation of CYP102A2, a self-sufficient P450 monooxygenase from *Bacillus subtilis*, *Appl. Microbiol. Biotechnol.*, 66 (2004) 180-186.

[61] T. Omura, R. Sato, The Carbon Monoxide-Binding Pigment of Liver Micorsomes, *J. Biol. Chem.*, 239 (1964) 2370-2378.

[62] T. Kitazume, A. Tanaka, N. Takaya, A. Nakamura, S. Matsuyama, T. Suzuki, H. Shoun, Kinetic analysis of hydroxylation of saturated fatty acids by recombinant P450foxy produced by an *Escherichia coli* expression system, *Eur. J. Biochem.*, 269 (2002) 2075-2082.

## Tables

**Table 1** The sequence identities of the CYP102 enzymes from *K. racemifer* DSM44963; (a) Krac0936 and (b) Krac9955 with selected other CYP102A subfamily members including CYP102A1 (CYP102A1; *B. megaterium*), CYP102A2 (*B. subtilis*), CYP102A3 (*B. subtilis*), CYP102A5 (*B. cereus*), CYP102A6 (*Bradyrhizobium diazoefficiens* USDA 110), CYP102A7 (*B. licheniformis*), CYP102A11 (*Erythrobacter* sp. NAP1), CYP102A13HaA2 and CYP102A13BisB5 (the HaA2 and BisB5 strains of *Rhodopseudomonas palustris*). Also included are CYP102D1 (*Streptomyces avermitilis*), CYP505 (*Fusarium oxysporum*) and CYP102 members from *Singulisphaera acidiphila*, *Pseudonocardia asaccharolytica*, *Smaragdicoccus niigatensis*, *Nakamurella multipartita*, *Afipia broomeae* and *Variovorax paradoxus*. Sequences and nomenclature were obtained from the NCBI database or P450 homepage (<http://drnelson.uthsc.edu/CytochromeP450.html>).[32]

(a)

Protein; Krac0936	Accession number	Identities	Positives	Gaps	Score/ (out of 2238)
<i>S. acidiphila</i>	WP_015249791	600/1082(55%)	764/1082(70%)	10/1082(0%)	1219
CYP102A6	BAC48147	592/1078(55%)	751/1078(69%)	18/1078(1%)	1200
<i>A. broomeae</i>	WP_006019711	583/1075(54%)	754/1075(70%)	9/1075(0%)	1199
CYP102A13BisB5	ABE39056	581/1070(54%)	747/1070(69%)	5/1070(0%)	1162
<i>V. paradoxus</i>	WP_019657533	588/1079(54%)	741/1079(68%)	20/1079(1%)	1152
CYP102A13HaA2	ABD08340	574/1072(54%)	734/1072(68%)	12/1072(1%)	1150
Krac9955	EFH88481	550/1059(52%)	744/1059(70%)	24/1059(2%)	1123
CYP102A11	WP_007165056	546/1072(51%)	715/1072(66%)	14/1072(1%)	1088
CYP102A7	WP_003183869	551/1067(52%)	711/1067(66%)	15/1067(1%)	1063
CYP102A3	O08336	534/1068(50%)	703/1068(65%)	19/1068(1%)	1055
CYP102A2	O08394	530/1075(49%)	705/1075(65%)	24/1075(2%)	1038
CYP102A1	P14779	532/1070(50%)	720/1070(67%)	26/1070(2%)	1040
CYP102A5	ADL27534	509/1047(49%)	684/1047(65%)	21/1047(2%)	986
CYP102D1	WP_010982013	456/1081(42%)	644/1081(59%)	27/1081(2%)	831
CYP505	Q9Y8G7	429/1089(39%)	604/1089(55%)	53/1089(4%)	717

(b)

<b>Protein; Krac9955</b>	<b>Accession number</b>	<b>Identities</b>	<b>Positives</b>	<b>Gaps</b>	<b>Score (out of 2254)</b>
Krac0936	EFH80345	550/1059(52%)	744/1059(70%)	24/1059(2%)	1118
<i>S. acidiphila</i>	WP_015249791	514/1076(48%)	697/1076(64%)	21/1076(1%)	1037
<i>P. asaccharolytica</i>	WP_028931449	512/1072(48%)	690/1072(64%)	18/1072(1%)	1004
<i>S. niigatensis</i>	WP_026306974	489/1071(46%)	681/1071(63%)	26/1071(2%)	971
<i>N. multipartita</i>	WP_015747845	488/1074(45%)	673/1074(62%)	21/1074(1%)	967
<i>A. broomeae</i>	WP_006019711	490/1089(45%)	676/1089(62%)	24/1089(2%)	956
<i>V. paradoxus</i>	WP_019657533	489/1079(45%)	682/1079(63%)	34/1079(3%)	955
CYP102A13HaA2	ABD08340	484/1072(45%)	673/1072(62%)	22/1072(2%)	943
CYP102A6	BAC48147	484/1090(44%)	666/1090(61%)	31/1090(2%)	938
CYP102A2	O08394	479/1076(45%)	665/1076(61%)	38/1076(3%)	935
CYP102A13BisB5	ABE39056	482/1086(44%)	669/1086(61%)	24/1086(2%)	934
CYP102A7	WP_003183869	477/1074(44%)	655/1074(60%)	37/1074(3%)	904
CYP102A3	O08336	460/1069(43%)	652/1069(60%)	33/1069(3%)	891
CYP102A5	ADL27534	463/1043(44%)	649/1043(62%)	27/1043(2%)	884
CYP102A1	P14779	444/1075(41%)	634/1075(58%)	41/1075(3%)	841
CYP102D1	WP_010982013	388/1042(37%)	583/1042(55%)	31/1042(2%)	722
CYP505A1	Q9Y8G7	384/1078(36%)	587/1078(54%)	62/1078(5%)	642

**Table 2** Comparative data and sequences for CYP102 family members. The sequences of conserved regions of the substrate entrance channel and the heme binding motif as well as the predicted pI and length of the amino acid chain are also provided. In the substrate entrance channel the Arg47 and Tyr51 residues of CYP102A1 have been highlighted in red. In the heme binding motif conserved residues are highlighted in bold. Sequences and nomenclature were obtained from the NCBI database or P450 homepage (<http://drnelson.uthsc.edu/CytochromeP450.html>).[32]

<b>Protein</b>	<b>Accession number</b>	<b>substrate entrance channel</b>	<b>Heme binding motif</b>	<b>pI</b>	<b>AA</b>
Krac0936	EFH80345	AREL-GPIYQLEI <b>PG</b> <b>R</b> RT-VIVS	<b>YK</b> <b>P</b> <b>F</b> <b>G</b> <b>N</b> <b>G</b> <b>E</b> <b>R</b> <b>A</b> <b>C</b> <b>I</b> <b>G</b>	5.2	1073
Krac9955	EFH88481	AQRH-GPIFQLPLPDGTRRIVLS	<b>YK</b> <b>P</b> <b>F</b> <b>G</b> <b>T</b> <b>G</b> <b>Q</b> <b>R</b> <b>S</b> <b>C</b> <b>I</b> <b>G</b>	5.3	1079
<i>S. acidiphila</i>	AGA30711	AKEH-GPIYRLDLP <b>R</b> KL-VILS	<b>YK</b> <b>P</b> <b>F</b> <b>G</b> <b>N</b> <b>G</b> <b>Q</b> <b>R</b> <b>A</b> <b>C</b> <b>I</b> <b>G</b>	6.0	1080
<i>P. asaccharolytica</i>	WP_028931449	AQEY-GPLYKLTPAGVR-LMVS	<b>YK</b> <b>P</b> <b>F</b> <b>G</b> <b>T</b> <b>G</b> <b>Q</b> <b>R</b> <b>A</b> <b>C</b> <b>I</b> <b>G</b>	4.9	1076
CYP102A1	P14779	ADEL-GEI <b>F</b> KFEAP <b>G</b> <b>R</b> VT- <b>R</b> <b>Y</b> LS	<b>F</b> <b>K</b> <b>P</b> <b>F</b> <b>G</b> <b>N</b> <b>G</b> <b>Q</b> <b>R</b> <b>A</b> <b>C</b> <b>I</b> <b>G</b>	5.0	1048
CYP102A2	O08394	AKEQ-GPIFQLHTPAGTI-IVVS	<b>YK</b> <b>P</b> <b>F</b> <b>G</b> <b>N</b> <b>G</b> <b>Q</b> <b>R</b> <b>A</b> <b>C</b> <b>I</b> <b>G</b>	5.4	1060
CYP102A3	O08336	ADEL-GPIFRDFDPGVSS-VFVS	<b>YK</b> <b>P</b> <b>F</b> <b>G</b> <b>N</b> <b>G</b> <b>Q</b> <b>R</b> <b>A</b> <b>C</b> <b>I</b> <b>G</b>	5.7	1053
CYP102A5	ADL27534	AEEY-GPIFQIQTLSDTI-IVVS	<b>YK</b> <b>P</b> <b>F</b> <b>G</b> <b>N</b> <b>G</b> <b>Q</b> <b>R</b> <b>A</b> <b>C</b> <b>I</b> <b>G</b>	5.7	1033
CYP102A7	WP_003183869	ADEM-GPIFQFKFADAIG-VFVS	<b>YK</b> <b>P</b> <b>F</b> <b>G</b> <b>N</b> <b>G</b> <b>Q</b> <b>R</b> <b>A</b> <b>C</b> <b>I</b> <b>G</b>	5.7	1073
CYP102A13HaA2	ABD08340	AEEL-GPIFWLDMMGAPI-VIVS	<b>W</b> <b>K</b> <b>P</b> <b>F</b> <b>G</b> <b>N</b> <b>G</b> <b>Q</b> <b>R</b> <b>A</b> <b>C</b> <b>I</b> <b>G</b>	6.2	1071
CYP102D1	WP_010982013	SKQFPEGLYGMEIAGIEQ-VFVW	<b>YK</b> <b>P</b> <b>F</b> <b>G</b> <b>N</b> <b>G</b> <b>V</b> <b>R</b> <b>A</b> <b>C</b> <b>I</b> <b>G</b>	4.9	1072
<i>A. broomeae</i>	WP_006019711	TKEL-GPIFWLDMMGQPL-VIVS	<b>W</b> <b>K</b> <b>P</b> <b>F</b> <b>G</b> <b>N</b> <b>G</b> <b>Q</b> <b>R</b> <b>A</b> <b>C</b> <b>I</b> <b>G</b>	6.4	1080
<i>V. paradoxus</i>	WP_019657533	AQDL-GGIYWLDMPGMPV-IVVS	<b>F</b> <b>K</b> <b>P</b> <b>F</b> <b>G</b> <b>N</b> <b>G</b> <b>Q</b> <b>R</b> <b>A</b> <b>C</b> <b>I</b> <b>G</b>	6.0	1072

**Table 3** Potential active site residues of the Krac0936 and Krac9955 enzymes compared to those of other CYP102 family enzymes. The SRSs are based on those of CYP102A1 and those proposed for other CYP enzymes; SRS1 (69–92), SRS2 (181–188), SRS3 (200–208), SRS4 (I-helix, residues 253–271), SRS5 (325–335) and SRS6 (434–441). The amino acids in Krac0936 and Krac9955 which have been considered are given in brackets to the left and right of the SRS site, respectively. The equivalent residues of CYP102A1 are given at the bottom of each column. The residues of CYP102A1 that have previously been investigated are highlighted in red and underlined. Conserved residues (to CYP102A1) are highlighted in bold.

<b>CYP enzyme</b>	(72-101) <b>SRS1</b> (64-93)	(182-201) <b>SRS2</b> (174-193)	(202-217) <b>SRS3</b> (194-209)
Krac0936	<b>FDKRVW-RPLQNV</b> - <b>AFAGDGLFTAYTQE</b> <b>ANW</b>	<b>MGNALGESMARGRRLALQEK</b>	<b>LMVKNHRQFQADVDFM</b>
<i>S. acidiphila</i>	<b>FDKLIS-TALGKVR-SFSGDGLFTAWTFEP</b> <b>NW</b>	<b>MVRSLGESLLQANRLPLQEA</b>	<b>LLFRTHRHEQDIAYM</b>
<i>A. broomeae</i>	<b>FDKAVR-GALRRVR-AIGDGLFTADTTEP</b> <b>NW</b>	<b>LVRSL-ETIMMTRGLPLEGL</b>	<b>FMRARRDMAADVAFM</b>
CYP102A13HaA2	<b>FDKAVR-GALRRVR-TVGGDGLFTADTSEP</b> <b>NW</b>	<b>LVRSL-ETIMMTRGLPLENL</b>	<b>WMKKRRDTLAEDVAFM</b>
<i>V. paradoxus</i>	<b>FDKSTR-GALRRLR-AAS-HGLFTSDTHEET</b> <b>W</b>	<b>MVRTL-ETVQNRRLGLPLEEL</b>	<b>MLKKELAQQRKDIRYM</b>
Krac9955	<b>FDKRVS-VNVRNLR-SAIGSGLFSADTSDP</b> <b>NW</b>	<b>MVHILSTTAARATRPPIEEK</b>	<b>LRFRQRRLQADLDLL</b>
<i>P. asaccharolytica</i>	<b>FDKAVG-GGLSSLRGSVVSTGLFTAETDDP</b> <b>LW</b>	<b>MVRSLLTAQARVRLPIQNR</b>	<b>LHIREQRQTEEDQIFM</b>
CYP102A2	<b>FDKSIE-GALEKVR-AFSGDGLFTSWTHEP</b> <b>NW</b>	<b>MVRALDEAMHQQRLDVQDK</b>	<b>LMIRTKRQFHHDIQAM</b>
CYP102A3	<b>FDKNLG-KGLQKVR-EFGDGLFTSWTHEP</b> <b>NW</b>	<b>MLRALKEAMNQSKRLGLQDK</b>	<b>MMVTKLQFQKDIEVM</b>
CYP102A1	<b>FDKNLS-QALKFVR-DEFAGDGLFTSWTHEK</b> <b>NW</b>	<b>MVRALDEAMNKLQRANPDDP</b>	<b>AYDENKRQFQEDIKVM</b>
CYP102D1	<b>FFKQIDKTPLAHVR-DYAGAGLFTAHQH</b> <b>EEEW</b>	<b>LLQALVESLRRSQELPVMTK</b>	<b>MRKADDKKYRENIRLM</b>
	(67-96)	(177-196)	(197-212)
<b>CYP enzyme</b>	(256-280) <b>SRS4</b> (249-273)	(331-352) <b>SRS5</b> (323-345)	(439-456) <b>SRS6</b> (445-452)
Krac0936	<b>LNDVNIRYQIITFLIAGHETTSGLL</b>	<b>ESRLWPTAPLFSVYPYEETVV</b>	<b>YELKIKQTLTLKPEGFTM</b>
<i>S. acidiphila</i>	<b>LDDVNIRYQILTFLIAGHETTSGLL</b>	<b>ESRLWPTAPAFALYPREETTL</b>	<b>YRLRVKETLTLKPDGFTM</b>
<i>A. broomeae</i>	<b>LDDVNIRYQINTFLIAGHETTSGLL</b>	<b>ESLRMWPPAPAYGVTPDKDETI</b>	<b>YQMHLKETLTIKPDGFKI</b>
CYP102A13HaA2	<b>LDDVNIRYQINTFLIAGHETTSGLL</b>	<b>ETLRLWPPAPAYGVAPIQDETI</b>	<b>YRMVLKETLTIKPEGFKI</b>
<i>V. paradoxus</i>	<b>LTDKMIRDECIEFLIAGHETTSGLL</b>	<b>ESRLYPTAPAI SMRAKEETKI</b>	<b>YKLLIKEALTIKPEGFRI</b>
Krac9955	<b>LDDVNIRNQIITFLIAGHETTSGLL</b>	<b>ETLRLCPTAPMFNRYPYEERPL</b>	<b>YQLKIREALTIKPADFYM</b>
<i>P. asaccharolytica</i>	<b>LPDSNIQAQCITFLIAGHETTSGLL</b>	<b>EALRLWPTAPAFTRAPFEDIVL</b>	<b>YRLKTKTTLTIKPDFFRI</b>
CYP102A2	<b>LDDENIRFQIITFLIAGHETTSGLL</b>	<b>ESRLWPTAPAFSLYAKEDTVI</b>	<b>YELDIKQTLTLKPNDFHI</b>
CYP102A3	<b>LDDENIRYQIITFLIAGHETTSGLL</b>	<b>ETLRLYPTAPAFSLYAKEDTVL</b>	<b>YELKIKEALTIKPDFFKI</b>
CYP102A1	<b>LDDENIRYQIITFLIAGHETTSGLL</b>	<b>EALRLWPTAPAFSLYAKEDTVL</b>	<b>YELDIKETLTLKPEGFVV</b>
CYP102D1	<b>LDDDNVRDQVVTFIAGHETTSGLL</b>	<b>ETLRLWAPIPMIGKSPLEDTV I</b>	<b>YKMDVKEALTRKPGGFEL</b>
	(249-273)	(321-342)	(429-446)

**Table 4** Product distribution data from the *in vitro* (in bold) and whole-cell turnover reactions (in italics) of Krac0936 and different fatty acids. The reported values for CYP102A1 are included (in brackets) for comparison.[8] The turnovers were analysed by GC-MS. The data are reported as the mean of at least three experiments.

<b>Substrate –</b>	<b>ω-1%</b>		<b>ω-2%</b>		<b>ω-3%</b>	
Krac0936-dodecanoic acid	<b>40</b>	<i>47</i>	<b>51</b>	<i>48</i>	<b>9</b>	<i>4</i>
Krac0936-tridecanoic acid	<b>50</b>		<b>46</b>		<b>4</b>	
Krac0936-tetradecanoic acid	<b>68</b>	<i>81</i>	<b>26</b>	<i>16</i>	<b>6</b>	<i>3</i>
Krac0935-pentadecanoic acid	<b>46</b>		<b>47</b>		<b>6</b>	
Krac0936-hexadecanoic acid	<b>37</b>	<i>67</i>	<b>54</b>	<i>27</i>	<b>9</b>	<i>6</i>
Krac0936-heptadecanoic acid	<b>44</b>		<b>43</b>		<b>13</b>	
Krac0936- <i>cis</i> -9-tetradecanoic acid	<b>93</b>		<b>7</b>		<b>-</b>	
Krac0936- <i>cis</i> -9-hexadecanoic acid	<b>12</b>		<b>18</b>		<b>70</b>	
Krac0936- <i>cis</i> -9-octadecanoic acid	<b>48</b>		<b>42</b>		<b>10</b>	
Krac0936- <i>cis,cis</i> -9,12-octadecanoic acid	<b>24</b>		<b>74</b>		<b>2</b>	
CYP102A1-dodecanoic acid	(36)	<i>35.5</i>	(30)	<i>29</i>	(34)	<i>35.3</i>
CYP102A1-tridecanoic acid	(17)		(65)		(18)	
CYP102A1-tetradecanoic acid	(44)	<i>51</i>	(28)	<i>27</i>	(28)	<i>22</i>
CYP102A1-pentadecanoic acid	(32)		(49)		(19)	
CYP102A1-hexadecanoic acid	(31)	<i>38.5</i>	(48)	<i>49</i>	(21)	<i>12.5</i>
CYP102A1-heptadecanoic acid	(49)		(35)		(16)	

**Table 5** Product distribution data from the *in vitro* and whole-cell turnover reactions (in italics) of Krac9955 with different fatty acids. The data are reported as the mean of at least three experiments.

<b>Substrate –</b>	<b><math>\omega</math>-1%</b>	<b><math>\omega</math>-2%</b>	<b><math>\omega</math>-3%</b>	<b><math>\omega</math>-4%</b>	<b><math>\omega</math>-5%</b>	<b><math>\omega</math>-6%</b>	<b><math>\omega</math>-7%</b>	<b><math>\omega</math>-8%</b>	<b>epox%</b>
decanoic	38	10	52	-	-	-	-	-	-
undecanoic	6	78	14	2	-	-	-	-	-
dodecanoic	6 ( <i>7</i> )	16 ( <i>18</i> )	72 ( <i>69</i> )	5 ( <i>4.5</i> )	1 ( <i>0.5</i> )	-	-	-	-
tridecanoic	2	17	20	56	3	2	-	-	-
tetradecanoic	1	7	14	34	41	2	2	-	-
pentadecanoic	1	2	10	26	20	37	2	2	-
<i>cis</i> -9-tetradecanoic	-	-	-	-	-	26	-	-	74
<i>cis</i> -9-hexadecanoic	-	-	-	-	-	-	-	16	84
10-undecanoic	-	100	-	-	-	-	-	-	-

**Table 6** The relative NADPH oxidation activities and product formation rates for the Krac0936 turnovers of fatty acids. The coupling efficiency of the turnover of the fatty acids with Krac0936 and Krac9955 is also reported. The turnover activities were measured using a Krac0936 concentration of 0.2  $\mu\text{M}$  in 50 mM Tris, pH 7.4. Coupling is the percentage efficiency of NADPH utilisation for the formation of products. The data are reported as a mean  $\pm$  S.D. ( $n \geq 3$ ) and the rates are given in  $\text{nmol.nmol-CYP}^{-1}.\text{min}^{-1}$ .

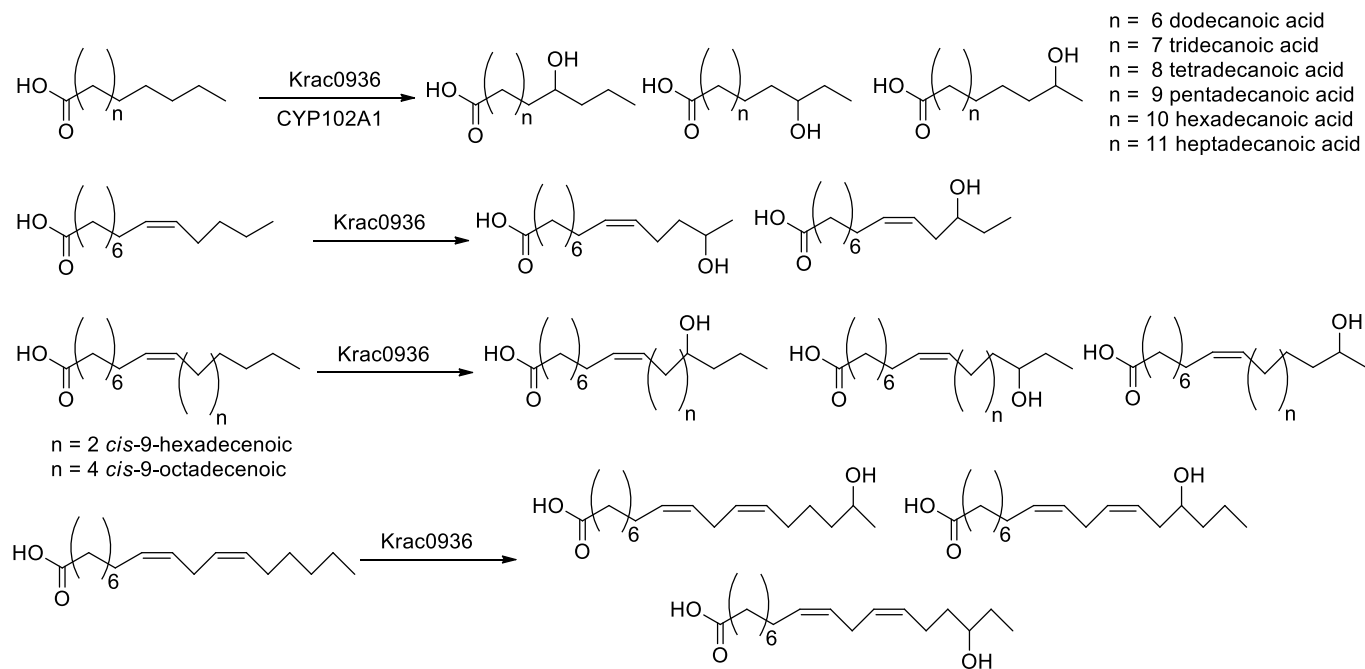
Substrate	Krac0936 <sup>a</sup>			Krac9955 <sup>a</sup>
	NADPH	PFR	Coupling	Coupling
decanoic	-	-	-	7 $\pm$ 1
undecanoic	-	-	-	34 $\pm$ 1
dodecanoic	15 $\pm$ 1	1 $\pm$ 0.2	8 $\pm$ 0.3	89 $\pm$ 6
tridecanoic	246 $\pm$ 11	139 $\pm$ 7	57 $\pm$ 1	93 $\pm$ 4
tetradecanoic	469 $\pm$ 7	392 $\pm$ 12	84 $\pm$ 2	69 $\pm$ 4
pentadecanoic	951 $\pm$ 24	861 $\pm$ 13	91 $\pm$ 2	73 $\pm$ 4
hexadecanoic	392 $\pm$ 5	236 $\pm$ 20	60 $\pm$ 4	-
heptadecanoic	38 $\pm$ 0.6	13 $\pm$ 0.4	35 $\pm$ 1	-
octadecanoic	-	-	-	-
<i>cis</i> -9-tetradecanoic	15 $\pm$ 1	2.7 $\pm$ 0.1	19 $\pm$ 1	14 $\pm$ 1
<i>cis</i> -9-hexadecanoic	210 $\pm$ 3	188 $\pm$ 9	89 $\pm$ 3	8 $\pm$ 1
<i>cis</i> -9-octadecanoic	924 $\pm$ 11	296 $\pm$ 11	32 $\pm$ 1	-
<i>cis</i> -9,12-octadecadienoic	1350 $\pm$ 19	880 $\pm$ 16	65 $\pm$ 1	-
10-undecanoic	-	-	-	12 $\pm$ 0.1

<sup>a</sup> The alcohols dodecanol, tetradecanol, hexadecanol and the methyl esters methyl dodecanoate, methyl tetradecanoate were tested with Krac0936 and resulted in little or no NADPH oxidation above the leak rate. GC-MS analysis of the turnovers of these esters and alcohols with both Krac9955 and Krac0936 showed no detectable levels of product. No increase in the NAD(P)H oxidation activity was observed when replacing NADPH with NADH for either Krac0936 and Krac9955 and little or no product was formed in these turnovers.

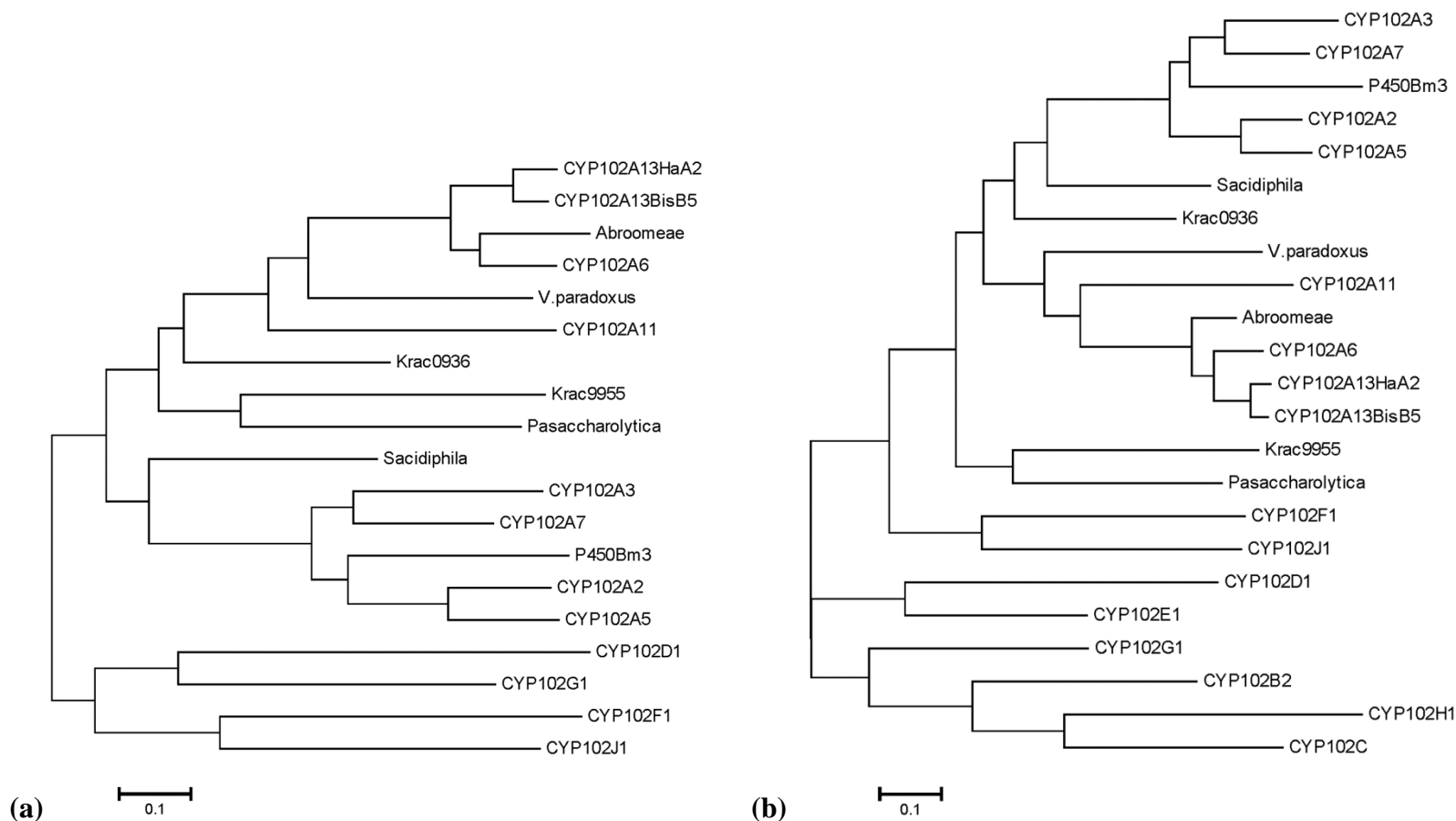


## Figures and Schemes

**Scheme 1** Product profiles of the Krac0936 catalysed turnovers of fatty acids. The same products are found in the turnovers of CYP102A1 with saturated fatty acids though the product distributions are different (Table 4).

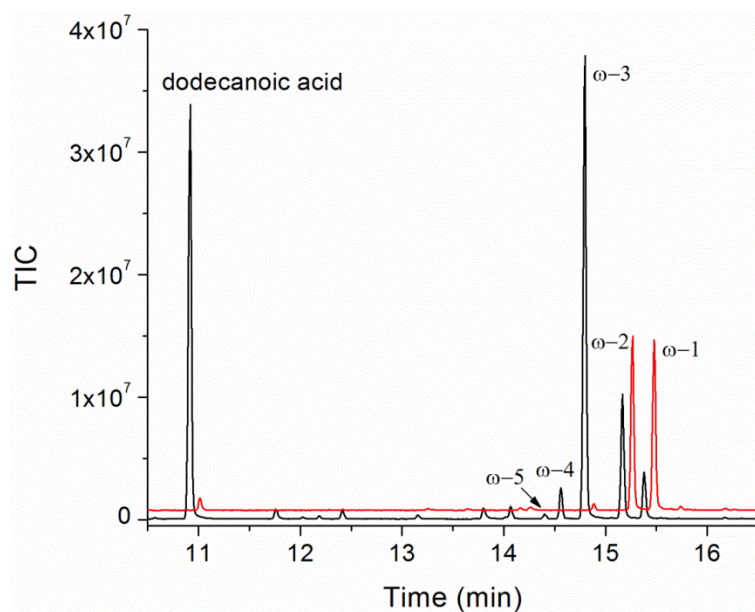






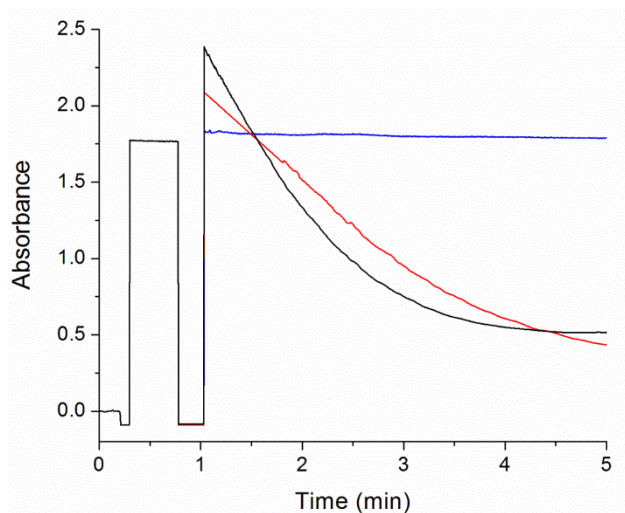
**Figure 1** Phylogenetic tree of the (a) holoprotein and (b) heme domain only (residues 1-473) of selected CYP102 family enzymes including Krac9955 and Krac0936 (*K. racemifer* DSM44963 ) CYP102A1 (CYP102A1; *B. megaterium*), CYP102A2 (*B. subtilis*), CYP102A3 (*B. subtilis*), CYP102A5 (*B. cereus*), CYP102A6 (*Bradyrhizobium diazoefficiens* USDA 110), CYP102A7 (*B. licheniformis*), CYP102A11 (*Erythrobacter* sp. NAP1), CYP102A13HaA2 and CYP102A13BisB5 (the HaA2 and BisB5 strains of *Rhodopseudomonas palustris*), CYP102B2 and CYP102D1 (both *Streptomyces avermitilis*), CYP101C2 (*Rhodococcus erythropolis* PR4), CYP102E1 (*Ralstonia metallidurans*), CYP102F1 (*Actinosynnema pretiosum* subsp. *Auranticum*) CYP102G1 (*Streptomyces scabies*) CYP102H1 (*Nocardia farcinica*), CYP102J1 (*Burkholderia* sp.) and the CYP102 members from *Singulisphaera acidiphila*, *Pseudonocardia asaccharolytica*, *Afipia broomeae* and *Variovorax paradoxus*. Sequences and nomenclature were obtained from the NCBI database or P450 homepage (<http://drnelson.uthsc.edu/CytochromeP450.html>).[32]

**Figure 2** GC-MS analyses of the BSTFA/TMSCl derivatised whole-cell oxidation turnovers of Krac0936 and Krac9955 (black) with dodecanoic acid. The substrate and products are labelled as are impurities (with an asterisk). The Krac0936 turnover has been offset slightly along the x- and y-axes for clarity.

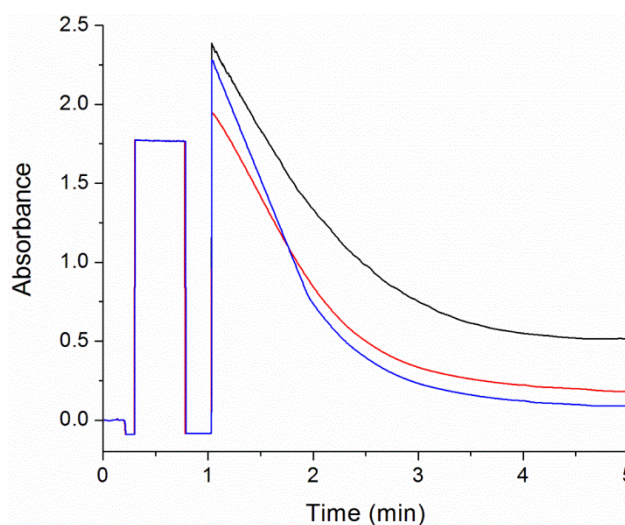


**Figure 3** NADPH oxidation assays of (a) Krac0936 with dodecanoic (blue), tetradecanoic (red) and pentadecanoic (black) acids, and (b) pentadecanoic (black), *cis*-9-octadecenoic (red) and *cis,cis*-9,12-octadecadienoic (blue) acids.

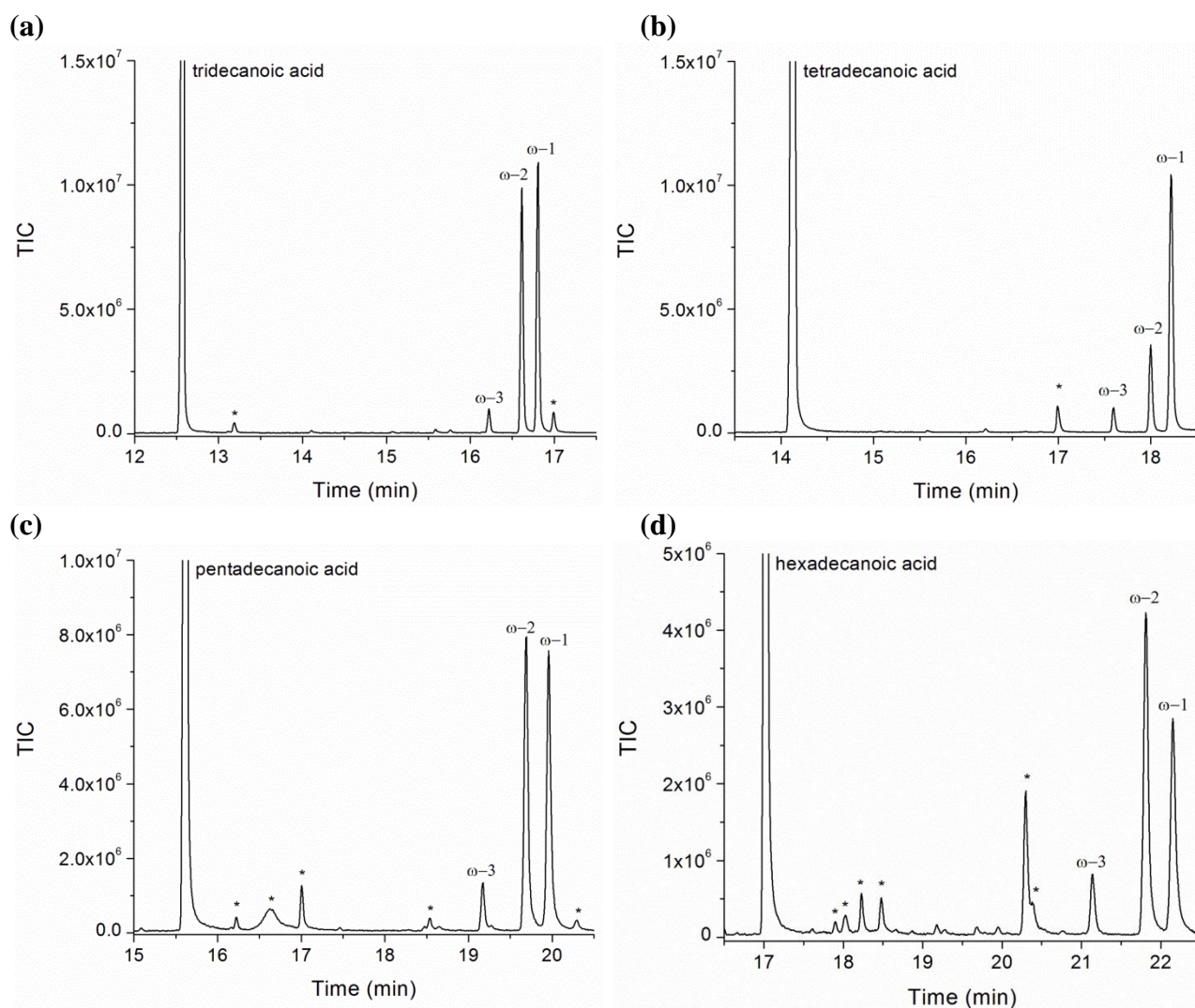
(a)



(b)

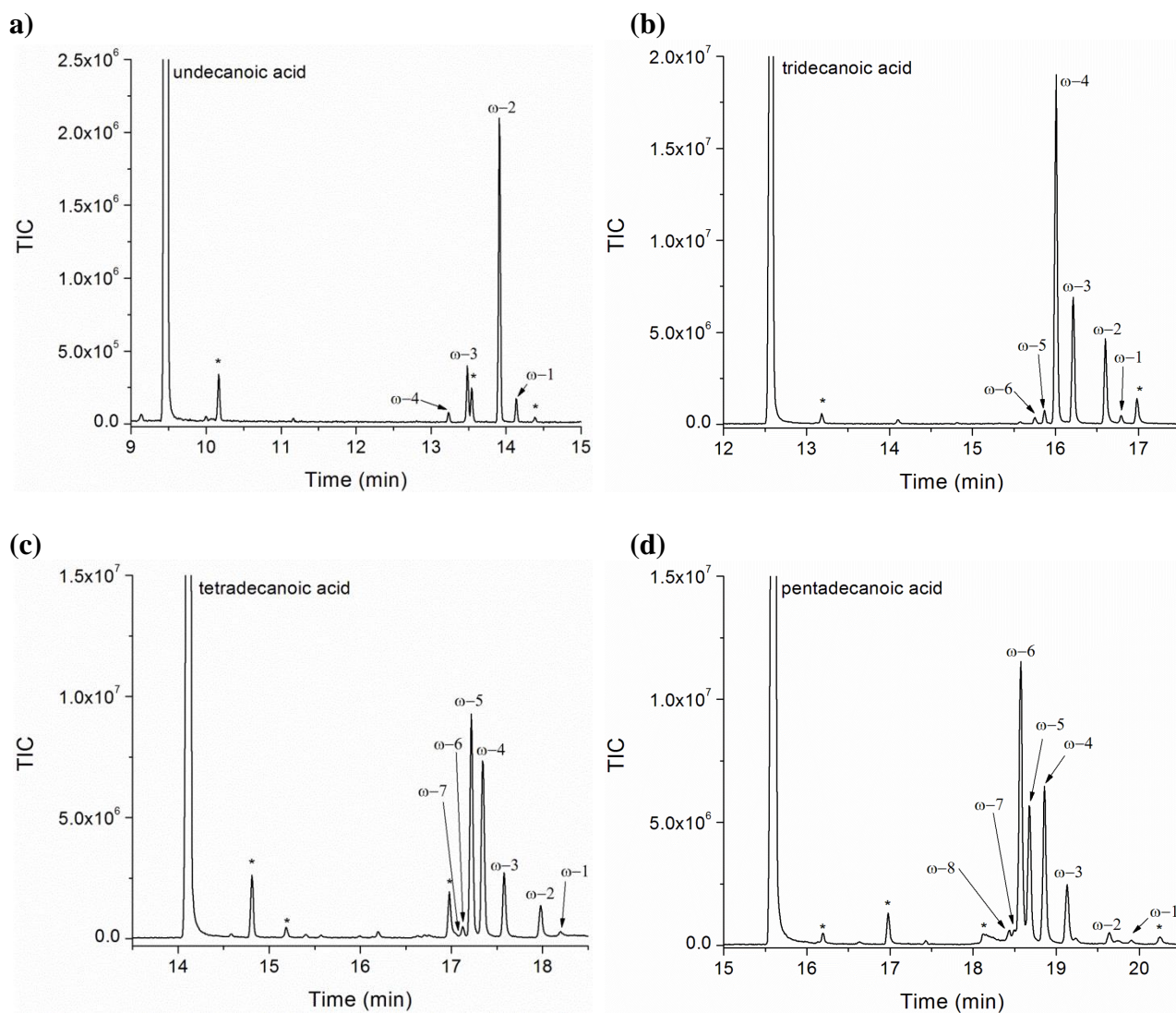


**Figure 4** GC-MS analyses of the BSTFA/TMSCl derivatised *in vitro* turnovers of Krac0936 with (a) tridecanoic acid, (b) tetradecanoic acid, (c) pentadecanoic acid and (d) hexadecanoic acid. The substrate and products are labelled as are impurities (with an asterisk).

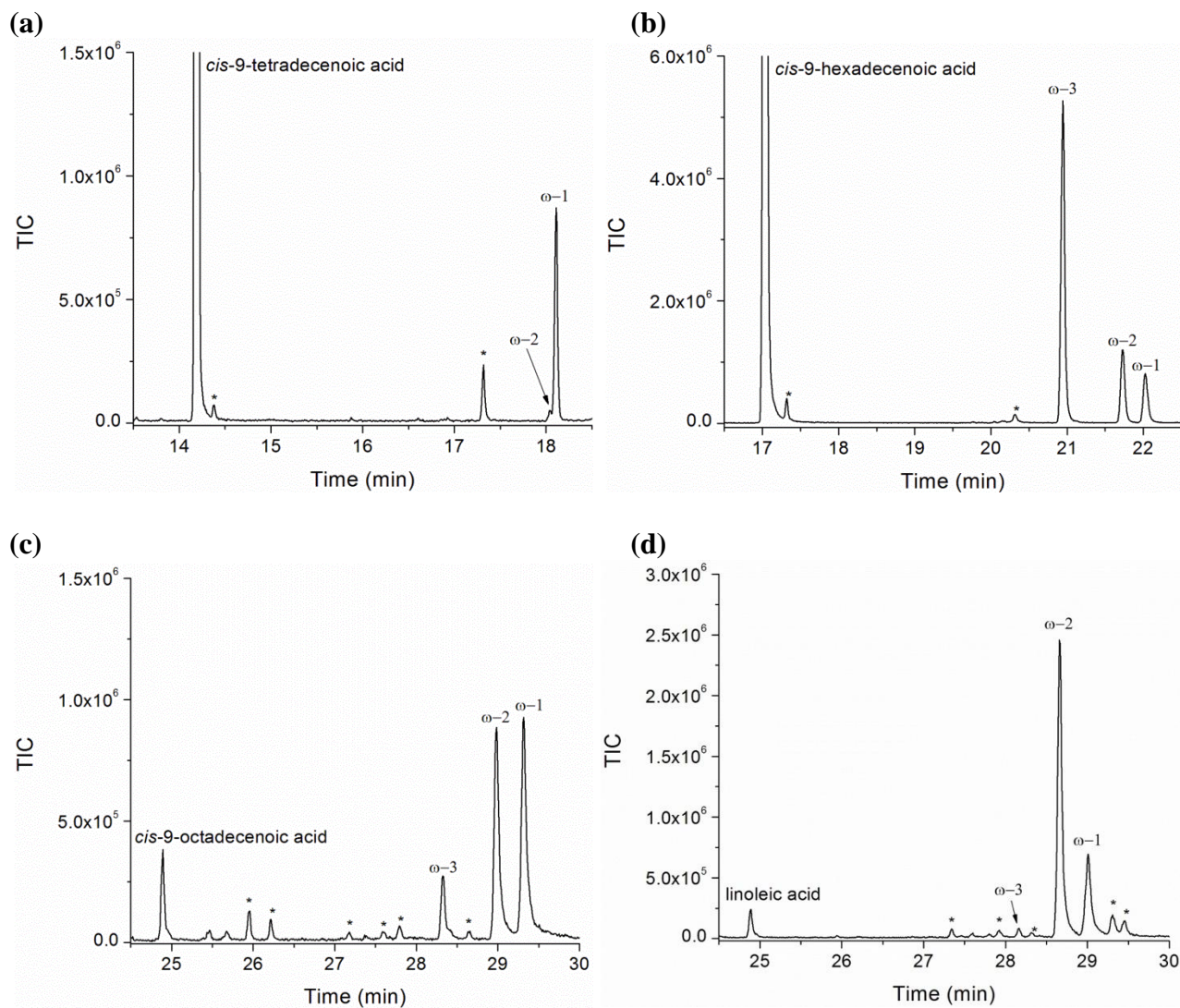




**Figure 5** GC-MS analyses of the BSTFA/TMSCl derivatised *in vitro* turnovers of Krac9955 with (a) undecanoic acid, (b) tridecanoic acid, (c) tetradecanoic acid and (d) pentadecanoic acid. The substrate and products are labelled as are impurities (with an asterisk).



**Figure 6** GC-MS analyses of the BSTFA/TMSCl derivatised *in vitro* turnovers of Krac0936 with (a) *cis*-9-tetradecenoic acid, (b) *cis*-9-hexadecenoic acid, (c) *cis*-9-octadecenoic acid and (d) *cis,cis*-9,12-octadecadienoic (linoleic) acid. The substrate and products are labelled as are impurities (with an asterisk).





**Figure 7** GC-MS analyses of the BSTFA/TMSCl derivatised *in vitro* turnovers of Krac9955 with (a) *cis*-9-tetradecenoic acid and (b) 10-undecenoic acid. The substrate and products are labelled, as are impurities (with an asterisk).

