UNIVERSITÀ DEGLI STUDI DI TORINO

CYP116B5: a new class VII catalytically self-sufficient cytochrome P450 from *Acinetobacter radioresistens* that enables growth on alkanes

Daniela Minerdi, Sheila J. Sadeghi, Giovanna Di Nardo, Francesco Rua, Silvia Castrignanò, Paola Allegra and Gianfranco Gilardi*

Department of Life Sciences and Systems Biology, University of Torino,

via Accademia Albertina 13, 10123 Torino, Italy

Corresponding author: Gianfranco Gilardi, Department of Life Sciences and Systems Biology, University of Torino, via Accademia Albertina 13, 10123 Torino, Italy, Email: gianfranco.gilardi@unito.it, Tel office: +39-011-6704593, Fax: +39-011-670 4643

Subject category: Microbial ecology and functional diversity of natural habitats

Keywords: *n*-alkane oxidation/ self-sufficient cytochrome P450 monooxygenase/ RT-PCR/*Acinetobacter radioresistens*/ gas chromatography

Running title: Oxidation of *n*-alkanes by a novel self-sufficient bacterial cytochrome P450.

Summary

In this work a gene coding for a novel self-sufficient class VII cytochrome P450 monooxygenase (CYP116B5) was identified in the genome of *Acinetobacter radioresistens* S13 that allows it to grow on media supplemented with medium (C14, C16) and long (C24, C36) chain alkanes as the sole energy source. Phylogenetic analysis of the N- and C-terminal domains of CYP116B5 suggests an evolutionary model of *cyp116b5* gene: a plasmid-mediated horizontal gene transfer probably occurred from the donor bacterium *Rhodococcus jostii* RHA1 to the receiving *A. radioresistens* S13. This event was followed by gene fusion and a subsequent integration of the new fused gene in *A. radioresistens* chromosome.

Heterologous expression of CYP116B5 in *E. coli* BL21, together with the *A. radioresistens* Baeyer-Villiger monooxygenase, allowed the recombinant bacteria to grow on long- and medium-chain alkanes, showing that CYP116B5 is involved in the first step of terminal oxidation of medium-chain alkanes overlapping AlkB and in the first step of sub-terminal oxidation of long-chain alkanes. Furthermore, it was also demonstrated for the first time that CYP116B5 is a naturally self-sufficient cytochrome P450 consisting of a haem domain (aa 1-392) involved in the oxidation step of *n*-alkanes degradation, and its reductase domain (aa 444-758) comprising the NADPH-, FMN- and [2Fe2S]-binding sites. To our knowledge CYP116B5 is the first member of this class to have its natural substrate and function identified.

Introduction

Bacterial oxidation of *n*-alkanes is a very common phenomenon in soil and water (Bühler and Schindler, 1984) and it is a major process in geochemical terms. Bacterial monooxygenases (MOs) are relevant to biogeochemistry (Hakemian and Rosenzweig, 2007), bioremediation (Wood, 2008) and biocatalysis (Leak *et al,* 2009). Based on cofactor and cellular location, five MO superfamilies can be distinguished in bacteria: soluble flavin-binding monooxygenases (FMO) (van Berkel *et al,* 2006), cytochromes P450 (CYP) (Bernhardt, 2006), soluble diiron monooxygenases (SDIMO) (Notomista *et al,* 2003), alkane hydroxylases (pAH/AlkB) (van Beilen and Funhoff, 2007) and copper-containing proteins (Lieberman and Rosenzweig, 2004). There are overlaps in the physiological substrate range of bacterial MOs (van Beilen and Funhoff, 2007): FMO and CYP enzymes act on alkanes, aromatics and heterocyclics, AlkB-type MO is typically involved in oxidation of liquid alkanes (C_5–C_{16}), SDIMOs act on small gaseous hydrocarbons (C_1–C_4 alkanes and alkenes) and single-ring aromatics (e.g. phenol, toluene), and copper MOs specifically oxidize small compounds, such as methane and ammonia.

Cytochromes P450 are a large family of haem-thiolate enzymes found in all biological kingdoms and they catalyze a wide range of oxidative reactions. Substrates include natural compounds, such as steroids, fatty acids, pheromones, leukotrienes and prostaglandins, as well as xenobiotics such as drugs and carcinogens (Dodhia and Gilardi, 2009). Their primary chemical reaction is the monooxygenation of the substrate involving the activation of molecular oxygen, followed by the insertion of a single atom of oxygen into an organic substrate with concomitant reduction of the other oxygen atom to water. Electron equivalents are supplied by NAD(P)H through an electron transfer chain involving different redox partners, the nature of which is

4

used to classify cytochromes P450 into ten classes (Hannemann *et al*, 2007; Sadeghi and Gilardi, 2013).

Despite the identification and cloning of an enormous number of cytochromes P450 from bacterial sources, there have been very few examples that do not fall into the general class I category of enzymes. An example of exception is the soluble fatty acid hydroxylase from *Bacillus megaterium*, P450-BM3, that is a catalytically self-sufficient single-polypeptide enzyme containing haem, FAD and FMN in equimolar ratios and requiring NADPH as a source of electrons (Narhi and Fulco, 1986). Two further P450 enzymes (CYP102A2 and CYP102A3) from *Bacillus subtilis* displaying an overall organization similar to that of P450-BM3 have also been identified (Gustafsson, 2000). Recently, a membrane-bound eukaryotic counterpart of P450-BM3 has been cloned from *Fusarium oxysporum* (Kitazume *et al*, 2000). A distinctive type of P450–redox partner fusion system was identified from genome analysis of pathogenic *Burkholderia* species, the heavy metal-tolerant bacterium *Cupriavidus metallidurans* CH34 and *Rhodococcus* sp. NCIMB 9784 (De Mot and Parret, 2002). In this system, P450 of undefined structure and substrate specificity, classified in the CYP116B P450 family, are fused to a reductase module with an aminoacid sequence resembling that of *Burkholderia cepacia* phthalate dioxygenase reductase (PDOR), an enzyme that provides electrons to phthalate dioxygenase (PDO) to enable this enzyme to oxygenate phthalate as a carbon source for growth (Gassner *et al*, 1995). As with P450-BM3, the P450 domain in the novel CYP116–PDOR fusion enzymes is fused at the N-terminus of the polypeptide chain. The *Rhodococcus* enzyme from this family (CYP116B2) exhibits P450-like spectral characteristics. Genetic dissection of the enzyme enabled the characterization of FMN and 2Fe-2S centres in the expressed PDOR module (Roberts *et al*, 2003; Hunter *et al*, 2005).

CYP116B2 was also shown to have weak activity towards 7-ethoxycoumarin, a prototypical fluorescent substrate for the P450 superfamily, but no physiologically relevant substrates were identified.

Bacterial cytochrome P450 systems involved in alkane degradation have been reported in *Rhodococcus rhodochrous* that contains an octane hydroxylating cytochrome P450 enzyme (Cardini and Jurtshuk, 1970). Another cytochrome P450 enzyme involved in alkane hydroxylation was isolated (Muller *et al*, 1989) and cloned from *Acinetobacter calcoaceticus* EB104 (Maier *et al*, 2001).

This paper investigates the presence of cytochrome P450 in *Acinetobacter radioresistens* S13, a bacterium isolated from soil surrounding an activated sludge pilot plant near Torino (Italy) (Giunta and Pessione, 1997) that can utilize medium- and long-chain alkanes as the sole carbon and energy source (Minerdi *et al,* 2012). We found that it possesses genes coding for terminal alkane hydroxylase (*alk*B) and a Baeyer-Villiger monooxygenase (*alm*A) involved in the initial steps of terminal and subterminal oxidation of medium- and long-chain alkanes, respectively. The two genes are differentially expressed according to the presence of medium- or long-chain alkanes (Minerdi *et al*, 2012).

Results

Screening for cytochrome P450 genes in *A. radioresistens* strain S13

In order to examine the presence of cytochrome P450 alkane hydroxylase in *A. radioresistens* S13, PCR experiments were performed with oligonucleotide primers based on the sequence of the P450 gene of *A. radioresistens* strain SK82. This led to a PCR product of 500 bp, a size compatible with the presence of a cytochrome P450 coding gene. Analysis of the full length nucleotide sequence obtained by a genome-walking strategy revealed the presence of one complete open reading frame (ORF) of 2,274 bp, hereinafter referred to as ORF2274. Sequence comparison using the BLAST programme (Altschul *et al*, 1997) against the non-redundant GenBank database showed that the amino acid sequence had very low E-values (the probability due to chance that there is another alignment with a similarity greater than the given score), when aligned with bacterial cytochrome P450 monooxygenases.

Phylogenetic and amino acid sequence analysis

A CLUSTALW multi-alignment was followed by visual inspection and manual editing of the *A. radioresistens* S13 P450 sequence. Phylogenetic tree of the whole amino acid sequence was inferred from the selected sequences by the Distant Matrix method. Results placed the sequence under investigation within the self-sufficient class VII P450 monooxygenases (Roberts *et al*, 2003) (Fig. 1).

The ORF2274 encodes a protein of 758 amino acids, with predicted molecular mass of 85,700 Da; A ClustalW multi-alignment of its amino acid sequence with cytochrome P450 sequences available in databases was used to identify conserved regions and to detect the degrees of sequence similarity and identity between the *A. radioresistens* S13 P450 protein and orthologous proteins. The sequence homology search using the aminoacid sequence coded by ORF2274 revealed significant

homology with the self-sufficient class VII cytochrome P450 (Roberts *et al*, 2002).

Comparison of the translated sequence from residue 1 to 392 with those of the SWISSPROT database shows a sequence homology of 62 % with the haem domain of CYP116B2 from *Rhodococcus* sp. NCIMB 9784 (Fig. 2A). This indicates that the *A. radioresistens* protein under investigation is a new member of CYP116B subfamily and the new enzyme was named CYP116B5 (http://drnelson.uthsc.edu/bacteria.html). The sequence [367]FGYGAHQCMG (Fig. 2A, red) matches the PROSITE consensus motif ([F/W][S/G/N/H]X[G/D]X[R/H/P/T]XC[L/I/V/M/F/A/P][G/A/D]) for cytochrome P450 enzymes, spanning the hydrophobic haem-ligand pocket showing that the sequence between amino acids 1-392 corresponds to a haem binding domain and the residue C374 is predicted to be the heam-iron fifth ligand. Another consensus motif is the [263]AAHETT (Fig. 2A, green) that corresponds to the putative oxygen binding sequence (Roberts *et al*, 2002).

The C-terminal portion from residue 444 to 758 of the translation product of the CYP116B5 gene shows sequence homology to reductase proteins. An alignment of this sequence with those of two representative reductase subunits of oxygenases from various sources is shown in Figure 2B. The C-terminal region of proteins of known function can be subdivided into three distinct functional parts: a FMN-binding domain, a NADH-binding domain and a ferredoxin-like [2Fe2S] domain. The sequence [518]SRGGS (Fig. 2B, grey) corresponds to the consensus motif [G/S]RGGS involved in binding the phosphate group of FMN. The fingerprint sequence GXGXXP for NADH binding, [559]GIGITP (Fig. 2B, purple) is a region highly conserved within the representative members of this class of enzymes. The crystal structure of the closely related phthalate dioxygenase reductase indicates that the proline residue in this sequence contacts the bound nicotinamide (Correll *et al*, 1992). Another signature

sequence is represented by the four highly conserved cysteine residues that are marked by asterisks in Figure 2B. Three of the four cysteines are closely clustered in the primary amino acid sequence [708]CTEGLCGSC, and it conforms to the PROSITE consensus motif CXX[G/A]XC[G/A/S/T]XC of a [2Fe2S] ferredoxin (Roberts *et al*, 2002). Taken together these results indicate that CYP116B5 is a class VII catalytically self-sufficient cytochrome P450 enzyme composed of a haem (residues 1-392) and a reductase domain (residues 444-758), the haem is ligated to C374 and the reductase domain comprises NADH and FMN binding sites, as well as a [2Fe2S] cluster involving cysteine residues 708, 713, 716 and 746.

A model for the evolution of the CYP116B5 gene

In order to investigate the evolutionary origin of A. *radioresistens* S13 *cyp116b5* coding gene, a phylogenetic analysis of the two separate N- and C-terminal domains of CYP116B5 was performed. In order to retrieve the most similar sequences the amino acid sequences of the two domains were used as a query to probe the protein database of completely sequenced proteobacterial genomes with the BLASTP and R-PSI BLASTP option of BLAST programme. As expected, neighbor joining tree positioned the CYP116B5 N-terminus-haem containing domain (N-HD) among class I cytochrome P450 monooxygenases but, surprisingly, far away from *Acinetobacter* class I P450s. As shown in Figure 3A, S13-HD clusters together with the class I P450 (CYP 116) from *R. jostii* strain RHA1 with a bootstrap value of 99 out of 100.

The C-terminal domain of CYP116B5 (C-OXD) shows high sequence similarity with the reductase domain of several monooxygenases, positioning C-OXD among 2Fe2S-flavin containing oxidoreductases. Interestingly, C-OXD clusters together with a 2Fe2S-flavin containing oxidoreductase from *R. jostii* RHA1, with a bootstrap value of 93 out of 100 (Fig. 3B). The gene coding for *R. jostii* RHA1 oxidoreductase is

adjacent to the gene coding for CYP116 of *R. jostii* RHA1 and both P450 and its reductase of this organism are located in the pRHL3 plasmid (http://www.ncbi.nlm.nih.gov/nuccore/NC_008271.1). For this reason the oxidoreductase of *R. jostii* is called OXRED116.

The results obtained from the phylogenetic analysis of the two separate N-HD and C-OXD domains of CYP116B5, together with the arrangement of the P450 and its reductase in the *R. jostii* and *A. radioresistens*, suggest a likely evolutionary model of the gene coding for CYP116B5. It can be speculated that a horizontal gene transfer event mediated by pRHL3 occurred from the donor bacterium *R. jostii* RHA1 to the receiving *A. radioresistens* S13. This event was followed by gene fusion and a subsequent integration of the new fused gene in *A. radioresistens* chromosome, probably as the result of recombination processes.

Expression of the CYP116B5 gene in media with alkanes as carbon source

In order to investigate the induction of the CYP116B5 gene by different alkanes, RT-PCR experiments were performed on RNA extracted from *A. radioresistens* S13 grown in minimal medium with the addition of C14 and C16 (medium length alkanes), and C24 and C36 (long-chain alkanes) as the sole energy and carbon source and from bacteria grown with sodium acetate (control). Using the *cyp116b5* gene specific primers an amplified product of the expected size of 500 bp, was obtained on RNA extracted from S13 grown with both medium- and long-chain alkanes (Fig. 4Aiii). No amplified product was obtained from the RNA of bacteria grown with sodium acetate or from the RT-negative controls (Fig. 4Aii). An amplified fragment of the expected size was also obtained using the *A. radioresistens* S13 16S rDNA specific primers on the RNA of the bacteria grown in the presence of medium- and long-chain length alkanes and sodium acetate (Fig. 4Ai). The amplicons obtained with

10

cyp116b5 and 16S rRNA primers were purified and sequenced. A nucleotide sequence of about 600 bp was obtained for all PCR products, showing a 100% sequence identity to S13 P450 and 16S ribosomal genes of *A. radioresistens* S13. These data suggest that the *cyp116b5* gene is indeed induced when the *A. radioresistens* S13 is grown both in medium- and long-chain alkanes.

Heterologous expression of CYP116B5 and co-expression of Ar-BVMO and CYP116B5 genes

In order to confirm the role of CYP116B5 in the first step of alkane oxidation, *E. coli* BL21 cells that are not able to grow on alkanes, were transformed with the expression plasmid pET-CYP116B5. Since pBAD and pT7 are incompatible plasmids, cyp116B5 coding gene was subloned in the expression vector pET (Novagen). A co-transformation of *E. coli* BL21 cells with pET-CYP115B5 and the expression vector pT7-Ar-BVMO harboring the *A. radioresistens* BVMO (Minerdi *et al*, 2012), was also carried out, in order to decipher the functions of the two enzymes in medium- and long-chain alkanes oxidation.

Spectrophotometric evidence for the presence of a P450 enzyme in the cell-free extract of both BL21 and TOP10 recombinant cells was obtained from the difference spectrum of the reduced and CO-bound form and the oxidised form (Fig. 4B). No absorption increase was observed at 450 nm in the non-transformed cells (data not shown).

The cell cultures containing both the expression vectors were analyzed by SDS-PAGE. A band at 57 kDa corresponding to the expected Ar-BVMO was detected on a Coomassie-stained gel from the cell extract of induced BL21 cells but was not observed in the negative control of the same strain harboring pT7-Ar-BVMO whose expression was not induced (Fig. 4Bii). Despite the presence of active CYP116B5 as

shown in the UV-vis spectrum (Fig. 4Bi) a band corresponding to this enzyme is not visible, and this is likely to be due to a low level of expression.

Heterologous expression proves CYP116B5 is an alkane monooxygenase

We used a heterologous expression approach to link the *cyp116b5* gene to metabolism of alkanes. The transformed *E. coli* BL21 growth on MSM medium supplemented with medium-length liquid (C14 and C16) and long-length solid (C24 and C36) alkanes was studied in liquid cultures. Growth was measured as an increase in OD_{600} for cultures supplemented with liquid alkanes (Fig. 5A), and for cultures supplemented with solid alkanes the number of colony-forming units were measured on LB agar plates on which 1 ml of liquid culture containing the alkane was plated (Fig. 5B). Recombinant *E. coli* BL21 growth in the presence of C14 and C16 alkanes was significantly higher than that measured in control experiments. On the contrary, no significant growth was detected when *E. coli*-pBAD-CYP116B5 was grown on minimal medium supplemented with C24 and C36 alkanes (Fig. 5B). This data demonstrate that in the case of subterminal oxidation of long-chain alkanes another monooxygenase is required.

In order to confirm the role of Ar-BVMO in the degradation of long-chain alkanes, *E. coli* BL21 was co-transformed with pET-CYP116B5 and pT7-Ar-BVMO. The growth of transformed BL21 cells was tested in MSM medium supplemented with C24 and C36 as the sole energy and carbon source. Results showed a significant growth both on C24 and C36 alkanes (Fig. 5C).

Activity of CYP116B5

The *A. radioresistens* S13 cytochrome P450 is proposed to be involved in terminal oxidation of medium-chain alkanes, and in order to confirm its activity partially purified CYP116B5 was incubated with tetradecane at 25°C for 26 hours (Chen *et al*,

2012) and the reaction products were analysed by gas chromatography. Gas chromatographic analysis of the reaction products showed a peak corresponding to tetradecanol (RT= 19.28 min). This peak was present only in the enzymatic reaction and not observed in the control (Fig. 6) demonstrating that the CYP116B5 is indeed capable of alkane oxidation.

Discussion

A. radioresistens S13 is capable of growing on minimal medium supplemented with medium- and long-chain alkanes, but relatively little is known about the genetic characteristics of its alkane-degradative system. We have previously reported the presence of *alm*A and *alk*B genes coding for a BVMO and AlkB hydroxylase, whose expression is induced by the presence of long- and medium-chain alkanes in the growth medium, respectively (Minerdi *et al*, 2012).

Enzymes of the AlkB and CYP153 families are known to catalyze the first step in the catabolism of medium-chain length alkanes, as they selectively oxidize the terminal carbon to produce the 1-alkanols enabling their host organisms to utilize alkanes as carbon sources (Koch *et al,* 2009).

Several bacterial strains can assimilate alkanes larger than C20. These strains usually contain several alkane hydroxylases. Those active on C10–C20 alkanes are usually related to *P. putida* GPo1 AlkB or to *Acinetobacter* sp. EB104 cytochrome P450. However, the enzymes that oxidize alkanes larger than C20 seem to be rather different. *Acinetobacter* sp. M1, that can grow on C13–C44 alkanes, contains a soluble Cu^{2+} dependent alkane hydroxylase that is active on C10–C30 alkanes. *Acinetobacter* strain DSM 17874 has been found to contain the *almA* gene that codes for a BVMO enzyme capable of oxidizing alkanes from C20 to >C32 in length (Throne-Holst *et al*, 2006). Several bacterial strains can degrade >C20 alkanes using enzyme systems that have still not been characterized and that may include new proteins unrelated to those currently known.

The presence of a gene coding for a self-suffcient P450 monooxygenase of class VII in the genome of *A. radioresistens* S13 is surprising because up to now only class I P450 enzymes have been identified in *Acinetobacter* species.

The evolutionary model we propose in this paper involves a plasmid-mediated horizontal gene transfer from *R. jostii* to *A. radioresistens* S13. This event was followed by gene fusion and a subsequent integration of the new fused gene in *A. radioresistens* chromosome. Many of these proteins comprise one or several domains. Gene fusions characteristically bring together proteins that function in a concerted manner, such as successive enzymes in metabolic pathways, enzymes and the domains involved in their regulation, or DNA-binding domains and ligand-binding domains in prokaryotic transcriptional regulators (Yanai *et al*, 2001). The selective advantage of domain fusion lies in the increased efficiency of coupling of the corresponding biochemical reaction or signal transduction step (Marcotte *et al*, 1999) and in the tight co-regulation of expression of the fused domains (Yanai *et al*, 2001). In some cases, the amount and source of horizontal gene transfer can be linked to an organism's lifestyle. Horizontal gene transfer has the capability of introducing, immediately upon integration, completely novel physiological traits (Koonin *et al*, 2001). The *Acinetobacter* genus has received particular attention because of its metabolic versatility and, in the case of clinical isolates, of its multidrug resistance pattern (Towner, 2009). *Acinetobacters* have the ability to take up extracellular DNA from the environment, a mechanism probably used for the acquisition of new functions by horizontal gene transfer (Barbe *et al*, 2004). *A. radioresistens* S13 was isolated from soil surrounding an activated sludge pilot plant for its fast phenol catabolism when used as the sole carbon and energy source (Pessione and Giunta, 1997). According to their capabilities to degrade xenobiotics, *A radioresistens* and *Rhodococcus jostii* can occupy the same ecological niche, thus it can be assumed that S13 CYP116B5 coding gene could be derived from *R. jostii* by gene fusion and

horizontal transfer event leading to the acquisition of new metabolic capabilities such as the oxidation of alkanes.

The *cyp116b5* gene is induced by the presence of both medium- and long-chain alkanes, while expression of the gene coding for alkane hydroxylase B is induced only by the presence of medium-chain alkanes (Minerdi *et al*, 2012). Some bacterial strains contain only one alkane hydroxylase, as is the case for the well-characterized alkane degrader *P. putida* GPo1. Many other strains have several alkane-degradation systems, each one being active on alkanes of a certain chain-length or being expressed under specific physiological conditions. *Alcanivorax borkumensis* has two AlkB-like alkane hydroxylases and three genes coding for cytochromes P450 believed to be involved in alkane oxidation (Schneiker *et al*, 2006). In addition, *A. borkumensis* seems to have other uncharacterized genes involved in oxidation of branched alkanes and phytane (Schneiker *et al*, 2006). The substrate range of the *A. borkumensis* AlkB-like alkane hydroxylases partially overlaps. AlkB1 oxidizes C_5–C_{12} *n*-alkanes, while AlkB2 is active on C_8–C_{16} *n*-alkanes (van Beilen *et al,* 2004).

Heterologous expression of novel bacterial alkane hydroxylases is complicated when there is a three component system where all the components are necessary for enzyme activity. Several alkane hydroxylase systems have been expressed heterologously. A DNA region of about 35 Kbp containing alkane hydroxylase system genes from *P. putida* Gpol was cloned into an *E. coli* strain and into a mutant strain of *P. putida* unable to grow on alkanes. These transformants metabolized *n*-alkanes as shown by mineralization and growth assays (Eggink *et al*, 1987). Heterologus expression of other alkane 1-monooxygenase genes from several bacteria such as *Rhodococcus* strains (Whyte *et al,* 2002), *Alkanicorax borkumensis* AP1 (Smits *et al,* 2002),

Prauserella rugosa NRRLB-2295 (Smits *et al,* 2002) and *Mycobacterium tuberculosis* H37Rv (Smits *et al,* 2002) were confirmed using this *P. putida* system.

By comparison to all the above-mentioned P450 systems we conclude that CYP116B5 from *A. radioresistens* is a self-sufficient enzyme. Heterologous expression of CYP116B5 in *E. coli* BL21 demonstrated that the bacterium is able to grow on minimal medium supplemented with medium chain alkanes but not in the presence of long chain alkanes. These data confirm that cytochrome P450 is involved in the first step of terminal oxidation of medium-chain alkanes overlapping AlkB and in the first step of sub-terminal oxidation of long-chain alkanes. In the case of long chain alkanes another monooxygenase is needed for their degradation and bacterial growth. Co-expression of both CYP116B5 and BVMO in *E. coli* allowed the cells to grow also with C24 and C36 alkanes, demonstrating that the long-chain alkane degradation in *A. radioresistens* S13 requires the additional activity of the BVMO.

In conclusion, we found a novel bacterial self-sufficient cytochrome P450 belonging to family CYP116 involved in the first steps of alkane oxidation and to our knowledge, it is the first enzyme of this class to have its natural substrate and function identified.

Experimental procedures

Bacterial strains and culture media

Acinetobacter radioresistens S13 was streaked onto Luria–Bertani (LB, 1.0% Tryptone 0.5% Yeast Extract 1.0% Sodium Chloride) agar medium and incubated at 30 °C. For long-term storage, bacterial culture was maintained at -80 °C in liquid LB containing 20% glycerol.

Escherichia coli strains TOP10 and BL21 (DE3) (Invitrogen, Karlsruhe, Germany) for vector propagation and heterologous expression of the recombinant proteins (using the pBAD and pT7 vectors) were grown at 37°C in LB medium containing 100 and 50 μg/ml ampicillin as a selection marker, respectively. In the case of BL21 cells transformed with pET-CYP116B5, 30 μg/ml kanamycin was used as the selection marker.

DNA preparation and PCR detection of cytochrome P450 gene

A. radioresistens S13 was grown on LB medium at 30°C for 24 hours and centrifuged for 10 min at 4,000 rpm. The pellet was resuspended in 250 μl of sterile water. This suspension was heated at 95°C for 5 min to release the DNA. The heated suspension (2 μl) was then used to PCR-amplify the gene coding for cytochrome P450. To amplify the genomic sequence of *A. radioresistens* S13 cytochrome P450 gene, primers SK82P450-Forward (5'-TTCCTCGTGAACAGCAACTG-3') and SK82P450-Reverse (5'-GCAGTGCCATATCAGCAAGA-3') were designed on the basis of the P450 gene of *A. radioresistens* strain SK82 sequence, available in databases (ZP_05361336.1). PCR reactions were carried out in a final volume of 20 μl and containing 0.5 mM of each primer, 2.0 μl of 10x buffer (Qiagen, Hilden, Germany), 2.5 mM $MgCl_2$, 250 mM of each dNTP, 2 μl of bacterial lysate preparation and 1U of *Taq* DNA polymerase (Qiagen, Hilden, Germany). The PCR cycling conditions were as follows: denaturation at 95°C for 3 min; 30 cycles at 94°C for 45 s, 55°C for 1 min and 72°C for 1 min; and a final extension at 72°C for 7 min using a Techne TC-312 thermal cycler. In order to obtain the full sequence of the gene, a genome walking strategy was adopted using primers specifically designed on the basis of *A. radioresistens* SK82 cytochrome P450 flanking gene sequences.

PCR product purification, sequencing and gene cloning in expression vector

The PCR product (about 700 bp) was excised from the gel and purified using the QIAquick PCR purification kit (Qiagen, Hilden, Germany). The purified product was then directly sequenced without cloning steps by using the PCR primers described above. The sequencing processes were conducted with an ABI model 3730 DNA sequencer by the Eurofins MWG operon sequencing service (Ebersberg, Germany). Two primers (S13P450-Forward: 5'-AAGTTCAGGATGAACGTAAAA-3'; S13-Reverse: 5'-TCATTTTCCCTGAATTTGTTGTTA-3') designed at the 5' and 3' of the sequence were used to amplify the full length S13 P450 gene using the PCR conditions described above except for the annealing temperature that was 62°C.

An amplified fragment of about 2,300 bp was obtained and cloned into pBAD-TOPO (Invitrogen) expression vector and subsequently *E. coli* BL21 were transformed with this plasmid. To verify the accuracy of amplification, the plasmid-DNA, propagated in *E. coli* BL21 and isolated with a NucleoSpin Plasmid Kit (Macherey-Nagel), was sequenced. Subsequently, for co-expression experiments, the *cyp*115b5 gene was subcloned in the pET30a expression vector (Novagene) using the restriction sites *EcoR*1 and *Nde*I. The correct insertion of the ORF was confirmed by DNA sequencing of the entire clone.

Nucleotide sequence accession number

The cytochrome P450 gene sequence obtained from *A. radioresistens* S13 was submitted to the GenBank database and assigned the Accession No. AET34456.1.

Sequence retrieval and alignment

Amino acid sequences from completely sequenced genomes of proteobacteria were retrieved from GenBank database (http://www.ncbi.nlm.nih.gov). BLAST (Altschul *et al*, 1997) probing of database was performed with the BLASTP option of this

program using default parameters. Only those sequences retrieved with an E-value below the 0.05 threshold were taken into account. Reverse Position-Specific BLAST (PSI) algorithm was used to search conserved domains in the *A. radioresistens* S13 P450 sequence.

The ClustalW (Thompson *et al*, 1994) program was used to perform pairwise and multiple amino acid sequence alignments. Alignments were manually checked and mis-aligned regions were removed.

Phylogenetic analysis

Phylogenetic analysis was performed using MEGA version 4 (Tamura *et al*, 2007), after multiple sequence alignment and truncation to the same length. Distances according to the Kimura two-parameter model (Kimura, 1980) and clustering with the neighbor-joining method (Saitou and Nei, 1987) were determined using bootstrap values based on 1,000 replications.

Growth of *A. radioresistens* S13 on *n*-alkanes

A. radioresistens S13 was grown in LB medium at 37°C on a rotary shaker at 180 rpm until growth reached the stationary phase (A_{600}= 3.8-4.0) and transferred to 500 ml flasks containing 100 ml of mineral salts medium (MSM) (Sakai *et al*, 1994). Alkanes were added either as 0.35% v/v liquid alkanes (C14 and C16, 99% purity, Sigma Aldrich, Italy) (Throne-Holst *et al,* 2006) or 3 g/l of solid alkanes (C24 and C36, 98% o purity, Sigma Aldrich, Italy) as the sole carbon and energy source. Absorbance at 600 nm was adjusted to 0.1. Flasks were incubated at 30°C in a rotary shaker operating at 220 rpm for 7 days. 1 ml samples were taken from each culture at regular intervals and were used to measure bacterial growth. For cultures supplemented with liquid alkanes, growth was measured spectrophotometrically (Agilent Technologies

model 8453E, Santa Clara, USA) at 600 nm using a cuvette with 1 cm path length. All the solid alkanes formed small micelles in the liquid medium, which made it impossible to accurately measure the OD_{600}. Therefore, growth for cultures supplemented with solid alkanes was measured as viable cell count on LB plates at 30°C. Identical medium and MSM plus sodium acetate (5g/l) as carbon source in place of alkanes were used as controls. All experiments were done in triplicates.

All alkanes of defined chain lengths were purchased from Sigma-Aldrich (Italy). For simplicity, alkanes of defined chain lengths are referred to by the number of carbon atoms they contain, i.e. tetradecane will be referred to as C14, hexadecane as C16, tetracosane as C24 and hexatriacontane as C36 throughout the article.

Semi-quantitative reverse transcription RT-PCR analyses

Total RNA for RT-PCR from *A. radioresistens* S13 grown in the presence of C14, C16, C24 and C36 alkanes and sodium acetate added as a control to the minimal growth medium was isolated using RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. cDNA was synthesized in a two-step process using Superscript II (Invitrogen). The primers used in RT-PCR reactions consisted of the following couple specifically designed on the basis of the sequence of *A. radioresistens* S13 P450 gene: S13P450-Forward (5': TGTCTGGCAGTTTGTTCTGG-3') and S13P450-Reverse (5': ACTACCGCACTAGCCTCTGC-3'). Single-stranded cDNA was obtained with the specific antisense primer S13P450-Reverse using total RNA as the template. RNA samples were denatured at 65°C for 5 min and then reverse transcribed at 50°C for 1 hour in a final volume of 20 μl containing 2 μl of total RNA, 1mM each of specific primers, 0.5 mM dNTPs, 10U RNase inhibitor, 4 μl buffer (Invitrogen), 2 μl 0.1M dithiothreitol and 1 μl of Superscript II (Invitrogen). The absence of undigested

genomic DNA was assessed by a control PCR carried out with Platinum *Taq* DNA polymerase (Invitrogen) and the *A. radioresistens* S13 16S rRNA gene specific primers (Minerdi *et al*, 2012) using the five total RNAs as templates and the PCR conditions described above. RT-PCR experiments were conducted using three replicates on three independent samples. The amplified products were analyzed by 1.4% agarose gel electrophoresis in a TAE running buffer (Sambrook *et al*, 1989).

Heterologous expression of cytochrome P450 in *E. coli* BL21 and TOP10 cells

Expression of *A. radioresistens* cytochrome P450 (CYP116B5) was achieved starting from both a colony of *E. coli* TOP10 and BL21 transformed with the pBAD-CYP116B5 (Ampicilline resistant) and pET-CYP116B5 (Kanamycin resistant) plasmid respectively, grown overnight in 5 ml of LB with 100 μg/ml of ampicillin and 30 μg/ml of kanamycin at 37°C. This was used as the inoculum of 500 ml of Terrific broth (TB), containing 100 μg/ml of ampicillin. The cultures were grown at 37°C, until an optical density at 600 nm of 0.4-0.6 was achieved. At this point protein production was induced by the addition of L-arabinose (0.2%) for pBAD-CYP116B5 and 1mM IPTG in the case of pET-CYP116B5. 0.5 mM δ−aminolevulinic acid (haem precursor) and 20 mg/l riboflavin (FMN precursor) were also added. The induced cells were then grown for 48 h at 24°C. After this time, the cells were harvested by centrifugation at 4,000 rpm for 20 min at 4°C and resuspended in lysis buffer (50 mM potassium phosphate pH 7.4 and 0.1 mM phenylmethyl sulfonyl fluoride). Lysozyme was added to the suspension at a final concentration of 1 mg/ml and the cell suspension was stirred for 30 min at 4°C. Lysis was performed by sonication (5 x 30s pulses) on ice with a Misonix Ultrasonic Sonicator (Teltow, Germany). The cell-free extracts were obtained by ultracentrifugation at 40,000 rpm for 20 min at 4°C. Expression of the *A. radioresistens* P450 (CYP116B5) was detected by using the

supernatant and measuring the formation of the reduced and carbon monoxide bound form of the P450 from the difference spectrum obtained at 450 nm using a Hewlett-Packard 8453 diode array spectrophotometer.

Co-transformation and co-expression of cytochrome P450 and BVMO in *E. coli* BL21

E. coli BL21 (DE3) was transformed with an equal concentration of pET-CYP116B5 and pT7-Ar-BVMO plasmids (Minerdi *et al*, 2012). In order to co-express CYP116B5 and Ar-BVMO a combination of the expression conditions described above and in Minerdi *et al* (2012) was used. After the co-expression the cells were collected by centrifugation, one aliquote was used for growth experiments and the other was lysed for spectroscopy and SDS-PAGE analysis.

Expression of the Ar-BVMO gene was detected by analyzing the cell lysate by SDS-PAGE and comparing the resulting protein banding pattern with that of a negative control. In the case of CYP116B5 expression, spectrophotometric analysis of the cell lysate upon reduction with sodium dithionite and the subsequent bubbling of carbon monoxide was used to obtain a difference spectrum at 450 nm, as mentioned earlier.

Growth of *E. coli* BL21 on *n*-alkanes

E. coli BL21 cells transformed with pET-CYP116B5 and co-transformed with pET-CYP116B5/pT7-Ar-BVMO were washed 5 times with MSM medium and were transferred to 500 ml flasks containing 100 ml MSM medium with 50 μg/ml of ampicillin, 25 μg/ml kanamycin, and 1 mM IPTG. Medium- and long-chain alkanes were added in the same conditions described above. Non induced *E. coli* BL21 cells were used as controls. The surfactant Alkanol (Sigma) was added at a concentration of 1% v/v.

1 ml samples were taken from each culture after 36 hours and were used to measure bacterial growth. These experiments were done to investigate the involvement of CYP116B5 and Ar-BVMO in the degradation of medium- and long-chain alkanes.

Partial purification of CYP116B5

Expression of *A. radioresistens* CYP116B5 was achieved as described above. The bacterial lysate from transformed TOP10 cells was centrifuged at 40,000 rpm for 20 min at 4°C and then the supernatant was loaded on a DEAE column pre-equilibrated with buffer A (500 mM Tris-HCl pH 7.2, 1mM EDTA). CYP116B5 was eluted using a gradient of 0-500 mM KCl. Eluted fractions that contained the protein (peak absorbance at 418 nm) were pooled and loaded on a Q-sepharose column pre-equilibrated with buffer A. The column was washed extensively with the same buffer and CYP116B5 was eluted using a gradient of 0-500 mM KCl. Fractions containing CYP116B5 were pooled and buffer exchanged to buffer A by Amicon Ultra centrifugal filters (Millipore, Watford, UK).

Enzyme-tetradecane incubation

In order to prove that alkanes are substrates of CYP116B5, the partially purified enzyme was incubated with the liquid alkane tetradecane for 26 hours at 25°C in 5 ml glass vials.

The concentration of properly folded P450 protein was determined from the CO difference spectrum of *E. coli* extract after sonication, removal of cell debris, and bubbling CO into the sample. The reactions were carried out in buffer A containing partially purified protein (5 μM). Substrate was added to this solution as 10 μl of 400 mM total substrate in 1% ethanol. After 1 minute, NADPH was added at a concentration of 1.5 mM. Reaction products were extracted by using hexane. The tubes were vortexed and then centrifuged at 14,000 rpm for 5 minutes in a

microcentrifuge. The hexane layer was removed with a pipette and analyzed by gas chromatography to determine product formation. Control reactions were performed by repeating these steps in the absence of the enzyme.

Gas chromatography

Identification of analytes was performed using tetradecane and tetradecanol standards (Sigma). All samples were injected at a volume of 2.0 μl and analyses were performed in triplicates. Qualitative analysis of transformation reaction product was performed on an Agilent 7890 A gas chromatograph with a flame ionization detector (FID). Direct analysis of tetradecane hydroxylation products were performed on a DB-WAX capillary column (cross-linked/surface bonded 100% polyethylene glycol, 30 m length, 0.32 mm ID, 0.25 mm film thickness) connected to the FID detector. The program used for separating the alcohol product was 250°C injector, 250°C detector, 50°C oven for 2 minutes, followed by 10°C/minute gradient up to 250°C, and finally 250°C for 3 minutes.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25:** 3389–3402.

Barbe, V., Vallenet, D., Fonknechten, N., Kreimeyer, A., Oztas, S., Labarre, L., *et al*. (2004) Unique features revealed by the genome sequence of *Acinetobacter* sp. ADP1, a versatile and naturally transformation competent bacterium. *Nucleic Acids Res* **32**: 5766-5779.

Bernhardt, R. (2006) Cytochromes P450 as versatile biocatalysts. *J Biotechnol* **124**: 128–145.

Bühler, M., and J. Schindler. (1984) Aliphatic hydrocarbons. In *Biotransformations*. Kieslich, K. (ed.). Verlag Chemie Weinheim, Weinheim, Germany. pp. 329-385.

Cardini, G., and Jurtshuk, P. (1970) The enzymatic hydroxylation of *n*-octane by *Corynebacterium* sp. strain 7E1C. *J Biol Chem* **245:** 2789-2796.

Correll, DL, Jordan, T.E., and Weller, D.E. (1992) Nutrient flux in a landscape: Effects of coastal land use and terrestrial community mosaic on nutrient transport to coastal waters. *Estuaries* **15**: 431-442.

Chen, M.M.Y., Snow, C.D., Vizcarra, C.L., Mayo, S.L., and Arnold, F.H. (2012) Comparison of random mutagenesis and semi-rational designed libraries for improved cytochrome P450 BM3-catalyzed hydroxylation of small alkanes. *Protein Eng Des Sel* **25**: 171-178.

De Mot, R., and Parret, A.H. (2002) A novel class of self-sufficient cytochrome P450 monooxygenases in prokaryotes. *Trends Microbiol* **10**: 502-508.

Dodhia, V.R. and Gilardi, G. (2009) Cytochromes P450: Tayloring a class of enzymes for biosensing in "Engineering the bioelectronic interface". Pp. 154-193. Davis, J. (ed.), RSC Publishing, London, UK.

Eggink, G., van Lelyveld, P. H., Arnberg, A., Arfman, N., Witteveen *et al. (*1987). Structure of the Pseudomonas putida alkBAC operon. Identification of transcription and translation products. *J Biol Chem* **262**: 6400-6406.

Gassner, G.T., Ludwig, M.L., Gatti, D.L., Correll, C.C., and Ballou, D.P. (1995) Structure and mechanism of the iron-sulfur flavoprotein phthalate dioxygenase reductase. *FASEB J* **9**: 1411-8.

Gustafsson, M.C.U. (2000) Ph.D. thesis. Bacterial cytochromes P450: studies on cytochrome P450 102A2 and P450 102A3 of *Bacillus subtilis.* Lund University, Lund, Sweden.

Hakemian, A.S., and Rosenzweig, A.C. (2007). The biochemistry of methane oxidation. *Annu Rev Biochem* **76**: 223-241.

Hannemann, F., Bichet, A., Ewen, K.M., Bernhardt, R. (2007) Cytochrome P450 systems-biological variations of electron transport chains. *Biochim Biophys Acta* **1770**: 330-344.

Hunter, D.J., Roberts, G.A., Ost, T.W., White, J.H., Muller, S., Turner, N.J., Flitsch, S.L., and Chapman, S.K. (2005) Analysis of the domain properties of the novel cytochrome P450 RhF. *FEBS Lett* 579: 2215-2220.

Kimura, M. (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16:** 111-120.

Kitazume, T., Takaya, N., Nakayama, N., and Shoun, H. (2000) *Fusarium oxysporum* fatty-acid subterminal hydroxylase (CYP505) is a membrane-bound eukaryotic

counterpart of *Bacillus megaterium* cytochrome P450-BM3. *J Biol Chem* **275:** 39734-39740.

Koch, DJ., Chen, M.M., van Beilen, J.B., and Arnold, F.H. (2009) In vivo evolution of butane oxidation by terminal alkane hydroxylases AlkB and CYP153A6. *Appl Environ Microbiol* **75** : 337-344.

Koonin, E.V., Makarova, K.S., and Aravind, L. (2001) Horizontal gene transfer in prokaryotes: quantification and classification. *Ann Rev Microbiol* **55**: 709-742.

Leak, D.J., Sheldon, R.A., Woodley,J.M., and Adlercreutz, P. (2009) Biocatalysts for selective introduction of oxygen. *Biocatal Biotransformation* **27**: 1-26.

Lieberman, R.L., and Rosenzweig, A.C. (2004) Biological methane oxidation: regulation, biochemistry, and active site structure of particulate methane monooxygenase. *Crit Rev Biochem Mol Bi ol* **39**: 147-64.

Maier, T., Foerster, H.H., Asperger, O., and Hahn, U. (2001) Molecular Characterization of the 56-kDa CYP153 from *Acinetobacter* sp. EB104. *Biochem Biophys Res Commun* **286:** 652-658.

Marcotte, E., Pellegrini, M., Yeates, T., and Eisenberg, D. (1999) A census of protein repeats. *J Mol Biol* **293**: 151-160.

Minerdi, D., Zgrablic, I., Sadeghi, S.J., and Gilardi, G. (2012) Identification of a novel Baeyer-Villiger monooxygenase from *Acinetobacter radioresistens*: close relationship to the *Mycobacterium tuberculosis* prodrug activator EtaA. *Microb Biotechnol*: **5**: 700-716.

Muller, R., Asperger, O., and Kleber, H.P. (1989) Purification of Cytochrome P-450 from *n*-hexadecane-grown *Acinetobacter Calcoaceticus*. *Biome Bioche Act.* **48:** 243-254.

Narhi, L.O., and Fulco, A.J. (1986) Characterization of a catalytically self-sufficient 119,000 dalton cytochrome P450 monooxygenase induced by barbiturates in *Bacillus megaterium*. *J Biol Chem* **261:** 7160-7169.

Notomista, E., Lahm, A., Di Donato, A., and Tramontano, A. (2003) Evolution of bacterial and archaeal multicomponent monooxygenases. *J Mol Evol* 56: 435-445.

Pessione, E., and Giunta, C. (1997) *Acinetobacter radioresistens* metabolizing aromatic compounds. Biochemical and microbiological characterization of the strain. *Microbios* **89:** 105-117.

Roberts, G.A., Grogan, G., Greter, A., Flitsch, S.L., and Turner, N.J. (2002) Identification of a new class of cytochrome P450 from a *Rhodococcus* sp. *J Bacteriol* **184**:3898-908.

Roberts, G.A., Celik, A., Hunter, D.J., Ost, T.W., White, J.H., Chapman, S.K., Turner, N.J., and Flitsch, S.L. (2003) A self-sufficient cytochrome P450 with a primary structural organization that includes a flavin domain and a [2Fe-2S] redox center. *J Biol Chem* **278**: 48914-48920.

Sadeghi, S.J. and Gilardi, G. (2013) Chimeric P450 Enzymes: activity of artificial redox fusions driven by different reductases for biotechnological applications. *Biotechnol Appl Biochem* **60**: 102-110.

Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4:** 406-425.

Sakai, Y., Maeng, J.H., Tani, Y., and Kato, N. (1994) Use oflong-chain n-alkanes (C13–C44) by an isolate, *Acinetobacter* sp. M-1. *Biosci Biotechnol Biochem* **58:** 2128-2130.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press.

Schneiker, S., Martins dos Santos, V.A., Bartels, D., Bekel,T., Brecht, M., Buhrmester, J., *et al* (2006) Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis*. *Nat Biotechnol* **24:** 997-1004.

Smits, T.H., 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**: 1733-1742.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* **24:** 1596-1599.

Thompson, J. D., Higgins, D.G., and Gibson, T. J. (1994) CLUSTALW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight-matrix choice. *Nucleic Acids Res* **22:** 4673-4680.

Throne-Holst, M., Markussen, S., Winnbergm, A., Ellingsen, T.E., and Kotlar, H.K,. Zotchev, S.B. (2006) Utilization of *n*-alkanes by a newly isolated strain of *Acinetobacter venetianus*: the role of two AlkB-type alkane hydroxylases. *Appl Microbiol Biotechnol* **72:** 353-360.

Towner, K.J. (2009) *Acinetobacter:* an old friend, but a new enemy. *J Hosp Infect* **73**: 355-363.

van Beilen, J.B., and Funhoff, E.G. (2007) Alkane hydroxylases involved in microbial alkane degradation. *Appl Microbiol Biotechnol* **74**: 13-21.

van Beilen, J.B., Marin, M.M., Smits,T.H., Rothlisberger, M.A.G., Witholt, B., Rojo, F. (2004) Characterization of two alkane hydroxylase genes from the marine hydrocarbonoclastic bacterium *Alcanivorax borkumensis*. *Environ Microbiol* **6:** 264-273.

van Berkel, W.J.H., Kamerbeek, N.M., and Fraaije, M.W. (2006) Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J Biotechnol* **124:** 670-6.

Wood, T.K. (2008) Molecular approaches in bioremediation. *Curr Opin Biotechnol* **19**: 57.

Whyte, L.G., Smits, T.H.M., Labbe, D., Witholt, B., Greer, C.W. and Van Beilen, J.B. (2002) Gene cloning and characterization of multiple alkane hydroxylase systems in *Rhodococcus* strains Q15 and NRRL B-16531. *Appl Environ Microbiol* **68**: 5933-5942.

Yanai, I., Derti, A., and DeLisi, C. (2001) Genes linked by fusion events are generally of the same functional category: a systematic analysis of 30 microbial genomes. *Proc Natl Acad Sci USA* **98:** 7940-7945.

.

Titles and legends to figures

Figure 1

Phylogenetic relationships within naturally self-sufficient cytochrome P450 enzymes. The sequence of CYP116B5 was aligned with class VII cytochrome P450 sequences from other bacteria and the un-rooted phylogenetic tree was calculated by using CLUSTALW and neighbour-joining method. The numbers at the nodes are the bootstrap confidence values obtained after 1000 replicates. The scale bar indicates distance in substitutions per nucleotide. The species and proteins are: *R. eutropha*= cytochrome P450 (YP_299468.1) from *Ralstonia eutropha* JMP134; *C. metallidurans*= cytochrome P450 from *Cupriadivus metallidurans* (YP_587063.1); *V. paradoxux*= cytochrome P450 from *Variovorax paradoxus* S110 (YP_002945479.1); *A. pieachaudii*= cytochrome P450 from *Achromobacter piechaudii* ATCC 43553 (ZP_ 06690070.1); *B. ubonensis*= cytochrome P450 from *Burkholeria ubonensis* (ZP_02376798.1); *B. thailandensis*= cytochrome P450 from *Burkholderia thailandensis* (ZP_438922.1); *B. oklahomensis*= cytochrome P450 from *Burkholderia oklahomensis* (ZP_02360413.1); *C. testosteroni*= cytochrome P450 from *Comamonas testosteroni* KF-1 (ZP_ 03546158.1); *D. acidovorans*= cytochrome P450 from *Delftia acidovorans* SPH-1 (YP_001561718.1); *R. equi*= cytochrome P450 from *Rhodococcus equi* ATCC 33707 (ZP_ 06830897); *R. ruber*= cytochrome P450 from *Rhodococcus ruber* (AAY1795001); *S. viridis*= cytochrome P450 from *Saccharomonospora viridis* DSM 43017 (YP_003132604.1); *T. bispora*= cytochrome P450 from *Thermobispora bispora* (YP_003651726.1); *K. sedentarius*= cytochrome P450 from *Kytococcus sedentarius* DSM20547 (YP_003149991.1); *S. witchii*= cytochrome P450 from *Sphingomonas witchii* RW1 ()YP_001263583.1*); S.*

aggregata= cytochrome P450 from *Stappia aggregata* AM 12614 (ZP_01549958.1);

A. radioresistens S13= cytochrome P450 from *Acinetobacter radioresistens* S13 (AET34456.1); *A. radioresistens* SK82= Cytochrome P450 from *Acinetobacter radioresistens* SK82 (ZP_05361336.1); *A. radioresistens* SH164= cytochromome P450 from *Acinetobacter radioresistens* SH164 (EEY86620.1).

Figure 2

A) Aminoacid sequence alignment of the N-terminal portion (residues 1-392) of CYP116B5 with those of other P450 enzymes. The conserved oxygen and haem binding sites are highlighted in green and red, respectively. The conserved cysteine residue that provides the fifth ligand to the haem iron is marked by an asterisk.

RHERY= thiocarbamate-inducible cytochrome P450 from *Rhodococcus erythropolis* NI86/21 (ZP_04386486.1); BACSUB= cytochrome P450 from *Bacillus subtilis* (AAC00266.1); PSSP= terpene-inducible cytochrome from *Pseudomonas* spp. (AAA25996.1); MYCOTUB= putative cytochrome P450 from *Mycobacterium tuberculosis* (CAA17439.1); ACRAD= cytochrome P450 from *A. radioresistens* S13 (AET34456.1).

B) Aminoacid sequence alignment of the C-terminal portion (residues 444-758) of CYP116B5 (ACRAD) with two different dioxygenase reductase subunits. The FMN-binding motif is highlighted in grey, and the NADH- binding motif is highlighted in purple. The four cysteine residues which are involved in binding the iron-sulfur cluster are marked with asterisks. POBB= phenoxibenzoate dioxygenase ??subunit from *Pseudomonas pseudoalcaligenes* (Q52186.1); VANB= vanillate O-demethylase oxidoreductase from *Pseudomonas aeruginosa* E2 (EKA50456.1).

Figure 3

Phylogenetic analysis of the N-terminal **(A)** and C-terminal **(B)** domain of cytochrome P450 of *A. radioresistens* S13. Neighbor joining tree was built using pairwise deletion with p- distance options. The numbers at the nodes are the bootstrap confidence values obtained after 1000 replicates. The scale bar indicates distance in substitutions per nucleotide. The species and proteins are:

(A) N-term. *A. radioresistens* SK82= N-terminal domain of cytochrome P450 class VII from *Acinetobacter radioresistens* SK82 (ZP_05361336.1); N-term. *A. radioresistens* SH164= N-terminal domain of cytochrome P450 class VII from *Acinetobacter radioresistens* SH164 (EEY86620.1); N-term. *A. radioresistens* S13= N-terminal domain of CYP116B5 from *Acinetobacter radioresistens* S13 (AET34456.1); CYP116 *R. josti* RHA1= cytochrome P450 116 class I from *Rhodococcus jostii* RHA1 (YP_709125.1); CYP *Roseobacter* sp. MED193= cytochrome P450 class I from *Roseobacter* sp. MED193 (ZP_01058383.1); CYP *Bradyrhizobium* sp. BTAi1= cytochrome P450 class I from *Bradyrhizobium* sp. BTAi1 (YP_001237726); CYP *C. glutamicum*= cytochrome P450 class I from *Corynebacterium glutamicum* AGC13032 (YP_224846); CYP *Arthrobacter* sp. FB24= cytochrome P450 class I (YP_831475.1) from *Arthrobacter* sp. FB24; CYP254 *R. jostii* RHA1= cytochrome P450 class I from *Rhodococcus jostii* RHA1 (ABG92259.1); CYP257 *R. jostii* RHA1= cytochrome P450 class I from *Rhodococcus jostii* RHA1 (YP_708874.1); CYP *C. segnis*= cytochrome P450 class I from *Caulobacter segnis* (YP_003594400.1); CYP *C. segnis*= cytochrome P450 class I from *Caulobacter segnis* (YP_003594400.1); CYP *Caulobacter* sp. K31= cytochrome P450 class I from *Caulobacter* sp. K31 (YP_001686023,1); CYP153 *Acinetobacter* sp. ADP1: cytochrome P450 class I from *Acinetobacter* sp. ADP1 (CAG68425.1);

CYP111 *Acinetobacter* sp. OC4= cytochrome P450 class I from *Acinetobacter* sp. OC4 (BAE78452.1);); CYP153 *Acinetobacter* sp. EB104= cytochrome P450 class I from *Acinetobacter* sp. EB104 (NP_114222.1); CYP147 *R. jostii* RHA1= cytochrome P450 class I from *Rhodococcus jostii* RHA1 (ABG94315); CYP256 *R. jostii* RHA1= cytochrome P450 class I from *Rhodococcus jostii* RHA1 (YP_708186.1).

(B) oxred *R. jostii* RHAI= 2Fe2S-oxidoreductase from *Rhodococcus jostii* RHA1 (YP_709124.1); oxred *Rhodococcus* sp. DK17= 2Fe2S, flavodoxin oxidoreductase from *Rhodococcus* sp. DK17 (ABD14375.1); oxred *R. opacus* B4= 2Fe2S, flavodoxin oxidoreductase from *Rhodococcus opacus* B4 (YP_002778964.1); oxred 116 *R. jostii*= 2Fe2S, flavodoxin 116 from *Rhodococcus jostii* (YP_709124.1); C-term *A. radioresistens* S13= C-terminal domain of CYP116B5 from *Acinetobacter radioresistens* S13 (AET34456.1); C-term *A. radioresistens* SH164= C-terminal domain of CYP116B5 from *Acinetobacter radioresistens* SH64 (EEY86620.1); C-term *A. radioresistens* SK82= C-terminal domain of CYP116B5 from *Acinetobacter radioresistens* SK82 (ZP_05361336.1); oxred *Janthinobacterium* sp. Marseille= 2Fe2S, 2Fe2S, flavodoxin oxidoreductase from *Janthinobacterium* sp. Marseille (YP_001354196.1); oxred *D. acidovorans*= 2Fe2S, flavodoxin oxidoreductase from *Delftia acidovorans* (YP_001562385.1); oxred *C. taiwanensis*= 2Fe2S, flavodoxin from *Cupriavidus taiwanensis* (YP_002007256.1); oxred *R. euthropha* H16= 2Fe2S, flavodoxin oxidoreductase from *Ralstonia eutropha* H16 (YP_728901.1); oxred *S. japonicum* UT26S= 2Fe2S, flavodoxin-oxidoreductase from *Sinorhizobium japonicum* UT26S (YP_003547113.1); oxred *Curvibacte*r sp. = 2Fe2S, flavodoxin oxidoreductase from *Curvibacter* sp. (CBA29570.1); oxred *Acinetobacter* sp. DR1= 2Fe2S, flavodoxin oxidoreductase from *Acinetobacter* sp. DR1 (YP_003731973.1); oxred *Acinetobacter* sp. SH024= 2Fe2S, flavodoxin oxidoreductase from

Acinetobacter sp. SH024 (ZP_06690260.1); oxred *Acinetobacter* sp. RUH2624=
2Fe2S, flavodoxin oxidoreductase from *Acinetobacter* sp. RUH2624
(ZP_05824483.1); oxred *A. baumanni* AYE = 2Fe2S, flavodoxin oxidoreductase
from *Acinetobacter baumanni* AYE (YP_001713598.1); oxred *A. baumanni* AB900=
2Fe2S, flavodoxin oxidoreductase from *Acinetobacter baumanni* AB900
(ZP_04663333.1).

Figure 4

A) Agarose gel of RT-PCR products amplified by using *A. radioresistens* S13 16S
rDNA. i): experiment; ii) control without RT and iii) CYP116B5 gene specific
primers. RNA from S13 grown in minimal medium supplemented with C14 (lane 1),
C16 (lane 2), C24 (lane 3) and C36 (lane 4) alkane; control experiment of RNA from
S13 grown in minimal medium supplemented with sodium acetate (lane 5); lane M=
DNA mass ladder (XL 1 Kb, Eppendorf), lane M'= DNA mass ladder (HyperLadder
I, Bioline).

B) i) Difference spectrum in the presence of carbon monoxide for cell-free extract of
E. coli TOP10-pBAD-CYP116B5 (black dotted line), *E. coli* BL21-pET-CYP116B5
(gray dashed line) and *E. coli* BL21 co-transformed with pET-CYP116B5 and pT7-
Ar-BVMO (black line).

(ii) SDS-PAGE of cell extract showing the expression of the Ar-BVMO gene in co-
transformed *E. coli* BL21. Lane 1: non induced crude extract; lane 2: purified Ar-
BVMO, lane 3: IPTG induced crude extract.

Figure 5

Growth of *E. coli* BL21-pET-CYP116B5 on *n*-alkanes.

A) Growth curve of non induced BL21 in minimal medium (open square), in minimal medium supplemented with C14 (open circle) and C16 (open triangle); *E. coli*-pBAD-CYP116B5 in minimal medium (filled circle) supplemented with C14 (filled square) and C16 (filled triangle) during the first 36 h.

B) Viable cell number per millilitre of culture when the bacteria were grown on minimal medium supplemented with the solid, long-chain C24 and C36 alkanes after 3 and 6 days, respectively. CTRL= control experiments bacteria grown on minimal medium without *n*-alkanes.

C) Growth of *E. coli* BL21 transformed with pET-CYP116B5 and pET-CYP116B5/pT7-Ar-BVMO on *n*-alkanes. Viable cell number per millilitre of cultures of bacteria were grown on minimal medium supplemented with the solid, long-chain C24 and C36 alkanes after 3 days. CTRL= control experiments, bacteria grown on minimal medium without *n*-alkanes.

Figure 6

Gas chromatogram following the incubation of CYP116B5 with tetradecane and NADH. Tetradecane R_T = 9.14 min.; tetradecanol R_T = 19.28 min.

Figure 7

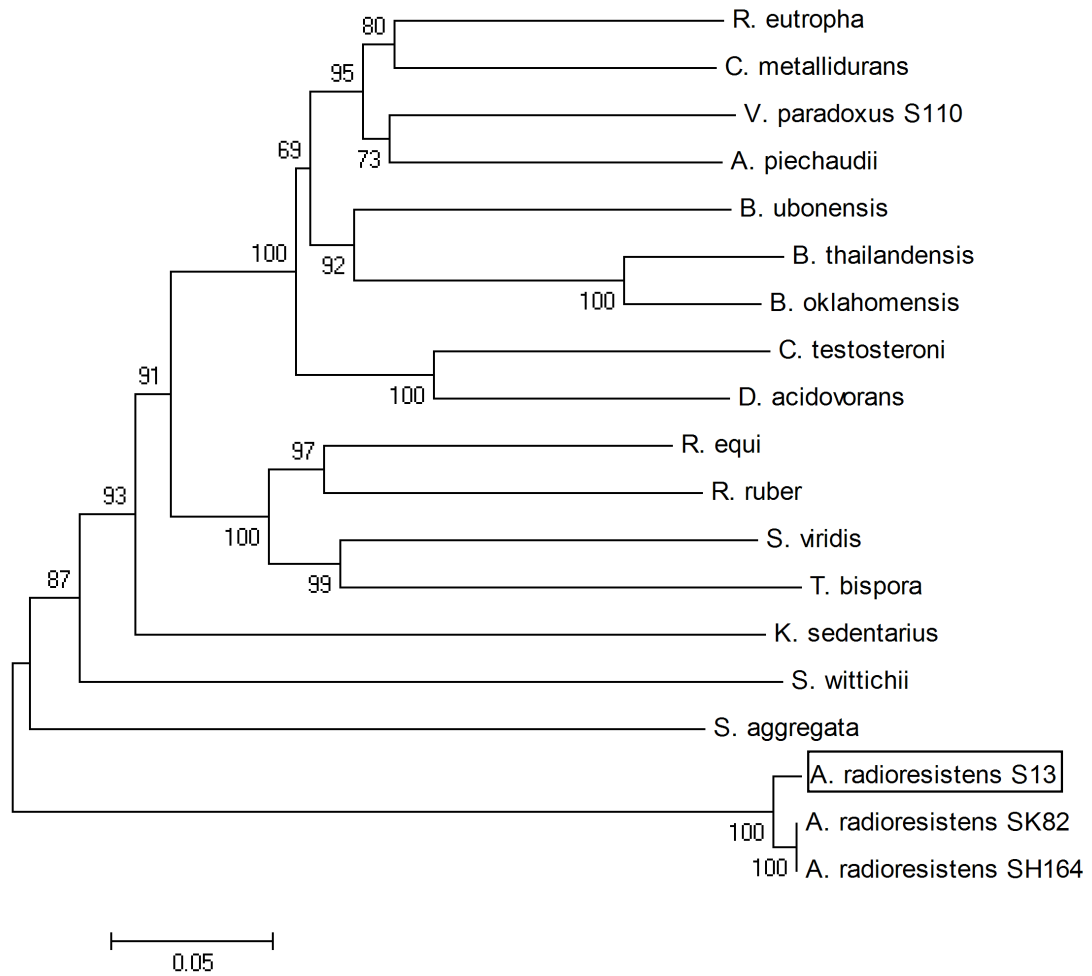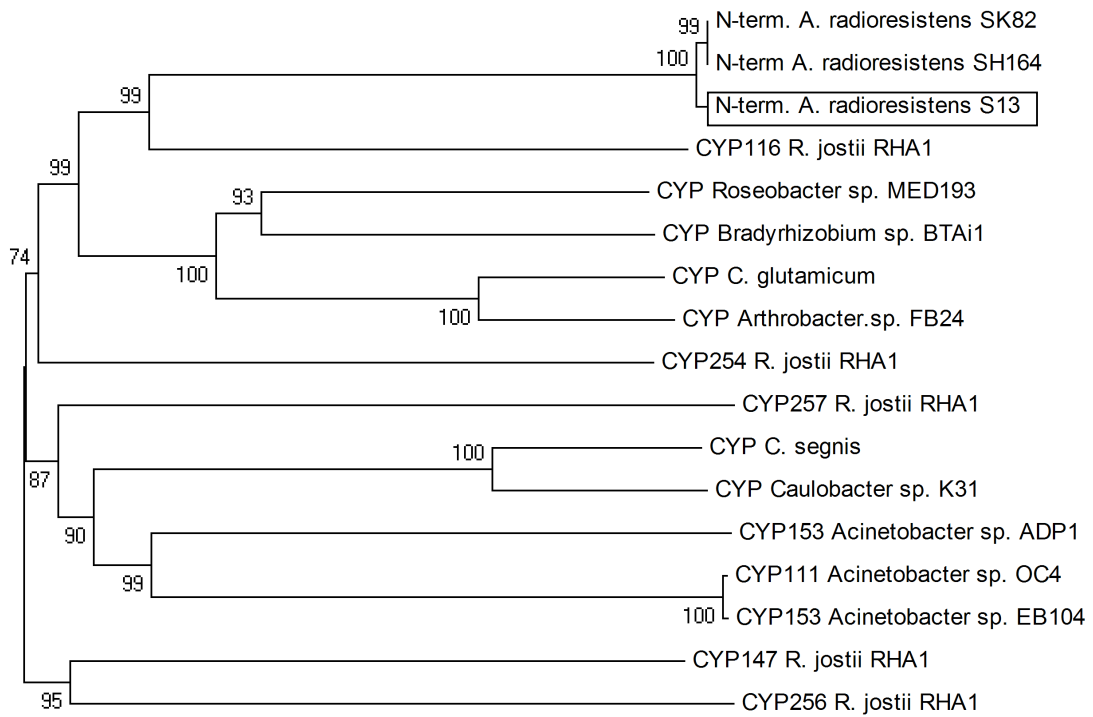Proposed pathway for alkane degradation in *A. radioresistens* S13.

FIGURE 1

```
                    10        20        30        40        50        60
                     |         |         |         |         |         |
RHERY      -------------MTEALIDPVVFNPYDYHFHDDPYPTYKRLREESPLYHNPDVGF---W
BACSUB     -------------------MTIASSTASSEFLKNPYSFYDTLRAVHPIYKGSFLKYP-GW
PSSP       --------MDARATIPEHIARTVILPQGYADDEVIYPAFKWLRDEQPLAMAHIEGYDPMW
MYCOTUB    --------------------------------------------SVFHAGEANS----
ACRAD      MNVKVPMHGQCSFHAELQQKGKNFDMFAVPYQQDPALALKEFRAQLPIFFSEAMGY---W

                    70        80        90       100       110       120
                     |         |         |         |         |         |
RHERY      ALSRYADVVDGFKDNKRLSSANGVSLDP---------AAYGPHAHYVMSFLAMDD-PRHM
BACSUB     YVTGYEETAAILKDAR-FKVRTPLPESS---------TKYQDLSHVQNQMMLFQNQPDHR
PSSP       IATKHADVMQIGKQPGLFSNAEGSEILYDQNNEAFMRSISGGCPHVIDSLTSMDP-PTHT
MYCOTUB    ------MLAGLLGDSSLLLIDDDVHRDR---------RRLMSPPFHRDAVARQAG--PIA
ACRAD      IVTRYEDVKAIFRDPITFSACNALEKLTPSC--PEALKILEKYKYGMNRTLVNEDEPVHM

                   130       140       150       160       170       180
                     |         |         |         |         |         |
RHERY      RLRQLVSRGFTPRRVAELDGRILELTKQHLIPAL-GAGEFDWITEVAGKLPMDVISELMG
BACSUB     RLRTLASGAFTPRTTESYQPYIIETVHHLLDQVQ-GKKKMEVISDFAFPLASFVIANIIG
PSSP       AYRGLTLNWFQPASIRKLEENIRRIAQASVQRLLDFDGECDFMTDCALYYPLHVVMTALG
MYCOTUB    EIAAANIAGWPMAKAFAVAPKMSEITLEVILRTV-----IGASDPVRLAALRKVMPRLLN
ACRAD      ERRRALMDAFTPQNLEEHQHFVRELVRKKVDGFI-YKGRADLVQEMLWEIPLMVALHFLG

                   190       200       210       220       230       240
                     |         |         |         |         |         |
RHERY      VPEPDRAELRRKADLVVHREAG-----------VLDVPQPAVEASISLMGYYKEMIAQRR
BACSUB     VPEEDREQLKEWAASLIQTIDFT---------RSRKALTEGNIMAVQAMAYFKELIQKRK
PSSP       VPEDDEPLMLKLTQDFFGVHEPDEQAVAAPRQSADEAAARRFHETIATFYDYFNGFTVDRR
MYCOTUB    VG--PWATLALANPSLLNNRLWS-----------R-LRRRIEEADALLYAEIADRRADPD
ACRAD      VPEDDMQELRKFAVAHTVNTWGR---------PTLEQQLEVAEGVGQFWEYSGRVLEKMK

                   250       260       270       280       290       300
                     |         |         |         |         |         |
RHERY      RSRTDD--LTSALLDAEIDGDKLSDQEILGFMFLMVVAGNETTTKLLGNALYWGSHHREQ
BACSUB     RHPQQD--MISMLLKGREK-DKLTEEEAASTCILLAIAGHETTVNLISNSVLCLLQHPEQ
PSSP       SCPKDD--VMSLLANSKLDGNYIDDKYINAYYVAIATAGHDTTSSSSGGAIIGLSRNPEQ
MYCOTUB    LAARTDT-LAMLVRAADEDGRTMTERELRDQLITLLVAGHDTTATGLSWALERLTRHPVT
ACRAD      NNPEGKGWMYDMIAKNRVMPEVVTDNYLHSMMMAIMVAAHETTALASANALKLLLADRKV
                                                    ‾‾‾‾‾‾‾‾‾

                   310       320       330       340       350       360
                     |         |         |         |         |         |
RHERY      IAPVLDDPER-----------ASEWVEETLRYDTSSQIVARTSVVDLEYHGRTIPAGEKV
BACSUB     LLKLRENPDL-----------IGTAVEECLRYESPTQMTARVASEDIDICGVTIRQGEQV
PSSP       LALALAKSDPAL-----------IPRLVDEAVRWTAPVKSFMRTALADTEVRGQNIKRGDRI
MYCOTUB    LAKAVQAADASAAGDPAGDEYLDAVAKETLRIRPVVYDVGRVLTEAVEVAGYRLPAGVMV
ACRAD      WKKICDNPQL-----------IPGAVEECLRHSGSVVAWRRQVTTESEVSGVKFRKGDKL

                   370       380       390       400       410       420
                     |         |         |         |         |         |
RHERY      LLLIGSANRDADVFDDADSFQIGRSSSGKLAS-----FGGGVHFCLGAHLAKLEAKIALA
BACSUB     YLLLGAANRDPSIFTNPDVFDITRSPNPHLS------FGHGHHVCLGSSLARLEAQIAIN
PSSP       MLSYPSANRDEEVFSNPDEFDITRFPNRHLG------FGWGAHMCLGQHLAKLEMKIFFE
MYCOTUB    VPAIGLVHASAQLYPDPERFDPDRMVGATLSPTTWLPFGGGNRRCLGATFAMVEMRVVLR
ACRAD      FLVSASANHDELHFENADELDIYRDNAIEHLT-----FGYGAHQCMGKNIGRMEMCIFIE
                                                      ≈≈≈≈≈≈≈*≈≈


RHERY      EFG
BACSUB     TLL
PSSP       ELL
MYCOTUB    EIL
```

FIGURE 2A

40

```
                    453         463         473         483         493         503
                     |           |           |           |           |           |
POBB     MSAAATMAPVSLRIHAIAYGADDVLLFDLRAPARDGLAPFDAGAHIDLRLPRGITRSYSL
VANB     --------MIEVIISAMRLVAQDIISLEFVRADGGLLPPVEAGAHVDVHLPGGLIRQYSL
ACRAD    ----LKAMTRPVRVQAVYPEAEDIIRIELAAIHGEELPRWSAGSHIELVLPNGLSRKYSL

                    513         523         533         543         553         563
                     |           |           |           |           |           |
POBB     LNDPAERHRYVIGVKREPESRGGSAWLHADARVGALIEVDGPSNHFALDESAPHAVFIAG
VANB     WNQPGVQSHYCIGVLKDPASRGGSKAVHENLRVGMRVQISEPRNLFPLEEGVERSLLFAG
ACRAD    CG-LATDQFYTIAVKREPVSRGGSQWIHQYLKAGEQIYIKGPKNFFKLNLQASQYVLIAG
                                                                        ==
                    573         583         593         603         613         623
                     |           |           |           |           |           |
POBB     GIGITPLWSMVQRLEHLGTPWTLHYRARSRRGAALLD--ELAGHGDRVHLSFSD-EGAPS
VANB     GIGITPILCMAQELAAREQDFELHYCARSADRAAFVEWLKVCDFADHVRFHFDNGPDQQK
ACRAD    GIGITPILSMANSLREQGRPYWLIYLSRQRASMALLK--EVAAHGSAAELYISS--EGKR
         ======
                    633         643         653         663         673         673
                     |           |           |           |           |           |
POBB     LDLAAIVAAAPEGAHFYCCGP-VPMLEAFEAAC-VGLDPARVHLEYFAAK---EAPATEG
VANB     LNAAALLAAEAEGTHLYVCGPGGFMGHVLDTAKEQGWADNRLHREYFAAAP--KVSADDG
ACRAD    IDLQQLLSALPAGTQVCACGP-EALLDTLTNYI-EDLSQVQLTVEHFGSGKNLFLYENDT

                    683         693         703         713         723         733
                     |           |           |           |           |           |
POBB     GFVVHLARSGRTIPIAAGCTILDALQAGGVAVPSSCQQGVCGICETAVLAGVPDHRDLVL
VANB     SFEVRIHSTGQVLQVPADQTVSQVLDEAGIIVPVSCEQGICGTCITRVVDGEPDHRDFFL
ACRAD    DFEVELLDSGLTLTVARDQTLLDCLLDKGIDVSFDCTEGLCGSCQLPVEEGEIDHRDKVL
                                                    *     *   *

                    743         753
                     |           |
POBB     SDQERAAGRTMMICCSGSKTAELTLDL
VANB     TDAEKAKNDQFTPCCSRAKSACLVLDL
ACRAD    TRAERDGMKSVISCCSRGK-GKLKLKL
                    *
```
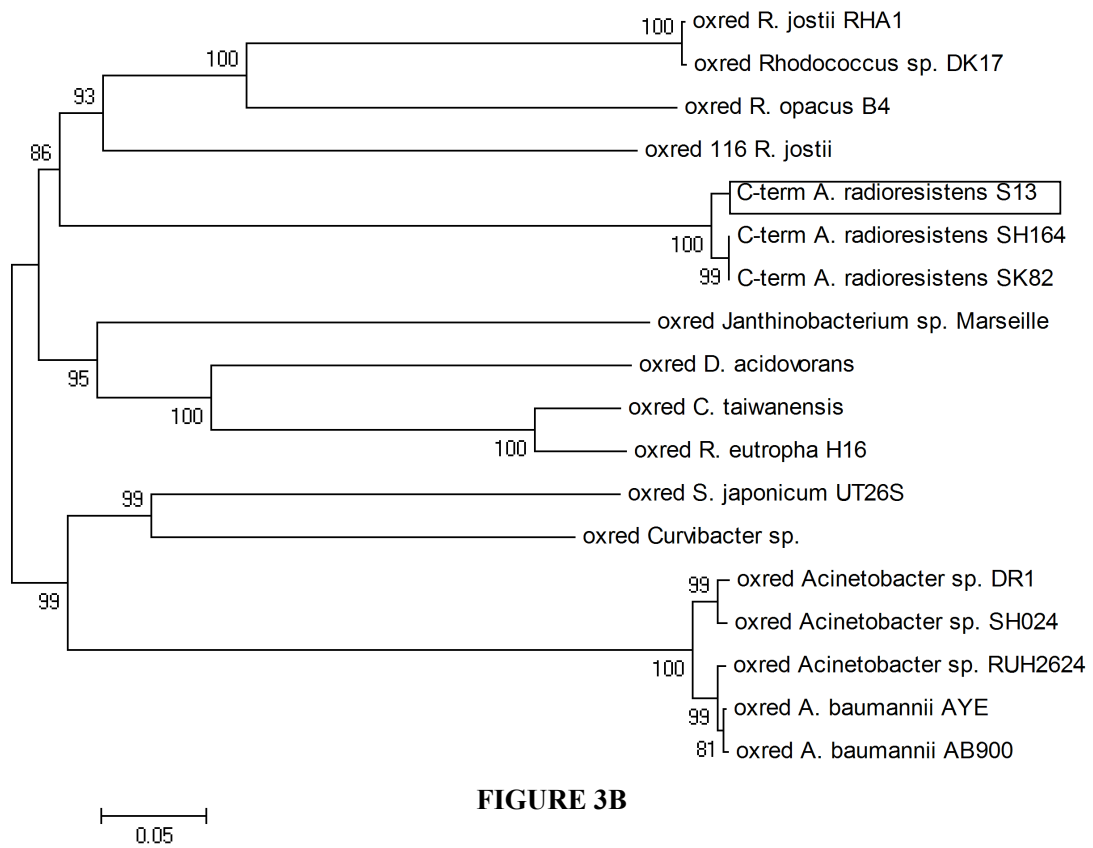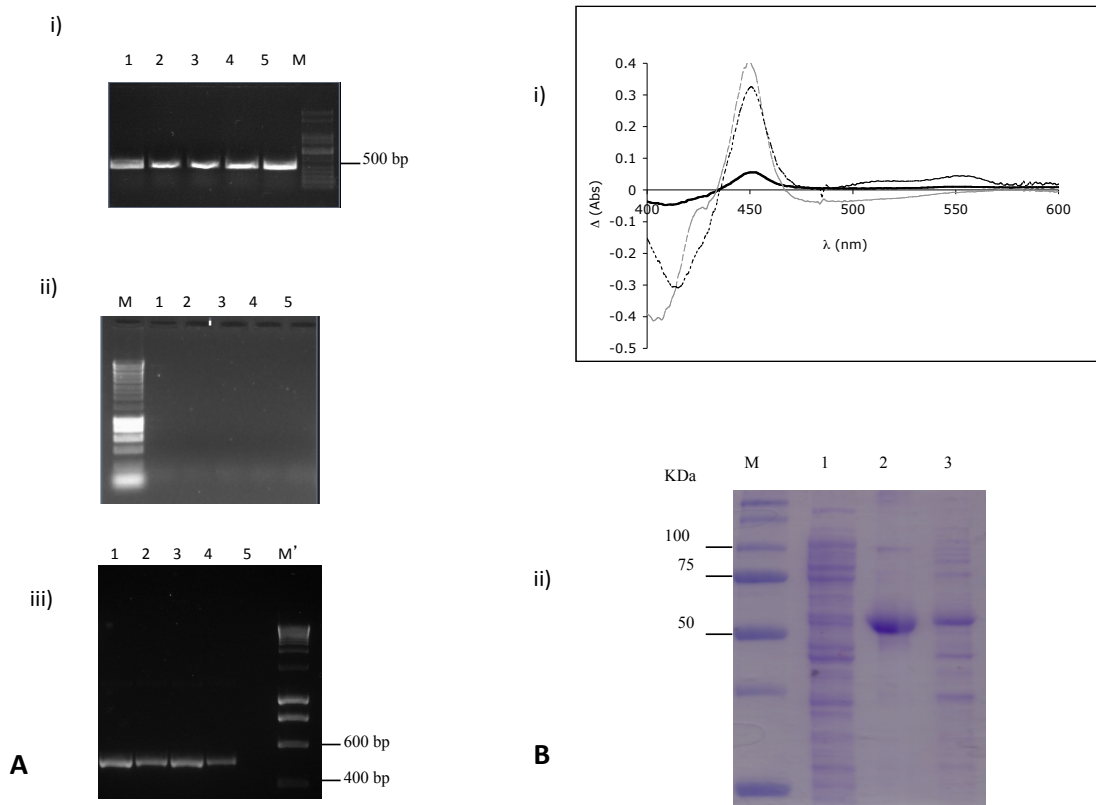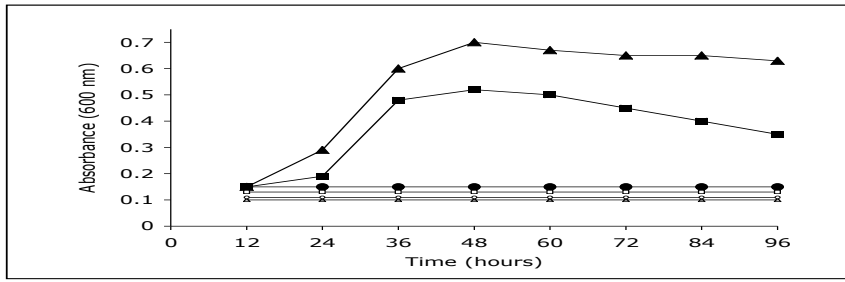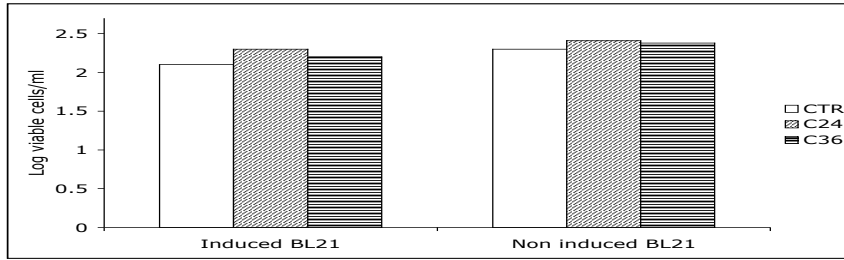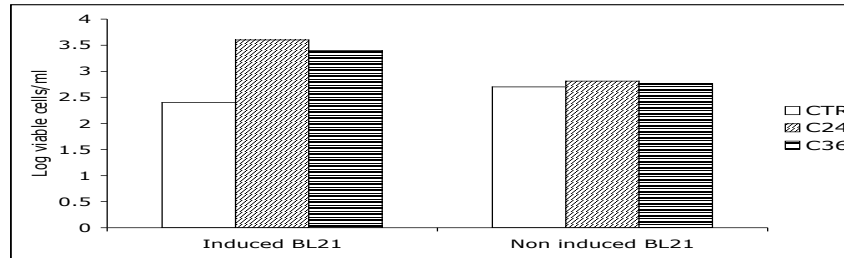
FIGURE 2B

FIGURE 3A

FIGURE 3B

0.05

A

B
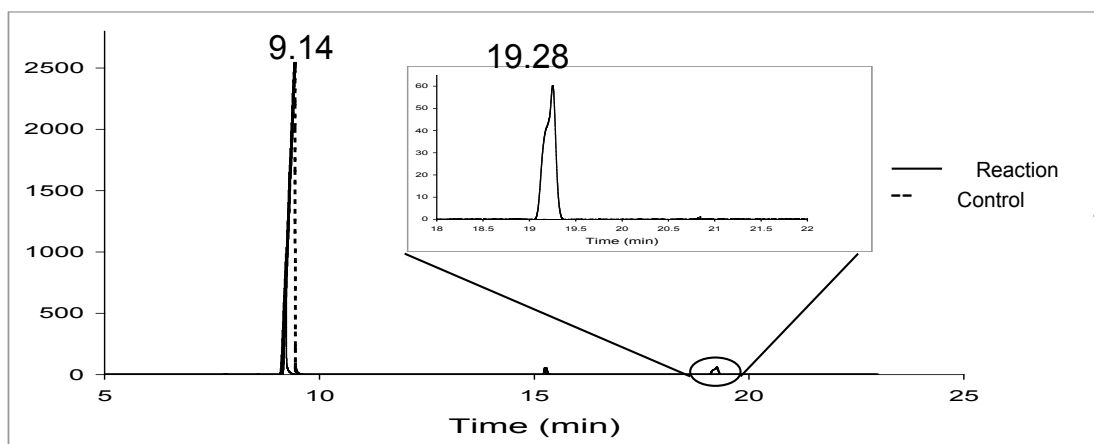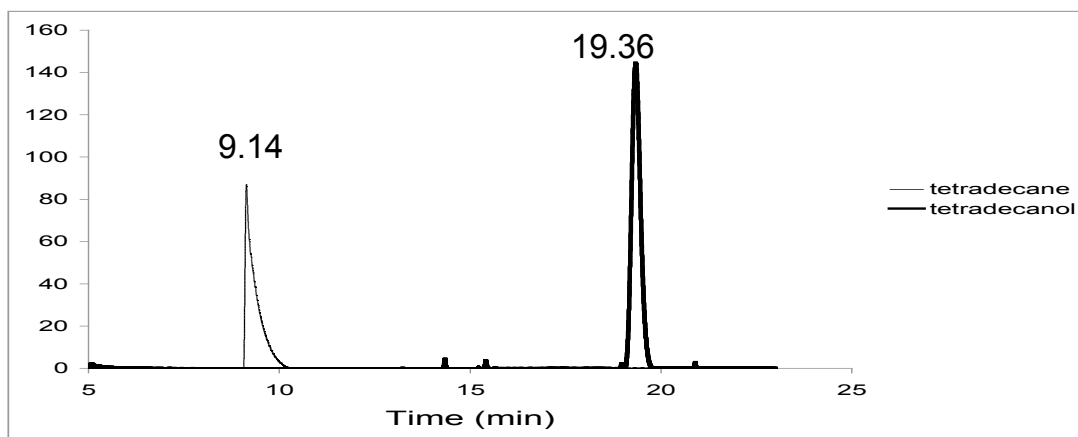
FIGURE 4

Figure 5

Figure 6

TERMINAL OXIDATION

CH_3-CH_2-$(CH_2)_n$-CH_2-CH_3

CYP1156B5   ALKB   CYP1156B5

CH_3-CH_2-$(CH_2)_n$-CH_2-CH_2-**OH**

Alcohol-DH ↓

CH_3-CH_2-$(CH_2)_n$-CH_2-CH_2-**CHO**

Aldehyde-DH ↓

CH_3-CH_2-$(CH_2)_n$-CH_2-CH_2-**COOH**

ω-oxidation ↓

HO-CH_2-CH_2-$(CH_2)_n$-CH_2-CH_2-**COOH**

Alcohol and Aldehyde-DH ↓

HOOC-CH_2-CH_2-$(CH_2)_n$-CH_2-CH_2-**COOH**

→ **β-OXIDATION**

| MEDIUM CHAIN ALKANES |

SUBTERMINAL OXIDATION

OH
|
CH_3-CH_2-$(CH_2)_n$-CH-CH_3

Alcohol-DH ↓   O
‖
CH_3-CH_2-$(CH_2)_n$-C-CH_3

Ar-BVMO ↓   O
‖
CH_3-CH_2-$(CH_2)_{n-1}$-CH_2-**O-C**-CH_3

Esterase ↓

CH_3-CH_2-$(CH_2)_{n-1}$-CH_2-**OH**   +   **HOOC**-CH_3

↓

TCA CYCLE

| LONG CHAIN ALKANES |

FIGURE 7