

IDENTIFICATION AND CHARACTERIZATION OF A NOVEL MONOOXYGENASE
FROM BURKHOLDERIA XENOVORANS LB400

By

Marie Carmen Montes-Matías

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

and

The Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Microbiology and Molecular Genetics

is written under the direction of

Dr. Gerben J. Zylstra

and approved by

New Brunswick, New Jersey

October, 2008

ABSTRACT OF THE DISSERTATION

Identification and Characterization of a Novel Monooxygenase from *B. xenovorans*

LB400

By MARIE CARMEN MONTES-MATIAS

Dissertation Director:
Dr. Gerben Zylstra

Whole genome sequences have been key elements in discovering new genes and in predicting protein function. A gene found in *Burkholderia xenovorans* LB400 encodes a protein that is 77% identical and 85% similar to the *p*-cymene monooxygenase from *Pseudomonas putida* F1. Phylogenetic analysis of this enzyme reveals that it is closely related to cymene, xylene and alkane monooxygenase. Protein alignment of these sequences identified 10 conserved histidine residues, of which 8 were reported to be essential for catalysis in alkane monooxygenase. The identified consensus sequence for these enzymes is [HX₍₃₎HX₍₂₅₎HX₍₃₎HHX₍₁₃₉₎HX₍₂₎HH], which is characteristic of membrane-bound hydroxylases with an active site that require iron and the activation of molecular oxygen for their catalytic cycle. In order to analyze the *B. xenovorans* LB400 monooxygenase enzyme in more detail, we cloned the cognate genes into the expression vector pQE.30 and performed biotransformation assays to determine the substrate specificity of the CymA1A2-like enzyme. The novel monooxygenase found in LB400 hydroxylates *p*-cymene to the product *p*-cymen-8-ol. The cymene monooxygenase found in *P. putida* F1 hydroxylated the methyl group of *p*-cymene, while the novel monooxygenase catalyzed the hydroxylation of the carbon of the isopropyl group

adjacent to the benzene ring. The enzyme also oxidized the substrates cumene, isobutylbenzene, *n*-butylbenzene, propylbenzene, 1, 4-diisopropylbenzene, *tert*-butyltoluene, *sec*-butylbenzene, and 4-ethyltoluene in addition to *p*-cymene. The collected data indicate that the LB400 monooxygenase described above is able to catalyze an interesting and novel reaction, but its physiological role is still unidentified. Steric hindrance and chain length definitely play an important role in determining the enzyme specificity. Metabolite analysis and RT-PCR data indicated that there is no correlation between the novel enzyme activity and growth of the strain on *p*-cymene, propylbenzene, 1,4-diisopropylbenzene, 4-ethyltoluene, *n*-butylbenzene, isobutylbenzene, *sec*-butylbenzene, *tert*-butylbenzene, and *tert*-butyltoluene. Other pathways and enzymes are involved in the degradation of these substrates. This data indicates that although bioinformatic analysis is a valuable tool in protein analysis and function prediction, it cannot always accurately predict the catalytic activity and physiological role of an enzyme.

Acknowledgements and Dedication

I would like to especially acknowledge my thesis advisor, Dr. Gerben Zylstra and my committee members, Dr. Bini, Dr. Chase and Dr. Kobayaishi for their guidance and help in the process of getting my degree. Dr. Zylstra, thank you for all the talks guiding my project, and thanks for believing in me. Without your help and guidance my dream of getting my philosophy doctorate could have never been accomplished. Thanks for been a great advisor and a great human being.

I will also like to thank all the Zylstra lab members for being excellent co-workers during all these years. Thanks for all the scientific help provided during my research. Thanks for always cheering me up and for believing. Sinèad, Karen, Sunny, Jon-Chang and Hung-Kuang, thanks for the chats and good moments spent together. Alan, Laurie and Michael, thanks for all the sequencing, ordering and your disposition to help. It was really appreciated!

I will also like to thank Dr. Haggblom and Dr. Ahn for training and lending me the GC-MS instrument used in this study. This instrument was essential for this research and without your help it would not be completed.

Thanks to the Pipeline Minority Program at Rutgers University and Robert Wood Johnson School of Medicine for financial support during my first two years as a graduate student. I would also like to mention Dr. Kathleen Scott and Mrs. Coletta for granting me the NSF fellowship “Building a Learning Community in Science and Mathematics through Educational Partnerships” that supported my work for two consecutive years. It was a great experience!

Last but not least, I would like to thank my beloved husband Orlando Navarro and my parents, Carmen Matías and Carlos Montes. Orlando, thank you for crying with me, laughing with me and worrying with me during all these years. You are my best friend and I wish someday I could support you the same way you did. Mami and papi, thanks a lot for your guidance, support and strength when times were difficult, but especially for your love.

I would like to especially dedicate my thesis to my family. To my dear daughter Sophia Alejandra, you were mommy's motivation and strength to get my degree. You are the most important person in my life and I love you with all my heart! To my husband Orlando, because no matter how difficult it was, you were always there supporting my dreams and walking with me to achieve them. To my dear parents and brother Carlos. You have always loved me and been there for me. To my dear nephew Carlos Andrés, because you are a very special person in my life. To my grandmother Mamita. You are an example of faith and perseverance. Thank you all for being supportive. I love you all!

Table of Contents

Abstract of the Dissertation	ii
Acknowledgements and Dedication	iv
1. Introduction.....	1
2. Literature Review.....	5
2.1 General Overview of Oxygenases.....	5
2.1.1 Environmental Importance of Oxygenases	9
2.1.2 Oxygenases as Industrial Catalysts	10
2.2 Iron-dependent Oxygenases.....	11
2.2.1 Heme-iron Dependent Oxygenases.....	12
2.2.2 Non-heme Iron-dependent Oxygenase.....	14
2.2.3 Oxygen Activation by Heme Containing Monooxygenases	32
2.2.4 Oxygen Activation by Non-heme Containing Monooxygenases.....	34
3. Materials and Methods.....	38
3.1 Bacterial Strains, Plasmids and Media.....	38
3.2 Chemicals and Reagents	38
3.3 Amplification of the Putative Cymene Monooxygenase	40
3.4 Cloning of the Putative Cymene Monooxygenase.....	42
3.5 Sequencing Analysis.....	43
3.6 Protein Expression and Biotransformation Assay.....	43
3.7 Extraction of the Biotransformed Products.....	44
3.8 Gas Chromatography-Mass Spectrometry Analysis	44
3.9 LB400 Growth Curves Using Different Carbon Sources	46
3.10 RNA Isolation	46
3.11 RT-PCR	47
4. Results and Discussion	49
4.1 Sequence and Phylogenetic Analysis of Putative Cymene Monooxygenase from <i>B. xenovorans</i> LB400	49
4.1.1 Comparative Analysis of the Genes Involved in the <i>p</i> -Cymene Degradation in <i>P. putida</i> F1 and the Genes Identified in <i>B. xenovorans</i> LB400	49
4.1.2 Phylogenetic Analysis of the Novel Monooxygenase from <i>B. xenovorans</i> LB400 ...	51
4.2 Cloning, Protein Expression and Substrate Specificity Analysis of a Novel Monooxygenase from <i>B. xenovorans</i> LB400	60

4.2.1	Amplification and Cloning Putative Cymene Monooxygenase Genes	60
4.2.2	Expression of <i>B. xenovorans</i> LB400 putative cymene monooxygenase	63
4.2.3	Substrate Specificity Analysis of the Putative Cymene Monooxygenase by GC-MS	66
4.3	Correlation of Monooxygenase Activity to Physiology of <i>B. xenovorans</i> LB400	93
4.3.1	Analysis of Different Carbon Sources that Support Growth of <i>B. xenovorans</i> LB400	94
4.3.2	Metabolite Analysis of Cultures Utilizing Aromatic Hydrocarbons as Their Carbon Source	102
4.3.3	Detection of Genes Expressed by Cell Cultures Grown on Aromatic Hydrocarbon Sources	112
5.	Conclusions.....	116
6.	Future Directions	118
7.	References.....	120
	Curriculum Vita	128

List of Figures

Figure 1.1: Representation of central aromatic pathways in LB400	3
Figure 2.1: Representation of the different classes of oxidoreductases.....	6
Figure 2.2: Examples of biotransformations catalyzed by different oxygenases	8
Figure 2.3: Schematic representation of membrane bound non-heme diiron monooxygenases	16
Figure 2.4: Degradation pathway of <i>n</i> -alkanes by <i>P. putida</i> Gp1	18
Figure 2.5: Degradation of <i>m</i> -xylene by <i>P. putida</i> mt-2 and <i>S. yanoikuyae</i> B1	22
Figure 2.6: Degradation pathway of <i>p</i> -cymene by <i>P. putida</i> F1	26
Figure 2.7: Chemical structures of substrates utilized for the substrate specificity analysis of the cymene monooxygenase from <i>P. putida</i> F1	31
Figure 2.8: Intermediate species formed in the reaction of P450 monooxygenases.....	33
Figure 2.9: The catalytic cycle of soluble methane monooxygenase	35
Figure 3.1: Methodology for protein expression and substrate specificity analysis.....	45
Figure 4.1: Gene organization around the cymene monooxygenase of <i>P. putida</i> F1 and the putative cymene monooxygenase from <i>B. xenovorans</i> LB400.....	52
Figure 4.2: Phylogenetic tree of selected alkane and aromatic hydrocarbon hydroxylases	54
Figure 4.3: Protein sequence alignment of the xylene, alkane, cymene and putative cymene monooxygenase component	1
Figure 4.5: Analysis of enzyme activity using different expression vectors	64
Figure 4.6: Oxidation of <i>p</i> -cymene catalyzed by the putative cymene monooxygenase from <i>B. xenovorans</i> LB400.....	67
Figure 4.7: Gas chromatograms and mass spectrum comparisons of standards <i>p</i> -cumeic alcohol and <i>p</i> -cymen-8-ol	69
Figure 4.8: Proposed mechanism for the degradation of <i>p</i> -cymene by the novel monooxygenase.....	70
Figure 4.9: Experimental controls used in the GC-MS analysis.....	72

Figure 4.10: Chemical structures of substrates utilized in the substrate specificity analysis of <i>B. xenovorans</i> LB400	73
Figure 4.11: Oxidation of cumene by the novel monooxygenase	75
Figure 4.12: Oxidation of 1,4-diisopropylbenzene by the novel monooxygenase	76
Figure 4.13: Oxidation of propylbenzene by the novel monooxygenase	78
Figure 4.14: Oxidation of <i>n</i> -butylbenzene by the novel monooxygenase	79
Figure 4.15: Oxidation of isobutylbenzene by the novel monooxygenase	81
Figure 4.16: Oxidation of <i>sec</i> -butylbenzene by the novel monooxygenase	82
Figure 4.17: Oxidation activity of 4-ethyltoluene by the novel monooxygenase	85
Figure 4.18: Oxidation activity of <i>tert</i> -butyltoluene by the novel monooxygenase	87
Figure 4.19: Gas chromatogram and mass spectrum comparisons between the standard 4-ethylbenzyl alcohol and the biotransformation product the substrate <i>tert</i> -butyltoluene...	89
Figure 4.20: Growth curves of LB400 utilizing <i>p</i> -cymene, isobutylbenzene and 4-ethyltoluene sole carbon source	96
Figure 4.21: Growth curves of LB400 utilizing 1,4-diisopropylbenzene and <i>tert</i> -butylbenzene as sole carbon source	98
Figure 4.22: Growth curves of LB400 utilizing <i>n</i> -butylbenzene and <i>sec</i> -butylbenzene as sole carbon source	99
Figure 4.23: Growth curves of LB400 utilizing <i>tert</i> -butyltoluene and propylbenzene as sole carbon source	101
Figure 4.24: Analysis of metabolites produced by LB400 cells growing on <i>p</i> -cymene	103
Figure 4.25: Analysis of metabolites produced by LB400 cells growing on propylbenzene	105
Figure 4.26: Analysis of metabolites produced by LB400 cells growing on isobutylbenzene	106
Figure 4.27: Analysis of metabolites produced by LB400 cells growing on 4-ethyltoluene	108
Figure 4.28: RT-PCR results for cell cultures grown on different hydrocarbons	113

Figure 4.29: Proposed pathway for the degradation of <i>tert</i> -butylbenzene by B. xenovorans LB400	115
--	-----

List of Tables

Table 2.1: List of chemicals and degradation rates by cymene monooxygenase from <i>P. putida</i> F1	30
Table 3.1: Strains and plasmids used in this study	39
Table 3.2: List of primers used in this study.....	1
Table 4.1: Proteins containing conserved histidine motifs common to membrane desaturases, hydroxylases and related proteins.....	56
Table 4.2: Summary of substrate specificity analysis indicating the product and its retention time measured by GC-MS	91
Table 4.3: Comparison between substrate specificity analysis of the putative cymene monooxygenase and growth physiology of LB400 on identified substrates	110

1. Introduction

Hydrocarbons are energy-rich compounds abundantly distributed on Earth. Microorganisms capable of utilizing these compounds as the only energy source can be found almost everywhere, even in non-contaminated sites (van Beilen 1994; van Beilen *et al.* 1994). For the past decades, research scientists have been focusing on studying the metabolism of these microorganisms and the oxyfunctionalization reactions that lead to the initial step in the degradation of hydrocarbons. More importantly, research has been focused on studying the genes/proteins involved in the activation of these compounds leading to their degradation. The oxyfunctionalization of hydrocarbons represents a very promising tool in the bioremediation of contaminated sites and in the production of chemicals and pharmaceutical products that could be too difficult or expensive to synthesize by chemical means.

Bacteria from the *Burkholderia* genus are very diverse and its species can be found in a wide variety of ecological niches such as soil, plants, animals, and humans among others. *Burkholderia xenovorans* LB400 is one of the most important species in the degradation of polychlorinated biphenyls described to date. It was first isolated from a PCB-containing landfill in upper New York state (Bopp 1986; Chain *et al.* 2006) and today, it represents the model organism for PCB degradation. It can oxidize more than 20 different PCB congeners with up to six chlorine substitutions on the biphenyl ring (Bedard *et al.* 1986; Seeger *et al.* 1995; Seeger *et al.* 1999; Chain *et al.* 2006). Tiedje and colleagues have recently published the complete genome sequence of *B. xenovorans* LB400 (Chain *et al.* 2006). Its genome size is 9.73 Mbp, and it harbors approximately 9,000 coding sequences distributed over three circular replicons, chromosome 1 (4.90

Mbp), chromosome 2 (3.36 Mbp) and a megaplasmid (1.47 Mbp) (Chain *et al.* 2006). This bacterium was the first non-pathogenic *Burkholderia* strain whose complete genome sequence became available.

Based on sequence analysis we have identified a putative cymene monooxygenase enzyme in *B. xenovorans* LB400 located in chromosome 1 of this strain. Its oxygenase component is 77% identical and 85% similar to the cymene monooxygenase from *P. putida* F1. Cymene monooxygenase from *P. putida* F1 has been already characterized (Eaton 1997) and its substrate specificity has been analyzed and described (Nishio *et al.* 2001). This enzyme catalyzes the oxidation of *p*-cymene to *p*-cymic alcohol (in addition to the catalysis of other chemicals), the initial step for the degradation of *p*-cymene to *p*-cumate, which is further metabolized to pyruvate, acetyl CoA and isobutyrate, intermediates of the aerobic respiratory chain (Figure 1.1). So far, the function and specificity of the putative cymene monooxygenase found in *B. xenovorans* LB400 enzyme has only been predicted by bioinformatics analysis based on sequence identity/similarity. Cloning and expression of this putative cymene monooxygenase is essential to confirm a relationship between sequence similarity and function. A schematic representation of the central aromatic pathways in *B. xenovorans* LB400 published by Tiedje and coworkers proposed that *p*-cymene is oxidized to its alcohol form *p*-cymic alcohol, and then to *p*-cumate, which can be further metabolized to isobutyrate, an intermediate of the respiratory pathway (Figure 1.1). Due to the great interest generated on the area of biotransformation for environmental and industrial purposes, this thesis research focuses on the identification and substrate specificity analysis of the putative cymene monooxygenase from *B. xenovorans* LB400. The goal of

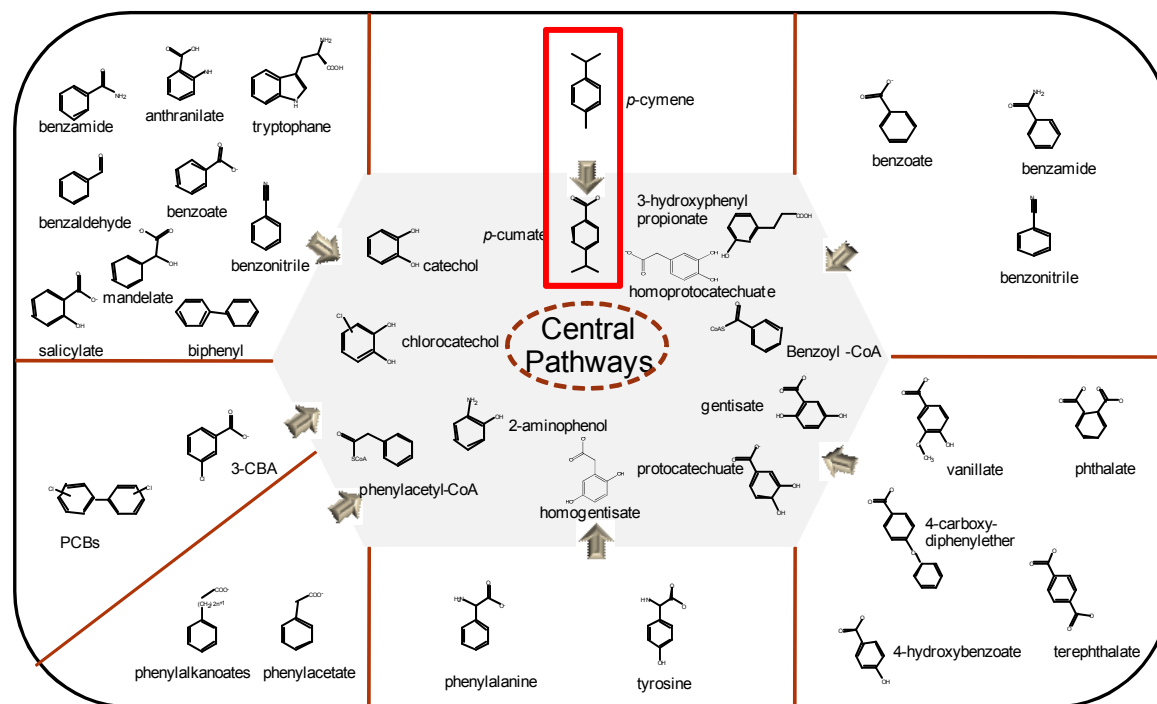


Figure 1.1: Representation of central aromatic pathways in LB400

The peripheral aromatic pathways are represented in the outer sectors. The boxed region represents the proposed pathway for the degradation of *p*-cymene by the putative cymene monooxygenase from LB400.

Figure adapted from Chain *et al.* (2006)

this thesis was to clone and characterize the putative cymene monooxygenase from this strain and compare it to the cymene monooxygenase enzyme from *P. putida* F1. In addition to identify the function of this enzyme, these findings will help to elucidate if function prediction by bioinformatics analysis can always be a reliable tool for protein function prediction.

2. Literature Review

2.1 General Overview of Oxygenases

Oxygenases specifically introduce one or both atoms of molecular oxygen into specific organic or inorganic substrates, and function in catabolic and detoxification reactions in both prokaryotic and eukaryotic organisms (Harayama *et al.* 1992; Leahy *et al.* 2003). They belong to the oxidoreductase class of enzymes, a group of enzymes that catalyze the transfer of electrons from the oxidant to the reductant. This group also includes the dehydrogenases, oxidases and peroxidases, in addition to the oxygenases. Figure 2.1 shows a tree where different classes of oxidoreductases are illustrated, with emphasis on the oxygenases (especially non-heme diiron monooxygenases) which are the research focus of this work. As shown in this Figure 2.1, the main classes of oxygenases are FAD- (which catalyze epoxidations and aromatic oxygenations), copper- (which catalyze the C-H oxifunctionalizations) and the iron-dependent (the most abundant group that catalyzes the C-H oxifunctionalizations) oxygenases (Bühler *et al.* 2000). Oxygenases can be subdivided into monooxygenases and dioxygenases, by whether they add one or two atoms of molecular oxygen into the substrate. Monooxygenases catalyze the insertion of a single oxygen atom into the substrate and require two electrons and two protons to cleave molecular oxygen, taking one atom for substrate oxidation and the second atom to produce a water molecule (Sligar *et al.* 2005). This group is often known as “mixed-function oxygenases” because they catalyze the addition of an oxygen atom into the substrate, and at the same time they act as an oxidase by producing a water molecule in the process. On the other hand, dioxygenases catalyze the incorporation of

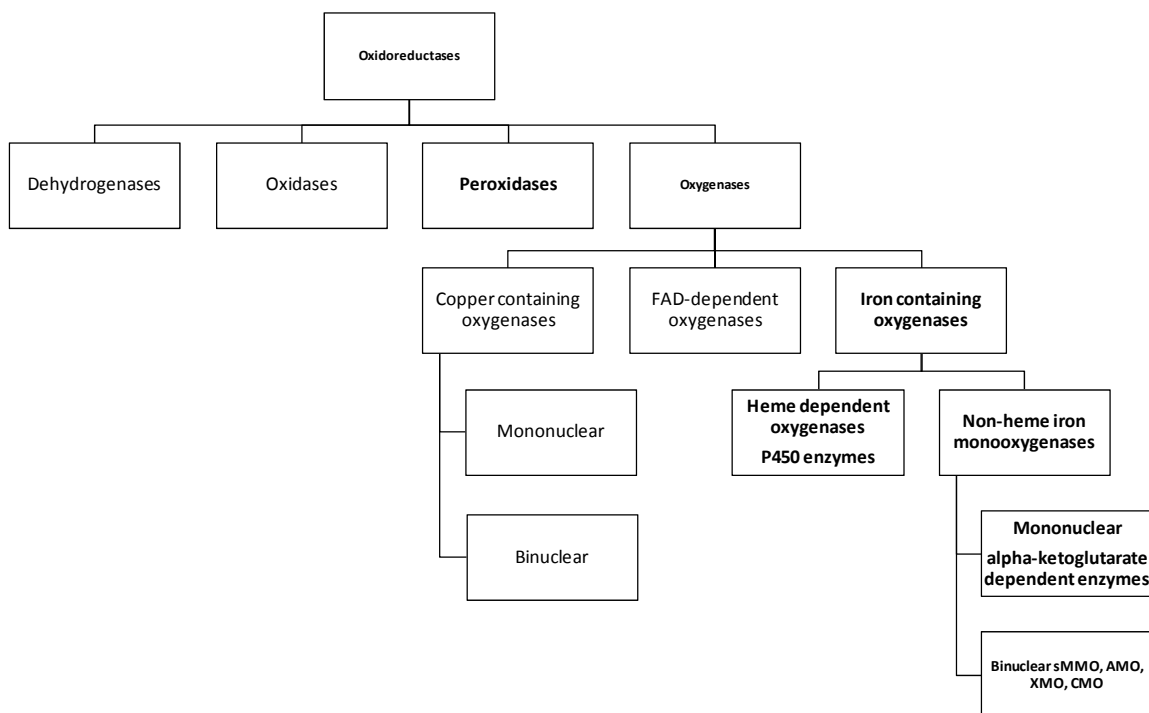


Figure 2.1: Representation of the different classes of oxidoreductases

The diagram exemplifies the different types of oxygenases, including the non-heme binuclear monooxygenase group, which includes cymene monooxygenase, soluble methane monooxygenase, alkane monooxygenase and xylene monooxygenase.

Adapted from Bühler (2003)

two oxygen atoms into the substrate(s) and require two electrons and two protons to cleave dioxygen. Figure 2.2 shows different examples of biotransformations catalyzed by different oxygenases including epoxidations, oxidations and mono- and dioxygenations reactions. Among the many interesting reactions performed by oxygenases, they catalyze the selective oxyfunctionalization of hydrocarbons (Figure 2.2E) reaction that is of great interest in the areas of industrial and environmental microbiology and is the focus of this research project.

Mono- and dioxygenases usually require the presence of cofactors, which can often be a transition metal, in order to overcome the kinetic barriers that make difficult the activation of molecular oxygen and the substrate. Enzymes that do not require cofactors have been already identified and described (Fetzner 2002), but these enzymes will not be discussed. The most common metals employed by the oxygenases are copper and iron because they can form complexes with dioxygen and/or substrate(s) affecting oxygen/substrate electronic structure and reactivity when they are in their lower oxidation state (Fetzner 2002). The electron transport chain associated with oxygenases contains at least two redox centers. The first redox center is usually a flavin and the second an iron-sulfur cluster (Harayama *et al.* 1992). The electron transport is initiated by a single two-electron transfer from NAD(P)H to a flavin, followed by two single-electron transfers from the flavin to an iron-sulfur cluster (Harayama *et al.* 1992). Our research focuses on the analysis of a non-heme diiron enzyme found in *B. xenovorans* LB400 with sequence similarity to the non-heme diiron cymene monooxygenase found in *P. putida* F1.

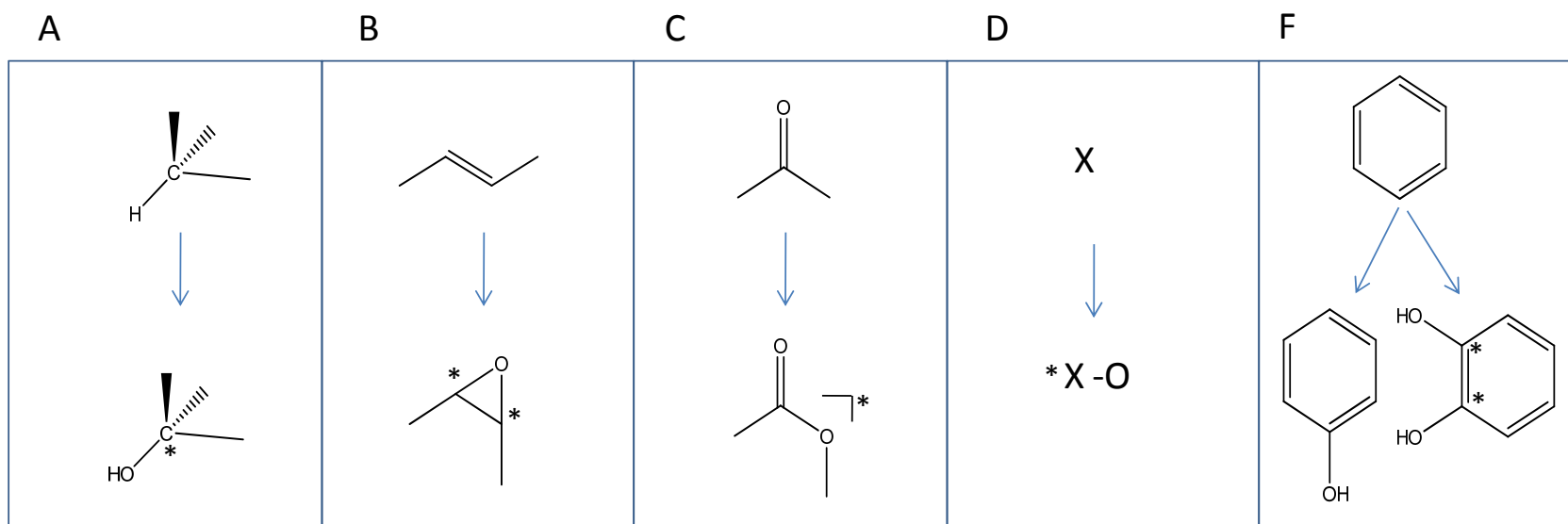


Figure 2.2: Examples of biotransformations catalyzed by different oxygenases

A. Representation of the C-H activation at sp^3 -hybridized carbon atoms

B. Vinyl group epoxidation

C. Bayer-Villiger oxidation

D. Heteroatom oxidation

E. Aromatic mono- and dioxygenation

Figure adapted from Bühler (2003)

2.1.1 Environmental Importance of Oxygenases

Generally, oxygenases are of great importance in nature. Their role can range from the hydroxylation of steroids and the synthesis of hormones in mammals to the degradation of toxic compounds by hydroxylation by microorganisms, where the hydroxylated product can be further metabolized to generate products that are directed to the respiratory pathways. The study and identification of bacterial species capable of degrading hydrocarbons has generated great interest, because these compounds can be found worldwide as environmental pollutants that could represent an issue for the environment and human health. Bioremediation represents a very promising tool for the decontamination of the environment because conventional remediation processes can be very expensive and cause further environmental concerns because of the dispersion of the chemical in the environment (Hilyard *et al.* 2007).

In contrast to higher organisms which can utilize a very limited number of carbon sources, organisms such as bacteria, archaea and fungi can utilize a wide variety of compounds as their source of carbon and energy. It has been found that many bacterial strains are capable of recycling fixed organic carbons to produce carbon dioxide and biomass that is no longer detrimental to the environment. By the addition of molecular oxygen, mono- and dioxygenases oxidize hydrocarbons to their respective alcohol forms, which are further metabolized by additional enzymes to compounds that are no longer harmful to the environment. This activity represents a useful environmental tool in the biodegradation of many chemicals of (Chain *et al.* 2006) both natural and anthropogenic origin that contaminate the environment and whose disposal is frequently very difficult, hazardous, and costly. The understanding of the metabolic processes and bacterial

responses involved in these reactions is essential in order to find effective strategies for the oxidation of hydrocarbons and their biotransformation to non-toxic compounds. Unfortunately, further research needs to be done in order to achieve biotransformation of these hydrocarbons at a larger scale significant enough to address the problem. The main focus of this research is the study of the hydroxylation of hydrocarbons from a novel monooxygenase identified in the very versatile hydrocarbon degrading species *B. xenovorans* LB400.

2.1.2 Oxygenases as Industrial Catalysts

In addition to their essential role in nature, the enzymatic activity of oxygenases plays an important role in industry by catalyzing the production of chemicals, polymers and pharmaceutical building blocks. The industrial use of these enzymes could represent a reduction in both the amount of waste produced and the emissions and the synthesis of undesirable by-products whose disposal can be very expensive and inefficient. Oxygenases are important to the chemical industry for three reasons: (1) Their non-biological counterparts either do not exist or do not have the required regioselectivity and stereoselectivity, (2) they use oxygen as a cheap, environmentally friendly oxidant in contrast to toxic chemical oxidants, and (3) they can be used to prepare chiral building blocks and pharmaceutical intermediates or to modify natural products with biological activities (Li *et al.* 2002). Although different enzymes can catalyze the incorporation of oxygen into a substrate, oxygenases can catalyze this reaction in a stereo- and regioselective way, thus representing a very promising tool in the area of synthetic organic chemistry. For this reason, research has focused on this group of enzymes and their industrial exploitation for many years. Examples of the industrial exploitation of

cytochrome P450 monooxygenases include the hydroxylation of Reichstein S, a progesterone derivative which leads to corticosteroids with anti-inflammatory activity by the fungus *Curvularia* (Urlacher *et al.* 2004). Also, the fungus *Pleurotus ostreatus* is capable of metabolizing phenanthrene to phenanthrene-trans-9,10-dihydrodiol and 2,2-diphenic acid as well as mineralizing it to CO₂ as part of the fermentation process (Bezalel *et al.* 1996; Urlacher *et al.* 2004). Today, 2% of the chemicals and polymers available in the market are synthesized by bioprocesses (Bühler *et al.* 2000). Unfortunately, the use of oxygenases at an industrial scale is often difficult and lot of research still needs to be conducted. The main problems that limit their use include the fact that these enzymes usually consist of multiple components, require expensive cofactors such as NAD(P)H and are often membrane-associated (van Beilen *et al.* 2003).

2.2 Iron-dependent Oxygenases

Proteins that require iron as a cofactor that are involved in the molecular oxygen reactions require the redox properties of metals for catalysis (Nordlund and Eklund 1995). They are usually involved in the oxidation of inactive organic molecules such as alkanes, olefins, arenes and heteroatoms (Fang *et al.* 1995). These enzyme systems are known to activate molecular oxygen by a sequential, two-electron reduction of oxygen to the formal oxidation state of hydrogen peroxide (Groves 2006). Di-metal proteins contain the unifying feature that when the iron ions are in their most stable oxidation state, they are bridged by carboxylates and oxide or hydroxide ions (Nordlund and Eklund 1995). Diiron enzymes can be subdivided into heme and non-heme enzymes, and although they have structural differences, it is thought that their mechanisms for

molecular oxygen activation are very similar. Diiron enzymes can be soluble or membrane-bound. Although the diiron centers of all the enzymes belonging to this category have not been structurally characterized yet, all the diiron centers described so far contain at least one bridging carboxylate ligand, suggesting that similar carboxylate bridges may be involved in the diiron centers of integral membrane histidine motif-containing enzymes (Shanklin and Whittle 2003). Different classes of diiron proteins that catalyze oxygen-mediated reactions have been described. Class I contains the heme soluble diiron proteins hemerythrin and myohemerythrin where the diiron center is coordinated via nitrogen ligands (Shanklin and Whittle 2003). Class II is composed of soluble methane monooxygenase, ribonucleotide reductase and the soluble plant Δ^9 desaturase in which the diiron center is coordinated by oxygen ligands (Fox *et al.* 1994; Wallar and Lipscomb 1996; Kurtz 1997; Shanklin and Whittle 2003). Since the analysis of the membrane-bound non-heme diiron enzyme AlkB show significant differences from the class I and class II of diiron proteins, it is suggested that there exist a Class III group of enzymes that differs from the other two classes already described (Shanklin and Whittle 2003). It has been proposed that members of this new class of diiron oxygenases are: alkane monooxygenase, xylene monooxygenase, cymene monooxygenase and the soluble methane monooxygenase.

2.2.1 Heme-iron Dependent Oxygenases

Heme-dependent monooxygenases contain a protoporphyrin IX tetrapyrrole ring with an iron nucleus, known as the heme group. One of the most characterized enzymes belonging to this group is the metalloenzyme cytochrome P450. The iron nucleus of P450 enzymes is coordinated by four nitrogen atoms, which are important for enzymatic

catalysis (Urlacher *et al.* 2004). This group of enzymes catalyze the oxidation of organic substrates that are generally hydrophobic and are mostly involved in biologically important processes, such as bile acid and steroid hormone synthesis, activation and deactivation of drugs, catabolite assimilation, xenobiotic detoxification and fatty acid hydroxylation (Carrondo *et al.* 2007). Cytochrome P450 contains an iron-protoporphyrin IX center coordinated to a cysteine thiolate capable of activating molecular oxygen at the iron center to incorporate one oxygen atom into its biological substrate (Groves 2006).

2.2.1.1 Cytochrome P450

Cytochrome P450 enzymes are capable of introducing atomic oxygen into allylic positions, double bonds, or even into non-activated C-H bonds (Urlacher *et al.* 2004). P450 enzymes ensure detoxification of exogenous compounds by rendering these mostly lipophilic compounds water-soluble to facilitate excretion (Bühler 2003). The catalytic mechanism for oxygen activation by P450 monooxygenases can be summarized into six steps: (1) The substrate is bound and a water molecule is displaced. (2) The ferric enzyme is reduced to a ferrous state by a one-electron transfer. (3) Oxygen is bound, resulting in a ferrous-dioxy species. (4) A second reduction occurs, followed by a proton transfer leading to an iron-hydroperoxo intermediate. (5) The O-O bond is cleaved and it releases water and activated iron-oxo ferryl species. (6) The iron-oxo ferryl species oxidizes the substrate and the product is subsequently released (Urlacher *et al.* 2004). It has been proposed that this enzyme initiate its chemistry by the oxidation of a resting Fe^{III} state to a reactive oxo- Fe^{IV} -porphyrin cation radical intermediate (Groves 2006).

P450 monooxygenase enzymes are present in many organisms and can be divided into four classes: Class I require a FAD-containing reductase and iron sulfur redoxin.

Class II requires FAD/FMN-containing P450 reductase for the transfer of electrons (they are self-sufficient enzymes containing the P450 monooxygenase and the FAD/FMN-reductase domain on a single peptide chain). Class III contains enzymes that do not require electron donors because they are self-sufficient to convert peroxygenated substrates that already contain oxygen. Finally, Class IV, which receives electrons from NAD(P)H (Urlacher *et al.* 2004). Most of the cytochrome P450 enzymes described so far require additional proteins (multi-component enzymes) for the transport of reducing equivalents from NAD(P)H to the P450 oxygenase component. An exception is cytochrome P450 BM-3 of *Bacillus megaterium* which consists of a single polypeptide with a P450 domain and an electron transport domain of the microsomal type (Ravichandran *et al.* 1993; Munro *et al.* 2002; Bühler 2003). Crystal structures of different P450 enzymes are already available facilitating the study of the structure, function and mechanisms catalyzed by these enzymes.

2.2.2 Non-heme Iron-dependent Oxygenase

In contrast to the heme iron-dependent proteins, non-heme iron-dependent proteins do not contain the heme active site for oxygen/substrate activation. Although models for oxygen activation are based on heme-containing systems, non-heme diiron enzymes can also mediate oxygen activation and transfer into the substrate. Non-heme diiron oxygenases are subdivided into mononuclear and binuclear (Figure 2.1). The mononuclear α -ketoglutarate-dependent dioxygenases catalyze the hydroxylation of its substrate and, at the same time, the decarboxylation of α -ketoglutarate to succinate and CO_2 (Solomon *et al.* 2003; Hoffart *et al.* 2006). These enzymes are present in

microorganisms and higher organisms and their oxygenation reactions are of interest for the synthesis of pharmaceuticals and other chemicals.

Binuclear non-heme iron oxygenases usually contain a bridged diiron center in the active site that, among other reactions, can catalyze the oxyfunctionalization of hydrocarbons. Figure 2.3 shows the diiron non-heme alkane, xylene, cymene and putative cymene monooxygenase systems including all their components. The active sites of non-heme diiron enzymes usually contain two histidines, one monodentate carboxylate and two to three ligands (Hegg and Que 1997; Solomon *et al.* 2003). Proteins belonging to this group of enzymes are methane monooxygenase (sMMO), alkane monooxygenase (AlkB), xylene monooxygenase (XylM) and toluene monooxygenase (TMO) (Bertrand *et al.* 2005). Most of the non-heme diiron oxygenases require the Fe^{II} active site for the activation of dioxygen (Solomon *et al.* 2003). Two different types of non-heme diiron active sites have been reported (Shan and Que 2006). The first type consists of the mononuclear iron center that is coordinated to two histidines and a carboxylate that occupy one face of an octahedron (2-His-1-carboxylate facial triad) (Shan and Que 2006). This motif confers mechanistic flexibility, as it allows three coordination sites on the metal center to be available for binding exogenous ligands such as substrate, cofactor, and/or oxygen (Shan and Que 2006). The second type of active site is a diiron center with two histidines and four carboxylates as ligands and is associated with the methane and toluene monooxygenases, fatty acid desaturases and ribonucleotide reductase (Shan and Que 2006). At least one coordination site on each iron is available for exogenous ligand binding and there is evidence that oxygen binds to both iron centers in the course of catalysis (Shan and Que 2006). The mechanism for

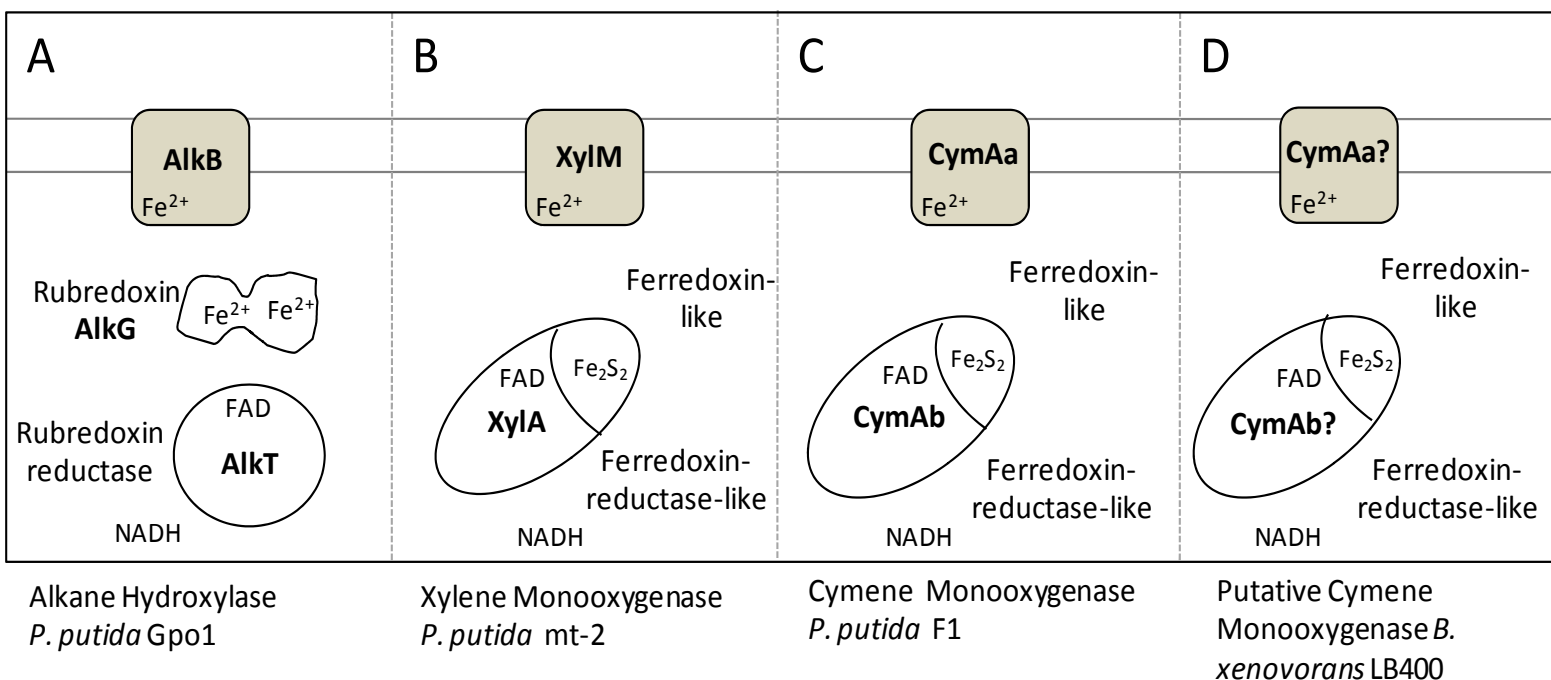


Figure 2.3: Schematic representation of membrane bound non-heme diiron monooxygenases

- A. Alkane Monooxygenase and its three components (hydroxylase, rubredoxin and rubredoxin reductase)
- B. Xylene Monooxygenase enzyme composed of two components: hydroxylase (XylM) and reductase (XylA)
- C. Cymene monooxygenase enzyme composed of a hydroxylase (CymAa) and reductase (CymAb)
- D. Putative Cymene monooxygenase composed of the hydroxylase (CymAa?) and reductase (CymAb?)

Figure adapted and modified from van Beilen (1994)

oxygen activation by non-heme diiron oxygenases was not fully understood, but crystallographic data of enzymes belonging to this group reveal that the non-heme Fe^{II} active site generally have two histidine residues, one monodentate carboxylate and two to three water ligands (Hegg and Que 1997; Solomon *et al.* 2003). In contrast to heme- Fe^{II} , non-heme Fe^{II} sites have an additional exchangeable position that allows the possibility of substrate and/or cofactor binding to the Fe^{II} and new coordination modes of oxygen activation (Shan and Que 2006).

2.2.2.1 Alkane Monooxygenase from *P. putida* GPo1

P. putida GPo1 (formerly *P. oleovorans*) can catalyze the degradation of *n*-alkanes. Enzymes responsible for this activity are encoded by the *alkBFGHJKL* and *alkST* operons located in the OCT plasmid of the GPo1 strain (van Beilen *et al.* 1994). Alkane monooxygenase catalyzes the first step of alkane metabolism by the oxidation of *n*-alkanes to *n*-alkanols. The octane degradation pathway initiated by alkane monooxygenase in *P. putida* GPo1 is summarized in Figure 2.4. In general, alkane degrading bacteria contain multiple alkane hydroxylase genes for the degradation of alkanes (van Beilen *et al.* 2002). The presence of these multiple genes is proposed to be essential for substrate recognition (van Beilen and Funhoff 2007). The alkane hydroxylase system consists of three components (see Fig 2.3A): a membrane-bound hydroxylase component and two soluble proteins, a rubredoxin and a rubredoxin reductase (Peterson *et al.* 1966; Benson *et al.* 1977; Kok *et al.* 1989). AlkB is the hydroxylase component, containing 25% sequence homology to XylM (xylene hydroxylase component) from *P. putida* mt-2 (Suzuki *et al.* 1991). The rubredoxin component (AlkG) of the alkane hydroxylase system belongs to a class of red-colored

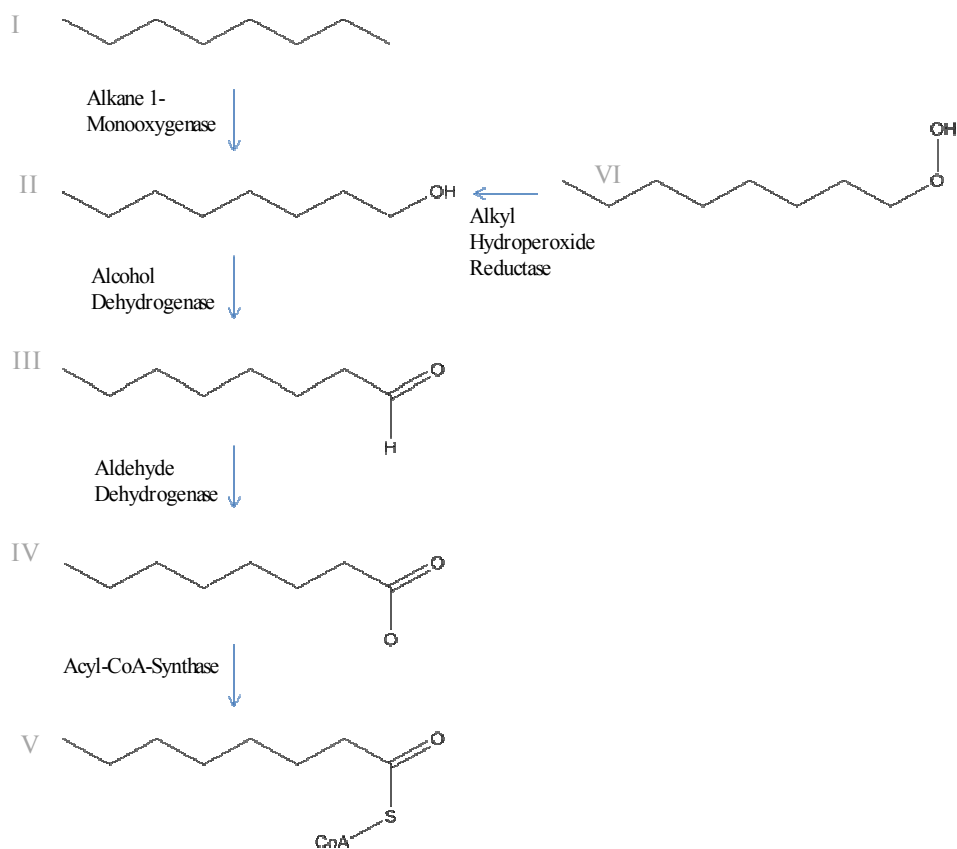


Figure 2.4: Degradation pathway of *n*-alkanes by *P. putida* Gpo1

I. *n*-Octane, II. 1-Octanol, III. 1-Octanal, IV. Octanoate, V. Octanoyl-CoA and VI. Octane hydroperoxide

electron transfer proteins that contain one or more active centers consisting of a single iron atom coordinated by four cysteine sulfur atom ligands in a tetrahedral structure (Lovenberg 1965; van Beilen 1994). The rubredoxin reductase (AlkT) is a flavoprotein that transfers electrons from NADH to a rubredoxin, which in turn transfers electrons to the reductase component (van Beilen 1994).

Approximately 100 different species of bacteria and 100 different species of fungi are known to metabolize alkanes (Leahy and Colwell 1990; Harayama *et al.* 1999; Austin *et al.* 2003). To date, only a few enzymes have been identified capable of the oxidation of alkanes to alkanols. Sequence similarity analysis have identified almost identical genes to the *alkB* of *P. putida* GPo1, other strains of *P. putida* and *P. mendocina*, suggesting evolutionary relationships (van Beilen *et al.* 2001). Although several bacterial strains have been identified capable of growing on medium and long-chain *n*-alkanes, *alkB* genes or *alkB* homologues responsible for this activity have not been identified in every strain. In order to identify the enzymes responsible for this activity, van Beilen and coworkers worked on the design of degenerate primers made possible by the identification of conserved regions for the identification of closely related genes to the *alkB* from *P. putida* GPo1 (Smits *et al.* 1999). Degenerate primers design made possible by the identification of conserved regions in the alkane monooxygenases from *P. putida* GPo1 (AlkB) and *Acinetobacter sp.* ADP1 (AlkM). Despite the similarity of these two enzymes, their *alk* gene organization is quite different from each other. The *alk* genes in *P. putida* GPo1 are clustered together in the OCT plasmid, while the *alk* genes in *Acinetobacter sp.* ADP1 are not grouped, clustered or localized on a plasmid as they are for GPo1 (Ratajczak *et al.* 1998). In fact, while the *alk* genes are located in the OCT

plasmid in strain GPo1, in other species these genes have been found to be inserted on the chromosome, suggesting that gene transfer occurred. Their findings indicate that the *Acinetobacter* sequences cluster in one group, while *Pseudomonas* sequences are divergent among the genus (Ratajczak *et al.* 1998).

Although the structure of the active site of the alkane monooxygenase is not known, spectroscopic and genetic evidence suggest that there is a nitrogen-rich environment located in the protein (Smits *et al.* 2002) containing eight or nine histidine residues coordinating the iron atoms (van Beilen *et al.* 2005). These histidine residues have been found to be conserved in related enzymes, and some of them have been proven to be required for catalysis, acting as ligands for the irons present in this protein (Shanklin *et al.* 1997). Studies have found that some non-heme, membrane-bound enzymes utilize the oxygen rebound mechanism for the hydroxylation of its substrate (Austin *et al.* 2000; Groves 2003; Bertrand *et al.* 2005) which involves the cleavage of the C-H bond by an electrophilic metal-oxo intermediate to generate a substrate-based radical (Bertrand *et al.* 2005). Identification of these intermediates can be accomplished by the use of diagnostic substrates (i.e. norcarane) and the production of ring-opened vs. ring-closed products depends on the rebound of the enzyme and the distribution of *endo* and *exo* isomers depend on active site constraints and thermodynamics (Bertrand *et al.* 2005). Austin and coworkers (2005) described two different possibilities where the substrate can change the active site. One is that due to the closeness of AlkB to the cell membrane, hydrophobic compounds in the cell membrane might affect the structure and reactivity of the enzyme. Another possibility could be that the accumulation of hydrocarbons in the lipid bilayer affects the membrane structure and function of the

enzyme. The AlkB diiron active site is suggested to be near the membrane interface and the substrate binding pocket is arranged to accommodate a medium sized hydrocarbon (Groves 2006) in order to have catalytic activity.

2.2.2.2 Xylene Monooxygenase (XMO)

Xylenes and toluenes are chemicals widely used in industry as solvents and as components for the industrial synthesis of rubber and paints. These chemicals have been found as contaminants of soil and water, generating a general concern among the community. Since early in the 1900s, finding biological ways for the degradation of these compounds has been the research interest of many scientists. The enzyme responsible for the first step of catalysis of these chemicals is xylene monooxygenase. The strain *P. putida* mt-2 is capable of degrading xylenes' and toluenes'. *P. putida* mt-2 contains a plasmid, pWW0, which contains genes coding for the “upper” and “meta” pathway for xylenes and toluenes degradation. The “upper” operon contains genes for the production of the carboxylic acid form of these compounds and the “meta” operon encoding genes responsible for the biotransformation of carboxylic acids to TCA cycle intermediates (Bühler *et al.* 2000). Xylene monooxygenase catalyzes the oxidation of toluenes and xylenes to (methyl) benzyl alcohols (Worsey and Williams 1975; Suzuki *et al.* 1991) (see Figure 2.5 for pathway information). It is encoded by two different genes, XylA and XylM. XylM is the enzyme hydroxylase component, and its protein sequence is similar to the AlkB and CymAa hydroxylase components involved in the degradation of octane and cymene, respectively. XylA is the reductase component, and it transfers reducing equivalents from NADH to XylM (Shaw and Harayama 1992; Bühler *et al.* 2002). It has been demonstrated that in addition to xylenes and toluenes, xylene

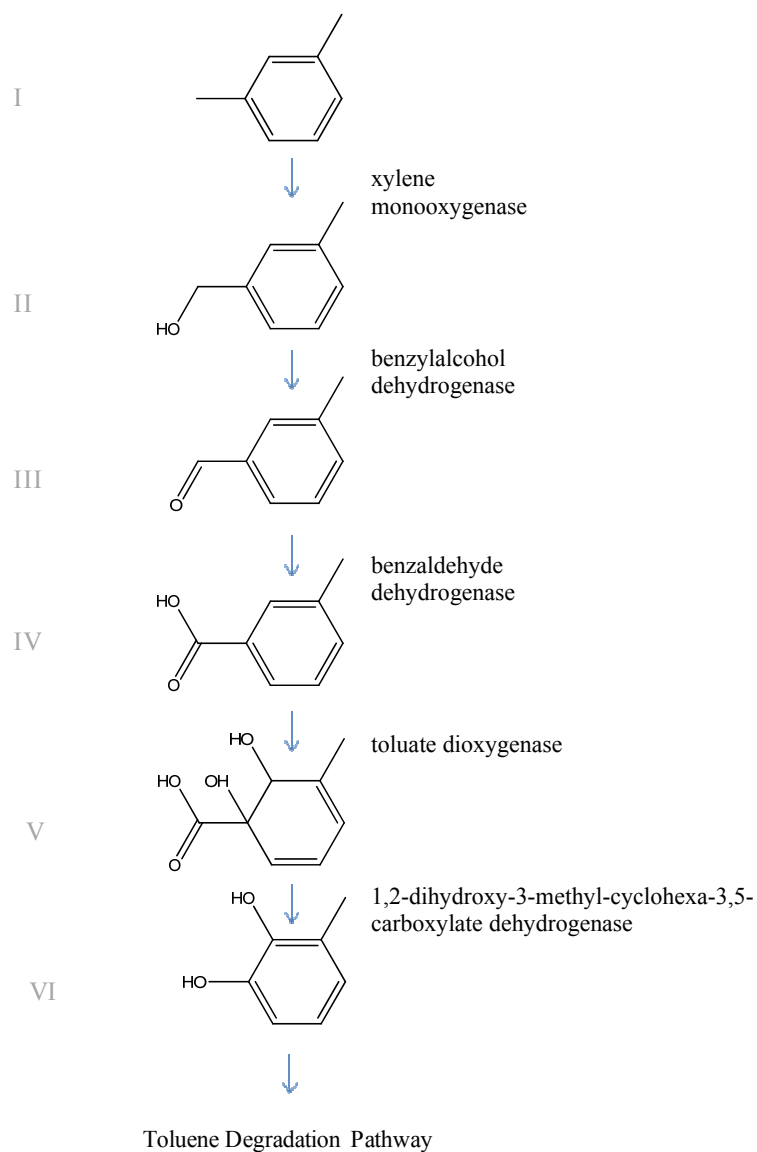


Figure 2.5: Degradation of *m*-xylene by *P. putida* mt-2 and *S. yanoikuyae* B1

I. *m*-Xylene, II. 3-Methylbenzylalcohol, III. 3-Methylbenzaldehyde, IV. *m*-Methylbenzoate, V. 1,2-Dihydroxy-3-methylcyclohexa-3,5-dienecarboxylate, VI. 3-Methylcatechol

monooxygenase can oxidize *m*- and *p*-ethyl- methoxy-, nitro-, bromo- and chloro-substituted toluenes to their respective benzyl alcohols (Kunz and Chapman 1981; Delgado *et al.* 1992; Wubbolts *et al.* 1994; Bühler *et al.* 2002). The *upper* operon of the pWW0 plasmid of *P. putida* mt-2 encodes three different enzymes: the xylene monooxygenase (XylMA), the benzyl alcohol dehydrogenase (XylB) (Shaw *et al.* 1992) and the benzaldehyde dehydrogenase (XylC) (Bühler *et al.* 2000). It also contains the *xy/N* and *xy/W* genes whose function is still unidentified (Bühler *et al.* 2000).

Research data published by Witholt and coworkers in 2000 reports that the recombinant xylene monooxygenase from *P. putida* mt-2 expressed in *E. coli* JM101 can also catalyze the formation of benzaldehyde and carboxylic acid in addition to the benzyl alcohol form. Although xylene monooxygenase can catalyze these reactions, the oxygenation rate for the conversion of *p*-cuminic aldehyde to *p*-cumate is low. Benzaldehyde dehydrogenase might be required by the native host to catalyze the reaction at a metabolically sufficient rate (Bühler *et al.* 2000). In addition, the alcohol and benzaldehyde dehydrogenases might be required by the host because dehydrogenation reactions are favorable for the host because reduced equivalents are produced rather than consumed, as occurs by the action of the xylene monooxygenase (Bühler *et al.* 2000).

Xylene monooxygenase is a non-heme diiron enzyme and it is closely related to the cymene monooxygenase from *P. putida* F1. Sequencing analysis results reveal that the cymene monooxygenase from *P. putida* F1 is most closely related to xylene monooxygenase from *P. putida* mt-2. Sequence identity for the hydroxylase component

CymAa/XylM is 40.4%, and percent identity of the reductase subunit CymAb/XylA is 41.5%. This study have identified a putative cymene monooxygenase in *B. xenovorans* LB400, so it will be interesting to compare its sequence similarity to XylM and CymAa.

2.2.2.3 Soluble Methane Monooxygenase (sMMO)

The soluble methane monooxygenase utilizes molecular oxygen to catalyze the hydroxylation of methane to methanol and water (Wallar and Lipscomb 1996). The enzyme can be found in methanotrophs and is the only enzyme capable of cleaving the very stable C-H bond of methane. It is a soluble non-heme diiron enzyme is composed of three different components: the hydroxylase component (MMOH), a reductase component (MMOR) that contains FAD and a [2Fe-2S] cluster, and a regulatory component known as Component B (MMOB) (Fox *et al.* 1989; Lipscomb 1994; Wallar and Lipscomb 1996; Zhang *et al.* 2006). The hydroxylase component contains a catalytic site with a diiron center and the reductase component is required to reduce the diferric state of the active site of the hydroxylase component to the diferrous state needed for oxygen activation and oxidation of the substrate. Studies show that the effector protein Component B controls the regioselectivity of hydroxylation of substrates, suggesting that it controls the shape of the hydroxylase active site (Froland *et al.* 1992; Zhang *et al.* 2006). The regulation catalyzed by Component B has been generating great interest because sMMO contains a regulatory system that ensures that the use of NADH is coupled to the hydrocarbon oxidation, showing a selective preference for methane as the primary substrate (Zhang *et al.* 2006). The hydroxylase component (MMOH) and the regulatory component (MMOB) are responsible for this regulation.

Resolution of the crystal structure of methane monooxygenase from *Methylococcus capsulatus* reveal that the coordination of the active site diiron cluster is coordinated to the α subunit, similarly to the coordination of the iron binding domain of ribonucleotide reductase R2 (RNR R2) (Rosenzweig *et al.* 1993; Nordlund and Eklund 1995). Rosenzweig *et al.* (1993) reported that the protein ligands to the diiron coordination site are two histidines and four carboxylate groups. Folding of the $\alpha\beta$ dimer is similar to the RNR R2 dimer indicating an evolutionary relationship between the two enzymes (Nordlund and Eklund 1995).

2.2.2.4 Cymene Monooxygenase (CMO)

The terpene *p*-cymene (also referred to as *p*-isopropyltoluene) is a natural organic product that contains an isopropyl group attached to a benzene ring with a methyl group oriented *para*. Although the degradation of *p*-cymene has been reported to be very common in different bacterial species (Eaton 1997), not much information is known about this enzyme. Previous research found that *P. putida* F1 genes involved in the degradation of *p*-cymene are located on the *cym* operon, which in addition to the cymene monooxygenase, encodes the *p*-cumin alcohol dehydrogenase, *p*-cumate alcohol dehydrogenase, a regulatory protein, an outer membrane protein and the acetyl-CoA synthetase, whose function is still unidentified. These enzymes catalyze the biotransformation of *p*-cymene to *p*-cumate, which is further metabolized to the intermediates isobutyrate, acetyl CoA and pyruvate, all intermediates of the respiratory pathway. Degradation of *p*-cymene and subsequent steps leading to the respiratory pathway and products are summarized in Figure 2.6. Cloning and sequence analysis of the *cym* operon was performed by Eaton in 1997 to describe the function of each gene

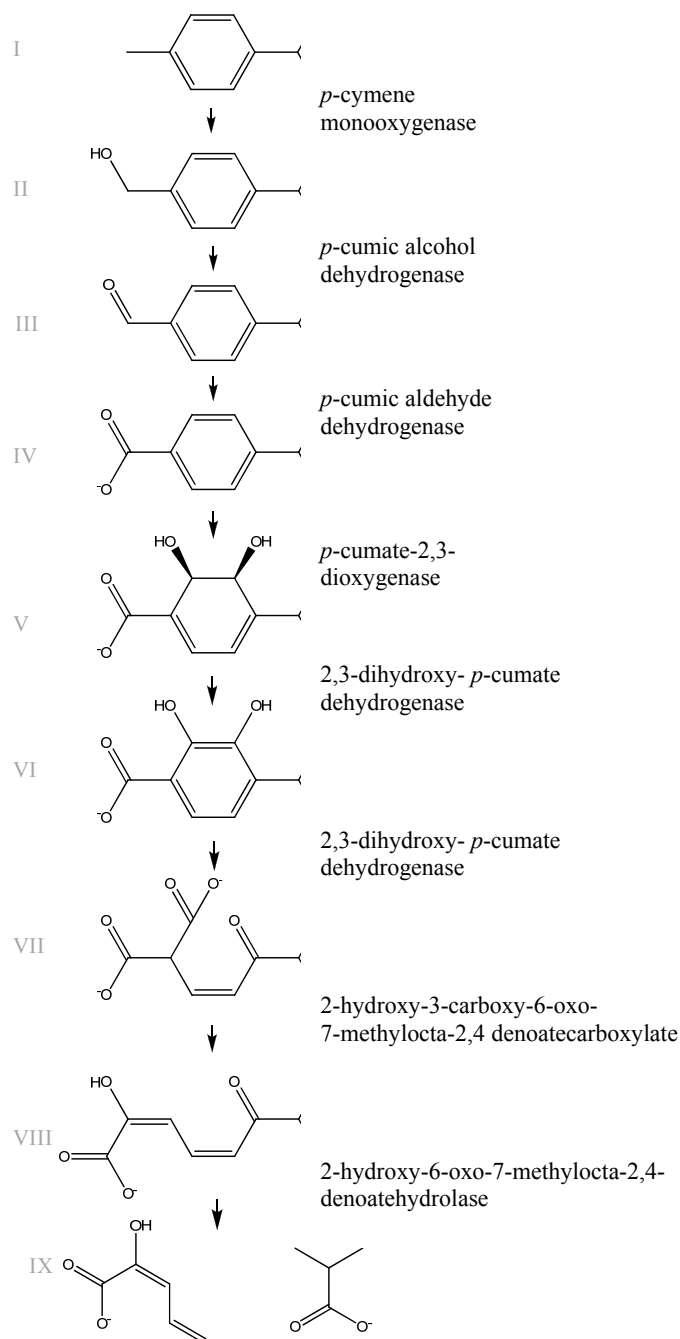


Figure 2.6: Degradation pathway of *p*-cymene by *P. putida* F1

I. *p*-Cymene, II. *p*-Cumic Alcohol, III. *p*-Cumic Aldehyde IV. *p*-Cumate V. *cis*-2,3-Dihydroxy-2, 3-Dihydro-*p*-Cumate VI. 2,3-Dehydroxy *p*-Cumate VII. 2-Hydroxy-3-carboxy-6-oxo-7-methylocta-2,4-Dienoate VIII. 2-Hydroxy-6-oxo-7-methylocta-2,4-Dienoate IX. *cis*-2-Hydroxy Penta-2,4-Ienoate/ Isobutyrate

present in the operon. For comparisons with the putative enzyme from LB400, we will focus on the cymene monooxygenase, which is encoded by the genes *cymAa* and *cymAb*. *cymAa* encodes the hydroxylase component of cymene monooxygenase and *cymAb* encode its reductase component. Cloning of these two genes into a recombinant *E. coli* and protein expression resulted in the conversion of *p*-cymene into a combination of *p*-cuminic alcohol, *p*-cuminic aldehyde and *p*-cuminic acid (which seems to be the result of a spontaneous oxidation of *p*-cuminic aldehyde) (Eaton 1997). CymAa of *P. putida* F1 has a molecular weight of 43 kDa and it is closely related to the xylene monooxygenase enzyme from *P. putida* mt-2, with 39 % amino acid sequence identity, and to the AlkB from OCT plasmid from *P. putida* Gpo1, with 22% of sequence identity. It has been found that these three enzymes contain a conserved region of 50 amino acids, where 11 of these amino acids are histidine residues. Eight of these 11 histidine residues have been identified in other enzymes, such as the rat liver stearoyl coenzyme A Δ^9 desaturase and other membrane-associated eukaryotic and cyanobacterial desaturases (Shanklin *et al.* 1997). It was found that the rat liver stearoyl coenzyme A Δ^9 desaturase requires all of these histidine residues for catalytic activity (Shanklin *et al.* 1994). The histidine residues found in this motif act as ligands for the two iron atoms that serve as cofactors for this group of enzymes (Shanklin *et al.* 1994).

cymAb encodes the reductase subunit of the cymene monooxygenase and has a molecular weight of ~38 kDa. Sequence analysis revealed that the gene contains the number and order of cysteine residues characteristic of the iron-sulfur flavin reductases which are closely related to the family of reductases where flavoprotein and ferredoxin are combined (Dutta and Gunsalus 1997). The amino terminus of this enzyme is closely

related to the amino terminus of a typical chloroplast-type ferredoxin which contains a [2Fe-2S] binding site that is conserved in all the reductases described belonging to the flavoprotein and ferredoxin family (Dutta and Gunsalus 1997).

In addition to *cymAa* and *cymAb* genes, the *P. putida* F1 *cym* operon contains other genes involved in the biotransformation of *p*-cymene to *p*-cumate and then to intermediates of the respiratory chain. The operon contains the *cymB* gene, which encodes the protein *p*-cuminic alcohol dehydrogenase that catalyzes the conversion of *p*-cuminic alcohol to *p*-cuminic aldehyde in the presence of the cofactor NAD⁺. This enzyme has a molecular weight of ~26 kDa and contains six conserved amino acids present in all the class II alcohol dehydrogenases (Eaton 1997). The most closely related enzymes are the bacterial oxidoreductases. The *cymC* gene product is *p*-cuminic aldehyde dehydrogenase with a molecular weight of ~53 kDa. It catalyzes the conversion of *p*-cuminic aldehyde to *p*-cumate in the presence of NAD⁺. This enzyme belongs to the family of NAD⁺-linked aldehyde oxidoreductases (Eaton 1997). The most closely related enzymes are the aldehyde dehydrogenases of eukaryotes. The *cymD* gene represents a putative outer membrane protein with a deduced molecular weight of ~48 kDa. It is similar to other proposed outer membrane proteins encoded by genes that are part of operons encoding the catabolism of aromatic compounds (Eaton 1997). The acetyl coenzyme A synthetase is encoded by *cymE*, and has a molecular weight of ~71 kDa. This enzyme is similar to a number of acetyl coenzyme A synthetases (Eaton 1997). The role of this enzyme in the metabolism of *p*-cymene it is still unidentified. A regulatory protein identified as CymR with molecular weight of ~23 kDa has been identified in this

operon, which shows very low similarity to other regulatory proteins already described (Eaton 1997).

The cymene monooxygenase from *P. putida* F1 was cloned and expressed in order to get some insight about the substrate range and specificity for this enzyme and to compare it to the closely related xylene monooxygenase from *P. putida* mt-2 (Nishio *et al.* 2001). Cymene monooxygenase activity towards a variety of substrates is summarized in Table 2.1. Chemical structures of the compounds tested are represented in Figure 2.7. Cymene monooxygenase enzyme was capable of oxidizing other substrates in addition to *p*-cymene, such as 4-chlorotoluene, 4-ethyltoluene and 4-methylthiotoluene. These results indicate that the position of the substituent and the size of the alkyl group attached affect the specificity of the enzyme, showing preferences for substitutions at the *para* position (Nishio *et al.* 2001). This enzyme is also capable of the epoxidation of styrene and the oxidation of 4-chlorostyrene oxide from 4-chlorostyrene (Nishio *et al.* 2001). It is our goal to clone, express and study the substrate specificity analysis of the putative cymene monooxygenase from *B. xenovorans* LB400.

Oxygen Activation by Heme and Non-heme Oxygenases

Since the discovery of oxygenases more than 50 years ago, great interest was generated concerning the activation of molecular oxygen by these enzymes. It all started with the work of Hayaishi *et al.* (1955), when they isolated *Pseudomonas* sp capable of the degradation of tryptophan to kynurenine and then, to anthranilic acid (Hayaishi *et al.* 1955). Interestingly, anthranilate was further oxidized to CO₂, NH₃ and H₂O (Hayaishi *et al.* 1955; Hayaishi 2005). At this time, it was thought that molecular oxygen was an

Compound	Expected Product	Rate of production (U/g)
Cumene	2-Phenyl-1-propanol	n.d.
<i>p</i> -Cymene	4-Isopropylbenzyl alcohol	14.2
Ethylbenzene	Phenyethyl alcohol	n.d.
4-Ethyltoluene	4-Ethylbenzyl alcohol	10.8
<i>n</i> -Hexane	1-Hexanol	n.d.
4-Phenyltoluene	4-Phenylbenzyl alcohol	n.d.
Styrene	Styrene oxide	1.30
Toluene	Benzyl alcohol	n.d.
<i>o</i> -Xylene	2-Methylbenzyl alcohol	n.d.
<i>m</i> -Xylene	3-Methylbenzyl alcohol	1.15
<i>p</i> -Xylene	4-Methylbenzyl alcohol	2.07
4-Chlorostyrene	4-Chlorostyrene oxide	33.3
4-(Methylthio)toluene	4-(Methylthio)benzyl alcohol	14.1
2-Chlorotoluene	2-Chlorobenzyl alcohol	n.d.
3-Chlorotoluene	3-Chlorobenzyl alcohol	2.10
4-Chlorotoluene	4-Chlorobenzyl alcohol	14.2
4-Fluorotoluene	4-Fluorobenzyl alcohol	2.48
4-Nitrotoluene	4-Nitrobenzyl alcohol	2.17

Table 2.1: List of chemicals and degradation rates by cymene monooxygenase from *P. putida* F1

[g- Grams cell dry weight; n.d.- no detectable activity; U- units of activity (nanomoles produced/minute)]

Table adapted from Nishio *et al.* (2001)

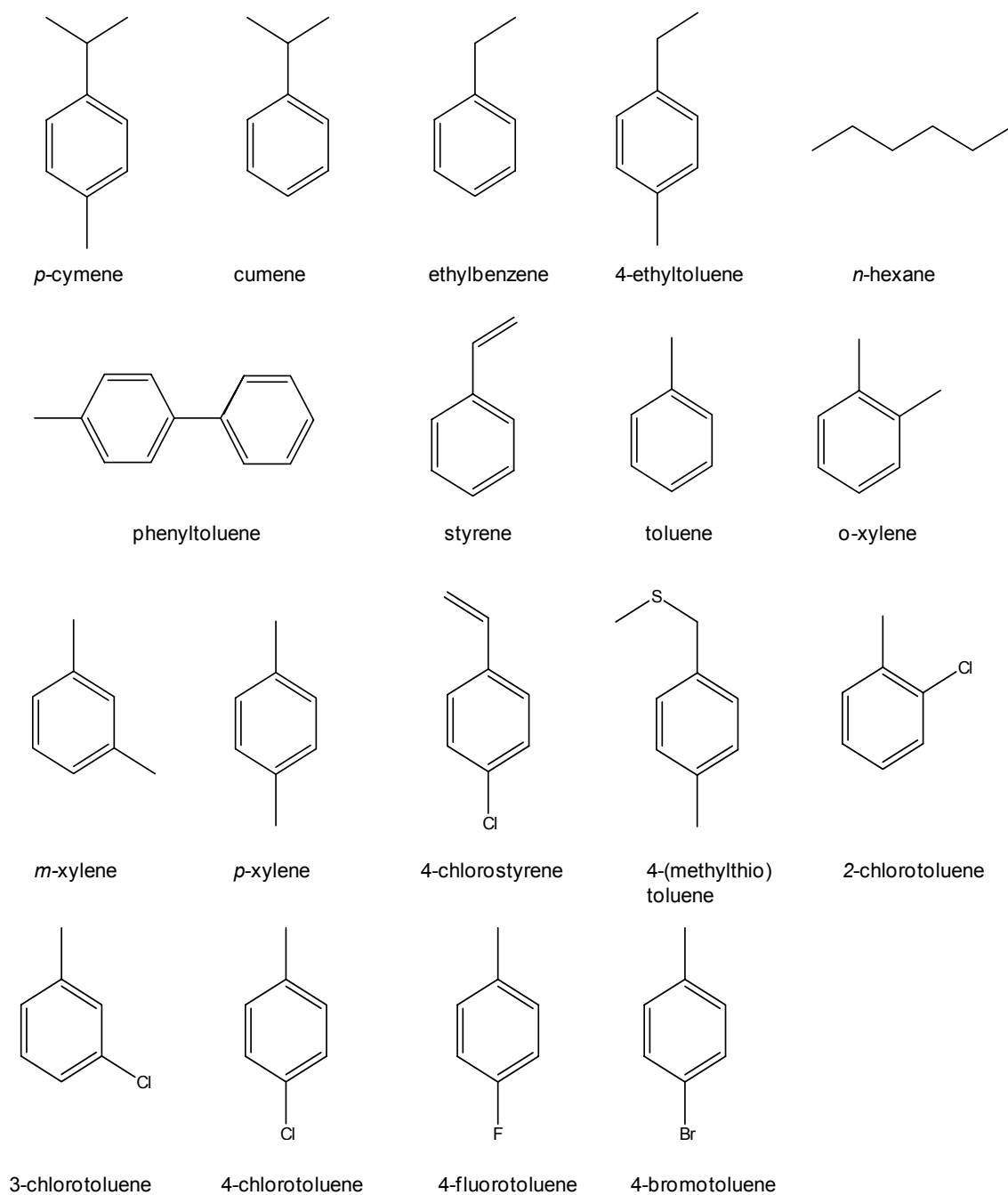


Figure 2.7: Chemical structures of substrates utilized for the substrate specificity analysis of the cymene monooxygenase from *P. putida* F1

electron acceptor that was reduced in the process to H_2O or H_2O_2 . It was never thought that oxygen was directly incorporated into the substrate(s) (Hayaishi 2005). However, further research using labeled isotopes confirmed that oxygen was indeed incorporated into the substrate and great interest was generated in the area of oxygen activation mechanisms by the oxygenases (Carrondo *et al.* 2007).

Heme and non-heme diiron oxygenases can activate molecular oxygen and initiate the oxidation of its substrate. Oxygen activation by heme diiron oxygenases has been widely studied spectroscopically, due to the fact that the porphyrin ring exhibits intense transitions from $\pi \rightarrow \pi^*$ (Solomon *et al.* 2003). Non-heme diiron oxygenases lack this $\pi \rightarrow \pi^*$ characteristic of the porphyrin ligand making the spectroscopic analysis of its active site and catalysis difficult. Resolution of crystal structures from different non-heme diiron oxygenases has provided some insight about the active site and oxygen activation.

2.2.3 Oxygen Activation by Heme Containing Monooxygenases

During heme catalysis, the porphyrin macrocycle leaves only one axial position for oxygen binding and activation (Solomon *et al.* 2003). In P450 enzymes, the heme group is directly involved in the activation of oxygen and substrate oxidation (Bühler 2003). The catalytic cycle of cytochrome P450 enzymes has been widely studied (Ortiz de Montellano 1995; Schlichting *et al.* 2000; Newcomb *et al.* 2003).

Figure 2.8 shows the intermediates formed in the process of oxygen and substrate activation by cytochrome P450 enzymes as described by Newcomb *et al.* (2003). First

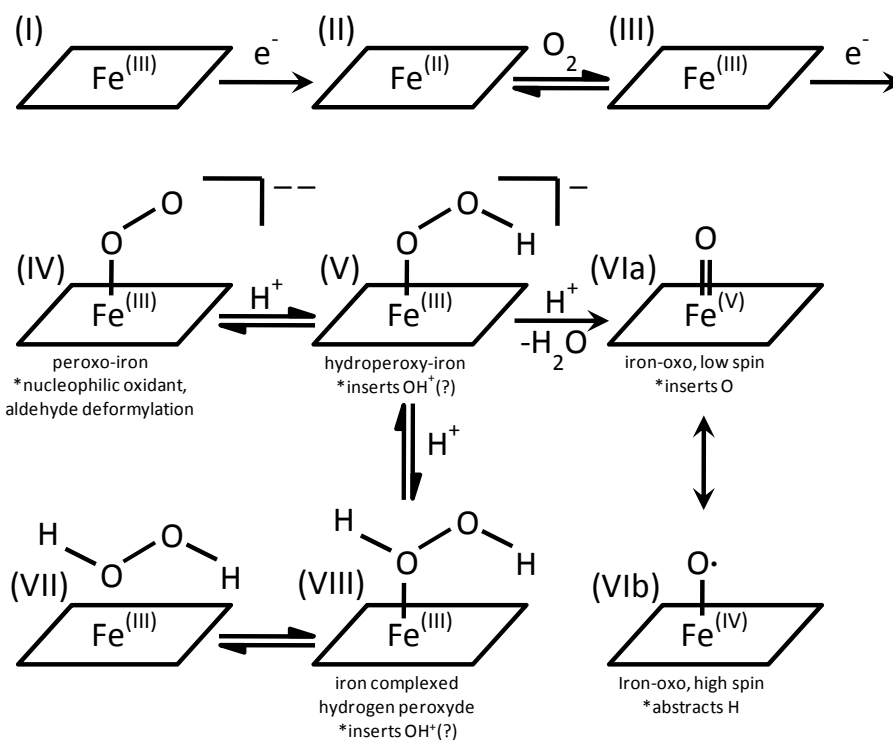


Figure 2.8: Intermediate species formed in the reaction of P450 monooxygenases

The parallelogram represents the heme group; * represents the proposed role of Fe-O intermediates as oxidants.

Adapted from Bühler *et al.* (2003) and Newcomb *et al.* (2003)

the enzyme binds the substrate when iron is in its resting state, reducing its reduction potential (I). Then, P450 accepts electrons from NAD(P)H to produce Fe^{II} (II). This allows the binding of oxygen to produce the superoxide-iron complex, followed by a reduction to produce the peroxo-iron where oxygen is in its formal oxidation state of hydrogen peroxide (IV). Protonation of the peroxo-iron on the distal oxygen produces the hydroperoxo-iron intermediate (V). A second addition of protons at the distal oxygen and the loss of a water molecule produce the iron-oxo intermediate (VIa) which is supposed to interact through different spin states (VIa and VIb). When protonation occurs at the proximal oxygen, iron-complexed hydrogen peroxide can be produced, which can be dissociated (VII). The exact mechanism for the insertion of oxygen is still not fully understood, but Newcomb *et al.* (2003) reviewed the reaction indicating that multiple mechanisms and oxidants are involved (Newcomb *et al.* 2003; Chandrasena *et al.* 2004). Figure 2.8 shows the iron-complexed hydrogen peroxide (VIII) as possible electrophilic oxidant to catalyze the insertion of OH into the C-H bond.

2.2.4 Oxygen Activation by Non-heme Containing Monooxygenases

In non-heme catalysis the iron sites have additional, exchangeable positions that allow the possibility of substrate and/or cofactor binding to the iron and new coordination modes for oxygen activation (Solomon *et al.* 2003). The active site of a non-heme diiron enzyme such as the methane monooxygenase has two histidine residues, one monodentate and two to three water molecules where the ligation is less covalent, which can greatly modify the electronic structure of the active site and the oxygen activation (Solomon *et al.* 2003). Oxygen activation in methane monooxygenase has been

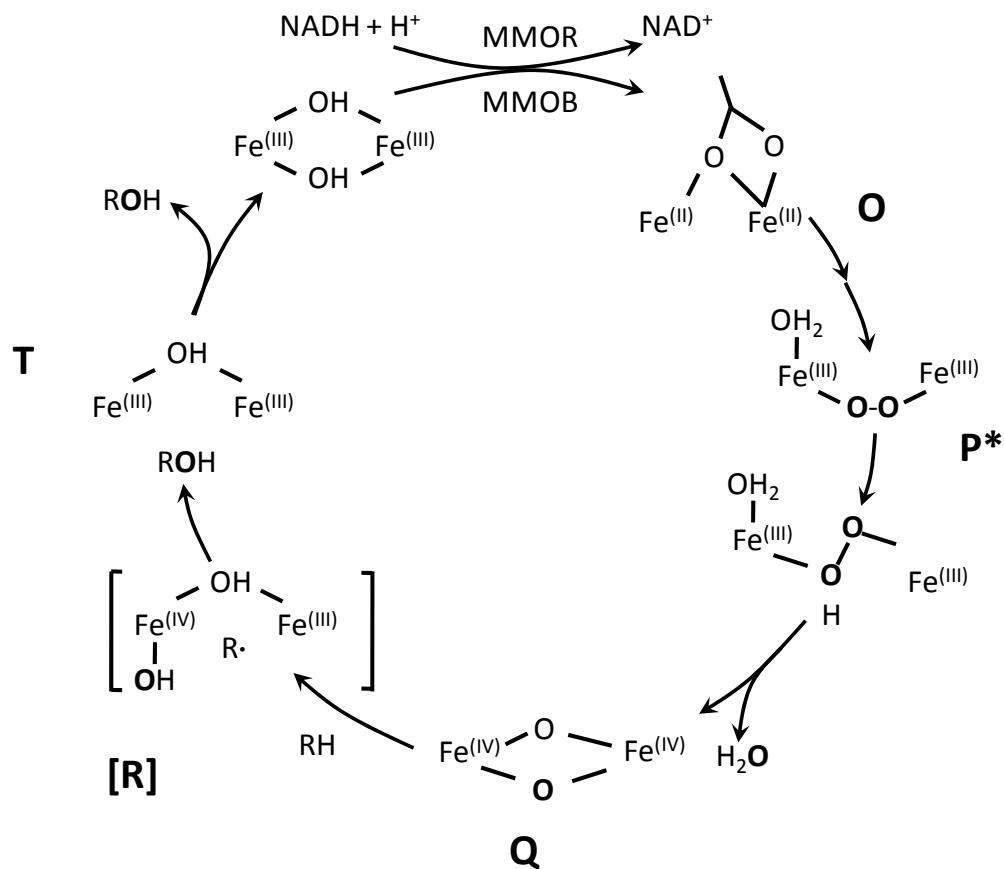


Figure 2.9: The catalytic cycle of soluble methane monooxygenase

Atoms derived from molecular oxygen are in bold. See text for description of O, P*, Q, [R] and T species

Adapted from Wallar and Lipscomb (1996); Brazeau *et al.* (2001); Bühler (2003)

described by Wallar and Lipscomb (1996). The active diiron site of methane monooxygenase involves the formation of several intermediates during the activation of oxygen. Figure 2.9 summarizes the steps involved in oxygen activation by the sMMO (Wallar and Lipscomb 2001), starting with the reduction of the diferric state (H^{ox}) of the active site by two electrons donated by NADH. Diferrous state (H^I) can react with oxygen through an intermediate (O) where the oxygen can be bound to the enzyme but not to the cluster. Then, the oxygen binds to the cluster to form the intermediate (P^*) which with addition of a proton is converted to the hydroperoxo intermediate (P). The addition of a second protons form intermediate (Q) followed by the cleavage of the highly stable C-H bond and oxygen insertion forming complex (T). When the product is released, the cycle is completed (Wallar and Lipscomb 2001; Brazeau and Lipscomb 2003).

Although the analysis of the mechanism of oxygen activation is more difficult to study in membrane bound enzymes, the hydroxylation mechanism of class III non-heme diiron enzymes alkane and xylene monooxygenase has been studied by using the diagnostic substrate norcarene (Austin *et al.* 2000; Austin *et al.* 2003). XylM and AlkB contain a histidine rich active site that catalyzes the insertion of an oxygen atom into a C-H bond (Bühler *et al.* 2000; Austin *et al.* 2003). The study was based on the use of a substrate that undergoes characteristic structural changes in response to forming specific reaction intermediates (Austin *et al.* 2000), providing evidence of the intermediates produced in catalysis. The data confirm that XylM and AlkB both utilize the oxygen rebound mechanism for oxygen activation and presumably, oxygen binding and reduction lead to the ferryl species $O=Fe^{IV}-Fe^{IV}$ (Austin *et al.* 2000; Austin *et al.* 2003). While

these enzymes reductively activate molecular oxygen to hydroxylate its substrates, they differ in their active site coordination chemistry, cellular location and cofactor content (Rozhkova-Novosad *et al.* 2007). Similar studies on the cymene monooxygenase and the novel monooxygenase discussed in this thesis could also provide some insight about the enzyme mechanics and the active site.

Although the degradation of alkanes has been identified in many organisms, the enzymes responsible for this activity have not been always identified. It has been proposed that alkane hydroxylases can be identified by their reaction mechanisms and unidentified alkane hydroxylases can be characterized based on their mechanisms facilitating the identification of these enzymes in whole cells (Rozhkova-Novosad *et al.* 2007). Since cymene monooxygenase from *P. putida* F1 has been identified as a membrane-bound non-heme diiron monooxygenase and our study identified the putative cymene monooxygenase from *B. xenovorans* LB400 also as a membrane-bound non-heme diiron monooxygenase, it would be interesting to determine if these enzymes follow the identified oxygen rebound mechanism described in alkane and xylene monooxygenase. Identification and characterization of these mechanisms could facilitate the identification and characterization of similar enzymes that have not been described yet by whole cell studies.

3. Materials and Methods

3.1 Bacterial Strains, Plasmids and Media

B. xenovorans LB400 strain was grown in liquid mineral salts medium (MSB) (Stanier *et al.* 1966) supplemented with 20 mM succinate as the carbon source at overnight at 30°C. *E. coli* DH5 α was used as plasmid host. For growth analysis in a variety of hydrocarbons as the sole carbon source, *B. xenovorans* LB400 was grown in 500 mL flasks by adding 100 mL liquid media supplemented with the appropriate hydrocarbon source provided in the vapor phase up to saturation. The expression vector used in this study was pQE-30 from Qiagen (Valencia, CA). The generated plasmid was cloned in *E. coli* BL21*(DE3) and *E. coli* clones containing the putative recombinant cymene monooxygenase were grown on LB-ampicillin (100 μ g/mL) plates at 37°C. *E. coli* BL21*(DE3) was used for protein expression of the construct containing the hydroxylase and reductase genes. For protein expression of the putative cymene monooxygenase clones were grown on minimal media supplemented with 20 mM glucose. Strains and generated plasmid in this study are described in Table 3.1.

3.2 Chemicals and Reagents

Chemicals used in this research project were obtained from Sigma-Aldrich (St. Louis, MO). The PCR and sequencing primers used were obtained from Sigma Chemical Co. (St. Louis, MO). RT-PCR reagents were obtained from Qiagen (Valencia, CA). Herculase Enhanced Polymerase utilized for amplification of putative cymene monooxygenase was obtained from Stratagene (La Jolla, CA.). The TA cloning vector

Strain or Plasmid	Relevant Characteristic (s)	Source
<i>E. coli</i>		
DH5 α	<i>F</i> 80dlacZ_M15_ (lacZYA-argF)U169 deoR recA1 endA1 hsdR17 (r _K m _K) phoA supE44 thi-1 gyrA96 relA1	Invitrogen
BL21* (DE3)	<i>F</i> ompT hsdSB (r _B m _B) gal dcm rne131 (DE3)	Invitrogen
<i>B. xenovorans</i> LB400	Wild Type	Accession numberYP_557479
Plasmids		
pCR2.1	Km ^r Ap ^r ; <i>E. coli</i> expression vector, source of <i>lac</i> promoter	Invitrogen
pTrc99A	Ap ^r ; <i>E. coli</i> expression vector, source of <i>trc</i> promoter	Amersham Pharmacia Biotech
pQE-30	Ap ^r ; <i>E. coli</i> expression vector, source of T5 promoter	Qiagen
pGEM-T	Ap ^r ; <i>E. coli</i> cloning vector, source of T7 and SP6 promoter	Promega
pET101-D/TOPO	Ap ^r ; <i>E. coli</i> expression vector, source of T7 promoter	Invitrogen
CMOpCR2.1	Km ^r Ap ^r ; pCR2.1 vector with a 2.7kb fragment containing putative cymene monooxygenase genes	This study
CMOpTrc99A	Ap ^r ; pTrc99A vector with a 2.7kb fragment containing putative cymene monooxygenase genes	This study
CMOpQE-30	Ap ^r ; pQE-30 vector with a 2.7kb fragment containing putative cymene monooxygenase genes	This study
CMOpGEM-T	Ap ^r ; pGEM-T vector with a 2.7kb fragment containing putative cymene monooxygenase genes	This study
CMOpET101-D/TOPO	Ap ^r ; pET101-D/TOPO vector with a 2.7kb fragment containing putative cymene monooxygenase genes	This study

Table 3.1: Strains and plasmids used in this study

kit (pGEM-T) was obtained from Promega. Restriction endonucleases were obtained from New England BioLabs.

3.3 Amplification of the Putative Cymene Monooxygenase

The genes coding for the putative cymene monooxygenase enzyme were amplified using the forward primer BamHI-CMO1-F and the reverse primer CMO2-R (see Table 3.2 for primer sequences). Gene amplification was obtained using the Herculase Enhanced DNA Polymerase (Stratagene, La Jolla, CA). The reaction was set up as follow: 100 ng of each primer, 4% dimethylsulfoxide, 150ng of DNA, 25 mM dNTPs, 5 μ L Herculase reaction buffer, and 5 U/ μ L of Herculase enzyme to a final volume of 50 μ L. The polymerase chain reaction was run on a GENE AMP PCR System 9700 from Applied Biosystems (Foster City, CA). The PCR program performed for this reaction was subdivided in three different steps. (1) Denaturation at 95°C for 2 minutes (2) 10 cycles [Denaturation at 95°C for 30 seconds, annealing at 52°C for 30 seconds and extension 68°C for 3 minutes] (3) 20 cycles [Denaturation at 95°C for 30 seconds, annealing at 52°C for 30 seconds and extension 68°C for 7 minutes] and finally, single cycle extension at 68°C for 10 minutes. The product size of this reaction was 2.7 kilobase pairs. Ten μ L of the PCR product was combined with 10 μ L of Ready-mix Taq PCR reaction mix with MgCl₂ (Sigma, St. Louis, MO) and was incubated at 72°C for 30 minutes for addition of A overhang nucleotide at the end as required for TA cloning.

Cymene Monooxygenase		
PCR Primers		
BamHI-CMO1-F	5'-CGGGATCCATGTGGGACTACATC-3'	Amplification of the putative cymene monooxygenase
CMO2-B	5'-GCAGCCTGACTGTACC-3'	
RT-PCR Primers		
CYO1 F	5'-ATGTGGGACTACATCAAATAC-3'	Amplification of the putative cymene hydroxylase region
CYO-NcoI-R	5'-AACCCGCACTCCGG-3'	
LB400 GYRB F	5'-GTCACGCTGATCAACGAG-3'	Amplification of positive control <i>gyrB</i>
LB400 GYRB R	5'-GCAGCACCTGAATATGCG-3'	
LB400bphA1 F	5'-ATGACGTGCTGGTCGAC-3'	Amplification of the biphenyl gene
LB400bphR R	5'-TCAGGTAGGTCTCCAGCG-3'	

Table 3.2: List of primers used in this study

3.4 Cloning of the Putative Cymene Monooxygenase

TA cloning of the putative monooxygenase was performed by combining 2 μ L of PCR product, 5 μ L of Rapid Ligation buffer, 50 ng of pGEM-T (Promega, Madison, WI) vector and 3 Weiss units/ μ L of T4 DNA ligase and incubated for 18 hr at 17°C. Two μ L of the ligated DNA was transformed into 25 μ L of *E. coli* DH5 α competent cells (Invitrogen, Carlsbad, CA) and incubated for 1 hour at 37°C and 200 rpm. Transformed cells were plated onto selective medium (LB plus 100 μ g/mL ampicillin) and incubated overnight at 37°C. For colony selection, plates were supplemented with 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-Gal) (Sigma-Aldrich, St. Louis, MO) for blue and white screening. The putative cymene monooxygenase genes were later transferred to the expression vector pQE-30 (Qiagen, Valencia, CA). Digestion of the pQE-30 vector and the CMO pGEM-T clone was performed by the addition of 0.5 μ g of DNA, restriction digestion buffer and 2 Units of the restriction enzymes BamHI and NotI (New England BioLabs, Ipswich, MA). Restriction digests were incubated at 37°C for three hours and the digested DNA was ligated by adding 5 μ L of ligation buffer and 0.5 μ L of T4 ligase (Invitrogen, Carlsbad, CA) and incubating overnight at 17°C. Three μ L of the ligation reaction was transformed into *E. coli* DH5 α competent cells (Invitrogen, Carlsbad, CA) and the cells incubated for 1 hour at 37°C and 200 rpm. Transformed cells were plated onto selective medium and incubated overnight at 37°C to select for colonies containing the CMOpQE-30 construct. Three additional constructs were created in a similar fashion in order to test for biotransformation efficiency with different promoters by using the following vectors: pCR2.1, pTrc99A and pET101-D/TOPO.

3.5 Sequencing Analysis

Sequencing reactions were performed to check for possible undesired mutations that might occur during amplification of the genes of interest using ABI 3100 Genetic Analyzer. Reactions were set up by the addition of 1.6 nmol of sequencing primer to 200 ng of CMOpQE-30 construct. Sequencing data were analyzed using the Chromas software to determine the quality of the run and screening for mutations was performed by comparing the published sequence in Genbank to the construct sequence with using the Lasergene Software version 5.0. Phylogenetic analysis and protein alignment of the most closely related genes was performed using the Lasergene Software version 5.0 MegAlign and sequence similarities were obtained by performing Blast searches at the National Center for Biotechnology Information (NCBI) website (www.ncbi.nlm.nih.gov).

3.6 Protein Expression and Biotransformation Assay

In order to obtain expression, the CMOpQE-30 construct DNA was transformed into the BL21(DE3) competent cells (Invitrogen, Carlsbad, CA). Cultures were grown overnight on MSB containing 20 mM glucose, 1.0 mM thiamine and 100 µg/mL ampicillin. Two percent of the overnight culture was added to a flask containing MSB media with the addition of 20 mM glucose, 1.0 mM thiamine and 100 µg/mL ampicillin for selection. Cells were grown at 37°C and 250 rpm until the culture reached an optical density (OD) of 0.5. One mM isopropyl-beta-D-thiogalactopyranoside [IPTG] (Sigma-Aldrich, St. Louis, MO) was added to induce expression. Cells were harvested by centrifugation at 8K for 10 minutes after one hour of incubation and resuspended in 50 mM Na/KPO₄ buffer pH 7.25 with 20 mM glucose. Volume was adjusted to an optical

density of 2.0 and the appropriate substrate was added. Hydrocarbons were provided in the vapor phase by adding 1.0 mL of the desired carbon source to a flask bulb inserted into a rubber stopper. Cells were incubated at 30°C and 200 rpm for 16 hours.

3.7 Extraction of the Biotransformed Products

E. coli cells exposed to the appropriate hydrocarbons were collected by adding 25 mL of cells into 50 mL centrifuge collection tubes and centrifuge cells using a Beckman Coulter Avanti J-25 centrifuge at 8K for 15 minutes. Twenty mL of the supernatant and 20 mL ethyl acetate were carefully transferred to a sterilized centrifuge tube and vortexed for one minute. The phases were separated by centrifugation at 8K for 15 minutes. The organic layer (ethyl acetate layer) was collected and transferred to a clean centrifuge tube. Anhydrous sodium sulfate was added to remove any water residue left in the sample. The extract was centrifuged at 8K for 5 minutes in order to be collected. The samples were dried using a Rotavapor (Buchi) for 15 minutes (or until sample completely dried out) and the dried residue resuspended in 2 mL of ethyl acetate. Figure 3.1 shows the methodology followed for the analysis.

3.8 Gas Chromatography-Mass Spectrometry Analysis

Analysis of substrate biotransformation and analysis of metabolites produced during bacterial growth was performed using a Hewlett Packard G1800C GCD System Series II Gas Chromatograph Electron ionization detector connected to a capillary column HP-5ms (25m x 0.25mm x 0.25µm). The chromatographic parameters for the detection were an inlet of 250°C, detector at 280°C and split/splitless injection of one

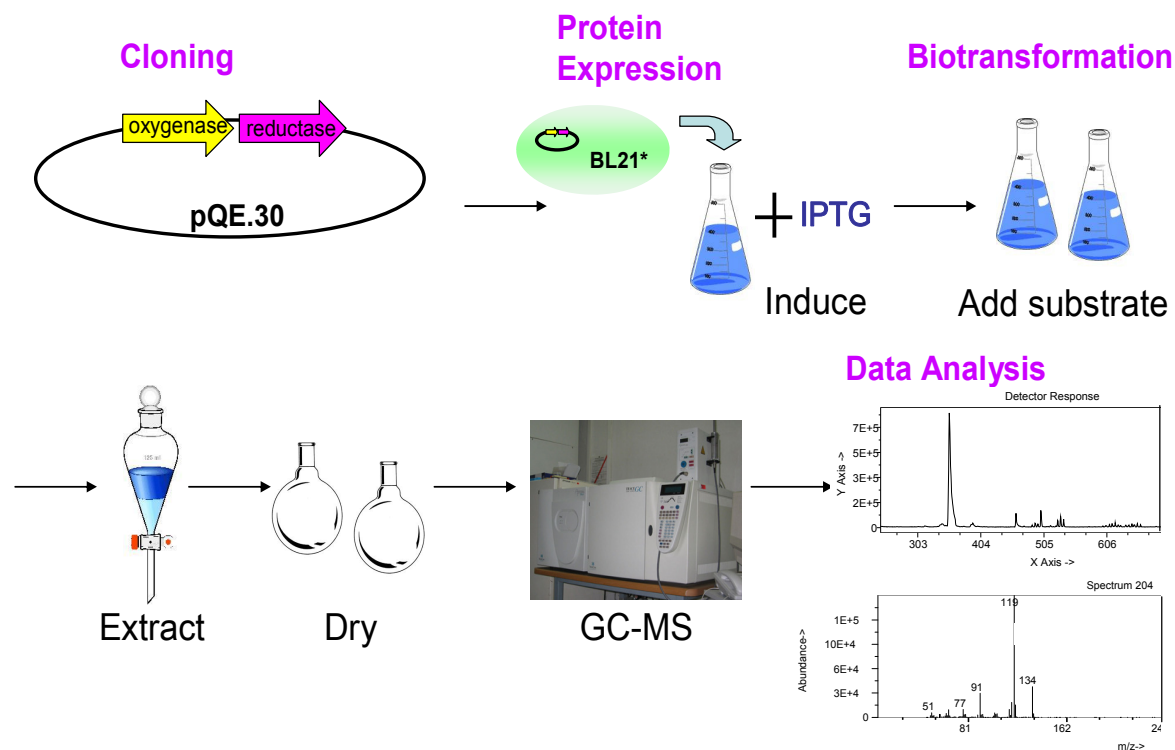


Figure 3.1: Methodology for protein expression and substrate specificity analysis

Once the monooxygenase system is cloned into an expression vector the construct has to be transferred to the competent cells BL21*(DE3). Cells were grown on minimal media with specific substrates to be extracted for GC-MS analysis in order to obtain the retention time of substrate/product and their mass spectrums.

minute. The temperature gradient was from 60°C to 240°C with an increment rate of 15°C/min for compound separation for a total time of 14 minutes. Product identification was achieved by using the Wiley MS spectral library (8th edition) and Agilent Chemstation.

3.9 LB400 Growth Curves Using Different Carbon Sources

B. xenovorans LB400 was tested for its ability to grow on various hydrocarbons as its sole carbon source. The substrate was provided in the vapor phase up to saturation. The strain was grown overnight at 30°C at 200 rpm in MSB liquid media supplemented with the 20 mM succinate. One-hundred µL were transferred to 250 mL flasks containing 50 mL of MSB liquid minimal media supplemented with the appropriate hydrocarbon source. The chemicals tested for this assay were: *p*-cymene, *n*-butylbenzene, ethyltoluene, *m*-xylene, *p*-xylene, pentylbenzene, toluene, isobutylbenzene, *tert*-butylbenzene, *sec*-butylbenzene, propylbenzene, 4-ethyltoluene, 1,4-diisopropylbenzene and *tert*-butyltoluene. Cultures were grown at 30°C and 200 rpm. Growth was supported for the required amount of time (varies depending on the substrate) and samples were taken at different time points to measure the optical density of the culture. Optical density was measured using a Beckman Coulter DU800 Spectrophotometer at a wavelength of 600 nm. Samples were measured every 8 hours.

3.10 RNA Isolation

B. xenovorans LB400 strain was grown overnight at 30°C at 200 rpm in MSB liquid media supplemented with the 20 mM of succinate as the carbon source. One

hundred μL of the overnight culture was inoculated into fresh medium containing the carbon source of interest. Cells were collected when the optical density (600 nm) reached an 0.2-0.3. One mL of the cells (from the desired carbon source) was centrifuged at 4°C at 8 K rpm for 10 minutes. Cell pellets were used for the RNA extraction. RNA was extracted using the RNeasy Mini kit (Qiagen, Valencia, CA) with the following modifications. The bacterial pellet was resuspended in 1.0 mL of TRIzol reagent (Invitrogen, Carlsbad, CA). Two hundred μL of chloroform (CHCl_3) was added to the sample and mixed by inverting the vials and incubated at room temperature for 5 minutes. Three hundred fifty μL of the sample (supernatant only without touching any debris on the interface) were transferred to a new and RNase free tube containing 350 μL of 70% ethanol. The mixture was added to the RNeasy mini column and the manufacturer's recommendations for extraction were followed. Samples were treated with DNase for one hour to remove any genomic DNA left in the sample and RNA was eluted with 50 μL of RNase free water.

3.11 RT-PCR

An RT-PCR approach was used in this study to determine gene expression when LB400 was grown on the different hydrocarbon sources. RNA from LB400 growing on different carbon sources was used for the RT-PCR reactions. The selected primer sequences used for the RT-PCR reaction are shown in Table 3.1. The *gyrB* gene was used as a positive control. Primers for a section of the hydroxylase component of putative cymene monooxygenase gene were used to determine whether mRNA was synthesized during cell growth on a given substrate to provide information to determine if

the enzyme is involved in the growth physiology of LB400. Primers for the *bphA1* gene were also used in this experiment to determine if the growth capabilities of the strain are conferred by the biphenyl pathway. Primers were added to a final concentration of 0.6 μ M. For RT-PCR, a total of 80 ng of extracted RNA was used in a final volume reaction of 25 μ L. RT-PCR reactions were performed using the Qiagen One Step RT-PCR kit (Qiagen, Valencia, CA). The PCR reaction cycle requires an initial step of 30 min at 50°C (reverse transcription reaction), then activation of the HotStarTaq DNA Polymerase takes place at 95°C for 15 minutes for inactivation of the reverse transcriptase, denaturation and linearization of the cDNA. Then, 30 additional cycles were performed as follows: 1 min at 95°C, 1 min at 50°C and 1 min at 72°C, and a final extension step of 10 min at 72°C, in an ABI GenAmp PCR System 9700 thermocycler. RT-PCR products were loaded on a 1% agarose gel and run for 1 hour at 80 volts. The gel was stained in a 1X solution of SYBR Green (Invitrogen) in TAE buffer for approximate one hour at room temperature and scanned using a Storm 860 Scanner (Amersham Biosciences) with blue 254 nm filter at the high voltage setting.

4. Results and Discussion

4.1 Sequence and Phylogenetic Analysis of Putative Cymene Monooxygenase from *B. xenovorans* LB400

The complete genome sequence of *B. xenovorans* LB400 was recently published (Chain *et al.* 2006). It contains a DNA region closely related to the cymene monooxygenase from *P. putida* F1. This enzyme has been identified as a putative cymene monooxygenase by sequence analysis, but no further research has been done on this putative protein. Phylogenetic and sequence analysis of the putative enzyme can provide information about its function and evolutionary relationship to other enzymes and probably about its catalytic mechanism. Cloning and expression of this novel enzyme is essential for its proper classification and function determination. The intention in this chapter is to perform a comparative analysis of the putative *B. xenovorans* LB400 cymene monooxygenase and the cymene monooxygenase from *P. putida* F1. In addition, I will determine the phylogenetic relationship of the putative cymene monooxygenase to other enzymes and use computer-based algorithms to identify conserved regions among hydroxylase enzymes that have already been described.

4.1.1 Comparative Analysis of the Genes Involved in the *p*-Cymene Degradation in *P. putida* F1 and the Genes Identified in *B. xenovorans* LB400

Sequence analysis of the putative *B. xenovorans* LB400 cymene monooxygenase reveals that it is 77% identical and 85% similar to the already described cymene monooxygenase from *P. putida* F1. I intend to compare the protein sequence of these

two enzymes and their upstream and downstream genes to determine if there is any similarity between the genes and how they are organized.

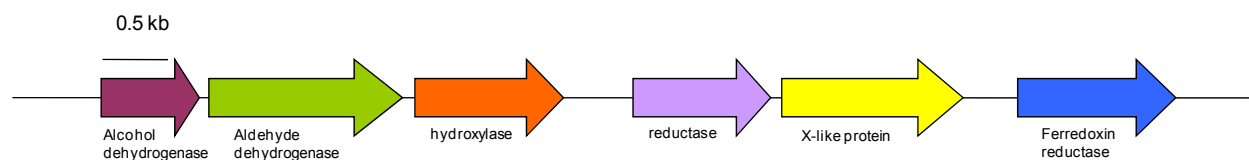
The hydroxylase component of the cymene monooxygenase, CymAa, is a protein of 376 amino acids (Eaton 1997). A homologous gene, with the same number of amino acids was found in *B. xenovorans* LB400, encoding the hydroxylase component of a putative cymene monooxygenase. The *cym* operon of *P. putida* F1 also includes the reductase component, CymAb, which contains 349 amino acids (Eaton 1997). The putative CymAb found in LB400 also contains 349 amino acids and its DNA sequence is 77% identical to the CymAb found in *P. putida* F1. In addition, LB400 encodes a protein of 253 amino acids with 77% DNA identity to CymB, a 235 amino acid alcohol dehydrogenase found in *P. putida* F1 (Eaton 1997). Also, we have identified a gene coding for a protein predicted to be a *p*-cumate aldehyde dehydrogenase. In *P. putida* F1 this protein is known as CymC (Eaton 1997) and its DNA sequence is 77% identical to the protein found in *B. xenovorans* LB400. The *cym* operon also encodes for three additional proteins that have not been identified in *B. xenovorans* LB400. These are an outer membrane protein (CymD), an acetyl CoA synthetase (CymE) with unidentified function and another outer membrane protein (CymR). In contrast to the *cym* operon of *P. putida* F1, LB400 encodes for an X-like protein and a putative ferredoxin reductase. Figure 4.1 shows the gene organization found in the *cym* operon from *P. putida* F1 and *B. xenovorans* LB400. Comparison of the hydroxylase and reductase genes coding for the cymene monooxygenase from *P. putida* F1 and the putative cymene monooxygenase from *B. xenovorans* LB400 reveal that the genes are organized in a similar fashion. The proteins CymD, CymE and CymR encoded by the *cym* operon from *P. putida* F1 were

not identified close to the putative cymene monooxygenase genes. However, we were able to identify a putative *p*-cumin alcohol dehydrogenase and a putative *p*-cumin aldehyde dehydrogenase, which suggest that these enzymes are responsible for catalysis of the substrate to its alcohol and aldehyde form, respectively.

4.1.2 Phylogenetic Analysis of the Novel Monooxygenase from *B. xenovorans* LB400

In order to determine the evolutionary relationship of the putative cymene monooxygenase to other non-heme diiron enzymes, we generated a phylogenetic tree of the enzyme by parsimony analysis. The dendrogram shown in Figure 4.2 identifies the phylogenetic relationships of the putative cymene monooxygenase found in *B. xenovorans* LB400 to other monooxygenases described, focusing on alkane and xylene monooxygenases and on the cymene monooxygenase from *P. putida* F1, which contains a high percent of identity and similarity to the novel enzyme described in this study. Phylogenetic analysis reveals that the putative cymene monooxygenase from *B. xenovorans* LB400 is evolutionarily related to the cymene monooxygenase found in *P. putida* F1. *P. putida* F1 CymAa and the putative cymene monooxygenase group on the same node of the phylogenetic tree indicating that both are closely related. Xylene monooxygenase is evolutionary related to the putative cymene monooxygenase described in this study as well, suggesting that xylene, cymene and the putative cymene monooxygenase share a common ancestor. The dendrogram shows two main groups of enzymes. The first group is composed of the xylene and cymene monooxygenases and our putative enzyme described in this study. These enzymes hydroxylate (or are proposed to hydroxylate in the case of the putative cymene monooxygenase) the methyl group of a benzene ring. The second group is composed of the alkane monooxygenases,

Gene Organization in *B. xenovorans* LB400



Gene Organization in *P. putida* F1

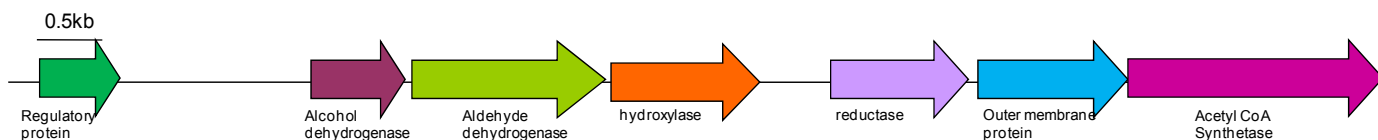


Figure 4.1: Gene organization around the cymene monooxygenase of *P. putida* F1 and the putative cymene monooxygenase from *B. xenovorans* LB400

- A. Gene organization of the putative cymene monooxygenase of *B. xenovorans* LB400 and upstream and downstream genes.
- B. Gene organization of the cymene monooxygenase operon of *P. putida* F1

which hydroxylate alkanes. The dendrogram shown on Figure 4.2 suggest that cymene and xylene monooxygenase share common ancestor and that alkane monooxygenase is related to the putative cymene monooxygenase from *B. xenovorans* LB400.

Since the putative monooxygenase is evolutionarily related to xylene, cymene and alkane monooxygenases, we decided to use computer algorithms to align the protein sequences of these enzymes and find conserved regions. Previous work on membrane-bound fatty acid desaturases (found in a variety of organisms) identified three conserved histidine motifs in their primary sequence, which were found to be essential for catalysis (Shanklin *et al.* 1994). These three histidine conserved motifs (eight histidine residues total) were also identified in bacterial alkane hydroxylases and xylene monooxygenases (Shanklin *et al.* 1994). Further research done on AlkB concluded that the eight histidine conserved residues reside in hydrophilic domains on the cytoplasmic side of the membrane and that they are essential for catalysis (van Beilen *et al.* 1992). Even though membrane-bound desaturases and membrane-bound hydroxylases are evolutionary divergent, a consensus sequence has been found among them: $[HX_{(3 \text{ or } 4)}HX_{(20-50)}HX_{(2 \text{ or } 3)}HHX_{(100-200)}HX_{(2 \text{ or } 3)}HH]$ (Shanklin *et al.* 1994). Mossbauer spectrometry on AlkB demonstrated that this motif is related to a diiron nuclear cluster (Shanklin *et al.* 1997) and that it is consistently found in enzymes that are proposed to have active sites that require iron and the activation of molecular oxygen as part of the catalytic cycle (Shanklin *et al.* 1994). Later studies based on the common motifs among integral membrane desaturases and bacterial hydroxylases found that the cymene monooxygenase from *P. putida* F1 also contains these histidine motifs (Shanklin and Cahoon 1998).

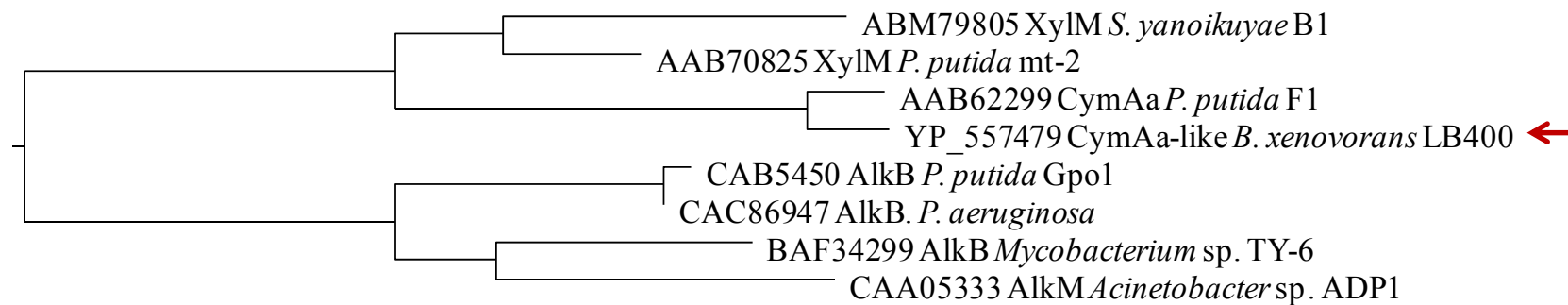


Figure 4.2: Phylogenetic tree of selected alkane and aromatic hydrocarbon hydroxylases

Accession numbers and strains are indicated on the tree. The red arrow points at the putative cymene monooxygenase analyzed on this chapter.

Although the enzymes that contain the histidine motif have different functions, it was proposed that these motifs were present on enzymes with similar chemistries (Shanklin *et al.* 1994; Shanklin and Cahoon 1998). Table 4.1 summarizes some of the proteins found to contain the common motifs, including integral-membrane desaturases, integral-membrane hydroxylases and related proteins. Given the conserved sequences found in these membrane-bound enzymes, the intention of this research was to investigate if the putative cymene monooxygenase contains the histidine-rich motif typically found on enzymes that require iron and the activation of molecular oxygen for catalysis.

We have identified by sequence analysis and protein alignment that the putative cymene monooxygenase from LB400 contains the conserved histidine motif found in membrane desaturases and identified in alkane, xylene and cymene monooxygenases (Shanklin and Cahoon 1998). Since the putative cymene monooxygenase found in LB400 is evolutionary related to alkane, cymene and xylene monooxygenase, we chose the three proteins for alignment studies (Figure 4.3). We have identified 10 conserved histidine residues in the putative cymene monooxygenase, xylene monooxygenase, cymene monooxygenase and alkane monooxygenase. The protein alignment matches the proposed consensus sequence for desaturases and hydroxylases that are membrane bound and form the Class III of diiron enzymes (Shanklin and Cahoon 1998). The proposed consensus sequence found the alkane, cymene, xylene and the putative cymene monooxygenase is $[HX_{(3)}HX_{(25)}HX_{(3)}HHX_{(139)}HX_{(2)}HH]$ (see Figure 4.3). In addition to the histidine residues proven to be essential for catalysis for the non-heme diiron membrane bound enzymes desaturases and hydroxylases, protein sequence alignment led us to identify two additional conserved histidine residues with the following consensus

Enzyme	Organism	Substrate	Product
18:0-CoA desaturase (scd/olel)	<i>Rattus/Saccharomyces</i>	Stearoyl-CoA	Oleoyle-CoA
Plastid-type Δ^9 -desaturase (desC)	<i>Synechocystis</i>	Plastid glycerolipid-18:0	Plastid glycerolipid-18:1
Plastid-type Δ^6 -desaturase (desD)	<i>Anabaena</i>	Plastid glycerolipid-18:2	Plastid glycerolipid-18:3(4)
Plant microsomal desaturase (fad2)	<i>Arabidopsis</i>	Phospholipid-18:1	Phospholipid-18:2
Plant microsomal desaturase (fad3)	<i>Arabidopsis</i>	Phospholipid-18:2	Phospholipid-18:3
Plastid-type desaturase (desA/fad6)	<i>Synechocystis/Arabidopsis</i> <i>Spinacia/Arabidopsis</i>	Plastid glycerolipid-18:1	Plastid glycerolipid-18:2
Plastid-type desaturase (fad7/8 desB)	<i>Arabidopsis/Synechocystis</i>	Plastid glycerolipid-18:2	Plastid glycerolipid-18:3
Plant microsomal Δ^6 -desaturase (BOD6)	<i>Borago</i>	Phospholipid-18:2	Phospholipid-18:3
Cytochrome b ₅ -desaturase fusion	<i>Helianthus</i>	Unidentified	Unidentified
Animal ω -3 desaturase (fat-1)	<i>Caenorhabditis</i>	Phospholipid-18(20):2	Phospholipid-18(20): 3
Animal fatty acid desaturase (MLD)	<i>Homo</i>	Unidentified	Unidentified
Sterol-C-5 desaturase (Erg3)	<i>Saccharomyces</i>	Episterol	Ergosta-5,7,24(28)-trienol
Fatty acid hydroxylase (FAH12)	<i>Ricinus</i>	Phospholipid-oleate	Phospholipid-ricinoleic acid
Fatty acid acetylenase	<i>Crepis</i>	Phospholipid-18:2	Phospholipid-crepenynic acid
Fatty acid epoxigenase	<i>Crepis/Vernonia</i>	Phospholipid-18:2	Phospholipid-vernolic acid
Sphingolipid α -hydroxylase (Fah1P)	<i>Saccharomyces</i>	Sphingolipid-26:0	Sphingolipid- α -OH 26:0
Sphingolipid 4-hydroxylase (Syr2)	<i>Saccharomyces</i>	Dihydroceramide	Phytoceramide
Alkane ω -hydroxylase/epoxidase (AlkB)	<i>Pseudomonas</i>	Alkanes/alkenes	Primary alcohols/epoxides
Xylene monooxygenase (XylM)	<i>Pseudomonas</i>	Toluene/xylenes	(methyl)Benzaldehydes
<i>p</i> -cymene monooxygenase (CymAa)	<i>Pseudomonas</i>	<i>p</i> -Cymene	<i>p</i> -Cumic alcohol
β -carotene hydroxylase (crtZ)	<i>Arabidopsis</i>	β -Carotene	Zeaxanthin
β -carotene oxygenase/ketolase (crtW)	<i>Haematococcus</i>	β -Carotene	Canthaxanthin
C-4 sterol methyl oxidase (Erg25)	<i>Saccharomyces</i>	4,4-Dimethylzymosterol	Zymosterol
Aldehyde decarbonylase (Cer1)	<i>Arabidopsis</i>	Long chain aldehydes	Alkanes

Table 4.1: Proteins containing conserved histidine motifs common to membrane desaturases, hydroxylases and related proteins

Adapted from (Shanklin and Cahoon 1998)

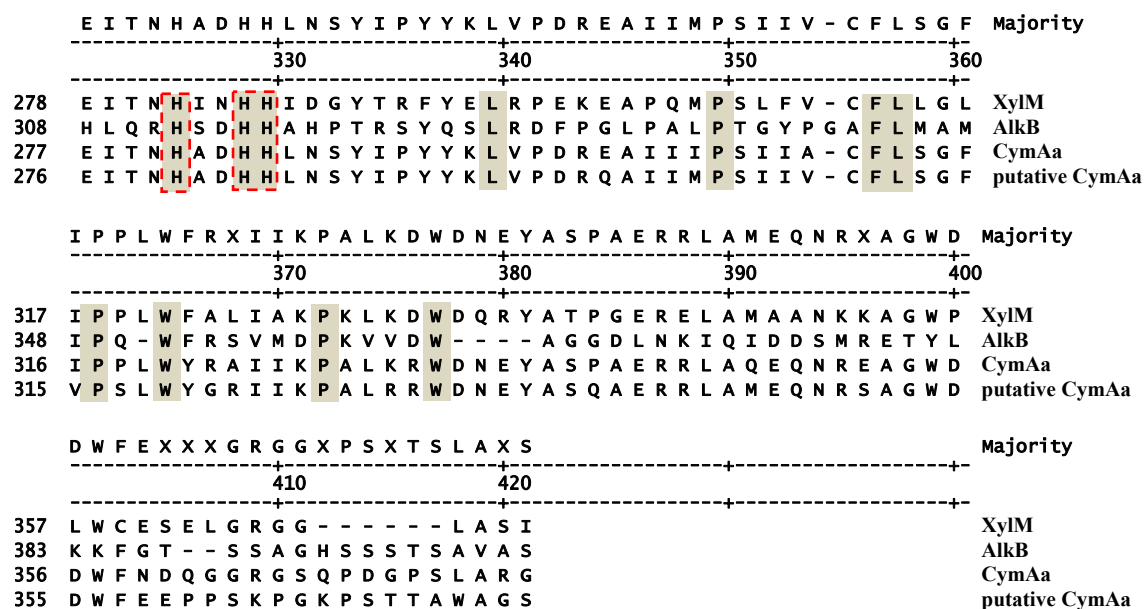


Figure 4.3: Protein sequence alignment of the xylene, alkane, cymene and putative cymene monooxygenase component

Shaded boxes represent the conserved amino acids and the red dashed lines indicate the location of the conserved histidine residues.

(xylene hydroxylase, accession number P21394; alkane hydroxylase, accession number CAB54050; cymene monooxygenase, accession number AAB62299; putative cymene hydroxylase, accession number YP_557479)

[HX₍₃₎HX₍₂₅₎HX₍₃₎HHX₍₉₉₎ HX₍₂₂₎HX₍₁₈₎HX₍₂₎HH]. Research analysis on AlkB and XylM identified 11 conserved histidine residues, of which 8 were proven to be essential for catalysis (Shanklin and Whittle 2003). We propose that the 8 conserved histidine residues found in the histidine motifs are also essential for catalysis of the enzyme described in this study, as they are essential for AlkB.

The phylogenetic and sequence analysis performed in this study suggest that the identified enzyme is a monooxygenase with an active site that require iron and the activation of molecular oxygen as part of the catalytic cycle, as was previously proposed for other enzymes containing this motif (Shanklin *et al.* 1994). Since research on the membrane-bound desaturases (Shanklin *et al.* 1994) and AlkB (Shanklin and Whittle 2003) reveal that these residues are required for enzymatic activity, we propose that the histidine motif found in the putative monooxygenase is essential for catalysis and is characteristic of diiron enzymes where the histidine residues serve as ligands for the iron atoms and coordinate the iron in place. Given that catalytic mechanisms are hard to study in membrane-bound enzymes, this analysis predict the enzyme's catalytic mechanism and the interactions required for activity.

4.2 Cloning, Protein Expression and Substrate Specificity Analysis of a Novel Monooxygenase from *B. xenovorans* LB400

In order to obtain information about what substrates are oxidized by the putative cymene monooxygenase, protein expression is required. Five different constructs were created to test the activity of the expressed recombinant putative cymene monooxygenase towards a variety of substrates. The substrates tested were: *p*-cymene, cumene, ethylbenzene, toluene, *m*-xylene, *p*-xylene (all identified substrates for the cymene monooxygenase from *P. putida* F1), *n*-dodecane, isobutylbenzene, propylbenzene, *n*-butylbenzene, 2,2-methylpropylbenzene, 1,4-diisopropylbenzene, pentylbenzene, *tert*-butylbenzene, *tert*-butyltoluene, *sec*-butylbenzene, and *m*-cymene. This chapter focuses on describing the biotransformation/protein activity of the putative cymene monooxygenase and comparing it to the cymene monooxygenase enzyme from *P. putida* F1.

4.2.1 Amplification and Cloning Putative Cymene Monooxygenase Genes

In order to identify the function and the substrate(s) utilized by the putative cymene monooxygenase from *B. xenovorans* LB400, we amplified the DNA region containing the hydroxylase and the reductase genes using Herculase Enhanced DNA Polymerase (Stratagene). PCR primers were designed with the addition of a BamHI site at the N-terminal end of the hydroxylase component to facilitate subsequent transfer to the appropriate vector. PCR amplification using Taq DNA Polymerase was not successful, even though different modifications such as the variation in the extension cycle (increasing to longer extension periods up to 7 minutes) and variation of the PCR mixture

(increasing the MgCl_2 concentration, addition of DMSO or the addition of PCR enhancer) were done. A possibility for these results could be the very high G + C content of the sequence in the region of interest. The product was cloned into the vector pGEM-T to create the construct CMOpGEM-T (Figure 4.4). The multiple cloning site of the pGEM-T vector and the added BamHI site at the N-terminal of the hydroxylase component of the enzyme are required for the transfer of the cymene monooxygenase genes to an appropriate expression vector by restriction digestion.

Since mutations might occur spontaneously during the PCR cycles, sequencing of the PCR product is required to insure against any effect on the protein function and substrate specificity. Sequencing data confirmed that the target DNA had no mutations or modifications, indicating that the polymerase used facilitated the amplification of the difficult target (GC-rich sequence) with greater sensitivity and as a proofreading enzyme. Five different constructs were tested in order to select which expression system works best expressing the putative cymene monooxygenase. In addition to the construct CMOpGEM-T (under the control of the T7 promoter), four additional constructs were created to measure biotransformations catalyzed by the enzyme by gas chromatography/mass spectrometry (GC-MS). The selected expression vectors are under the control of different promoters in order to vary the amount of protein being produced by the different expression systems. This is required because in some cases, over expression of the protein could have a toxic effect to the cells or cause the formation of inclusion bodies. In other cases, too little expression could not be enough for the enzymatic analysis. These different constructs will help elucidating which enzyme/promoter combination works best for the expression of our novel enzyme.

A

Plasmid tested	Protein Expression Activity
CMO pCR2.1	n.d.
CMOpQE-30	see Figure 4.6
CMOpTrc99A	n.d.
CMOpET101-D/TOPO	See Figure 4.5B
CMOpGEM-T	n.d.

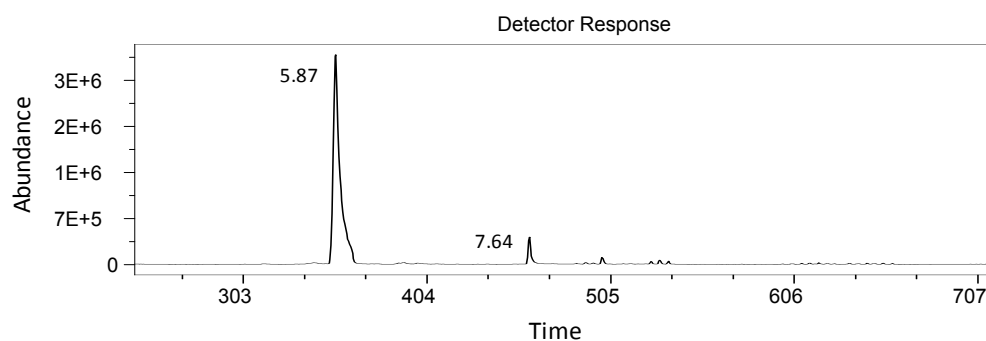
B **CMOpET101-D /TOPO + *p*-cymene**

Figure 4.4: PCR amplification and cloning of the putative cymene monooxygenase

- Amplification of the 2.7 kilobase pairs product of the hydroxylase and reductase gene.
- TA cloning of the putative genes into vector pGEM-T including the multiple cloning site for transfer into the expression vector pQE-30.

The vector pTrc99A (Amersham Pharmacia Biotech) was used to create the CMOpTrc99A construct. It is an *E. coli* expression vector under the control of the *trc* promoter which contains the -35 region of the *trp* promoter together with the -10 region of the *lac* promoter (Brosius *et al.* 1985). The construct CMO-pET101-D was created by using the vector pET101-D/TOPO (Invitrogen). Since this vector does not require the (A) overhang essential for TA cloning, cloning into this directional vector requires the addition of a CACC sequence at the N-terminal end of the PCR product to facilitate ligation of monooxygenase and reductase genes to the pET101-D/TOPO vector directionally. This construct is under the control of the T7 promoter, which contains an upstream sequence that acts as a translational enhancer to make protein synthesis very efficient. Another construct designed for protein expression was CMOpQE-30. This construct is under the control of the T5 promoter, which provides high-level protein expression regulated by double *lac* operator repression. In addition, the construct CMOpCR2.1 was created, which is also under the control of the T7 promoter. All these different vectors and promoters were used to test protein expression and substrate specificity analysis of the putative cymene monooxygenase from *B. xenovorans* LB400.

4.2.2 Expression of *B. xenovorans* LB400 putative cymene monooxygenase

All the created constructs (CMOpQE-30, CMOpCR2.1, CMOpET101-D/TOPO, CMOpTrc99A and CMOpGEM-T) were tested for protein activity, measuring transformation of product to substrate rates by GC-MS. Figure 4.5 summarizes the GC-MS data detecting the protein expression of the putative monooxygenase when using

A

Plasmid tested	Protein Expression Activity
CMO pCR2.1	n.d.
CMOpQE-30	see Figure 4.6
CMOpTrc99A	n.d.
CMOpET101-D/TOPO	See Figure 4.5B
CMOpGEM-T	n.d.

B CMOpET101-D /TOPO + *p*-cymene

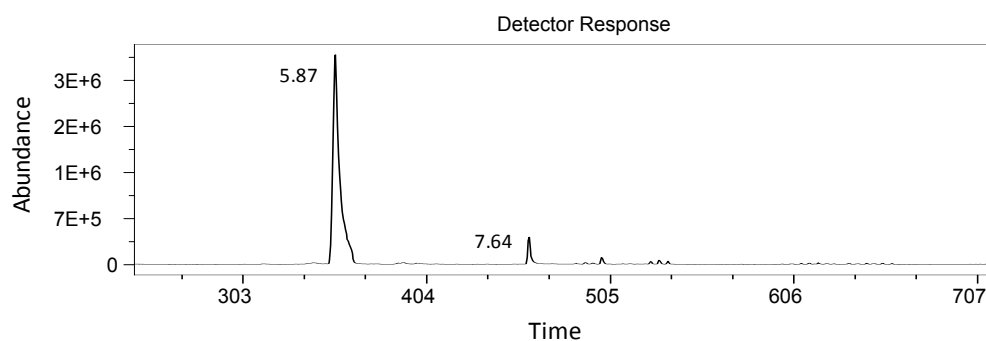


Figure 4.5: Analysis of enzyme activity using different expression vectors

- A. Represents all the constructs created and protein activity detected measured by GC-MS.
- B. Example of low biotransformation activity by the construct CMOpET101-D.
n.d.-not detected

different expression vectors. The constructs CMOpCR2.1, CMOpTrc99A and CMOpGEM-T showed no detectable substrate to product biotransformation activity.

Since the expression could overcome these problems and provide better results. This vector contains possibility of protein toxicity with protein over expression exists, this vector was tested because lower levels of protein the T7 and the SP6 promoters. However, it does not contain a ribosome binding site, resulting in poor protein translation and lower amounts of protein production. Plasmid CMOpCR2.1 showed no detectable biotransformation indicating no protein activity. The construct contains the *lac* ribosome binding site that is driven by the strong hybrid *trp/lacZ* promoter (*trc*) under the control of the *lac^q* allele of the *lac* repressor gene (Lomri *et al.* 1993). Attempts to express the putative cymene monooxygenase using the pTrc99A vector failed. Very low biotransformation activity was detected when the CMOpET101-D/TOPO construct was tested (Figure 4.5B). The low biotransformation levels that resulted from the analysis suggest that the robust T7 promoter might cause toxic levels of protein expression. When the CMOpQE-30 construct was tested, biotransformation of substrate to product was detected. Clear and high peaks were obtained, indicating that this construct is suitable for the substrate specificity analysis of the putative cymene monooxygenase. The differences in the biotransformation rate may be related to differences in the promoters used for this analysis. The reason the T5 promoter worked better than the other promoters tested could be related to the double *lac* operator repression module present in this system that tightly regulates the high-level expression of recombinant proteins in *E. coli*. As described by the manufacturer, the protein synthesis in this system is effectively

blocked in the presence of high levels of *lac* repressor and the stability of cytotoxic constructs is enhanced.

4.2.3 Substrate Specificity Analysis of the Putative Cymene Monooxygenase by GC-MS

Given the high percent identity (77%) and similarity (85%) between the putative cymene monooxygenase from *B. xenovorans* LB400 and the cymene monooxygenase from *P. putida* F1, I decided to investigate if these two enzymes catalyze the oxidation of similar substrates. Initial substrates were chosen based on the work done by Nishio et al. in 2001 describing the substrate specificity of the cymene monooxygenase from *P. putida* F1. Cymene monooxygenase from *P. putida* F1 catalyzed the oxidation of the substrates *p*-cymene, 4-ethyltoluene, styrene, *m*-xylene, *p*-xylene, 4-chlorostyrene, 4-(methylthio)toluene, 3-chlorotoluene, 4-chlorotoluene, 4-fluorotoluene and 4-nitrotoluene (see Figure 2.6 for chemical structures). Their findings indicate an enzyme preference for substrates substituted by an alkyl group or heteroatom at the *para* position followed by substitutions at positions *meta* and subsequently *ortho* (Nishio et al. 2001). The fastest oxidation rate identified was toward the substrate 4-chlorotoluene, which is over 2 fold higher than the oxidation rate of *p*-cymene (Nishio et al. 2001). The intention of this chapter was to determine if the putative cymene monooxygenase from *B. xenovorans* LB400 catalyzes similar oxidations to those catalyzed by the cymene monooxygenase found in *P. putida* F1.

Cymene monooxygenase from *P. putida* F1 oxidizes *p*-cymene to its alcohol form 4-isopropylbenzyl alcohol. GC-MS analysis revealed that the putative cymene monooxygenase is capable of oxidizing *p*-cymene, but the product formed in this reaction

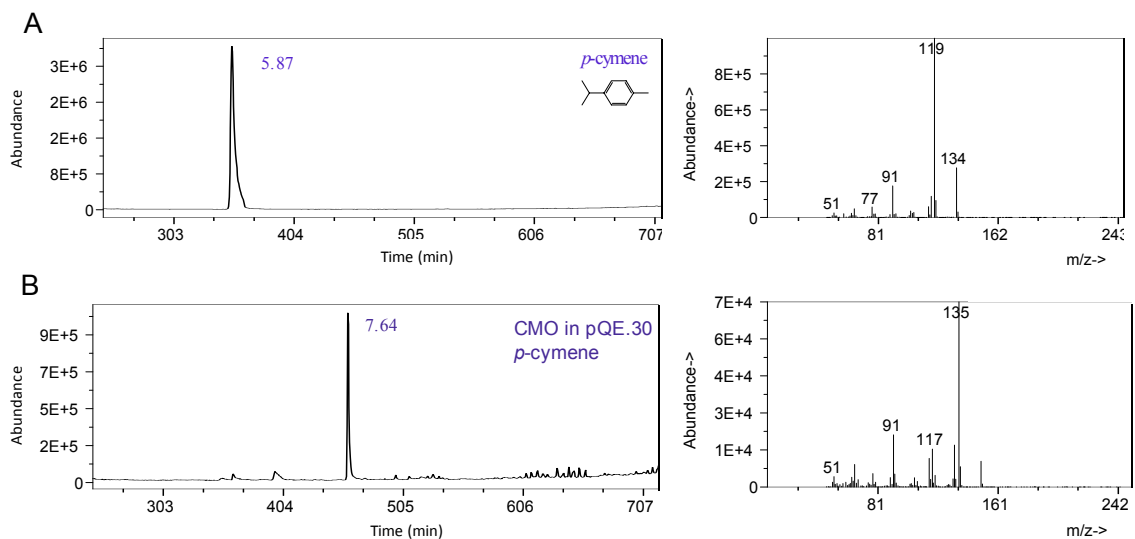


Figure 4.6: Oxidation of *p*-cymene catalyzed by the putative cymene monooxygenase from *B. xenovorans* LB400

- A. Gas chromatogram of substrate *p*-cymene with peak eluting at 5.87 minutes and its mass spectrum showing a molecular weight of 134 g/mole.
- B. Gas chromatogram of biotransformation reaction showing the substrate *p*-cymene with peak eluting at 5.87 minutes and the product of the reaction with a peak eluting at 7.64 minutes. Product mass spectrum is included showing a molecular weight of 150 g/mole.

was identified as *p*-cymen-8-ol, and not as 4-isopropylbenzyl alcohol as originally expected (Figure 4.6 and 4.7). *p*-Cymen-8-ol is a natural oil often found as a component of plant leaves. The retention time of standard 4-isopropylbenzyl alcohol was 7.01 minutes, while the obtained product has a retention time of 7.64 minutes (see Figure 4.7). The mass spectrum of the 4-isopropylbenzyl alcohol standard does not match the spectrum of the obtained product. The GC-MS library identified the product as *p*-cymen-8-ol by a perfect library match, indicating that they are different products (Figure 4.7). A commercially available standard for *p*-cymen-8-ol was run and its retention time and mass spectrum matched the reaction product confirming the data obtained from the GC-MS library. These results indicate that an oxidation reaction occurred, but instead of occurring at the methyl group of *p*-cymene, it occurs at the central carbon of the isopropyl group of the substrate *p*-cymene (Figure 4.8). This reaction is rather unusual because the attack occurs at a sterically hindered group. Although this monooxygenase is active toward the substrate *p*-cymene, its chemical isomer with orientation at the *meta* position (*m*-cymene) did not show substrate activity, perhaps because substitutions at the *meta* position cause steric hindrance that prevents enzyme catalysis. These results suggest that the enzyme shows preference toward substrates with *para* orientation.

In order to confirm that the GC-MS results are not affected by the *E. coli* cells, Na/KPO₄ buffer or by the expression vector itself, negative controls were analyzed. Figure 4.9 shows the GC-MS results for the Na/KPO₄ buffer (no cells added) incubated with the substrate *p*-cymene, pEQ-30 vector with no insert incubated with the *p*-cymene substrate and the expressed *E. coli* cells without aromatic hydrocarbons provided as substrate. The clear and sharp peak at 5.87 minutes in Figures 4.6A and 4.6B was

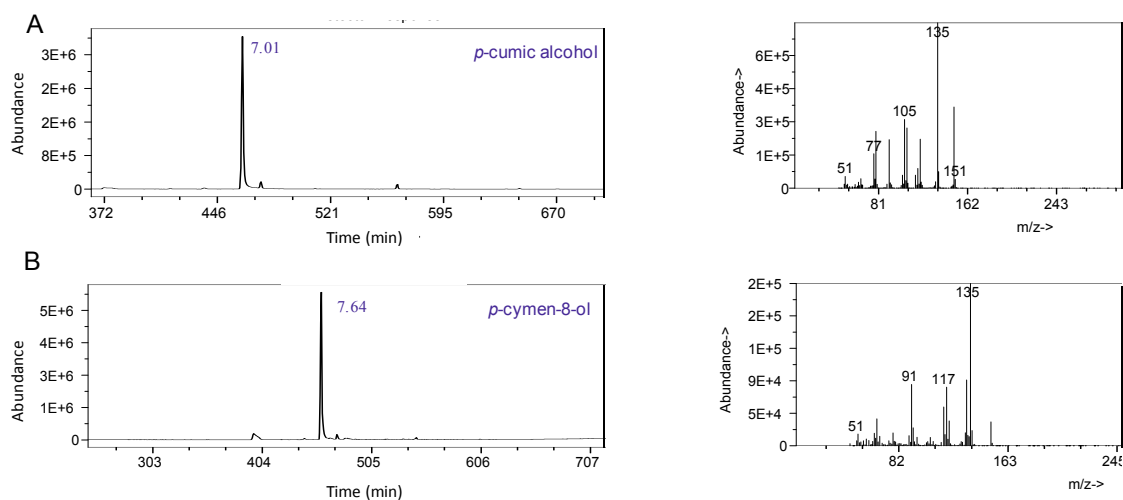


Figure 4.7: Gas chromatograms and mass spectrum comparisons of standards *p*-cumic alcohol and *p*-cymen-8-ol

- A. Gas chromatogram of standard *p*-cumic alcohol with peak eluting at 7.01 minutes and its mass spectrum showing a molecular weight of 150 g/mole.
- Gas chromatogram of standard *p*-cymen-8-ol with peak eluting at 7.64 minutes and its mass spectrum showing a molecular weight of 150 g/mole.

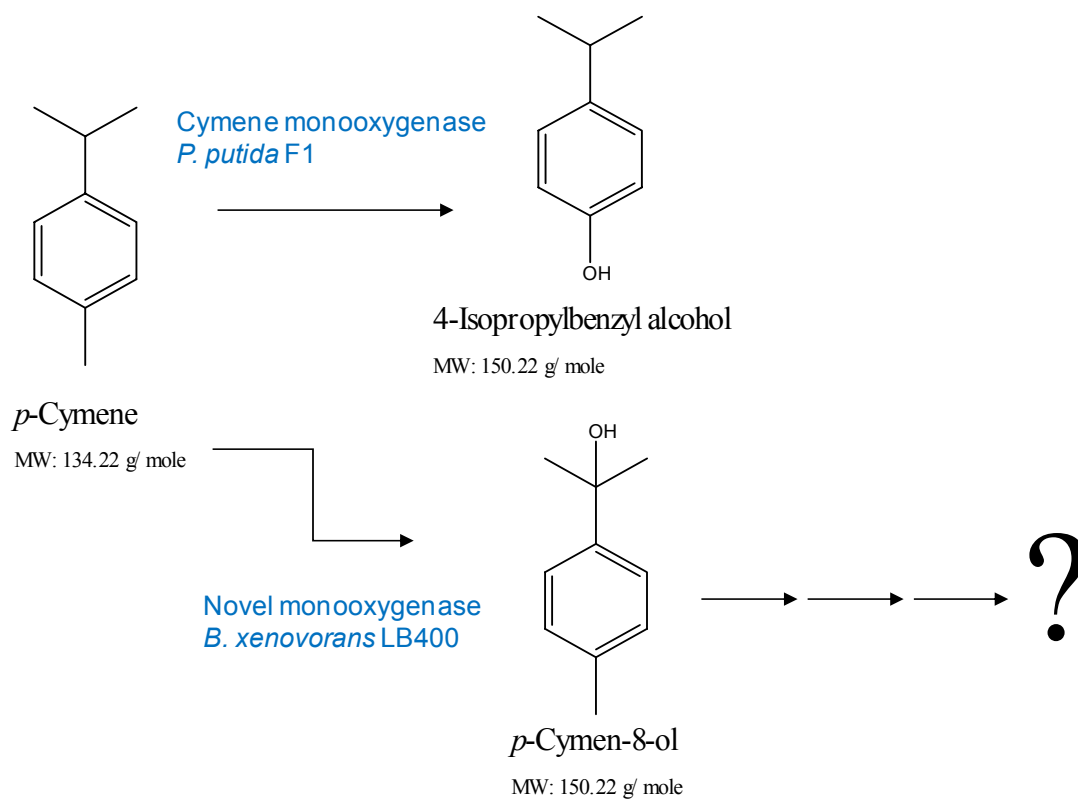


Figure 4.8: Proposed mechanism for the degradation of *p*-cymene by the novel monooxygenase

Cymene monooxygenase catalyzes a hydroxylation at the methyl group substituent of the *p*-cymene substrate. The novel monooxygenase catalyzes the hydroxylation of the isopropyl group (at the carbon located next to the benzene ring)

identified as the substrate *p*-cymene and no product or substrate were detected when the plasmid was incubated without the substrate (Figure 4.9 B). These data confirm that the peaks present in the analysis are the result of the novel monooxygenase activity, and not the result of *E. coli* cells or the vector.

Since the hydroxylation of the carbon occurs at an unexpected position, it was our intention to determine the substrate range of this novel enzyme and see if it oxidizes substrates similar to those oxidized by cymene monooxygenase found in *P. putida* F1. In addition to *p*-cymene, I analyzed the substrates cumene, isobutylbenzene, *n*-butylbenzene, propylbenzene, *n*-dodecane, benzene, 2,2-methylpropylbenzene, 1,4-diisopropylbenzene, pentylbenzene, ethylbenzene, *tert*-butyltoluene, *tert*-butylbenzene, *sec*-butylbenzene, toluene, *m*-xylene, *p*-xylene, *m*-cymene, and 4-ethyltoluene (see Figure 4.10 for chemical structures). The alkyl groups of these substrates have different lengths, different chemical orientation and some contain a methyl group in the *para* position. The lack of the methyl group on some of the substrates is essential to confirm that the hydroxylation occurs at the alkyl group of the substrate (chain length of two carbons or more) and not at the methyl substituent. Based on these initial results (biotransformation of *p*-cymene to *p*-cymen-8-ol), my intention was to determine the enzyme activity towards a variety of substrates, in order to determine if the oxidation reaction always occurs at the same carbon regardless of the substrate being utilized.

Since the monooxygenase oxidizes the carbon next to the benzene ring, we decided to test the enzyme activity toward the substrate cumene, which lacks the methyl group at the *para* position present in *p*-cymene. Our data confirms that the enzyme is active toward

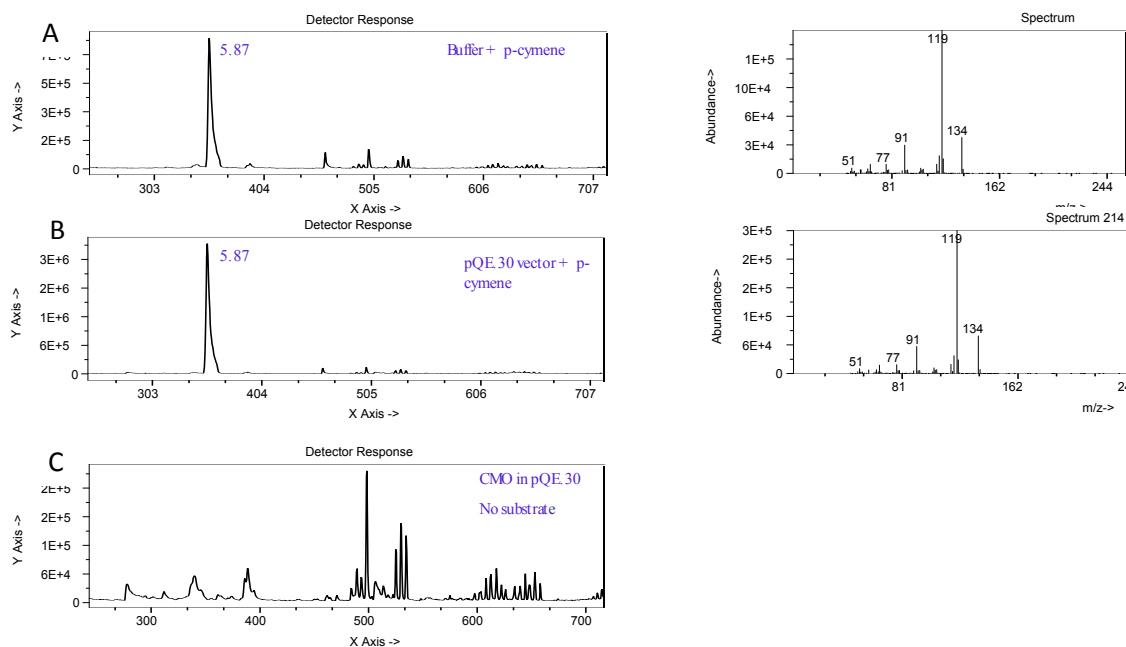


Figure 4.9: Experimental controls used in the GC-MS analysis

- Gas chromatogram of phosphate buffer incubated with *p*-cymene. The peak eluting at 5.87 minutes is identified as the substrate.
- Gas chromatogram of cloned vector with no insert into BL21 cells incubated *p*-cymene. The peak eluting at 5.87 minutes is identified as the substrate.
- Gas chromatogram of the CMOpEQ-30 construct with no substrate.

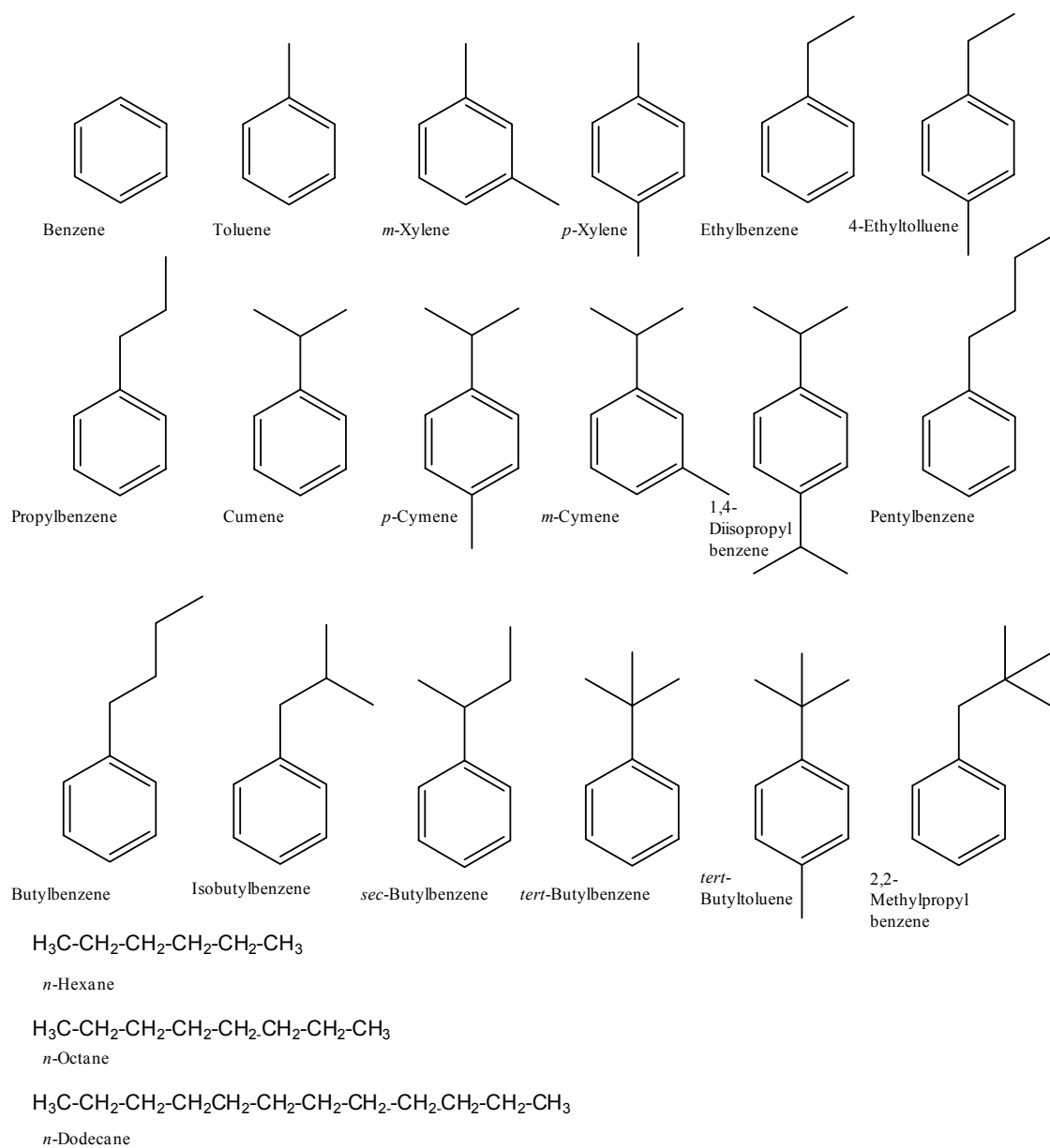


Figure 4.10: Chemical structures of substrates utilized in the substrate specificity analysis of *B. xenovorans* LB400

this substrate. The substrate retention time is 4.65 minutes (Figure 4.11A). GC-MS results show a product with retention time of 6.58 minutes. Mass spectrometry analysis revealed that the product molecular weight of the product is 136 g/mole, while the molecular weight of the substrate is 120 g/mole (Figure 4.11B). The addition of 16 g/mole to the molecular weight of the substrate indicates the addition of an oxygen atom. Contrary to the cymene monooxygenase (which shows no activity towards the substrate cumene) (Nishio *et al.* 2001), our novel monooxygenase catalyze the oxidation of cumene, but which carbon is hydroxylated is still unidentified. We propose that the hydroxylation of cumene occurs at the isopropyl group of the substrate, at the carbon proximal to the benzene ring of cumene. Unfortunately, standards are not commercially available to confirm which carbon is hydroxylated and the GC-MS library is not capable of identifying the product of the reaction. Since the putative enzyme is active toward cumene, it indicates that the presence of the methyl group present in *p*-cymene is not required for enzymatic activity as it is required for the cymene monooxygenase found in *P. putida* F1.

The substrate 1,4-diisopropylbenzene has two isopropyl groups oriented at the *para* position. This compound is structurally similar to *p*-cymene, but the methyl group present in *p*-cymene is substituted by a second isopropyl group. Given that *p*-cymene is a substrate for the novel monooxygenase, testing if the enzyme can attack this substrate will provide more information about its specificity. Substrate specificity analysis of the substrate 1,4-diisopropylbenzene indicates that a hydroxylation reaction occurred. 1,4-Diisopropylbenzene (molecular weight of 162 g/mole) has a retention time of 7.51 minutes, while the product of the analysis has a retention time of 9.10 minutes and the

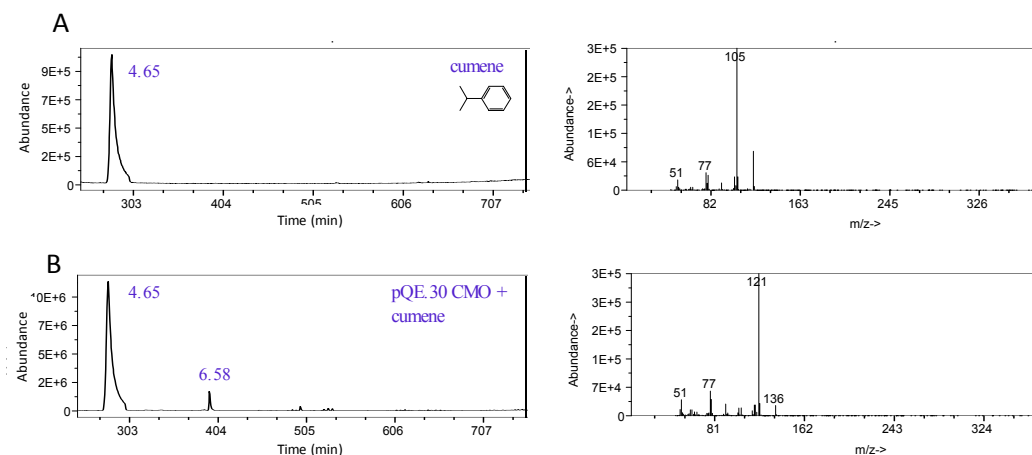


Figure 4.11: Oxidation of cumene by the novel monooxygenase

- A. Gas chromatogram of cumene with peak eluting at 4.65 minutes and its mass spectrum showing a molecular weight of 120 g/mole.
- B. Gas chromatogram of biotransformation reaction showing the substrate cumene with peak eluting at 4.65 minutes and the product of the reaction with a peak eluting at 6.58 minutes. Product mass spectrum is included showing a molecular weight of 136 g/mole.

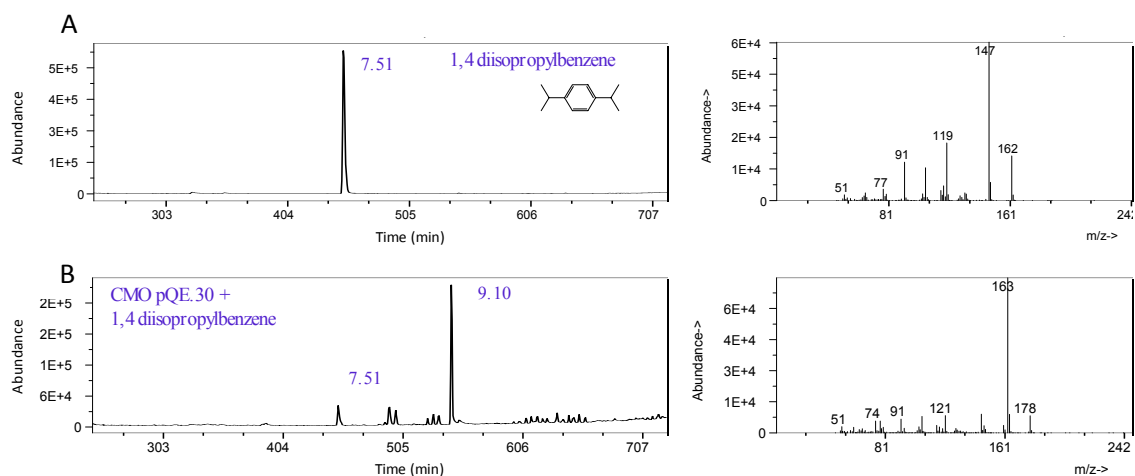


Figure 4.12: Oxidation of 1,4-diisopropylbenzene by the novel monooxygenase

- A. Gas chromatogram of 1,4-diisopropylbenzene with peak eluting at 7.51 minutes and its mass spectrum showing a molecular weight of 162 g/mole.
- B. Gas chromatogram of biotransformation reaction showing the substrate 1,4-diisopropylbenzene with peak eluting at 7.51 minutes and the product of the reaction with a peak eluting at 9.10 minutes. Product mass spectrum is included showing a molecular weight of 178 g/mole.

addition of 16 g/mole to the original substrate molecular weight indicates that one atom of oxygen was added to the substrate. Since the substrate contains two isopropyl groups available for hydroxylation at the carbon next to the benzene ring, it is uncertain how the enzyme selects which carbon to oxidize, but our data indicate that only one of the two available carbon atoms was oxidized and a dihydroxy compound was not formed. The lack of commercial standards for the proposed product and the absence of a GC-MS library match make difficult the identification of the product formed by the enzyme.

In order to determine if the substituent itself had any effect on hydroxylation, we decided to test the substrate propylbenzene which has a molecular weight of 120 g/mole. Two different peaks were observed by GC-MS, the first peak elutes at GC 5.01 minutes and correspond to the substrate propylbenzene (Figure 4.13A). The second peak, with a retention time of 7.21 minutes had a molecular weight of 136 g/mole (Figure 4.13B). This product represents the addition of an oxygen atom to propylbenzene. I propose that the attack occurs at the propyl group at the carbon located next to the benzene ring. To date, there is no commercially available standard to confirm that the suggested product was produced in the reaction and the GC-MS library was unable to correctly identify the product of the reaction.

To determine the activity of the enzyme towards substrates with variations on the alkyl group, we decided to test substrates with a butyl group attached to the benzene ring. The substrates tested in this category were *n*-butylbenzene, isobutylbenzene, *sec*-butylbenzene and *tert*-butylbenzene, all with a molecular weight of 134 g/mole. When *n*-butylbenzene was tested, the retention time for the substrate is 6.22 minutes. The GC-MS results show a product with a peak eluting at 8.25 minutes (Figure 4.14A). Mass

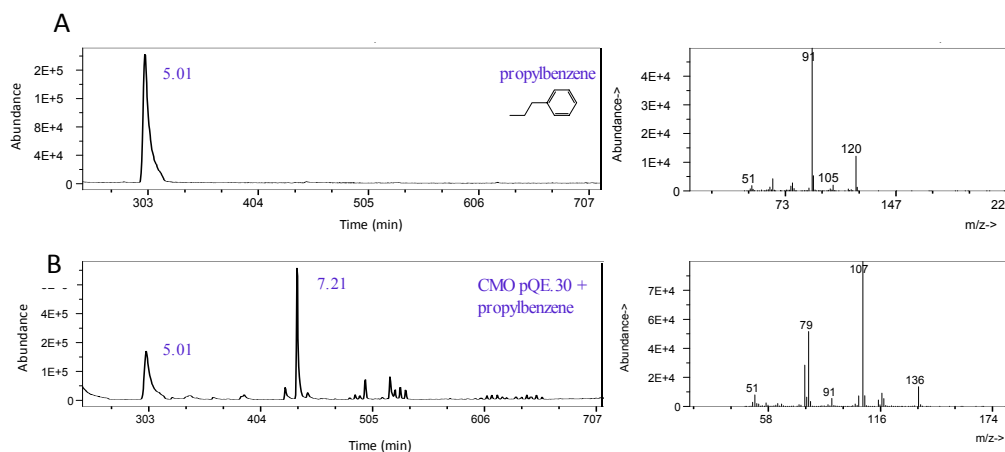


Figure 4.13: Oxidation of propylbenzene by the novel monooxygenase

- A. Gas chromatogram of propylbenzene with peak eluting at 5.01 minutes and its mass spectrum showing a molecular weight of 120 g/mole.
- B. Gas chromatogram of biotransformation reaction showing the substrate propylbenzene with peak eluting at 5.01 minutes and the product of the reaction with a peak eluting at 7.21 minutes. Product mass spectrum is included showing a molecular weight of 136 g/mole.

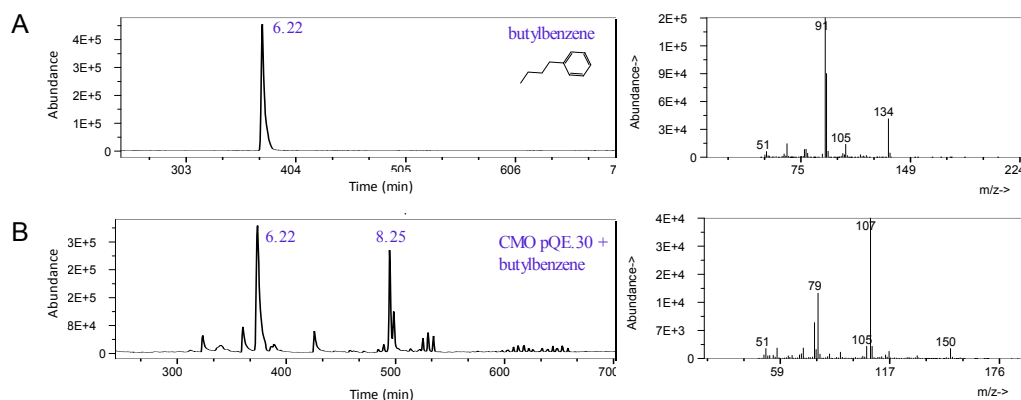


Figure 4.14: Oxidation of *n*-butylbenzene by the novel monooxygenase

- A. Gas chromatogram of *n*-butylbenzene with peak eluting at 6.22 minutes and its mass spectrum showing a molecular weight of 136 g/mole.
- B. Gas chromatogram after of biotransformation reaction showing the substrate *n*-butylbenzene with peak eluting at 6.22 minutes and the product of the reaction with a peak eluting at 8.25 minutes. Product mass spectrum is included showing a molecular weight of 150 g/mole.

spectrum of this peak reveals that the product has a molecular weight of 150 g/mole, indicating the addition of an oxygen atom to *n*-butylbenzene (Figure 4.14B). The peak mass spectrum does not match any of the compounds of the GC-MS. In order to identify the products of the reaction advanced identification techniques are required. Analysis using the substrate isobutylbenzene indicates that the enzyme can hydroxylate the substrate producing a compound to produce a product with molecular weight of 150 g/mole. Isobutylbenzene has a retention time of 5.66 minutes and its product has a retention time of 8.68 minutes (Figure 4.15A), which does not match any of the compounds available in the GC-MS library. The molecular weight of the product is 150 g/mole (Figure 4.15B) indicating the addition of an oxygen atom to the substrate. When *sec*-butylbenzene was used as a substrate for the novel enzyme, enzyme activity was detected. The molecular weight of the substrate is 134 g/mole and the product molecular weight is 150 g/mole. The obtained reaction product has a peak eluting at 7.63 minutes, while the substrate retention time is 5.71 minutes (Figure 4.16A). Mass spectrum analysis shows that the product has a molecular weight of 150 g/mole (Figure 4.16B), indicating the addition of an oxygen atom to the *sec*-butylbenzene substrate. Enzyme activity was not detected when *tert*-butylbenzene (retention time of 5.49 minutes) was used as a substrate (data not shown). A possible reason for this likely result might be that the carbon proposed to be hydroxylated by the novel enzyme is no longer available for hydroxylation because it is saturated by three methyl groups and a benzene ring. If this is the case, this data confirm that the novel monooxygenase catalyzes the hydroxylation of

the

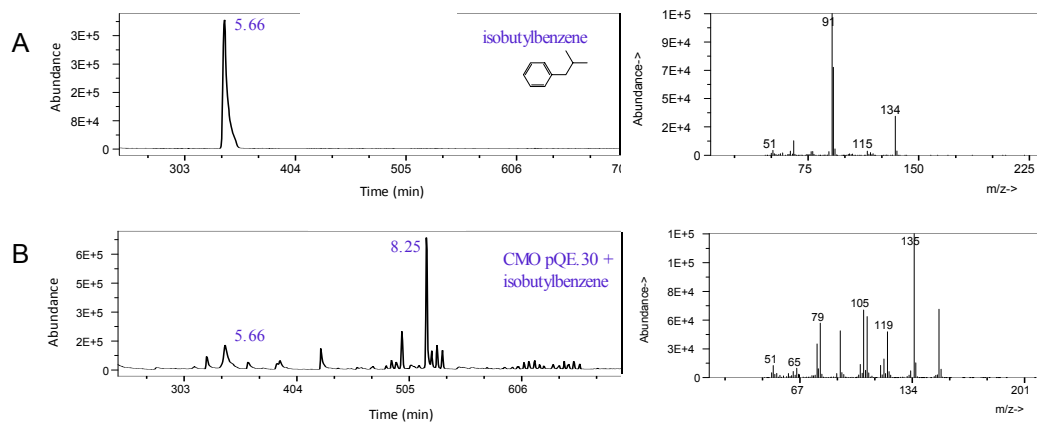


Figure 4.15: Oxidation of isobutylbenzene by the novel monooxygenase

- Gas chromatogram of isobutylbenzene with peak eluting at 5.66 minutes and its mass spectrum showing a molecular weight of 136 g/mole.
- Gas chromatogram of biotransformation reaction showing the substrate isobutylbenzene with peak eluting at 5.66 minutes and the product of the reaction with a peak eluting at 8.25 minutes. Product mass spectrum is included showing a molecular weight of 136 g/mole.

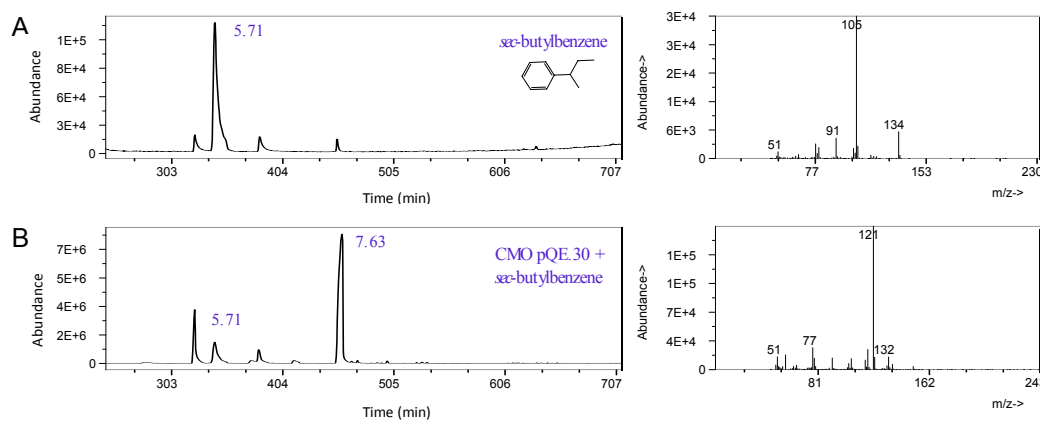


Figure 4.16: Oxidation of *sec*-butylbenzene by the novel monooxygenase

- A. Gas chromatogram of *sec*-butylbenzene with peak eluting at 5.71 minutes and its mass spectrum showing a molecular weight of 134 g/mole.
- B. Gas chromatogram of biotransformation reaction showing the substrate *sec*-butylbenzene with peak eluting at 5.71 minutes and the product of the reaction with a peak eluting at 7.63 minutes. Product mass spectrum is included showing a molecular weight of 136 g/mole. The chromatogram shows an unidentified peak of molecular weight of 94 g/mole.

substrate at the carbon of the substituent group located next to the benzyl ring and when this carbon is not available, the enzyme is not active.

Pentylbenzene was tested to investigate if a longer substituent chain can be a substrate for the novel enzyme. GC-MS data analysis did not identify any hydroxylated product when pentylbenzene was added as a substrate for the enzyme. Since no activity was detected with longer side chain (a pentyl group), we decided to try compounds with shorter chain substituents, such as ethylbenzene (which contains an ethyl group attached to benzene) and toluene (which contains a methyl group attached to benzene). No activity was detected when toluene was tested as a possible substrate for the enzyme. No enzyme activity was detected when ethylbenzene was used as a substrate either. Cymene monooxygenase from *P. putida* F1 showed no detectable activity toward these two substrates either (Nishio *et al.* 2001). To determine if the addition of a methyl group to the substrate at the *para* or *meta* position affect the specificity of the enzyme, we tested the substrates *p*-xylene and *m*-xylene. No activity was detected toward these two substrates by our novel enzyme. These results differ from the cymene monooxygenase from *P. putida* F1 which is capable of catalyzing the hydroxylation of these substrates. The substrate toluene, which only contains a methyl group attached to the benzene ring, was not identified as a substrate for the novel enzyme nor for the cymene monooxygenase enzyme from *P. putida* F1.

No activity was detected when ethylbenzene was used as a substrate for the enzyme. However, we decided to try the catalytic effect of the addition of a methyl group to the substrate (4-ethyltoluene), which is structurally similar to *p*-cymene, lacking only a methyl group. Data showed that the substrate 4-ethyltoluene was hydroxylated by

the enzyme. 4-Ethyltoluene has a molecular weight of 120 g/mole and a retention time of 5.12 minutes (Figure 4.17). The product of the reaction has a molecular weight of 134 g/mole and a retention time of 7.38 minutes (Figure 4.17). Molecular weight calculations indicate the addition of only 14 g/mole and further analysis of this product identifies it as methylacetophenone, which is the aldehyde form of the expected alcohol produced in the reaction. Previous studies for xylene monooxygenase identified the capability of this enzyme of catalyzing the multistep catalysis of toluene and pseudocumene to its alcohol, aldehyde and acid forms (Bühler *et al.* 2000). The biotransformation of the substrate to its aldehyde form suggests that the enzyme described in this study is also capable of catalyzing dehydrogenation of the alcohol product to form the aldehyde, as was previously described for the xylene monooxygenase (Bühler *et al.* 2000). Similarly, the expression of the recombinant cymene monooxygenase from *P. putida* F1 identified the formation of *p*-cuminic aldehyde, which is oxidized spontaneously to *p*-cuminic acid (Eaton 1997). Identification of the product by the GC-MS library provides supporting evidence for our hypothesis that the substitution occurs at the carbon next to the benzene ring, usually from alkyl groups containing 2-4 carbons. Since the enzyme shows no activity toward the substrate ethylbenzene, but it does toward 4-ethyltoluene, we can conclude that the presence of the methyl group at the *para* position is essential for enzyme catalysis of this substrate.

Incubation of the clone with *tert*-butylbenzene yield no product. However, when the substrate *tert*-butyltoluene (a methyl group added at the *para* position of the substrate *tert*-butylbenzene) was tested, enzyme activity was detected. The substrate *tert*-butyltoluene (molecular weight of 148 g/mole) has a retention time of 6.63 minutes. The

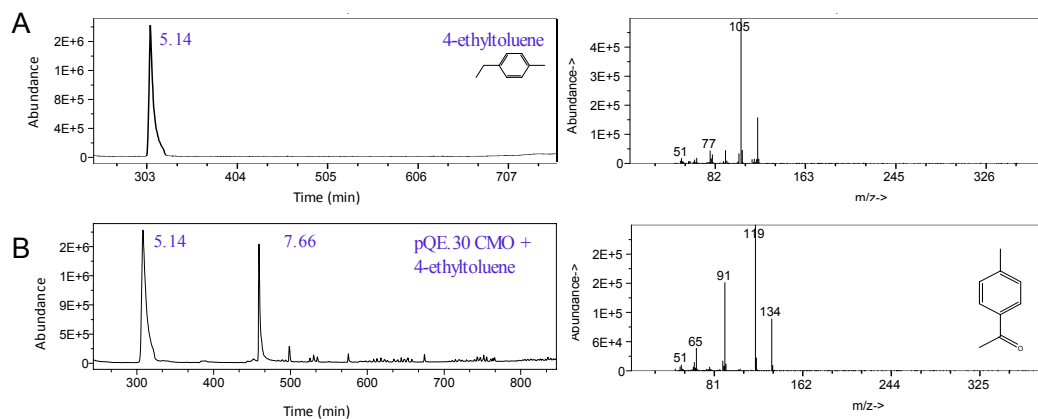


Figure 4.17: Oxidation activity of 4-ethyltoluene by the novel monooxygenase

- A. Gas chromatogram of 4-ethyltoluene with peak eluting at 5.14 minutes and its mass spectrum showing a molecular weight of 120 g/mole.
- B. Gas chromatogram of biotransformation reaction showing the substrate 4-ethyltoluene with peak eluting at 5.14 minutes and the product of the reaction with a peak eluting at 7.66 minutes. Product mass spectrum is included showing a molecular weight of 134 g/mole. GC-MS library identifies the product as methylacetophenone.

product of the substrate *tert*-butyltoluene has molecular weight of 164 g/mole and retention time of 9.35 minutes (Figure 4.18). Interestingly, the novel monooxygenase showed activity towards this substrate, when it was not expected based on our hypothesis. The carbon proposed to be hydroxylated by the enzyme is saturated in this substrate, causing steric hindrance for hydroxylation of this carbon. Since this substrate was hydroxylated, the standard 4-*tert*-butylbenzyl alcohol was run in order to determine if the hydroxylation reaction occurs at the methyl group of the *tert*-butyltoluene substrate. The standard has a retention time of 9.34 minutes and the mass spectrum patterns of the standard 4-*tert*-butylbenzyl alcohol is a perfect match for the mass spectrum of the product formed from *tert*-butyltoluene (Figure 4.19A and 4.19B confirming that the hydroxylation of *tert*-butyltoluene occurs at the methyl group of the substrate (see Figure 4.19C).

The data shown so far indicates that the enzyme has activity on substrates that have a benzene ring with alkyl substituents attached. We used the substrates *n*-hexane, *n*-octane and *n*-dodecane to determine if linear alkanes are substrates for the enzyme. No hydroxylation activity was detected when these chemicals were used, indicating that linear alkanes are not substrates for this monooxygenase. Since it seems that the benzene ring is required for the enzyme activity, we decided to test benzene as a substrate itself. Our data indicate that there is no enzyme activity toward benzene, which of course lacks any alkyl substituents. Neither toluene, nor *m*-xylene, nor *p*-xylene, which has a methyl group available for a possible hydroxylation were substrates for the enzyme, indicating that the enzyme shows a preference for longer alkyl group substituents.

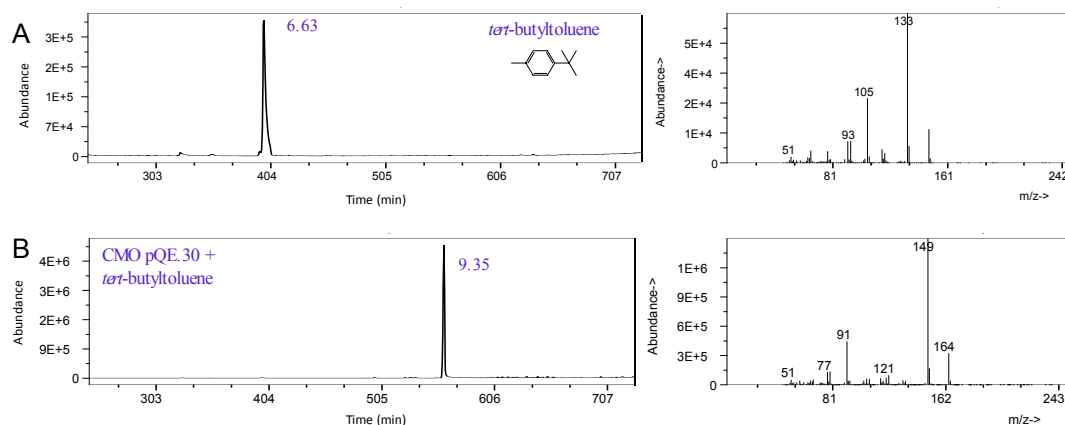


Figure 4.18: Oxidation activity of *tert*-butyltoluene by the novel monooxygenase

- Gas chromatogram of *tert*-butyltoluene with peak eluting at 6.63 minutes and its mass spectrum showing a molecular weight of 148 g/mole.
- Gas chromatogram of biotransformation reaction showing the substrate *tert*-butyltoluene with peak eluting at 6.63 minutes and the product of the reaction with a peak eluting at 9.35 minutes. Product mass spectrum is included showing a molecular weight of 164 g/mole.

Although the novel enzyme utilizes the compound *p*-cymene as a substrate, the hydroxylation reaction occurs at a different position than the attack described for the *P. putida* F1 cymene monooxygenase, which attacks the methyl group of *p*-cymene. In fact, this enzyme oxidizes substrates that lack methyl substituents, such as propylbenzene, *n*-butylbenzene, *sec*-butylbenzene and cumene, contrary to cymene monooxygenase which does not hydroxylate these substrates because it requires the methyl group substituent for activity. The collected data indicates that what was originally predicted to function as a cymene monooxygenase by bioinformatic analysis, is a novel enzyme catalyzing the unusual hydroxylation reaction of aromatic hydrocarbons at the *alpha* carbon adjacent to the benzene ring as substituent. Interestingly, *tert*-butyltoluene was identified as a substrate for this enzyme hydroxylating the methyl group at the *para* position. The novel enzyme was also shown to be capable of catalyzing a hydroxylation reaction at the methyl group when *tert*-butyltoluene was added as a substrate. It seems that the addition of a methyl group at the *para* position of *tert*-butylbenzene (*tert*-butyltoluene) changes the enzyme specificity and it becomes a substrate for catalysis. These results perhaps suggest that the extra substituent aids in binding of the substrate to the enzyme, whereas with toluene itself the K_m is too far above the concentration at saturation in water for appreciable hydroxylation to occur. Even though the attack occurs at the methyl group of *tert*-butyltoluene (contrary to the attack proposed for the other substrates), we can conclude that this attack still occurs at the carbon next to the benzene ring. It is still unidentified why the enzyme's attack occurs at the methyl group of *tert*-butyltoluene, but in *tert*-butyltoluene the *alpha* carbon is saturated and not available for catalysis. Possibly, the enzyme has some preference for hindered groups (*alpha* carbon) and not for

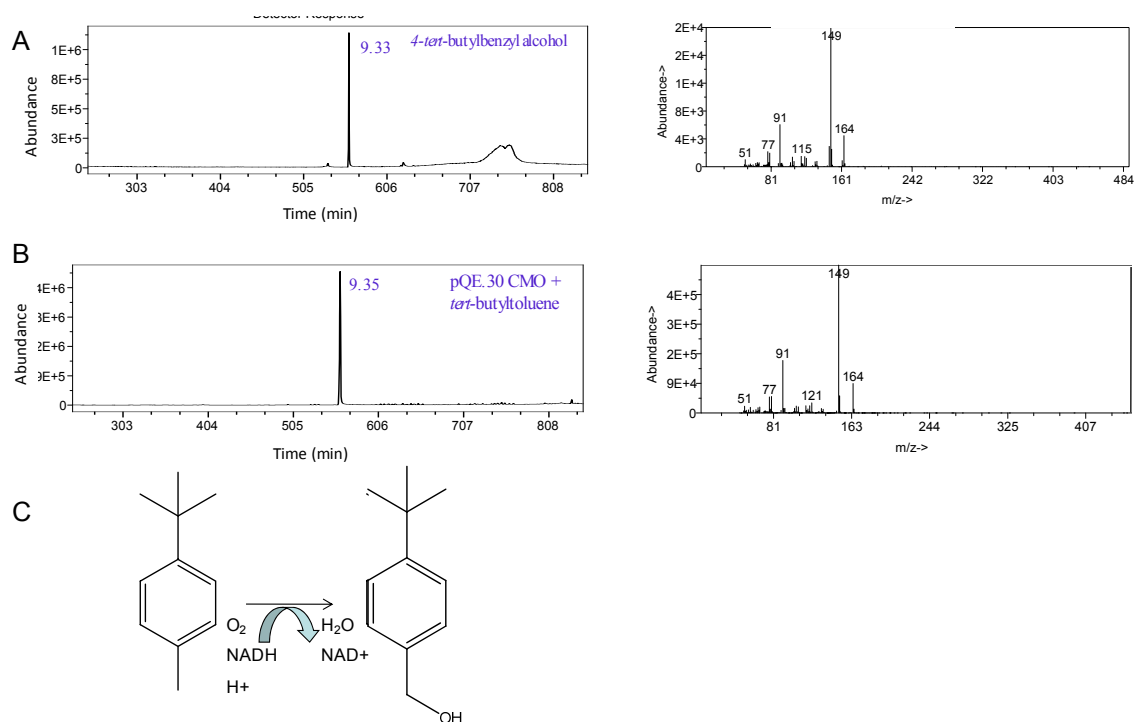
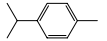

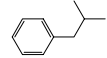
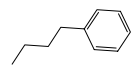
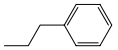
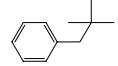
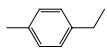


Figure 4.19: Gas chromatogram and mass spectrum comparisons between the standard 4-ethylbenzyl alcohol and the biotransformation product the substrate *tert*-butyltoluene

- Gas chromatogram of the standard 4-*tert*-butylbenzyl alcohol with elution time at 9.33 minutes and its mass spectrum with the molecular weight of 164 g/mole.
- Gas chromatogram of the product of biotransformation product of enzyme catalysis with elution time of 9.35 minutes and mass spectrum of 164 g/mole.
- Mechanism of enzyme reaction when using *tert*-butyltoluene as its substrate. the methyl group available in substrates such as *p*-cymene and 4-ethyltoluene.

Substrate	Structure	Substrate Molecular Weight (g/mole)	Substrate Retention Time (min)	Product Molecular Weight (g/mole)	Product Retention Time (min)	Product
<i>p</i> -Cymene		134	5.87	150	7.64	<i>p</i> -Cymen-8-ol
Cumene		120	4.65	136	6.58	Unidentified
Isobutylbenzene		134	5.66	150	8.68	Unidentified
<i>n</i> -Butylbenzene		134	6.22	150	8.25	Unidentified
Propylbenzene		120	5.01	136	7.21	Unidentified
2,2-Dimethylpropyl benzene		148	6.32	n. d.	n. d.	n. d.
4-Ethyltoluene		120	5.12	136	7.38	4- Methylacetopheno ne



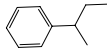



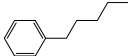
Substrate	Structure	Substrate Molecular Weight (g/mole)	Substrate Retention Time (min)	Product Molecular Weight (g/mole)	Product Retention Time (min)	Product
<i>tert</i> -Butyltoluene		148	6.63	164	9.35	4-Isopropylbenzyl alcohol
<i>tert</i> -Butylbenzene		134	5.49	n. d.	n. d. -	n. d.
<i>sec</i> -Butylbenzene		134	5.71	150	7.63	Unidentified
Toluene		n. d.	n. d.	n. d. -	n. d. -	n. d. -
<i>m</i> -Xylene		n. d.	n. d.	n. d. -	n. d. -	n. d. -
<i>p</i> -Xylene		n. d.	n. d.	n. d. -	n. d. -	n. d. -
Pentylbenzene		148	7.34	n. d.	n. d.	n. d.

Table 4.2: Summary of substrate specificity analysis indicating the product and its retention time measured by GC-MS
n. d. not detected

Our data indicates that the novel enzyme found in *B. xenovorans* LB400 can catalyze the hydroxylation of aromatic hydrocarbons that contain alkyl substituents attached to the benzene ring (usually 3-4 carbons, with the exception of *tert*-butyltoluene which contains five carbons attached to the benzene) for catalysis (see Table 4.2 for summary of substrates tested and enzymatic activity). Based on our results we can conclude that although the monooxygenase found in *B. xenovorans* LB400 catalyzes the hydroxylation of *p*-cymene, it is not a cymene monooxygenase as the one described in *P. putida* F1. By these data we cannot identify if the product of the reaction is accumulated or if it is further metabolized by the strain. However, although the analysis of genes found upstream of the novel monooxygenase (see Figure 4.1) reveal that two additional genes in the operon are similar to *p*-cumin alcohol dehydrogenase and *p*-cumin aldehyde dehydrogenase, *p*-cymen-8-ol cannot be oxidized further suggesting that the product gets accumulated.

4.3 Correlation of Monooxygenase Activity to Physiology of *B. xenovorans* LB400

The novel monooxygenase discussed previously has the ability to hydroxylate substrates that have not previously been described as carbon sources that support growth of *B. xenovorans* LB400. Many studies have reported the ability of LB400 to metabolize a wide variety of compounds, including PCBs (Denef *et al.* 2004; Denef *et al.* 2005; Denef *et al.* 2006; Parnell *et al.* 2006). Although it was thought that the biphenyl pathway was involved in the degradation of PCBs, recent data reveal that cell systems other than the biphenyl pathway contribute to the efficient degradation of PCBs and biphenyls (Parnell *et al.* 2006). Elucidating what pathways are involved in the metabolism of the aromatic hydrocarbons used in this study will provide insight into the metabolic routes contributing to the efficient biodegradation of these chemicals. In order to determine if there is a correlation between the monooxygenase activity and the physiology of *B. xenovorans* LB400, we decided to determine whether the chemicals indentified as substrates for the novel monooxygenase can be utilized by the strain LB400 as its only source of carbon and energy. The analysis of bacterial growth on aromatic hydrocarbons and the identification of the metabolic pathways that are involved in the process is essential to understand the physiology of this very versatile strain. Our efforts were directed at the identification of bacterial growth on a variety of aromatic hydrocarbons, analysis of any metabolites produced during growth, and the identification of the genes involved in the metabolism of these substrates by the strain.

4.3.1 Analysis of Different Carbon Sources that Support Growth of *B. xenovorans* LB400

It has been reported that *B. xenovorans* LB400 is enriched in aromatic catabolic pathways which are contained in genomic islands that confer a high versatility to the strain (Chain *et al.* 2006) (see Figure 1.1 for details). Eleven central and 20 peripheral aromatic pathways were identified in LB400 (Figure 1.1). Analysis of the capabilities of aromatic substrates to support bacterial growth will provide supporting evidence of the versatility and distinctiveness of *B. xenovorans* LB400 to adapt to different environments and offer the possibility of using this strain as a tool for the bioremediation of contaminated sites.

In order to further analyze the capabilities of LB400, we decided to investigate its ability to grow on a variety of aromatic hydrocarbons that were identified as substrates for the novel monooxygenase found in *B. xenovorans* LB400. In addition, we decided to test its ability to grow on the aromatic hydrocarbon *tert*-butylbenzene, which, although it was not reported as a substrate for the novel monooxygenase, has structural similarities to *tert*-butyltoluene, which was identified as a substrate for the novel enzyme. Analysis of growth capabilities were measured using growth curves to obtain a description of the growth pattern of the strain. The ratio of the initial inoculum was 1:10,000, trying to avoid a rapid shift between the lag phase to the log phase. As a positive control, *B. xenovorans* LB400 cells were grown on succinate, a carbon source that supports bacterial growth by the donation of electrons to produce fumarate and FADH₂ to generate energy.

LB400 is capable of utilizing *p*-cymene as its sole carbon source. Growth was detected after 36 hours of incubation, a relatively short time considering that the cells

required a period of adjustment to the new environment using aromatic hydrocarbons as their source of carbon and energy. Although growth was detected on *p*-cymene shortly after the initial inoculation, the growth rate was slower compared to other substrates, taking approximately 10 days to enter into the stationary phase (see Figure 4.20B). Isobutylbenzene can also support bacterial growth of *B. xenovorans* LB400 (Figure 4.20C). As with *p*-cymene, growth of LB400 on isobutylbenzene can be detected after approximately 36 hours of incubation. Cells grown on isobutylbenzene adapt relatively fast to the changes that the strain undergoes in adapting to the new surroundings, and the exponential phase is reached approximately 5 days after the initial inoculum. The lag phase for cultures grown on succinate (positive control) is less than an hour, while bacterial cells grown on isobutylbenzene and *p*-cymene showed a lag phase of approximately 30 hours. After three days of incubation, growth was detected on cells grown on 4-ethyltoluene (Figure 4.20D). It took approximately 6 days until the cells reached stationary phase when using 4-ethyltoluene as its sole source of carbon. Comparisons of the growth curves indicate that growth on these substrates is very similar, but growth on the substrates 4-ethyltoluene and isobutylbenzene seems to enter into the exponential phase earlier compared to growth on *p*-cymene. To our knowledge, growth of LB400 on isobutylbenzene and 4-ethyltoluene has not been reported before, so it will be interesting to determine if the novel enzyme described in this study is involved in the growth physiology of the strain or, as is the case for PCBs degradation, other metabolic pathways are involved in the growth physiology of the strain.

LB400 can also utilize 1,4-diisopropylbenzene as its sole source of carbon and energy. Growth was detected after 4 days of incubation and it took approximately 10

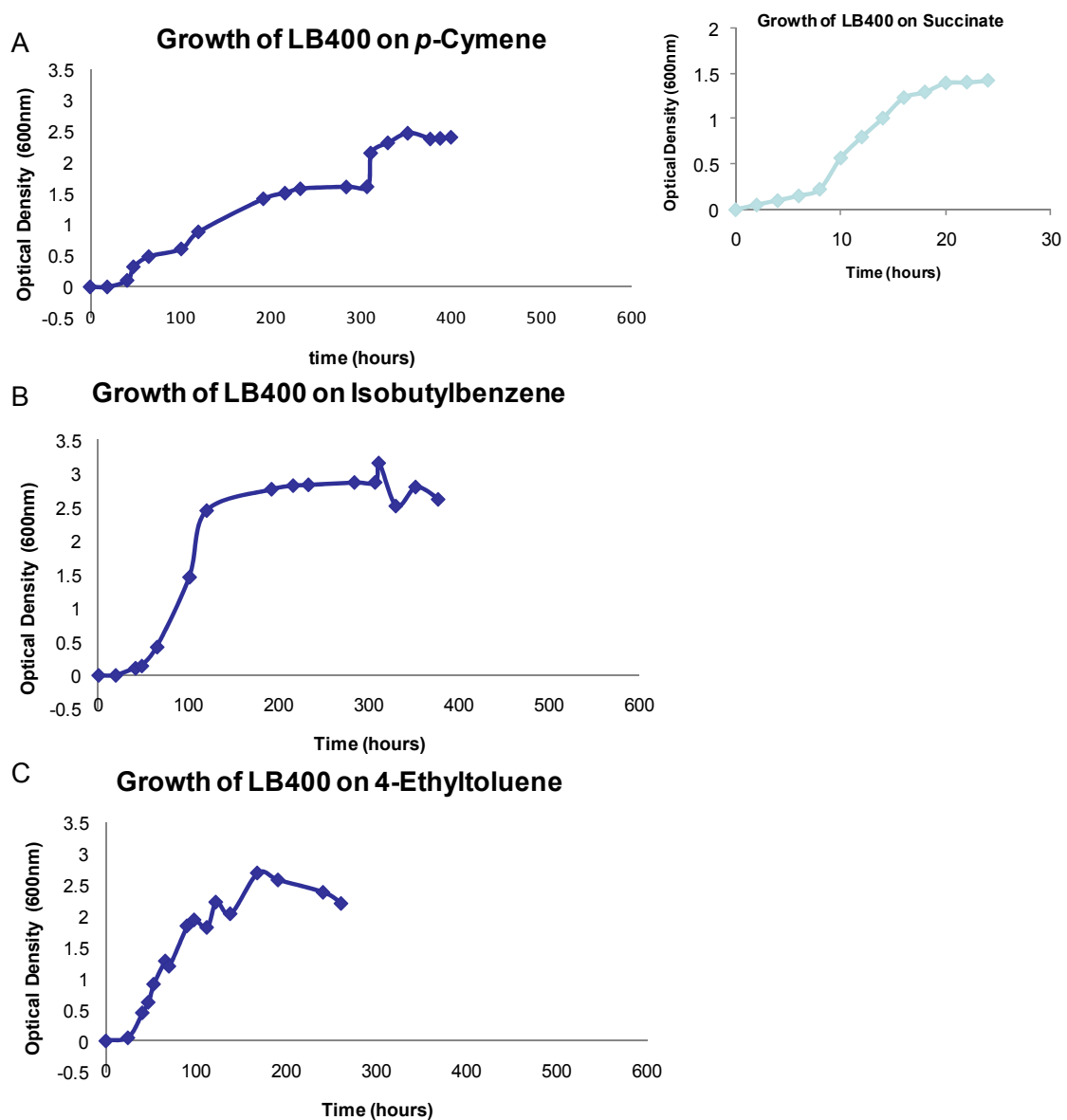


Figure 4.20: Growth curves of LB400 utilizing *p*-cymene, isobutylbenzene and 4-ethyltoluene sole carbon source

- A. Growth Curve of LB400 grown on *p*-cymene
 - B. Growth Curve of LB400 grown on isobutylbenzene
 - C. Growth Curve of LB400 grown on 4-ethyltoluene
- Right panel: LB400 grown on succinate was used as an experimental control

days to reach the stationary phase (Figure 4.21A). Similarly, growth was detected on the hydrocarbon *tert*-butylbenzene 4 days after inoculation (Figure 4.21B). No activity toward *tert*-butylbenzene was detected with the novel monooxygenase expression clone, suggesting that a different pathway is responsible for the degradation of this compound. Cultures of LB400 grown on *tert*-butylbenzene reaches an optical density value of 1.2, indicating that this is a less efficient substrate for the strain, if compared to *p*-cymene, isobutylbenzene, 4-ethyltoluene, 1,4-diisopropylbenzene, *n*-butylbenzene, *tert*-butyltoluene and propylbenzene which reaches OD values close to 3.0. Since *tert*-butylbenzene is not a substrate for the novel monooxygenase, it will be interesting to identify if bacterial growth is supported by the same metabolic routes as for those substrates that are catalyzed by the enzyme. *n*-Butylbenzene and *sec*-butylbenzene support bacterial growth of LB400 as well. The growth curve of *n*-butylbenzene shows a longer lag phase, requiring approximately 8 days to growth (Figure 4.22A). At day 17, the cultures entered stationary phase. Growth on *sec*-butylbenzene were also detected 8 days after inoculation, but compared to cells grown on *n*-butylbenzene, cells grown on *sec*-butylbenzene reach the lowest optical density value reported in this study (approximately 0.5) (Figure 4.22B). These results suggest that *sec*-butylbenzene is not a good substrate to support the growth physiology of LB400. It might be that *sec*-butylbenzene or its downstream metabolites are toxic to the cell or inhibit as the concentration of metabolites accumulate in the growth media.

LB400 growth was also supported by the substrates *tert*-butyltoluene and propylbenzene. Growth on *tert*-butyltoluene was detected 9 days after initial inoculation, reaching the exponential phase 15 days after (Figure 4.23A). Interestingly, when cells

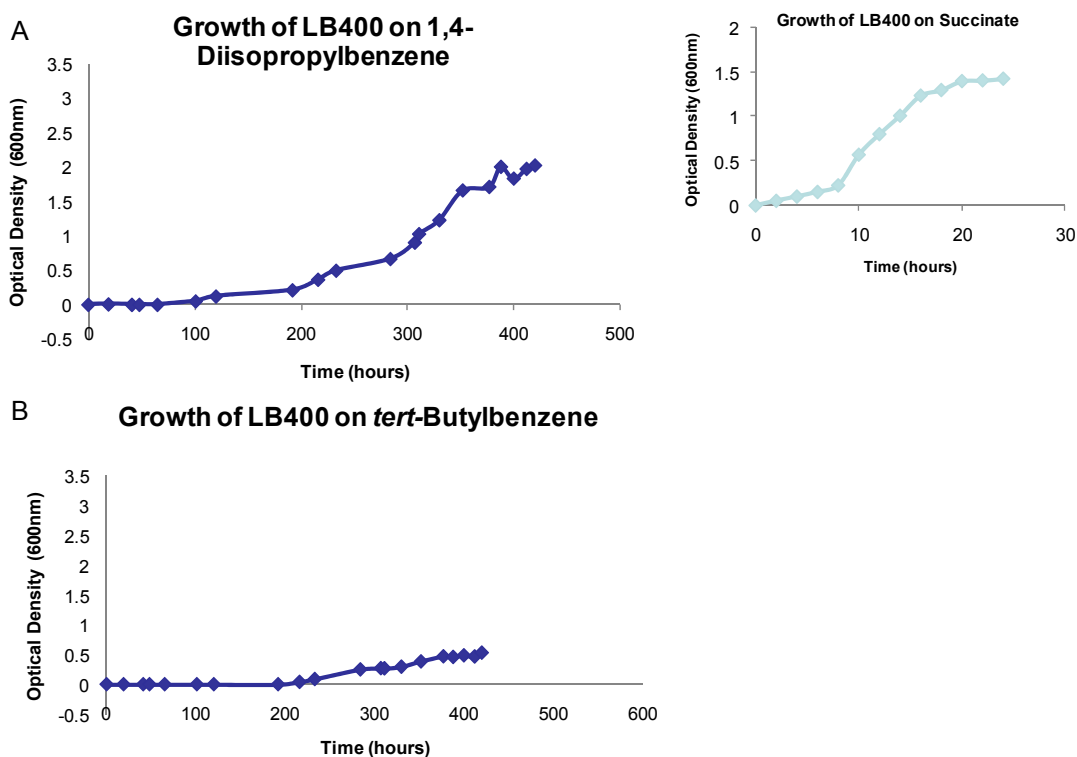


Figure 4.21: Growth curves of LB400 utilizing 1,4-diisopropylbenzene and *tert*-butylbenzene as sole carbon source

- A. Growth Curve of LB400 grown on 1,4-diisopropylbenzene
 - B. Growth Curve of LB400 grown on *tert*-butylbenzene
- Right panel: LB400 grown on succinate was used as an experimental control

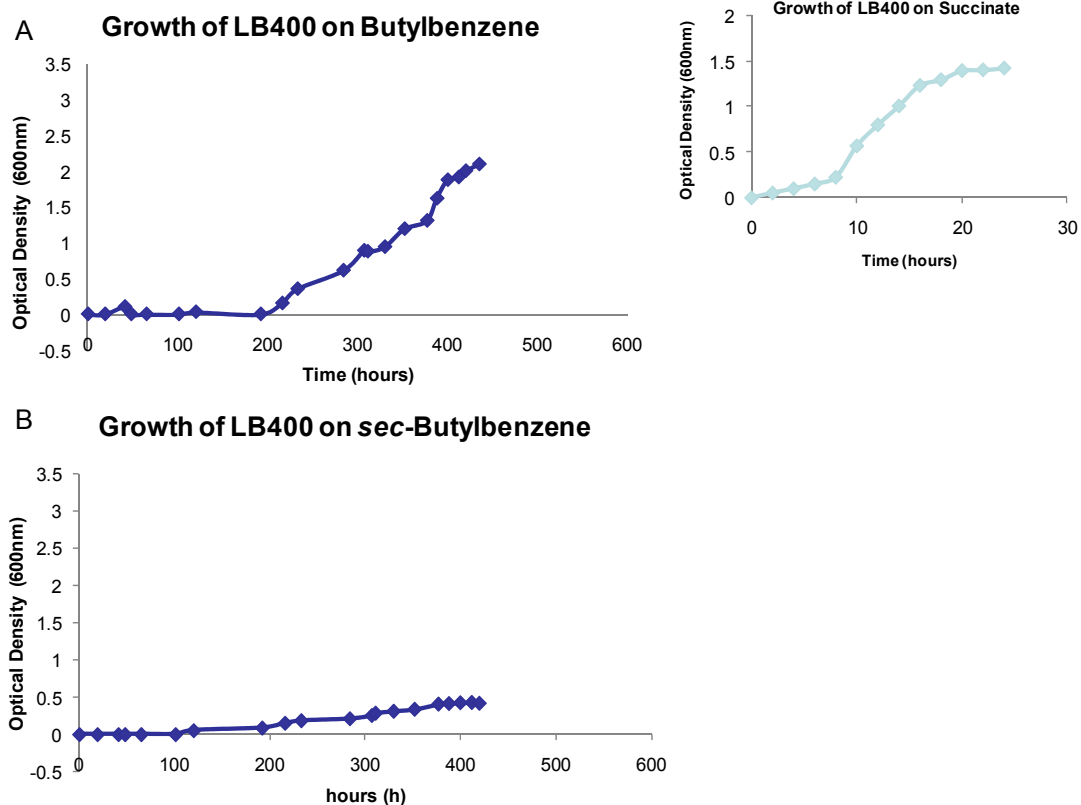


Figure 4.22: Growth curves of LB400 utilizing *n*-butylbenzene and *sec*-butylbenzene as sole carbon source

A. Growth Curve of LB400 grown on *n*-butylbenzene

B. Growth Curve of LB400 grown on *sec*-butylbenzene

Right panel: LB400 grown on succinate was used as an experimental control

were grown on this compound, the media changed its color to an intense pink (Figure 4.23A). Attempts to identify what metabolite(s) produced this interesting change in the culture medium failed, thus no explanation could be provided for the changes observed. Growth on propylbenzene was detected 14 days after the initial inoculum (Figure 4.23B). This is an unusual long lag phase for an aerobic organism. This was the longest lag phase detected for all the substrates tested. To determine if growth on this substrate was the result of a spontaneous mutation that occurred during incubation, we transferred an inoculum of the culture to fresh media with propylbenzene provided to the culture in the vapor phase to determine if the extension of the lag phase of the cells changes. If a spontaneous mutation occurred during incubation of the first culture, the lag phase of the second culture should be shorter. The lag phase of the second culture was 15 days as in the original culture, indicating that the delay in growth was the result of a period of adaptation of the cells to the substrate and not of a mutation in the strain. LB400 grown on propylbenzene shows a change of color in the culture medium (Figure 4.23B). The same pink color detected with the substrate *tert*-butyltoluene was produced during growth on propylbenzene. Unfortunately, the metabolite(s) responsible for this change have not been identified yet. Identification of this product could provide some insight about pathways/metabolic activity involved in the degradation of the aromatic hydrocarbons. The change in color and the similarities in growth pattern when LB400 is grown on *tert*-butyltoluene and on propylbenzene suggests that a similar pathway is involved in the degradation of these compounds.

Our results indicate that *B. xenovorans* is able to utilize the hydrocarbons *p*-cymene, *n*-butylbenzene, isobutylbenzene, *sec*-butylbenzene, propylbenzene, 4-ethyltoluene, *tert*-

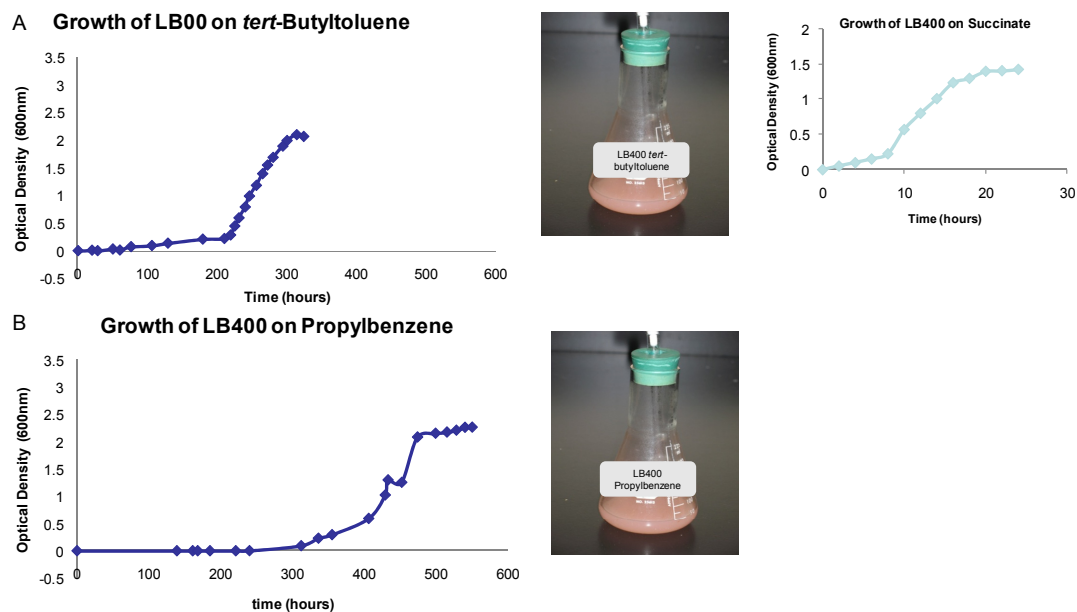


Figure 4.23: Growth curves of LB400 utilizing *tert*-butyltoluene and propylbenzene as sole carbon source

A. LB400 cultures grown on *tert*-butyltoluene

B. Growth Curve of LB400 grown on propylbenzene

Right panel: LB400 grown on succinate was used as an experimental control

butylbenzene, and 1,4-diisopropylbenzene as the sole carbon source to support bacterial growth. All of these compounds, with the exception of *tert*-butylbenzene, were also identified as substrates for the novel monooxygenase. Although cumene was identified as a substrate for the novel monooxygenase as well, no growth was detected. Lack of growth on cumene might suggest that although bacterial growth can be detected for most of the hydrocarbons reported as enzyme substrates, there is no correlation between the enzyme activity and the cell physiology of the strain. Another possibility could be that the metabolites produced in the reaction are toxic to the *B. xenovorans* LB400 cells.

4.3.2 Metabolite Analysis of Cultures Utilizing Aromatic Hydrocarbons as Their Carbon Source

Identification of the metabolites produced by the cultures grown on the different aromatic hydrocarbons could provide information about the cell systems involved in the degradation of these chemicals. Since *B. xenovorans* LB400 is a versatile microorganism capable of growing and tolerating a variety of chemicals, the identification of the cell systems that confer these capabilities are of great importance in the area of bioremediation. The analysis of these metabolites was performed by GC-MS. The products of the cell cultures grown on *p*-cymene, propylbenzene, *n*-butylbenzene, *sec*-butylbenzene, *tert*-butylbenzene, isobutylbenzene, 1,4-diisopropylbenzene, 4-ethyltoluene and *tert*-butyltoluene were extracted and analyzed by GC-MS.

GC-MS analysis of cell cultures grown on *p*-cymene revealed a peak at 8.02 minutes identified as a metabolite. The gas chromatograph shows another peak at 5.87 minutes

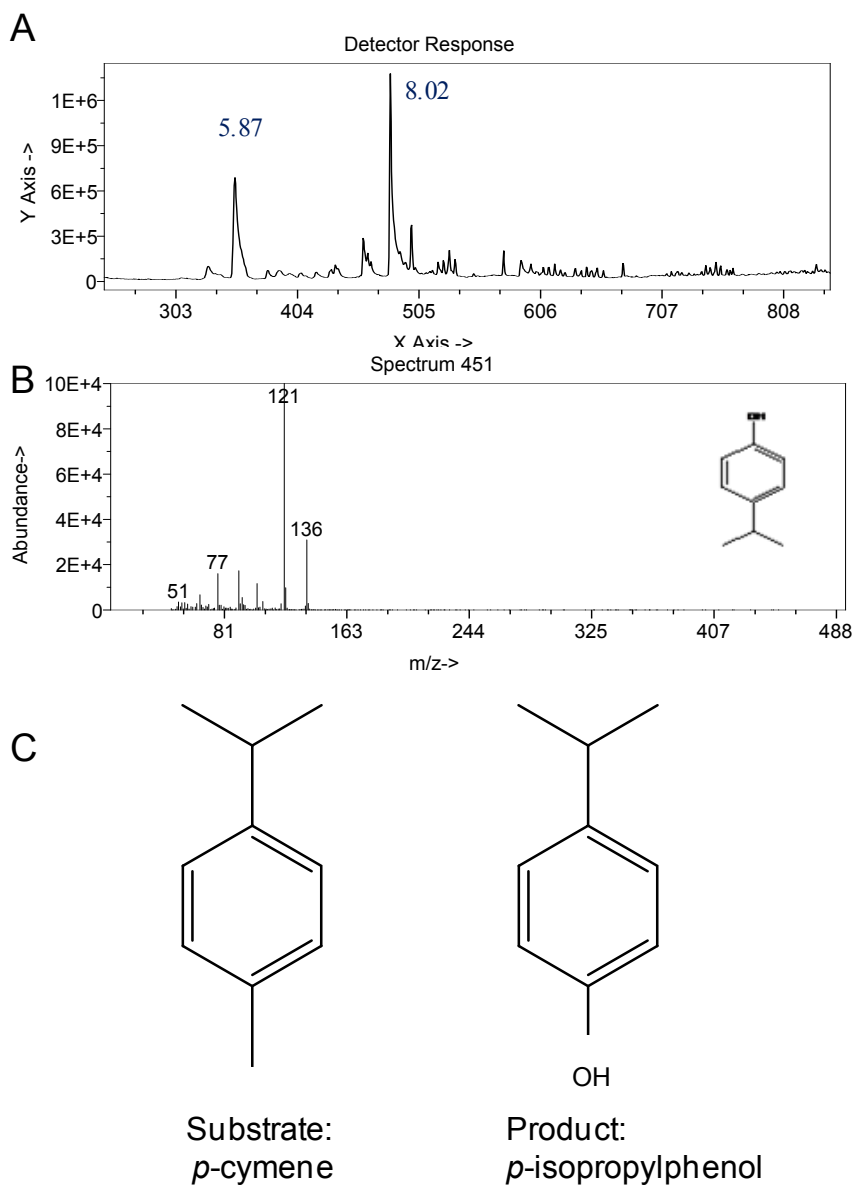


Figure 4.24: Analysis of metabolites produced by LB400 cells growing on *p*-cymene

- Gas chromatogram of metabolites produced by LB400 when growing on *p*-cymene
- Mass spectrum of the product *p*-isopropylphenol
- Chemical structures of substrate and product of the reaction

corresponding to the substrate *p*-cymene (Figure 4.24). The molecular weight of the product is 136 g/mole and the molecular weight of the substrate is 134 g/mole. The GC-MS library identifies the product as *p*-isopropylphenol (see Figure 4.24C for chemical structure). Comparison of the chemical structures of the substrate provided and the product reveal that the only difference between the two compounds is the substitution of the methyl group present in the substrate for the hydroxyl group present on *p*-isopropylphenol. Based on these results, no correlation between the activity of the novel enzyme and the growth physiology of the LB400 strain can be made.

The gas chromatogram of LB400 cultures grown on propylbenzene shows two different peaks eluting at 5.03 minutes and 7.31 minutes (Figure 4.25A). The peak eluting at 5.03 minutes correspond to propylbenzene with a molecular weight of 120 g/mole. The peak eluting at 7.31 minutes correspond to the metabolite produced in the reaction. It has a molecular weight of 136 g/mole (Figure 4.25B). The peak was identified by the GC-MS library as 1-phenyl-2-propanol (Figure 4.25C). This differs from the product of propylbenzene oxidation by the novel monooxygenase clone. This data suggest that a monooxygenase is involved in the degradation of propylbenzene, but the pathway utilized for this activity cannot be identified yet.

Analysis of the chromatogram of the cultures grown on isobutylbenzene identifies the substrate eluting at 5.67 minutes and its mass spectrum confirms a molecular weight of 134 g/mole (Figure 4.26). A metabolite eluting at 7.36 minutes and of molecular weight of 150 g/mole was identified in the reaction. The GC-MS library did not identify the product, but its mass spectrum confirms the addition of an oxygen atom (Figure 4.26B) as

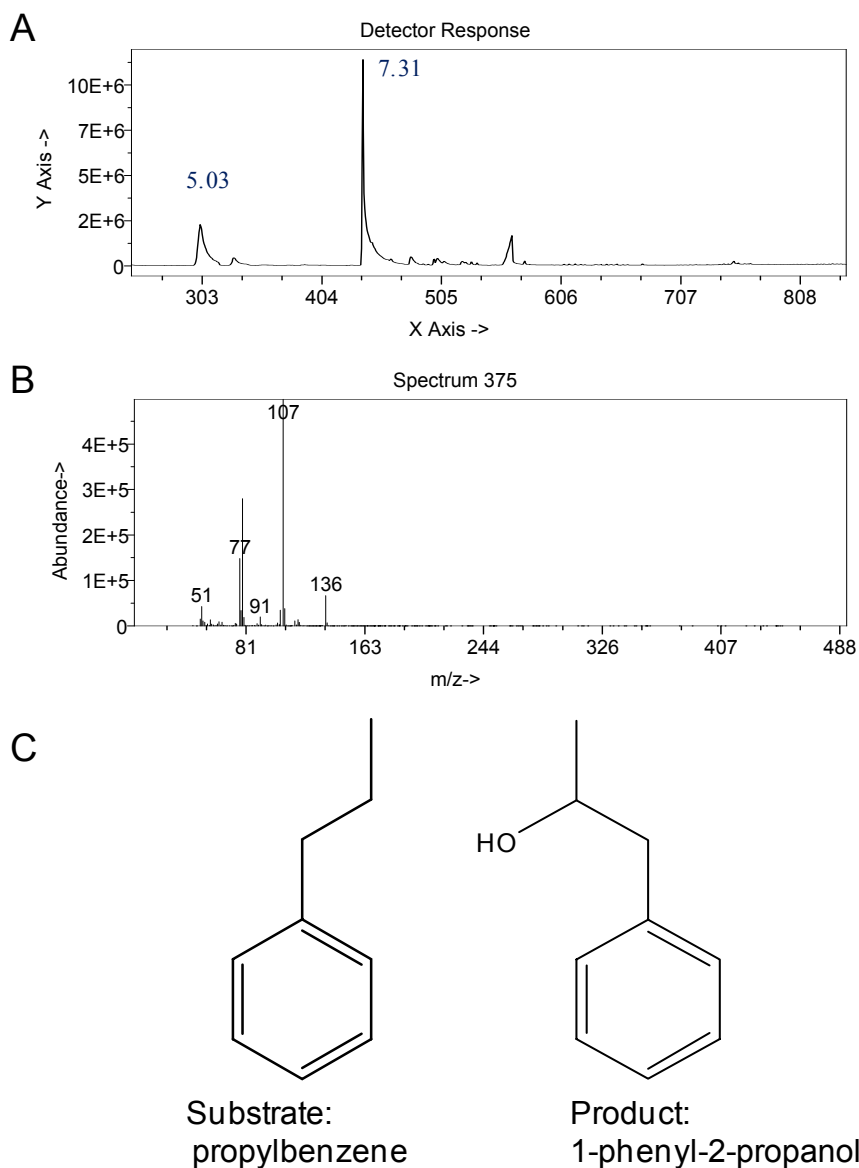


Figure 4.25: Analysis of metabolites produced by LB400 cells growing on propylbenzene

- Gas chromatogram of metabolites produced by LB400 when growing on propylbenzene
- Mass spectrum of the product 1-phenyl-2-propanol
- Chemical structures of substrate and product of the reaction

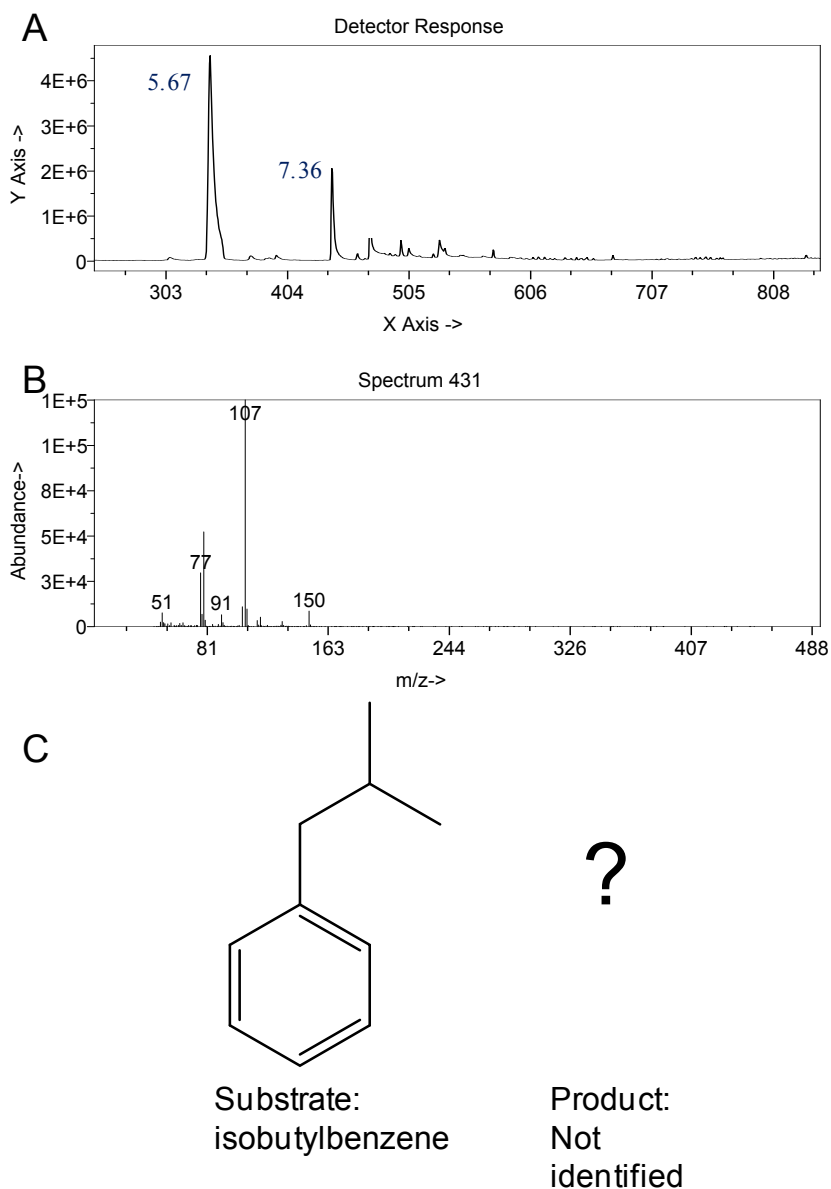


Figure 4.26: Analysis of metabolites produced by LB400 cells growing on isobutylbenzene

- Gas chromatogram of metabolites produced by LB400 when growing on isobutylbenzene
- Mass spectrum of the product (unidentified)
- Chemical structures of substrate

occurs when propylbenzene is used as the substrate.

When 4-ethyltoluene was provided as a substrate to support the growth physiology of LB400, we identified one peak eluting at 10.02 minutes with a molecular weight of 150 g/mole (Figure 4.27). The product of the reaction was identified by the GC-MS library as ethylbenzoic acid. Interestingly, this product suggests that a monooxygenase is involved in the pathway by hydroxylating the methyl group of the substrate, then followed by two dehydrogenations to produce the carboxylic acid form of the substrate. Based on the results obtained by this analysis, the hydroxylation reaction occurs at the methyl group of 4-ethyltoluene, and not at the *alpha* carbon of the ethyl group as occurs by the catalysis of our novel enzyme. These data suggest that the novel enzyme is not correlated to the growth physiology of LB400 on 4-ethyltoluene and another metabolic route is responsible for degradation of this substrate.

B. xenovorans LB400 can utilize *tert*-butylbenzene as a substrate to support its growth physiology, however, this compound was not identified as a substrate for the novel monooxygenase described in this work. The metabolite analysis of cells grown on different aromatic hydrocarbons does not indicate that the novel enzyme is involved in the efficient degradation of these compounds. Table 4.3 shows a comparison between the substrates catalyzed by the enzyme and the products of growth physiology. Although the enzyme is capable of growing on the aromatic hydrocarbons *sec*-butylbenzene, *n*-butylbenzene, *tert*-butyltoluene and 1,4-diisopropylbenzene, we were not able to detect any of the metabolites produced in the enzymatic reaction by GC-MS analysis of *B. xenovorans* LB400 cells grown on these substrates. Metabolite analysis of cultures grown on 4-ethyltoluene identifies a product which is the carboxylic acid form of the

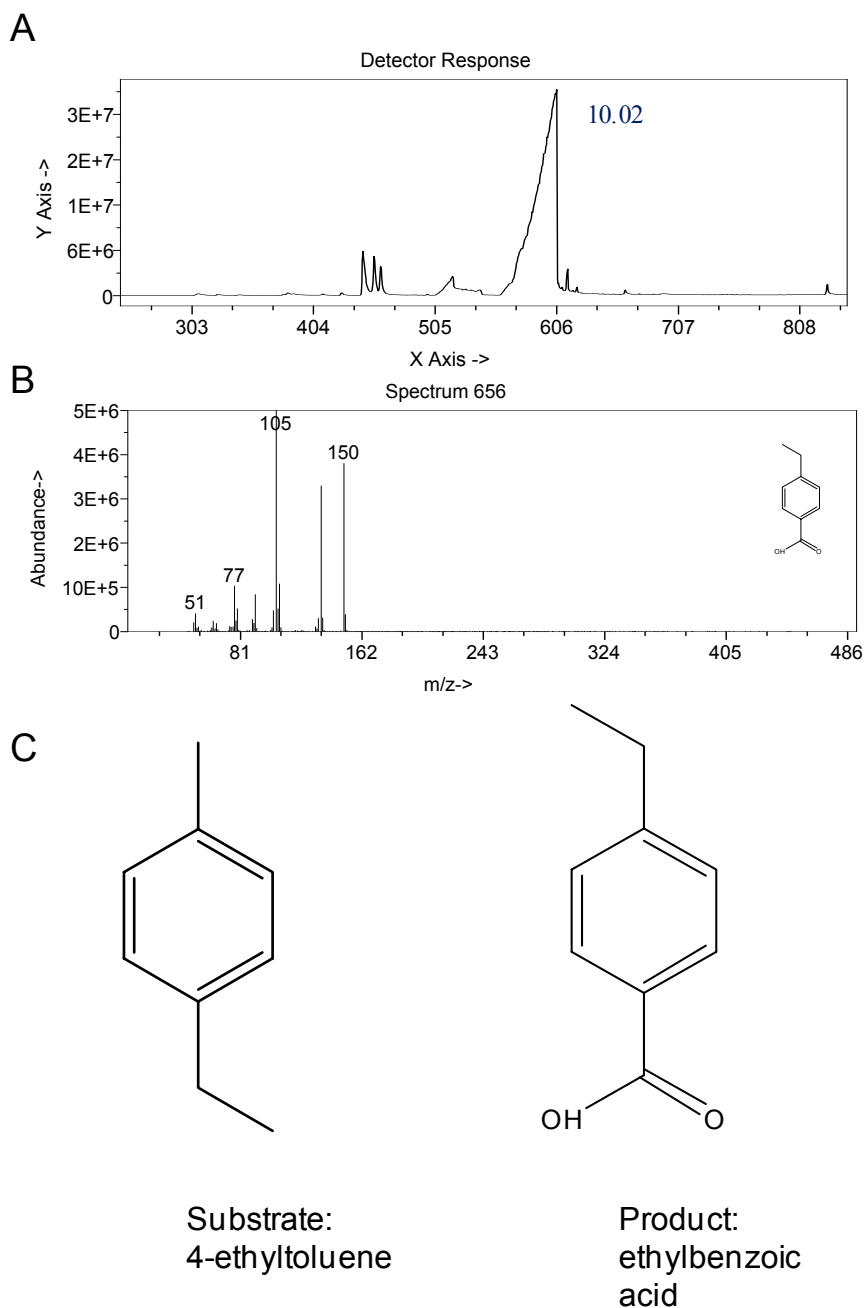


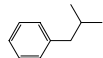
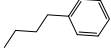
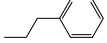
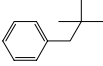
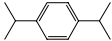
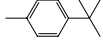


Figure 4.27: Analysis of metabolites produced by LB400 cells growing on 4-ethyltoluene

- Gas chromatogram of metabolites produced by LB400 when growing on 4-ethyltoluene
- Mass spectrum of the product ethylbenzoic acid
- Chemical structures of substrate and product of the reaction

Substrate	Structure	Product Molecular Weight (g/mole)	Possible Product	Enzyme's Product Retention Time (min)	Growth as only carbon source	Metabolite Molecular Weight (g/mole)	Metabolite Retention Time (min)	Metabolite Product
<i>p</i> -Cymene		150	<i>p</i> -Cymen-8-ol	7.64	Yes	136	8.02	<i>p</i> -Isopropylphenol
Cumene		136	Unidentified	6.58	No	n. d.	n. d.	n. d.
Isobutylbenzene		150	Unidentified	8.68	Yes	150	7.36	Unidentified
<i>n</i> -Butylbenzene		150	Unidentified	8.25	Yes	n. d.	n. d.	n. d.
Propylbenzene		136	Unidentified	7.21	Yes	136	7.31	1-Phenyl-2-propanol
2,2-Methylpropyl-Benzene		n. d.	n. d.	n. d.	Yes	n. d.	n. d.	n. d.
1,4-Diisopropylbenzene		178	Unidentified	9.10	Yes	n. d.	n. d.	n. d.
<i>tert</i> -Butyltoluene		164	<i>tert</i> -butylbenzyl alcohol	9.35	Yes	n. d.	n. d.	n. d.


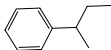
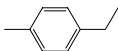
Substrate	Structure	Product Molecular Weight (g/mole)	Possible Product	Enzyme's Product Retention Time (min)	Growth as only carbon source	Metabolite Molecular Weight (g/mole)	Metabolite Retention Time (min)	Metabolite Product
<i>tert</i> -Butylbenzene		n. d.	n. d.	n. d.	Yes	87	4.22	Isobutyric acid
<i>sec</i> -Butylbenzene		150	Unidentified	7.63	Yes	n. d.	n. d.	n. d.
4-Ethyltoluene		136	Methylacetophenone	7.38	Yes	150	10.02	Ethylbenzoic acid

Table 4.3: Comparison between substrate specificity analysis of the putative cymene monooxygenase and growth physiology of LB400 on identified substrates

Product retention times and molecular weights are included for comparisons.

substrate. The analysis of the cultures grown on 4-ethyltoluene, isobutylbenzene and propylbenzene identify that an oxygen atom was added to the substrate. These results suggest that 4-ethyltoluene, isobutylbenzene and propylbenzene use a similar pathway to support its growth physiology, but the enzymes and pathway involved in their utilization has not yet been identified. Metabolite analysis suggested that the degradation of *tert*-butyltoluene occurs via the biphenyl pathway.

It was previously described that the biphenyl pathway is not always directly involved in the degradation of biphenyls and PCBs by *B. xenovorans* LB400 (Parnell *et al.* 2006). We suggest that, as for the case of biphenyls and PCBs, the activity of the enzyme described in this study is not involved in the degradation of these compounds by LB400. Degradation of these aromatic hydrocarbons might be supported by other enzymes of *B. xenovorans* LB400. Since the metabolite analysis of the products of the substrates propylbenzene, 4-ethyltoluene and isobutylbenzene share chemical structure similarities we suggest that the degradation of these substrates occur via a similar pathway. The chemical structure of the substrate *tert*-butyltoluene differs from the previous substrates, so the pathway utilized for the degradation of this compound might differ from the one used by the cells when other aromatic hydrocarbons are provided as the carbon source. In fact, the chemical structure of *tert*-butyltoluene is similar to the structure of *tert*-butylbenzene, so it would be interesting to determine whether degradation of these substrates is catalyzed by a common pathway. Unfortunately, no metabolite product was detected when 1,4-diisopropylbenzene, *n*-butylbenzene, *sec*-butylbenzene and *tert*-butylbenzene were used as a substrate for growth. Further analysis in this area could

provide information of the pathway responsible for the degradation in order to determine if the chemistry of the substrate affects the pathway involved in its degradation.

4.3.3 Detection of Genes Expressed by Cell Cultures Grown on Aromatic Hydrocarbon Sources

Reverse transcriptase polymerase chain reaction (RT-PCR) was used to determine if the hydroxylase gene of the novel monooxygenase was expressed when LB400 was grown on different carbon sources. Cells cultures of LB400 grown on succinate, *p*-cymene, propylbenzene, *n*-butylbenzene, isobutylbenzene, *sec*-butylbenzene, *tert*-butylbenzene, 1,4-diisopropylbenzene, 4-ethyltoluene, and *tert*-butyltoluene were selected for RNA extraction when cell cultures reached an optical density of 0.3-0.4 (measured at 600nm). As an experimental positive control for the RT-PCR reaction we designed primers to amplify the *gyrB* gene, which encodes a DNA topoisomerase II, which is constitutively expressed. The expected size of the *gyrB* amplicon is 1042 bases. To determine if the novel monooxygenase is expressed during growth on the aromatic hydrocarbons, we designed primers to amplify a region of the hydroxylase component which gives a product of 1057 bases. Additional primers were designed in this study for the amplification of *bphA1*, which codes for the large subunit of the biphenyl dioxygenase. The expected product size of this amplification is 954 bases.

The RT-PCR data obtained from the cultures grown on *p*-cymene (Figure 4.28B) showed no expression of the novel monooxygenase, confirming that this enzyme is not involved in the catabolism of this substrate by LB400 strain. Amplification of the novel hydroxylase and the biphenyl oxygenase genes was not detected when cells were grown on succinate (Figure 4.28A), propylbenzene (Figure 4.28C), *n*-butylbenzene (Figure

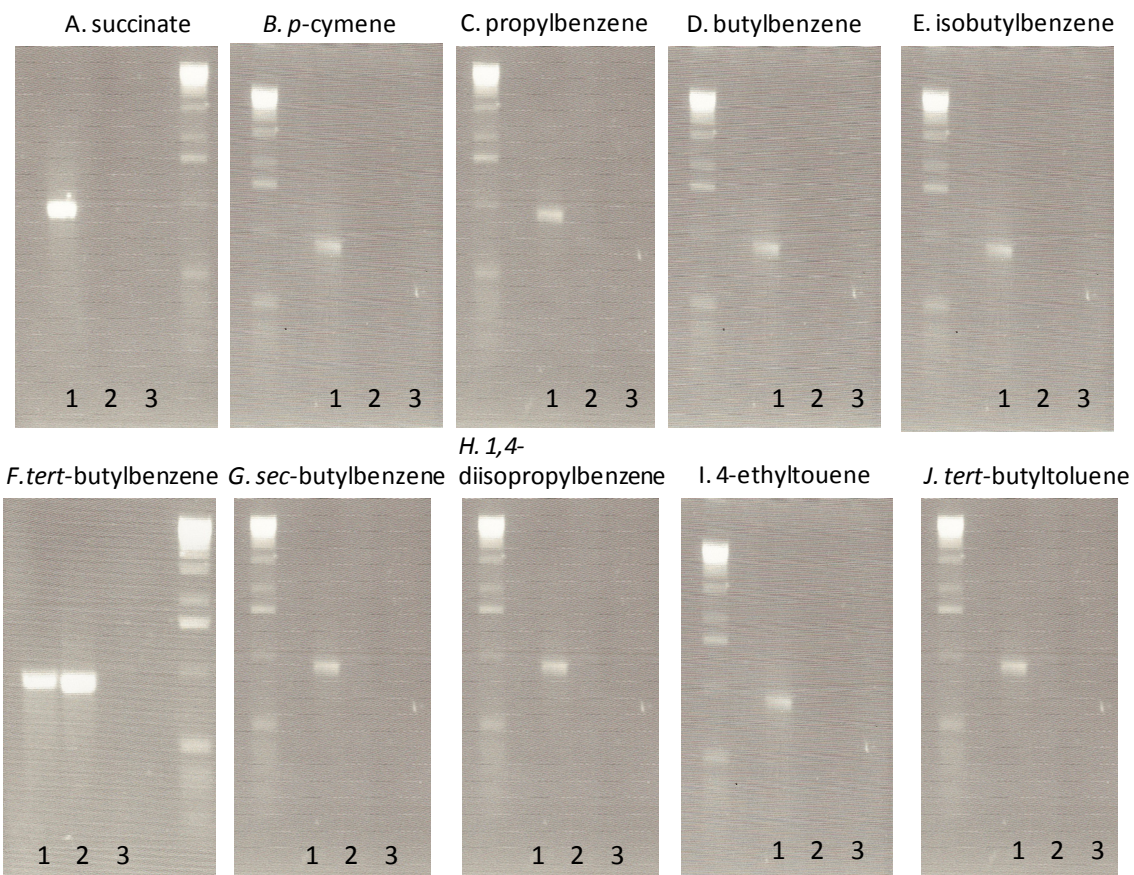


Figure 4.28: RT-PCR results for cell cultures grown on different hydrocarbons.

Ladder 1Kb+; lane 1: *gyrB* primers; lane 2: biphenyl dioxygenase primers; lane 3: putative cymene monooxygenase primers

4.28D), isobutylbenzene (Figure 4.28E), *tert*-butylbenzene (Figure 4.28F), *sec*-butylbenzene (Figure 4.28G), 1,4-diisopropylbenzene (Figure 4.28H), 4-ethyltoluene (Figure 4.28I), and *tert*-butyltoluene (Figure 4.28J). The negative results for the cells grown on the aromatic hydrocarbons indicate that the activity of the novel monooxygenase is not correlated to the growth physiology of LB400. The negative RT-PCR results for LB400 novel monooxygenase grown on these substrates indicate that while the enzyme is active towards these substrates, it is not expressed during their degradation by LB400.

cDNA was detected when the RNA from cells grown on *tert*-butylbenzene was added as a template to detect the amplification of the biphenyl dioxygenase (Figure 4.28K). No product was detected when the novel hydroxylase primers were used as a template. RT-PCR results showed that the *bphA1* gene was expressed when the LB400 cells used *tert*-butylbenzene as a substrate. Evidence from RT-PCR data confirms that the biphenyl genes are involved in the degradation of *tert*-butylbenzene (Figure 4.29) and although this compound is structurally similar to *tert*-butyltoluene, the strain utilizes different enzymatic pathways for their degradation.

Our data indicates that although the novel enzyme is capable of catalyzing the hydroxylation of a variety of aromatic hydrocarbons, it is not involved in the growth physiology of the strain on these compounds. As was previously published for the growth of *B. xenovorans* LB400 on biphenyls and PCBs (Parnell *et al.* 2006), we suggest that other enzymes, and not the novel monooxygenase, are involved in the degradation of these compounds.

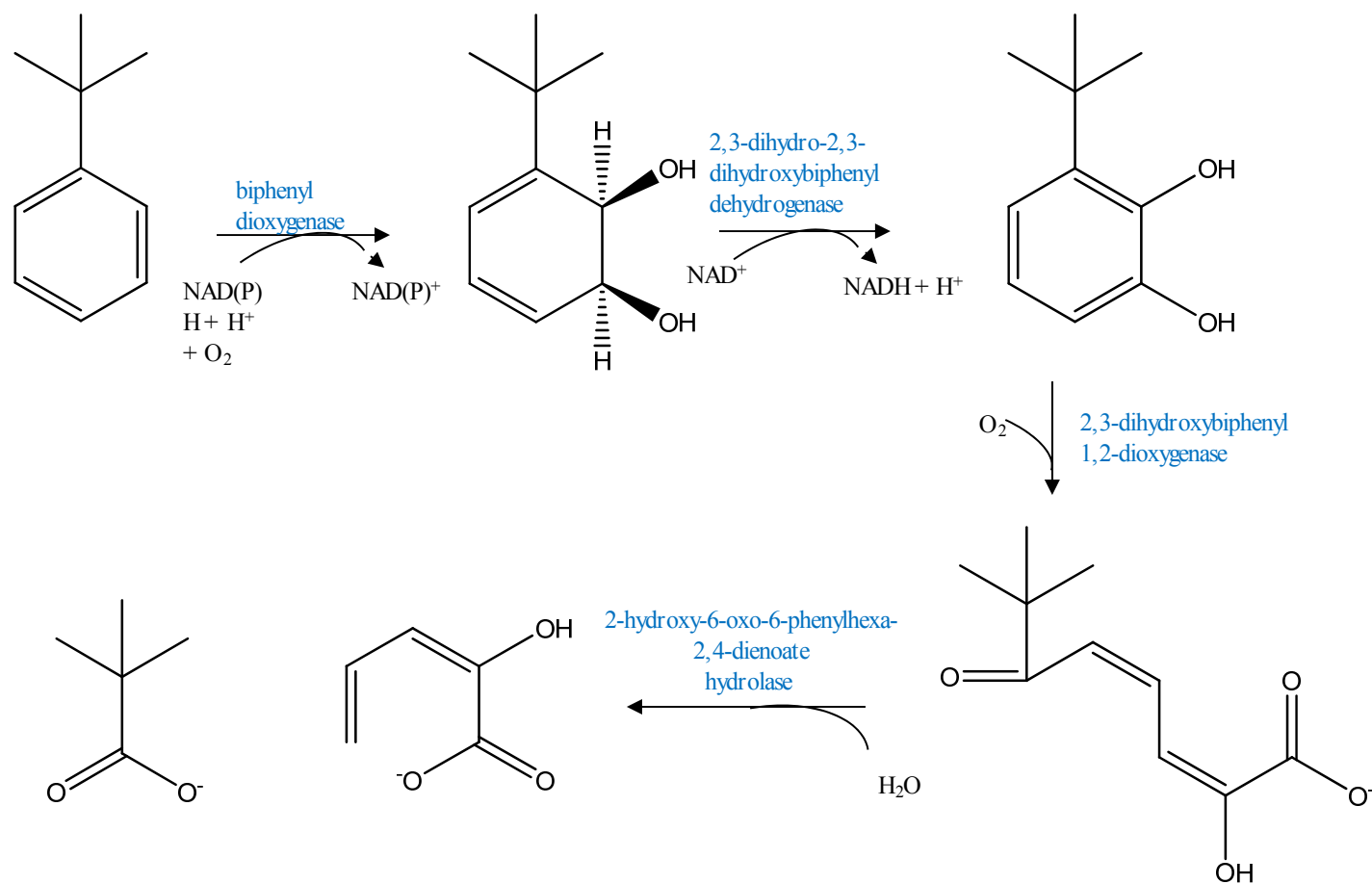


Figure 4.29: Proposed pathway for the degradation of tert-butylbenzene by *B. xenovorans* LB400

5. Conclusions

Protein expression and substrate specificity analysis of the putative cymene monooxygenase found in *B. xenovorans* LB400 revealed that although the enzyme recognizes *p*-cymene as a substrate, its product differs from the one described for the cymene monooxygenase from *P. putida* F1 (Nishio *et al.* 2001). The novel enzyme found in *B. xenovorans* LB400 attacks *p*-cymene at the hindered carbon from the isopropyl group next to the benzene ring. This attack is quite unusual, because similar enzymes such as cymene and xylene monooxygenase have been identified to hydroxylate the methyl group substituent of their substrates. In fact, this novel enzyme catalyzes the hydroxylation of substrates that do not have the methyl group required for hydroxylation by methyl hydroxylating enzymes. Even though these differences have been identified, phylogenetic analysis revealed that this enzyme contains the histidine motifs conserved in membrane-bound enzymes that are proposed to have active sites that require iron and the activation of molecular oxygen as part of the catalytic cycle (Shanklin *et al.* 1997). We have identified by substrate specificity analysis that the enzyme is active towards a variety of substrates, but in order to determine how broad is its substrate range, further research is required. It would be interesting to identify if the enzyme is active toward halogenated aromatic hydrocarbons, as was reported for xylene and cymene monooxygenase previously. Although we have identified 10 different substrates catalyzed by this enzyme, the mechanism of catalysis of this enzyme is still uncertain and further research is required for its characterization.

Growth is the most global manifestation of how all bacterial cells adapt to changes in the environment, how they compete with others and how to ensure their survival (Schaechter 2006). The identification of growth of *B. xenovorans* LB400 on different aromatic hydrocarbons confirms how versatile is this strain in adapting to different growth conditions and utilizing hydrocarbons as its only source of carbon and energy. By this analysis we identified that the novel monooxygenase described in this study does not support growth of the strain on the substrates identified. The metabolite analysis data obtained from the cell cultures grown on 4-ethyltoluene, propylbenzene and isobutylbenzene indicated a monooxygenation reaction different from that catalyzed by the novel monooxygenase studied after cloning in *E. coli*. The metabolite produced by isobutylbenzene was not identified, but the molecular weight suggest that cells grown on isobutylbenzene involve a monooxygenation reaction that also differs from the one catalyzed by the novel enzyme. RT-PCR data on *tert*-butylbenzene suggest that the biphenyl pathway is involved in the degradation of *tert*-butylbenzene, but this compound was not identified as a substrate for the novel enzyme. It is still unknown if there is a correlation between the substrates utilized by the enzyme and the pathway utilized by the strain for the degradation of these substrates as the only source of carbon and energy.

6. Future Directions

Understanding the enzyme mechanisms and the pathways involved in the degradation of hydrocarbons represent a very useful tool in the areas of bioremediation and the organic synthesis of chemicals. In order to understand the chemical interactions and the oxygen/substrate activation mechanisms, it would be interesting to determine the function on catalysis of the conserved histidine regions and the identification of the amino acids involved in catalysis. This can be achieved by site directed mutagenesis experiments, as was previously described for the alkane hydroxylase from *P. putida* GPo1 (van Beilen *et al.* 2005). Since the mechanism of catalysis of the novel enzyme found on *B. xenovorans* LB400 differs from that of *P. putida* F1 cymene monooxygenase, and the well-characterized xylene and alkane monooxygenase, it will be interesting to analyze the hydrophobicity of this enzyme to predict its structure and topology, as was performed on the alkane monooxygenase (van Beilen *et al.* 1992).

In this study, we report enzyme activity toward 9 different substrates. It will be interesting to further analyze the substrate range of this novel enzyme. By our data, we can conclude that the enzyme is specific for substrates containing a benzene ring with an alkyl group of at least two carbon length, but it would be interesting to determine if the presence of other chemical groups on the substrate affect the enzyme specificity. A possibility for this study could be the addition of halogenated groups to the substrates already identified as substrates for the enzyme as was reported previously for the xylene and cymene monooxygenase (Kunz and Chapman 1981; Delgado *et al.* 1992; Wubbolts *et al.* 1994; Nishio *et al.* 2001; Bühler *et al.* 2002). Since the product of the hydroxylation reaction by the novel enzyme was not identified in all the reactions

because of the lack of commercially available standards and the inability of the GC-MS library to find a match, further identification of the product of the reaction is required. Probably, the use of chemical techniques, such as nuclear magnetic resonance (NMR) could be useful identifying the product.

We were able to measure growth of LB400 on a variety of hydrocarbons. Based on our results, the novel enzyme is not involved in the degradation and growth physiology of *B. xenovorans* LB400. RT-PCR data suggest that LB400 metabolizes the substrate *tert*-butylbenzene by the biphenyl pathway. In order to confirm this data, a gene knockout of the biphenyl pathway will corroborate this information. It would be interesting to determine whether the other chemicals tested utilize a common biochemical route for metabolism and if substrate chemical structure affects the metabolic route utilized by LB400. Since the metabolite analysis of the cell cultures reveals the addition of an oxygen atom, a gene knockout of the novel monooxygenase will provide further evidence that this enzyme is not involved in the degradation of the aromatic hydrocarbons.

7. References

- Austin, R. N., K. Buzzi, E. Kim, G. J. Zystra and J. T. Groves (2003). "Xylene monooxygenase, a membrane-spanning non-heme diiron enzyme that hydroxylates hydrocarbons via a substrate radical intermediate." J. Inorg. Chem. **8**: 733-740.
- Austin, R. N., H. K. Chang, G. J. Zystra and J. T. Groves (2000). "The non-heme diiron alkane monooxygenase of *Pseudomonas oleovorans* (AlkB) hydroxylates via a substrate radical intermediate." J. Am. Chem. Soc. **122**(47): 11747-11748.
- Bedard, D. L., R. Unterman, L. H. Bopp, M. J. Brennan, M. L. Haberl and C. Johnson (1986). "Rapid assay for screening and characterizing microorganisms for the ability to degrade polychlorinated biphenyls." Appl. Environ. Microbiol. **4**: 761-768.
- Benson, S., M. Fennewald, J. Shapiro and C. Huettner (1977). "Fractionation of inducible alkane hydroxylase activity in *Pseudomonas putida* and characterization of hydroxylase-negative plasmid mutations." J. Bacteriol. **132**(2): 614-621.
- Bertrand, E., R. Sakai, E. Rozhkova-Novosad, L. Moe, G. G. Fox, J. T. Groves and R. N. Austin (2005). "Reaction mechanisms of non-heme diiron hydroxylases characterized in whole cells." J. Inorg. Biochem. **99**: 1998-2006.
- Bezalel, L., Y. Hadar, P. P. Fu, J. P. Freeman and C. E. Cerniglia (1996). "Initial oxidation products in the metabolism of pyrene, anthracene, fluorene, and dibenzothiophene by the white rot fungus *Pleurotus ostreatus*." Appl. Environ. Microbiol. **62**(7): 2247-2253.
- Bopp, L. (1986). "Degradation of highly chlorinated PCBs by *Pseudomonas* strain LB400." J. Ind. Microbiol. **1**: 23-29.
- Brazeau, B. J. and J. D. Lipscomb (2003). "Key amino acid residues in the regulation of soluble methane monooxygenase catalysis by component B." Biochemistry **42**(19): 5618-5631.
- Brosius, J., M. Erfle and J. Storella (1985). "Spacing the -10 and -35 regions in the *tac* promoter." J. Biol. Chem. **260**: 3539-3541.
- Bühler, B. (2003). Specific oxyfunctionalization of pseudocumene by multistep biocatalysis - Mechanism and scale up Zurich, Swiss Federal Institute of Technology Zurich. **PhD**.
- Bühler, B., A. Schmid, B. Hauer and B. Witholt (2000). "Xylene monooxygenase catalyzes the multistep oxygenation of toluene and pseudocumene to corresponding alcohols, aldehydes, and acids in *Escherichia coli* JM101." J. Biol. Chem. **275**(14): 10085-10092.

- Bühler, B., B. Witholt, B. Hauer and A. Schmid (2002). "Characterization and application of xylene monooxygenase for multistep biocatalysis " Appl. Environ. Microbiol. **68**(2): 560-568.
- Carrondo, M. A., I. Bento, P. M. Matias and P. F. Lindley (2007). "Crystallographic evidence for dioxygen interactions with iron proteins " J. Biol. Inorg. Chem. **12**(4): 429-442.
- Chain, P. S., V. J. Denef, K. T. Konstantinidis, L. M. Vergez, L. Agullo, V. L. Reyes, L. Hauser, M. Cordova, L. Gomez, M. Gonzalez, M. Land, V. Lao, F. Larimer, J. J. LiPuma, E. Mahenthiralingam, S. A. Malfatti, C. J. Marx, J. J. Parnell, A. Ramette, P. Richardson, M. Seeger, D. Smith, T. Spilker, W. J. Sul, T. V. Tsoi, L. E. Ulrich, I. B. Zhulin and J. M. Tiedje (2006). "*Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility." Proc. Natl. Acad. Sci. U S A **103**(42): 15280-7.
- Chandrasena, R. E., K. P. Vatsis, M. J. Coon, P. F. Hollenberg and M. Newcomb (2004). "Hydroxylation by the hydroperoxy-iron species in cytochrome P450 enzymes." J. Am. Chem. Soc. **126**: 115-126.
- Delgado, A., M. G. Wubbolts, M.-A. Abril and J. L. Ramos (1992). "Nitroaromatics are substrates for the TOL plasmid upper-pathway enzymes." Appl. Environ. Microbiol. **58**(1): 415-417.
- Denef, V. J., J. A. Klappenbach, M. A. Patrauchan, C. Florizone, J. L. M. Rodrigues, T. V. Tsoi, W. Verstraete, L. D. Eltis and J. M. Tiedje (2006). "Genetic and genomic insights into the role of benzoate-catabolic pathway redundancy in *Burkholderia xenovorans* LB400." Appl. Environ. Microbiol. **72**: 585-595.
- Denef, V. J., J. Park, T. V. Tsoi, J.-M. Rouillard, H. Zhang, J. A. Wibbenmeyer, W. Verstraete, E. Gulari, S. A. Hashsham and J. M. Tiedje (2004). "Biphenyl and benzoate metabolism in a genomic conte outlining genome-wide metabolic networks in *Burkholderia xenovorans* LB400." Appl. Environ. Microbiol. **70**: 4961-4970.
- Denef, V. J., M. A. Patrauchan, C. Florizone and J. Park (2005). "Growth substrate- and phase specific expression of biphenyl, benzoate and C₁ metabolic pathwas in *B. xenovorans* LB400." J. Bacteriol. **187**: 7996-8005.
- Dutta, T. K. and I. C. Gunsalus (1997). "Reductase Gene Sequences and Protein Structures *p*-Cymene Methyl Hydroxylase." Biochem. Biophys. Res. Commun. **233**: 502-506.
- Eaton, R. W. (1997). "*p*-Cymene catabolic pathway in *Pseudomonas putida* F1: cloning and characterization of DNA encoding conversion of *p*-cymene to *p*-cumate." J. Bacteriol. **179**(10): 3171-80.

- Fang, J. M., C. H. Lin, C. W. Bradshaw and C. H. Wong (1995). "Enzymes in organic synthesis: oxidoreductions." J. Chem. Soc. **1**: 967 - 978.
- Fetzner, S. (2002). "Oxygenases without requirement for cofactors or metal ions." Appl. Microbiol. Biotechnol. **60**: 243-257.
- Fox, B. G., W. A. Froland, J. E. Dege and J. D. Lipscomb (1989). "Methane monooxygenase from *Methylosinus trichosporium* OB3b. Purification and properties of a three-component system with high specific activity from a type II methanotroph." J. Biol. Chem. **264**: 10023-10033.
- Fox, B. G., J. Shanklin, J. Ai, T. M. Loehr and J. Sanders-Loehr (1994). "Resonance raman evidence for an Fe-O-Fe center in stearoyl-ACP desaturase. Primary sequence identity with other diiron-oxo proteins." Biochemistry **33**(43): 12776 - 12786.
- Froland, W. A., K. K. Andersson, S.-K. Lee, Y. Liu and J. D. Lipscomb (1992). "Methane monooxygenase component B and reductase alter the regioselectivity of the hydroxylase component-catalyzed reactions. A novel role for protein-protein interactions in an oxygenase mechanism." J. Biol. Chem. **267**: 17588-17597.
- Groves, J. T. (2003). "The bioinorganic chemistry of iron in oxygenases and supramolecular assemblies." Proc. Natl. Acad. Sci. USA **100**: 3569-3574.
- Groves, J. T. (2006). "High-valent iron in chemical and biological oxidations." J. of Inorg. Biochem. **100**: 434-447.
- Harayama, S., H. Kishira, Y. Kasai and K. Shutsubo (1999). "Petroleum biodegradation in marine environments." Microbiol. Biotechnol. **1**: 63-70.
- Harayama, S., M. Kok and E. L. and Neidle (1992). "Functional and Evolutionary Relationships Among Diverse Oxygenases." Annu. Rev. Microbiol. **46**: 565-601.
- Hayaishi, O. (2005). "Fifty years of oxygen activation." J. Biol. Inorg. Chem. **10**: 1-2.
- Hayaishi, O., M. Katagari and S. Rothberg (1955). "Mechanism of the pyrocatechase reaction." J. Am. Chem. Soc. **77**: 5450.
- Hegg, E. L. and L. J. Que (1997). "The 2-His-1-Carboxylate Facial Triad — An Emerging Structural Motif in Mononuclear Non-Heme Iron(II) Enzymes." Eur. J. Biochem. **250**(3): 625-629.
- Hilyard, E. J., J. M. Jones-Meehan, B. J. Spargo and R. T. Hill (2007). "Enrichment, isolation, and phylogenetic identification of polycyclic aromatic hydrocarbon-degrading bacteria from Elizabeth River sediments." Appl. Environ. Microbiol. **74**(4): 1176-1182.

- Hoffart, L. M., E. W. Barr, R. B. Guyer, J. M. J. Bollinger and C. Krebs (2006). "Direct spectroscopic detection of a C-H-cleaving high-spin Fe(IV) complex in a prolyl-4-hydroxylase." Proc. Natl. Acad. Sci. U S A **103**(40): 14738-14743.
- Kok, M., R. Oldenhuis, M. P. van der Linden, P. Raatjes, J. Kingma, P. H. van Lelyveld and B. Witholt (1989). "The *Pseudomonas oleovorans* alkane hydroxylase gene. Sequence and expression." J. Biol. Chem. **264**(10): 5435-5441.
- Kunz, D. A. and P. J. Chapman (1981). "Isolation and characterization of spontaneously occurring TOL plasmid mutants of *Pseudomonas putida* HS1." J. Bacteriol. **146**(3): 952-964.
- Kurtz, D. M. (1997). "Structural similarity and functional diversity in diiron-oxo proteins" J. Biol. Inorg. Chem. **2**(2): 159-167.
- Leahy, J. G., P. J. Batchelor and S. M. Morcomb (2003). "Evolution of the soluble diiron monooxygenases." FEMS Microbiol. Rev. **27**: 449-479.
- Leahy, J. G. and R. R. Colwell (1990). "Microbial degradation of hydrocarbons in the environment." Microbiol. Rev. **54**(3): 305-15.
- Li, Z., J. B. van Beilen, W. A. Duetz, A. Schmid, A. de Raadt, H. Grieng and B. Witholt (2002). "Oxidative biotransformations using oxygenases." Curr. Opin. Chem. Biol. **6**: 136-144.
- Lipscomb, J. D. (1994). "Biochemistry of the soluble methane monooxygenase." Annu. Rev. Microbiol. **48**: 371-399.
- Lomri, N., J. Thomas and J. R. Cashman (1993). "Expression in *Escherichia coli* of the cloned flavin-containing monooxygenase from pig liver." J. Biol. Chem. **268**(7): 5048-5054.
- Lovenberg, W., Sobel, B.E. (1965). "Rubredoxin: a new electron transfer protein from *Clostridium pasteurianum*." Proc. Natl. Acad. Sci. U S A **54**(1): 193-199.
- Munro, A. W., D. G. Leys, K. J. McLean, K. R. Marshall, T. W. B. Ost, S. Daff, C. S. Miles, S. K. Chapman, D. A. Lysek, C. C. Moser, C. C. Page and P. L. Dutton (2002). "P450 BM-3 the very model of a modern flavocytochrome." TRENDS Biochem. Sci. **27**: 250-257.
- Newcomb, M., D. Aebisher, R. Shen, R. E. Chandrasena, P. F. Hollenberg and M. J. Coon (2003). "Kinetic isotope effects implicate two electrophilic oxidants in cytochrome p450-catalyzed hydroxylations." J. Am. Chem. Soc. **125**(20): 6064-6065.
- Nishio, T., A. Patel, Y. Wang and P. C. Lau (2001). "Biotransformations catalyzed by cloned *p*-cymene monooxygenase from *Pseudomonas putida* F1." Appl. Microbiol. Biotechnol. **55**(3): 321-5.

- Nordlund, P. and H. Eklund (1995). "Di-iron-carboxylate proteins." Curr. Opin. Struct. Biol. **5**: 758-766.
- Ortiz de Montellano, P. R. (1995). "Arylhydrazines as probes of hemoprotein structure and function." Biochimie **77**: 581-593.
- Parnell, J. J., J. Park, V. Denef, T. Tsoi, S. Hashsham, J. Quensen, 3rd and J. M. Tiedje (2006). "Coping with polychlorinated biphenyl (PCB) toxicity: Physiological and genome-wide responses of *Burkholderia xenovorans* LB400 to PCB-mediated stress." Appl. Environ. Microbiol. **72**(10): 6607-14.
- Peterson, J. A., D. Basu and C. M. J. (1966). "Enzymatic omega-oxidation. I. Electron carriers in fatty acid and hydrocarbon hydroxylation." J. Biol. Chem. **241**: 5162-5164.
- Ratajczak, A., W. Geibdorfer and W. Hillen (1998). "Alkane Hydroxylase from *Acinetobacter* sp Strain ADP1 Is Encoded by *alkM* and Belongs to a New Family of Bacterial Integral-Membrane Hydrocarbon Hydroxylases." Appl. Environ. Microbiol. **64**(4): 1175-1179.
- Ravichandran, K. G., S. S. Boddupalli, C. A. Hasermann, J. A. Peterson and J. Deisenhofer (1993). "Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450's." Science **261**: 731-736.
- Rosenzweig, A. C., C. A. Frederick, S. J. Lippard and P. Nordlund (1993). "Crystal structure of bacterial non-haem iron hydroxylase that catalyses the biological oxidation of methane." Nature **366**: 537-543.
- Rozhkova-Novosad, E. A., J. C. Chae, G. J. Zylstra, E. M. Bertrand, M. Alexander-Ozinskas, D. Deng, L. A. Moe, J. B. van Beilen, M. Danahy, J. T. Groves and R. N. Austin (2007). "Profiling mechanisms of alkane hydroxylase activity in vivo using the diagnostic substrate norcarane." Chem. Biol. **14**(2): 165-172.
- Schaechter, M. (2006). "From growth physiology to systems biology." Int. Microbiol. **9**: 157-161.
- Schlichting, I., J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko and S. G. Sligar (2000). "The catalytic pathway of cytochrome p450cam at atomic resolution." Science **287**: 1615-1622.
- Seeger, M., K. N. Timmis and B. Hofer (1995). "Degradation of chlorobiphenyls catalyzed by the bph-encoded biphenyl-2,3-dioxygenase and biphenyl-2,3-dihydrodiol-2,3-dehydrogenase of *Pseudomonas* sp. LB400." Appl. Environ. Microbiol. **61**: 2654-2658.
- Seeger, M., M. Zielinski, K. N. Timmis and B. Hofer (1999). "Regiospecificity of dioxygenation of di- to pentachlorobiphenyls and their degradation to

- chlorobenzoates by the bph-encoded catabolic pathway of *Burkholderia sp.* strain LB400." Appl. Environ. Microbiol. **65**: 3614-3621.
- Shan, X. and L. J. Que (2006). "High-valent nonheme iron-oxo species in biomimetic oxidations." J. Inorg. Biochem. **100**(4): 421-433.
- Shanklin, J., C. Achim, H. Schmidt, B. G. Fox and E. Münck (1997). "Mössbauer studies of alkane omega-hydroxylase: evidence for a diiron cluster in an integral-membrane enzyme." Proc. Natl. Acad. Sci. U S A **94**(7): 2981-2986.
- Shanklin, J. and E. B. Cahoon (1998). "Desaturation and related modifications of fatty acids." Annu. Rev. Plant Physiol. Plant Mol. Biol. **49**: 611-641.
- Shanklin, J. and E. Whittle (2003). "Evidence linking the *Pseudomonas oleovorans* alkane ω -hydroxylase, an integral membrane diiron enzyme, and the fatty acid desaturase family." FEBS Lett. **545**: 188-192.
- Shanklin, J., E. Whittle and B. G. Fox (1994). "Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in Alkane hydroxylase and xylene monooxygenase." Biochemistry **33**(43): 12787 - 12794.
- Shaw, J. F. and S. Harayama (1992). "Purification and characterisation of the NADH-acceptor reductase component of xylene monooxygenase encoding by the TOL plasmid pWW0 from *Pseudomonas putida* mt-2." Eur. J. Biochem. **209**: 51-61.
- Shaw, J. P., F. Schwager and S. Harayama (1992). "Substrate-specificity of benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase encoded by TOL plasmid pWW0. Metabolic and mechanistic implications." J. Biochem. **283**: 789-794.
- Sligar, S. G., T. M. Makris and I. G. Denisov (2005). "Thirty years of microbial P450 monooxygenase research: Peroxo-heme intermediates-- The central bus station in heme oxygenase catalysis." Biochem. Biophys. Res. Commun. **338**: 346-354.
- Smits, T. H., S. B. Balada, B. Witholt and J. B. van Beilen (2002). "Functional analysis of alkane hydroxylases from gram-negative and gram-positive bacteria." J. Bacteriol. **184**: 1735-1742.
- Smits, T. H. M., M. Rothlisberger, B. Witholt and J. B. van Beilen (1999). "Molecular screening for alkane hydroxylase genes in Gram-positive and Gram-negative strains." Environ. Microbiol. **1**(4): 307-317.
- Solomon, E. I., A. Decker and N. Lehnert (2003). "Non-heme iron enzymes: Contrast to heme catalysis " Proc. Natl. Acad. Sci. U S A **100**(7): 3589-3594.
- Stanier, R. Y., N. J. Palleroni and M. Duodoroff (1966). "The aerobic pseudomonads: a taxonomic study. ." J. Gen. Microbiol. **43**: 159-271.

- Stubna, A., D. H. Jo, M. Costas, W. W. Brennessel, H. Andres, E. L. Bominaar, E. Munck and L. Que, Jr. (2004). "A structural and Mossbauer study of complexes with $\text{Fe(2)(micro-O(H))(2)}$ cores: stepwise oxidation from $\text{Fe(II)(micro-OH)(2)Fe(II)}$ through $\text{Fe(II)(micro-OH)(2)Fe(III)}$ to $\text{Fe(III)(micro-O)(micro-OH)Fe(III)}$." Inorg. Chem. **43**(10): 3067-79.
- Suzuki, M., T. Hayakawa, J. F. Shaw, M. Rekik and S. Harayama (1991). "Primary structure of xylene monooxygenase: Similarities to and differences from the alkane hydroxylation system." J. Bacteriol. **173**(5): 1690-1695.
- Urlacher, V. B., S. Lutz-Wahl and R. D. Schmid (2004). "Microbial P450 enzymes in biotechnology." Appl. Microbiol. Biotechnol. **64**: 317-325.
- van Beilen, J. B. (1994). Alkane oxidation by *Pseudomonas oleovorans*: genes and proteins. Groningen, University of Groningen. **PhD**.
- van Beilen, J. B., W. A. Duetz, A. Schmid and B. Witholt (2003). "Practical issues in the application of oxygenases." TRENDS Biotechnol. **21**(4): 170-176.
- van Beilen, J. B. and E. G. Funhoff (2007). "Alkane hydroxylases involved in microbial alkane degradation." Appl. Microbiol. Biotechnol. **74**: 13-21.
- van Beilen, J. B., S. Panke, S. Lucchini, A. G. Franchini, M. Rothlisberger and B. Witholt (2001). "Analysis of *Pseudomonas putida* alkane-degradation gene clusters and flanking insertion sequences: evolution and regulation of the alk genes." Microbiology **147**(6): 1621-30.
- van Beilen, J. B., D. Penninga and B. Witholt (1992). "Topology of the membrane-bound alkane hydroxylase of *Pseudomonas oleovorans*." J. Biol. Chem. **267**(13): 9194-9201.
- van Beilen, J. B., T. H. Smits, F. F. Roos, T. Bruner, S. B. Balada, M. Röthlisberger and B. Witholt (2005). "Identification of an amino acid position that determines the substrate range of integral membrane alkane hydroxylases." J. Bacteriol. **187**: 85-91.
- van Beilen, J. B., M. Wubbolts, G. and B. Witholt (1994). "Genetics of alkane oxidation by *Pseudomonas oleovorans*." Biodegradation **5**: 161-164.
- van Beilen, L. B., T. H. Smits, L. G. Whyte, S. Schorcht, M. Rothlisberger, T. Plaggemeier, K. H. Engesser and B. Witholt (2002). "Alkane hydroxylase homologs in Gram-positive bacteria." Environ. Microbiol. **4**: 676-682.
- Wallar, B. J. and J. D. Lipscomb (1996). "Dioxygen activation by enzymes containing binuclear non-heme iron clusters." Chem. Rev. **96**: 2625-2658.

- Wallar, B. J. and J. D. Lipscomb (2001). "Methane monooxygenase component B mutants alter the kinetics of steps throughout the catalytic cycle." Biochemistry **40**(7): 2220-2033.
- Worsey, M. J. and P. A. Williams (1975). "Metabolism of toluene and xylenes by *Pseudomonas putida* (arvilla) mt-2: evidence for a new function of the TOL plasmid." J. Bacteriol. **124**: 7-13.
- Wubbolts, M., G., P. Reuvekamp and B. Witholt (1994). "TOL plasmid-specified xylene oxygenase is a wide substrate range monooxygenase capable of olefin epoxidation." Enz. Microb. Technol. **16**(7): 608-615.
- Zhang, J., B. J. Wallar, C. V. Popescu, D. B. Renner, D. D. Thomas and J. D. Lipscomb (2006). "Methane Monooxygenase Hydroxylase and B Component Interactions." Biochemistry **45**(9): 2913-2926.

Curriculum Vita

Marie Carmen Montes-Matías

October 2008	Rutgers, the State University of New Jersey (New Brunswick) Ph.D. Microbiology and Molecular Genetics
June 2001	University of Puerto Rico (Mayaguez) Industrial Microbiology