University of Tennessee, Knoxville
Trace: Tennessee Research and Creative Exchange

Masters Theses | Graduate School

8-2004

Cloning and Expression of Cytochrome $P450_{cam}$

Sirisha Chada
University of Tennessee, Knoxville

Recommended Citation

Chada, Sirisha, "Cloning and Expression of Cytochrome $P450_{cam}$. " Master's Thesis, University of Tennessee, 2004.
https://trace.tennessee.edu/utk_gradthes/4320

This Thesis is brought to you for free and open access by the Graduate School at Trace: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Masters Theses by an authorized administrator of Trace: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

To the Graduate Council:

I am submitting herewith a thesis written by Sirisha Chada entitled "Cloning and Expression of Cytochrome $P450_{cam}$." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Chemical Engineering.

Abhijeet P. Borole, Major Professor

We have read this thesis and recommend its acceptance:

Brian H. Davison, Paul D. Frymier

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a thesis written by Sirisha Chada entitled "Cloning and expression of cytochrome $P450_{cam}$." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Chemical Engineering.

Abhijeet P. Borole

Major Professor

We have read this thesis
and recommend its acceptance

Brian H. Davison

Paul D. Frymier

Accepted for the Council

Anne Mayhew

Vice Chancellor and
Dean of Graduate Studies

(Original Signatures are on file with official student records.)

CLONING AND EXPRESION OF CYTOCHROME $P450_{CAM}$

A thesis presented for the Master of Science Degree

University of Tennessee, Knoxville

Sirisha Chada

August 2004

Copyright© 2004 by Sirisha Chada

All rights reserved

To

My Beloved Parents

ACKNOWLEDGMENTS

I would like to thank my parents Chada Sri Rama Chandra Murthy and Satya Swarajya Lakshmi by dedicating this thesis to them. Their unconditional love and encouragement have helped me get this far in life. They have inspired me to succeed in every endeavor of mine by lending their support. They are my teachers; my best friends and they are everything for me.

I would like to thank my major professor Dr. Abhijeet P. Borole for providing me with an opportunity to work with him on this project at the esteemed Oak Ridge National Laboratory. Under his facile guidance I have learned many new techniques and evolved as an engineer. He has been a great mentor and has helped me improve my intellectual faculties by constantly challenging me with his high standards. I cannot thank him enough for his support during my initial days as a graduate student. Dr. Brian H. Davison and Dr. Paul D. Frymier have been very supportive throughout the process and have rendered their valuable guidance, which I am grateful for. I would like to thank Dr. Paul R. Beinkowski for supporting my studies for the last five months.

I would like to thank Choo Y. Hamilton for her patience and support while working with me in the lab and Miguel Rodriguez for being such a good friend. I would like to thank all the former and the current staff of Biochemical Engineering Research Group for their support and encouragement for the last two years. Working with all of you has been an

enriching experience. I'm indebted to all my teachers from my childhood to date who have imparted wisdom and have contributed to my acumen.

My dear friends Deepti, Patanjali, Sridhar, Kiran, K.R., have been there for me always. Thank you very much for always believing in me and encouraging me to succeed. I would like to express my profound gratitude to Kalyana Ganti for being my friend, philosopher and guide. This thesis would not have been complete if it were not for you. Your patience, kindness and affection have helped me survive the most difficult phase of my life.

ABSTRACT

Polyaromatic hydrocarbons (PAHs) are ubiquitously present in the environment and have deleterious affects on humans. Cytochrome P450 enzymes have been found to be useful in the biodegradation of the PAHs through an initial hydroxylation step. This makes them more water-soluble and accessible to the action of other oxidases. Cytochrome $P450_{cam}$ is a model P450 enzyme that has been mutated for enabling PAH and alkane hydroxylation. P450s therefore have potential for use in commercial process for production of primary alcohols.

Rhodococcus organisms have a special mechanism for transporting the hydrophobic hydrocarbons into their cells. Cloning of the $P450_{cam}$ gene into *Rhodococcus* can potentially provide a hydroxylation system with high conversion rates. To enable this, cytochrome $P450_{cam}$ gene (*CYP101*) was cloned into a *Rhodococcus-Escherichia coli* shuttle vector. This would provide tools to modify the gene and target specific PAHs. This study is aimed at developing a *Rhodococcus*-based biocatalyst using cytochrome P450 enzymes.

The cytochrome $P450_{cam}$ gene was synthesized using PCR (polymerase chain reaction); the product was cloned into pKSD6-1 (a *Rhodococcus-Escherichia coli* shuttle vector) and transformed into *Escherichia coli*. The recombinant protein in *Escherichia coli* was produced by induction and the protein extract was analyzed using spectrophotometric analysis, camphor binding and carbon monoxide binding assays.

The results showed a Soret band at 414nm but no carbon monoxide or substrate binding was observed. The control culture grown under similar conditions with only the pKSD6-

1 gene did not show any absorption band at 414nm. These results together indicate that the gene has been cloned into the recombinant plasmid, but the enzyme being produced might be in an inactive form.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Background

Proteins are the molecular instruments through which the genetic information is expressed. They are the most abundant biomolecules and are diverse in their functions. Enzymes are highly specialized proteins and are central to every biochemical process; they serve as reaction catalysts of the biological systems. Enzymes are more desirable than their chemical catalysts because of their superior catalytic power, high degree of specificity for their substrates and their ability to function in aqueous solutions under mild conditions of temperature and pressure. These properties of the enzymes have led to important applications in the fields of medicine, chemical industry, food processing and agriculture. [1]

Oxygenases are enzymes that catalyze oxidative reactions in which the oxygen atoms are directly incorporated into the substrate molecule forming a new hydroxyl or a carboxyl group. Among these are dioxygenases, which catalyze reactions in which both the oxygen atoms are incorporated into the organic substrate molecule and monooxygenases, that are more abundant in nature, catalyze reactions in which only one of the oxygen atoms is incorporated into the

substrate while the other is reduced to water. The main substrate accepts one of the oxygen atoms and the co-substrate furnishes the Hydrogen atom to reduce the other oxygen atom. The general reaction catalyzed by the monooxygenases is

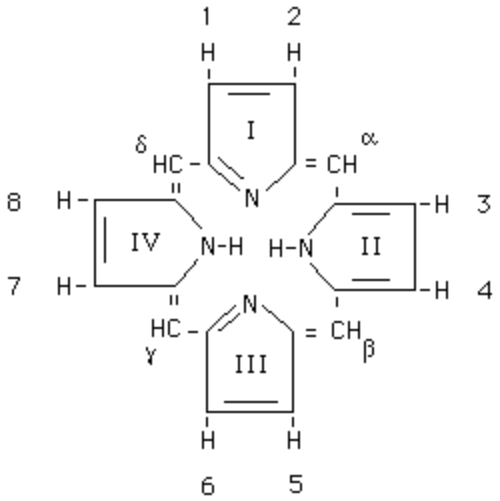
$$AH + BH_2 + O - O \longrightarrow A - OH + B + H_2O \qquad (1)$$

In the above equation AH is the main substrate and BH_2 is the co-substrate. Monooxygenases are also known as hydroxylases as they catalyze reactions in which the main substrate becomes hydroxylated. They are also known as mixed function oxidases or mixed-function oxygenases to indicate that they oxidize two different substrates simultaneously. The co-substrate used by the monooxygenases is a reduced flavin nucleotide such as $FADH_2$ or $FMNH_2$, some enzymes use NADH or NADPH while a very few use α-ketoglutarate. [1]

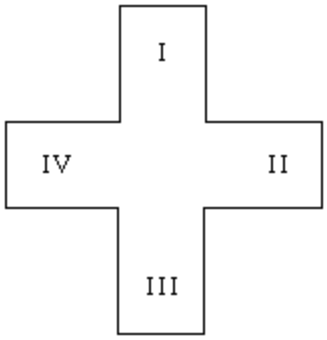
A type of heme protein called cytochrome P450 (CYP) catalyzes a vast number of very complex monooxygenation reactions. It is capable of catalyzing the oxidation of the C-H bonds to alcohol functionality. The P450 enzymes have the ability to selectively oxidize an array of endogenous and exogenous organic compounds ranging from medium chain alkanes such as n-heptane and n-octane to steroidal, polyaromatic and very large molecules such as triterpenes without further oxidation to aldehydes and ketones. This is therefore a novel alternative to the industry where all current alkane oxidation processes use elevated temperatures and pressures and metal oxide catalysts, which do not have the required activity, selectivity or specificity. [2]

All heme proteins consist of a porphyrin ring, which binds the iron atom. Porphyrins are planar ligands with a central ring structure and various substituent groups at the edges of the ring as shown in Figure 1. The central ring structure of the porphyrins consists of four pyrrole rings and the structure of a pyrrole ring is shown in Figure 2. Heme is a metalloporphyrin with the complex organic ring structure of a protoporphyrin with the iron atom bound in its ferrous (Fe^{2+}) state. Metalloporphyrins, contain a metal atom bound to a porphyrin and are synthesized naturally and utilized biologically as redox catalysts. These metalloporphyrins possess characteristic absorption spectra that can be used for identification and quantitative estimation of the porphyrins and porphyrin bound compounds. These metal complexes are incorporated into many proteins as prosthetic groups. Within the protein framework the prosthetic group is a versatile tool with highly specialized functions.

Heme serves as a prosthetic group among various classes of proteins such as the globins, catalases, hydrogen-binding peroxidases, cytochrome P450 etc. As seen in the Figure 3 the iron atom of the heme has six coordination bonds, four to the nitrogen atoms of the porphyrin ring system and two, perpendicular to the plane of the porphyrin .The unique C-H activation activity of cytochrome P450 has no equivalent in classical synthetic methodologies.

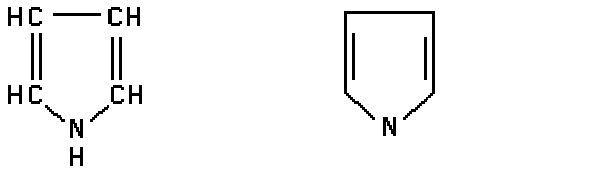


Porphin, showing the four pyrrole rings and the Roman numerals which designate them. Arabic numbers indicate positions at which substituents may be attached. Greek letters denote the methene bridges.



Schematic representation of porphin.

Figure 1: Structure of porphyrin



Pyrrole

Abbreviated version of pyrrole

Figure 2: Structure of Pyrrole Ring

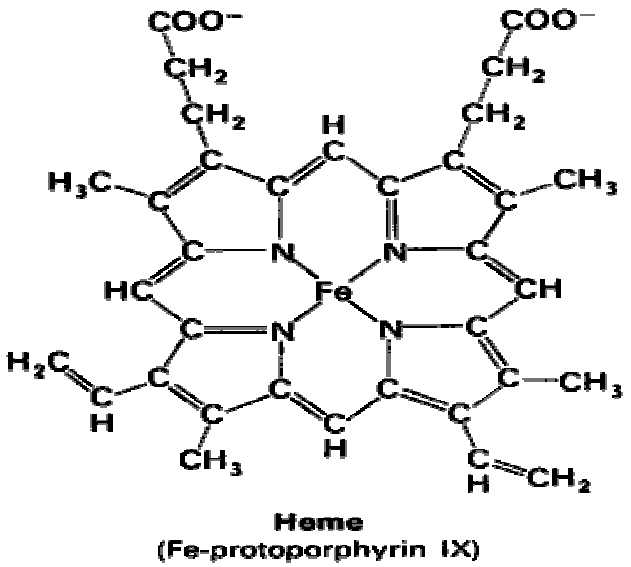


Figure 3: Structure of Heme

The electron-donating nitrogen atoms to which the iron atom is bound in the heme prevent the oxidation of the iron to ferric state thereby preserving its reversible oxygen binding capacity. This is one of the reasons why the heme group is embedded deep inside the proteins that contain it. They also generally have a His (Histidine) residue occupying one of the coordinate bonds leaving only one site for the oxygen to bind. Small molecules such as carbon monoxide (CO) and Nitric oxide (NO) coordinate to the heme iron with an affinity greater than that of oxygen. [1]

Among the cytochrome P450 enzymes, cytochrome $P450_{cam}$ was the first P450 enzyme to be structurally characterized and has been studied in detail. Cytochrome $P450_{cam}$ from the soil bacterium *Pseudomonas putida* is responsible for

the stereo specific methylene hydroxylation of camphor as the first step in the compound's catabolism when used as the sole carbon source. Cytochrome $P450_{cam}$ is a soluble enzyme that is obtainable in relatively large amounts with a high purity. It has been used extensively as a model to investigate the chemical and biophysical mechanisms of P-450 dependent mechanisms. [2]

It has been shown that cytochrome $P450_{cam}$ can be engineered to oxidize many different classes of compounds that are structurally unrelated to its natural substrate camphor. Increasing the hydrophobicity of the active site in $P450_{cam}$ by random mutagenesis is shown to increase its activity towards hydrophobic molecules that are both smaller than camphor like simple alkanes and larger ones such as naphthalene and pyrene. [3]

Polycyclic aromatic hydrocarbons (PAHs) are hazardous and are persistent in the environment. They are procarcinogens and their carcinogenicity is expressed by their oxidation by the human cytochrome P450 enzymes such as CYP1A1. The inability of the PAHs to be soluble in water and the lack of naturally occurring microbial enzyme systems for their efficient degradation are the reasons for their persistence in the environment. Initial oxidation by the P450 monooxygenases and non-heme dependent dioxygenases has been reported to be the probable rate-limiting step after the microorganism sequesters the PAHs. All the P450 enzymes follow the same mechanism of generating the ferryl intermediate

species, which attacks a variety of substrates. This combined with the fact that P450 enzymes are involved in the PAH carcinogenesis of the mammals and the PAH bioremediation by the microorganisms provides reason enough to investigate methods which will enable the use of these enzymes to degrade the PAHs. [3]

1.2 Overview

This section introduces objective, concept and experimental philosophy of the cloning of the cytochrome $P450_{cam}$ gene into the *Rhodococcus-Escherichia coli* shuttle vector pKSD6-1 (obtained from Peter C.K. Lau and Matthew J. Grossman, Biotechnology Research Institute, National Research Council, Quebec, Canada). The basic idea behind cloning cytochrome $P450_{cam}$ into a *Rhodococcus-Escherichia coli* shuttle vector is to produce the camphor-binding enzyme in the new host *Rhodococcus* and modify the enzyme genetically using random mutagenesis so that the enzyme will be able to act on polycyclic aromatic hydrocarbons and in the production of primary alcohols. This involved firstly cloning the recombinant pKSD6-1 +*CYP101* plasmid into *Escherichia coli* and then transforming it into *Rhodococcus*.

The hydrocarbon-degrading enzymes are a part of the cytoplasmic membrane system, so the hydrocarbon substrate has to be transported through the hydrophilic cell wall into the cell to be metabolized. This step of hydrocarbon

intake is strongly limited by the hydrophobic character of these substrates. *Rhodococcus* is one of the organisms that are able to utilize hydrocarbons; it is able to do so by changing the physical-chemical character of its cell wall by inserting amphiphilic molecules into its surface layers thereby assuring the efficient transport by a strong cell attachment to the hydrophobic substrate. [4]

Rhodococcus organism therefore has an extraordinary capacity for metabolizing recalcitrant organic compounds and has a demonstrated solvent tolerance, which make it a suitable host for cloning *CYP101*. The full exploitation of *Rhodococcus* is limited by the lack of genetic tools for manipulating the organism. The lack of such knowledge in *Rhodoccocus* is compensated by the extensive knowledge on the genetics and the molecular biology of *Escherichia coli*, which provide all the necessary tools to facilitate its use while performing random mutagenesis to create a biocatalyst, which can act on substrates other than camphor. [5]

Cloning of the cytochrome $P450_{cam}$ gene into the *Rhodococcus-Escherichia coli* shuttle vector pKSD6-1 consisted of producing the gene from pcWoriP450_{cam} plasmid by polymerase chain reaction (PCR), cloning the gene into the commercially available pCR2.1 TOPO vector (Invitrogen Corporation, Carlsbad, CA), obtaining the *CYP101* gene from the recombinant Topo+*CYP101* plasmid ligating it with the pKSD6-1 plasmid and transforming it into *Escherichia coli*.

Analysis of the pKSD6-1 plasmid showed that cloning the *CYP101* gene from the plasmid pCWori+*CYP101* into pKSD6-1 was not a feasible task because of the unavailability of appropriate restriction sites to directly insert the gene into the vector. It was therefore essential to use an intermediate vector, in this case the commercially available pCR2.1 TOPO in order to achieve this end. The *CYP101* gene was produced from the pCWori+*CYP101* plasmid using Polymerase chain reaction using primers P450 1F and P4501R that were designed based on the standard design conventions stated in Sambrook Cloning Manual. [6]

The PCR product of *CYP101* gene was analyzed using gel electrophoresis and the gene was cloned into the pCR2.1 TOPO vector and transformed into the Top10F' chemically competent *Escherichia coli* cells (Invitrogen, Carlsbad, CA). Among the various positive clones the one which yielded the *CYP101* gene (1.4 kbp size fragment) upon restriction analysis was selected. This recombinant plasmid was named 3-TOPO+*CYP101*.

The sequence of *CYP101* that was amplified using P4501F and P4501R primers was inserted into the gene sequence of the commercially available pCR2.1 TOPO and a recombinant gene sequence of the 3-TOPO+*CYP101* was created. This recombinant gene sequence was analyzed using the NEBcutter V2.0 tool for the appropriate restriction sites to clone the *CYP101* gene into pKSD6-1.

Upon analysis, the ApaI site was found to be the most suitable site for use in both the vector and the insert. Both the vector pKSD6-1 and the plasmid containing the insert 3-TOPO+*CYP101* were restricted with ApaI to get the desired "sticky" ends. The recombinant cytochrome $P450_{cam}$ and pKSD6-1 was then transformed into *Escherichia coli* (DH5α) cells and a positive clone was isolated. The double resistance of pKSD6-1 to ampicillin and chloramphenicol antibiotics was used in selecting the positive clone.

The positive clone containing the recombinant *CYP101* and pKSD6-1 was grown and the cytochrome $P450_{cam}$ enzyme production was induced in the cells using IPTG (Isopropyl β-D-thiogalactopyranoside). The protein extract from the cells was analyzed using spectrophotometry, camphor binding and carbon monoxide binding assays. These analyses revealed that the *CYP101* gene has been cloned into the pKSD6-1 shuttle vector and the enzyme that was produced was in its inactive form. This present work is a detailed compilation of the experimental procedures undertaken and analysis of the results.

Chapter 2

LITERATURE REVIEW

2.1 Cytochrome P450

Background

Cytochrome P450 enzymes are a family of b-type heme containing proteins, their ubiquity is evident by their presence everywhere in nature from Archaebacteria to humans with the exception of enterobacteria. As of January 18, 2004 the number of named cytochrome P450s include 1277 animal sequences, 1098 plant sequences, 207 lower eukaryote sequences and 461 bacterial sequences. That is a total of 3043 different named sequences and there are many more to be named**.** [7,8]

The name P450 was given to indicate a pigment ('P') with a heme Fe^{2+}-carbon monoxide complex having an absorption spectral band at 450nm. The cytochrome P450 enzymes catalyze monooxygenation of a number of aromatic and aliphatic compounds in numerous metabolic and biosynthetic pathways. In the biodegradative pathways of the eukaryotes they are responsible for the metabolism of most xenobiotics as well as pharmaceuticals. This is important

because though the oxidation of the xenobiotics is a prelude to their excretion sometimes they get converted into reactive metabolites that cause deleterious side effects. [9,10]

In plants they play active roles in herbicide resistance and biosynthetic pathways such as morphine biosynthesis, in fungi they are involved in the antibiotic synthesis; they synthesize resistance to plant toxins in insects and in the metabolic and synthetic pathways of bacteria. The most common reaction catalyzed by the cytochrome P450 is hydroxylation but forms that catalyze epoxidation, dealkylation and sulfoxidation are also known. The P450 reaction generally requires two electrons (which are provided by its redox partner), dioxygen and two protons. [7]

There are three classes of P450 systems (shown in Figure 4); Class I are three component protein systems in which the redox partners consist of NADH-dependent FAD containing ferredoxin reductase and ferredoxin, examples of the class I P450s are most bacterial and all eukaryotic mitochondrial P450s. Class II P450 systems such as the eukaryotic microsomal (like those associated with the endoplasmic reticulum) are two component systems in which the redox partner is the NADPH-dependent diflavin (containing FAD and FADH) reductase. Class III consists of a single component, which is a fusion of both the components, representative of such a system, is the bacterial P450BM3. [7]

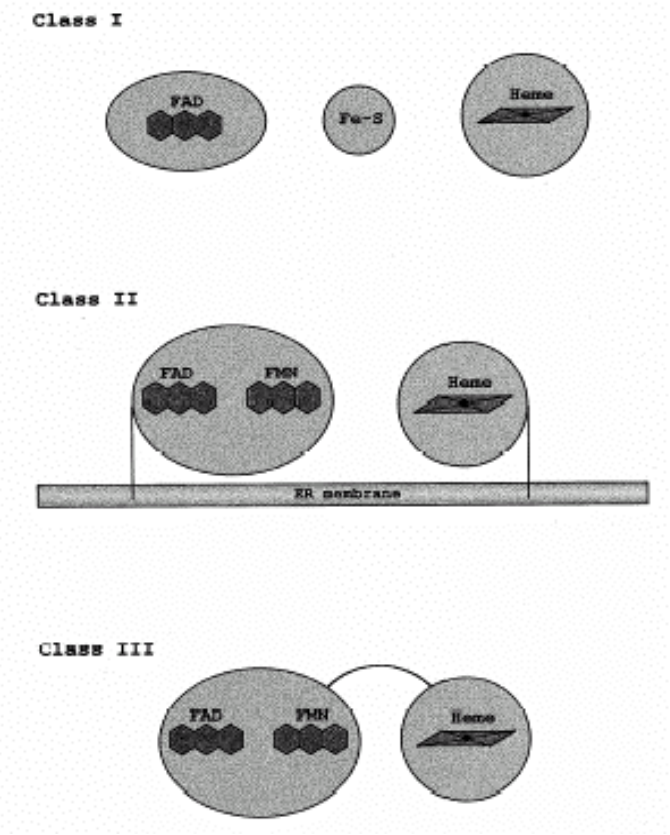


Figure 4: Various classes of P450 systems. [7]

Reactions catalyzed by Bacterial P450s

P450s contain an obligate heme cofactor that is linked to the peptide backbone of the protein via a thiolate ligand provided by a cysteine residue. This cofactor allows the enzyme to activate molecular oxygen to produce a high valent iron-oxo species that is a powerful oxidant and enables these enzymes to perform a range of oxidative transformations. Differences in the substrate specificities of the various P-450s come from the different peptide backbones. *Escherichia coli* and *Salmonella typhymurium* contain no P450s at all. The Table 1 lists the various types of reactions catalyzed by the bacterial P450 enzymes, the species of their origin and the various substrates they act upon. [11]

Comparison of the bacterial and mammalian cytochrome P450s

Rheinwald and Gunsalus (in 1973) were the first to show that the genetic information for the proteins involved in camphor degradation is located on a plasmid. There are some similarities between the membrane bound mammalian mitochondrial steroid hydroxylating systems such as *CYP11A1*, *CYP11B1*, *CYP11B2* and the bacterial *CYP101* systems. The mammalian systems receive reduction equivalents from NADPH via membrane–associated adrenoredoxin reductase (AdR) and adrenodoxin (Adx), a soluble [2Fe-2S] cluster containing ferredoxin while the CYP101 system receives NADH via an FAD-containing putidaredoxin reductase (PdR) and a [2Fe-2S] cluster-containing ferredoxin,

Table 1: Reactions catalyzed by bacterial P450s [11]

Enzyme	Gene	Original Host Organism	Reaction catalyzed	Substrate	Product
$P450_{cam}$	CYP101	*Pseudomonas putida*	Stereo specific hydroxylation	(1R)-(+)-camphor	5-*exo*-hydroxycamphor
$P450_{terp}$	CYP108	*Pseudomonas species*	Stereo selective hydroxylation	α-terpineol	
$P450_{lin}$	CYP111	*Pseudomonas species*	Stereo selective hydroxylation	Linalool	
$P450_{cin}$	CYP176 A1	*Citrobacter braakii*	Stereo selective hydroxylation	Cineole	
$P450_{camr}$		*Rhodococcus sp*		(1R)-(+)-camphor	6-endo hydroxycamphor
$P450_{PicK}$	CYP107L 1	*Streptomyces venzuelae*	1. Hydroxylation Of the C-10 and/or C-12 position . 2. Dehydroxylation of the 12-membered macrolide at C-10 position followed by the hydroxylation at the C-12 position.	Macrolides	Methymycin Neomethymycin Novamethymycin 14-membered macrolide ,picromycin
$P450_{NovI}$		*Streptomyces spheroids Streptomyces niveus*	Hydroxylation		L-tyrosine β-hydroxytyrosine
Thermop-hilic P450	CYP119	*Sulfolobus solfataricus*	Hydroxylation	Long-chain fatty acids	
P450BM 3	CYP102 A1	*Bacillus megaterium*	Hydroxylation	Long chain fatty acids	
P450Eth B		*Rhodococcus ruber*	Bioremediation of industrial wastes by hydroxylation.	Methyl and ethyl *t*-butyl etherand methyl *t*-pentyl ether.	
P450sca		*Streptomyces carbophilus*	Lowering cholestrol by Hydroxylation using pravastatin	cholestrol	
$P450_{EpoK}$		*Sorangium cellulosum*SMP44	*In vitro* epoxidation	Epithilone D	Epithilone B

putidaredoxin. Figure 5 shows a schematic presentation of the soluble bacterial *CYP101* system and of the mammalian mitochondrial *CYP11A1* system. Arrows indicate the possibility of transferring of first electron to the heterologous P450. [10]

Though the electrons can be transferred to a heterologous P450 (i.e. CYP101 or *CYP11A1*) by putidaredoxin and adrenodoxin, the ferredoxins in the systems cannot substitute for each other. Even though the ferredoxins cannot substitute each other it has been shown that a system containing PdR, Pdx and microsomal P450 *CYP2B4* converts substrate effectively. Observed reaction rates of both the systems range from 300 per minute for the CYP101 to 100 per minute for the

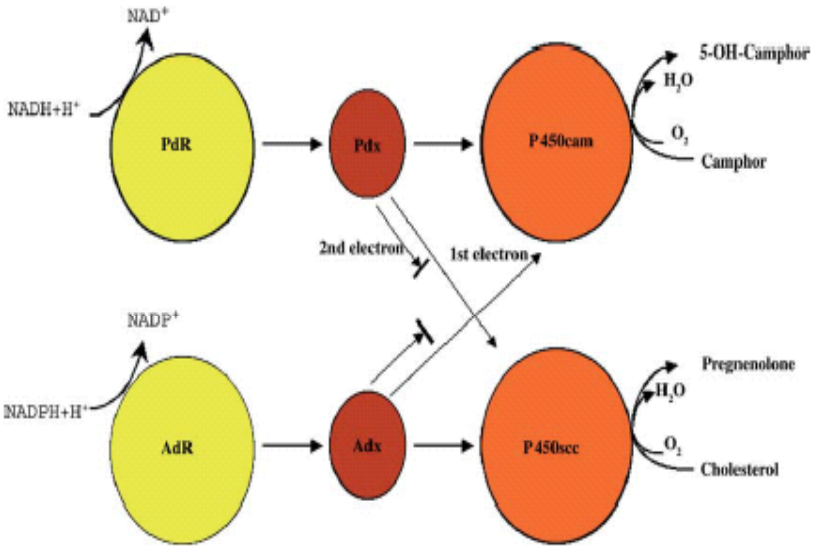


Figure 5: Electron transfer systems in *CYP101* and *CYP11A1* [10]

CYP11A1, (catalyzes the first step in the conversion of cholesterol to prenenolone). The difference in the reaction rates can be attributed to the different velocities, which are the defining point of the fast bacterial and the slow mammalian P-450 systems. [10]

Primary structures of the two ferredoxins reveal a 40% structural similarity with 38 identical amino acids and 18 conservative substitutions. When the primary structures of both the ferredoxins are aligned, the [2Fe-2S] cluster-coordinating residues that consist of four residues in both the proteins are found at the same relative positions. Both Adx and Pdx are examples of vertebrate-type ferredoxins and their tertiary structure reveals the same overall topology. The Figure 6 gives

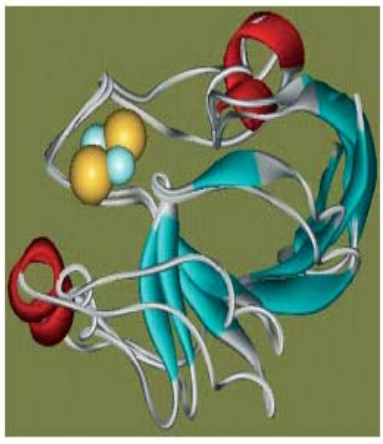


Figure 6: 3D structures of Adx (top)
and Pdx (bottom). [10]

the 3D structures of Adx and Pdx. Interaction sites of the mitochondrial cytochrome P-450s are different than those of the CYP101. There is a major difference in the C-terminal of both the ferredoxins, Adx has an extended C-terminal region and it does have an aromatic residue, which is the very important W106 (Tryptophan) in Pdx. These are a few of the differences and similarities between the two P-450 systems. [10]

Properties

Crystalline $P450_{cam}$ contains one heme per 46,000 molecular weight; it has one mole of aspartate or asparagines as its NH_2 (amino) terminal amino acid. It is completely homogenous when examined by electrophoretic mobility and hydrodynamic properties. It gives a single protein band over a pH range of 4.0 to 8.8 in a cellulose acetate strip electrophoresis. It has an isoelectric point at a pH of 4.7. Molecular weight of $P450_{cam}$ varies between 44,000 and 46,000. The purified $P450_{cam}$ possesses a single molecule of ferriprotoporphyrin IX (heme) as a prosthetic group. The heme can be removed by several methods and the apoprotein thus obtained can be reconstituted to active $P450_{cam}$ with a commercial or ^{57}Fe-replaced heme. [12]

The α and β bands of oxidized form of $P450_{cam}$ shows absorption spectra at 571 and 538nm respectively. The Soret band is at 417nm; a shoulder or γ band is at

365nm, the aromatic amino acid content at 280nm with a shoulder at 290nm. Reduction of $P450_{cam}$ with dithionite causes a shift in the Soret band from 417nm to 411 nm, the α and β bands are fused into a single absorption band at 540nm. The amount of active $P450_{cam}$ in the sample is directly proportional to the Soret absorption band. [12]

Reduced $P450_{cam}$ reacts with carbon monoxide to form a complex and causes the Soret maximum to shift to 446nm; the α and β bands are fused at 550nm. Carbon monoxide doesn't alter the absorption spectra of the ferric (oxidized) $P450_{cam}$. Cytochrome $P450_{cam}$ is stabilized by the addition of camphor both during purification and storage. [12]

Camphor free enzyme solutions show a gradual loss of enzymatic activity on storage for long periods of time compared to the camphor bound ones. The camphor bound enzymes are more stable because the binding of camphor the enzyme's natural substrate helps it preserve its tertiary structure. Lower protein concentrations augment the loss in enzymatic activity. Camphor bound enzyme has absorption maxima at 280, 392 and 643 nm. Addition of dithionite to an anaerobic solution reduces the heme and the spectral maxima shift to 409nm and 545nm. When carbon monoxide is bubbled through the reduced enzyme, it results in the absorption maxima to 446nm and 552nm. [12, 13]

Phenylimidazoles inhibit $P450_{cam}$ activity at low concentrations and replace the substrate. DPNH (diphosphopyridine nucleotide dehydrogenase) and KCN (Potassium Cyanide) at high concentrations are inhibitory to the enzyme. In the absence of camphor, a high concentration of DPNH causes permanent damage to $P450_{cam}$ that is irreversible. In the presence of camphor $P450_{cam}$ is slowly reduced by camphor. Activation of $P450_{cam}$ by glycerol is directly proportional to its concentration. Sulfhydryl compounds such as cysteine, homocysteine, and thioglycolic acid, glutathione and cystathione increase the activity of the enzyme which probably is due to the conversion of the small amounts of enzymatically inactive P-420 present to $P\text{-}450_{cam}$. [12]

Crystal Structure

Crystals obtained from the ammonium sulfate solutions in the orthorhombic II form provide good details about the structure of the enzyme. $P450_{cam}$ is an asymmetrically shaped protein resembling a triangular prism of approximately 30 Å with a maximum dimension of 60 Å. There are twelve helical segments labeled from A to L which dominate the structure and account for ~45% of the residues. Four anti-parallel β pairs make up about 15% of the structure. There are no parallel β pairs or extended sheet structure. The helices are arranged in three layers stacked on top of one another and the heme is embedded in the layers. The regular triangle shape of the enzyme is a result of the efficient packing of four molecules in a unit cell; a reason for the low water content of the crystals. [14]

The helices E, F, G, I form an antiparallel Greek key helical bundle; a common structural motif found in globins, cytochrome *c*, etc. Helices L and I are almost parallel to one another; helices D and L are also in approximate parallel alignment. The remainder of the structure is a mix of helices interspersed with β pairs. The number of residues in helical conformation is equal in the N and the C terminal halves of the molecule while majority of the β structure is in the C terminal domain. [14]

The heme is located between the proximal L helix and the distal I helix. The proximal helix runs along the molecular surface while the distal helix runs through the center of the center of the molecule. The heme is approximately 8Å away from the surface but it is not exposed to the bulk solvent. The hydrogen donors such as Arg112, Arg299, His35, Gln108, Asp297 and a water molecule surround the heme propionate groups. One of the heme propionate groups forms an unusual interaction of carboxyl-carboxyl hydrogen bond with the side chain of Asp297 and the Arg299 forming an ion pair with the propionate with its guanidium group only 4.0 Å away from the Asp297 carboxyl group, its positive charged counterion thus aiding in stabilization of the carboxyl-carboxyl hydrogen bond. [14]

Catalytic cycle of cytochrome $P450_{cam}$

Cytochrome $P450_{cam}$ (CYP101) catalyzes the insertion of one oxygen atom from O_2 into camphor thereby converting it into 5-*exo*- hydroxycamphor; this is the first step in the camphor catabolism by the soil bacterium *Pseudomonas putida*. There are two electrons required for the camphor hydroxylation by CYP101, which come from the oxidation of NADH, which is in turn catalyzed by putidaredoxin reductase (PdxR). A $Cys_4Fe_2S_2$ ferredoxin, putidaredoxin (Pdx) shuttles the electrons from PdxR to CYP101. [15]

The catalytic cycle of *Pseudomonas putida* $P450_{cam}$ (as shown in Figure 7) consists of reversible substrate binding, which converts the six-coordinate, low-spin met form of the protein to the five-coordinate, high-spin Fe^{III} camphor complex, thereby increasing the reduction potential of the heme iron atom. The addition of the first electron reduces the enzyme to the five-coordinate Fe^{II} camphor complex. Oxygen binds to the five-coordinate Fe^{II} camphor complex and gives the six-coordinate ferrous-dioxy (Fe^{II}-O_2) intermediate. Addition of a second electron and two protons followed by cleavage of the oxygen-oxygen bond produces a molecule of water and an oxidizing species, the so-called activated oxygen intermediate; and insertion of the iron-bound oxygen into the substrate produces 5-*exo*-hydroxycamphor and product is released. [16]

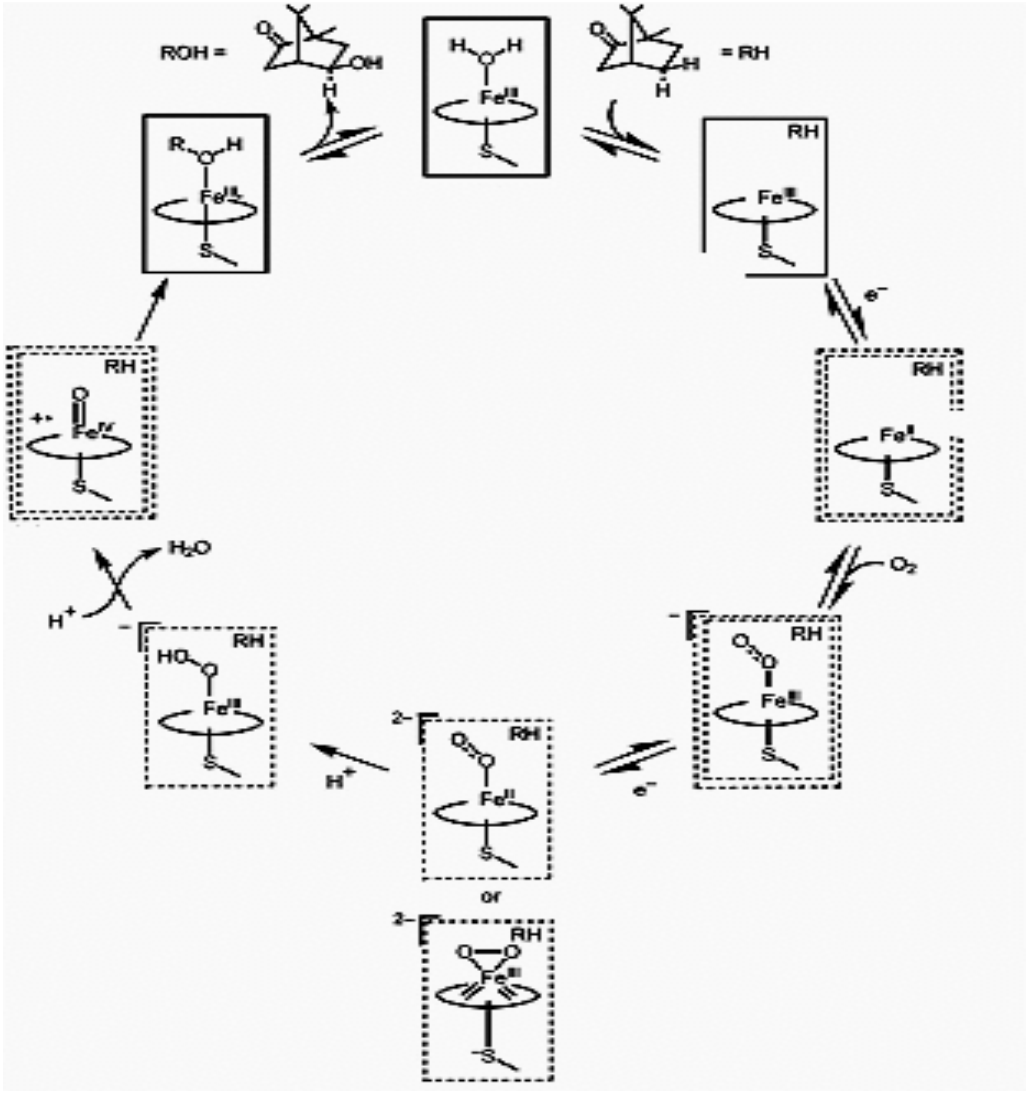


Figure 7: Catalytic cycle of *Pseudomonas putida* $P450_{cam}$ [17]

Putidaredoxin reductase and Putidaredoxin

There are two types of electron transfer systems associated with monooxygenation reactions catalyzed by P-450s 1) those requiring a flavoprotein dehydrogenase, an iron-sulfur protein, and the P-450; and 2) those requiring a FAD and FMN containing flavoprotein dehydrogenase and the P-450. An example of the former class is $P\text{-}450_{cam}$ and they are usually associated with mitochondrial steroidogenesis, while the latter is associated with microsomal drug or steroid metabolism and P-450BM3. $P\text{-}450_{cam}$ has been cloned, sequenced, and expressed in an active form in *E. coli* thereby providing a useful tool for studies of the electron transfer and monooxygenation reaction utilizing site-directed mutagenesis. [18]

Active site mutants of CYP101 have been shown to oxidize a wide range of compounds. CYP101 activity requires two proteins, a flavoprotein putidaredoxin reductase (PdxR) that accepts two electrons from NADH and one at a timetransfers these to cytochrome $P450_{cam}$ through the mediation of an iron-sulfur protein putidaredoxin (Pdx) whose role is to transfer the electrons from NADH cofactor to CYP101 enzyme. [19]

The Figure 8 illustrates the flow of electrons from NADH to $P450_{cam}$. In the first electron transfer reduced Pdx combines with a ferric-camphor bound $P450_{cam}$ and reduces it to the ferrous form. In the second electron transfer the reduced

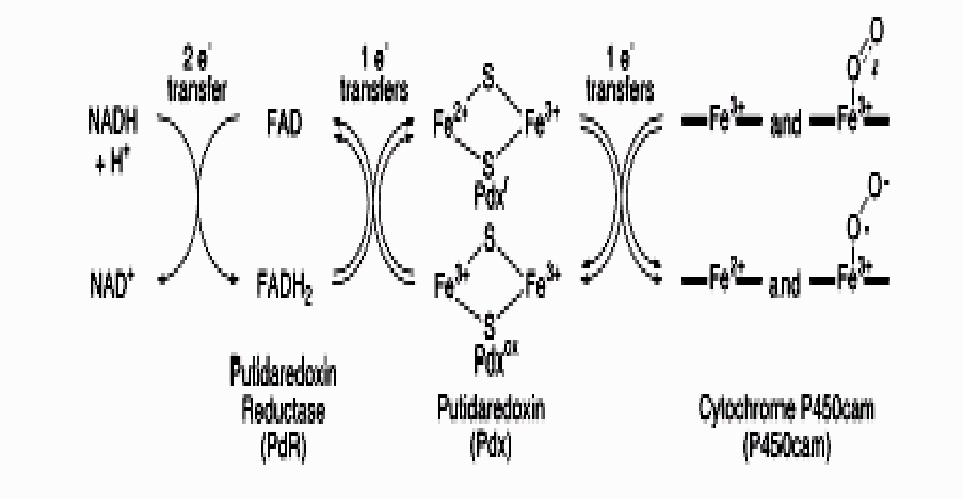


Figure 8: Electron Transport from PdR to $P450_{cam}$ Mediated by Pdx. [20]

Pdx forms a binary complex with oxyferrous, D-camphor bound $P450_{cam}$ and leads to the formation of the products 5-*exo*-hydroxycamphor, water, oxidized Pdx and D-camphor free ferric P450. Pdx plays the role of an 'effector' in the second electron transfer in addition to being an electron carrier. [19]

There are three possible mechanisms for the flow of electrons from PdxR to $P450_{cam}$; these include the cluster, Pdx shuttle and the transient cluster mechanisms. In the cluster mechanism the PdxR/Pdx/$P450_{cam}$ complex forms and survives for one or more turnovers. This indicates that both the electrons from one molecule of NADH move through the flavin of PdR to sequentially reduce the same Pdx and $P450_{cam}$ molecules. [20]

The shuttle mechanism doesn't require the formation of a complex of all the three proteins during the catalytic cycle; it involves the reduction of the oxidized Pdx^{ox} by PdxR and release of the reduced form of Pdx^{r} into the solution and reduction of $P450_{cam}$ by the binding of Pdx^{r}. The transient cluster mechanism is a hybrid of the prior two mechanisms. In this one molecule of Pdx remains bound to the $P450_{cam}$ throughout one or more catalytic cycles. The cycling of Pdx^{ox} to Pdx^{r} occurs via a transient association of fully reduced or one electron reduced PdR to Pdx^{ox} /$P450_{cam}$ complex. Two possible mechanisms for transfer of electrons from PdR to $P450_{cam}$ are shown in Figure 9. [20]

There is evidence in support of the shuttle mechanism which includes the ability of PdxR to reduce Pdx in the absence of $P450_{cam}$ and also the fact that Pdx^{r} is able to carry out both the electron transfer steps to $P450_{cam}$ in the absence of PdxR. In addition mutagenesis studies have revealed that the binding sites on Pdx for the PdxR and $P450_{cam}$ overlap, indicating that PdxR and $P450_{cam}$ cannot bind simultaneously to Pdx. Though the fusion proteins of PdxR, Pdx and $P450_{cam}$ have an efficient turn over, the shuttle mechanism is probably what fits best in the overall picture. [20]

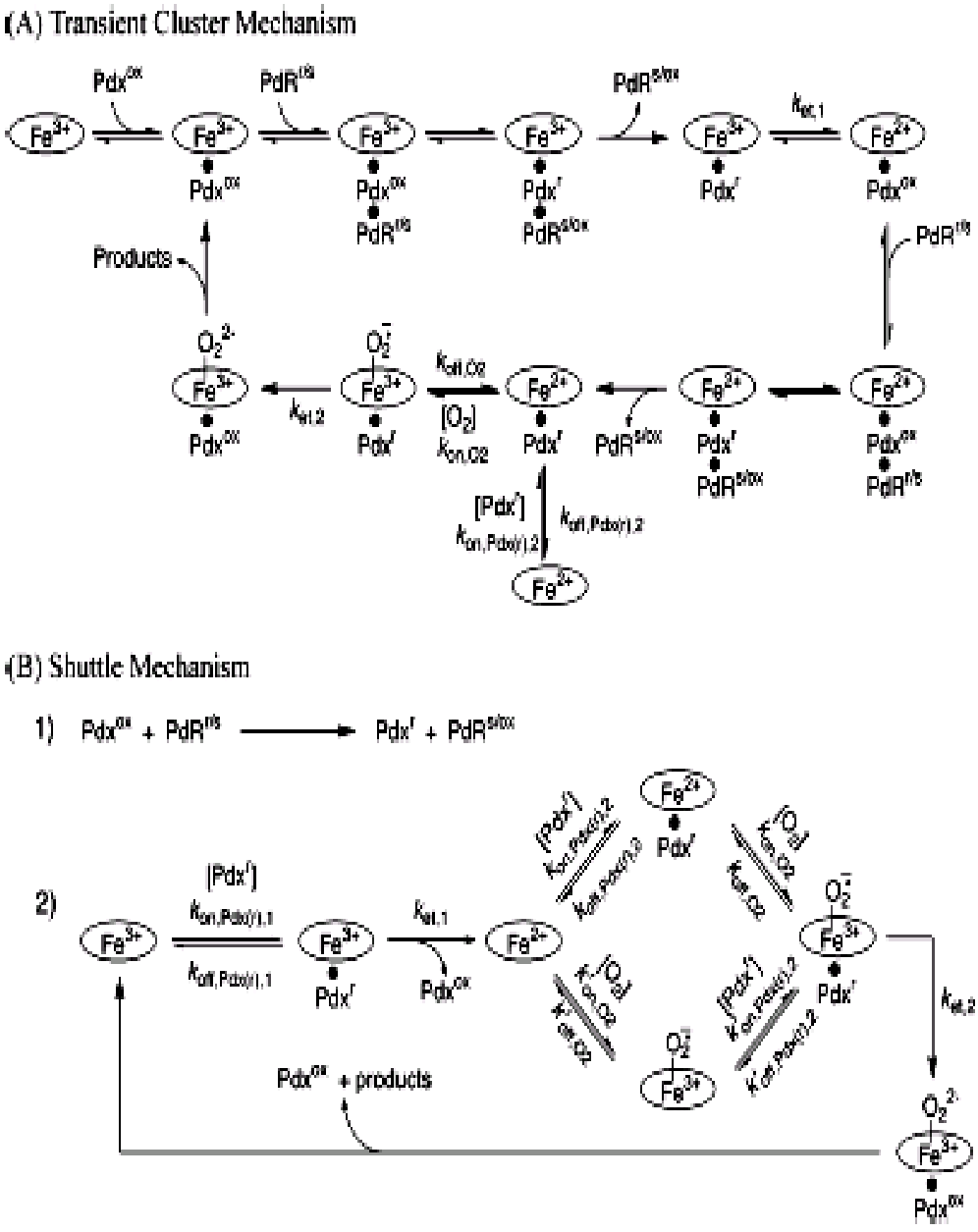


Figure 9: Possible mechanisms for transfer of electrons from PdR to $P450_{cam}$ The oval represents the porphyrin ring of camphor-bound $P450_{cam}$ [20]

Substrate specificity of $P450_{cam}$

Cytochrome P450s catalyze a range of oxidative transformations such as carbon hydroxylation, heteroatom oxidation, Π (pi) bond oxidation, and hydrocarbon desaturation. Despite differences in phylogenetic origin, cellular localization, electron donor partners and substrate specificity of P450 enzymes their active sites and catalytic mechanisms are remarkably similar. In the catalytic cycle of the P450 enzymes, the primary function of the enzymes is to promote the formation of ferryl species and to bind and orient the substrate in its proximity. The chemo and the regio-specificity of substrate oxidation are determined by the intrinsic reactivities of the substrate sites that are accessible to the ferryl species in the enzyme substrate complex. [21]

The oxidation of the organic substrate is caused by the coupled turnover, in addition to this the P450 enzymes undergo uncoupled turnover to produce oxygen, hydrogen peroxide and water. The uncoupling of the NAD(P)H, (chemically reduced form of Nicotinamide Adenine Dinucleotide Phosphate) consumption from the organic substrate oxidation is due to dissociation of super oxide (O_2^{-*}) from the ferrous dioxy complex, and/or reduction of the final ferryl species to water by electrons provided by NAD(P)H. Uncoupled turnover is low in resting enzyme but it greatly increases with the binding of some ligands and substrates; it is a substrate dependent catalytic process that produces reduced

oxygen rather than organic metabolites. Figure 10 shows possible places in the *Pseudomonas putida* catalytic cycle where uncoupling can occur [21]

The P450 enzymes that have hydrophobic active sites have very broad substrate specificity with preference for lipophilic substrates; these facts suggest that the binding of the P450 substrates is governed by hydrophobic and steric factors. Catalytic turnover is defined as the NADH-and substrate dependent formation of either organic or reduced oxygen metabolites. Binding of the substrates to the cytochrome $P450_{cam}$ active site is a prerequisite for catalytic turnover. Saturation of the active site of $P450_{cam}$ with camphor results in the complete conversion of the enzyme from low spin state to high spin state.

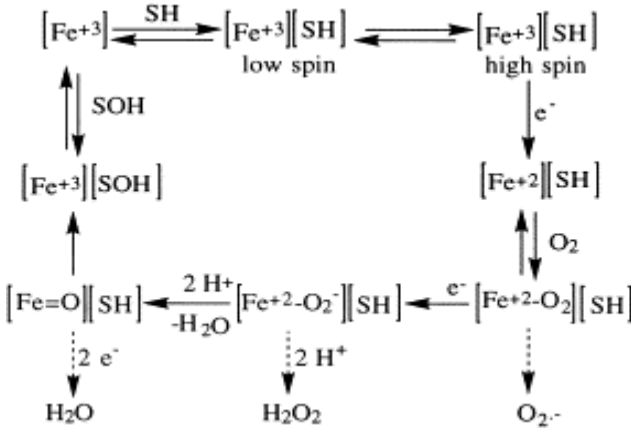


Figure 10: Catalytic cycle of cytochrome P450, showing the individual steps at which uncoupling can occur. [21]

Also the binding affinity and the extent to which the enzyme is converted to the high spin state by a saturating concentration of the substrate are correlated. NADH consumption provides a direct measure of the extent of catalytic turnover engendered by the binding of a substrate to $P450_{cam.}$ [21]

The diversity of the substrates for the P450 enzymes is not only due to the multiplicity of the isoforms but because of the fact that they are able to catalyze a broad range of P450 dependent transformations with a broad range of substrates. It is important to understand the substrate specificities of P450 enzymes for the design and construct of P450 enzymes with tailored specificities .The availability of the high resolution crystal structures of the four microbial P450 enzymes makes the bacterial enzymes amenable to analysis of the features that control substrate and reaction specificity. [22]

Two methods have been used to predict the substrate specificities of the well characterized $P450_{cam}$; one involves the use of molecular dynamic simulations to predict the regio-and stereo specificity of the products and the other includes the three-dimensional active site constrained screening procedures which are used to identify the compounds that fit within the active site and therefore are possible substrates. Both these processes necessitate the dissection of the catalytic process into discrete steps in order to correlate the effects of structure on individual branch points in the process. [22]

Heavily chlorinated compounds such as chlorinated benzenes such as hexachlorobenzene (HCB), pentachlorobenzene (PeCB), biphenyls and dioxins are some of the most problematic environmental pollutants. Degradation of the most heavily chlorinated benzenes requires initial reductive dehalogenation followed by oxidation by the dioxygenase pathway. Pathways that involve oxidation of chlorohydroquinones by action of flavin dependent monoxygenases degrade polychlorinated phenols. Then these reactive chlorohydroquinones are readily dehalogenated by hydrolytic or reductive mechanisms. Ring cleavage in polychlorinated benzenes occurs only after most of the chlorines are removed resulting in a more facile pathway. Cytochrome $P450_{cam}$ has been genetically engineered to act on compounds such as 1,2-DCB (dichlorobenzene), 1, 3-DCB, 1,4-DCB, 1,3,5-TCB (trichlorobenzene), 1,2,4,5-TeCB (tetrachlorobenzene), PeCB (pentachlorobenzene) and HCB (hexachlorobenzene). [23]

The wild type monooxygenase cytochrome $P450_{cam}$ has a low activity for oxidation dichlorobenzene and trichlorobenzene to chlorophenols and no activity for heavily chlorinated benzenes. Increasing the active site hydrophobicity with Y96F (Tyr to Phe) increased the activity of CYP101 causing the conversion of the PeCB and HCB slowly. Also decreasing the space available at the top of the active site with the F87W (Phe to Trp) mutation forces the substrate to be bound closer to the heme and results in further increase in activity. [23]

Cytochrome $P450_{cam}$ has been also engineered to oxidize gaseous alkanes like butane and propane to butan-2-ol and propan-2-ol by reducing the substrate-binding pocket. Using bulky amino acid substitutions reduced the size of the substrate pocket and the enzyme-substrate binding was improved. A picture of the active site of $P450_{cam}$ is shown in Figure 11. [24]

Engineering the active site topology for accommodation of different alkanes can enhance activity, selectivity and coupling efficiency of alkane oxidation by CYP101. Reduction of the active site volume using the mutations Y96F (Tyr to Phe)–V247L (Val to Leu) showed a four-fold greater activity for hexane compared to 3-methylpentane. When the active site was larger as a result of the Y96A (Phe to Ala) –V247A (Valine to Ala) double mutation 3-methylpentane was oxidized twice as fast as hexane. [25]

Styrene was weakly bound to the wild-type $P450_{cam}$; the binding was strengthened in the Y96F (Tyr to Phe) mutant, and again in Y96A (Tyr to Ala) mutant. This strengthened binding in mutants is a result of the removal of the hydroxyl group of tyrosine to generate a more active hydrophobic site that can interact better with styrene. [26]

Selective catalytic oxidation of simple alkanes under mild conditions is one of the more difficult chemical reactions to achieve using conventional synthetic methods. Cytochrome $P450_{cam}$ was engineered to oxidize various simple alkanes

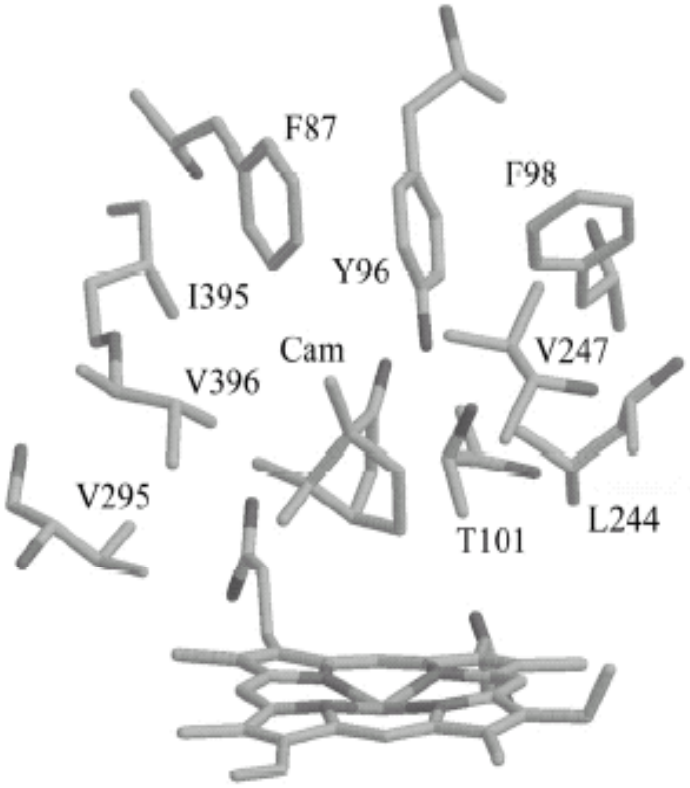


Figure 11: Active site of the P450cam enzyme with camphor bound in it [24]

like hexane, heptane, 2-methylpentane, 3-methyl pentane and 2-methylhexane. Since the substrate-binding pocket of cytochrome $P450_{cam}$ is mostly lined with hydrophobic residues with the exception of tyrosine 96, this was changed to alanine (Y96A) and phenylalanine (Y96F) to suit the hydrophobic nature of alkanes. The mutants showed an increased activity towards the alkanes proving the fact that increasing the hydrophobicity promotes alkane oxidation. Linear alkanes showed higher product formation rates with the Y96F and Y96A mutants than the branched ones. [27]

Polyaromatic hydrocarbons (PAHs) are hazardous and recalcitrant environmental contaminants mainly derived from anthropogenic pyrolysis of organic matter such as fossil fuel consumption and coal refining processes. These are procarcinogens and they are converted into highly carcinogenic oxidation products by heme-dependent cytochrome P450 enzymes such as CP1A1.PAH oxidation activity of P450 enzymes is of wide interest because P450s are utilized in the initial oxidation step of biodegradation of PAHs. [28]

The cytochrome $P450_{cam}$ has been used to oxidize a variety of unnatural substrates. Its mutant Y96A not only increases the hydrophobicity of the substrate pocket but also creates space for the binding of substrates larger than camphor. These two are important factors that enhance the ability of $P450_{cam}$ to act on PAHs. Replacing Y96 with hydrophobic residues like Gly, Ala, Val, Phe and further testing the mutants on naphthalene and pyrene have revealed that the

mutants oxidize the substrates faster than the wild type. Naphthalene is oxidized 1- and 2-naphthol via 1,2-oxide intermediate and pyrene gave a 1,6- and 1,8-pyrenequinone. [29]

The wild type cytochrome $P450_{cam}$ enzyme has very low activity with polycyclic aromatic hydrocarbons like phenanthrene, fluoranthene, pyrene and benzo [a] – pyrene where as the mutants (with mutations at F87 and Y96) show increased activity on the same substrates. Hydrophobic substitutions at Y96 broadens the $P450_{cam}$ substrate range through two effects 1) increasing the active site hydrophobicity promotes binding of organic molecules and it also reduces uncoupling by not favoring the entry of water molecules into the active site 2) Y96 also forms a part of the substrate access channel, therefore hydrophobic substitutions at this residue may promote binding of hydrophobic molecules like pyrene. [29]

Cloning of the Cytochrome $P450_{cam}$ gene into *Escherichia coli*

Cloning of the cytochrome $P450_{cam}$ gene involved the localizing of the *camC* gene within the PstI and HindIII fragment as shown in Figure 12, this gene is responsible for the production of the $P450_{cam}$ enzyme. The primary amino acid sequence of P-450_{cam} has a 412 amino acid sequence that corresponds to a gene length of 1236bp. Top panel of the restriction map of the 2.2-kilobase pair

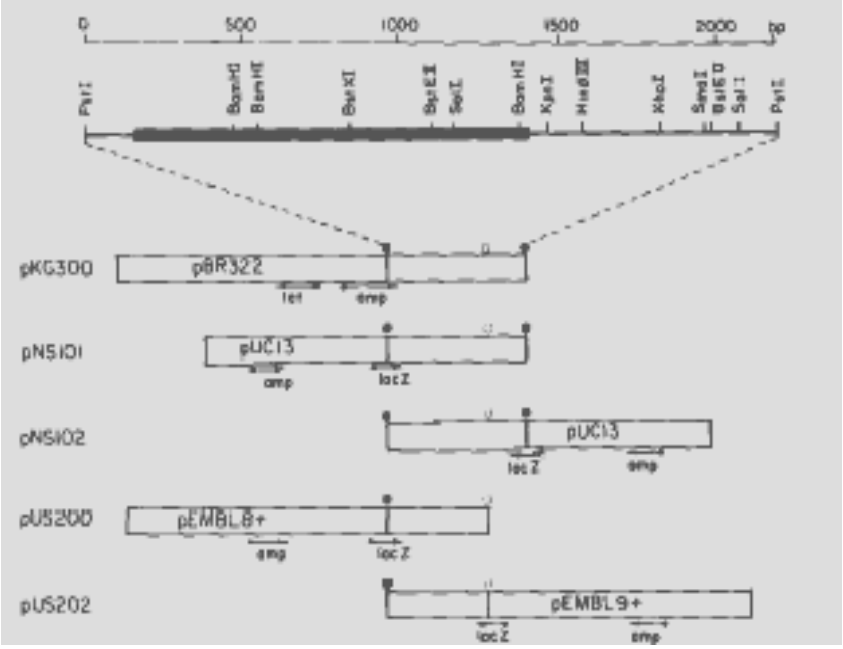


Figure 12: Restriction maps of pKG3OO and its derivatives. [3]

CAM plasmid fragment containing the cytochrome P-450_{cam} gene shows a heavy line representing the cytochrome P-450_{cam} gene in Figure 12. The bottom panel indicates the various constructions used. Regions between the circles represent CAM plasmid DNA, where *0* and 0 are symbols for PstI and HindIII sites, respectively. Gene symbols are as follows: tet-tetracycline resistance gene; amp-ampicillin resistance gene, 0-lactamase gene; *lacZ,* β-galactosidase gene. Arrows indicate the direction of transcription. [3]

CAM plasmid was sub cloned into pUC13, pEMBL8+ and pEMBL9+ so as to increase the yield of the plasmid DNA because these vectors had a higher copy number than their parental vector pBR322 from which pKG300 was derived. In

these various sub clones the *camC* gene was transcribed from the *lac* promoter, thereby increasing the yield of the cloned protein than that produced from the β lactamase gene promoter in pBR322. [3]

The entire cytochrome $P450_{cam}$ gene from the PstI to the HindIII site that was 1578bp was sequenced; this also included a flanking sequence and some DNA on the 3′ end. The Figure 13 shows a heavy line representing the cytochrome P-450_{cam} gene. Arrows in Figure 13 indicate the direction and extent of sequencing from each restriction site. Only restriction sites used for sequencing are shown, *PstI (P),* BglI (Bg), *RsaI (R),* BamHI (B), EcoRV *(E),* BstXI *(B X), BstEII* (BE), Sal1 (SI, and HindIII *(H).* [3]

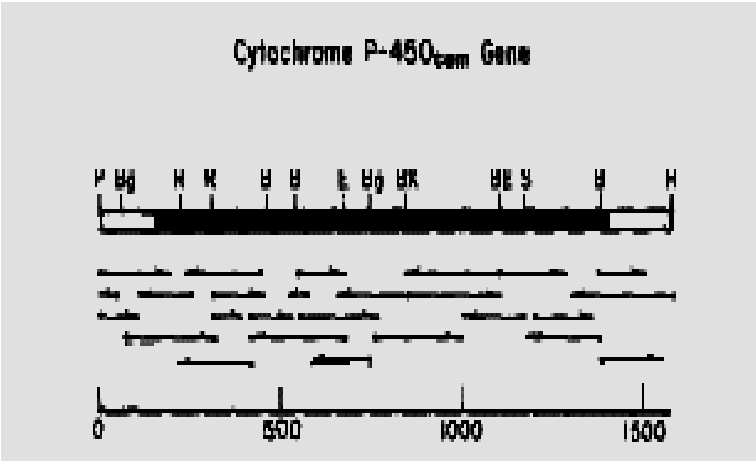


Figure 13: Sequencing strategy of the cytochrome P-450 gene.[3]

Figure 14 shows nucleotide numbering begins with the first nucleotide of the *PstI* site which is 155 nucleotides 5' to the cytochrome P-450_{cam} gene. Amino acid numbering begins with the first Thr (Threonine) and the initiating Met (Methionine) of the cytochrome P-450_{cam} and putidaredoxin reductase genes, respectively. The underlined nucleotides refer to putative Shine-Dalgarno sequences. The boxed amino acids and nucleotides indicate differences between the protein sequence and the nucleotide-derived amino acid sequence

Thus the cytochrome P-450_{cam} gene was cloned into *Escherichia coli*, its amino acid and nucleotide sequence were determined and it was demonstrated through various analysis that the cytochrome P-450_{cam} enzyme from *Escherichia coli* and *Pseudomonas putida* is indistinguishable with respect to its optical spectrum, molecular weight, catalytic activity and N-terminal sequence. Placement of the *cam*C gene behind the *lac* promoter in *Escherichia coli* vectors enabled the *cam*C gene behind the *lac* promoter in *Escherichia coli* vectors enabled the production of the enzyme in large quantities. When the gene was oriented opposite of the direction of the vector promoter in *Escherichia coli*, P-450_{cam} protein was undetectable by the reduced CO (Carbon monoxide)-difference spectral analysis. [3]

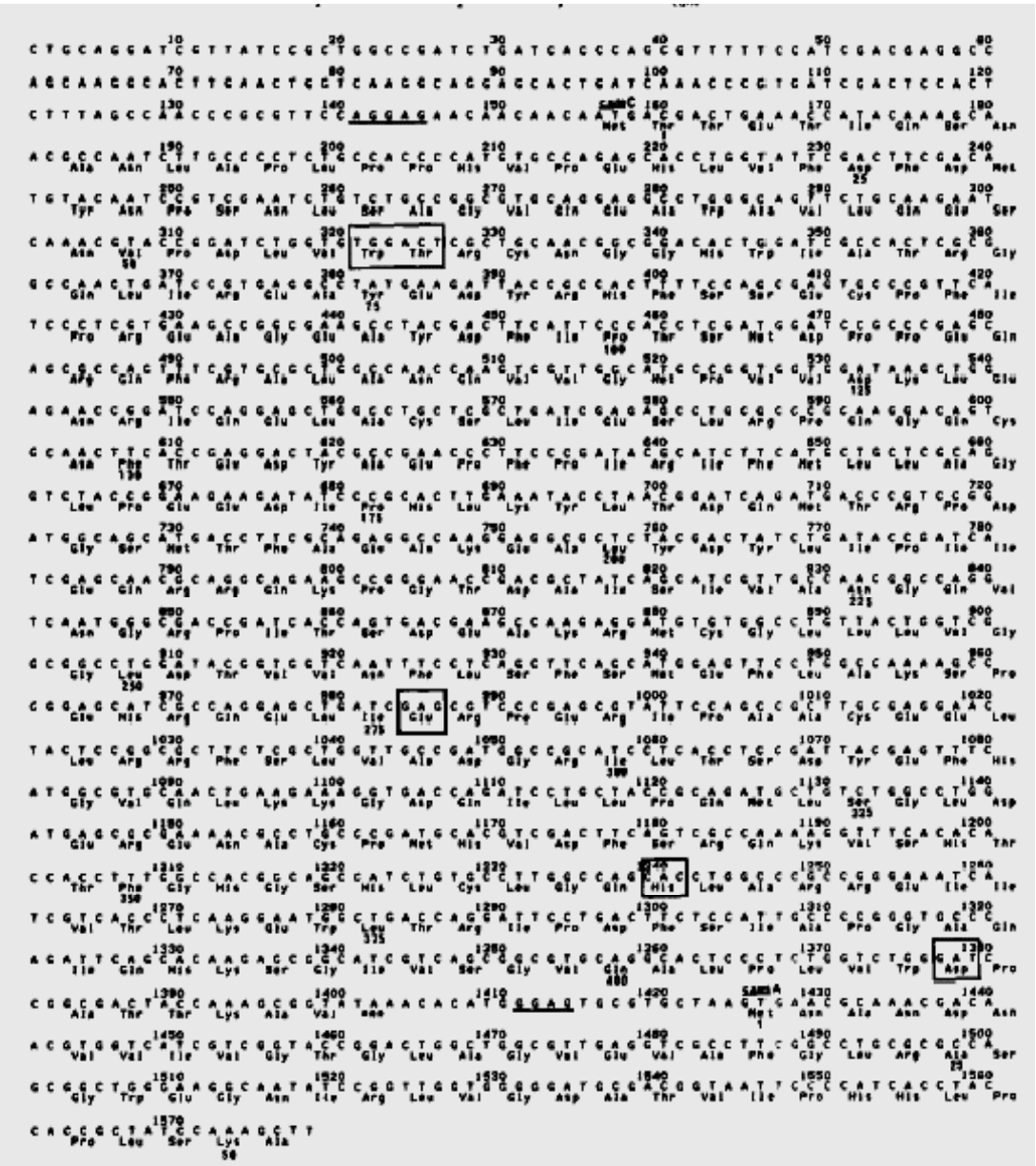


Figure 14: The complete nucleotide sequence of the cytochrome P-450 gene and a 5' portion of the putidaredoxin reductase gene. [3]

2.3 Shuttle vectors

Introduction

Shuttle vectors consist of double replicons, in which one half of the plasmid allows replication in an enteric bacterial host and the other half allows replication in any other host from which the other replicative region has been derived. These vectors are used in transferring recombinant plasmids from *Escherichia coli* into non-purple Gram (+) bacteria and most eukaryotic hosts. In these shuttle vectors, one half of the vector is derived from *Escherichia coli* specific vectors, the second half comes from often small cryptic (but ideally multicopy) replicons, which were isolated from the individual host strain; they provide the oriV specific to the host from which they were isolated. The two halves are then joined by DNA manipulation, which is the reason why this approach is also referred to as double replicon strategy. [30]

Rhodococcus-Escherichia coli shuttle vectors

The lack of plasmid vectors suitable for use in *Rhodococcus* has led several groups to develop binary vectors, in which extant *Rhodococcus* plasmid is combined with an *Escherichia coli* plasmid forming a vector that replicates in both cell types. Some of these *Rhodococcus-Escherichia coli* shuttle vectors are given in the Table 2.

Table 2: *Rhodococcus-Escherichia coli* shuttle vectors

Rhodococcus-Escherichia coli Shuttle vector	Features	Reference
pRHK1	Has a cryptic plasmid pRC4, derived from *Rhodococcus rhodochrous* and a vector from *Escherichia coli* with a kanamycin resistance gene, pHSG298	[31]
pNC9503	*Escherichia coli-Rhodococcus erythropolis* shuttle vector with a unique restriction site for XbaI. It has a kanamycin resistance gene and a thiostrepton resistance gene .The origin of replication is from *Rhodococcus rhodochrous* plasmid pNC903	[32]
pNC9501	In addition to all the features in the pNC9503 it has two other unique restriction sites for KpnI and EcoRI	[32]
pRF29	Cointegrate of pRF28 (StuI linearized) and pUC13 (SmaI linearized). Has ampicillin resistance gene and chloramphenicol resistance gene	[33]
pMVS301 and pMVS302	*Escherichia coli–Rhodococcus* shuttle plasmids have ampicillin resistance gene from pBR322 and thiostrepton resistance gene	[34]
pKSD6-1	*Rhodococcus-Escherichia coli* shuttle vector containing ampicillin and chloramphenicol resistance genes.	[18]

Construction of the *Rhodococcus-Escherichia coli* shuttle vector pKSD6-1

The *Rhodococcus-Escherichia coli* shuttle vector pKSD6-1 is capable of replication in both *Rhodococcus* and *Escherichia coli* and it was constructed as a deletion derivative of the plasmid containing the *sox* genes The *Rhodococcus* sp. strain X309-11-15 was used as a source of a 100-kb pSOX plasmid, and *Rhodococcus* sp. strain X309-10-2 (referred as 10-2) was the plasmid free derivative of the *Rhodococcus* sp strain X309. [18]

- pSOXΔ1(52 kbp in size) was produced as the first deletion derivative of the pSOX plasmid by selecting from the electrotransformed 10-2 strain using DBT(Dibenzothiophene).Restriction of pSOXΔ1 with EcoR1 produced five fragments, which were selected for the *sox* containing genes using the Southern hybridization test. This pSOXΔ1 was introduced into 10-2 and further selection on DBT plates resulted in pSOXΔ2. [18]

- pSOXΔ2 was a 42kb fragment, a series of self defined manipulations like the self ligation of a 30kb *sox* containing EcoRI fragment from pSOXΔ2 gave pSOXΔ3 which was capable of replicating in strain 10-2. A chloramphenicol resistance (Cm^r) gene was added to the rightmost 15-kb XbaI-EcoRI fragment at the unique EcoRI site of pSOXΔ3.To facilitate

cloning, both the linearized vector and the XbaI-XhoI DNA fragment containing the Cm^r gene were blunt ended with dNTPs and Klenow DNA polymerase I. [18]

- Digestion with EcoRV and Xba I gave an 18.5-kb fragment which when transformed into 10-2 and selection on chloramphenicol plates gave several colonies that were analyzed. The analysis revealed the existence of two populations of cells one containing the expected 18.5 kb plasmid (pSOXΔ4) and the other having a 9kb plasmid (pSOXΔ5). [18]

- pSOXΔ6 was obtained by ligating the 4-kb KpnI fragment to the 3.5 kb Cm^r gene cassette, which was a blunt end ligation. Screening for the 7.5 kb pSOXΔ6 revealed the deletion of 400bp in one derivative, which was named as pSOXΔ6-1. The entire process of constructing the deletion derivatives is given in the Figure 15 below. In the Figure 15 the ability of each plasmid to replicate (*ori*) or possess desulphurization activity (*sox*) is shown at the right by the plus and minus signs. Cm^r indicates the chloramphenicol resistance marker gene derived from pRF29 plasmid and --- in Figure 15 is the undefined boundary of deletion. [18]

- pSOXΔ6-1 was linearized by SphI and ligated to the Nae I site (within the f1 origin) of the pBluescript II KS –plasmid. The pBluescript II KS – plasmid provides the colE1 replication origin, a multiple cloning site and

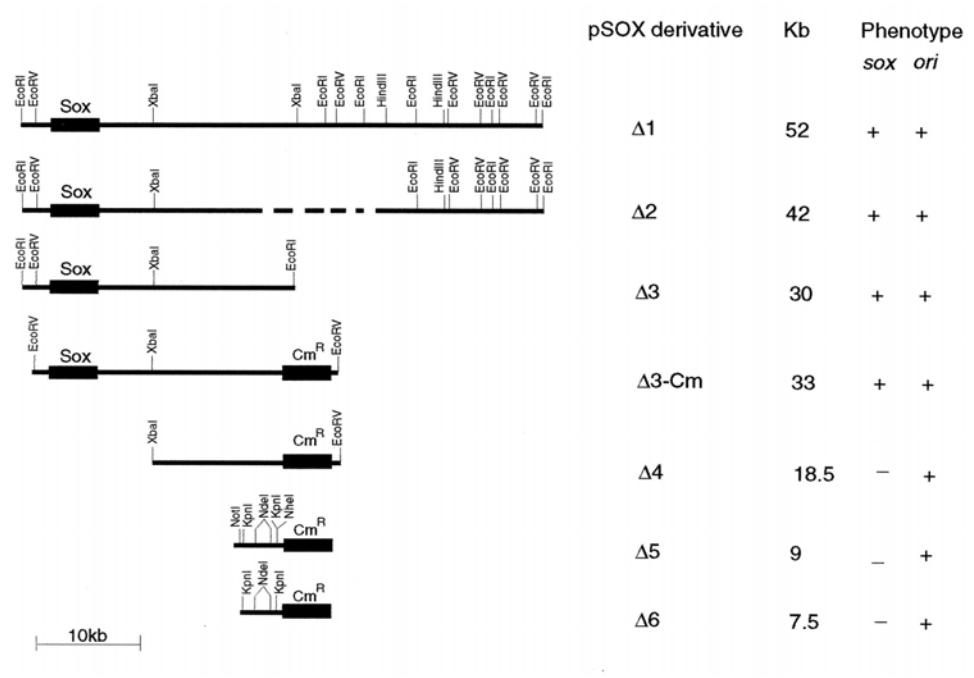


Figure 15: Construction of the deletion derivatives of the pSOX plasmid. [18]

the *lacZα* reporter gene for screening inserts in *Escherichia coli* (blue-white selection) and Ampicillin resistance (Ap^r). The resultant plasmid was named pKSD6-1. Figure 16 shows the shuttle vector that has been used for *sacB* expression in *Rhodococcus* strain 10-2, the MCS, multiple-cloning site (expanded region).It also shows the blunt-end cloning of the end-filled 3.8-kb *Bam*HI fragment of pUM28 plasmid containing *sacB* at the *Eco*RV site. [18]

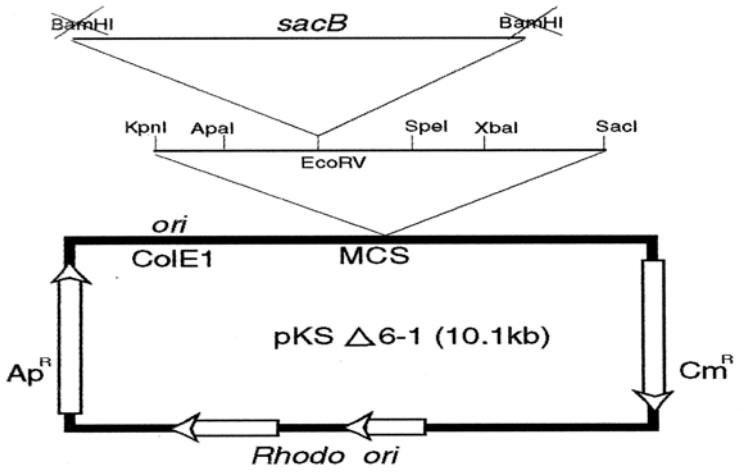


Figure 16: Construction of the pKSD 6-1 *Rhodococcus-Escherichia coli* shuttle vector.

[18]

2.4 pCWori plasmid

pCWori+ is a derivative of the plasmid pHSe5, it contains a pBR322 origin of plasmid replication, the *lac* I gene, the β lactamase gene for conferring ampicillin resistance and a bacteriophage origin of replication. The transcription and translation regions contain a *lacUV*5 promoter and two copies of a *tac* promoter cassette followed by a translation initiation region derived from the phage T4 lysozyme gene. This DNA sequence contains a 7 bp Shine-Dalgarno element located three nucleotides 5′ of the initiation codon contained within an *NdeI* restriction site. A *trp*A transcription termination cassette is located downstream of the unique 3′ restriction sites.[39] pCW is a plasmid that over expresses the

che W protein, it contains the *che W* gene from the pJL63 plasmid that has been introduced into the pHSe5. In pCW the expression is under the control of the *tac* promoter. [35]

C h a p t e r 3

METHODS AND RESULTS

3.1 Introduction

The main aim of the project was to clone the CYP101 gene into a *Rhodococcus-Escherichia coli* shuttle vector and express the protein in a *Rhodococcus* bacterium. On examining the MCS (multiple cloning site) of pKSD6-1 and sequence of the cytochrome $P450_{cam}$ (*CYP101*) gene, it was found that cloning of the cytochrome $P450_{cam}$ gene directly into the chosen vector pKSD6-1 was not possible because of the unavailability of compatible restriction sites on the insert and the vector. This necessitated the usage of the commercially available pCR2.1-TOPO vector as an intermediary. The CYP101 was first cloned into pCR2.1-TOPO, and then the CYP101 gene was restricted with a suitable restriction enzyme and cloned into pKSD6-1.

3.2 Production of the gene from pCWori plasmid containing CYP101 using PCR

The pCWori+CYP101 plasmid (Figure 17) was obtained from Prof. Paul R. Oritz de Montellano (University of California, San Fransisco, CA). The following factors were taken into consideration while designing the primers. The MCS of

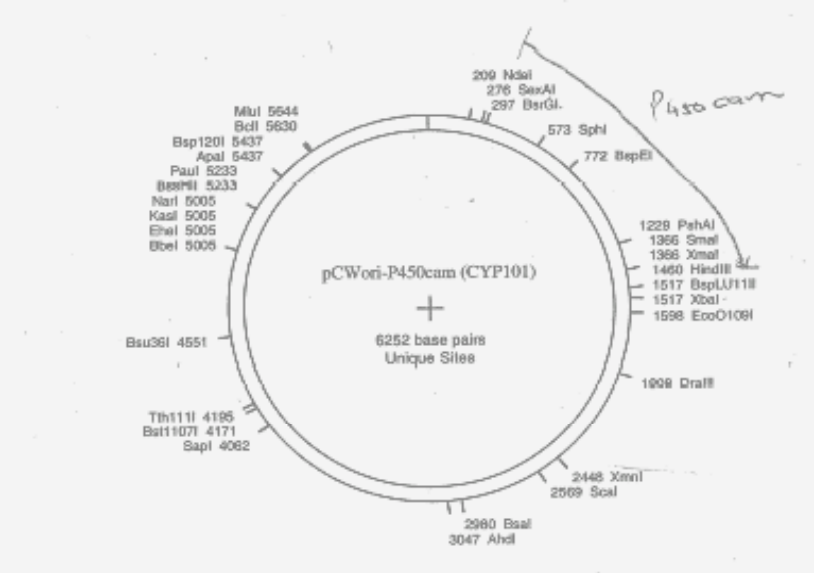


Figure 17: pCWori+P450$_{cam}$

the vector pKSD6-1 was observed and all the available sites on that were analyzed for suitability. The suitable restriction sites on the MCS met the following conditions

- They were not present anywhere on the insert, i.e. *CYP101* sequence
- They should have sufficient activity in the same NEB buffer that will be used while setting up the restriction reaction.

Using the *CYP101* sequence from the gene bank the forward (P450 1F) and the reverse primers (P450 1R) were designed with ApaI and Sac I restriction sites included in the P4501F and P4501R respectively as they met all the

aforementioned conditions. The feasibility of the designed sequences was then tested using the primer design link in the Invitrogen website (see [36])

The sequences of the designed primers are given below:

P4501F (5′-3′): TAG GGG CCC ACA ATG ACG ACT GAA ACC ATA CA

P4501R (5′-3′): CAT GAG CTC AAG CTT TGG ATA GCG GTGG

The primers were obtained from Invitrogen and their (10pmol/μL) solutions were prepared. The PCR for the production of the $P450_{cam}$ gene was optimized by setting up seven 30μL PCRs using buffers A, B, C, D, F, J, and N from the Invitrogen PCR optimization kit. The protocol for this set of PCRs is given below.

Protocol for the PCR to produce CYP101

- 5X Buffer: 6μL
- Primers $\begin{Bmatrix} 10\,pmol P4501F(Forward primer) : 1.5\,\mu L \\ 10\,pmol P4501R(\mathrm{Re}verse primer) : 1.5\,\mu L \end{Bmatrix}$
- DNA (CYP101 1:50 dilution of 383.25μg/mL): 3μL
- 25 mM dNTP: 0.6μL
- Sterile water: 17.1 μL
- Taq Polymerase: 0. 3μL
- Total volume: 30μL mix contents by gently tapping

- Add Nujol oil: 10μL and mix the contents again by tapping.

The contents in the PCR vial were run in a thermocycler (Perkin Elmer Cetus DNA THERMAL CYCLER, MA), using program files 90, 91 and 18, which were programmed as follows:

- File 90: 2.5 min at 94°C
- File 91: $\left\{\begin{matrix} 1\,\text{min}\,at\,94^{\circ}C \\ 2\,\text{min}\,at\,55^{\circ}C \\ 4\,\text{min}\,at\,72^{\circ}C \end{matrix}\right\}$ 30 cycles of file 91
- File 18: delay time 10 min at 72ºC

After all the programs in the thermocycler were completed, then the PCR vials were removed and the reaction solution of 30μL was carefully pipetted using a manual pipette and placed in another vial. 15μL of each of the PCR reactions performed was run on a gel to test the presence of the desired *CYP101* band. Gel electrophoresis using agarose gel (0.7%) showed the fragment equal in size to the $P450_{cam}$ gene, 1.4 kilobasepair (kbp) in six lanes; of these C, B, A, D, F, J were the lanes on which the gene was seen better (in that order as seen in Figure 18) than on others. The gene was excised from the gel using a clean sterile blade and DNA was extracted from the gel using Wizard PCR preps DNA purification system (Promega, CA). Both the insert and vector DNA were restricted with ApaI and SacI to get the desired 'sticky' ends. The protocol for the restriction reaction is given in Table 3.

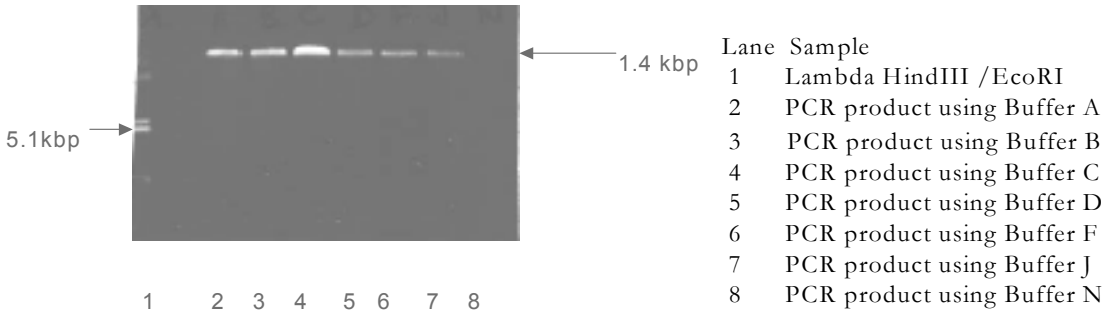


Figure 18:Gel showing the 1.4 kbp P450 gene from the PCR optimization buffers.

Table 3:Protocol for the restriction reactions of insert and vector

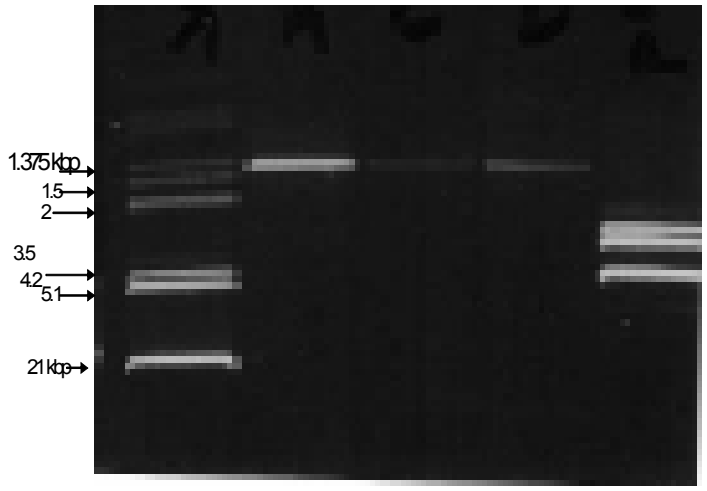
Insert	Vector
DNA (from C, D and A): 10µL	3.83µg/mL Vector DNA: 2µL
ApaI* (10000u/ml): 1µL	ApaI(10000u/ml): 1µL
SacI *(20000u/ml): 1µL	SacI(20000u/ml): 1µL
NEB4* (10X): 3µL	NEB4 (10X): 3µL
BSA* (100X): 0. 3µL	BSA (100X): 0. 3µL
Sterile water: 14.7µL(for PCR products)	Sterile water: 22.7µL(for vector DNA)
Total: 30µL	Total: 30µL

*All the restriction enzymes and buffers used for the reactions were obtained from the New England Biolabs, MA.

3.3 Attempt to directly clone cytochrome $P450_{cam}$ into pKSD6-1

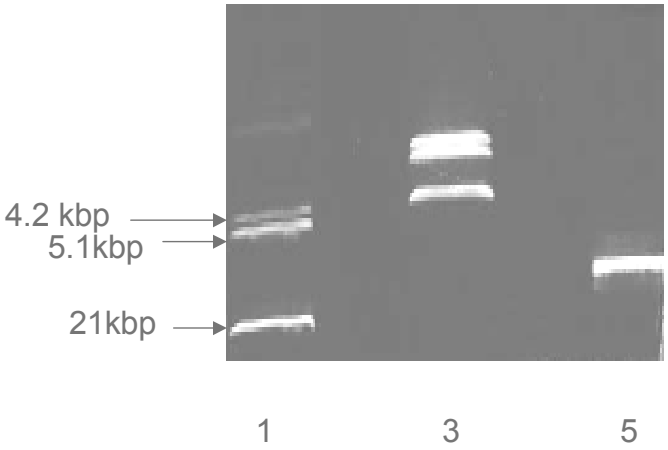
Since the complete restriction map of the vector pKSD6-1 was not known, the experiments were devised on the basis of the available partial restriction map of pKSD6-1. The DNA of the PCR products that was run and purified from the gel and the vector DNA (pKSD6-1) were restricted with ApaI and SacI enzymes

The contents were mixed and placed in an incubator at 37°C overnight; the next day they were subjected to gel electrophoresis. The insert gave the desired single fragment of 1.4kbp but the vector gave three bands instead of the expected one 10.1kbp band (seen in Figure 19). It was therefore concluded that ApaI + SacI together had three restriction sites on the pKSD6-1 plasmid. It was necessary to investigate the number of restriction sites for each of the enzymes ApaI and SacI so as to determine which restriction enzyme was cutting the vector more than once to omit its use further during experimentation; therefore pKSD6-1 was restricted with ApaI and SacI separately. ApaI was found have a single restriction site on pKSD6-1 at the MCS where as SacI had three restriction sites on the entire plasmid (shown in Figure 20)



Lane	Sample
1	Lambda HindIII/EcoRI
2	CYP101(C) w/ApaI+SacI
3	CYP101 (A) w/ApaI+SacI
4	CYP101 (D) w/ApaI+SacI
5	pKSD6-1 w/ApaI+SacI

Figure 19: Gel analysis of the insert and vector restricted with ApaI and SacI



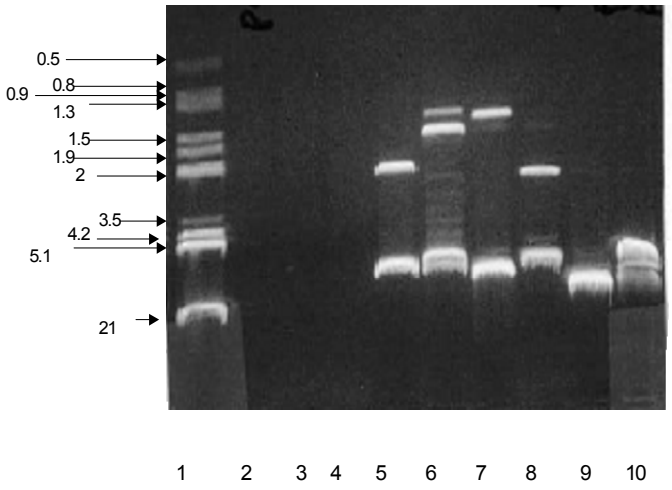
Lane	Sample
1	Lambda HindIII /EcoRI
2	-
3	pKSD6-1 w/ SacI
4	-
5	pKSD6-1 w/ ApaI

Figure 20: Gel showing the restriction of pKSD6-1 w/ SacI (3) and Apa I (5)

Also an attempt was made to construct a partial restriction map of the shuttle vector. To this end the vector was digested with different enzymes available in the laboratory. The gel analysis of the restriction reactions is given Figure 21.

3.4 Cloning of cytochrome $P450_{cam}$ into pCR 2.1 TOPO

After finding that *CYP101* couldn't be cloned into pKSD6-1 directly using current primers/restriction sites; the alternative was to design new primers with different restriction sites or to clone via a cloning vector such as commercially available pCR 2.1-TOPO. A 50 μL PCR was performed this time different restriction sites or to clone via a cloning vector such as commercially available pCR 2.1-TOPO.



Lane	Sample
1	Lambda HindIII/EcoRI
2	-
3	-
4	-
5	pKSD6-1 w/ NdeI
6	pKSD6-1 w/ EcoRI
7	pKSD6-1 w/ HindIII
8	pKSD6-1 w/ XhoI
9	pKSD6-1 w/ NotI
10	pKSD6-1 w/ NcoI

Figure 21: Gel analysis of pKSD6-1 restricted with various enzymes.

A 50 μL PCR was performed this time using only buffer C (that produced the clearest band on the gel) for the production of the *CYP101* gene; also another PCR was set up for the control DNA given in the cloning kit. The control PCR was set up as given in the TOPO TA cloning

Protocol for the 50μL PCR reactions

- Buffer: 10μL
- Primers $\begin{Bmatrix} 10\,pmol P4501F(Forward primer) : 2.5\mu L \\ 10\,pmol P4501R(\text{Re}verse primer) : 2.5\mu L \end{Bmatrix}$
- DNA (CYP101 1:100 dilution from midi prep): 5μL
- dNTP (2.5Mm): 1μL
- Sterile water: 28 μL
- Taq Polymerase: 1μL
- Total volume: 50μL mix contents by gently tapping

The products were tested for the presence of the respective genes using gel electrophoresis as shown in Figure 22; the P450 band is not very clearly visible in the Figure 22. After verification the bands were excised from the gel and purified using the Wizard prep PCR purification system.50 μL of each DNA was eluted after purification. The purified *CYP101* and the control genes thus obtained from PCR were cloned into the TOPO cloning vector, pCR 2.1-TOPO and

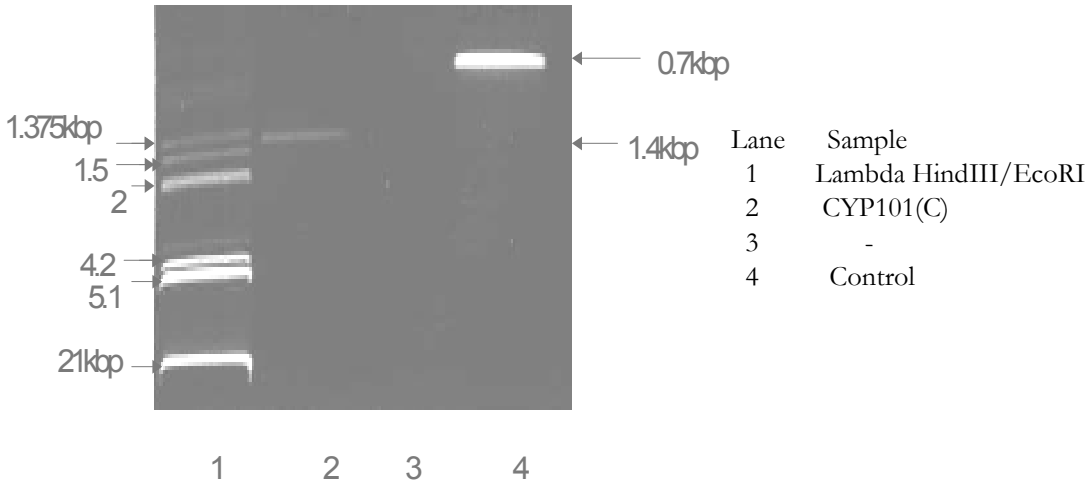


Figure 22: Gel showing the *CYP101* and the control genes

transformed into Top10F′ cells using the one shot chemical transformation protocol given in the TOPO TA cloning manual.

100 μL of each transformation solution was plated on one LB (Luria broth) agar plate (1/10 plate) containing 50 μg/mL ampicillin; the rest of each solution was plated on to another one .All the four plates were incubated over night at 37°C. The transformation was successful resulting in colonies on all plates.

Twenty colonies were picked from the CYP101 transformation containing agar plates and two colonies were picked from the control plates and grown in 5mL of LB containing 50 μg/mL ampicillin. The colonies were allowed to grow overnight in the incubator at 37°C and 150 rpm. The next day, 3mL of each culture was used and DNA from the culture was obtained using lysis by boiling

method (boiling prep). The size of the recombinant pCR2.1 TOPO+*CYP101* (shown in Figure 23) is 5.4 kbp; this was estimated using the sequences of the plasmid pCR 2.1 TOPO and the PCR product *CYP101*.

Then this recombinant plasmid was analyzed using NEBcutter V2.0 software for a suitable restriction site. The desired site was required to cut the recombinant plasmid on the $P450_{cam}$ gene only and linearize the plasmid (Topo+*CYP101*). NotI was the only site that was present once on the TOPO gene and absent on the *CYP101* gene (as shown in Figure 24), therefore it was chosen to restrict all the recombinant DNA isolated from the twenty colonies.

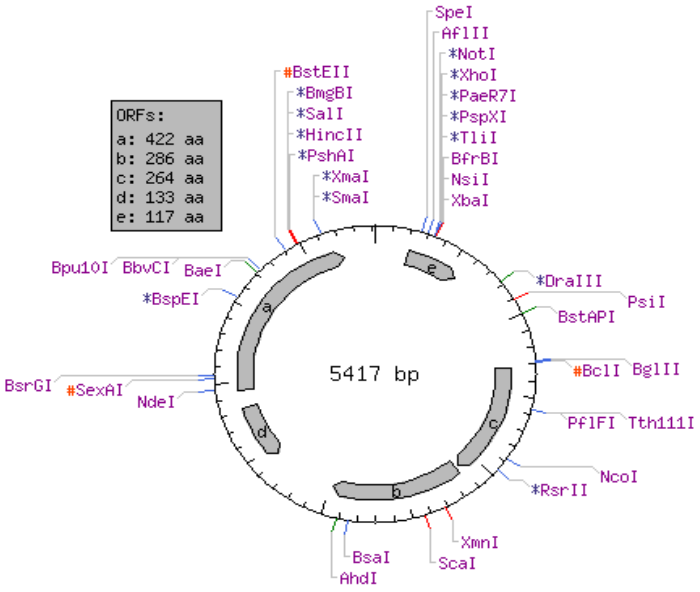


Figure 23: pCR2.1 TOPO+CYP101 plasmid

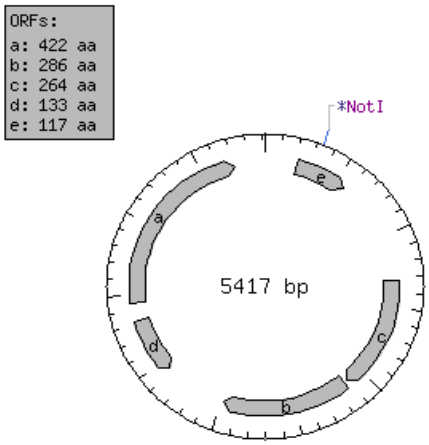
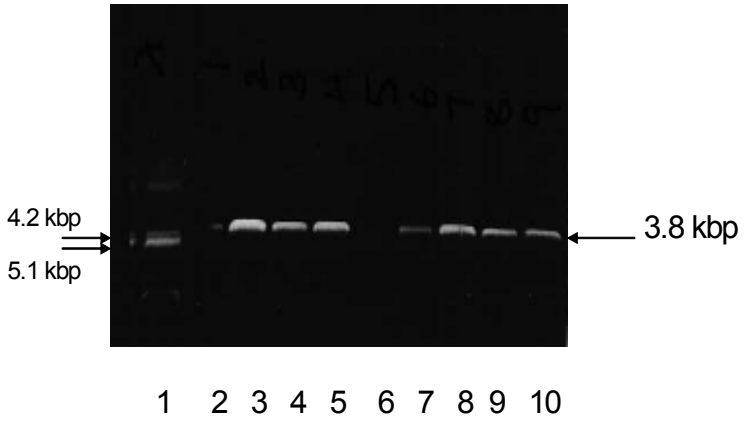


Figure 24: pCR 2.1 TOPO+CYP101 with NotI restriction

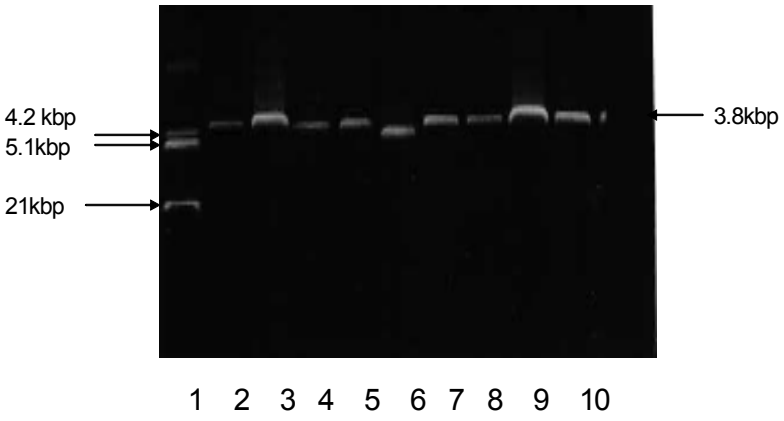
After a two-hour restriction reaction using the standard protocol all the samples were analyzed on a gel. The gel revealed an unexpected 3.8 kbp band (shown in Figure 25 (a) and (b)) instead of the expected 5.4 kbp. This meant that there was no insert in any of the colonies. This fact was further confirmed by restricting the recombinant plasmids with EcoRV, which has two restriction sites on the recombinant plasmid pCR2.1TOPO+CYP101.

Another 50μL PCR was set up for the production of the CYP101 gene using the protocol given previously. The CYP101 gene formed was verified using an agarose gel. To clone the P450 gene into the TOPO vector 4μL of the direct PCR reaction mixture was used; it was immediately transformed into Top10F′ cells. The transformants were plated on LB agar plates containing 50μg/mL ampicillin, X-gal (5-bromo-4-chloro-3-indolyl-bD-galactoside) and 100mM IPTG to facilitate blue/white selection. The plates were incubated overnight at 37°C.



Lane	Sample
1	Lambda HindIII/EcoRI
2	Colony 1 DNA
3	Colony 2 DNA
4	Colony 3 DNA
5	Colony 4 DNA
6	Colony 5 DNA
7	Colony 6 DNA
8	Colony 7 DNA
9	Colony 8 DNA
10	Colony 9 DNA

(a)



Lane	Sample
1	Lambda HindIII/EcoRI
2	Colony 10 DNA
3	Colony 11 DNA
4	Colony 12 DNA
5	Colony 13 DNA
6	Colony 14 DNA
7	Colony 15 DNA
8	Colony 16 DNA
9	Colony 17 DNA
10	Colony 18 DNA

(b)

Figure 25: Gels showing the 3.8kbp band instead of the expected 5.4kbp in the DNA isolated from twenty different colonies.

- After 24 hrs many colonies were found on each plate. Five colonies from each plate were picked and grown in 5mL LB cultures containing 50μg /mL of ampicillin. One mL of each of these cultures was stored at-80°C in 20% glycerol. The rest of each of the culture was subjected to lysis by boiling (adapted from Holmes and Quigley).

- Using restriction analysis with NotI enzyme, the presence of the insert was verified. Gel electrophoresis after restriction analysis revealed the presence of the expected 5.4 kbp fragment (seen in Figure 26) on the gel.

The expected fragment of 5.4 kbp was seen only in the lanes containing the DNA sample from colony 3, which was selected from the agar plates containing the

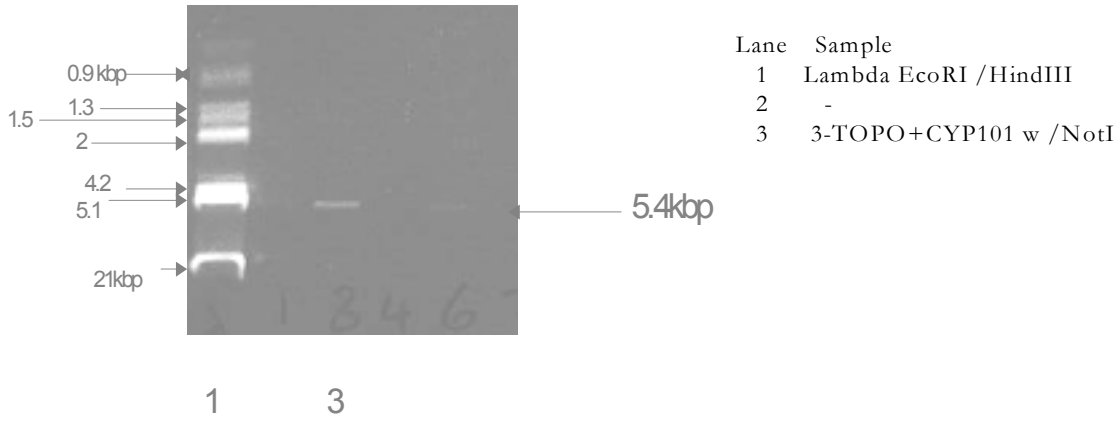


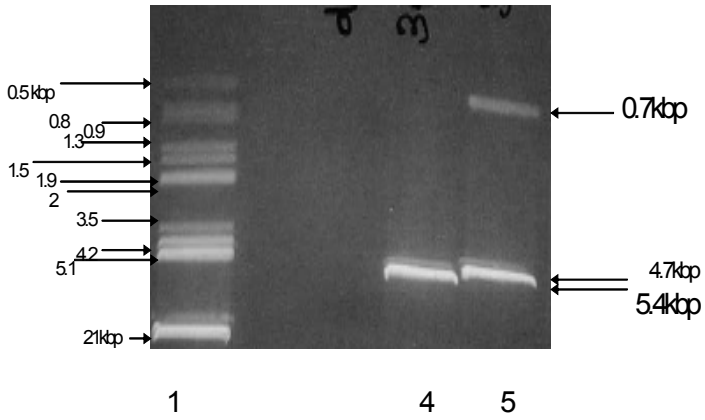
Figure 26: Gel showing the presence of the 5.4 kbp linear TOPO+*CYP101* plasmid

recombinant TOPO+*CYP101* transformants thereby proving the presence of the gene in the recombinant plasmid of this colony. This colony was grown again in a 5mL LB culture containing ampicillin and a frozen stock containing 20% glycerol was made from this culture using 20% glycerol and was stored at –80°C; this stock was named 3-TOPO+*CYP101*.

3.5 Cloning of cytochrome P450 from 3-TOPO+*CYP101* to pKSD6-1

A culture of the frozen stock of the 3- TOPO+*CYP101* DNA was inoculated into 5mL of LB containing ampicillin and was allowed to grow overnight in the incubator with a shaker at 225rpm and 37°C. The culture was used as inoculum for the 100mL culture containing 50 μg/mL ampicillin that was grown overnight on the shaker at 225rpm and 37°C. The DNA was extracted from the 100mL culture using midiprep. The presence of the insert was confirmed again by restricting 3-TOPO+*CYP101* using NotI and EcoRV enzymes (shown in Figure 27) All the restriction sites in the multiple cloning site of the vector were analyzed for an enzyme suitable to act with ApaI for cloning the cytochrome $P450_{cam}$ gene from 3-TOPO+*CYP101*. None of the sites on the MCS of the vector were found to be suitable enough to act with ApaI for cloning.

Also analysis of the 3- TOPO+*CYP101* plasmid revealed that there were two ApaI sites flanking the cytochrome $P450_{cam}$ gene (shown in Figure 28), therefore ApaI was chosen for restricting the cytochrome $P450_{cam}$ gene and ligating into



Lane	Sample
1	Lambda HindIII/EcoRI
2	-
3	-
4	3-TOPO+CYP101 w /NotI
5	3-TOPO+CYP101 w /EcoRV

Figure 27: Gel showing 3- TOPO+P450 w/ NotI and EcoRV

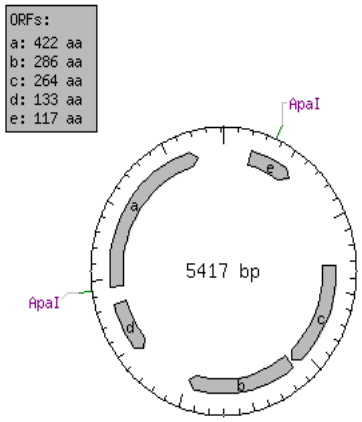


Figure 28: 3-TOPO+CYP101 plasmid showing ApaI sites flanking the cytochrome $P450_{cam}$ gene.

pKSD6-1. TOPO+*CYP101* with ApaI gave two distinct bands on the gel one of 1.4 kbp (containing CYP101) and the other 4.0 kbp (shown in Figure 29).

3.6 Ligation of the insert into the vector

100μL solutions of the vector (pKSD6-1) and the insert (3-TOPO+CYP101) were restricted with ApaI on a 0.7% agarose gel. After an overnight restriction, the solutions were subjected to gel electrophoresis to get the desired 10.1-kbp vector and 1.4+4.0 kbp insert band (Figure 30). The insert (1.4 kbp) and the vector 10.1kbp bands were excised and were using Wizard PCR preps DNA Purification Kit (Promega, WI). 50μL of DNA were eluted for each of the vector and insert DNA from their respective bands.

The concentration of DNA was measured in the eluted solutions using SPECTRAmax Plus^{384} (Molecular Devices, Sunnyvale, CA). The amounts of DNA present in the vector and the insert solutions were 327.5ng and 249.75ngrespectively from the calculations based on the absorption spectrum of the DNA samples.

- 300ng of the vector was the basis for the ligation reaction. Depending on this the amount of insert required was calculated using the recommended 1:1 vector-insert ratio.

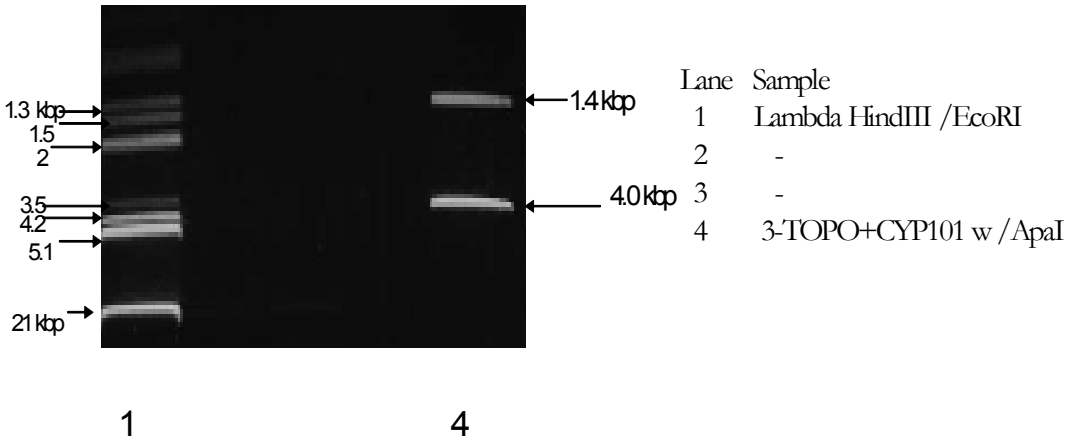


Figure 29: 3-TOPO+CYP101 with /ApaI restriction

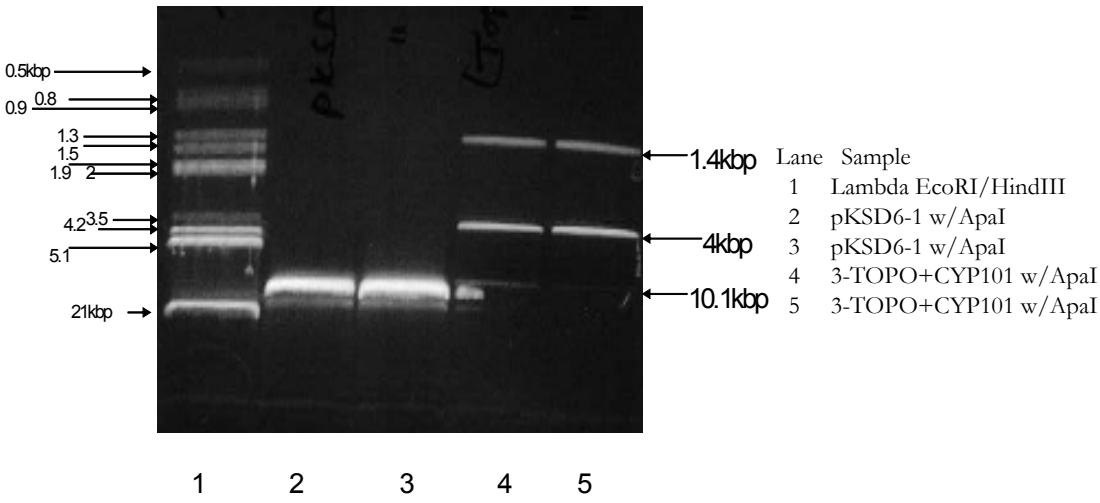


Figure 30: Gel with insert and vector DNA.

Basis: 300ng vector

$$\frac{300ngvector \times 1.8kbp}{10.1kbp} \times \frac{1}{1} \quad = 53.465\text{ng insert} \qquad (2)$$

The required volumes of the insert and vector were then reduced to much lesser volumes in a DNA mini (JOUAN Nordic A/S, Denmark) at 1mbar pressure to suit the total required volume of the ligation reaction (10μL). The ligation reaction was placed on a vortex for overnight incubation at 4°C.

The protocol for the ligation reaction is given below

- Vector: 6μL (6.55ng/mL)
- Insert: 2μL (4.995ng/mL)
- T4 ligase: 1μL
- 10X ligase buffer: 1μL
- Total Volume: 10μL

After vortexing at 4°C overnight the ligation reaction was transformed into chemically competent *Escherichia coli* (DH5α) cells. Transformed 3μL of the reaction mixture into three samples of 50μL DH5α cells using chemical transformation protocol. The transformants were plated on LB agar plates

containing 30μg/mL chloramphenicol and 50μg/mL ampicillin. The plates were incubated at 37°C for overnight growth.

After 24 hrs there were no colonies found on any of the plates. A single white colony was observed after 48hrs, which was immediately picked and grown in 5mL LB containing 30μg/mL chloramphenicol and 50μg/mL ampicillin. Again after another 48hr period at 37°C, growth was observed in the test tube. 2mL of this culture was stored in 1mL vials with 20% glycerol at -80°C. The remaining 3mL culture was purified using Wizard Plus Miniprep (Promega, WI). The uncut DNA was run on the gel to verify the presence of DNA. No DNA was found on the gel.

- Various attempts to clone the *CYP101* from 3-TOPO+*CYP101* to pKSD6-1 were made after the first one failed. These included ligating the vector and insert in ratios of 1:3, 3:1 (vector: insert), increasing the quantities of the vector and insert ligated. These were unsuccessful as well.

3.6 Production of the $P450_{cam}$ enzyme

When the frozen cells that were stocked from the first transformation were cultured in 100mL LB containing 30μg/mL chloramphenicol and 50μg/mL ampicillin to give another attempt at purifying the gene, it was observed that the

culture was pink in color. Purified $P450_{cam}$ enzyme has a brown color and a comparison to the color of the recombinant culture indicated that further investigation for the presence of the enzyme was needed.

The cytochrome $P450_{cam}$ enzyme was extracted from the *Escherichia coli* containing the recombinant gene by using the following protocol.

- The *Escherichia coli* (DH5α) cells were grown to an O.D. (optical density) of 0.6 at 37ºC, the flask containing the culture was transferred briefly to an incubator at 30ºC for fifteen minutes. Adding 1mM IPTG and incubating the cell culture at 30ºC and 90rpm for 36 hours induced the production of cytochrome P450cam from the recombinant pKSD6-1+CYP101 plasmid.

- Then the cells were harvested by centrifuging in the Sorvall RC-5B Refrigerated Super speed centrifuge (Dupont Instruments, DE) in an SS-34 rotor at 13.5K for twenty minutes to obtain the cell mass. Excess supernatant was removed from the cells using a pipet.

    - The cell mass obtained: 1.21g

    - Volume of the 50mM Tris HCL added at pH 8.0: 4.84 mL

    - Amount of 0.5mM EDTA added to make an overall 1mM solution: 9.68μL

- 19.36mg of lysosyme (Sigma-Aldrich, MO) was added to this mixture and then the mixture was incubated on ice for an hour. After incubation the cells were sonicated using Sonicator-Cell disruptor Model W-225R (Heat systems-Ultrasonics, Inc. currently known as Misonix, Inc, NY). The sonicator settings were as follows:

  - ✓ Stirrer speed: 1.5rpm
  - ✓ Output control on '5'
  - ✓ Time the solution was sonicated for: 1min
  - ✓ Interval between the sonications was for: 45secs
  - ✓ % Duty cycle: 42%
  - ✓ Under these conditions the frequency dial is at 30

While sonicating the cells the height of the stand on which the beaker was placed was adjusted accordingly. The cells for sonication were kept in a test tube in the beaker containing ice. Also sufficient care was taken so that the probe was immersed in the cell solution but didn't touch the walls or the bottom of the tube.

After sonicating the cells 1.382μL (i.e.19.36 u) of Rnase A (QIAGEN, CA) was added from a 14u/μL solution to make up a 4u/μL solution. Then 38.72μL of 2u/μL DnaseI solution was added to the lysed cells containing Rnase A to make a 16u/μL DnaseI containing solution. This mixture was stirred at 4°C for an hour at a stirrer speed of 4rpm. This mixture was then centrifuged in Sorvall RC-5B Refrigerated Super speed centrifuge (Dupont Instruments, DE) in an SS-34 rotor at 13.5K for twenty minutes.

The centrifuged mixture was removed and the cell extract was carefully decanted into another clean sterile 20mL scintillation vial. 50μL of this extract was used for spectral analysis. The spectral analysis revealed the presence of an enzyme with a peak at 414 nm (as seen in Figure 31). It was confirmed from Figure 31 that a heme protein was present. To observe the peaks accurately (as seen in Figure 32) the enzyme was further purified using ultra filtration with a 30000 molecular weight membrane.

3.7 Spectral analysis of $P450_{cam}$ after Carbon monoxide binding.

In order to prove the presence of the $P450_{cam}$ enzyme in the cell extract the first and the most important assay is the carbon monoxide binding assay. This assay has another special requirement, that the enzyme solution be reduced with dithionite anaerobically. Dithionite gets quickly oxidized on exposure to air. Also the dissolved oxygen in water can oxidize the dithionite when the required 1mM

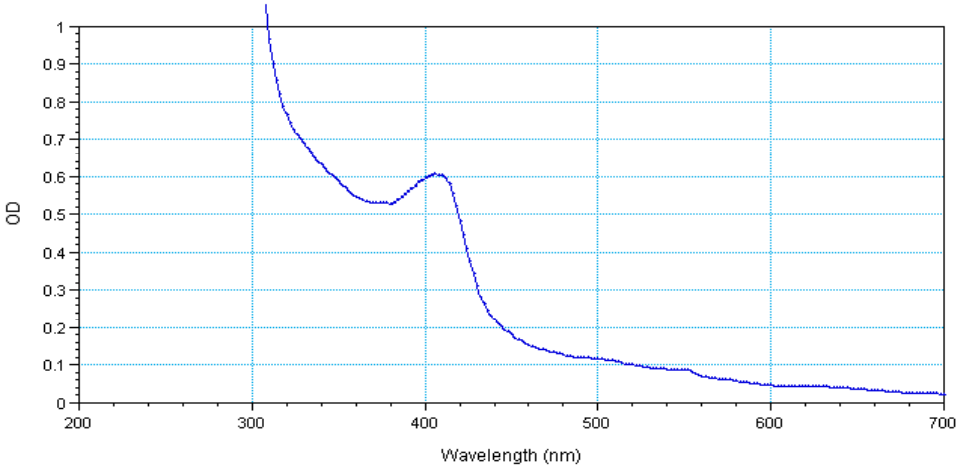


Figure 31: Absorbance spectra of the oxidized $P450_{cam}$

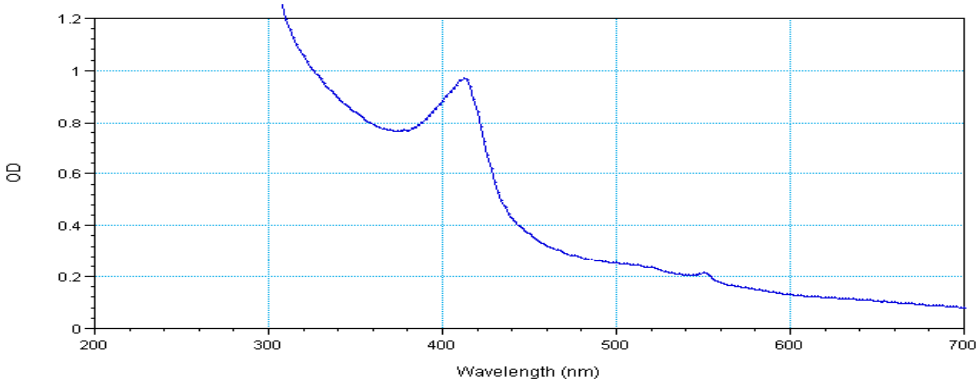


Figure 32: Absorbance spectra of the concentrated $P450_{cam}$

solution is prepared. Therefore it is essential to keep the dithionite oxygen-free in all forms and add it to an oxygen-free enzyme solution for the assay to be successful. If the enzyme solution is not oxygen free it oxidizes the dithionite and produces acid and other oxidation products that denature the enzyme.

Anaerobic conditions were established in the laboratory by purging the enzyme with nitrogen. After a number of trials it was found that this was not sufficient to provide the requisite condition, therefore another method was devised for better results. A vial of water was crimp-capped, the headspace was vacuumed and was purged with nitrogen alternately, and this vacuum-gas method was repeated thrice to deaerate the headspace above the water level in the vial. The dissolved oxygen in the water was then removed by repeating the same procedure. Thus the water was deaerated. This water was used for the purpose of preparing deaerated dithionite solution.

The enzyme solution was placed in a cuvette with a rubber stopper, it was made oxygen-free by using the vacuum-gas method (gas used for purging was nitrogen). The anaerobic dithionite solution was then carefully added to the oxygen free enzyme solution in a ratio of 5:1(dithionite: enzyme). The spectrum was then measured immediately to check if the enzyme was reduced. There was a shift in the Soret band from 414nm to 411nm (shown in Figure 33) as expected. Carbon monoxide gas was bubbled through the solution for five minutes using a

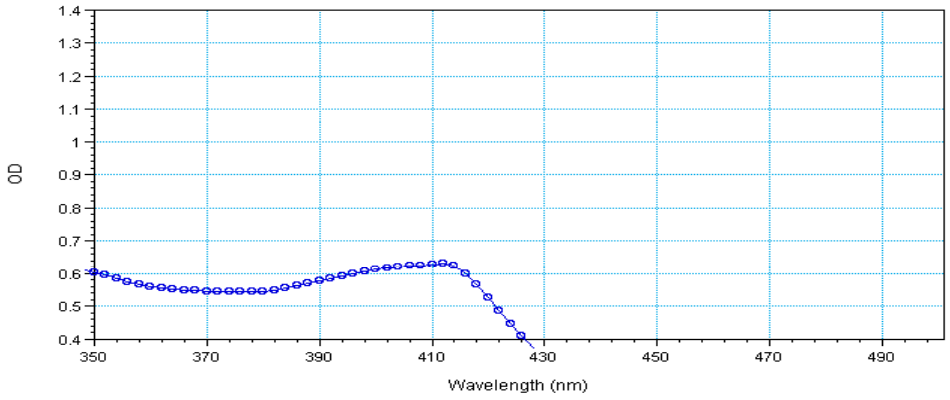


Figure 33: Absorption spectra on adding dithionite (shift from 414nm to 411nm)

needle to maintain the anaerobicity; addition of carbon monoxide didn't show the expected shift in the Soret band from 411nm to 446nm.

3. 8 Assay using the camphor bound cytochrome $P450_{cam}$

The cytochrome $P450_{cam}$ isolated was not in the camphor bound form. To cytochrome $P450_{cam}$ at pH 7.4 (pH was adjusted by adding 50Mm Tris HCl buffer) KCl and camphor were added such that the final enzyme solution had 0.1M KCl and 5mM camphor. The absorption spectrum of this solution was measured from 200nm to 700nm. The absorption spectrum was expected to shift from 414nm to 392nm; instead the shift was from 414nm to 405nm (as seen in Figure 34).

Then dithionite was added anaerobically as before to the camphor bound enzyme and the spectrum were measured but it didn't show the expected shift of the

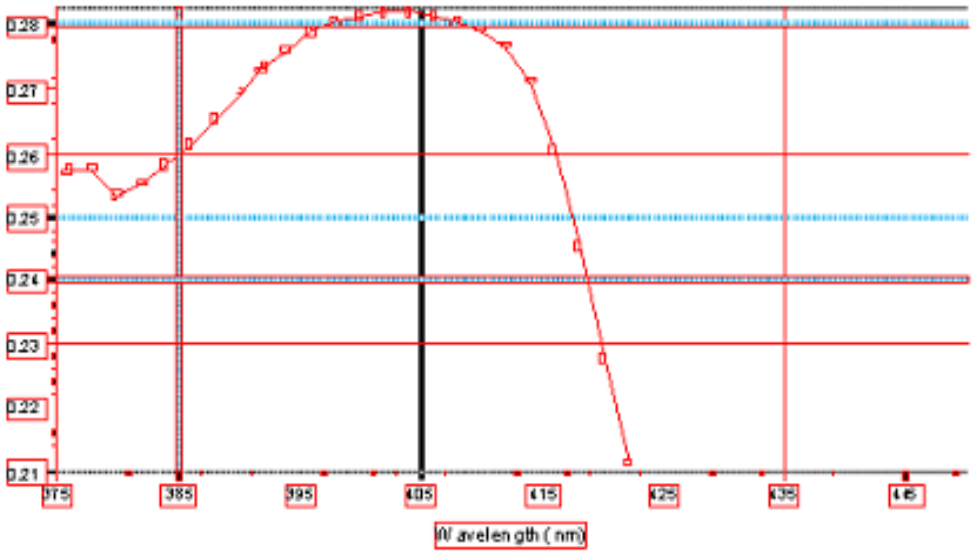


Figure 34: Absorption spectra of the Oxidized $P450_{cam}$ containing 5mM camphor

Soret band to 409nm (Figure 35). Neither did bubbling carbon monoxide through the dithionite reduced solution shift the peak to 446nm (Figure 36).

3.9 Control experiment

Escherichia coli (DH5α) cells containing pKSD6-1 were grown in LB containing 50μg/mL ampicillin and 30μg/mL chloramphenicol at 37°C overnight. These cells showed no growth. Upon verification it was found that these *Escherichia coli*

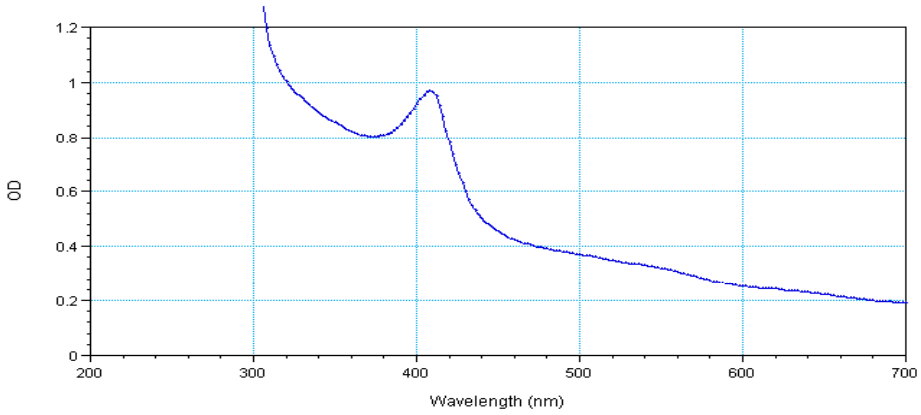


Figure 35: Absorbance spectra of oxidized $P450_{cam}$ and camphor with dithionite

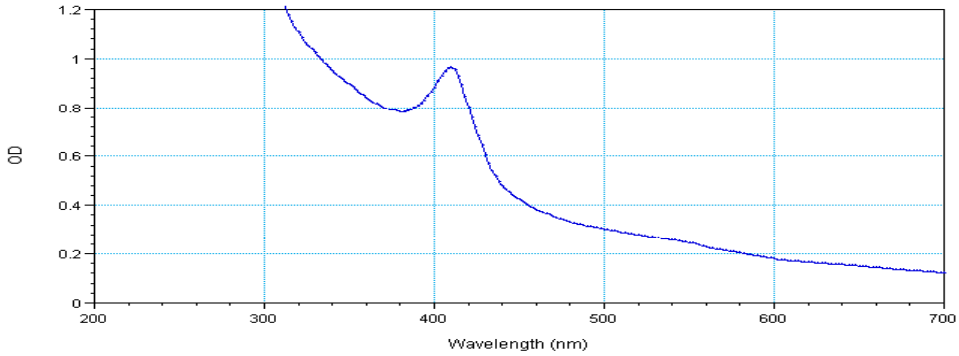


Figure 36: Absorbance spectra of $P450_{cam}$+ Camphor+ Dithionite +Carbon monoxide

cells containing the pKSD6-1 plasmid had always been grown only in 50μg/mL ampicillin. Therefore the control cells were grown in LB containing only.

50μg/mL ampicillin whereas the recombinant cell culture (*Escherichia coli* containing pKSD6-1 +P450) was grown in LB (Luria Broth) containing 50μg/mL ampicillin and 30μg/mL chloramphenicol and the protein was expressed in each culture according to the same protocol as before. An absorption spectrum was taken for the protein extract from each cell culture (Figure 37), which proves that only the recombinant pKSD6-1+*CYP101* cells produce a peak at 420nm.

As seen from Figure 37 it is evident that the control doesn't produce the cytochrome $P450_{cam}$ enzyme and therefore the recombinant pKSD6-1+CYP101 DNA contain the cloned $P450_{cam}$ gene. To make it certain that the control doesn't produce any peak at all, the protein extract from the cell culture was concentrated to thrice its original concentration and its absorption spectra was again measured (shown in Figure 38).

3.10 Analysis of Recombinant DNA.

It was also necessary to prove the presence of the cytochrome $P450_{cam}$ gene in the recombinant plasmid using molecular biology tools like restriction and gel

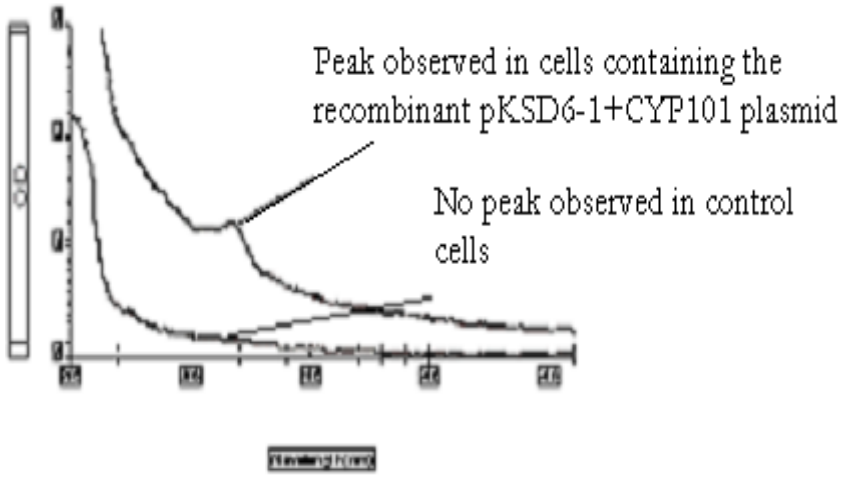


Figure 37: Absorption spectra of the protein extract control and recombinant culture

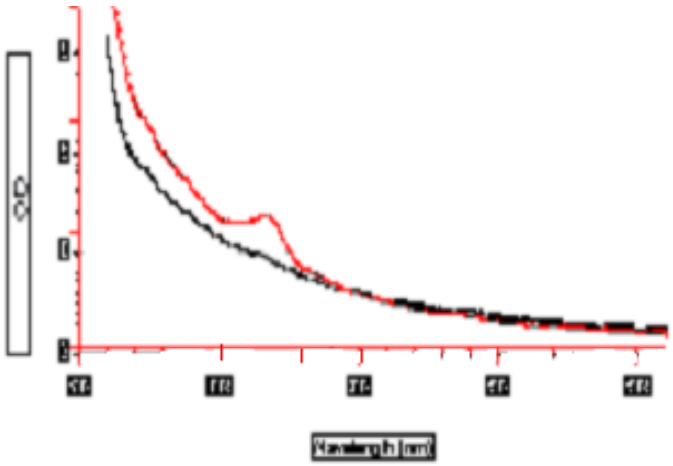


Figure 38: Difference in spectra of control (concentrated) and recombinant protein extract

analysis. Isolation of the recombinant pKSD6-1 containing $P450_{cam}$ was a difficult task. The DNA couldn't be isolated using the standard boiling prep protocol. After many trials the recombinant DNA was isolated .The uncut band was run on the gel and its size was observed to be 20kbp (seen in Figure 39). When this recombinant DNA was restricted with NotI, ApaI and BspEI, NotI +BspeI, NotI+ApaI the solutions were analyzed on an agarose gel (shown in Figure 40), they showed no DNA at all. Repetitive attempts produced similar results. This led to the conclusion that the DNA might have been degraded by Dnase activity.

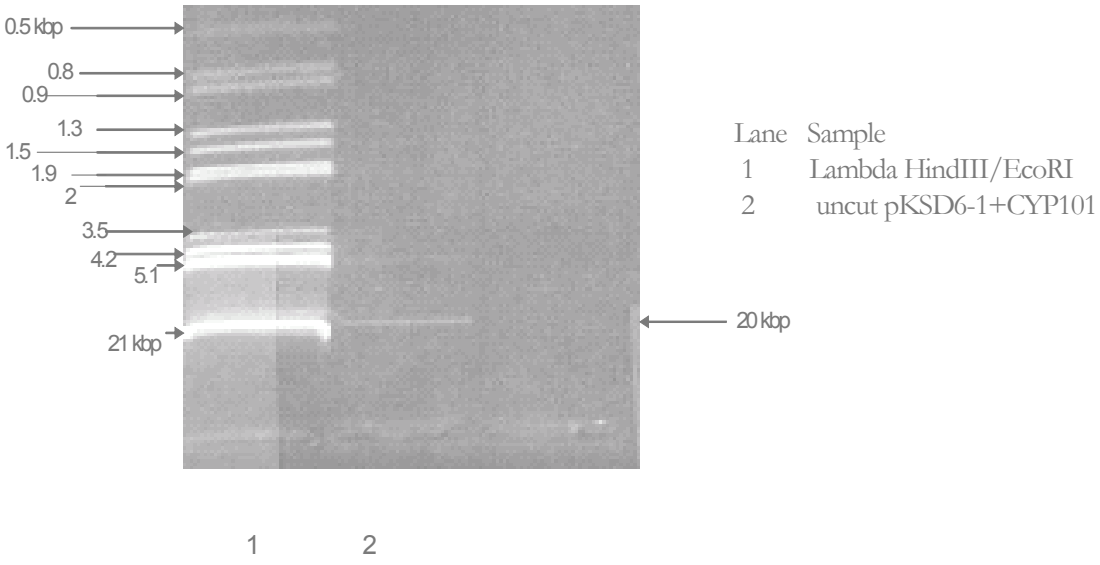


Figure 39: Agarose gel showing the uncut recombinant pKSD6-1+P450 at 20kbp.

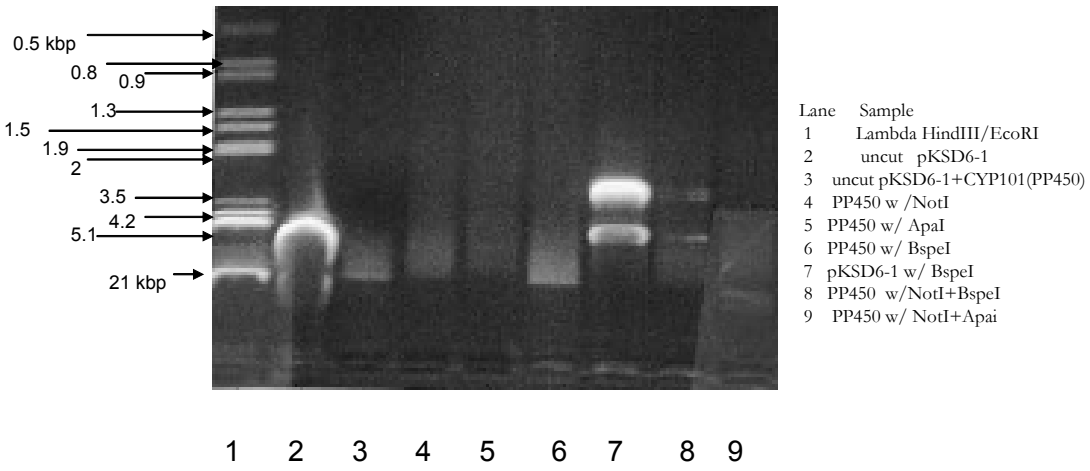


Figure 40: Restriction digests of recombinant pKSD6-1+P450 with various enzymes.

C h a p t e r 4

ANALYSIS AND DISCUSSION

4.1 Recombinant Protein Expression in *Escherichia coli*

The Gram-negative bacterium *Escherichia coli* is the most widely- used organism for producing heterologous proteins because of its ability to grow rapidly and at high density on inexpensive substrates, its well-characterized genetics and the availability of a large number of cloning vectors. The unique and subtle structural features of the gene sequence, the stability and the translational efficiency of the mRNA, the ease of protein folding, degradation of the protein by host cell proteases, major differences in codon usage between the foreign gene and native *Escherichia coli* and potential toxicity of the protein to the host *Escherichia coli* are some of the probable reasons why a gene might not be efficiently expressed in *Escherichia coli*. [37, 38]

Several elements need to be considered for efficient expression of genes and the highest level of protein synthesis. Figure 41 below shows the essential architecture of an *Escherichia coli* expression vector. The promoters in *Escherichia coli* are 10-100bp upstream of the ribosome-binding site (RBS) and the promoter

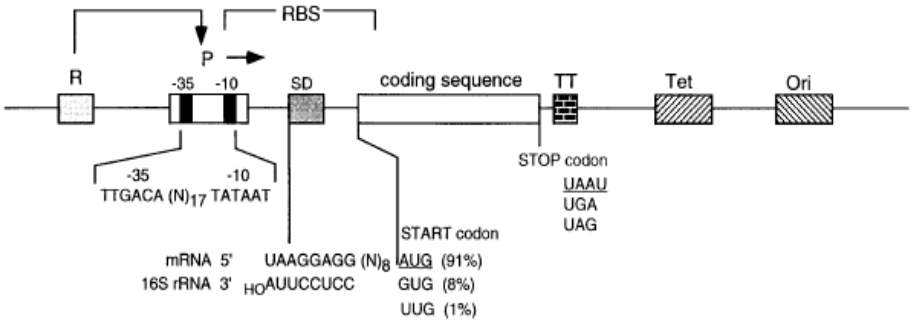


Figure 41: Schematic presentation of the salient features and sequence elements of a prokaryotic expression vector. [38]

is under the control of a regulatory gene (R), which may be present either on the vector or on the host chromosome. The promoters of *Escherichia coli* consist of a hexanucleotide sequence that is about 35bp upstream of the transcription initiation base (-35 region), separated by a spacer region from another hexanucleotide sequence (-10 region). [38]

The RBS (ribosome binding site), which spans approximately 54 nucleotides bound by positions –35 (±2) and +19 and +22 of the mRNA coding sequence is located downstream of the promoter. The Shine-Dalgarno (SD) site interacts with the 3′ end of the 16SrRNA during translation initiation. The range of the distance between the SD site and the initiation codon should be 5-13bp. The transcription terminator is located downstream of the coding sequence and serves as a signal to terminate

transcription. Vectors also contain a gene that confers resistance on the host to aid in plasmid selection and propagation. [38]

4.2 Requirements for expression of cytochrome $P450_{cam}$ in *Escherichia coli*

The active enzyme cytochrome $P450_{cam}$ requires the proper coordination of the non-covalently bound heme, which must be produced and incorporated into the nascent polypeptide.

For successful synthesis of cytochrome$P450_{cam}$ in *Escherichia coli*

- The gene encoding the particular P450 should be cloned downstream of an efficient promoter and the ribosome-binding site should be present on a multicopy number plasmid.

- An efficient promoter should direct the synthesis of hybrid mRNA molecules that contain signals immediately upstream of the initiation codon that directs *Escherichia coli* ribosomes to initiate translation of the inserted gene sequence.

- On being translated the nascent polypeptides should bind a molecule of heme and fold into the final tertiary structure.

- Production of the recombinant protein in *Escherichia coli* places additional burden on the cellular metabolism in *Escherichia coli* and therefore it is necessary for the cell to produce the proteins (Gro ES and Gro EL) required for folding. [39]

4.3 Synthesis of cytochrome $P450_{cam}$ enzyme in *Escherichia coli* using the vector pKSD6-1

Cytochrome $P450_{cam}$ gene was cloned at the ApaI site in the multiple cloning site of the shuttle vector pKSD6-1. The MCS (multiple cloning site) of pKSD6-1 has been derived from the pBluescript II KS (-) plasmid. [18] Since the entire sequence of the pKSD6-1 is not known (not given in [18]) the sequence of the gene-containing fragment of recombinant pKSD6-1+*CYP101* was constructed using some of the information that is given in [18]. First, the 3-TOPO+CYP101 sequence was constructed using pCR2.1 TOPO and the CYP101 sequence (the gene had appropriate primer ends). From this sequence the ApaI-ApaI fragment containing the cytochrome $P450_{cam}$ gene was taken and was added at the ApaI site in the MCS of pBluescript II KS (-)(Statagene). The various factors such as the promoter, the Shine-Dalgarno sequence etc that affect the expression of the protein in *Escherichia coli* were analyzed with the above sequence.

When the 5'ApaI-ApaI 3' was inserted into the ApaI in the MCS of the shuttle vector, the cytochrome P450 gene was found to be expressed under the T7

promoter. Some of the essential features for expression of the protein were found to be missing in the sequence. Though the -10 consensus sequence (TATA) was present, the -35 consensus sequence (TTGACA) was not as it should be, the very important Shine-Dalgarno sequence was not present between the promoter and gene and finally when this sequence was checked for ORFs (Open reading frames) using NEB cutter the protein being translated by the gene was 451aa(amino acid residues) instead of 414aa and the protein start codon was TTG(present on the vector) instead of ATG (present on the plasmid). These factors led to the conclusion that the gene might have been ligated in the reverse direction.

The inverse sequence of the 5'ApaI-ApaI 3' was obtained from [40] and inserted into the ApaI site of the shuttle vector (sequence given as Figure 42). This ligation puts the gene under the control of the T3 promoter. The sequences when analyzed using the NEB cutter showed the production of 415aa with an ATG start codon. The way in which the sequence was ligated causes the 3'-5' strand to be the coding gene (Figure 43). The distance between the start codon ATG and the T3 promoter was found to be only 26 nucleotides. The -10 sequence and -35 sequence for the *lac* promoter in *Escherichia coli* are 5'TATGTT 3' and 5'TTTACA 3' since the gene was ligated in the inverse direction the-10 and the -35 regions found on the 5'-3' of the gene are 5' AACATA 3' and 5' TGTAAA 3'; these are situated 69bp (base pairs) and 93bp upstream of the promoter on the coding gene

CTGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTAC
ACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTC
CCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCA
AAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTG
ACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCG
GTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAAC
AAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTCCATTCGCCATTCAGGCTGCGCAAC
TGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTG
CAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAG
CGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAG
TGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGG**GGGCCC**T
CTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCGCCCTTCATGAGCTCAAG
CTTTGGATAGCGGTGGTAGGTGATGGGGAATTACCGTCGCATCCCCCACCAACCGGATATTGCCTTCC
CAGCCGCTGGCGCGCAGGCCGAAGGCGACCTCAACGCCAGCCAGTCCGGTACCGACGATGACCACGTT
GTCGTTTGCGTTCACTTAGCACGCACTCCCATGTGT***TTA***TACCGCTTTGGTAGTCGCCGGATCCCAGAC
CAGAGGGAGTGCCTGCACGCCGCTGACGATGCCGCTCTTGTGCTGAATCTGGGCACCCGGGGCAATG
GAGAAGTCAGGAATCCTGGTCAGCCATTCCTTGAGGGTGACGATGATTTCCCGGCGGGCCAGGTGCT
GGCCAAGGCACAGATGGCTGCCGTGGCCAAAGGTGGTGTGTGAAACCTTTTGGCGACTGAAGTCGAC
GTGCATCGGGCAGGCGTTTTCGCGCTCATCCAGGCCAGACAGCATCTGCGGTAGCAGGATCTGGTCAC
CTTTCTTCAGTTGCACGCCATGAAACTCGTAATCGGAGGTGAGGATGCGGCCATCGGCAACCAGCGAG
AAGCGCCGGAGTAGTTCCTCGCAAGCGGCTGGAATACGCTCGGGACGCTCGATCAGCTCCTGGCGAT
GCTCCGGGCTTTTGGCCAGGAACTCCATGCTGAAGCTGAGGAAATTGACCACCGTATCCAGGCCGCCG
ACCAGTAACAGGCCACACATCCTCTTGGCTTCGTCACTGGTGATCGGTCGCCCATTGACCTGGCCGTTG
GCAACGATGCTGATAGCGTCGGTTCCCGGCTTCTGCCTGCGTTGCTCGATGATCGGTATCAGATAGTC
GTAGAGCGCCTCCTTGGCCTCTGCGAAGGTCATGCTGCCATCCGGACGGGTCATCTGATCCGTTAGGT
ATTTCAAGTGCGGGATATCTTCTTCCGGTAGACCTGCGAGCAGCATGAAGATGCGTATCGGGAAGGGT
TCGGCGTAGTCCTCGGTGAAGTTGCACTGTCCTTGCGGGCGCAGGCTCTCGATCAGCGAGCAGGCCA
GCTCCTGGATCCGGTTCTCCAGCTTATCCACCACCGGCATGCCAACCACTTGGTTGGCCAGCGCACGAA
ACTGGCGCTGCTCGGGCGGATCCATCGAGGTGGGAATGAAGTCGTAGGCTTCGCCGGCTTCACGAGG
GATGAACGGGCACTCGCTGGAAAAGTGGCGGTAATCTTCATAGGCCTCACGGATCAGTTGGCCGCGA
GTGGCGATCCAGTGTCCGCCGTTGCAGCGAGTCCACACCAGATCCGGTACGTTTGATTCTTGCAGAAC
TGCCCAGGCCTCCTGCACGCCGGCAGACAGATTCGACGGATTGTACATGTCGAAGTCGAATACCAGGT
GCTCTGGCACATGGGGTGGCAGAGGGGCAAGATTGGCGTTGCTTTGTATGGTTTCAGTCGT***CAT***TGT
GGGCCCGGTACCCAGCTTTTGTT**CCCTTTAGTGAGGGTTAATT**GCGCGCTTGGCGTAATCATGGTCAT
AGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACAC**AACATA**CGAGCCGGAAGCATAAAG**T**
GTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCC
AGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCG
TATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGG
TATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG
TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCT
CCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT
AAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCG
GATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCA
GTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGC
GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC
ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA
ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAA
GAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAG
CAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCA
GTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCT
TTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAA
TGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCG
TCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC
CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGG
TCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCC
AGTTAATAGTTTGCGCAACGT

CCCTTTAGTGAGGGTTAATT-T3 Promoter ,GGGCCC-ApaI site,*TTA C-terminal,*
AACATA -10 SITE,*CAT N-terminal* TGTAAA -35 SITE

Figure 42: Sequence of a part of pKSD6-1 containing the *CYP101* gene in its MCS

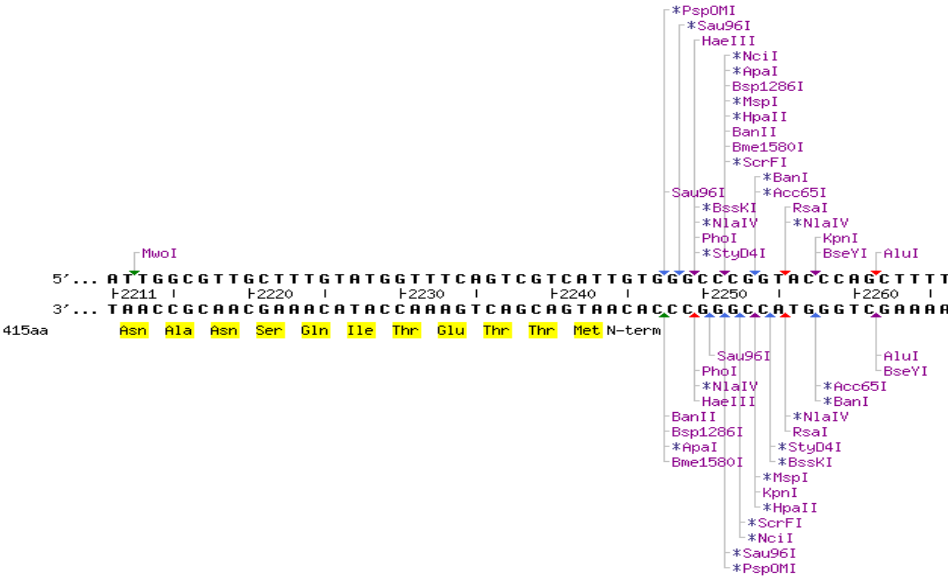


Figure 43: The direction of translation of the protein

respectively. The Shine-Dalgarno (SD) sequence for the *lac* I is 5' GGUGGU 3' which should be found on the sequence in Figure 42 as 5' ACCACC5' -3' between the promoter and the coding gene, but such a sequence was not present. This suggests that there might be another possible SD sequence as they vary from one promoter to another. When the sequence given in Figure 42 was analyzed in NEB cutter for ORFs, the sequence encoding the cloned gene was found to produce a protein of 415 aa (Figure 44). From the above analysis it is clear that the enzyme being produced is a P450.

```
MTTETIQSNA NLAPLPPHVP EHLVFDFDMY NPSNLSAGVQ
EAWAVLQESN
VPDLVWTRCN GGHWIATRGQ LIREAYEDYR HFSSECPFIP
REAGEAYDFI
PTSMDPPEQR QFRALANQVV GMPVVDKLEN RIQELACSLI
ESLRPQGQCN
FTEDYAEPFP IRIFMLLAGL PEEDIPHLKY LTDQMTRPDG
SMTFAEAKEA
LYDYLIPIIE QRRQKPGTDA ISIVANGQVN GRPITSDEAK
RMCGLLLVGG
LDTVVNFLSF SMEFLAKSPE HRQELIERPE RIPAACEELL
RRFSLVADGR
ILTSDYEFHG VQLKKGDQIL LPQMLSGLDE RENACPMHVD
FSRQKVSHTT
FGHGSHLCLG QHLARREIIV TLKEWLTRIP DFSIAPGAQI
QHKSGIVSGV
QALPLVWDPA TTKAV
```

Figure 44: 415 aa sequence encoded by the gene.

4.4 The chloramphenicol resistance gene

The multitude of attempts made at cloning the cytochrome $P450_{cam}$ gene from 3-TOPO+*CYP101* into pKSD6-1 were met with consistent failures. It was always presumed that there was something wrong with the vector: insert ratio and lots of attempts were made in trying to improve the quantities and the ratios of the vector and insert used whereas in fact the problem might have actually been with the expression of the resistance genes in the vector.

While performing the control experiment it was observed that the *Escherichia coli* (DH5α) cells containing the shuttle vector pKSD6-1 (without the *CYP101* gene) was unable to grow in a LB containing both chloramphenicol and ampicillin,

though it has both the resistance genes as seen in Figure 16.But the same cells were capable of growing in LB containing only ampicillin.

The *Escherichia coli* (DH5α) cells containing only pKSD6-1 were cultured previously to isolate vector DNA. This vector DNA was used in constructing the recombinant plasmid of pKSD6-1 containing $P450_{cam}$. The cells of the single positive clone isolated after transforming the recombinant plasmid initially took about 48hrs to grow in media containing both antibiotics.

Requirements of some gene products change over time. The need for enzymes in certain metabolic pathways wax and wane as the food sources change or are depleted. From the above observations, it is possible that the chloramphenicol resistance gene was lost due to a possible mutation. It was cultured only in media with ampicillin in the lab for over a year and it could have been lost due to a lack of selective pressure over generations. This could be one of the main reasons why the clones didn't grow on agar plates containing ampicillin and chloramphenicol.

4.5 Spectral Analysis of the $P450_{cam}$ enzyme from the recombinant plasmid.

The oxidized enzyme in the camphor-free form has two absorption maxima; one at 414nm and the other at 550nm instead of the expected maxima at 418 and 535nm, the other expected peaks at 280nm and 570nm were not observed. Also the absorption maxima did shift from 414nm to 411nm when reduced with

dithionite ($Na_2S_2O_4$) and gassed with carbon monoxide. Camphor binding assay showed absorption maxima shift to 405nm instead of the expected 392nm, but the other shifts were not observed.

Theoretically the absorption spectra of the oxidized camphor-free form of the cytochrome $P450_{cam}$ enzyme has α andβ bands at 571nm and 538nm, the Soret band is 417nm, a γ band at 365nm and aromatic acid content is seen at 280nm with a shoulder at 290nm, in the experiments only the Soret band was observed at 414nm and instead of the two α and β bands a single peak was observed at 550nm.

The presence of the Soret band at 414nm confirms the presence of the heme protein, cytochrome $P450_{cam}$. Actually both α and β bands fuse into a single band at 550nm because of the formation of a complex between the reduced cytochrome $P450_{cam}$ and carbon monoxide. The single band formed at 550nm in the experiments. The other bands like γ band at 365nm and others could not be observed because the sample was not very pure. Addition of camphor to the enzyme solution shifts the Soret band to 405nm instead of 392nm.Whereas reduction with dithionite and later addition of carbon monoxide didn't show the expected bands.

4.6 Possibility of the enzyme being cytochrome P420

Cytochrome P420 is an inactive form of cytochrome P450 that can be formed from all types of known P450s by incubation with acetone or exposure to high temperature and pressure. The P420 has a Soret absorption band at 420nm and the α and β bands coalesce into a single band in the α band region (at571nm), on adding sodium dithionite the Soret band shifts to 423nm and the α band is at556nm. Soret absorption of the ferrous-CO complex is more intense with a 420nm maximum. Addition of a substrate or inhibitor doesn't alter the P420 absorption bands. P420 is a collective term as some forms can be converted back into the active enzyme and show the ferrous-CO spectrum of the native protein while others cannot. Also the spectra of the P420 samples can vary in the Soret and α regions. [41]

The Soret band at 414nm, the coalescence of the α and β bands in a single band at 550nm observed in the results suggest the presence of P420, the inactive form $P450_{cam}$ in the enzyme solution. But there is no shift in the Soret or the α band on reducing the enzyme with dithionite; this could be because the reduction was not done under appropriate (anaerobic) conditions. Again addition of camphor to the enzyme solution shouldn't produce a shift in the absorption bands if the enzyme were P420, but the enzyme solution did show a shift in its Soret band

from 414nm to 405nm on addition of camphor, so the enzyme may not be a P420.

4.7 Other Possible reasons for the lack of enzymatic activity

Lack of proper thiolate coordination: The heterolytic oxygen scission of the heme bound oxygen is supposed to be facilitated by the strong electron donation of the anionic thiolate axial heme ligand. The inactivated form P420 lacks such thiolate coordination. Any mutation at the Cys357 residue is bound to destroy the heme environmental structure resulting in the loss of enzymatic activity.

Errors in the PCR: While producing the *CYP101* gene from PCR using the pCWori+*CYP101* as template there is a possibility that there were some mutations because of the limited fidelity of the *Taq* polymerase enzyme. If there were any such errors these have not been investigated in any way. Even a single mutation in the gene can alter the amino acid residues produced, hence the protein and thereby its spectral characteristics.

The enzyme is being produced in an inactive form, the causes for the lack of activity in the enzyme may be because of some deleterious mutation in the gene encoding for the enzyme. Also the disappearance of chloramphenicol resistance activity from the shuttle vector played an important role, in the sense better clones could have been isolated if this problem was observed earlier.

Chapter 5

SUMMARY AND FUTURE WORK

5.1 Summary

The fundamental idea of the project was to clone the *CYP101* gene into the *Rhodococcus-Escherichia coli* shuttle vector pKSD6-1, so that the enzyme can be expressed in *Rhodococcus* and it can be tailored using random mutagenesis to act on the carcinogenic PAHs, which are not readily degraded by naturally occurring enzymes or on the alkanes for producing commercially useful alcohols.

The cytochrome $P450_{cam}$ (*CYP101*) gene was produced with appropriate ends using pCWori +CYP101 as a template for polymerase chain reaction (PCR), and employing suitable primers that were designed using sites ApaI and SacI for the forward and the backward primers respectively. Direct cloning of the PCR product into the shuttle vector pKSD6-1 was not possible, because SacI site was found to have three restriction sites on the shuttle vector pKSD6-1. It was therefore necessary to find an alternative way to clone the *CYP101* gene into the shuttle vector pKSD6-1.

The pCR 2.1 TOPO is a commercially available vector with 'T' endings into which the cytochrome $P450_{cam}$ gene that was produced from the PCR with

P4501F and P4501R primers was cloned. This recombinant plasmid of TOPO+*CYP101* was transformed into *Escherichia coli.* The positive clones were restricted with NotI enzyme to test for the presence of the CYP101 gene. The presence of a 5.4 kbp product from the colony 3 during gel analysis confirmed the presence of the gene

ApaI site was found to be the only restriction site that was compatible for cloning the cytochrome $P450_{cam}$ gene from 3-TOPO+*CYP101* to pKSD6-1. The gene was restricted with ApaI enzyme from 3-TOPO+*CYP101* and was cloned into pKSD6-1 that was also restricted with ApaI and the recombinant pKSD6-1+*CYP101* plasmid was transformed into *Escherichia coli (DH5α).* Only one positive clone was found, which was grown in suitable medium containing the suitable antibiotics, chloramphenicol and ampicillin and stored.

The cytochrome $P450_{cam}$ enzyme was expressed from the aforementioned positive clone containing the $P450_{cam}$ gene in pKSD6-1. Spectral analysis of the protein extract revealed an enzyme showing a peak at 414nm instead of the expected 417nm and one at 550nm. Carbon monoxide binding assay did not show the expected shift in the peaks. The camphor-bound enzyme showed a shift in the Soret band from 414nm to 405nm instead of the expected 392nnm.

The results from the camphor and the carbon monoxide binding assays suggest that the enzyme present may be the inactive form of the cytochrome $P450_{cam}$

enzyme; cytochrome P420 or the cytochrome $P450_{cam}$ enzyme with some mutation. The DNA analysis of the recombinant pKSD6-1+*CYP101* plasmid was unsuccessful. Difficulties in the isolation of the recombinant DNA and degradation of the DNA during restriction analysis have caused speculation that the degradation might be a result of Dnase action.

5.2 Future work

The work presented in this thesis is just a step towards achieving the final goal of expressing cytochrome$P450_{cam}$ gene in a *Rhodoccocus-Escherichia coli* shuttle vector and tailoring the enzyme to act on the hydrocarbons of environmental and industrial interest. The first step towards achieving this end would be to generate a better clone than the available one.

In order to generate a better clone, it is essential to know more about the existing one. Therefore sequencing the present clone would provide a lot of information. From the sequence it could be found out if the gene was inserted properly and whether there are any mutations in the gene. If there are no mutations present in the gene then experiments should be carried out to investigate why the enzyme is not exhibiting its characteristic spectral properties.

Also the sequence of the clone will give information about the chloramphenicol resistance gene. Whether or not another attempt should be made at cloning the

cytochrome $P450_{cam}$ gene into pKSD6-1 will depend on the results of the aforementioned analysis of the existing clone. Then the better clone can be transformed into *Rhodococcus* cells through electroporation, the protein should be expressed and enzyme assays should be performed to test the presence of the cytochrome $P450_{cam}$. If the enzyme is present then it can be tested on polyaromatic hydrocarbons and alkanes for its activity and later engineered to suit the particular process.

BIBLIOGRAPHY

1. David L. Nelson, Michael M. Cox. *Lehninger: Principles of Biochemistry*. (2000) Third Edition, Worth Publishers, 782-783.

2. Stephen G. Bell, Erica Orton, Helen Boyd, Julie-Anne Stevenson, Austin Riddle, Sophie Campbell and Luet-Lok Wong. *Engineering cytochrome $P450_{cam}$ into an alkane hydroxylase.* (2003) Dalton Transactions, 11: 2133-2140.

3. Benjamin P. Unger, I.C. Gunsalus and Stephen G. Sligar. *Nucleotide sequence of the Pseudomonas putida cytochrome $P\text{-}450c_{am}$ gene and its expression in Escherichia coli.* The Journal of Biological Chemistry, 1986, 261: 1158-1163.

4. Harald Bredholt, Per Bruheim, Martin Potocky, and Kjell Eimhjellen. *Hydrophobicity development, alkane oxidation, and crude-oil emulsification in a Rhodococcus species.* (2002) Canadian Journal of Microbiology, 48: 295-304.

5. Philip A. Lessard, Xian M. O' Brien, Devin H Currie and Anthony J Sinskey. *pB264, a small, mobilizable, temperature sensitive plasmid from Rhodococcus.*http://www.biomedcentral.com/1471-2180/4/15.

6. Joseph Sambrook, David W. Russel. *Molecular cloning: a laboratory manual.* Third Edition, Cold Spring Harbor Laboratory Press, New York. Chapter 8, 8.18-8.113.

7. Caroline S. Miles, Tobias W. B. Ost, Michael A. Noble, Andrew W. Munro, Stephen K. Chapman. *Protein engineering of cytochromes P-450.*Biochimica et Biophysica Acta (2000), 1543: 383-407.

8. http://drnelson.utmem.edu/CytochromeP450.html

9. F. Peter Guengerich. *Cytochrome P450.Enzyme systems that metabolize drugs and other xenobiotics.* Edited by C. Loannides. (2002) John Wiley & Sons Ltd, 34-65.

10. Burkhard Schiffler and Rita Bernhardt. *Bacterial (CYP101) and mitochondrial P450 systems-how comparable are they?* Biochemical and Biophysical research communications. (2003) 312: 223-228.

11. Max J. Cryle, Jeanette E. Stok and James J. De Voss. *Reactions catalyzed by bacterial cytochromes P450.*Australian Journal of Chemistry. (2003) 56: 749-762.

12. C. A. Yu and I.C. Gunsalus. *Cytochrome $P450_{cam}$.crystalline structure and properties.* The Journal of Biological Chemistry (1974), 249: 94-101.

13. Julian A. Peterson. *Camphor binding Pseudomonas putida cytochrome P450.*Archives of Biochemistry and Biophysics (1971), 144: 678-693.

14. Thomas L. Poulos. *The crystal structure of cytochrome P450$_{cam}$.Cytochrome P-450: Structure, Mechanism and Biochemistry* (1986), Chapter 13.

15. Thomas L. Poulos, Barry C. Finzel, Andrew J. Howard. *High- resolution crystal structure of cytochrome P450$_{cam}$.*Journal of molecular biology (1987), 195: 687-700.

16. Susan Sondej Pochapsky, Thomas C. Pochapsky and Julie W. Wei. *A model for effector activity in a highly specific biological electron transfer complex: the cytochrome P450$_{cam}$-Putidaredoxin couple*. Biochemistry (2003), 42: 5649-5656.

17. Ilme Schlichting, Joel Berendzen, Kelvin Chu, Ann M. Stock, Shelley A. Maves, David E. Benson, Robert M. Sweet, Dagmar Ringe, Gregory A. Petsko, Stephen G. Sligar. *The catalytic pathway of cytochrome P450 cam at atomic resolution*. Science (2000), 287:1615-1622.

18. Claude Denis-Larose, Helene Bergeron, Diane Labbe, Charles W. Greer, Jalal hawari, Matthew J. Grossman, Bruce M. Sankey and Peter C. K. Lau. *Characterization of the basic replicon of Rhodococcus plasmid pSOX and development of a Rhodococcus-Escherichia coli shuttle vector*. (1998) Applied and Environmental Microbiology, 64: 4363-4367.

19. Hideo Shimada, Shingo Nagano, Hiroshi Hori, Yuzuru Ishimura. Putidaredoxin-cytochrome $P450_{cam}$ interaction. (2001) Journal of Inorganic biochemistry, 83: 255-260.

20. Matthew M. Purdy, Laura S. Koo, Paul Oritz de Montellano and Judith P. Klinman. *Steady state kinetic investigation of cytochrome $P450_{cam}$: interaction with redox partners and reaction with molecular oxygen*. (2004) Biochemistry, 43: 271-281.

21. Zhoupeng Zhang, Ole Sibbesen, Roy A. Johnson and Paul Oritz de Montellano. *The substrate specificity of cytochrome $P450_{cam}$.* (1998) Bioorganic and medicinal chemistry, 6: 1501-1508.

22. Ole Sibbesen, Zhoupeng Zhang and Paul Oritz de Montellano. *Cytochrome $P450_{cam}$ substrate specificity: relationship between structure and catalytic oxidation of alkylbenzenes.* (1998) Archives of biochemistry and biophysics, 353: 285-296.

23. Jonathan P. Jones, Ellen J O′ Hare and Luet-Lok Wong. *Oxidation of polychlorinated benzenes by genetically engineered CYP101 (cytochrome $P450_{cam}$).* 2001, European Journal of Biochemistry, 268: 1460-1467.

24. Stephen G. Bell, Julie Anne Stevenson, Helen D. Boyd, Sophie Campbell, Austin D. Riddle, Erica L. Orton and Luet-Lok Wong. *Butane and propane*

oxidation by engineering cytochrome P450$_{cam}$. (2002) Chemical Communications, 5: 490-491.

25. Julie-Anne Stevenson. *Engineering molecular recognition in alkane oxidation catalyzed by cytochrome P450$_{cam}$.*1998, New Journal of Chemistry, 6: 551-552.

26. Darren P. Nickerson et al. *The catalytic activity of cytochrome P450$_{cam}$ towards styrene oxidation is increased by site-specific mutagenesis.*1997, FEBS Letters, 405: 153-156.

27. Julie-Anne Stevenson, Andrew C.G. Westlake, Catherine Whittock and Luet-Lok Wong. *The catalytic oxidation of linear and branched alkanes by cytochrome P450$_{cam}$.* 1996, Journal of American Chemical Society, 118: 12846-12847.

28. Charles F. Harford-Cross, Angus B. Carmicheal, Fiona K.Allan, Paul A. Englad, Duncan A. Rouch and Luet-Lok Wong. *Protein engineering of cytochrome P450$_{cam}$ (CYP101) for the oxidation of polycyclic aromatic hydrocarbons*. (2000) Protein Engineering, 13 : 121-128.

29. Paul A. England, Charles F. Harford Cross, Julie Anne Stevenson, Duncan A. Rouch and Luet-Lok Wong. *The oxidation of naphthalene and pyrene by cytochrome P450$_{cam}$.* (1998) FEBS letters, 424: 271-274.

30. http://www.brunel.ac.uk/depts/bl/project/genome/moltec/vectors/shuttle.htm

31. T.Matsui, K.Hirasawa, J.Konishi, Y.Tanaka, K.Maruhashi and R.Kurane. *Microbial desulfurization of alkylated dibenzothiophene and alkylated benzothiophene by recombinant Rhodococcus sp. strain T09*. (2001) Applied Microbiology and Biotechnology, 56: 196-200.

32. R. Kalscheuer, M. Arenskotter and A. Steinbuchel. *Establishment of a gene transfer system for Rhodococcus opacus PD630 based on electroporation and its application for recombinant biosynthesis of poly (3-hydroxyalkanoic acids)*. (1999) Applied Microbiology and Biotechnology, 52: 508-515.

33. Jan Desomer, Patrick Dhaese and Marc Van Montagu. *Transformation of Rhodococcus fascians by high-voltage electroporation and development of R.fascians cloning vectors*. (1990) Applied and Environmental Microbiology, 56: 2818-2825.

34. M.E. Vogt Singer and W.R. Finnerty. *Construction of an Escherichia coli-Rhodococcus shuttle vector and plasmid transformation in Rhodococcus spp.* (1988) Journal of Bacteriology, 170: 638-645.

35. Julie A. Gegner. *Signal transduction in bacteria: CheW forms a reversible complex with the protein kinase CheA.* (1991) Proccedings at the national academy of science USA, 88: 750-754.

36. http://www.invitrogen.com/content.cfm?pageid=9716

37. Francois Baneyx. *Recombinant protein expression in Escherichia coli.* (1999), Current Opinion in Biotechnology, 10:411-421.

38. Savvas C. Makrides. *Stratergies for achieving high-level expression of genes in Escherichia coli.* (1996) Microbiological Reviews, 512-538.

39. Henry J. Barnes. *Maximizing the expression of eukaryotic cytochrome P450s in Escherichia coli.* (1996) Methods in Enzymology, 272: 3-14.

40. http://www.entelechon.com/eng/seqinv.html

41. C. A. Yu and I. C. Gunsalus. *Cytochrome $P450_{cam}$ II Interconversion with P420.* (1972) The Journal of biological chemistry, 249: 102-106.

42. Julian A. Peterson, Matthew C. Lorence and Bilal Amarneh. Putidaredoxin Reductase and Putidaredoxin. *Cloning, sequence determination and heterologous expression of the proteins.* (1990) The Journal of biological chemistry, 265: 6066-6073.

43. Paul A. England, Charles F. Harford Cross, Julie Anne Stevenson, Duncan A. Rouch and Luet-Lok Wong. *The oxidation of naphthalene and pyrene by cytochrome P450$_{cam}$.* (1998) FEBS letters, 424: 271-274.

VITA

Sirisha Chada was born in Visakhapatnam, India on the 16th of January 1981. She completed her schooling at St.Joseph's Girls' High School, Visakhapatnam in 1996 and her intermediate from B.V.K Junior College in 1998. She had great experiences at both the places. Her life at school has influenced her to become the person that she is today.

Sirisha's career in engineering started in 1998 when she enlisted herself as a student in the Chemical Engineering department at Andhra University, Visakhapatnam. She graduated in June 2002. Her interest in the field of Biochemical Engineering has driven her to pursue her Masters in Chemical Engineering at the University of Tennessee, Knoxville. She will be graduating in August 2004.