Cloning, expression, characterization and engineering of cytochrome P450 CYP116B3 from *Rhodococcus ruber* DSM 44319

Von der Fakultät Geo- und Biowissenschaften der Universität Stuttgart

zur Erlangung der Würde eines Doktors der

Naturwissenschaften (Dr. rer. Nat.) genehmigte Abhandlung

Vorgelegt von

Luo Liu

aus Peking (China)

Hauptberichter: Prof. Dr. Rolf D. Schmid

Mitberichter: Prof. Dr.-Ing. Dr. h.c. Matthias Reuss

Tag der mündlichen Prüfung: 29. August 2007

Institut für Technische Biochemie der Universität Stuttgart

2007

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig und nur unter Angabe der angegebenen Hilfsmittel und Literatur angefertigt habe.

Stuttgart, den

Danksagung

Mein besonderer Dank gilt Prof. Dr. Rolf D. Schmid für die Überlassung des interessanten und herausfordernden Themas, für seine Unterstützung und für sein Interesse am Fortgang der Arbeit.

Herzlich danke möchte ich Dr. Vlada Urlacher für ihre Betreuung und ihre wertvollen Ratschläge.

Der Biokatalysegruppe möchte ich für die großartige und nette Arbeitsatmosphäre und für die freundliche Zusammenarbeit danken.

Alle Mitarbeiter des Instituts für Technische Biochemie danke ich sehr für die angenehme und gemeinsam verbrachte Zeit.

Darüber hinaus danke ich

Dr. Kai Doderer, Dr. Steffen Maurer, Sabine Eiben, Michael Budde, Matthias Dietrich, Katja Kühnel für die interessanten Diskussionen,

Dipl. Biol. Alexander Seifert für die Modellierung des Enzyms P450 CYP116B3,

Lorraine Ling für ihre Hilfe bei der Korrektur der englischen Sprache,

Evelyne Weber für ihre Hilfe bei der Korrektur der deutsche Sprache.

Mein ganz besonderer Dank gilt meiner Mutter und meinem Vater für ihre Verständnis und endlose Unterstützung.

Abbreviations

A	absorbance
APS	ammonium persulfate
bp	base pair(s)
CIAP	calf intersinal alkaline phosphatase
CYP	Cytochrome P450 monooxygenase
CYP116B3	Cytochrome P450 monooxygenase from *Rhodococcus ruber* DSM 44319
DMSO	dimethyl sulfoxide
DNA	deoxyribnucleic acid
dNTP	dexynucleoside-5'-triphosphate
EDTA	ethylendiaminetetraacetic acid
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide, also known as riboflavin
GC	gas chromatography
GC-FID	gas chromatography-flame ionization
GC-MS	gas chromatography-mss spectroscopy
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl-β-D-thiogalactoside
kb	kilobases
kDa	kilo Dalton
KPi	potassium phosphate buffer
LB	Luria-Bertani
M	mole per liter
mg	milligram
min	minute
ml	milliliter
mol	mole
NADH	β-nicotinamide adenine dinucleotide
NADPH	β-nicotinamide adenine dinucleotide phosphate
nm	nanometer

OD	optical density
PAH	polycyclic aromatic hydrocarbon
PCR	polymerase chain reaction
PDB	Protein Data Bank
*p*NCA	*para*-nitrophenoxycarboxylic acid
R_f	retardation factor
rpm	revolutions per minute
RT	retention time
sec	second
SDS	sodium dodecyl sulfate
TAE	Tris/acetate/EDTA buffer
TE	Tris/EDTA buffer
TEMED	N,N,N',N'-tetramethyl-ethylendiamine
THF	tetrahydrofuran
TLC	thin layer chromatography
Tris	Tris-(hydroxymethyl)-aminomethane
UV	ultraviolet
°C	degree Celsius
λ	Wavelength
µg	microgram
µl	microliter

1. Zusammenfassung

Die P450-Enzyme weisen, im Komplex mit Kohlenmonoxid, ein charakteristisches Absorptionsmaximum bei 450 nm auf, welches zur Benennung der Enzyme beiträgt. Sie sind eine der größten Enzymsuperfamilien und gehören zur Klasse der Monooxygenasen (E.C.1.14.-.-.). Die P450-Enzyme kommen bei fast allen Organismen vor: Bakterien, Pilze, Pflanzen, Insekten bis zu den Säugetieren. Sie enthalten eine Hämgruppe im aktiven Zentrum, dessen Eisenatom während des Elektonentransportes seine Wertigkeit reversibel ändern kann. Diese Hämgruppe ist auch für die Farbe der Cytochrome verantwortlich. Die Enzyme können molekularen Sauerstoff binden und aktivieren und dadurch Oxidationen katalysieren, vor allem die für den Organismus wichtigen Hydroxylierungen. P450 Monooxygenasen katalysieren nicht nur Hydroxylierungen, sondern auch Epoxidierungen von Doppelbindungen, sowie Hydroxylierungen und Dealkylierungen von heteroatomischen Molekülen und einige untypische für die P450 Enzyme Reaktionen, wie C-C-Kopplung.

Bakterielle P450-Enzyme besitzen eine relativ hohe Stabilität. Sie sind, mit nur wenigen Ausnahmen, keine Membranproteine und somit wasserlöslich. In dieser Arbeit wurden mehrere bakterielle Stämme zur Untersuchung der Monooxygenasenaktivität getestet. Während der *in vivo* Untersuchung, in den Stämmen *Rhodoccocus ruber* DSM 44319 und *Rhodococcus erythropolis* DSM 43066 konnten oxidative Aktivitäten gegen Cyclohexan nachgewiesen werden. Da die Sequenz der genomische DNA der zwei *Rhodococcus* Stämme noch nicht bekannt ist, wurde ein Sequenzalignment bekannter P450-Gen-Sequenzen aus anderen *Rhodococcus* Stämmen, zur Betrachtung homologer Bereiche, angewendet. Basierend auf den homologen Bereichen des Sequenzalignments von P450RhF aus *Rhodococcus* sp. NCIMB 9784 und P450 CYP116 aus *Rhodococcus* sp. NI86/21 wurden degenerierte Primer entworfen. Mit diesen Primern wurde ein P450-ähnliches DNA-Fragment mit der Größe von 740 Basenpaaren aus *Rhodococcus ruber* DSM 44319 amplifiziert. Die flankierten Bereiche des P450-ähnlichen DNA-Fragmentes wurden mit der „Directional Genome Walking using PCR“ Methode sequenziert. Das vollständige Gen hat eine Größe von 2313 Basenpaaren und kodiert ein Protein mit 771 Aminosäuren. Dieses neue P450-Gen aus *Rhodocuccus ruber* DSM 44319 wurde von der „P450 Nomenclature committee“ als *CYP116B3* benannt und kann als ein neues Mitglied der Klasse IV von P450-

Monooxygenasen betrachtet werden. Die P450-Monooxygenasen der Klasse IV sind natürliche Fusionsprotein und selbstgenügend, d.h. sie brauchen keine zusätzlichen Proteine um Elektronen von NAD(P)H auf Häm zu übertragen. Sie bestehen aus drei verschiedenen Domänen, einer P450-Monooxygenase-Domäne, einer Flavin-enthaltenden Reduktase-Domäne und einer [2Fe2S]-enthaltenden Ferredoxin-Domäne. Die Aminosäurensequenz der Häm-Domäne von CYP116B3 wurde mit CYP116 und der Häm-Domäne von P450RhF verglichen und zeigte typische P450-Merkmale (Abbildung 1).

```
p450RhF         MSASVPASAPACPVDHAALAGGCPVSANAAAFDPFGSAYQTDPAESLRWSRDEEPVFYSP 60
CYP116B3        MSASVPASA--CPVDHAALAGGCPVSTNAAAFDPFGPAYQADPAESLRWSRDEEPVFYSP
cyp116          MTVDHAPEG-------VKSPTGCPVSGMAADFDPFRGAYQVDPSSSLRQARKDEPVFFSP
                *:.. ....       .  . *****  ** ****  ***.**:.*** :*.:****:**

p450RhF         ELGYWVVTRYEDVKAVFRDNILFSPAIALEKITPVSAEATATLARYDYAMARTLVNEDEP 120
CYP116B3        ELGYWVVTRYEDVKAVFRDNLVFSPAIALEKITPVSEEATATLARYDYAMARTLVNEDEP
cyp116          LLDYWVVTRYEDIKQIFKTPSVFSPSITVDQITPISDEALQILGSYQFAAGRMLVNEDEP
                 *.*********:* :*:   :***:*:::***:* **   *. *::* .* *******

p450RhF         AHMPRRRALMDPFTPKELAHHEAMVRRLTREYVDRFVESGKADLVDEMLWEVPLTVALHF 180
CYP116B3        AHMPRRRALMDPFTPKELAHHEAMVRRLTREYVDRFVESGKADLVDEMLWEVPLTVALHF
cyp116          IHTERRRLLMQPFEADNVATLEPKIREVVNTYLDRVIKDGRADLIGDLLYEVPCIVALIF
                 *  *** **:** ..::*  *. :*.:.. *:**.::.*:***:.::*:***  *** *

p450RhF         LGVPEEDMATMRKYSIAHTVNTWGRPAPEEQVAVAEAVGRFWQYAGTVLEKMRQDPSGHG 240
CYP116B3        LGVPEEDMATMRKYSIAHTVNTWGRPAPEEQVAVAEAVGRFWQYAGTVLEKMRQDPSGHG
cyp116          LGVPDEDIETCRQYGMQQTLFTWGHPTGDEQTRVATGMGKFWEFAGGLVDKLKADPNAKG
                ****:**: * *:*.: :*: ***:*: :**. ** .:*:**::** :::*:: **..:*

p450RhF         WMPYGIRKQREMPDVVTDSYLHSMMMAGIVAAHETTANASANAFKLLLENRAVWEEICAD 300
CYP116B3        WMPYGIRMQQQMPDVVTDSYLHSMMMAGIVAAHETTANASANAFKLLLENRPVWEEICAD
cyp116          WIPHAIEMQRQHPDLFDDNYLQNIMFGGVFAAHETTTNATGNAFRTLLENRSSWDEICAD
                *:*:.*. *:: **:. *.**:.:*:.*:.******:**:.***: *****. *:*****

p450RhF         PSLIPNAVEECLRHSGSVAAWRRVATADTRIGDVDIPAGAKLLVVNASANHDERHFERPD 360
CYP116B3        PSLIPNAVEECLRHSGSVAAWRRVATTDTRIGDVDIPAGAKLLVVNASANHDERHFDRPD
cyp116          PTLIPKAIEECLRYSGSVVAWRRKAVVDTTVGEVDIPAGGRLLIVMASANRDDSMFPEPD
                *:***:*:*****:****.**** *..** :*:******.:**:* ****:*:  * .**

p450RhF         EFDIRRPNSSDHLTFGYGSHQCMGKNLARMEMQIFLEELTTRLPHMELVPDQEFTYLPNT 420
CYP116B3        EFDIRRPNSSDHLTFGYGSHQCMGKNLARMEMQIFLEELTTRLPHMELVPDQEFTYLPNT
cyp116          DFDIHRGNAQRHLTFGIGSHTCLGATLARLEMKVFLEEVSRRLPHMSLVAGQEFSYLPNT
                :***:* *:. ***** *** *:* .***:**::****:: *****.**..***:*****

p450RhF         SFRGPDHVWVQWDPQANPERTDPAVLHRHQPVTIGEPAARAVSRTVTVERLDRIADDVLR 480
CYP116B3        SFRGPDHVWVQWDPQANPERTDPAVLQRQHPVTIGEPSTRSVSRTVTVERLDRIVDDVLR
cyp116          SFRGPEHVLVEWDPQQNP-----------------VPADRP------------------
                *****:** *:**** **                   *: *.
```

Abbildung 1: Vergleich der Aminosäurensequenzen von CYP116, der Häm-Domäne von CYP116B3 aus *Rhodococcus ruber* und P450RhF, erstellt durch ClustalX. Mit (*) gekennzeichnete Aminosäuren sind identisch, (:) kennzeichnen starke Ähnlichkeit, (.) kennzeichnen schwache Ähnlichkeit. Der konservierte Sauerstoff-Bindungs-Bereich wird in gelb hervorgehoben. Der konservierte Cystein-Rest ist in grün dargestellt.

Die konservierten Bereiche in der Reduktase-Domäne von CYP116B3 konnten durch Sequenzalignment der Reduktase- und Ferredoxin-Domäne zwischen der P450 CYP116B3, P450RhF und Phthalate Dioxygenase Reduktase aus *Pseudomonas cepacia* identifiziert werden (Abbildung 2). Das Alignment deutet an, dass die Reduktase-Domäne von CYP116B3 ein FMN enthält; und gilt als ein Beweis für ihre Reduktase-Funktion.

```
CYP116B3        SFRGPDHVWVQWDPQANPERTDPAVLQRQHPVTIGEPSTRSVSRTVTVERLDRIVDDVLR 480
p450RhF         SFRGPDHVWVQWDPQANPERTDPAVLHRHQPVTIGEPAARAVSRTVTVERLDRIADDVLR
PDR             -----------------------------------TTPQEDGFLRLKIASKEKIARDIWS
                                                    .. .     :.:   ::*. *:

CYP116B3        VVLRAPAGNALPAWTPGAHIDVDLG-ALSRQYSLCG-APDAPTYEIAVLLDPESRGGSRY 540
p450RhF         LVLRDAGGKTLPTWTPGAHIDLDLG-ALSRQYSLCG-APDAPSYEIAVHLDPESRGGSRY
PDR             FELTDPQGAPLPPFEAGANLTVAVPNGSRRTYSLCNDSQERNRYVIAVKRDSNGRGGSIS
                . *  . * .**.: .**:: : :  .  * ****. : :   * ***  *.:.****

CYP116B3        VHEQLRVGGSLRIRGPRNHFALDPDAEHYVFVAGGIGITPVLAMADHARARG-WSYELHY 600
p450RhF         IHEQLEVGSPLRMRGPRNHFALDPGAEHYVFVAGGIGITPVLAMADHARARG-WSYELHY
PDR             FIDDTSEGDAVEVSLPRNEFPLDKRAKSFILVAGGIGITPMLSMARQLRAEGLRSFRLYY
                . ::   *..:.:  ***.*.**  *: :::*********:*:** : **.*  *:.*:*

CYP116B3        CGRNRSGMAYLERVAGHGDRAALHVSAEG----TRVDLAALLATPVSGTQIYACGPGRLL 660
p450RhF         CGRNRSGMAYLERVAGHGDRAALHVSEEG----TRIDLAALLAEPAPGVQIYACGPGRLL
PDR             LTRDPEGTAFFDELTSDEWRSDVKIHHDHGDPTKAFDFWSVFEKSKPAQHVYCCGPQALM
                  *: .* *:::.::..  *: :::  :     . .*: :::  . .. ::*.***  *:

CYP116B3        AGLEDASRHWPDGALHVEHFTSSLTALDPDVEHAFDLDLRDSGLTVRVEPTQTVLDALRA 720
p450RhF         AGLEDASRNWPDGALHVEHFTSSLAALDPDVEHAFDLELRDSGLTVRVEPTQTVLDALRA
PDR             DTVRDMTGHWPSGTVHFESFG--ATNTNARENTPFTVRLSRSGTSFEIPANRSILEVLRD
                  :.* : :**.*::*.* *    :  :.  : .* : *  ** :..: ..:::*:.**

CYP116B3        NNIDVPSDCEEGLCGSCEVTVLEGEVDHRDTVLTKAERAANRQMMTCCSRACGDRLTLRL 780
p450RhF         NNIDVPSDCEEGLCGSCEVAVLDGEVDHRDTVLTKAERAANRQMMTCCSRACGDRLALRL
PDR             ANVRVPSSCESGTCGSCKTALCSGEADHRDMVLRDDEKGT--QIMVCVSRAKSAELVLDL
                 *: ***.**.* ****:.:: .**.**** ** . *:.:  *:*.* *** . .*.* *
```

Abbildung 2: Das Alignment zeigt die Reduktase-Domäne von CYP116B3 aus *Rhodococcus ruber* (Aminosäuren 419 bis 771), sowie von P450RhF (Aminosäuren 421 bis 773). PDR bezieht sich auf die Phthalate Dioxygenase Reduktase aus *Pseudomonas cepacia.* Mit (*) gekennzeichnete Aminosäuren sind identisch, (:) kennzeichnen starke Ähnlichkeit, (.) kennzeichnen schwache Ähnlichkeit. Der FMN-Bindungsbereich ist in blau hervorgehoben. Der Bindungsbereich von NAD(P)H in gelb. Die vier Cystein-Reste, die an dem Eisen-Schwefel-Cluster binden, sind in grün dargestellt.

Die Cytochrom P450-Monooxygenase CYP116B3 wurde erfolgreich in den Vektor pET28a(+) einkloniert und in *Escherichia coli* BL21(DE3) exprimiert. Anschließend wurde das Protein mittels IMAC (immobilized metal affinity chromatography) aufgereinigt. Die

gesamte Ausbeute des Proteins CYP116B3 betrug 18,6 mg l^{-1} mit einer Reinheit von 85%. Nur FMN wurde über Dünnschichtchromatographie in der Reduktase-Domäne nachgewiesen. Das Absorptionsspektrum von oxidiertem Protein weist ein Maximum bei 418 nm auf, welches typisch ist für Cytochrom P450 Enzyme (Abbildung 3). Die breite Absorptionsschulter zwischen 450 und 510 nm weist auf die Anwesenheit von Flavin hin. Die Zugabe von NADPH führte zu einer Senkung dieser Schulter, hervorgerufen durch die Reduktion von oxidiertem Flavin. Die Zugabe von CO zu dem Dithionit-reduziertem Protein führte zu einer Verschiebung des Absorptionsmaximums von 418 nm zu 449 nm durch die Bildung von Fe^{2+}-CO-Komplexen (Abbildung 4).

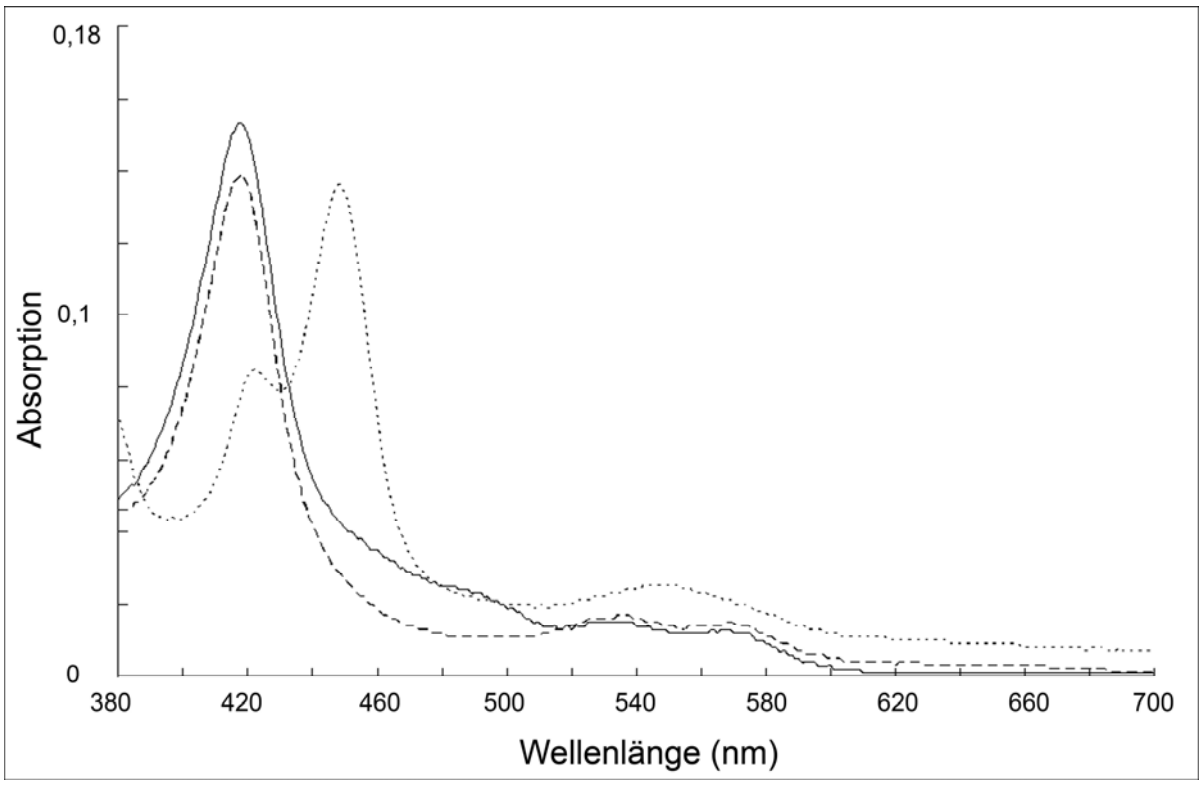

Abbildung 3: Dargestellt sind die Spektraleigenschaften des Cytochroms P450 CYP116B3 aus *Rhodococcus ruber* DSM 44319. Die durchgezogene Linie zeigt das Spektrum des oxidierten Enzyms, die gestrichelte Linie die des NADPH-reduziertem Enzyms. Die Punktlinie gibt das Spektrum des reduzierten/CO-gebundenem Enzyms wieder.

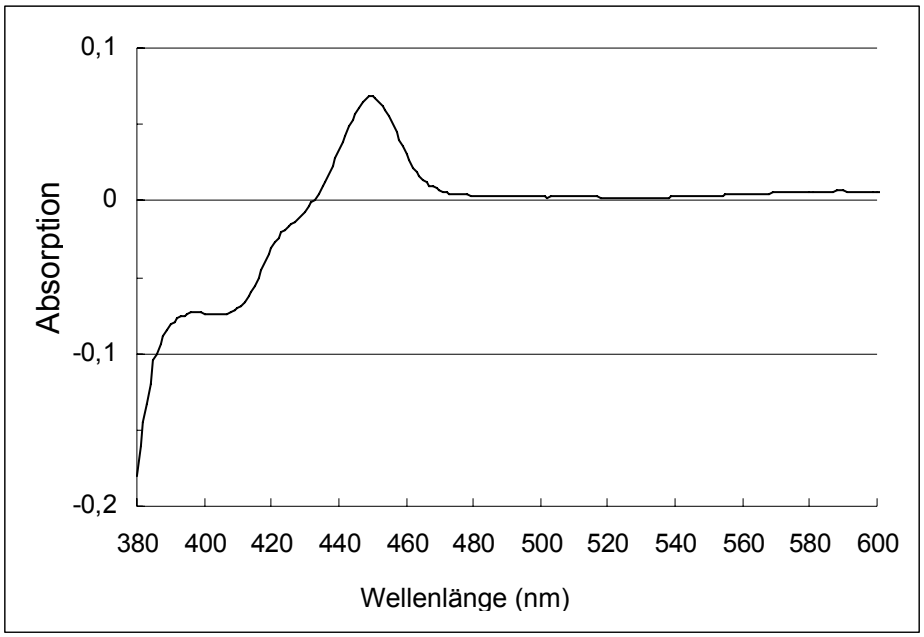

Abbildung 4: CO-Differenzspektrum von P450 CYP116B3 aus *Rhodococcus ruber* DSM 44319.

Die Aktivität der Reduktase wurde mit dem exogenen Elektronenakzeptor Cytochrom *c* bestimmt. Bei der NADPH-abhängigen Reaktion beträgt k_{cat} = 765 ± 144 und K_M = 3.4 ± 0.8. Bei der NADH-abhängigen Reaktion beträgt k_{cat} = 359.8 ± 11.5 and K_M = 126 ± 11. Aufgrund der ermittelten Werte ist ersichtlich, dass die Reduktase-Domäne NADPH vor NADH bevorzugt. So wie P450RhF, katalysiert P450 CYP116B3 auch die *O*-Dealkylierung von 7-Ethoxycumarin und produziert 7-Hydroxycumarin. In der Anwesenheit von NADPH, zeigt P450 CYP116B3 außerdem Hydroxylierungsaktivität gegen aromatische Kohlenwasserstoffe, wie Naphthalin, Inden, Acenaphthen, Toluol, Fluoren, und Ethylbenzol. Naphthalin, Acenaphthen und Fluoren werden an den aromatischen Ringen hydroxyliert, während Alkylbenzole, wie Toluol, *m*-Xylol und Ethylbenzol ausschließlich an den Seitenketten hydroxyliert werden.

Tabelle 1: Aufgelistet sind die eingesetzten Substrate, die synthetisierten Produkte von P450 CYP116B3 aus *Rhodococcus ruber* DSM 44319 und deren Produktbildungsrate.

Substrate	Produkte	Produktbildungsrate (nmol x nmol P450 $^{-1}$ x min^{-1})
7-Ethoxycumarin	7-Hydroxycumarin	0,917 ± 0,05
Acenaphthen	1-Acenaphthenol	0,079 ± 0,01
Fluoren	9-Fluorenol	0,04 ± 0,005
Naphthalin	1-Naphthol	0,106 ± 0,01
Inden	1-Indenol	n. d.
Toluol	Benzyl-alcohol	0,301 ± 0,01
Ethylbenzol	1-Phenylethyl-alcohol	n. d.
	2-Phenylethyl-alcohol	n. d.
m-Xylol	3-Methylbenzyl-alcohol	n. d.

n. d. – nicht bestimmt

Um die Struktur-Funktions-Beziehung zu untersuchen, wurde eine Struktur durch homologische Modellierung erstellt. Nach der Analyse des Enzym-Models wurde der Alanin-Rest an Position 109 gegen Phenylalanin ausgetauscht. Die Position 109 befindet sich über der Häm-Gruppe. Der Austausch des Alanin-Restes an Position 109 gegen Phenylalanin verkleinert den Substrat-Kanal, was die Bindung und Oxidation kleiner Kohlenwasserstoffe, wie Cyclohexan oder Alpha-Pinen, begünstigen sollte. Jedoch wurde nur eine Erhöhung der Aktivität gegen 7-Ethoxycumarin und PAHs von dieser Mutante beobachtet.

Mithilfe gerichteter Evolution können Optimierungen der Enzymfunktionen, ohne strukturelle Kenntnisse, vorgenommen werden. Um das Enzym P450 CYP116B3 zu verbessern wurde gerichtete Evolution, eingeschränkt auf die Monooxygenase-Domäne, in dieser Arbeit durchgeführt. Eine Mutationsbibliothek wurde durch Error-Prone PCR erstellt und durch ein entwickeltes Hochdurchsatzselektions-System auf erhöhte Aktivität untersucht. Wegen hoher fluoreszenter Empfindlichkeit von 7-Hydroxycumarin, das Produkt der *O*-Dealkylierung von 7-Ethoxycumarin, wurde 7-Ethoxycumarin als Test-Substrat eingesetzt. Die Messung wurde mit ganzen *E. coli* Zellen in Mikrotiter-Platten mit 96 Vertiefungen durchgeführt. Der Vorteil der Umsetzung mit ganzen Zellen ist, dass der Aufwand von

Zellaufschluss und die hohen Kosten des Co-Faktors NADPH vermieden werden konnten. Insgesamt wurden vier Mutanten-Generationen nach der Error-Prone PCR durchgemustert. Schließlich konnten die zwei besten Mutanten, 70A08 (A86T/T91S/A109F/I179F/I267L) und 74H10 (T91S/A109L/I179F/I267L), aus der vierten Error-Prone PCR Bibliothek, mit 100-fach erhöhter Aktivität gegen 7-Ethoxycumarin, identifiziert werden (siehe Tabelle 2).

Tabelle 2: Zusammenfassung der besten Mutanten nach vier Runden der Error-Prone PCR.

Runde von Error-Prone PCR	Variante	Nucleotid-Ersetzung	Aminosäure-Ersetzung	Produktbildungsrate (nmol x nmol P450 $^{-1}$ x min^{-1})
1	07F08	799 A → C	I267L	7
2	30A06	256 G → A	A86T	30
3	44G08	535 C → T	L179F	100
4	70A08	271 A → T	T91S	223
	74H10	325 T → C	A109L	141

2. Summary

The Cytochrome P450 monooxygenases are ubiquitous heme-containing proteins, which catalyze regio- and stereo-selective oxidations of non-activated hydrocarbon. Bacterial P450 enzymes are in most cases not membrane-associated, water soluble and exhibit relatively high stability. Several bacterial strains have been tested for their monooxygenase activity. During the *in vivo* screening, *Rhodoccocus ruber* DSM 44319 and *Rhodococcus erythropolis* DSM 43066 strains exhibited oxidation activity towards cyclohexane. Because the genome DNA sequence of the two *Rhodococcus* strains has not been sequenced, a homology search was applied using known P450 gene sequences from other *Rhodococcus* strains. A set of degenerate primers was designed to the most similar regions, identified through the DNA sequence alignment of P450RhF from *Rhodococcus* sp. NCIMB 9784 and P450 CYP116 from *Rhodococcus* sp. NI86/21. A P450-like gene fragment with 740 base pairs was amplified from *Rhodococcus ruber* DSM 44319 by PCR using these degenerate primers. The flanking regions of the P450-like DNA fragment were explored by directional genome walking using PCR combined TA-cloning. The entire P450 gene with 2313 bp was isolated. It encodes a protein of 771 amino acids. The primary protein structure suggests that it is a natural self-sufficient fusion protein consisting of a P450 monooxygenase domain, a flavin-containing reductase domain, and a [2Fe2S] ferredoxin domain. This new P450 gene from *Rhodococcus ruber* DSM 44319 was named by P450 nomenclature committee *CYP116B3* and can be considered as a new member of class IV of cytochrome P450 monooxygenases.

The cytochrome P450 monooxygenase CYP116B3 was successfully cloned into vector pET28a(+), and expressed in *Escherichia coli* BL21(DE3). Subsequently, the enzyme was purified using immobilized metal affinity chromatography with a total yield of 18.6 mg l^{-1} and purity of 85%. Thin layer chromatography detected only FMN within the reductase domain, as the sequence alignment had predicted. The reductase activity was determined using an exogenous electron acceptor cytochrome *c*. The reductase domain of this P450 CYP116B3 demonstrated a strong preference for NADPH over NADH. As well as P450RhF, P450 CYP116B3 catalyzed *O*-dealkylation of 7-ethoxycoumarin gave product 7-hydroxycoumarin. Furthermore, in the presence of NADPH, the P450 CYP116B3 demonstrated hydroxylation activity towards aromatic hydrocarbons, naphthalene, indene,

acenaphthene, toluene, fluorene, *m*-xylene and ethyl benzene. The conversion of naphthalene, acenaphthene and fluorene resulted in respective ring monohydroxylated metabolites. Alkyl aromatics like toluene, *m*-xylene and ethyl benzene were hydroxylated exclusively at the side chains.

The highest turnover rate catalyzed by wild-type P450 CYP116B3 is about 1 nmol of 7-ethoxyxoumarin converted per 1 nmol of P450 CYP116B3 per minute. Compare to P450 BM-3, the activity of P450 CYP116B3 is very low. In order to investigate the structure-function relationship, a structure model of P450 CYP116B3 was designed based on the known homologous structures. The position 109 was identified which is located on the ceiling of the substrate binding pocket. The substitution of alanine residue at position 109 by phenylalanine may promote binding and oxidation of smaller alkyl substrates, such as cyclohexane or alpha-pinene. However after the substitution no oxidation activity towards smaller substrates was detected. Only an increase of activity towards 7-ethoxycoumarin and PAHs was observed.

Directed evolution provides a useful tool to explore enzyme functions without structural information. Therefore, directed evolution was carried out in this study to improve the enzyme activity. The mutagenesis was limited to the monooxygenase domain of CYP116B3. A mutation library was generated by error-prone PCR. A high throughput screening system was developed based on the *O*-dealkylation of 7-ethoxycoumarin using whole *E. coli* cells, which expressed the enzyme in 96-well microtiter plates. Finally, the two best mutants, 70A08 (A86T/T91S/A109F/I179F/I267L) and 74H10 (T91S/A109L/I179F/I267L) were identified after four error-prone PCR rounds with 100-fold increased *O*-dealkylation activity towards 7-ethoxycoumarin. To investigate the alteration of the substrate spectrum of the two mutants, a wide rang of potential substrates was tested, including known substrates of the wild-type CYP116B3. However, 70A08 and 74H10 did not show increased activity towards any substrate besides 7-ethoxycoumarin.

3. Introduction

3.1. Cytochrome P450 monooxygenases

Cytochrome P450 monooxygenases (also called cytochrome P450 mixed-function monooxygenase or mixed-function oxidase, because they function as both oxygenase and oxidase) belong to enzyme class monooxygenases (E.C.1.14.-.-.). They are ubiquitous heme-containing proteins found in almost all lineages of life including prokaryotes (archaea, bacteria), lower eukaryotes (fungi, insects) and higher eukaryotes (plants and animals). Plants have the highest number of P450 genes, while *Escherichia coli* has none. Although the amino acid sequence identity of P450s is around 16-20%, their structural fold is conserved throughout evolution. In the active site, the heme with a cysteine residue as fifth ligand to the heme iron is a feature that contributes to the unusual spectral and catalytic properties of P450 monooxygenases. Monooxygenases require two electrons to reduce the heme iron and the second oxygen atom of dioxygen to water. They usually use NADH or NADPH as an electron donor to activate molecular oxygen (Werck-Reichhart and Feyereisen 2000). Catalyzing a wide range of oxidative reactions, P450 monooxygenases have critical and specific roles in the metabolism of exogenous and endogenous compounds. They are essential for utilization of carbon compounds as energy sources, detoxification of xenobiotics, biosynthesis of hormones and production of important natural products such as antibiotics.

3.1.1. Cytochrome P450 monooxygenases superfamily

The cytochrome P450 monooxygenase was first discovered in 1955 in rat liver microsomes (Klingenberg 1958), and further characterized by Sato and Omura (Omura and Sato 1964a; Omura and Sato 1964b). The name P450 refers to “pigment at 450 nm”, so named for the characteristic absorption maximum (Soret peak) formed by absorbance at a wavelength near 450 nm when the heme iron is reduced (e.g. with sodium dithionite) and complexed to carbon monoxide. This property is also used to develop a quantitative method for determination of P450 monooxygenases (Omura and Sato 1964a; Omura and Sato 1964b).

Currently, more than 6000 cytochrome P450 sequences are known and officially named (July 2006, http://drnelson.utmem.edu/CytochromeP450.html). P450s constitute a superfamily of heme-thiolate proteins existing in prokaryotes and eukaryotes. P450 superfamily genes are subdivided and classified according to the recommendation of a nomenclature committee (Nelson, Koymans et al. 1996) on the basis of amino acid identity, phylogenetic criteria and gene organization. Genes encoding for the P450 enzymes, and the enzymes themselves, are designated with the root symbol CYP, followed by an arabic number for the gene family (with more than 40% amino acid identity), a capital letter indicating the subfamily (with more than 55% amino acid identity), and another number for the individual gene. The convention is to italicize when referring to the gene. For example, *CYP101A1* is the gene that encodes for the enzyme CYP101A1 - the well known P450cam, which catalyzes hydroxylation of camphor to 5-*exo*-hydroxy-camphor.

3.1.2. Structural features

Since the first crystal structure of P450cam (CYP101A1) from *Pseudomonas putida* was determined in 1985 (Poulos, Finzel et al. 1985; Poulos, Finzel et al. 1987), the number of crystal structures is increasing rapidly. Sequence identity of P450 proteins is often extremely low. However, comparison of the crystal structures of P450 proteins shows a high conservation of their general topography and structural fold (Graham and Peterson 1999).

In general, P450 proteins contain approximately thirteen α helices and four β sheets (Figure 1.1) (Peterson and Graham 1998). The highest conserved structure is found in the core of the protein around the heme and reflects a common mechanism of electron and proton transfer and oxygen activation. The conserved core consists of a four-helix (D, E, I and L) bundle, helices J and K, two sets of β sheets, and a coil called the ‘meander’. Helix I contains a highly conserved threonine residue with an acidic residue which corresponds to the proton transfer groove on the distal side of the heme. Helix K contains the absolutely conserved Glu-X-X-Arg motif which may be needed to stabilize the core structure, located on the proximal side of the heme (the putative redox partner binding site). Helix L forms part of the heme-binding region. The heme-binding loop on the proximal face of the heme contains the most characteristic P450 consensus sequence Phe-X-X-Gly-X-Arg-Cys-X-Gly with the absolutely conserved cysteine residue. This cysteine residue serves as the fifth ligand to the heme iron

and is the reason for the characteristic absorbance of Soret band at 450 nm when heme is reduced and carbon monoxide-bound. The two sets of β sheets 1 and 2 are structurally conserved and may contribute to form the hydrophobic substrate access channel.

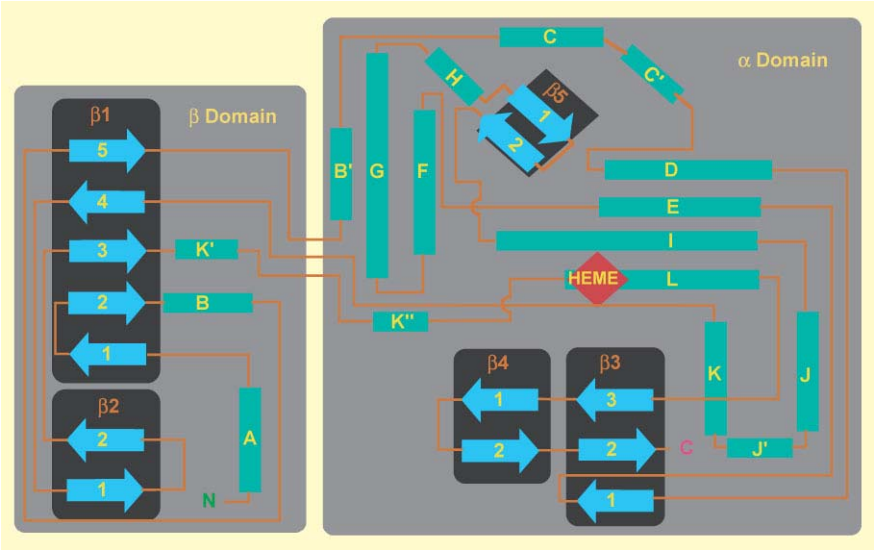

Figure 1.1: Topology drawing shows the secondary structure elements of the P450 proteins. Green rectangles represent helices, blue arrows represent β sheets, and orange lines represent the connection of the helices and sheets (Peterson and Graham 1998).

The most variable regions are usually associated with amino-terminal anchoring or redox partner binding, for substrate binding and recognition. Substrate recognition sites (SRS) were identified by Gotoh et al. (Gotoh 1992) on the basis of sequence alignments between members of the CYP2 family and P450cam. The variable regions associated with substrate binding sites are helices A, B and B', F and G, and their adjacent loops. Most eukaryotic P450s contain a hydrophobic amino-terminal membrane anchoring segment and are associated with microsomal membranes. An additional membrane association region may be located between F and G helices, which show high hydrophobicity.

3.1.3. Reactions catalyzed by P450s

P450 monooxygenases belong to Phase I enzymes, which convert insoluble (hydrophobic) compounds into more polar substances for further conjugation reactions that are catalyzed by

Phase II enzymes, for example, N-acetyltransferase or sulfotransferase. P450s catalyze a wide range of oxidation reactions, including aliphatic and aromatic hydroxylation, epoxidation of double bonds, heteroatom oxidation, heteroatom dealkylation and multiple oxidation including carbon-carbon bond cleavage, even reduction (Mansuy 1994; Sono, Roach et al. 1996; Cryle 2003).

Hydroxylation of non-activated carbon atoms is the most common reaction catalyzed by P450s. The function of such hydroxylases can be divided into two categories: biodegradative transformations that often initiate a cascade of reactions allowing an organism to use unusual carbon as its sole energy or carbon source; biosynthetic reactions essential to the production of complex primary or secondary metabolites. Many biodegradative P450s have been identified that allow microorganisms to use unusual chemicals in environment as energy source, for example, P450cam (Tyson, Lipscomb et al. 1972) and P450terp (CYP108) (Boddupalli, Hasemann et al. 1992) from different *Pseudomonas* sp. Aliphatic hydroxylation by P450 monooxygenases are also important in the biosynthesis of many complex compounds, especially antibiotics. P450pick (CYP107L1) from *Streptomyces venzuelae* is responsible for formacrolide hydroxylation. Aliphatic hydroxylation, As well as aromatic hydroxylation catalyzed by P450s takes place in both biosynthetic and biodegredative pathways (Cryle 2003).

Epoxidation of double bond is another major reaction type catalyzed by P450s. Epoxides are reactive functionalities and relevant to the biological activity of the natural products that contain them. This reactivity makes epoxides useful synthetic intermediates for products of pharmaceutical or agrochemical interest (Martinez and Stewart 2000). Stereoselective epoxidation of styrene-type substrates has been investigated for P450cam (Nickerson, Harford-Cross et al. 1997), P450terp (Fruetel, Mackman et al. 1994), and P450 BM-3 (Fruetel, Mackman et al. 1994), whereas P450terp shows an e.e. of 82% (Fruetel, Mackman et al. 1994).

P450 enzymes catalyze also heteroatom (N, S, O) oxidation and dealkylation leading to hydroxylation of the α-carbon or to formation of the corresponding oxide in the case of N or S. Oxidation of the α-carbon results in heteroatom dealkylation and formation of an aldehyde. P450SUI and P450SU2 from *Streptomyces griseolus* showed N-dealkylation and O-demethylation activity towards various sulfonylurea herbicides (O'Keefe, Lau et al. 1993).

Other reaction types, due to less relevance to this study, will be not described here. Dawson and coworkers (Sono, Roach et al. 1996) summarized more than 20 different reactions, which can be catalyzed by P450 monooxygenases (Figure 1.2). Some more uncommon reactions were recently summarized by Guengerich (Guengerich 2001b).

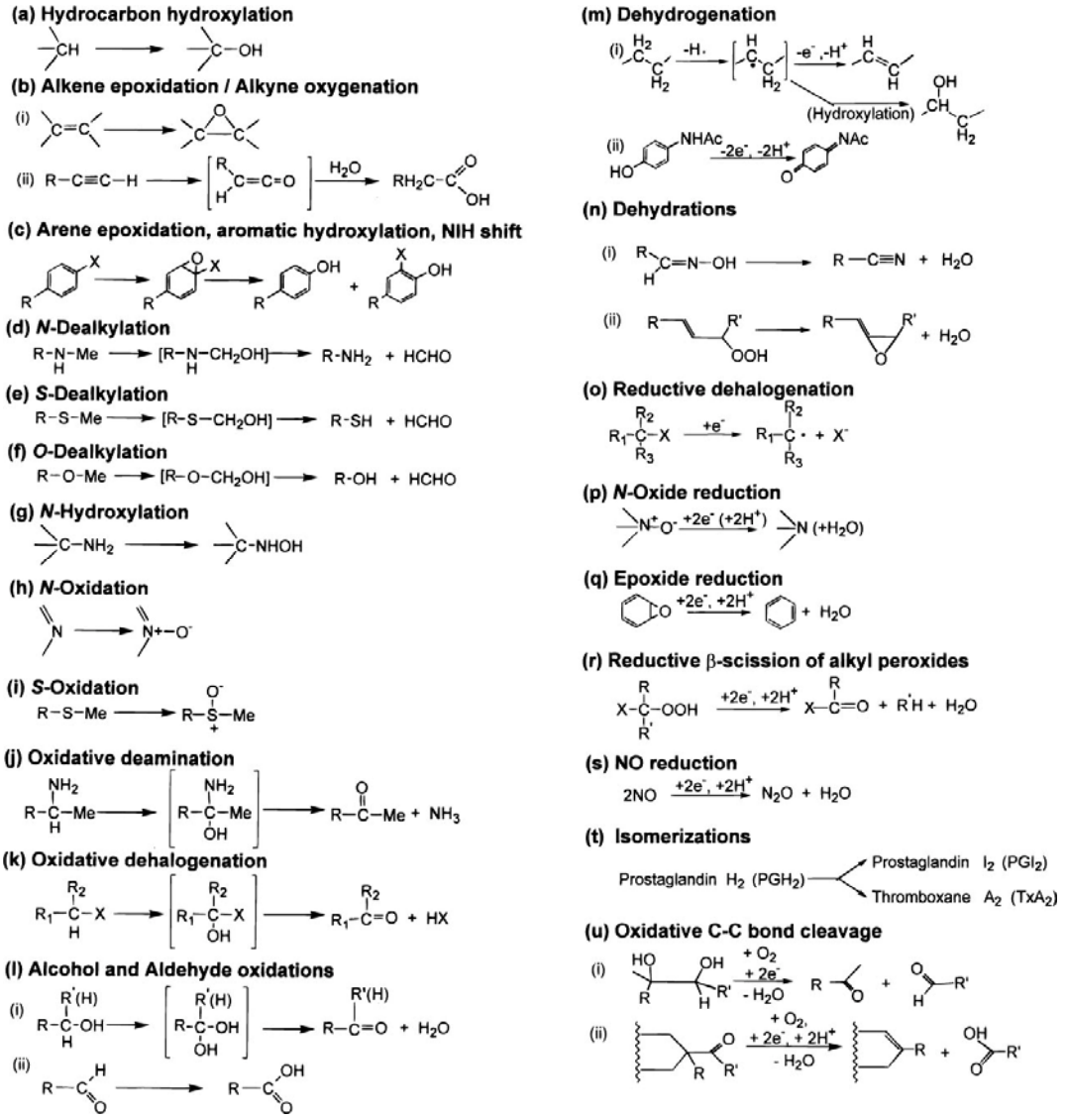

Figure 1.2: Summary of common reactions catalyzed by P450s, according to (Sono, Roach et al. 1996).

3.1.4. Functions of P450 monooxygenases in biological systems

Prokaryotic P450s are soluble proteins. P450s enable prokaryotes to catabolyze compounds used as energy or carbon sources or to detoxify xenobiotics. Other functions described for

prokaryotic P450s include fatty acids metabolism and biosynthesis of antibiotics. Eukaryotic class I (found in mitochondria) P450s catalyze several steps in the biosynthesis of steroid hormones and vitamin D3 in mammals. Class II P450s are the most common P450s in eukaryotes. Eukaryotic P450s and NADPH-P450 reductases are anchored on the outer face of the endoplasmic reticulum by amino-terminal hydrophobic anchors. Functions of class II P450s are extremely diverse. In fungi, they synthesize membrane sterols and mycotoxins, detoxification of phytoalexins, and metabolism of lipid carbon sources. Plant cytochrome P450s are involved in a wide range of biosynthesis or catabolism of lignin components, UV protectants, flower pigments, plant hormones, defensive compounds, and oxigenation of fatty acids for cutins and suberins (Schuler and Werck-Reichhart 2003). In insects, P450s are involved in the synthesis and catabolism of ecdysteroids and juvenile hormones and detoxification of foreign chemicals (Feyereisen 1999). In animals, physiological functions include biosynthesis and catabolism of signaling molecules, steroid hormones, retinoic acid and oxylipins (Hasler 1999). P450s are responsible for the removal or sometimes activation of exogenous compounds. Furthermore, P450s are regulated in the metabolism of steroid hormones and are often inducible by exogenous chemicals (Honkakoski and Negishi 2000).

3.1.5. Catalytic mechanism

In general, the reaction catalyzed by P450 monooxygenases involves substrate binding, reduction of the ferric heme, binding of molecular oxygen, activation of oxygen, insertion of an iron-bound oxygen into the substrate, and product release (Schlichting, Berendzen et al. 2000).

P450 monooxygenases are able to catalyze regiospecific and stereospecific hydroxylation of non-activated hydrocarbons at physiological temperatures. Such a reaction, uncatalyzed, requires extremely high temperature and even nonspecific. The mechanism by which P450s activate molecular oxygen and carry out all those reactions have been partially understood and documented. The active site of P450 contains an iron protoporphyrin IX (heme). The fifth ligand is a thiolate provided by the conserved cysteine residue. Resting P450 is in the low-spin ferric state and the sixth coordination position may be occupied by a water molecule.

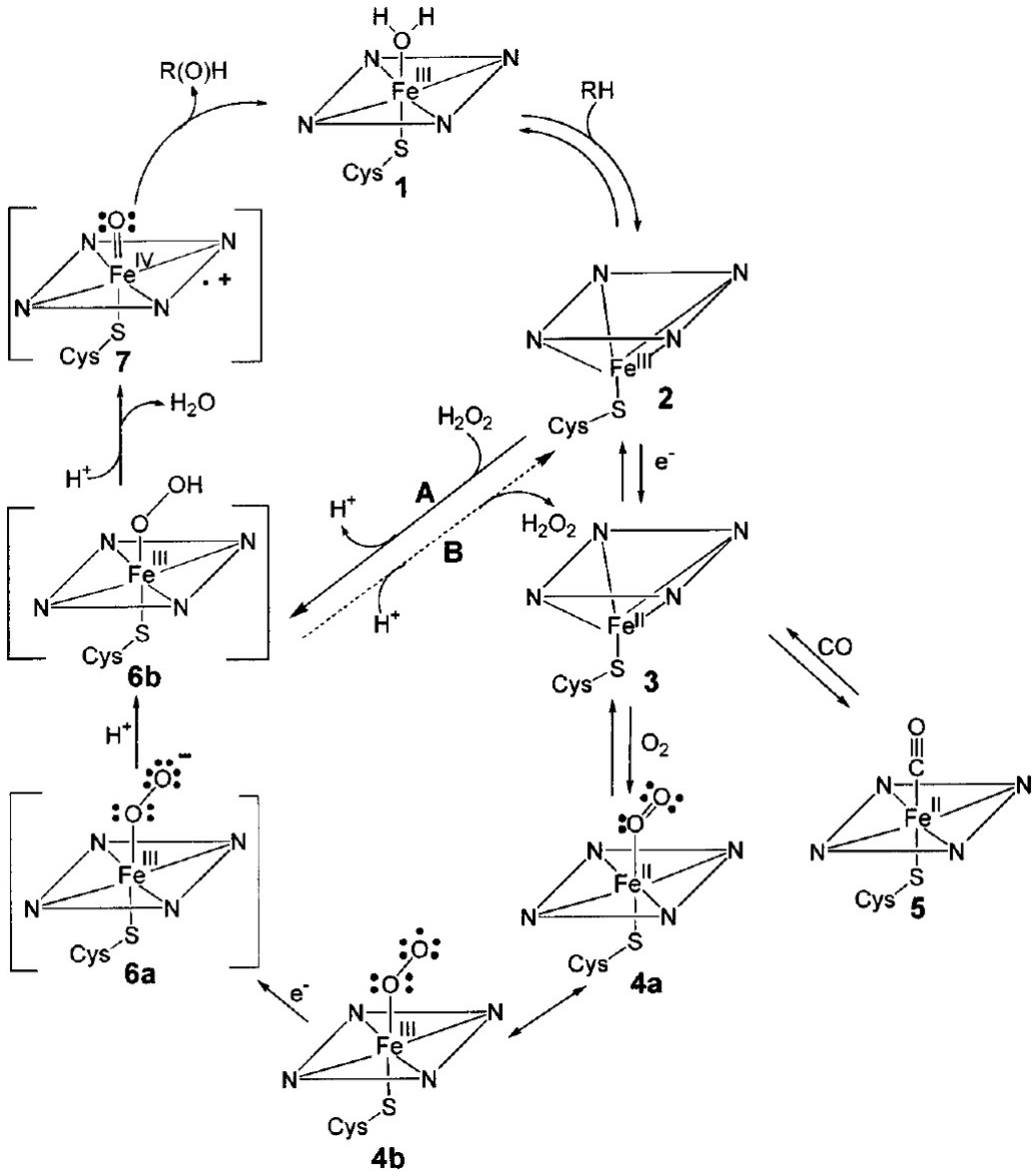

Figure 1.3: Catalytic cycles of P450 monooxygenase. Modified from (Sono, Roach et al. 1996).

Figure 1.3 shows the common enzyme mechanism of P450 monooxygenases. The first step in the reaction cycle is substrate binding. Displacement of the sixth ligand water molecule by substrate induces a low- to high-spin shift of the heme iron and an increase of reduction potential of the enzyme heme (from -300 to -173 mV), which facilitates uptake of the first electron from redox partner to the ferric P450 heme. Whereas the five-coordinate iron atom is out of the porphyrin plane. The second step is a one electron reduction of the complex to a high-spin ferrous state. In the third step, molecular oxygen is bound to the ferrous P450, to generate a superoxide complex with in of the prophyrin plane coordination (Figure 1.3, State

4a). The fourth step, a second reduction is the rate limit step in the reaction cycle. Addition of the second electron to the superoxide complex yields a ferric peroxide adduct (Figure 1.3, State 6a), which can be protonated to give a hydroperoxide complex (Figure 1.3, State 6b). Protonation of the unprotonated oxygen releases hydrogen peroxide and generates ferric state without oxygenated product (Figure 1.3, Pathway A). NAD(P)H and dioxygen is consumed in a 1:1 ratio for this uncoupling reation. After second protonation of the same oxygen the oxygen-oxygen bound is split to release a water molecule and to generate the ferryl-oxene (O = Fe^{IV}) activated oxygen intermediate (Figure 1.3, State 7). The active ferryl-oxene species inserts an oxygen atom into substrate by the oxygen rebound mechanism suggested by Groves et al. (Groves 1985). The monooxygenated product can be released from the enzyme, regenerating the initial ferric state of P450. Carbon monoxide can bind ferrous P450 (Figure 1.3, State 3) to form ferrous-CO complex (Figure 1.3, State 5) which is responsible for the unusual spectroscopic properties of P450. The rebound mechanism is generally accepted and shown in figure 1.4. A hydrogen is abstracted from the hydrocarbon (RH) by the active ferryl-oxene species (Figure 1.4, 1, $Por^{+}Fe^{IV}$-O), followed by hydroxyl radical recombination ("oxygen rebound") (Figure 1.4, 2), to generate the hydroxylated hydrocarbon and ferric complex (Figure 1.4, 3), then the hydroxylated hydrocarbon is released (Ogliaro, Harris et al. 2000).

Numerous peroxagenated compounds may substitute NADH or NADPH as well as O_2 which are required for the normal catalytic cycle and react directly with ferric state via a "peroxide shunt" (Figure 1.3, Pathway B) (Nordblom, White et al. 1976) path to generate hydroperoxide complex (Figure 1.3, 6b). Some P450s may carry out a reaction using hydrogen peroxide as the source of oxygen (Wolfe and Lipscomb 2003), or can be engineered to be able to use hydrogen peroxide (Joo, Lin et al. 1999).

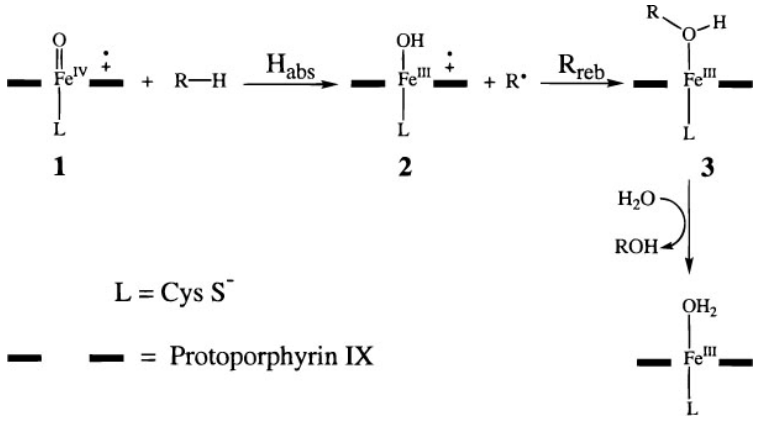

Figure 1.4. Rebound mechanism of hydroxylation catalyzed by P450 monooxygenases (Ogliaro, Harris et al. 2000).

3.1.6. Redox systems for cytochrome P450 monooxygenases

All known P450s are multi-component enzymes consisting of a P450 component with associated reductase components, also called P450-containing monooxygenase systems. As mentioned before, P450s catalyze incorporation of only one atom of molecular dioxygen into substrate while reducing the second to water with the following stochiometry:

$$RH + O_2 + NAD(P)H + H^+ \rightarrow ROH + H_2O + NAD(P)^+$$

In this equation, RH and ROH are the substrate and product, respectively. This reaction requires two electrons to activate an oxygen molecule. The two electrons are usually supplied from the reduced pyridine nucleotide coenzymes, NADH or NADPH. However, the two electrons must be transferred to P450 via electron transport protein(s).

P450s can be divided into four classes (Figure 1.5). Mitochondrial as well as most bacterial P450s are three component systems, in which electrons flow from NAD(P)H to an FAD-containing protein to an iron-sulfur protein to P450 (class I). FAD can accept two electrons from NADPH. The transfer occurs one electron at a time through adrenodoxin, which is a one electron carrier (Lambeth, Seybert et al. 1982). In mitochondria, the P450s are anchored to the membrane, whereas the electron transfer proteins are soluble (Mitani 1979). The bacterial P450cam (CYP101A1) from *Pseudomonas putida* (the first well characterized bacterial P450)

is a soluble protein which catalyzes 5-*exo* hydroxylation of camphor (Atkins and Sligar 1988). Its redox partners were identified as the NADH-dependent FAD-containing reductase, putidaredoxin reducase, and the 2Fe-2S cluster-containing ferredoxin, putidaredoxin. Both are soluble proteins and functionally similar to the mitochondrial electron transfer proteins. In contrast, the microsomal P450s (class II), are two component systems, both are membrane bound, with a NADPH-dependent diflavin reductase containing both FAD and FMN, to transport electrons from NADPH to P450 (Smith, Tew et al. 1994). FAD can be reduced by two electrons from NADPH, and FMN acts as the single electron carrier. The midpoint potentials of FAD and FMN are -328 mV and -190 mV, respectively. Thus, the sequence of electron transfer is NADPH → FAD → DMN → P450 (Vermilion, Ballou et al. 1981). In some microsomal P450 reactions, the NADH-cytochrome *b5* reductase/cytochrome *b5* electron transport system may participate in delivering electrons (White 1991). Class III P450s, such as CYP102A1 (P450 BM-3) from *Bacillus megaterium*, contain the same cofactors as the class II P450s but are soluble and fused into one single polypeptide (Narhi and Fulco 1986). Class IV P450s, recently discovered in *Rhodococcus*, are also soluble, one polypeptide enzymes but contain an NADPH-dependent, FMN-containing reductase and ferredoxin fused to the heme domain (Roberts, Grogan et al. 2002; Liu, Schmid et al. 2006).

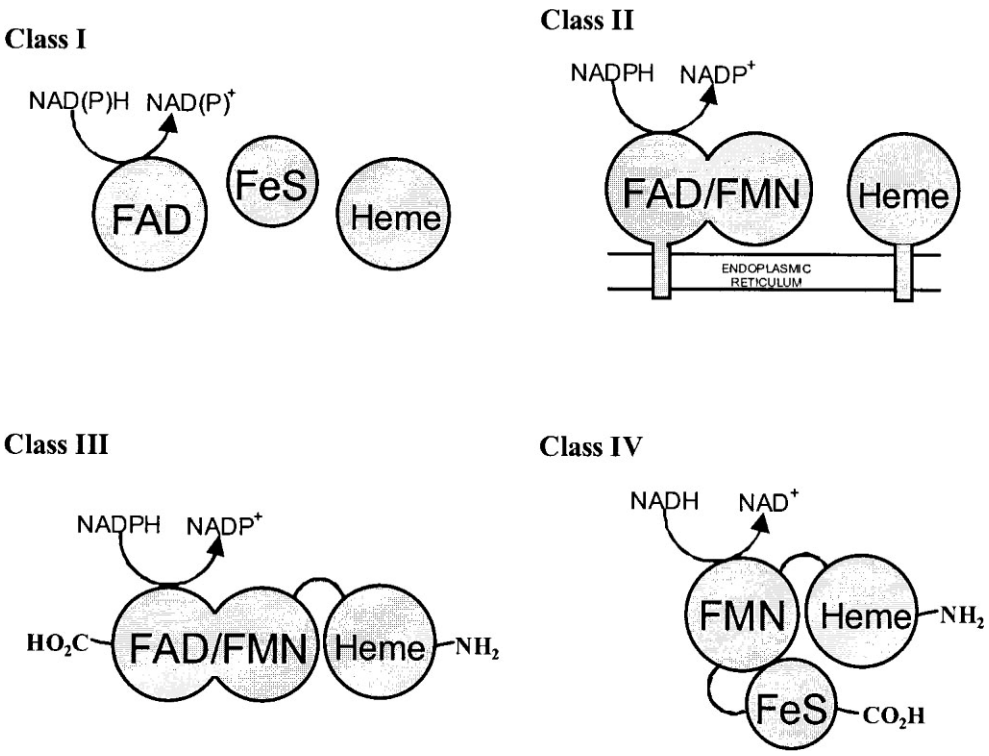

Figure 1.5. Schematic representation of the different classes of cytochrome P450 systems (Roberts, Grogan et al. 2002).

3.1.7. Self-sufficient P450 monooxygenases

In most cases, the cytochrome P450 redox partners are produced *in vivo* as separate polypeptides. However, fatty acid hydroxylase P450 BM-3 (CYP102A1) from *Bacillus megaterium,* which was described first in 1986 (Narhi and Fulco 1986), is a fusion protein inducible by barbiturates and consisting of a P450 domain and a reductase domain containing FAD and FMN. P450 BM-3 is the first known soluble natural fusion protein and was cloned and overexpressed in *E. coli* (Boddupalli, Estabrook et al. 1990). Two additional genes encoding P450 BM-3 like proteins (CYP102A2 and CP102A3) were predicted from genome sequence of *Bacillus subtilis* (Kunst, Ogasawara et al. 1997). The two P450s show 60% homology to P450 BM-3 at amino acids sequence level and were cloned, expressed and characterized (Budde, Maurer et al. 2004; Gustafsson, Roitel et al. 2004). They are fatty acid hydroxylases as well. An eukaryotic fusion protein containing P450 and its reductase was identified in fungus *Fusarium oxysporum* (Nakayama, Takemae et al. 1996). Different to P450 BM-3, this fatty acid hydroxylase (CYP505A1) is a membrane-bound protein (Kitazume, Takaya et al. 2000). A new class of P450 (P450RhF or CYP116B2) was identified from *Rhodococcus* sp. NCIMB 9784 (Roberts, Grogan et al. 2002). This is the first member of class IV. It consists of a P450 domain, a reductase domain and a ferredoxin domain. P450RhF was found to catalyze *O*-dealkylation of 7-ethoxycoumarin but the biological function and the natural substrates are unknown.

All known fusion P450 monooxygenase systems are self-sufficient, as they do not require any additional proteins for an electron transfer from NAD(P)H to the heme iron. A single peptide represents an easier component for industrial applications than a multi-component protein complex in terms of high-performance biotransformation or enzyme bioreactor. The self-sufficient P450 monooxygenases are therefore good candidates for industrial applications.

3.1.8. Industrial and environmental application of P450s

Chemical oxidation reactions are often unspecific and lead to the formation of additional undesired by-products. Besides that, they frequently require extremely high temperatures. The P450 enzymes, which catalyze the stereo- and regiospecific hydroxylation of non-activated hydroxycarbons under mild conditions, have therefore become interesting

candidates for industrial applications. Such biotransformations are especially useful in the food and pharmaceutical industries where high reaction selectivity on complex substrates is mandatory (Schmid, Dordick et al. 2001). Some commercial applications of oxygenases including production of pravastatin were summarized by van Beilen et al. in Table 1.1 (van Beilen, Duetz et al. 2003).

Table 1.1 Industrial applications involving oxygenases. Summarized by van Beilen et al. (van Beilen, Duetz et al. 2003).

Company	Reaction	Scale or status
Upjohn (1952), now Pharmacia (Peapack, New Yersey, USA)	Cortexone to corticosterone	Almost all centres of the steroid nucleus can be hydroxylated stereospecifically
Sankyo Pharma (Parsippany, NJ, USA) & Bristol-Myers Squibb Company (New York, NY, USA)	Compactin (Mevastatin) to Pravastatin	Annual market value US \$3.6 10^9, product concentration 5 g l^{-1}, conversion 70%
BASF (Ludwigshafen, Germany)	Hydroxylation of R-2-phenoxy propionic acid at the 4-position	Volumetric productivity 7 g l^{-1}.d, scale 120 000 l

The development of oxygenases for industrial application is more difficult than of hydrolases, isomerases, and lyases. Oxygenases are usually unstable, multiple components. They may be membrane-bound and require expensive co-factors such as NAD(P)H (Duetz, van Beilen et al. 2001). Currently, all the known applications of monooxygenases are carried out using whole cells systems. One important advantage is that the cells provide a sufficient amount of reduced co-factors (NADPH or NADH) and maintainable production of enzymes, which may be unstable *in vitro*. However, some substrates are toxic for the cells. Large or charged compounds may not easily be taken up by cells.

P450s can also be utilized in the bioremediation of pollutants in the environment, such as polycyclic aromatic hydrocarbons, which are harmful, chemically inert ubiquitous in the environment. These enzymes convert chemically inert compounds (alkanes, polyaromatic hydrocarbons, etc.) into more water-soluble, hydroxylated derivatives, which may be suitable

substrates for many other enzymes (Harford-Cross, Carmichael et al. 2000; Li, Ogawa et al. 2001a).

3.2. Modification of enzyme properties

The efficient biocatalysis has demand on suitable enzymes with high activity and stability under process conditions, desired substrate specificity and high regio- and stereoselectivity. However, wild-type enzymes often need to be improved to fulfill these requirements. There are two general strategies for protein engineering (Figure 1.6): rational redesign (Bornscheuer and Pohl 2001; Bornscheuer 2002) and directed evolution (Farinas, Schwaneberg et al. 2000).

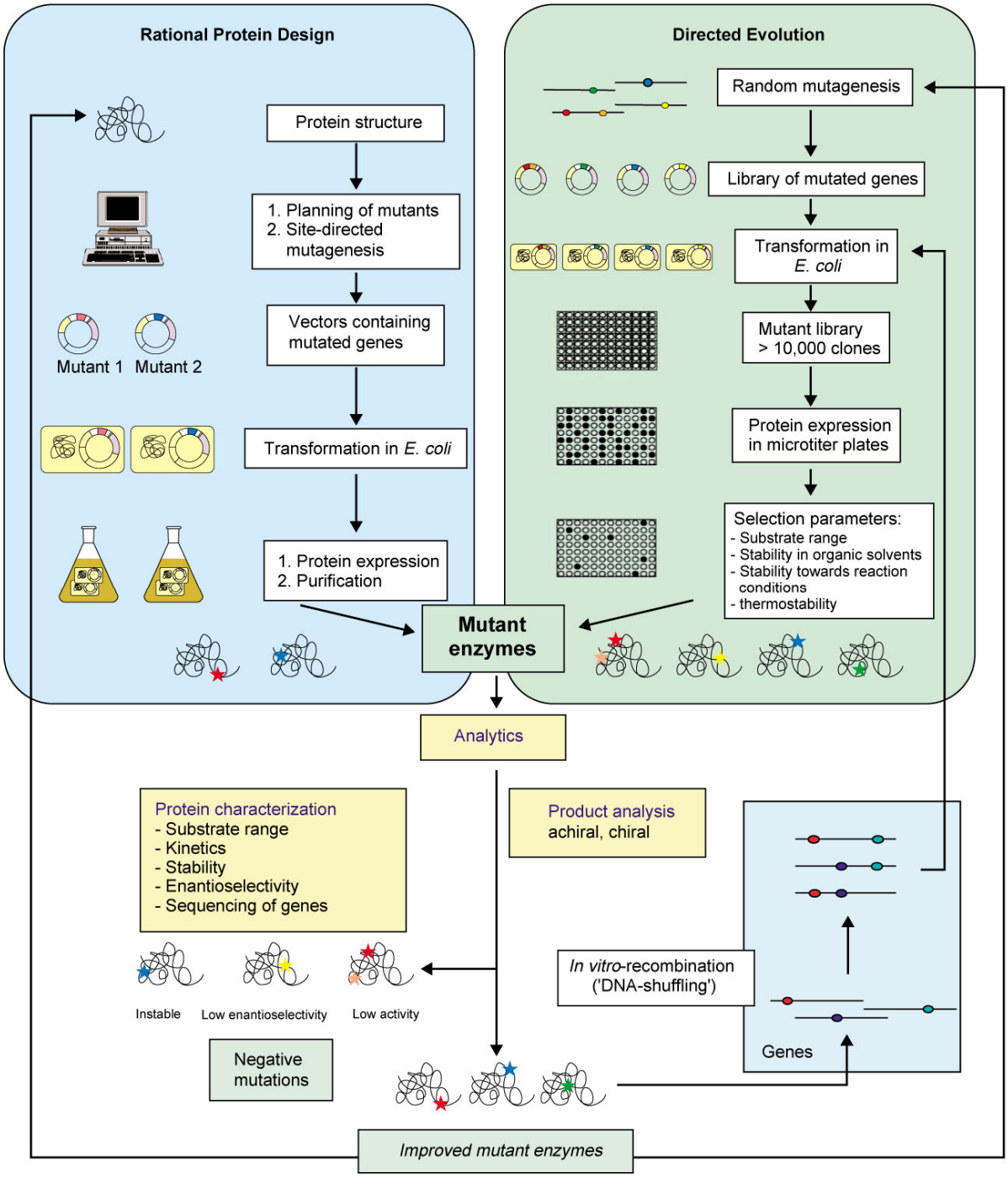

Figure 1.6: Comparison of rational protein design and directed evolution. During rational design mutants are designed on the basis of their structure and prepared by site directed mutagenesis. After transformation in host cells, the variant is expressed, purified and analyzed. Directed evolution starts with generation of mutation libraries by random mutagenesis, which are then expressed in host cells. Mutation libraries are usually screened in microtiter plates using a range of selection parameters. (Bornscheuer and Pohl 2001)

3.2.1. Rational protein design

Rational protein design is still a very effective strategy to elaborate improved enzymes. In rational design, precise changes in amino acid sequence are preconceived based on detailed knowledge about relationship between protein sequence, structure, function and mechanism, and are then introduced using site-directed mutagenesis (Chen 1999). In addition, the number of protein structures and sequences is increasing rapidly in public data bases. Using molecular modelling techniques, it is possible to predict how to introduce desired properties into a protein of interest, even if there are no structural data available because the structure of a homologous protein can be used as a model. Rational protein design has successfully been used for stabilization of enzymes towards thermoinactivation and oxidation. For example, the 3-isopropylmalate dehydrogenase was stabilized by the introduction of additional hydrophobic contacts (Akanuma, Yamagishi et al. 1999). The stability towards oxidation of formate dehydrogenase was increased by removal of cysteine and methionine residues (Slusarczyk, Felber et al. 2000). The substrate selectivity can be also altered by rational protein design. The substrate spectrum of P450cam was extended from camphor to polycyclic aromatic hydrocarbons by mutation of two aromatic residues (Phe87 and Tyr96) in the substrate-access channel (Harford-Cross, Carmichael et al. 2000). Another successful example of P450s has been reported by Li et al. The chain-length specificity of P450 BM-3 was modulated by rational design: in contrast to the wild-type, a variant with five mutations efficiently hydrolyzes a *p*-nitrophenoxyoctanoic acid (C8-pNCA), while the catalytic efficiency towards C10- and C12-pNCA is comparable to that of wild-type P450 BM-3 (Li, Schwaneberg et al. 2001).

3.2.2. Directed evolution

Directed evolution has rapidly become the method of choice for developing enzyme- and microorganism-based biocatalysts. The first examples of directed molecular evolution was the pioneering experiment of S. Spiegelman et al. and M. Eigen and W.Gardiner, who proposed that evolutionary approaches can be adapted for the engineering of biomolecules (Spiegelman, Haruna et al. 1965; Eigen and Gardiner 1984). Directed evolution has emerged as a powerful tool for biocatalyst engineering. The success of the evolutionary approach not only depends on the potency of the method itself but is also a result of the limitations of

alternative approaches, as our lack of understanding of the structure-function relationship of proteins in general hinders the rational design of proteins with new functions. Generally, directed evolution can be considered as repeated cycles of variation followed by selection. By coupling various protocols for generating large variant libraries of genes, together with high-throughput screens that select for specific properties of an enzyme, such as thermostability, catalytic activity and substrate specificity, it is now possible to optimize biocatalysts for specific applications.

In contrast to rational protein design, directed evolution does not require information about how enzyme structure relates to function. It mimics Darwinian evolution in the test tube, and involves the generation and selection of a variant library with sufficient diversity for the altered function (Zhao and Arnold 1997; Arnold 1998). Sequence diversity is typically created by either various random mutagenesis or *in vitro* recombination of gene fragments, such as error-prone PCR or DNA shuffling.

Error-prone PCR introduces random errors into DNA sequences during PCR under mutagenic reaction conditions (e.g. imbalanced nucleotides or by addition of Mn^{2+} or Mg^{2+} to the reaction mixture) (Leung, Chen et al. 1989; Cadwell and Joyce 1992), yielding randomly mutated products. The error rate can be increased more than thousand fold. Addition of nucleotide analogues such as 8-oxo-dGTP or dITP may also influence the mutation rate (Spee, de Vos et al. 1993). Mutator strains with genetic deficiencies in DNA proofreading and editing machinery, such as XL1-RED (Coia, Ayres et al. 1997), can be used to generate large, diverse plasmid libraries. Random errors are introduced into plasmids during plasmid replication. As well as single base substitutions, insertions, and deletions can be observed throughout the entire plasmid. However, the mutator phenotype is unstable and the growth rate of mutator strains is slow (Greener, Callahan et al. 1997).

DNA shuffling, as a method for *in vitro* recombination, was developed by Stemmer in 1994 (Stemmer 1994a; Stemmer 1994b). It was used to improve biocatalysts (Reetz and Jaeger 1999) and has resulted in the improvement of proteins of therapeutic interest (Kurtzman, Govindarajan et al. 2001). It consists of a three-step process that begins with the enzymatic digestion of genes, yielding smaller fragments of DNA. The small fragments are then randomly hybridized and are filled in to create longer fragments. Finally, any full-length, recombined genes that are recreated are amplified via polymerase chain reaction. If a series

of alleles or mutated genes is used as a starting point for DNA shuffling, the result is a library of recombined genes that can be translated into novel proteins, which can be screened for novel functions. Genes with beneficial mutations can be shuffled further, both to bring together these independent, beneficial mutations in a single gene and to eliminate any deleterious mutations.

Some successful examples of directed evolution have been reported. Polar organic cosolvents are used to increase substrate solubility and achieve high catalytic efficiency. Wong et al. evolved the cosolvent-resistance of P450 BM-3 using a pNCA assay (Wong, Arnold et al. 2004). The activity of P450 BM-3 was significantly improved in the presence of DMSO and THF. Lentz et al. modified the fatty acid specificity of P450 BM-3 by directed evolution using random mutagenesis and pNCA assay (Lentz, Li et al. 2001). The best mutants have the ability to hydroxylate 8-pNCA which the wild-type enzyme is unable to do. In another example, Joo et al. used digital-image screening to identify P450cam variants with enhanced activity in naphthalene hydroxylation, in the absence of the cofactor NADPH, via a 'peroxide shunt' pathway (Joo, Lin et al. 1999). Co-expression of P450cam with horseradish peroxidase from *E. coli* converted hydroxylation products into fluorescent products amenable by digital screening.

The recent development of new efficient diversity-generation methods and high-throughput methods will further accelerate the development of biocatalysts. Furthermore, the ever-expanding capabilities of rational design will lead to more powerful biocatalyst design strategies that combine the best of both approaches. Advances in other related fields such as bioinformatics, functional genomics, and functional proteomics will also extend the applications of directed evolution, rational design or a combination of both to many more industrial biocatalysts.

3.2.3. High throughput screening system

Mutation libraries can be generated by a number of methods. These methods rapidly combine beneficial mutations that arise from random mutagenesis and significantly expand the sequence diversity. Functionally improved variants should be identified by a high-throughput screening method and then used as the parents for the next round of evolution. As a result, the

success of directed evolution experiments often depends on the efficiency of the methods used to screen libraries for mutants with desired properties (Olsen, Iverson et al. 2000).

Schwaneberg et al. described a colorimetric assay using a surrogate substrate, para-nitrophenoxycarboxylic acid, pNCA (Schwaneberg, Schmidt-Dannert et al. 1999). This P450 activity detection was based on split of pNCA, leading to formation of spectroscopically detectable *p*-nitrophenolate. The principle is that terminal hydroxylation of pNCA generates an unstable hemiacetal that dissociates to form *p*-nitrophenolate, which can be directly monitored in a spectrophotometer at a wavelength of 410 nm. Like fatty acids, the same strategy can be used to screen for oxidation activity towards alkane (Farinas, Schwaneberg et al. 2000).

P450s may also be coupled to a second enzyme which converts the product into an easily detectable compound. Digital-image screening was reported to be used in directed evolution of P450 enzymes (Joo, Lin et al. 1999). A P450cam variant was identified showing enhanced hydroxylation activity towards naphthalene in the absence of cofactors through the "peroxide shunt" pathway. Co-expression of P450cam with horseradish peroxidise converted the products of the P450 reaction into fluorescent compounds amenable to digital imaging screening.

3.3. Rhodococci

The genus *Rhodococcus* belongs to the family of actinomycetes. Rhodococci are aerobic, gram-positive, non-sporulating bacteria with high G+C content in genomic DNA (60 – 70%), and exhibit nocardioform. The term "nocardioform" describes the morphology and refer to mycelial growth with fragmentation into rod-shaped or coccoid elements (Lechevalier 1989). Currently, there are twelve established *Rhodococcus* species (Bell, Philp et al. 1998; Goodfellow, Alderson et al. 1998): *R. coprophilus*, *R. equi*, *R. erythropolis*, *R. fascians*, *R. globerulus*, *R. marinonascens*, *R. opacus*, *R. percolatus*, *R. rhodnii*, *R. rhodochrous*, *R. ruber* and *R. zopfii*. Rhodococci are ubiquitous in nature and have frequently been isolated from a large variety of sources including soils, rocks, boreholes, groundwater, marine sediments, animal dung, the guts of insects and from healthy and diseased animals and plants. (Goodfellow 1989). The rhodoccoci cause diseases of human, animals and plants. *R. equi* is

the important veterinary pathogen causing broncho pneumonia, enteritis, lymphadenitis, abortion, and other diseases in animals (Scott, Graham et al. 1995). *R. fascians* is a plant pathogen causing fasciation disease (Murai, Skoog et al. 1980).

The *Rhodococcus* species are becoming increasingly interesting in environmental and industrial biotechnology. The rhodococci exhibit a broad substrate diversity for the degradation of phenols, aromatic acids, halogenated phenols, halogenated alkanes, substituted benzene, anilines, and quinolines (Finnerty 1992). *Rhodoccocus* sp. IGTA8 is able to covert dibenzothiophene in petroleum with release of inorganic sulfur. The desulfurization of dibenzothiophene is catalyzed by a multienzyme pathway consisting of two monooxygenases and a desulfinase (Gray, Pogrebinsky et al. 1996). They are also able to synthesize several products such as surfactants, flocculants, amides and polymers. The most important biotechnological process using a *Rhodococcus* is the production of acrylamide from acetonitrile using *R. rhodochrous* J1 by Nitto Chemistry Industry Company ltd (Japan) (Yamada and Kobayashi 1996). This is the first example of successful industrial production of a commercial chemical using a microbe. *R. ruber* is able to produce poly(3-hydroxyalkanoic) acid (Pieper and Steinbuchel 1992). A copolyester of 3-hydroxybutyrate and 3-hydroxyvalerate (3HB-co3HV), is a biodegradable plastic, commercially named "BIOPOL", produced by Zeneca Bio Products using the bacterium *Alcaligenes eutrophus* (Lee 1996). *R. ruber* can produce 3HB-co3HV with a different monomer and has an advantage growing on cheaper substrates than those used for BIOPOL production.

3.4. Aim of this work

The focus of this work was isolation, cloning, expression, characterization and engineering of new cytochrome P450s from bacterial sources.

Cytochrome P450 enzymes belong to the monooxygenase superfamily and can be found throughout all kingdoms of life. The ability of cytochrome P450s to catalyze difficult oxidation of non-activated hydrocarbons with high specificity and selectivity makes them attractive for use in biocatalysis and bioremediation. In addition, most bacterial P450 enzymes are water soluble and exhibit relatively high stability. Therefore, in this study we were interested in finding new cytochrome P450 activities from bacterial sources.

Furthermore, fusion cytochrome P450s are self-sufficient and especially interesting in this study, because no additional proteins are necessary for electron transfer from NAD(P)H to the heme iron. Thus, the self-sufficient cytochrome P450s are attractive candidates for industrial application.

However, wild-type enzymes are generally limited by low activity, low stability, or low selectivity. There are two common strategies to overcome this limitation. Rational design is an effective method to expand enzyme properties based on protein structure. Directed evolution is a powerful tool to improve enzyme properties without any structural information. A high-throughput screening system had to be developed to identify variants with desired properties from mutation libraries.

4. Materials and methods

4.1. Instruments

Instrument	Characteristics	Manufacturer
Agarose gel electrophoresis	DNA Sub Cell™	BioRad
	Mini Sub™ DNA Cell	BioRad
	Mini Sub™ Cell GT	BioRad
	Video Copy Processor P66E	Mitsubishi
	BWM 9X Monitor	Javelin Electronics
	UV-lamp table	NEG-Biotech
Balances	Basic, MCI Research RC 210 D	Sartorius
	Precision Advanced	OHAU®
Centrifuges	Eppendorf Centrifuge 5417 C	Eppendorf
	Eppendorf Centrifuge 5417 R	Eppendorf
	Eppendorf Centrifuge 5810 R	Eppendorf
	Universal 30F	Hettich
	KR 22 I (Rotor: AK 500-11, 155 mm)	Jouan
	G412	Jouan
	Sorvall RC-5B (Rotor: SA 600)	Du Pont Instruments
DNA-Sequencer	377 DNA Sequencer	Applied Biosystems
Fluorimeter	Fluostar	BMG
Homogenisator	Ultra-Turrax T25	Janke & Kunkel
Incubators	Certomat R Incubator	Braun
	HT-Incubator	Infors AG
	UM 500	Memmert
	WTE	Binder
Microwave oven	Micro-Chef FM A935	Moulinex
Mixer Mill	MM 2000	Retsch, Haan
PCR-cycler	Master Cycler Gradient	Eppendorf
	Robocycle Gradient 40	Stratagene®
pH-Meter	Digital pH Meter pH525	WTW

Pickroboter	BioPick	BioRobotics Ltd., Cambridge
Polyacrylamide gel electrophoresis	Minigel-Twin G42	Biometra®
	Model 583 Gel Dryer	BioRad
Power supplies	Power Pac 3000	BioRad
	Power Pac 300	BioRad
	Model 200/2.0 Power supply	BioRad
Rotavapor	Rotavapor R-134	Büchi
Sonifier	Sonifier 250	Branson
	Sonorex Super RK 514 H	Brandelin
Spectrophotometer UV/VIS	Ultrospec 3000	Pharmacia Biotech
	Cary 3E UV-VIS	Pharmacia Biotech
Thermomixer	Thermomixer 5436	Eppendorf
Ultrafiltration system	8050 Amicon	Amicon
Vacuum concentrator	SpeedVac Concentrator 5310	Eppendorf
Vortex	Vortex Genie 2	Scientific Industries
Water bath	Water bath B3	Haake-Fisons

4.2. Chemicals and enzymes

Chemicals	Company
Acenaphthylene	Lancaster
Fluorene	Lancaster
9-Fluorenone	Lancaster
9-Fluorenol	Lancaster
Acenaphthene	Sigma-Aldrich
7-Hydroxycoumarin	Sigma-Aldrich
7-Ethoxycoumarin	Fluka
1 kb-Ladder	MBI Fermentas
Acetic acid	
Acetone	
Acrylamide (30%)	Roth

Agar	
Agarose	
Ampicillin (Na-salt)	Fluka
Benzyl alcohol	
Calf intestinal alkaline phosphatase (CIAP)	
Coomassie Brilliant Blue R250	
Cyclohexane	
Diethyl ether	
Dimethyl sulfoxide (DMSO)	
N,N'-dimethyl formamide (DMF)	
dNTPs	
*EcoR*I enzyme	Fermentas
Ethanol	
Ethidium bromide	
Ethylene diamine tetraacetate (EDTA)	
Glucose	
Glycerol	
Kanamycin	
Imidazole	
indene	
Isopropyl-beta-D-thiogalactopyranoside (IPTG)	
*Hind*III enzyme	Fermentas
Magnesium chloride	
Magnesium sulfate anhydride	
Manganese chloride	
Naphthalene	
1-naphthol	
*Nde*I enzyme	Fermentas
Pfu DNA-polymerase	Fermentas
PstI enzyme	Fermentas
Rubidium chloride	
Sodium chloride	
Sodium dithionite	

T4 ligase	
Taq DNA-polymerase	
Tetramethylethylenediamine (TEMED)	
Toluene	
Tris-(hydroxymethyl)-aminomethane	
Triton® X-100	
5-Bromo-4-Chloro-3-Indolyl-β-D-galactoside (X-gal)	
Trypton	Roth
Xylene	
Yeast extract	Roth

4.3. Primers used

Name	Sequence 5' to 3'
Cyp_F1	CTACTGGGTSGTCACSCGSTACGA
Cyp_F2	CTCGTGAACGARGACGA
Cyp_R1	GCAYTCCTCGAYGGCSTTGGGGAT
Cyp_R2	GGGATGTCGACKTCGCCGA
WP1	CTAATACGACTCACTATAGGGNNNNATGC
WP2	CTAATACGACTCACTATAGGGNNNNGATC
WP3	CTAATACGACTCACTATAGGGNNNNTAGC
WP4	CTAATACGACTCACTATAGGGNNNNCTAG
NWP	CTAATACGACTCACTATAGGG
Walk_bio_P1	Biotin-TCGAGGGCGATGGCCGGTGA
Walk _P2	AACACGAGGTTGTCGCGGAA
Walk _Bio_P3	Biotin-TTCAAGCTGCTGCTCGAGAA
Walk _P4	TGGGAGGAGATCTGCGCGGA
Walk_bio_P5	Biotin-CTTCCTCGAGGAACTGACCA
Walk_P6	CGTACCCGATCAGGAGTTCA
Walk_bio_P7	Biotin-GGTGCGCTGTCGCGGCAGTA
Walk_P8	TGCGCCCGACGCGCCCACCTA
Walk_bio_P9	GAACCGTTCCGGGATGGCCTA

Walk_P10	GCACGTCTCGGCGGAAGGCA
Ruber-F	CTGGAATTCATGAGTGCATCAGTTCCGGCGT
Ruber-R	CATCAAGCTTTCAGAGTCGCAGGGCCA
ruber-FS1	GTGAAGGCGGTGTTCCGCGA
ruber-FS2	GTCGGCAGGTTCTGGCAGTA
ruber-FS3	GAACTCGAGCGACCACCTCA
ruber-FS4	GGTGCGCTGTCGCGGCAGTA
ruber-FS5	GCACTTCACCTCGTCCCTCA
ruber-RS1	AGGGTCCGGGCCATGGCGTA
ruber-RS2	GTCGGTGACGACGTCCGGCA
ruber-RS3	CTGATCGGGTACGAGTTCCA
ruber-RS4	TGTTCGTGGACGTAGCGCGA
ruber-RS5	GGCGCGCAACGCGTCGAGGA
T7-Promotor	TAATACGACTCACTATAGGG
T7-Terminator	GTTAGTTATTGCTCAGCGC
A109F-F	CGACTACGCCATGTTCCGGACCCTCGTGAACG
A109F-R	CGTTCACGAGGGTCCGGAACATGGCGTAGTCG

(S = G or C; Y = C or T; R = A or G; K = G or T)

4.4. Microorganisms and plasmids

Microorganism	Strain
Escherichia coli	DH5α
Escherichia coli	BL21(DE3)
Escherichia coli	Novablue(DE3)
Pseudomonas citronellolis	DSM 50332
Pseudomonas fluorescens	DSM 50090
Pseudomonas savastanoi	DSM 50267
Sphingomonas echinoides	DSM 1805
Rhodococcus rhodochrous	NCIMB 11216
Rhodococcus erythropolis	DSM 43066
Rhodococcus ruber	DSM 44319

Vector:

pET28a(+) with selective marker kanamycin resistance. (Novagene)

4.5. General buffers, media and solutions

4.5.1. Buffers and media

Buffer or Medium	Compound	Amount
Potassium phosphate buffer	1 M KH_2PO4	10 ml
Added water to final volume 1 liter	1 M K_2HPO4	40 ml
and pH value adjusted to pH 7.5		
1 x TE buffer	Tris	10 mM
with HCl	EDTA	1 mM
LB medium	tryptone	10 g/l
The pH was adjusted to 7.4 with	yeast extract	5 g/l
NaOH.	sodium chloride	5 g/l
	agar (Only for solid media added)	15 g/l

4.5.2. Antibiotics and others

Compound	Stock solution	Working concentration
Ampicillin (Amp)	100 mg/ml	100 µg/ml
Kanamycin (Kan)	30 mg/ml	30 µg/ml
IPTG	0.5 M	0.5 mM
PMSF	100 mM (in ethanol)	0.1 mM
X-gal	20 mg/ml (in DMF)	

4.5.3. Commercial kits

Kit	manufacturer
Spin Mini preparation kit	Qiagen GmbH
Midi preparation kit	Qiagen GmbH
QIAquick gel extraction kit	Qiagen GmbH
GenElute plasmid miniprep kit	Sigma
TA Cloning kit	Stratagene

4.6. Storage and cultivation of bacteria

4.6.1. Storage of *E. coli* stains

For long-term storage, 1 volume sterile 87% glycerol was added to 1 volume of overnight culture. The mixture was homogenized by vortexing and stored at – 80 °C. To recover the bacteria, an inoculum from the glycerol stock was streaked with a sterile inoculation needle onto an LB agar plate containing appropriate antibiotics.

4.6.2. Cultivation of *E. coli* stains in Erlenmeyer shake-flasks

To prepare a pre-culture, a single colony was picked from an LB agar plate and inoculated in 5 ml liquid LB medium containing appropriate antibiotics and grown at 37 °C and 180 rpm overnight. The overnight culture was used to inoculate fresh LB medium containing selection antibiotics with dilution 1 : 100. Main cultures were incubated at 37 °C and 180 rpm for *E. coli* DH5α and at 30 °C and 120 rpm for *E. coli* BL21(DE3) and *E. coli* Novablue(DE3) in shake-flasks.

4.6.3. Heterologous protein expression in *E. coli*

Protein over-expression was carried out in *E. coli* BL21(DE3) or *E. coli* Novablue(DE3) with vector pET28a(+), which has a T7 promoter and is controlled by the *lac* operator, which can be induced by addition of 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG). An overnight

culture was used to inoculate 1 : 100 fresh LB medium containing kanamycin (30 µg/ml). The cells were incubated at 30 °C and 120 rpm. At $OD_{600} \approx 0.8$ the cells were induced with 0.5 mM IPTG, incubation temperature was reduced to 25°C, and expression was carried out overnight.

4.6.4. Storage and cultivation of other bacteria

An ampoule was opened and the dried culture was resuspended in 1 ml liquid medium. One drop of the suspension was streaked onto an LB agar plate and incubated at 30 °C for two days. A single colony was picked into 5 ml LB medium and incubated at 30 °C overnight. The overnight culture was used to inoculate fresh LB medium with dilution 1 : 100. Main culture was incubated at 30 °C and 150 rpm overnight. For long-term storage all strains were stored in 50% glycerol at – 80 °C.

4.6.5. Automatic colony picking with a BioPick Robotics

In order to screen a mutation library, single colonies should be cultured in discrete reaction compartments. The BioPick is capable of picking single bacteria colonies using a single solid needle. The BioPick recognized single colonies on an agar plate using a CCD-camera, which took a picture and the picture was analyzed and recognized by computer. The recognized colonies were picked into a 96-well microtiter plate, which was filled with 200 µl LB medium containing 5% DMSO and appropriate antibiotics. The parameters were adjusted before picking each time.

4.6.6. Cultivation and expression of *E. coli* stains in 96-well microtiter plates

Microtiter plate was incubated at 37 °C and 600 rpm overnight. The microtiter plate can be stored at – 80 °C for a long-term storage. The cells were inoculated with a 96-pins replicator into a deep-well microtiter plate containing 600 µl LB medium and appropriate antibiotics. The deep-well microtiter plate was covered with a gas-permeable microplate sealer (Breathseal™, Greiner Bio One) and than incubated at 30 °C and 600 rpm overnight. The

cells were induced by addition of IPTG (final concentration 0.25 mM) and incubated at 25 °C and 600 rpm overnight.

4.7. Transformation of *E. coli* strains by heat shock

Transformation is the genetic alteration of a cell by the introduction, uptake and expression of foreign DNA. Competent cells are cells which are capable of uptaking extra-chromosomal DNA such as plasmids.

Solutions:

TfbI (per 200 ml)

Compound	Amount	Final concentration
Potassium acetate	0.558 g	30 mM
Rubidium chloride	2.42 g	100 mM
Calcium chloride	0.294 g	10 mM
Manganese chloride	2.0 g	50 mM
Glycerol	30 ml	15% v/v
pH 5.8 with dilute acetic acid		

TfbII (per 100 ml)

Compound	Amount	Final concentration
MOPS	0.21 g	10 mM
Calcium chloride	1.1 g	75 mM
Rubidium chloride	0.121 g	10 mM
Glycerol	15 ml	15 % v/v
pH 6.5 with dilute NaOH		

Preparation of *E. coli* competent cells by rubidium chloride method:
1 ml overnight culture of *E. coli* was transferred into 100 ml LB medium and the cells were incubated at 37 °C and 180 rpm to an OD_{600} 0.4 – 0.6. The cells were chilled on ice for 15 min and centrifuged at 4 °C and 4000 rpm for 10 min. The supernatant was discarded and the cells were resuspended in 40 ml pre-chilled TfbI gently by pipetting up and down. The cells were incubated on ice for 15 min. The cells were pelletted by centrifugation (4 °C, 4000 rpm

and 10 min) and the supernatant was discarded. The pellet was resuspended in 4 ml pre-chilled TfbII gently. For long-term storage, cells were aliquoted (200 µl), frozen in liquid nitrogen and stored at – 80 °C.

Transformation procedure:
1 µl plasmid DNA or 20 µl ligation mixture were added into an aliquot of competent *E. coli* cells, mixed gently and incubated on ice for 30 min. The heat-shock of the cells was carried out at 42 °C for 40 sec and after that the tube was transferred immediately onto ice. 800 µl room temperature LB medium were added into the tube and the tube was incubated at 37 °C and 200 rpm for 1 hour to allow expression of antibiotic resistance gene. The transformed cells were centrifuged at 4000 rpm for 2 min and the supernatant was discarded. The pellet was resuspended in the remaining supernatant and spread on a pre-warmed selective LB-agar plate containing appropriate antibiotics. The plate was incubated at 37 °C overnight for formation of colonies.

4.8. Molecular biological methods

4.8.1. Isolation of genomic DNA

To isolate genomic DNA from *Rhodococcus ruber* DSM 44319 or *Rhodococcus erythropolis* DSM 43066, 25 ml overnight culture was centrifuged at 4000 rpm for 15 min. The cell pellet was resuspended in 5 ml LI-solution (Tris-HCl 25 mM; EDTA pH 8.0 25 mM; saccpharose 300 mM) and a few milligram lysozyme (Sigma). After incubation at 37°C for 30 min 1 ml 10% SDS was added to the suspension and vortexed. After incubation at 60°C for 10 min 2 ml phenol/chloroform were added and mixed gently. The mixture was chilled on ice for 5 min and centrifuged at 4 °C and 4000 rpm for 25 min. The supernatant was removed in a fresh falcon tube. Phenol/chloroform extraction should be repeated if the supernatant is unclear. To precipitate the DNA, 1/10 volume 3 M sodium acetate and 1 volume pre-chilled isopropanol were added, and the mixture was subsequently centrifugated at 4°C and 4000 rpm for 40 min. The pellet was twice washed with 70% ethanol and solved in TE buffer finally.

DNA precipitation with ethanol and isopropanol

One of the simple methods of purification and concentration of DNA is precipitation with alcohols. In the presence of alcohol and monovalent cations, DNA is precipitated at low temperature. The DNA is recovered by rinsing in 70% alcohol, centrifugation, and resuspension in appropriate buffer.

Ethanol precipitation
Procedure:
To 1 volume of DNA solution 1/10 volume 3 M sodium acetate (pH 4.8) and 2.5 volume 100% pre-cooled ethanol (– 20 °C) were added and mixed gently. The mixture was placed at – 20 °C for 30 min. Then the mixture was centrifuged at 14000 rpm and 4 °C for 30 min. The supernatant was removed and the pellet was rinsed gently with 200 µl 70% ethanol and centrifuged as above. Finally the DNA was air dried and resuspended in water or TE buffer.

Isopropanol precipitation
Procedure:
This method was used to precipitate DNA from large volumes.
To 1 volume of DNA solution 0.7 volume isopropanol was added and mixed gently. After the mixture was placed at room temperature for 15 min, it was centrifuged at 14000 rpm for 15 min. The supernatant was removed and the pellet was rinsed gently with 200 µl 70% ethanol, and recovered by centrifugation as above. Finally the DNA was dried in a vacuum centrifuge and resuspended in water or TE buffer.

4.8.2. Polymerase chain reaction (PCR)

The polymerase chain reaction is used for enzymatic replication of a short fragment of DNA which lies between two parts of a known sequence. Two oligonucleotides are used as primers to determine the DNA fragment to be amplified. The DNA template is denatured by heating to break the hydrogen bonds between the DNA strands. When the solution cools to the annealing temperature, the primers bind to their target sequence, and DNA polymerase extends the annealed primers by joining the free nucleotide bases to the primers. The product

can be used as template for the next amplification, when the process is repeated, resulting an exponential increase of the desired DNA product.

Procedure:
Genomic DNA was used as template DNA. Original polymerase buffer was used. The components of the reaction mixture were pipetted together according to the following scheme:

Component	Volume (µl)	Final concentration
10 x polymerase buffer	5	1 x
DMSO	0 – 5	0 – 10%
dNTPs (10 mM)	4	0.8 mM
Forwards primer (100 pM)	1	2 pM
Reverse primer (100 pM)	1	2 pM
DNA template	variable	appx. 70 ng
Polymerase	1	1 U
Added H_2O to final volume of 50 µl		

The PCR reaction was carried out in a thermal cycler using the following program:

Step	Temperature (°C)	Duration (min)
Step 1. initial denaturation	96	4
Step 2. denaturation	96	1
Step 3. annealing	58	1
Step 4. extension	72	4
Step 5. final extenstion	72	4 *

Step 2 to step 4 was repeated for 30 cycles.

4.8.3. PCR with Dynazyme™ EXT

Due to strong secondary structure, GC-rich templates can be difficult to amplify using PCR. Dynazyme™ EXT polymerase from Finnzymes can be used with PCR additive, like DMSO, which makes the denaturation easier by relaxing the template DNA. In this study it was

applied for DNA amplification with degenerated primers, to get high yield and low nonspecific products after PCR on GC-rich template (up to 70% GC-content).

Procedure:

Genomic DNA was used as template DNA. Original polymerase buffer was used. The components of the reaction mixture were pipetted together according to the following scheme:

Component	Volume (µl)	Final concentration
10 x Dynazyme buffer (500 mM Tris-HCl (pH 9.0 at 25 °C), 15 mM $MgCl_2$, 150 mM $(NH_4)_2SO_4$ and 1 % Triton® X-100)	5	1 x (50 mM Tris-HCl (pH 9.0 at 25 °C), 1.5 mM $MgCl_2$, 15 mM $(NH_4)_2SO_4$ and 0.1 % Triton® X-100)
DMSO	2.5	5%
dNTPs (10 mM)	4	0.8 mM
Forwards primer (100 pM)	1	2 pM
Reverse primer (100 pM)	1	2 pM
DNA template	variable	appx. 70 ng
Dynazyme™ EXT	1	1 U
Added H_2O to final volume of 50 µl		

The PCR reaction was carried out in a thermal cycler using the following program:

Step	Temperature (°C)	Duration (min)
Step 1. initial denaturation	96	4
Step 2. denaturation	96	1
Step 3. annealing	58	1
Step 4. extension	72	4
Step 5. final extenstion	72	4 *

Step 2 to step 4 was repeated for 30 cycles.

*If the PCR product was further used for TA-Cloning, 20 min extension at 72°C completes the program to ensure full length PCR products and the 3´ adenylation, otherwise set extension at 72°C only for 4 min after the last cycle.

4.8.4. TA-Cloning

TA-cloning was carried out with TOPO TA Cloning® kit (Invitrogen). *Tag* polymerase has an adenylation activity, which adds a single deoxyadenosine (A) to the 3' end of the PCR product. The linearized vector supplied in the kit has a single deoxythymidine (T) overhang at the 3' end, it allows PCR insert to ligate effective with the vector. The ligation is catalyzed by topoisomerase I.

Solution and medium:

Salt solution	NaCl	1.2 M
	$MgCl_2$	0.06 M
S.O.C. medium	Trypton	2%
	Yeast extract	0.5%
	NaCl	10 mM
	KCl	2.5 mM
	$MgCl_2$	10 mM
	$MgSO_4$	10 mM
	Glucose	20 mM

Procedure:

The components of the reaction mixture were pipetted together according to the following scheme:

Component	Amount
Fresh PCR product	0.5 to 4 µl
Salt solution	1 µl
Sterile water	Add to 5 µl
TOPO® vector	1 µl
Final volume	6 µl

The cloning reaction mixture was mixed gently and incubated at room temperature for 5 min. 2 µl from the reaction mixture were added into a tube of One Shot® chemically competent *E.*

coli cells, mixed gently and incubated on ice for 20 min. The heat-shock of the cells was performed at 42 °C for 30 sec and after that the tube was transferred immediately onto ice. 250 µl room temperature S.O.C. medium were added into the tube and the tube was incubated at 37 °C and 200 rpm for 1 hour. 50 µl of the transformation mixture were spread on a prewarmed selective LB-Agar plate containing X-gal and appropriate antibiotics. The plate was incubated at 37 °C overnight. About 10 white colonies were picked in liquid LB medium and incubated at 37 °C overnight. Plasmid was isolated from the overnight culture for sequencing of the DNA insert.

4.8.5. Error-prone PCR

Error-prone PCR is the most commonly used method to generate variants with random mutations. Normally Tag polymerase is used because of its naturally high error rate, with errors biased towards AT to GC changes. Error-prone PCR reactions commonly contain higher concentrations of $MgCl_2$ than in basic PCR, in order to stabilize non-complementary pairs. Using unequal dNTPs and an addition of $MnCl_2$ and dITP increase the error-rate in PCR.

Procedure:
Plasmid carrying *CYP116B3* gene was used as DNA template. Original polymerase buffer was used. The components of the reaction mixture were put together according to the following scheme:

Component	Volume (µl)	Final concentration
10 x polymerase buffer	5	1 x
Unequal dNTPs	4	0.8 mM
Forwards primer (100 pM)	1	2 pM
Reverse primer (100 pM)	1	2 pM
DNA template	1	appx. 10 ng
Taq Polymerase	1	5 U
$MgCl_2$ (25 mM)	3	1.5 mM
$MnCl_2$ (5 mM)	0.5	0.05 mM
dITP (5 mM)	0.5	0.05 mM
Added H_2O to final volume of 50 µl		

The PCR reaction was carried out in a thermal cycler using the following program:

Step	Temperature (°C)	Duration (min)
Step 1. initial denaturation	96	4
Step 2. denaturation	96	1
Step 3. annealing	50	1
Step 4. extension	72	1.5
Step 5. final extension	72	4

Step 2 to step 4 was repeated for 30 cycles.

4.8.6 Isolation of plasmid DNA from *E. coli*

Minipreparation of plasmid for fast tests

Minipreparation is a technique to extract small amount (up to 5 µg) of plasmid DNA from bacterial cells. This method is based on alkaline lysis (Birnboim and Doly 1979). After completely lysis of cells with SDS, RNA was hydrolyzed by NaOH and RNaseI. Plasmid DNA after denaturation by NaOH was immediately renatured properly in a neutralization buffer whereas genomic DNA precipitated out. After centrifugation the plasmid DNA in supernatant was concentrated and purified by alcohol precipitation.

Buffers:

Resuspension buffer	Tris/HCl	100 mM
	EDTA	10 mM
	RNAseI	250 µg/ml
Lysis buffer	NaOH	200 mM
	SDS	1 % (w/v)
Neutralization buffer	Potassium acetate	29.4 g
	Acetic acid	11.5 ml
	H_2O	28.5 ml

Procedure:
1.5 ml overnight culture was centrifuged at 14000 rpm for 1 min. The supernatant was discarded and the cell pellet was resuspended in 200 µl resuspension buffer. To lyse the cells, 200 µl lysis buffer was added and mixed gently. After neutralization by addition of 200 µl neutralization buffer the proteins and genomic DNA were separated by centrifugation at 14000 rpm for 10 min. The supernatant was transferred in a fresh tube. The plasmid DNA from the supernatant was concentrated and purified by alcohol precipitation.

Minipreparation with a spin-column

This method was used to isolate highly pure plasmid DNA for sequencing or enzymatic digestion. Spin Miniprep kit from Qiagen or Sigma was used. It is also based on alkaline lysis, followed by purification of the DNA by ion-exchange chromatography.

Procedure:
2 ml overnight culture was harvested by centrifugation. The plasmid DNA isolation was performed according to the manual of the manufacturer.

Large amount plasmid isolation with midipreparation

In order to isolate up to 100 µl highly pure plasmid DNA, a midipreparation was used. This method is based on alkaline lysis, followed by ion-exchange chromatography and isopropanol precipitation.

Procedure:
50 ml overnight culture was harvested by centrifugation. The Midiprep kit from Qiagen was used. The plasmid DNA isolation was carried out refers to the manual of the manufacturer.

4.8.7. DNA digestion with restriction endonucleases

Restriction endonuclease cuts double-stranded DNA at specific position, known as recognition sequence. Many in laboratory used restriction enzymes produce sticky ends of

DNA molecules, which can be further ligated by T4 ligase to other DNA molecules with the same sticky end.

Procedure:
About 0.5 µg purified DNA was mixed with 5 µl 10 x restriction enzyme buffer, 3 U restriction endonuclease, and water was added to a final volume of 50 µl and mixed gently. The reaction mixture was than incubated at 37 °C for 3 hours. To inactivate the restriction endonuclease the reaction mixture was incubated at 80 °C for 20 min. Afterwards the digested DNA was purified by agarose gel extraction.

4.8.8. Dephosphorylation of DNA

Recircularization of vector without insertion of foreign DNA during ligation decreases cloning efficiency. Because DNA ligation absolutely needs a 5' phosphate group, the self-ligation can be blocked by remove the 5' phosphate group from the vector with calf intestinal alkaline phosphatase (CIAP). This treatment was used to decrease vector background in cloning.

Procedure:
After digestion of vector with restriction endonucleases 1 U CIAP was added to the reaction mixture and incubated at 37 °C for 1 hour. To denature CIAP after dephosphorylation the reaction mixture was incubated at 80 °C for 20 minutes. The DNA was than purified with agarose gel extraction to remove undigested plasmid, salts and proteins for further usage.

4.8.9. Ligation of DNA with T4 DNA ligase

DNA ligase joins two DNA molecules by forming a covalent phosphodiester bond between 3' hydroxyl end and 5' phosphate end of DNA, using ATP as energy donor, result a new DNA molecule.

Procedure:
About 10 fmol digested vector were mixed with 3 – 5 fold molar excess of foreign DNA. 2 µl ligase buffer containing 10 mM ATP (10 x), 1 µl liagse (1 U/µl) and water were added to the mixture, with a final volume of 20 µl and was incubated at 16 °C overnight.

4.8.10. Agarose gel electrophoresis

Electrophoresis uses a mechanism similar to sifting objects through a sieve. In the case of DNA-based gel electrophoresis, an electric field is used to push negatively charged DNA molecules through a gel matrix towards the anode. The migration rate is affected by a number of parameters. The concentration of agarose, voltage, conformation of DNA and size of DNA are the most important parameters. Under the same condition, larger and or DNA molecules move more slowly than smaller or supercoiled DNA because of stronger frictional drag. Ethidium bromide, a fluorescent dye, causes DNA to move slower, because ethidium bromide intercalates and uncoils DNA.

Buffers:

Buffer	Compound	Amount
50 x TAE buffer	Tris	242 g
ddH_2O was added to final 1000 ml.	Acetate acid	57.1 ml
	0.5 M EDTA, pH 8.5	100 ml
1% Agarose gel solution	Agarose	4 g
	1 x TAE buffer	400 ml
The agarose was boiled in a microwave oven with 1 x TAE buffer until the agarose dissolved. The agarose solution was cooled and stored at 60 °C.		
6 x loading buffer	Glycerol	30 % (m/v)
(solved in water)	Bromphenol blue	0.2 % (m/v)
	EDTA, pH 7.5	25 mM

Procedure:

Ethidium bromide was added to the warm agarose solution (1µl 1% solution to 10 ml agarose). The warm agarose solution was stirred to disperse ethidium bromide and than filled into a gel rack with a comb. After the gel has cooled down and solidified, the comb was removed. The gel with the rack was put into a chamber and covered with 1 x TAE buffer. 1 kb DNA ladder (molecular weight marker) was applied into the first gel slot. DNA samples were mixed 1:1 with DNA stop-buffer and applied into the rest gel slots. Electrophoresis was

performed at 120 V for 20 – 30 min. In the present of ethidium bromide DNA was illuminated under ultraviolet light and photographed.

4.8.11. Isolation of DNA from agarose gel

In order to isolate a desired DNA fragment after PCR or restriction enzyme digestion, an agarose gel electrophoresis was carried out. The band corresponding to the desired DNA fragment was cut using a scalpel on a UV-lamp table. To avoid damage of DNA by UV this step should be done very quickly. The DNA fragment was extracted as described in the manual of QIAquick Gel Extraction kit.

4.8.12. Quantitation of DNA

The DNA concentration can be determined by spectrophotometric measurement of the absorption in ultraviolet range. The readings were taken at 260 nm and 280 nm. The OD at 260 nm allows calculation of the concentration of nucleic acid in the sample. An OD of 1 corresponds to 50 µg/ml for double stranded DNA, 40 µg/ml for single stranded DNA and 33 µg/ml for RNA. The ratio of OD_{260}/OD_{280} gives an estimate of the purity of the DNA. Pure DNA preparations are from 1.8. Pure RNA is 2.0. If there is contamination of protein or phenol the ratios will be much lower than 1.8 for DNA.

4.8.13. Site-directed mutagenesis

Site-directed mutagenesis is a technique in which a mutation is created at a defined position in DNA molecule with known sequence. It was carried out by using PCR. Usually plasmid containing a wild-type gene was used as template and two synthetic complementary oligonucleotide primers containing the desired mutation. Because the newly synthesized DNA strands were based on the primers, a mutated plasmid was produced. To avoid unexpected mutations in the plasmid, *Pfu* polymerase was used in PCR because of its high accuracy towards *Taq* polymerase. In order to remove wild-type plasmid, the PCR product was treated by *Dpn*I, which digests specifically methylated DNA. DNA isolated from *E. coli* is methylated, in contrast to the by PCR synthesized DNA.

Procedure:

Plasmid DNA was used as template DNA. Original polymerase buffer was used. The components of the reaction mixture were pipetted together according to the following scheme:

Component	Volume (µl)	Final concentration
10 x polymerase buffer	5	1 x
DMSO	5	10%
dNTPs (10 mM)	4	0.8 mM
Forwards primer (100 pM)	1	2 pM
Reverse primer (100 pM)	1	2 pM
DNA template	variable	appx. 70 ng
Pfu Polymerase	1	1 U
Added H_2O to final volume of 50 µl		

The PCR reaction was carried out in a thermal cycler using the following program:

Step	Temperature (°C)	Duration (min)
Step 1. initial denaturation	96	4
Step 2. denaturation	96	1
Step 3. annealing	55 + 0.6	1
Step 4. extension	72	16
Step 5. final extenstion	72	4

Due to mismatch in binding the template DNA strands, the annealing temperature was set at a low level in the initial cycle, to allow the annealing between primers and template. To avoid amplifying of nonspecific product, the annealing temperature was increased by 0.6 °C each cycle. Step 2 to step 4 was repeated for 15 cycles.

4.8.14. Directional Genome Walking

In order to explore the flanking genomic sequences from a known DNA, we employed a method called “Directional Genome Walking Using PCR”, which was described by M.K.

Reddy *et al*. 2002 (Mishra, Singla-Pareek et al. 2002). One walking step consists of two PCRs. A primary PCR amplified a DNA fragment from one end of the known sequence to the unknown flanking region using a biotinylated primer, which was designed to the known sequence; and walking primers, which bound anywhere in the flanking region. With the biotinylated end the primary PCR product can be easily purified due to it's affinity to streptavidin to eliminate nonspecific PCR products. A secondary PCR was carried out based on the purified primary PCR product to reduce the noise for further sequencing.

Buffers:

Buffer	Compound	Concentration
Binding buffer TEN 100	Tris-HCl	10 mM
pH 7.5	EDTA	1 mM
	NaCl	100 mM
Wash buffer TEN 1000	Tris-HCl	10 mM
pH 7.5	EDTA	1 mM
	NaCl	1 M
1 x PCR buffer	KCl	50 mM
pH 8.3	Tris-HCl	10 mM
	$MgCl_2$	1.5 mM

Procedure:

Two primers were designed to the known sequence. The first primer was biotinylated and lies upstream of the second primer. If the reverse walking was carried out, the biotinylated first primer should lie downstream of the second primer. Both primers were designed to bind to the same DNA strand closed to the 5'end of the known DNA sequence. Four walking primers and nested primer were taken from original publication ("Directional Genome Walking Using PCR"). A primary PCR was performed with genomic DNA as template using the biotinylated first primer and each of four walking primers separately. The composition and the temperature program was the same as for PCR with Dynazyme.

For preparation of streptavidin magnetic beads (Roche), 25 µl of those beads were washed with 100 µl binding buffer TEN100 3 times and resuspended in 50 µl binding buffer. The

PCR product was mixed with the streptavidin magnetic beads. The mixture was incubated at room temperature for 30 minutes to immobilize the biotinylated primary PCR product on the streptavidin magnetic beads. During the incubation the settlement of beads was avoided by gentle shaking or pipetting from time to time. After incubation the beads were washed 3 times with washing buffer TEN1000 in order to remove all nonspecific amplification products. After washing the beads were resuspended in 25 µl 1 x PCR buffer. 1 µl beads was used as template for the second PCR using the primer 2 and nested primer (Mishra, Singla-Pareek et al. 2002). This step should increase the specificity of the PCR product. The PCR was performed under the same conditions as described above including 20 minutes extension at 72°C at the end. All PCR products were loaded onto 1% agarose gel for electrophoresis and subsequent purification. Bright bands about 300~400 bps were cut out and purified using QIAquick® kit for TA-cloning and sequencing. The resulting sequence was compared with the known sequence; an overlapping was required for successful sequence extension. The next primers could be designed to the new sequence for the next step of genomic "walking". After a few steps of walking the total sequence of a gene was completed successfully.

4.9. Biochemical methods

4.9.1. Cell disruption by sonication

The cells were yielded by centrifugation and resuspended in 20 ml 50 mM potassium phosphate buffer pH 7.5 containing 1 mM PMSF. The cells were disrupted by sonication using a Sonifier 250 sonicator. The cell suspension was kept on ice and sonicated for 1 min with output 40 W, 40% work interval and 1 min break. This process was repeated 5 times. The cell extract was centrifuged at 18000 rpm for 20 min at 4°C. The cell lysate was carefully removed in a new tube.

4.9.2 Protein purification with nickel sepharose column

In order to facilitate the purification of a protein, a His-tag (Hengen 1995) was added at the N-terminus of the recombinant protein. The separation is based on the affinity of His-tagged proteins towards nickel ions. Nickel ions were immobilized in a resin and packed in a column.

Since the recombinant protein is the only component with a His-tag, the His-tagged protein bound to the resin, whereas all other proteins pass through the column. The protein was released from the column by high concentration of imidazole, which competed with the His-tag for nickel binding.

Buffers:

Buffer	Compound	Concentration
Loading buffer	KPi	50 mM
	NaCl	300 mM
Wash buffer	KPi	50 mM
	NaCl	300 mM
	imidazole	30 mM
Elution buffer	KPi	50 mM
	NaCl	300 mM
	imidazole	200 mM
All the buffers were adjusted to pH 7.5.		

Procedure:

After a Nickel sepharose column was equilibrated in loading buffer, the cell lysate was loaded onto this column. The flow rate was kept at 1 ml/min during the whole purification process. The process was performed at room temperature. The column was washed with loading buffer (4 column volumes) to remove unbound material. The weak bound host proteins were removed by washing with wash buffer (4 column volumes). The His-tagged P450 protein was eluted with elution buffer.

The resulting fraction was dialysed against potassium phosphate buffer, pH 7.5 using a Vivaspin 15R membrane of 30000 (Vivascience). After that, the sample was concentrated 3-fold and stored at –20°C with 10% glycerol until use.

4.9.3. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

The sodium dodecyl sulphate polyacrylamide gel electrophoresis is used to separate proteins according to their molecular weight to check the purity and molecular weight. SDS denatures and binds to the proteins, and therefore, SDS gives them a negative charge proportional to the mass. Proteins migrate to the anode of gel in an electric field. Polyacrylamide gel works here as a molecular sieve, big molecules run slower through the gel than small molecules due to stronger frictional drag.

Buffers:

Buffer	Compound	
4 x Lower Tris buffer (200 ml) The pH value was adjusted to 8.8 with HCl and added H_2O to a final volume of 200 ml.	Tris SDS	36.34 g 0.8 g
4 x upper Tris buffer (200 ml) The pH value was adjusted to 6.8 with HCl and added H_2O to a final volume of 200 ml.	Tris SDS	12.11 g 0.8 g
Resolving gel (12.5%)	4 x Lower Tris Acrylamide bisacrylamide 30% (v/v) H_2O Ammonium persulfate 10% (m/v) TEMED	2.00 ml 3.33 ml 2.67 ml 40 µl 4 µl
Stacking gel (3.9%)	4 x Upper Tris Acrylamide bisacrylamide 30% (v/v) H_2O Ammonium persulfate 10% (m/v) TEMED	1.0 ml 0.52 ml 2.47 ml 40 µl 4 µl
Staining solution	Coomassie Brilliant Blue	1 g

	Glacial acetic acid	100 ml
	Methanol	300 ml
	H_2O	600 ml
Destaining solution	Methanol	300 ml
	Glacial acetic acid	100 ml
	H_2O	600 ml
5 x Tris glycine buffer	Tris	15 g/l
The pH value was adjusted to 8.3 with	Glycine	72 g/l
HCl	SDS	10% (w/v)
5 x loading buffer	Tris/HCl (pH 6.8)	320 mM
	Glycerol	50% (v/v)
	SDS	10% (m/v)
	β-mercaptoethanol	25% (m/v)
	Bromophenol blue	0.1% (m/v)

LMW standard (Bio-Rad)	
Protein	Molecular weight [kDa]
Phosphorylase B	97.4
Serum albumine	66.2
Ovalbumine	45.0
Carbonic anhydrase	31.0
Trypsin inhibitor	21.5
Lysozyme	14.0

Procedure:

After assembling two glass plates the resolving gel was poured into the gap between the glass plates. The gel was overlaid with isopropanol to ensure a flat surface and to exclude air. Isopropanol was washed off with water after gel had polymerized. The stacking gel was poured onto the polymerized resolving gel and a comb was inserted in the stacking gel solution. After the stacking gel had polymerized, the comb was removed and the gel was put in an electrophoresis chamber, which was filled with electrophoresis buffer. The protein samples were mixed with loading buffer and heated at 95 °C for 5 min. After that the samples

were loaded on stacking gel with the low molecular weight (LMW) standard as reference. For each gel the electrophoresis was performed at 10 mA for 10 min and 25 mA for approximately 50 min. After electrophoresis the gel was covered with staining solution on a shaker overnight and destained three times with destaining solution. Finally the gel was dried at 80 °C in vacuum between filter paper and cellophane foil to be stored.

4.9.4. Determination of cytochrome P450 concentration

Carbon monoxide binds ferrous P450 and induces a shift of the maximum absorbance of the heme to 450 nm. The concentration of P450 protein was determinated by differential spectrophotometry with the coefficient $\varepsilon_{450\text{-}490}$ = 91 $mM^{-1}cm^{-1}$ in CO-difference spectrum, which was developed by Omura and Sato (Omura and Sato 1964a)

$$C_{P450}\ [\mu M] = (A_{450} - A_{490}) \text{ x dil. x } 1000/\ \varepsilon$$

Procedure:
The sample of P450 solution was reduced by adding a few crystals of sodium dithionite and divided in two cuvettes. One of the cuvettes was bubbled with CO. The absorption of the sample between 400 nm and 500 nm was measured using the cuvette without bubbling as reference.

4.9.5. Spectroscopic characterization of the cytochrome P450 monooxygenase

Absorption spectra of the purified P450 monooxygenase were recorded with a spectrophtometer (Ultraspec 3000, Pharmacia Biotech) in optic glass cuvettes of path length 1 cm. Protein concentration was approximate 1 µM. The P450 monooxygenase was reduced by 50 µM NADPH and formation of carbon monoxide complex of P450 monooxygenase was performed by bubbling carbon monoxide gas in reduced enzyme solution for approximately 1 min.

4.9.6. Chromatographic determination of flavin

To release flavin compound, the P450 protein was denatured at 45 °C for 20 min and than was centrifuged at 14000 rpm for 30 min. 2 µl supernatant were analyzed by thin-layer chromatography (TLC), performed on silica gel plate (Kieselgel 60, Merck). Solvent system used for chromatography was *n*-butanol-acetic acid-water (5:2:3, v/v) (Gliszczynska and Koziolowa 1998). Authentic FMN (Riboflavin 5'-monophosphate) and FAD (Flavin adenine dinucleotide) was used as control. The TLC plate was visualized by exposing the plate to UV light 280 nm. The concentration of FMN in supernatant can be determined by fluorescence spectrophotometry in a black 96-well plate. A microtiter plate reader (FLUOstar BMG) was set up as followed: excitation wavelength, 470 nm; emission wavelength, 510 nm. The concentration of FMN was calculated using an extinction-emission coefficient ε = 23740 mM^{-1}, which was determined using FMN standards.

4.10. *In vitro* tests: GC-MS analysis and activity determination

In the following experiments the purified enzyme was used, which was stored in 50 mM potassium phosphate pH 7.5 with 10% glycerol at -20°C with concentration of 95.6 µM.

4.10.1. Reductase domain activity determination

Tests for reductase domain-dependent electron transfer to exogenous electron acceptor cytochrome *c* were performed at room temperature in 50 mM potassium phosphate buffer (pH 7.5) as described by Gustafsson (Gustafsson, Roitel et al. 2004). The K_M values can be graphically determined using standard methods (Fersht 1999).

4.10.2. Determination of deethylation activity

A 190 µl mixture in 96-well plate contained 50 mM potassium phosphate buffer with 1 µM P450. 7-ethoxycoumarin was added as a 150 mM stock solution in DMSO to a final concentration of 1.25 mM. The mixture was incubated at room temperature for 5 min. The same mixture without enzyme was used as negative control. An addition of 10 µl NADPH (6.7 mM) initiated the reaction. The 96 well plate was immediately put in a microtiter plate

reader (FLUOstar, BMG) to measure the fluorescence intensity of 7-hydroxycoumarin. The activity towards 7-ethoxycoumarin was determined by formation rate of 7-hydroxycoumarin. The excitation wavelength and the emission wavelength were 405nm and 460 nm, respectively. The formation rate of 7-hydroxcoumarin was calculated using an extinction-emission coefficient ε = 42970 mM^{-1}, which was determined using 7-hydroxycoumarin standards.

4.10.3. Determination of hydroxylation activity

Substrate was added as a 50 mM - 250 mM stock solution in acetone to a final concentration 1.25 of mM. The mixture was incubated at room temperature for 5 min. Negative controls were the same composition but without enzyme. An addition of 10 µl NADPH (6.7 mM) initiated the reaction. After 2 hours incubation at 25°C the mixture was extracted by thorough vortexing with 350 µl diethyl ether. After separation by centrifugation the diethyl ether was dried over magnesium sulfate anhydrous. 1 µl sample was analyzed by GC-MS (GC-MS QP2010, Shimadzu). Calibration curve was created by measurements of a concentration gradient of the expected products to quantify the results.

GC-MS temperature programs:

Substance	Temperature program
Acenaphthene	150 °C (1 min) – 150-270 °C (20 °C min^{-1})
Alpha-pinene	90 °C (5 min) – 90-250 °C (15 °C min^{-1})
Cyclohexane	50 °C (5 min) – 50-250 °C (20 °C min^{-1})
Ethyl benzene	70 °C (2 min) – 70-250 °C (15 °C min^{-1})
Fluorene	150 °C (1 min) – 150-270 °C (20 °C min^{-1})
Indene	100 °C (5 min) – 100-250 °C (15 °C min^{-1})
Naphthalene	100 °C (5 min) – 100-275 °C (10 °C min^{-1})
Toluene	70 °C (2 min) – 70-250 °C (15 °C min^{-1})
Terpineol	100 °C (1 min) – 100-250 °C (20 °C min^{-1})
m-Xylene	70 °C (2 min) – 70-250 °C (15 °C min^{-1})

4.11. *In vivo* tests

4.11.1. Biotransformation

50 ml overnight culture of each of the tested strains was transferred into 50 ml falcon tube and centrifuged at 4000 rpm and 4 °C with centrifuge of eppendorf 5810 R. The supernatant was discarded and the pellet was resuspended in 25 ml potassium phosphate buffer (50 mM, pH 7.5) containing 1% glucose (dilute from 20% autoclaved stock solution). The substrate was added to 5 mM final concentration. Aliquot of 1 ml sample was taken each hour during 7 hours and then an overnight sample (appx. 16 h). To remove the cells the samples were centrifuged and the supernatants were extracted with 250 µl toluene. The organic phase was dried over sodium sulfate anhydrous. 1 µl sample was analyzed by GC-FID (HRGC MEGA 2 series, FISONS Instruments), using a 60 m ZB-5 column. For cyclohexane, the temperature program used was 110°C for 5 min, the first temperature gradient of 5 °C min^{-1} to 150°C, second temperature gradient of 10 °C min^{-1} to 200°C.

4.11.2. Biotransformation using recombinant *E. coli*

To determine activity towards *m*-xylene and ethyl benzene, the *E. coli* BL21(DE3) expression strain carrying the P450 gene was used as a catalyst. The same *E. coli* strain containing an empty vector pET28(a+) was used as a negative control. The substrate was added to a final concentration of 2 mM in 200 ml overnight induced culture. Cells were incubated for 24 hours at 25°C and 120 rpm. The product formed was extracted in 50 ml dichloromethane and concentrated by evaporation to 1 ml. 1 µl sample was analyzed by GC-MS.

4.11.3. Biotransformation using recombinant *E. coli* in microtiter plate

After expression in deep-well microtiter plate the culture was centrifuged at 4000 rpm and 4 °C for 30 min. The supernatant was discarded and the pellet was resuspended in 190 µl KPi buffer (50 mM, pH 7.5). The cells were transferred into a black microtiter plate. Each well was added 10 µl 7-ethoxycoumarin (25 mM, in DMSO). The 96 well plate was immediately put into a microtiter plate reader (FLUOstar, BMG) to measure the fluorescence intensity of 7-hydroxycoumarin. The measure was carried out in nine cycles with 90 sec interval and

before each cycle the microtiter plate was shaken to mix the sample. The activity towards 7-ethoxycoumarin was determined by formation rate of 7-hydroxycoumarin.

4.12. Protein homology-modelling

Protein structures provide useful insights into the molecular basis of protein function; allow studies of structure-function relationship and effective protein design, such as site-directed mutagenesis. However, it is very time-consuming to obtain experimental structure by X-ray crystallography and nuclear magnetic resonance spectroscopy (NMR). A structure of a protein can be generated from the known structures of its related family members (Schwede, Kopp et al. 2003). This method is called homology-modelling or comparative modelling.

Procedure:
Homology-modelling requires at least one sequence of known 3D-structure with significant similarity to the target sequence. The amino acids sequence of the protein was submitted to NCBI-Blast (http://www.ncbi.nlm.nih.gov/BLAST) to identify 3D-structure template for homology-modelling. Sequences of the hits with highest identity to the target sequence were downloaded and a multiple sequence alignment file was created using ClustalX with default parameters. The advantage of multiple sequence alignment is that the gap region of target sequence can be better assigned than using only pair-wise alignment. The SWISS-MODEL alignment interface allows submitting custom alignment. The alignment file was subsequently submitted to SWISS-MODEL alignment interface (http://swissmodel.expasy.org). A model of target sequence was automatically built by SWISS-MODEL based on the chosen templates and saved in PDB-format. This structure can be further displayed and analyzed in PDB-Viewer (http://expasy.org/spdbv/mainpage.htm).

5. Results

5.1 Screening for strains with monooxygenase activity

For screening of bacterial strains with monooxygenase activity, cyclohexane was chosen as a test substrate due to its high stability, rather low toxicity, and industrial importance (Steyer and Sundmacher 2004). Cyclohexane can be hydroxylated by P450 in the presence of oxygen, leading to cyclohexanol (Figure 5.1).

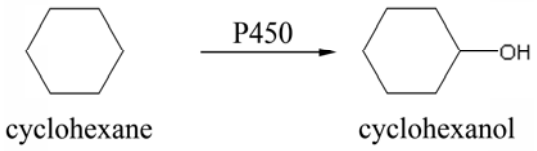

Figure 5.1: P450 catalyzes hydroxylation of cyclohexane to cyclohexanol.

Seven bacterial strains were tested for their oxidation activity towards cyclohexane (Table 5.1). They originate from three genuses: *Rhodococcus*, *Pseudomonas* und *Sphingomonas.* Microorganisms of genus *Rhodococcus* are widely spread in nature and known to transform a vast range of xenobiotics (Finnerty 1992). The typical *Pseudomonas* bacteria in nature can be found in a biofilm; they are ubiquitous in soil and water, and have been shown to grow in the presence of alkanes (van Beilen, Wubbolts et al. 1994). Due to their biodegradative and biosynthetic capabilities, several *Sphingomonas* strains have also been tested in this study. Furthermore, several cytochrome P450 monooxygenases have already been identified and isolated from these three genera (Hedegaard and Gunsalus 1965; Tyson, Lipscomb et al. 1972); (Ropp, Gunsalus et al. 1993; Fruetel, Mackman et al. 1994).

All strains grew in LB medium very well. The OD_{600} of overnight cultures reached approximately 2. The cells were then centrifuged and resuspended in potassium phosphate buffer, pH 7.5. Because the hydroxylation by P450 requires NAD(P)H, the cells need glucose to regenerate NAD(P)H. Therefore, 1% glucose was added in the buffer as an energy source. (Walton and Stewart 2004)

Table 5.1: Summary of all tested bacterial strains in this study. (+) indicates activity. (-) indicates no activity.

Strain	Activity towards cyclohexane
Rhodococcus erythropolis DSM 43066	+
Rhodococcus ruber DSM 44319	+
Pseudomonas citronellolis DSM 50332	-
Pseudomonas fluorescens DSM 50090	-
Pseudomonas savastanoi DSM 50267	-
Sphingomonas echinoides DSM 1805	-
Rhodococcus rhodochrous NCIMB 11216	-

GC analysis of the products of the biotransformation revealed only two *Rhodococcus* strains with activity towards cyclohexane, namely *Rhodococcus erythropolis* DSM 43066 and *Rhodococcus ruber* DSM 44319. Cyclohexanone was the only product detected (data not shown). Therefore, we assumed that the microorganism contains cyclohexane hydroxylase, which hydroxylates cyclohexane to cyclohexanol and further oxidizes cyclohexanol to cyclohexanone via gem-diol formation. Another possibility is an additional dehydrogenase which catalyzes the second oxidation of cyclohexanol to cyclohexanone (Figure 5.2). The *Pseudomonas citronellolis* DSM 50332 and *Pseudomonas fluorescens* DSM 50090 strains did not display any activity towards cyclohexane. *Pseudomonas savastanoi* DSM 50267, *Sphingomonas echinoides* DSM 1805 and *Rhodococcus rhodochrous* NCIMB 11216 strains did not survive in the presence of cyclohexane.

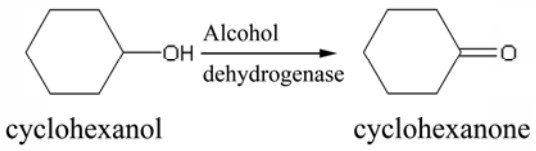

Figure 5.2: Cyclohexanol may be further oxidized to cyclohexanone by an alcohol dehydrogenase from the *Rhodococcus* strains.

5.2. Gene isolation

5.2.1 Strategy

As genomes of both strains *Rhodococcus ruber* DSM 44319 and *Rhodococcus erythropolis* DSM 43066 are unknown, the genes responsible for detected activity should be identified and isolated. Monooxygenation is generally catalyzed by flavin monooxygenases or cytochrome P450s. From the literature is known that many Rhodococci have numerous P450 monooxygenase genes. On this background, we decided to construct degenerate primers, based on the sequences of two known P450 genes, CYP116 (Nagy, Schoofs et al. 1995) and P450RhF (Roberts, Grogan et al. 2002). The strategy for isolation of homologous genes is demonstrated in the following schema (Figure 5.3).

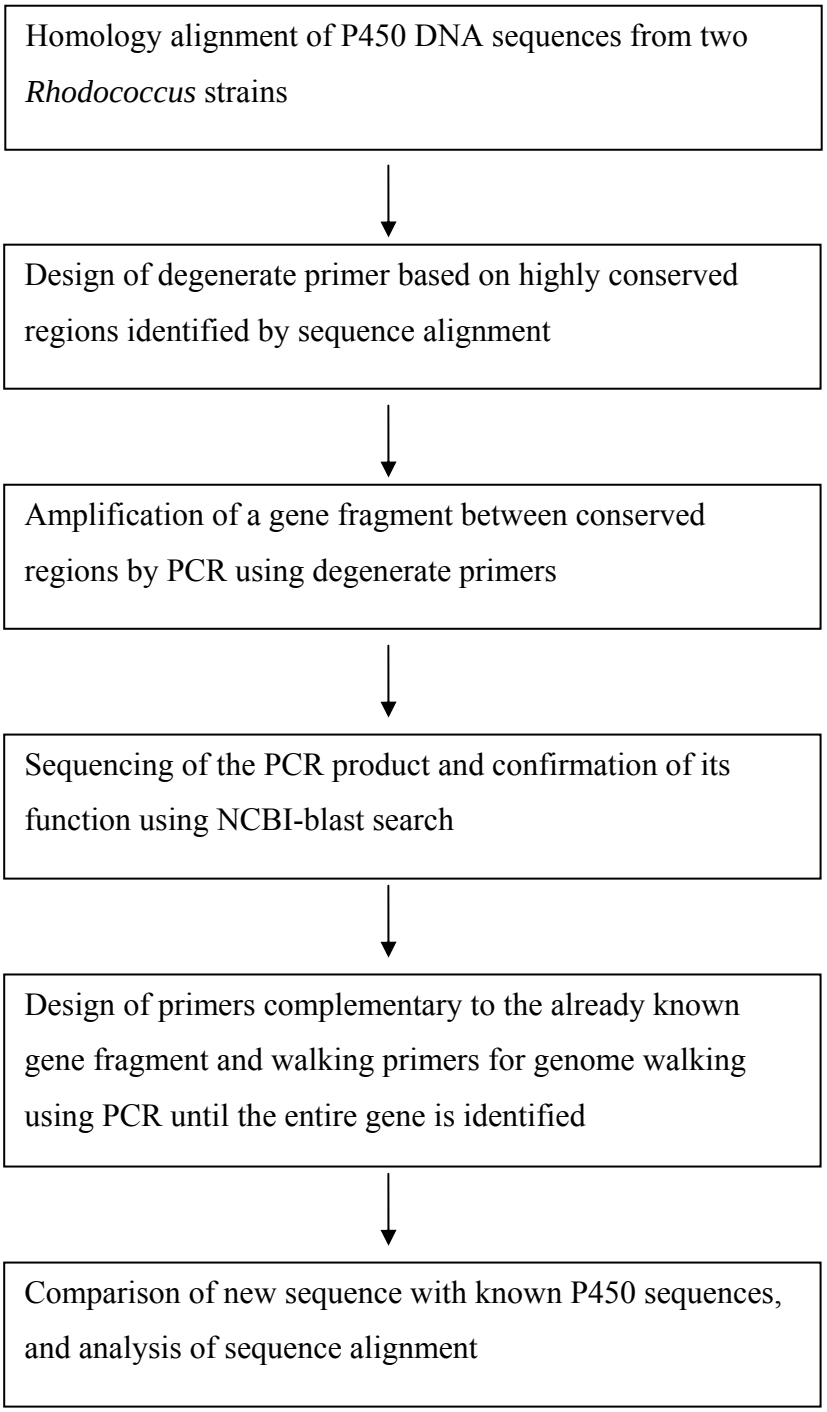

Figure 5.3: Strategy for identification of a new P450 gene.

5.2.2 Primer design

DNA sequences of CYP116 and P450RhF were aligned by ClustalX using defined parameters (Figure 5.4). Four degenerate primers were designed complementary to the most similar regions. Two forward primers should bind about 200 bp and 340 bp downstream of initial codon, and two reverse primers - about 900 bp and 1000 bp downstream of initial codon (Table 5.2). The oxygen-binding site is also highly conserved (highlighted in yellow, Figure 5.4); however it is located in the middle of the gene and therefore, not convenient for primer design.

```
CYP116          --------------------ATGACCGTCGATCACGCGCCGGAAGGCGTGAAAAGCC--C
p450RhF         ATGAGTGCATCAGTTCCGGCGTCGGCGCCGG-CGTGTCCCGTCGACCACGCGGCCCTGGC 60
                            *   ** **  *  *  ***     *  *     *   *

CYP116          GACCGGATGCCCCGTTTCAGGTATGGCGGCAGACTTCGATCCGTTCCGGGGTGCCTACCA
p450RhF         GGGCGGCTGCCCGGTGTCGGCGAACGCCGCGGCGTTCGATCCGTTCGGTTCCGCGTACCA 120
                *  *** ***** ** ** *  *  ** ** *  ************ *    ** *****

CYP116          GGTCGATCCATCCTCGTCGCTGCGGCAGGCCCGCAAAGACGAACCAGTGTTCTTCAGTCC
p450RhF         GACCGATCCGGCCGAGTCGCTGCGCTGGTCCCGCGACGAGGAGCCGGTGTTCTACAGCCC 180
                *  ******  **  *********   * ***** * ** ** ** ******* *** **

CYP116          ACTGCTCGACTACTGGGTGGTCACGCGCTACGAGGACATCAAGCAGATCTTCAA-GACAC
p450RhF         CGAACTCGGCTACTGGGTCGTCACCCGGTACGAGGATGTGAAGGCGGTGTTCCGCGACAA 240
                    **** ********* ***** ** ********  * ***  * * ***   ****

CYP116          CGTCGGTGTTCTCCCCGTCGATCACGGTCGACCAGATCACCCCGATCAGTGACGAGGCGC
p450RhF         CATCC-TGTTCTCGCCGGCGATCGCGCTGGAGAAGATCACTCCCGTCTCGGCGGAGGCCA 300
                * **  ******* *** ***** ** * **  ******* **  **   *  *****

CYP116          TGCAGATCCTCGGCAGCTATCAGTTCGCTGCCGGCCGGATGCTCGTGAACGAAGACGAGC
p450RhF         CCGCCACCCTCGCCCGGTACGACTACGCCATGGCCCGGACCCTCGTGAACGAGGACGAGC 360
                     * ***** * * **  * * ***    * *****  *********** *******

CYP116          CGATCCACACCGAACGTCGCCGCCTGCTGATGCAACCGTTCGAGGCCGATAACGTCGCCA
p450RhF         CCGCCCACATGCCGCGCCGCCGCGCGCTCATGGATCCGTTC-ACCCCGAAGGAACTGGCG 420
                *   *****     ** ******  *** *** * ****** *  ****       * *

CYP116          CGCT-CGAACCGAAGATCCGCGAGGTCGTCAACACCTATCTGGACCGGGTCATCAAGGAT
p450RhF         CACCACGAGGCGATGGTGCGACGGCTCACGCGCGAATACGTCGACCGCTTCGTCGAATCC 480
                * *  ***  *** * * **   * **     *   **  * *****  ** ** *

CYP116          GGCCGGGCCGACCTGATCGGCGATCTGCTCTACGAAGTTCCGTGCATCGTCGCGCTCATC
p450RhF         GGCAAGGCCGACCTGGTGGACGAGATGCTGTGGGAGGTTCCGCTCACCGTCGCCCTGCAC 540
                ***  ********** * * ***  **** *  ** ******  ** ****** **   *

CYP116          TTCCTCGGCGTGCCCGACGAGGACATCGAGACCTGCCGCCAATACGGGATGCAGCAGACC
p450RhF         TTCCTCGGCGTGCCGGAGGAGGACATGGCGACGATGCGCAAGTACTCGATCGCGCACACC 600
                ************** ** ******** * ***    *** * ***  ***   *** ***

CYP116          CTGTTCACCTGGGGACACCCGACCGGAGACGAACAAACTCGGGTCGCGACCGGGATGGGG
p450RhF         GTGAACACCTGGGGCCGCCCCGCGCCCGAGGAGCAGGTGGCCGTCGCCGAGGCGGTCGGC 660
                 **  ********* * ***  *    ** ** **       *****    * * * **

CYP116          AAGTTCTGGGAGTTCGCCGGCGGACTGGTCGACAAGCTCAAGGCCGACCCGAACGCGAAG
p450RhF         AGGTTCTGGCAGTACGCGGGCACGGTGCTCGAGAAGATGCGGCAGGACCCGTCGGGACAC 720
                * ******* *** *** ***    ** **** *** *   *   ******   *   *
```

```
CYP116          GGGTGGATTCCTCACGCGATCGAGATGCAGCGGCAGCACCCCGACCTCTTCGACGACAAC
p450RhF         GGCTGGATGCCCTACGGGATCCGCAAGCAGCGGGAGATGCCGGACGTCGTCACCGACTCC 780
                ** ***** **  *** ****   * ******* **   ** *** ** **  **** *

CYP116          TACCTGCAGAACATCATGTTCGGTGGTGTGTTTGCCGCGCACGAAACCACCACCAACGCC
p450RhF         TACCTGCACTCGATGATGATGGCCGGCATCGTCGCCGCGCACGAGACCACGGCCAACGCG 840
                ********    ** *** * *  **  *  * ********** *****  *******

CYP116          ACCGGTAACGCGTTCCGTACCCTGCTCGAGAACCGCAGTTCCTGGGACGAGATCTGCGCC
p450RhF         TCCGCGAACGCGTTCAAGCTGCTGCTCGAGAACCGCGCGGTGTGGGAGGAGATCTGCGCG 900
                 ***  *********      ***************      ***** ***********

CYP116          GACCCCACGCTGATCCCCAAGGCCATCGAGGAATGCCTCCGCTACAGCGGATCCGTCGTC
p450RhF         GATCCGTCGCTGATCCCCAACGCCGTCGAGGAGTGCCTGCGCCACTCCGGGTCCGTGGCG 960
                ** **  ************* *** ******* ***** *** **  *** ***** *

CYP116          GCCTGGCGCCGCAAGGCCGTGGTGGACACCACTGTCGGCGAAGTCGACATCCCGGCCGGC
p450RhF         GCGTGGCGACGGGTGGCCACCGCCGACACCCGCATCGGCGACGTCGACATCCCCGCCGGC 1020
                ** ***** **   ****   *  ******    ******* ********** ******

CYP116          GGTCGACTGCTGATCGTCATGGCATCGGCGAACCGCGACGACTCGATGTTCCCCGAACCC
p450RhF         GCCAAGCTGCTCGTCGTCAACGCGTCCGCCAACCACGACGAGCGCCACTTCGAGCGCCCC 1080
                *     *****  ******  ** ** ** **** ******       ***      ***

CYP116          GACGACTTCGACATCCACCGCGGAAATGCCCAGCGCCACTTGACATTCGGCATCGGCAGC
p450RhF         GACGAGTTCGACATCCGGCGCCCGAACTCGAGCGACCATCTCACCTTCGGGTACGGCAGC 1140
                ***** **********  ***   **  *      ***  * ** *****   *******

CYP116          CACACCTGCCTGGGCGCGACACTGGCGCGGTTGGAGATGAAGGTCTTCCTCGAAGAGGTC
p450RhF         CACCAGTGCATGGGCAAGAACCTGGCCCGCATGGAGATGCAGATCTTCCTCGAGGAACTC 1200
                ***   *** *****  **  ***** **  ******** ** ********** **  **

CYP116          TCGCGACGCCTGCCGCACATGTCACTCGTTGCCGGACAGGAATTCTCCTACCTTCCCAAC
p450RhF         ACCACGCGGCTTCCCCACATGGAACTCGTACCCGATCAGGAGTTCACCTACCTGCCGAAT 1260
                 *    ** ** ** ******  ******  ***  ***** *** ******* ** **

CYP116          ACCTCGTTCCGGGGACCCGAGCACGTGTTGGTCGAGTGGGATCCCCAGCAGAATCCCGTC
p450RhF         ACGTCCTTCCGCGGACCCGACCACGTGTGGGTGCAGTGGGATCCGCAGGCGAATCCCGAG 1320
                ** ** ***** ******** ******* ***  ********** ***  ********

CYP116          CCCGCCGA-CCGGCCCTGA----------------------------------------
p450RhF         CGCACCGATCCTGCTGTGCTGCACCGGCATCAACCGGTCACCATCGGAGAACCCGCCGCC 1380
```

Figure 5.4: DNA sequence alignment between CYP116 and P450RhF, generated by ClustalX. The stars indicate identical nucleotides. The degenerated primers were based on the sequences which are highlighted in blue. The conserved cysteine is highlighted in green. The oxygen binding site is highlighted in yellow. Because P450RhF is a fusion protein, longer than CYP116, only the heme domain of P450RhF was aligned to CYP116.

Table 5.2: Sequences of degenerated primers

Name	Sequence 5' to 3'
Cyp_F1	CTACTGGGTSGTCACSCGSTACGA
Cyp_F2	CTCGTGAACGARGACGA
Cyp_R1	GCAYTCCTCGAYGGCSTTGGGGAT
Cyp_R2	GGGATGTCGACKTCGCCGA

(S = G or C; Y = C or T; R = A or G; K = G or T)

5.2.3 Homology search

Genomic DNAs of *Rhodococcus ruber* DSM 44319 and *Rhodococcus erythropolis* DSM 43066 were isolated and used as templates for amplifying P450-like DNA fragments. The DNA concentrations were 165 ng/µl and 75 ng/µl for *Rhodococcus erythropolis* DSM 43066 and *Rhodococcus ruber* DSM 44319 respectively (Figure 5.5).

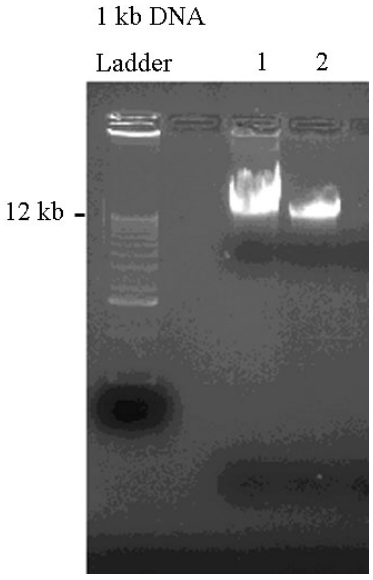

Figure 5.5: Genomic DNA of *Rhodococcus erythropolis* DSM 43066 (lane 1) and *Rhodococcus ruber* DSM 44319 (lane 2). The DNA was fragmented and the size was about 12 kb.

Four PCRs were performed using several primer combinations where each forward primer was combined with each reverse primer. The expected sizes of PCR products are shown in Table 5.3.

Table 5.3: The combinations of primer and expected sizes of PCR products

		Forward primers	
		Cyp_F1	Cyp_F2
Reverse primers	Cyp_R1	740 bp	600 bp
	Cyp_R2	800 bp	660 bp

The PCR products were amplified from genomic DNA from *Rhodococcus ruber* DSM 44319 by all the primer combinations but not from *Rhodococcus erythropolis* DSM 43066. All primer combinations resulted in products of expected length. However, PCRs with primer R2 produced also non-specific PCR products as seen in Figure 5.6, lane 2, 3. Thus, the PCR product obtained by degenerate PCR primer F1 and R1 was chosen for sequencing.

In order to obtain a high quality sequence, highly pure DNA is required for sequencing. Therefore, the PCR product was primarily cloned into the vector PCR®2.1-TOPO® by using the TA-Cloning® kit. The sequencing was carried out using purified plasmid DNA containing the target DNA fragment. The sequencing was run in two directions using M13 forward and M13 reverse primers, which are located upstream and downstream of the insert in the vector. The resulting sequence PCR-F1R1 was compared with CYP116 and P450RhF. (Figure 5.7) The sequence alignment revealed the nucleotide sequence identity between the DNA fragment from *Rhodococcus ruber* DSM 44319 and the P450 monooxygenase domain of P450RhF from *Rhodococcus* sp. NCIMB 9784 of about 96% in a pair-wise alignment. The sequence identity between the DNA fragment from *Rhodococcus ruber* DSM 44319 and CYP116 from *Rhodococcus* sp. NI86/21 was about 66% in a pair-wise alignment. This high homology enables assignment of the new enzyme to the CYP116 family.

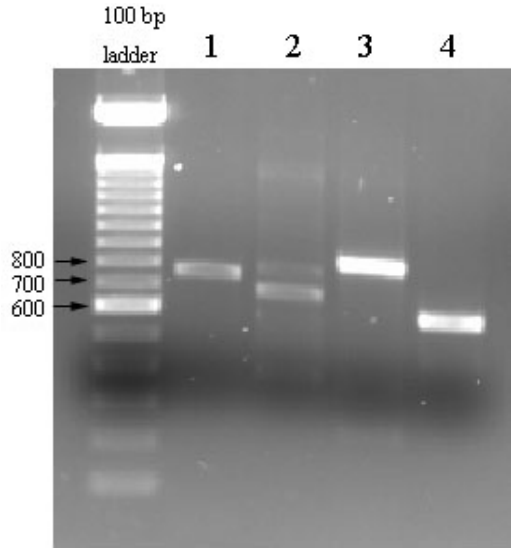

Figure 5.6: The PCR products from four primer combinations. Lane 1, 2, 3 and 4 were from the combinations F1-R1, F2-R2, F1-R2 and F2-R1, respectively. The DNA size of all major bands is consistent with expected sizes.

```
p450RhF         GAACTCGGCTACTGGGTCGTCACCCGGTACGAGGATGTGAAGGCGGTGTTCCGCGACAAC 240
PCR-F1R1        --------CTACTGGGTGGTCACGCGGTACGAGGATGTGAAGGCGGTGTTCCGCGACAAC
CYP116          CTGCTCGACTACTGGGTGGTCACGCGCTACGAGGACATCAAG-CAGATCTTCAAGACACC
                         ********* ***** ** ********  * *** * *   * *  **** *

p450RhF         ATCC-TGTTCTCGCCGGCGATCGCGCTGGAGAAGATCACTCCCGTCTCGGCGGAGGCCAC 300
PCR-F1R1        CTCG-TGTTCTCACCGGCCATCGCCCTCGAGAAGATCACCCCGGTCTCCGAGGAGGCCAC
CYP116          GTCGGTGTTCTCCCCGTCGATCACGGTCGACCAGATCACCCCGATCAGTGACGAGGCGCT
                 **  ******* *** * *** *  * **  ******* **  **   *  *****

p450RhF         CGCCACCCTCGCCCGGTACGACTACGCCATGGCCCGGACCCTCGTGAACGAGGACGAGCC 360
PCR-F1R1        CGCCACCCTCGCCCGCTACGACTACGCCATGGCCCGGACCCTCGTGAACGAGGACGAGCC
CYP116          GCAGATCCTCGGCAGCTATCAGTTCGCTGCCGGCCGGATGCTCGTGAACGAAGACGAGCC
                    * ***** * * **  * * ***    * *****  ********** ********

p450RhF         CGCCCACATGCCGCGCCGCCGCGCGCTCATGGATCCGTTC-ACCCCGAAGGAACTGGCGC 420
PCR-F1R1        CGCCCACATGCCGCGCCGCCGCGCACTCATGGACCCGTTC-ACCCCGAAGGAACTGGCGC
CYP116          GATCCACACCGAACGTCGCCGCCTGCTGATGCAACCGTTCGAGGCCGATAACGTCGCCAC
                   *****     ** ******   ** *** * ****** *  ****       * * *

p450RhF         ACCACGAGGCGATGGTGCGACGGCTCACGCGCGAATACGTCGACCGCTTCGTCGAATCCG 480
PCR-F1R1        ACCACGAGGCGATGGTGCGACGGCTCACGCGCGAATACGTCGACCGCTTCGTCGAATCCG
CYP116          GCT-CGAACCGAAGATCCGCGAGGTCGTCAACACCTATCTGGACCGGGTCATCAAGGATG
                 *  ***  *** * * **   * **     *   **  * *****  ** ** *    *

p450RhF         GCAAGGCCGACCTGGTGGACGAGATGCTGTGGGAGGTTCCGCTCACCGTCGCCCTGCACT 540
PCR-F1R1        GCAAGGCCGACCTGGTGGACGAGATGCTGTGGGAGGTACCGCTCACCGTCGCCCTGCACT
CYP116          GCCGGGCCGACCTGATCGGCGATCTGCTCTACGAAGTTCCGTGCATCGTCGCGCTCATCT
                **  ********** * * ***  **** *  ** ** ***  ** ****** **   **

p450RhF         TCCTCGGCGTGCCGGAGGAGGACATGGCGACGATGCGCAAGTACTCGATCGCGCACACCG 600
PCR-F1R1        TCCTCGGCGTGCCGGAGGAGGACATGGCGACGATGCGCAAGTACTCGATCGCCCACACCG
CYP116          TCCTCGGCGTGCCCGACGAGGACATCGAGACCTGCCGCCAATACGGGATGCAGCAGACCC
                ************* ** ******* * ***    *** * ***  ***    ** ***

p450RhF         TGAACACCTGGGGCCGCCCCGCGCCCGAGGAGCAGGTGGCCGTCGCCGAGGCGGTCGGCA 660
PCR-F1R1        TGAACACCTGGGGCCGCCCCGCGCCCGAGGAGCAGGTCGCCGTCGCCGAGGCGGTCGGCA
```

```
CYP116          TGTTCACCTGGGGACACCCGACCGGAGACGAACAAACTCGGGTCGCGACCGGGATGGGGA
                **  ********* * ***  *    ** ** **       *****    * * * ** *

p450RhF         GGTTCTGGCAGTACGCGGGCACGGTGCTCGAGAAGATGCGGCAGGACCCGTCGGGACACG 720
PCR-F1R1        GGTTCTGGCAGTACGCGGGCACGGTGCTCGAGAAGATGCGCCAGGACCCCTCGGGGCACG
CYP116          AGTTCTGGGAGTTCGCCGGCGGACTGGTCGACAAGCTCAAGGCCGACCCGAACGCGAAGG
                 ******* *** *** ***    ** **** *** *       *****    *   * *

p450RhF         GCTGGATGCCCTACGGGATCCGCAAGCAGCGGGAGATGCCGGACGTCGTCACCGACTCCT 780
PCR-F1R1        GCTGGATGCCCTACGGGATCCGCATGCAGCAGCAGATGCCGGACGTCGTCACCGACTCCT
CYP116          GGTGGATTCCTCACGCGATCGAGATGCAGCGGCAGCACCCCGACCTCTTCGACGACAACT
                * ***** **  *** ****   * ***** * **   ** *** ** **  ****  **

p450RhF         ACCTGCACTCGATGATGATGGCCGGCATCGTCGCCGCGCACGAGACCACGGCCAACGCGT 840
PCR-F1R1        ACCTGCACTCGATGATGATGGGCGGCATCGTCGCCGTGCACGAGACCACGGCCAACGCGT
CYP116          ACCTGCAGAACATCATGTTCGGTGGTGTGTTTGCCGCGCACGAAACCACCACCAACGCCA
                *******    ** *** * *  **  *  * **** ****** *****  *******

p450RhF         CCGCGAACGCGTTCAAGCTGCTGCTCGAGAACCGCGCGGTGTGGGAGGAGATCTGCGCGG 900
PCR-F1R1        CCGCGAACGCGTTCAAGCTGCTGCTCGAGAACCGCCCGGTGTGGGAGGAGATCTGCGCGG
CYP116          CCGGTAACGCGTTCCGTACCCTGCTCGAGAACCGCAGTTCCTGGGACGAGATCTGCGCCG
                ***  *********      ***************      ***** *********** *

p450RhF         ATCCGTCGCTGATCCCCAACGCCGTCGAGGAGTGCCTGCGCCACTCCGGGTCCGTGGCGG 960
PCR-F1R1        ATCCGTCGCTGATCCCCAACGCCGTCGAGGAGTGC------------------------
CYP116          ACCCCACGCTGATCCCCAAGGCCATCGAGGAATGCCTCCGCTACAGCGGATCCGTCGTCG
                * **  ************* *** ******* ***
```

Figure 5.7: DNA sequence alignment between PCR product obtained by primers F1 and R1 (PCR-F1R1), CYP116 from *Rhodococcus* sp. NI86/21 (CYP116) and P450RhF from *Rhodococcus* sp. NCIMB 9784 (p450RhF), generated by ClustalX. The stars indicate the identical nucleotides in all three DNA sequences. Because the PCR product was smaller in size than CYP116 and P450RhF, the residual sequences of CYP116 and P450RhF DNA sequences are not shown.

3.2.4 Genome Walking

After identification of the 740 bp DNA fragment, the entire new P450 gene was found by directional genome walking using PCR (Mishra, Singla-Pareek et al. 2002). The strategy is demonstrated in Figure 5.8. Two locus-specific primers were designed to the known DNA fragment close to the 3'-end of the fragment to obtain new sequences as long as possible. The non-biotinylated nested locus-specific primer P2 was located downstream of the biotinylated locus-specific primer P1. The primary four PCRs were carried out using genomic DNA as template and the biotinylated primer P1 combined with each of the walking primers taken as in the publication by Mishra et al. Due to a universal 3'-end, the walking primers theoretically bound at least each 256 bases in the genomic DNA. The resulting PCR products contained not only the target PCR product, but also many non-specific PCR products (Figure

5.9). Since biotin can be tightly bound to streptavidin, the biotinylated target PCR products were mixed with streptavidin-coated beads and then washed. Due to the strong affinity between the biotinylated PCR product and the streptevidin, the PCR product remained on the streptavidin-coated beads after washing while the non-biotinylated products were removed. To amplify a sufficient amount of highly pure DNA for TA-cloning and sequencing, a secondary PCR was carried out using DNA associated streptavidin-coated beads as template and non-biotinylated nested primer P2 and walking primer-2 without the universal 3'-end. To avoid interference by non-specific PCR products in the subsequent TA-cloning, the PCR product was purified using a Gel-Purification Kit. Only the brightest bands were cut out and purified (Figure 5.10). Because the DNA was amplified by *Taq* DNA polymerase, the 3'-end was adenylated. The purified DNA was inserted into the TA-cloning TOPO® vector and transformed into *E. coli*. After cultivation the resulting plasmid was isolated and the insert was sequenced. The newly obtained DNA sequence was checked by comparison with DNA sequences of CYP116 and P450RhF. The new DNA sequence with significant identity to each of the two genes was added to the original sequence. In this study, usually up to 400 bp new DNA sequence could be identified in one step. New primers were designed to the newly identified DNA sequence for the next step of genome walking. For backward genome walking (from 5' end of the known DNA fragment upstream), the primers were designed to the complementary DNA sequence and the non-biotinylated primer P2 located upstream of the biotinylated primer P1 close to the 5' end. After five walking steps the total sequence of the gene was completed successfully.

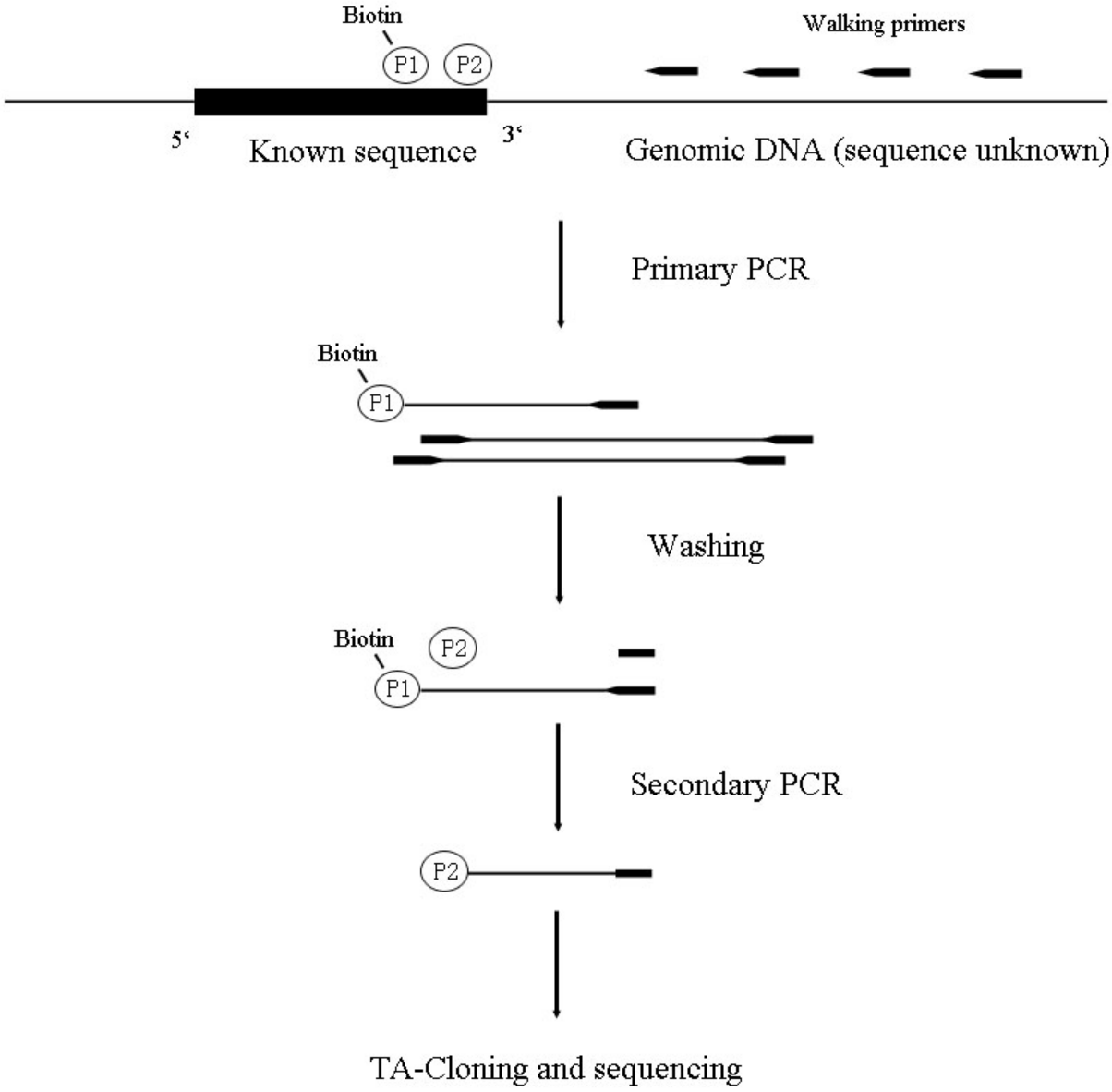

Figure 5.8: The concept of genome walking using PCR. P1 - biotinylated locus-specific primer; P2 – nested locus-specific primer.

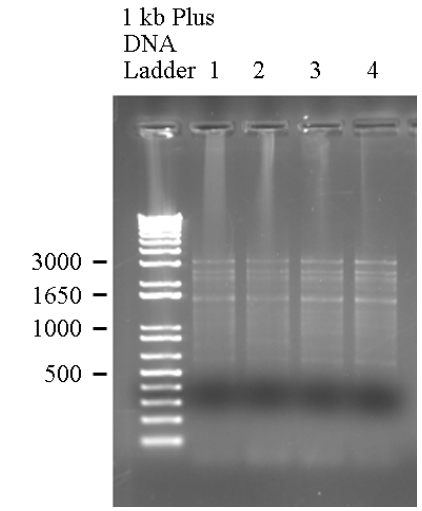

Figure 5.9: The primary PCRs were carried out using genomic DNA as template. The PCR products in lanes 1 to 4 were amplified using the same biotinylated primer and four different walking primers. Since walking primers can bind almost everywhere to the DNA template, products of any size were possible and the PCR products smeared on agarose gel between 4000 bp and 500 bp. The specific products are not seen on the agarose gel.

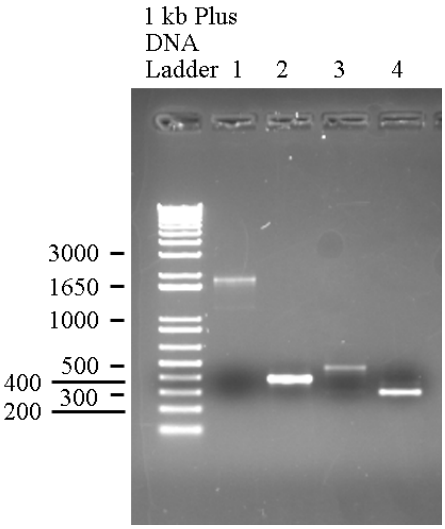

Figure 5.10: The secondary PCRs were carried out using the washed PCR products from the four primary PCRs as templates. Lanes 1 to 4 show the products from four independent PCRs. PCR products of different sizes distributed from 2000 bp to 200 bp were amplified. Only the brightest bands were chosen for further purification and sequencing.

5.3 Sequence analysis

The P450 sequence obtained was translated into a protein of 771 amino acids. A BLAST search using the BLOSUM62 standard scoring matrices (Altschul 1991) (www.ncbi.nlm.nih.gov/BLAST) revealed that this protein is a natural fusion protein consisting of a heme domain, a flavin-reductase domain, and a ferredoxin domain. This protein organization is the same as in P450RhF from *Rhodococcus* sp. NCIMB 9784 (Roberts, Celik et al. 2003). The domain architecture was graphically demonstrated by Conserved Domain Database (CDD) (Figure 5.11) (Marchler-Bauer, Anderson et al. 2005). The overall nucleotide and amino acids sequence identity between the new P450 enzyme and P450RhF is 93% in pair-wise alignment. No doubt, the newly identified protein is a cytochrome P450 monooxygenase, and is also a natural fusion protein like P450RhF. To be sure, the heme domain of the new cytochrome P450 was compared to CYP116 and the heme domain of P450RhF (Figure 5.12). The protein sequences of the reductase and ferredoxin domains of the new cytochrome P450 were compared to P450RhF and an authentic and well studied phthalate dioxygenase reductase from *Pseudomonas cepacia* (Correll, Batie et al. 1992). The amino acids sequence identity between the new cytochrome P450 reductase domain and phthalate dioxygenase reductase is 35%. The conserved motifs were identified in the sequence alignment (Figure 5.13). This alignment is an evidence for the function of the new cytochrome P450 reductase domain. De Mot and Parret also identified four putative fusion P450 enzymes of this type in three pathogenic *Burkholderia* species and in a heavy metal-resistant bacterium *R. metallidurans* (De Mot and Parret 2002). The new P450 sequence was deposited in the GenBank database with the accession number AY957485 and was named CYP116B3 by David Nelson.

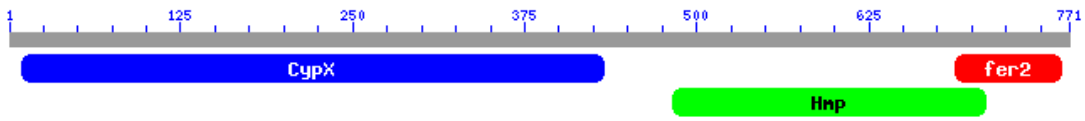

Figure 5.11: Putative conserved domains have been detected by Conserved Domain Database (CDD). The upper bar indicates the entire protein of CYP116B3. The blue bar suggests a cytochrome P450 domain that begins from the N-terminus and ends after 430 amino acid residues. The green bar suggests a possible flavodoxin reductase domain from 475 to 675 amino acid residues. The red bar suggests a possible ferredoxin domain from 680 amino acid residue to C-terminus.

```
p450RhF         MSASVPASAPACPVDHAALAGGCPVSANAAAFDPFGSAYQTDPAESLRWSRDEEPVFYSP 60
CYP116B3        MSASVPASA--CPVDHAALAGGCPVSTNAAAFDPFGPAYQADPAESLRWSRDEEPVFYSP
cyp116          MTVDHAPEG-------VKSPTGCPVSGMAADFDPFRGAYQVDPSSSLRQARKDEPVFFSP
                *:.. ....       .  . *****  ** ****  ***.**:.*** :*.:****:**

p450RhF         ELGYWVVTRYEDVKAVFRDNILFSPAIALEKITPVSAEATATLARYDYAMARTLVNEDEP 120
CYP116B3        ELGYWVVTRYEDVKAVFRDNLVFSPAIALEKITPVSEEATATLARYDYAMARTLVNEDEP
cyp116          LLDYWVVTRYEDIKQIFKTPSVFSPSITVDQITPISDEALQILGSYQFAAGRMLVNEDEP
                 *.*********:* :*:   :***:*:::***:* **   *. *::* .* *******

p450RhF         AHMPRRRALMDPFTPKELAHHEAMVRRLTREYVDRFVESGKADLVDEMLWEVPLTVALHF 180
CYP116B3        AHMPRRRALMDPFTPKELAHHEAMVRRLTREYVDRFVESGKADLVDEMLWEVPLTVALHF
cyp116          IHTERRRLLMQPFEADNVATLEPKIREVVNTYLDRVIKDGRADLIGDLLYEVPCIVALIF
                 *  *** **:** ..::*  *. :*.:.. *:**.::.*:***:.::*:***  *** *

p450RhF         LGVPEEDMATMRKYSIAHTVNTWGRPAPEEQVAVAEAVGRFWQYAGTVLEKMRQDPSGHG 240
CYP116B3        LGVPEEDMATMRKYSIAHTVNTWGRPAPEEQVAVAEAVGRFWQYAGTVLEKMRQDPSGHG
cyp116          LGVPDEDIETCRQYGMQQTLFTWGHPTGDEQTRVATGMGKFWEFAGGLVDKLKADPNAKG
                ****:**: * *:*.: :*: ***:*: :**. ** .:*:**::** :::*:: **..:*

p450RhF         WMPYGIRKQREMPDVVTDSYLHSMMMAGIVAAHETTANASANAFKLLLENRAVWEEICAD 300
CYP116B3        WMPYGIRMQQQMPDVVTDSYLHSMMMAGIVAAHETTANASANAFKLLLENRPVWEEICAD
cyp116          WIPHAIEMQRQHPDLFDDNYLQNIMFGGVFAAHETTTNATGNAFRTLLENRSSWDEICAD
                *:*:.*. *:: **:. *.**:.:*:.*:.******:**:.***: *****. *:*****

p450RhF         PSLIPNAVEECLRHSGSVAAWRRVATADTRIGDVDIPAGAKLLVVNASANHDERHFERPD 360
CYP116B3        PSLIPNAVEECLRHSGSVAAWRRVATTDTRIGDVDIPAGAKLLVVNASANHDERHFDRPD
cyp116          PTLIPKAIEECLRYSGSVVAWRRKAVVDTTVGEVDIPAGGRLLIVMASANRDDSMFPEPD
                *:***:*:*****:****.**** *..** :*:******.:**:* ****:*:  * .**

p450RhF         EFDIRRPNSSDHLTFGYGSHQCMGKNLARMEMQIFLEELTTRLPHMELVPDQEFTYLPNT 420
CYP116B3        EFDIRRPNSSDHLTFGYGSHQCMGKNLARMEMQIFLEELTTRLPHMELVPDQEFTYLPNT
cyp116          DFDIHRGNAQRHLTFGIGSHTCLGATLARLEMKVFLEEVSRRLPHMSLVAGQEFSYLPNT
                :***:* *:. ***** *** *:* .***:**::****:: *****.**..***:*****

p450RhF         SFRGPDHVWVQWDPQANPERTDPAVLHRHQPVTIGEPAARAVSRTVTVERLDRIADDVLR 480
CYP116B3        SFRGPDHVWVQWDPQANPERTDPAVLQRQHPVTIGEPSTRSVSRTVTVERLDRIVDDVLR
cyp116          SFRGPEHVLVEWDPQQNP----------------VPADRP------------------
                *****:** *:**** **                  *: *.
```

Figure 5.12: Amino acids sequence alignment between CYP116, heme domain of CYP116B3 from *Rhodococcus ruber*, and P450RhF, generated by ClustalX. Residues marked with (*) are identical, while (:) indicate strong similarity, (.) indicate weak similarity respectively. The conserved oxygen binding motif is highlighted in yellow. The conserved cysteine residue is highlighted in green.

```
CCYP116B3       SFRGPDHVWVQWDPQANPERTDPAVLQRQHPVTIGEPSTRSVSRTVTVERLDRIVDDVLR 480
p450RhF         SFRGPDHVWVQWDPQANPERTDPAVLHRHQPVTIGEPAARAVSRTVTVERLDRIADDVLR
PDR             ----------------------------------TTPQEDGFLRLKIASKEKIARDIWS
                                                  .. .     :.:   ::*. *:

CCYP116B3       VVLRAPAGNALPAWTPGAHIDVDLG-ALSRQYSLCG-APDAPTYEIAVLLDPESRGGSRY 540
p450RhF         LVLRDAGGKTLPTWTPGAHIDLDLG-ALSRQYSLCG-APDAPSYEIAVHLDPESRGGSRY
PDR             FELTDPQGAPLPPFEAGANLTVAVPNGSRRTYSLCNDSQERNRYVIAVKRDSNGRGGSIS
                . *  . * .**.: .**:: : :  .  * ****. : :   * ***  *.:.****

CCYP116B3       VHEQLRVGGSLRIRGPRNHFALDPDAEHYVFVAGGIGITPVLAMADHARARG-WSYELHY 600
p450RhF         IHEQLEVGSPLRMRGPRNHFALDPGAEHYVFVAGGIGITPVLAMADHARARG-WSYELHY
PDR             FIDDTSEGDAVEVSLPRNEFPLDKRAKSFILVAGGIGITPMLSMARQLRAEGLRSFRLYY
                . ::   *..:.:  ***.*.**  *: :::*********:*:** : **.*  *:.*:*

CCYP116B3       CGRNRSGMAYLERVAGHGDRAALHVSAEG----TRVDLAALLATPVSGTQIYACGPGRLL 660
p450RhF         CGRNRSGMAYLERVAGHGDRAALHVSEEG----TRIDLAALLAEPAPGVQIYACGPGRLL
PDR             LTRDPEGTAFFDELTSDEWRSDVKIHHDHGDPTKAFDFWSVFEKSKPAQHVYCCGPQALM
                  *: .* *:::.::..  *: :::  :     . .*: :::  . .. ::*.***  *:

CCYP116B3       AGLEDASRHWPDGALHVEHFTSSLTALDPDVEHAFDLDLRDSGLTVRVEPTQTVLDALRA 720
p450RhF         AGLEDASRNWPDGALHVEHFTSSLAALDPDVEHAFDLELRDSGLTVRVEPTQTVLDALRA
PDR             DTVRDMTGHWPSGTVHFESFG--ATNTNARENTPFTVRLSRSGTSFEIPANRSILEVLRD
                  :.* : :**.*::*.* *    :  :.  : .* : *  ** :..: ..:::*:.**

CCYP116B3       NNIDVPSDCEEGLCGSCEVTVLEGEVDHRDTVLTKAERAANRQMMTCCSRACGDRLTLRL 780
p450RhF         NNIDVPSDCEEGLCGSCEVAVLDGEVDHRDTVLTKAERAANRQMMTCCSRACGDRLALRL
PDR             ANVRVPSSCESGTCGSCKTALCSGEADHRDMVLRDDEKGT--QIMVCVSRAKSAELVLDL
                 *: ***.**.* ****:.:: .**.**** ** . *:.:  *:*.* *** . .*.* *
```

Figure 5.13: Only the reductase domain of CYP116B3 (amino acids residues 419 to 771) from *Rhodococcus ruber* and reductase domain of P450RhF (amino acids residues 421 to 773) are shown in this alignment (indicated by CYP116B3 and P450RhF, respectively). PDR is the phthalate dioxygenase reductase from *Pseudomonas cepacia.* Residues marked with (*) are identical, while (:) indicate strong similarity and (.) indicate weak similarity. FMN binding motif is highlighted in blue. NAD(P)H binding motif is highlighted in yellow. The four cysteine residues that bind to the iron-sulfur cluster are highlighted in green.

5.4 Recombinant expression and protein purification

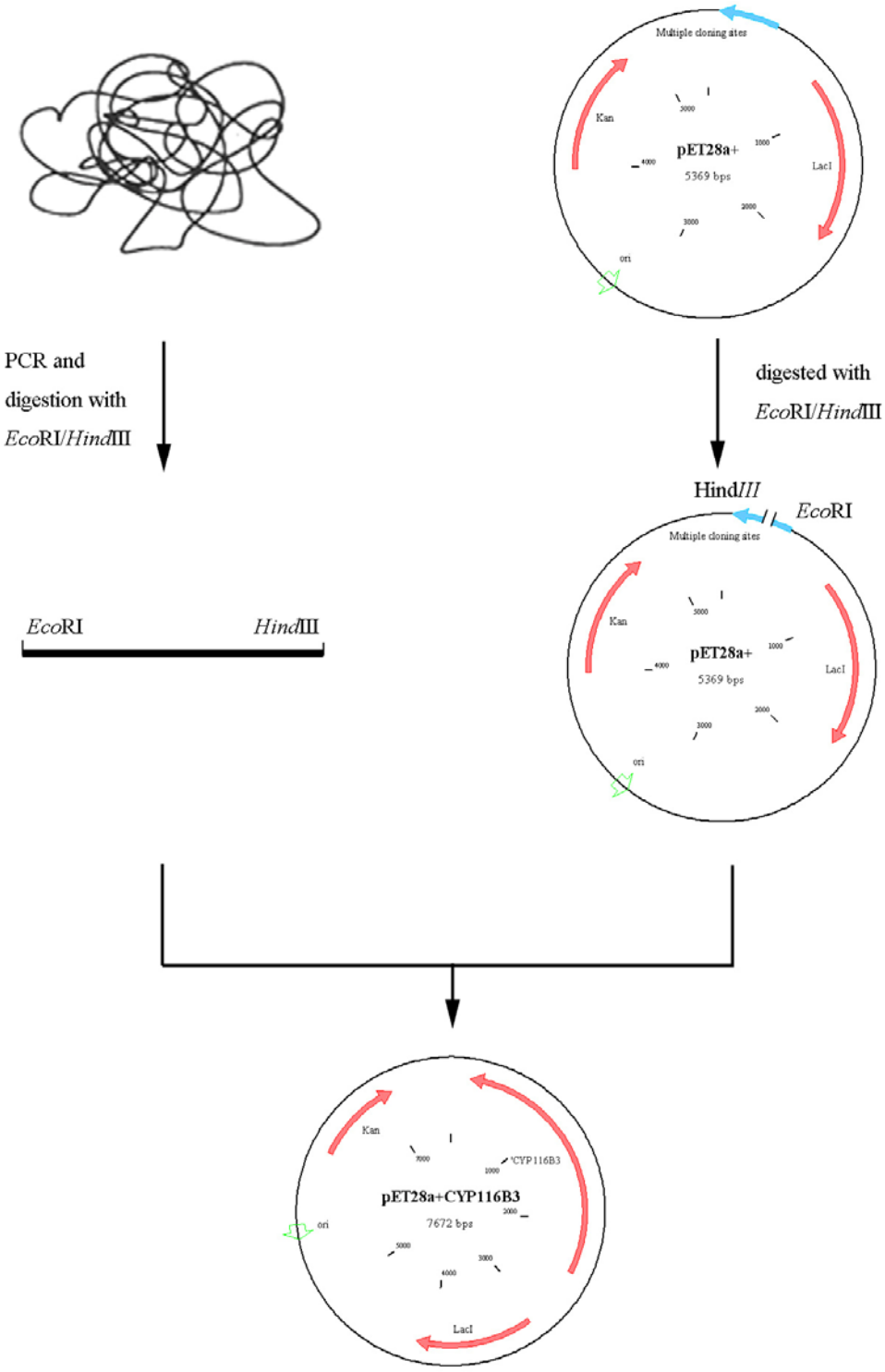

Figure 5.14: Schematic demonstration of cloning of CYP116B3 in the vector pET28a(+).

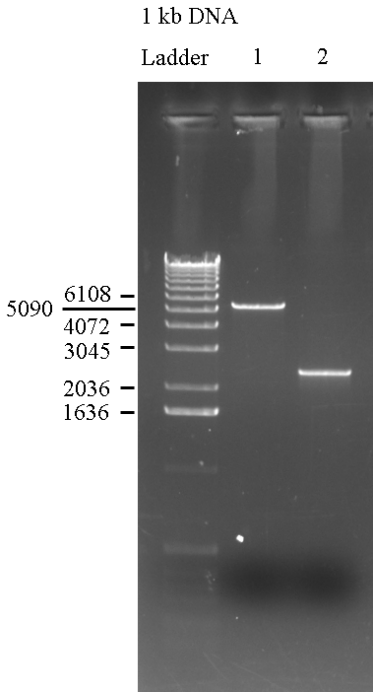

Figure 5.15: The entire gene was amplified by PCR using specific primers introducing restriction endonucleases sites upstreams and downstreams of the gene for further cloning. Lane 2 shows the DNA after digestion with *Eco*RI and *Hind*III. The pET28a(+) vector was also digested with the same restriction enzymes, *Eco*RI and *Hind*III (lane 1).

The expression strategy is schematically demonstrated in Figure 5.14. *Eco*RI and *Hind*III restriction sites were introduced by specific primers to clone P450 gene from *Rhodococcus ruber* DSM 44319 into the multiple cloning site of pET28a(+) expression vector. The pET28a(+) vector is IPTG inducible and the expression is under the control of the strong bacteriophage T7 promoter, which benefits a high yield expression of target protein. The PCR product and the vector were digested and purified (Figure 5.15). Because of the high transformation efficiency of *E. coli* DH5α, the ligation product was transformed in *E. coli* DH5α competent cells at first. The plasmid was isolated and the insert was checked by sequencing to avoid any replication error by DNA polymerase during PCR amplification. Because T7 promoter requires T7 RNA polymerase and *E. coli* DH5α lacks the T7 RNA polymerase, the expression construct was transformed in *E. coli* BL21(DE3) afterwards for protein expression. The resulting expression product contained 807 amino acids including a His_6-tag at N-terminus. The expression level after 0, 1, 2, 3, 4 and 16 hours after induction was analyzed by SDS-PAGE (Figure 5.16). Already after one hour induction a new band was visible at 89 KDa which corresponded to the target P450 protein (Figure 5.16, lane 2). The

maximal overexpression was reached after 16 hours. Yield of active P450 monooxygenase, calculated from the CO-difference spectrum, amounted to 18.6 mg l^{-1}.

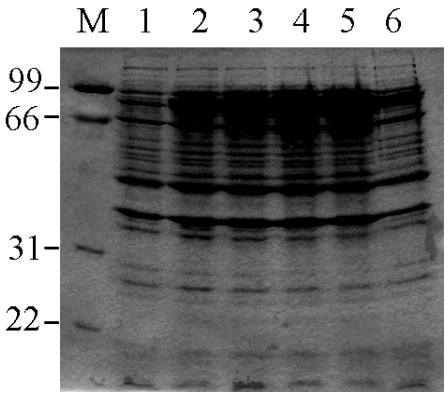

Figure 5.16: The protein expression of P450 from *Rhodococcus ruber* DSM 44319 visualized by SDS-PAGE. The overexpressed protein is visible at 89 kDa after induction with IPTG. Lane 1: before induction; lane 2: 1 hour after induction; lane 3: 2 hours after induction; lane 4: 3 hours after induction; lane 5: 4 hours after induction; lane 6: 16 hours after induction; M-molecular weight standard.

His_6-tag allowed for an easy high-performance one-step purification of a fusion protein by immobilized metal affinity chromatography (figure 5.17). Most cell proteins were not bound to the column (in flow though fraction, figure 5.17, lane 2), or were washed out by wash buffer with a low concentration of imidazole (30 mM) (figure 5.17, lane 3). The main band in lane 4 (indicated by an arrow) corresponds to the expected expression product at 89,000 Da that is identical to the molecular weight estimated from the protein sequence. The minor band is an *E. coli* protein non-specifically bound by the Ni-sepharose column. The purity was 85% after one step purification. The active enzyme recovery after purification was 73% (determined by CO-difference spectrum).

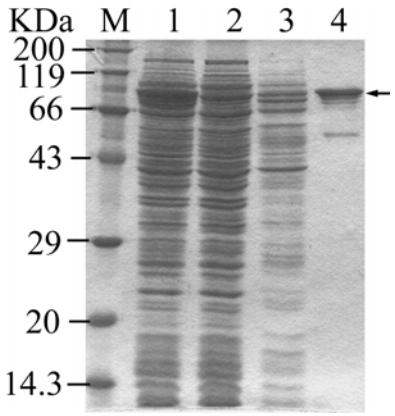

Figure 5.17: Protein purification of P450 from *Rhodococcus ruber* DSM 44319 visualized by SDS-PAGE. The target protein is visible at 89 kDa. Lanes: 1 - cell extracts; 2 - column flow-through fraction; 3 - washing fraction; 4 - purified enzyme; M - molecular weight standard.

5.5 Protein characterization

5.5.1 Spectroscopic characterization of the cytochrome P450 CYP116B3

Absorption spectrum of the oxidized protein showed a heme Soret band at 418 nm, which is typical for cytochrome P450 enzymes (Figure 5.18). A broad absorption shoulder between 450 and 510 nm indicates presence of flavins. The smaller α and β absorption bands are located at approximately 567 nm and 535 nm. The addition of NADPH led to a decrease of this shoulder due to reduction of the oxidized flavin component. The heme Soret maximum remained at 418 nm. The addition of carbon monoxide to the dithionite-reduced protein resulted in a Soret band shift from 418 nm to 449 nm due to formation of Fe^{2+}-CO complex, as expected for cytochrome P450 enzymes. This led also to a formation of a single peak of α and β absorption bands at approximately 550 nm. A CO-difference spectrum is subtraction of the spectrum of reduced P450 from the ferrous-CO complex typical for cytochrome P450 enzymes (5.19).

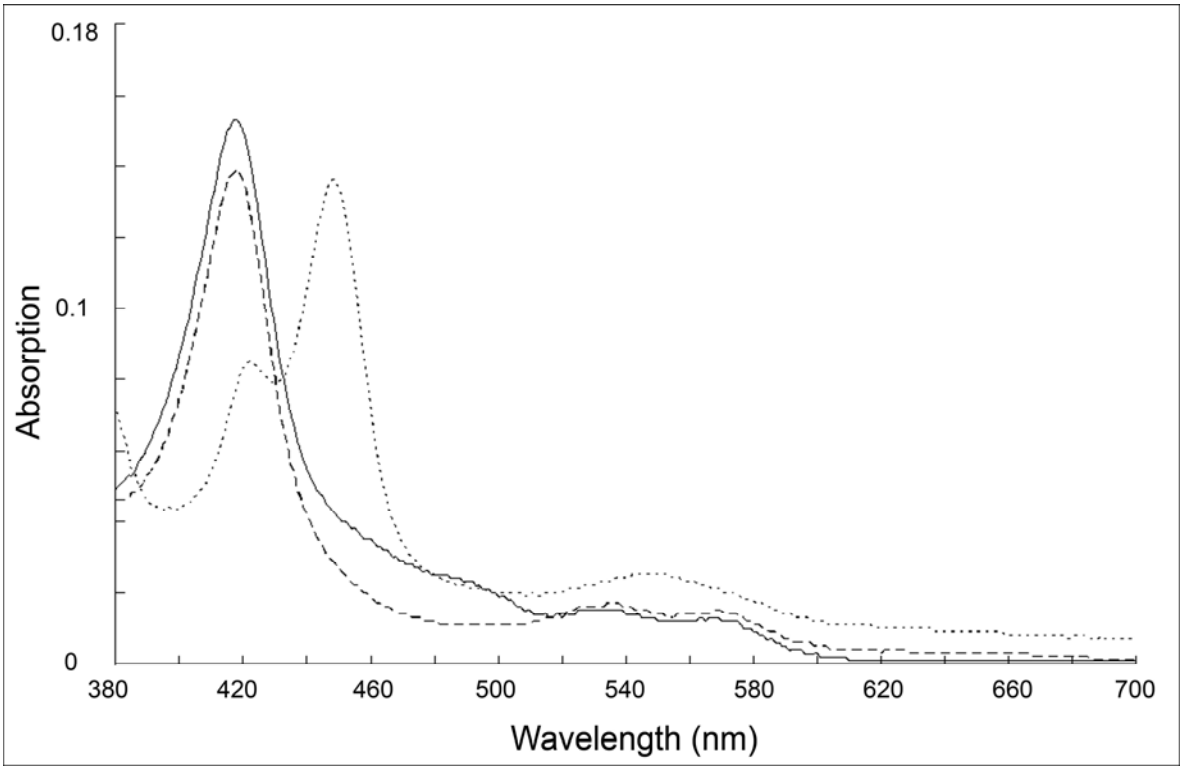

Figure 5.18: Spectral features of the cytochrome P450 CYP116B3 from *Rhodococcus ruber* DSM 44319. Spectra are shown for the oxidized (solid line), NADPH reduced (dashed line) and reduced/carbon monoxide-bound (dotted line) forms of the enzyme.

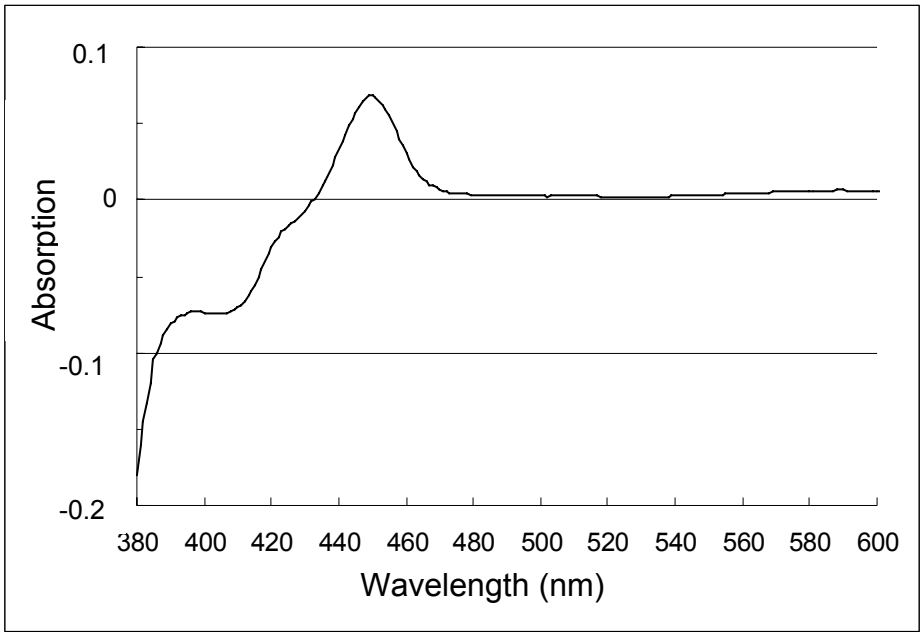

Figure 5.19: CO-difference spectrum of P450 CYP116B3.

5.5.2 Chromatographic determination of the flavin component of the cytochrome P450 CYP116B3

Identification of the flavin cofactor bound to the reductase was carried out using thin-layer chromatography (TLC). Because flavin cofactor is non-covalently bound to the protein, it can be released from the protein by heat treatment. The denatured protein was removed by centrifugation while the soluble flavin cofactor remained in supernatant. The analyzed sample had the same R_f-value of 0.41 as an authentic FMN, while FAD showed a different R_f-value of 0.24. No FAD was detected in the sample (Figure 5.20). The concentration of released FMN from enzyme was 0.12 mM as estimated from fluorescence measurements. Protein concentration in the analyzed sample was 0.11 mM; the molecular ratio between FMN and the protein is therefore 1.1 : 1, which is an indication that each P450 monooxygenase molecule contains one FMN molecule.

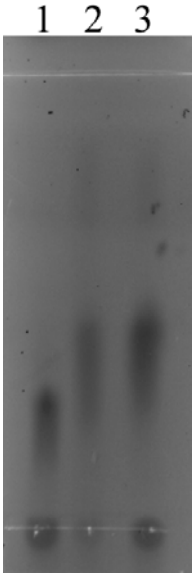

Figure 5.20: Thin-layer chromatography. Lane 1, 2 and 3 are authentic FAD, authentic FMN and sample, respectively.

5.5.3 Reductase activity determination under Steady-State conditions

In order to determine the reductase activity of the enzyme, cytochrome c was used as exogenous electron acceptor. Cytochrome c was reduced by cytochrome P450 CYP116B3 in the presence of NADPH or NADH. The enzyme kinetic follows the Michealis-Mantel Model. The maximum velocities (V_{max}) and Michealis constant (K_M) values were determined by Lineweaver-Burk plots (Figure 5.21 using NADPH and Figure 5.22 using NADH as the electron donor). The result showed that the cytochrome P450 CYP116B3 exhibited a preference for NADPH over NADH during this reduction, similar to P450RhF. The k_{cat} value for NADPH as an electron donor is more than twice that for NADH, and the K_M value for NADPH is 30-fold lower than that for NADH (Table 5.4).

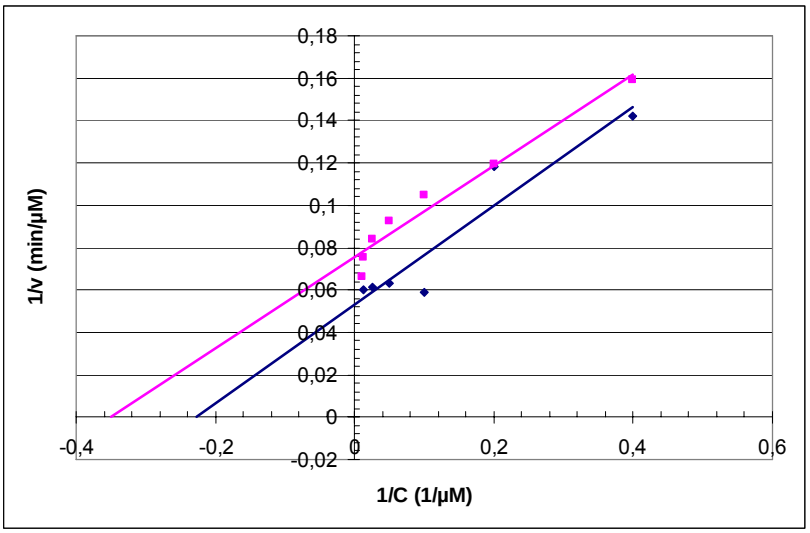

Figure 5.21: A double determination of reductase activity was performed using NADPH as electron donor. Data points and trend line from the same reaction are in the same color in the Lineweaver-Burk plots.

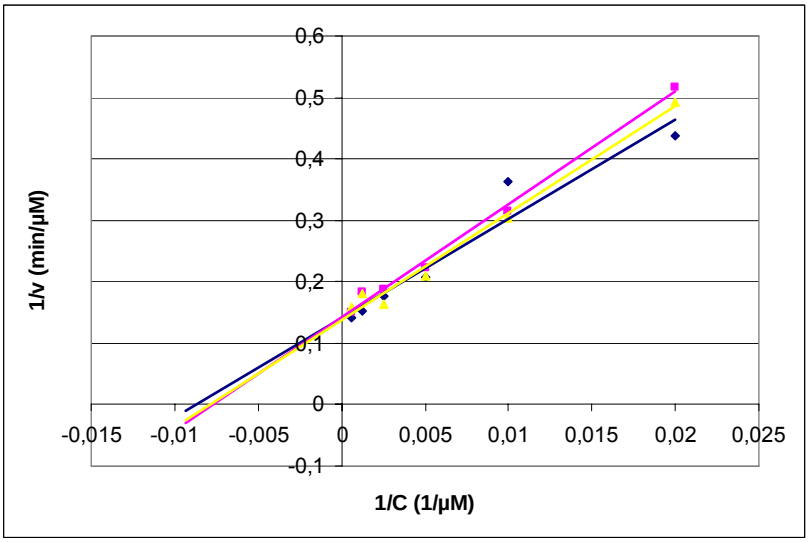

Figure 5.22: A triple determination of reductase activity was performed using NADH as the electron donor. Data points and trend line from the same reaction are in the same color in the Lineweaver-Burk plots.

Table 5.4: Kinetic parameters of the P450 monooxygenase from *Rhodococcus ruber* DSM 44319 with cytochrome *c* as the electron acceptor.

cofactor	K_M (µM)	v_{max} (µM min^{-1})	k_{cat} (min^{-1})
NADPH	3.4 ± 1.0	15.3 ± 3.0	765 ± 144
NADH	126.2 ± 11	7.2 ± 0.3	360 ± 12

5.5.4 Determination of substrate spectra and oxidation activity

After expression and purification, P450 from *Rhodococcus ruber* DSM 44319 was tested for its oxidation activity towards several different chemical compounds. 7-Ethoxycoumarin was chosen because the highly homologous P450RhF catalyzes the O-dealkylation, leading to 7-hydroxycoumarin. Both 7-ethoxycoumarin and 7-hydroxycoumarin are strongly fluorescent and can be separately identified by fluorescence spectroscopy. The rate of O-dealkylation of 7-ethoxycoumarin by the P450 monooxygenase was determined by measuring 7-hydroxycoumarin (Figure 5.23, Table 5.5).

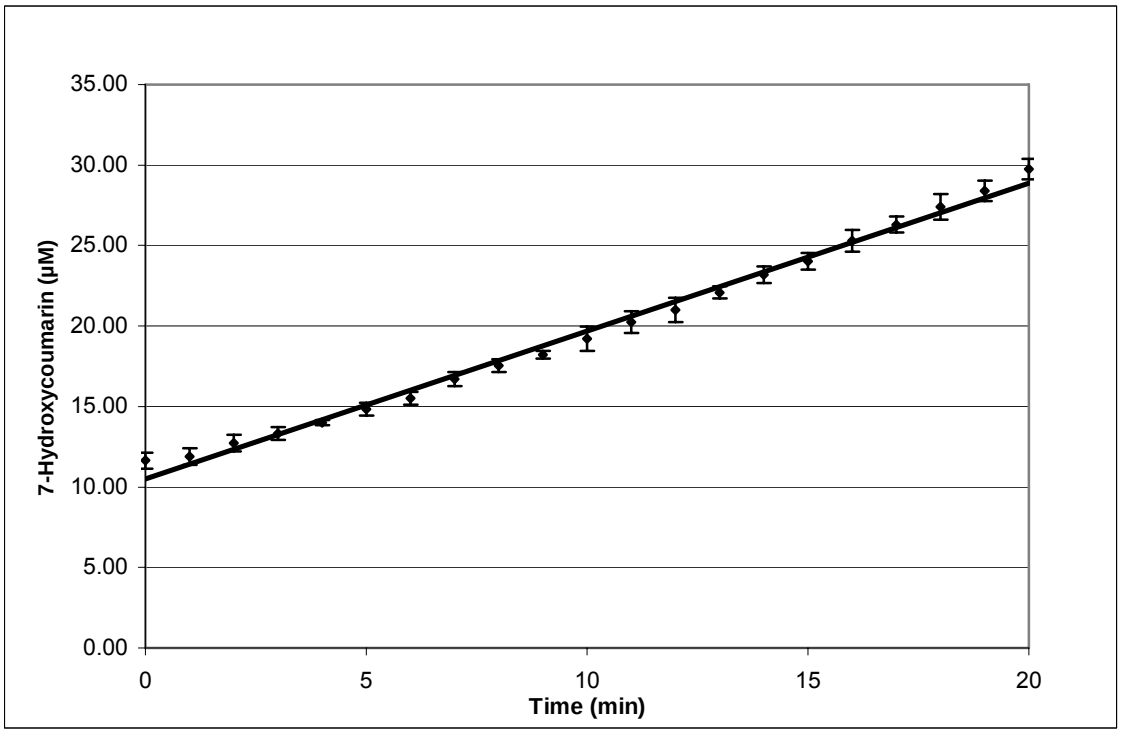

Figure 5.23: CYP116B3 catalyzes dealkylation of 7-ethoxycoumarin and the product, 7-hydroxycoumarin, was measured using fluorescence spectroscopy. A triple determination was performed to determine the product formation rate.

It is known that toluene-degrading biofilms used for microbial water purification contain numerous *Rhodococcus* strains. Therefore toluene was also chosen as substrate. GC-MS analysis revealed that toluene was hydroxylated by CYP116B3 at the methyl group to form benzyl alcohol (retention time, RT, 5.73 min; M^+ = 108, figure 5.24). However,

hydroxylation at the benzene ring was not observed. The activity was approximately threefold lower than towards 7-ethoxycoumarin.

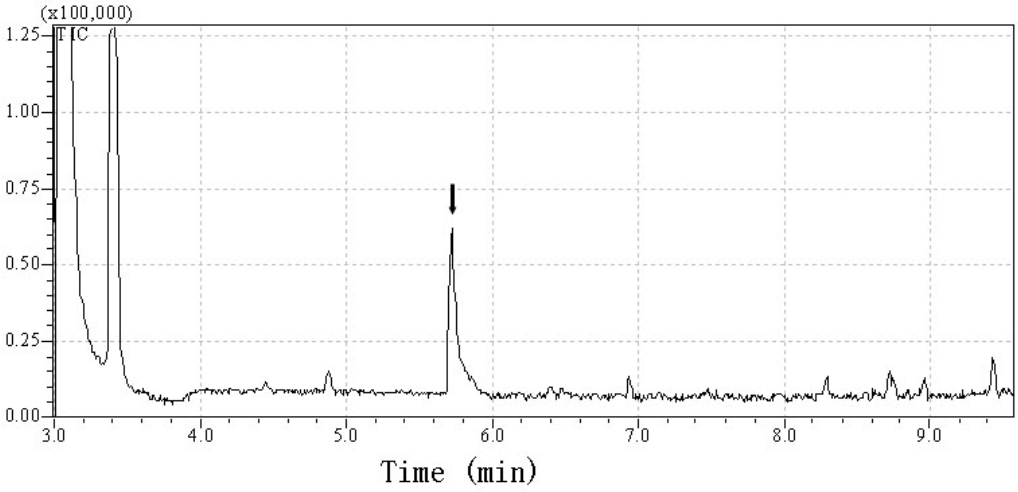

Figure 5.24: Toluene was hydroxylated to benzyl alcohol by cytochrome P450 CYP116B3. The arrow indicates the hydroxylation product benzyl alcohol in the GC-MS chromatogram.

Furthermore, some polycyclic aromatic hydrocarbons naphthalene, acenaphthene, fluorene, and indene were chosen as substrates. Aromatic hydrocarbons were oxidized by P450 CYP116B3 monooxygenase from *Rhodococcus ruber* DSM 44319 far slower than 7-ethoxycoumarin. The products derived from the corresponding substrates are summarized in Figure 5.27. Naphthalene was oxidized regioselectively to 1-naphthol (RT, 13.03 min; M^+ = 144; Figure 5.25, A). Acenaphthene was oxidized to 1-acenaphthenol (RT, 3.99 min; M^+ = 169; Figure 5.25, B). A single product of fluorene oxidation was identified as 9-fluorenol (RT, 4.43 min; M^+ = 181; Figure 5.25, C).

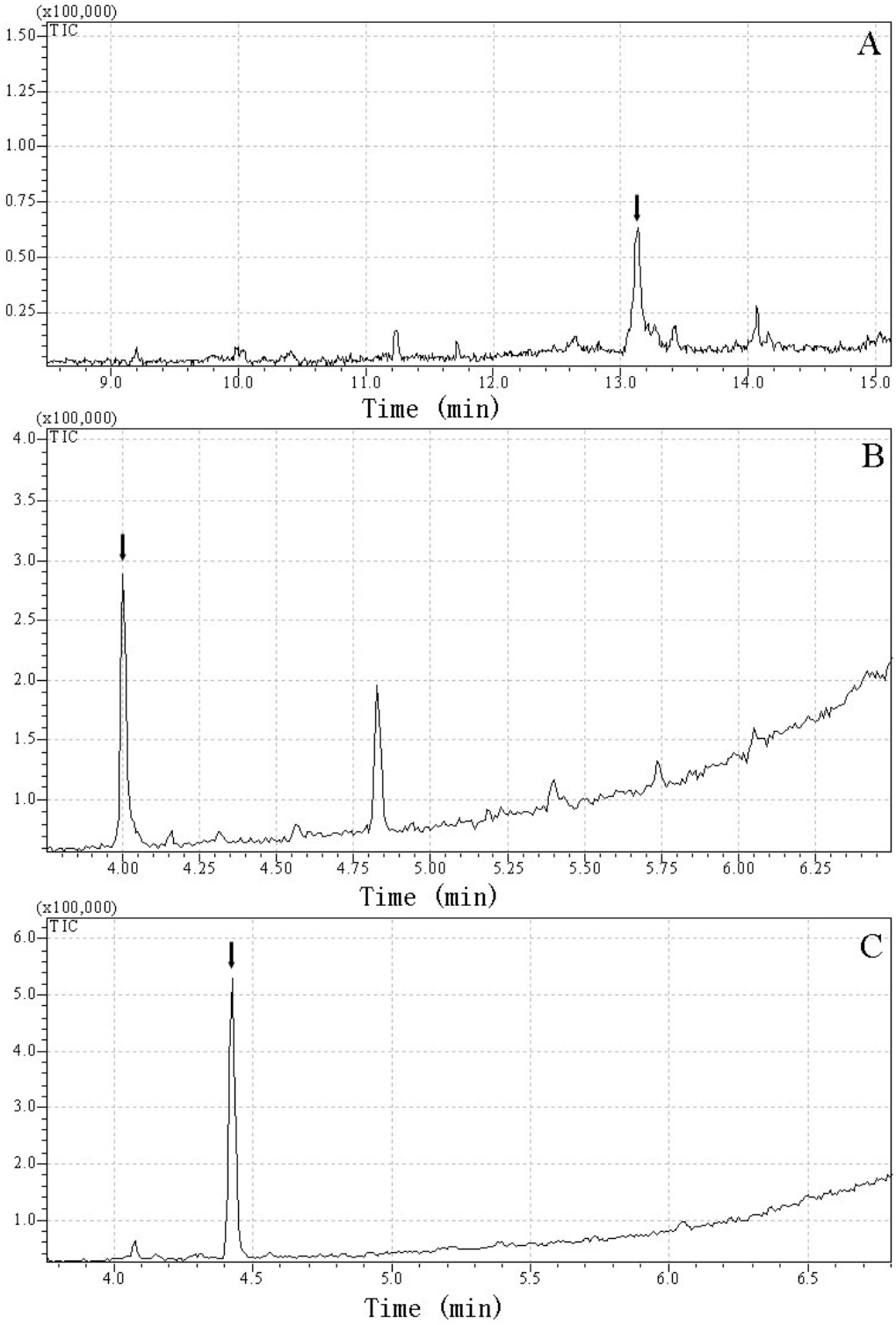

Figure 5.25: GC-MS results: A – naphthalene, B – acenaphthene and C – fluorene. The arrows indicate the products: 1-naphthol, acenaphthenol, and 9-fluorenol, respectively. The substrates are not shown on chromatograms, because the concentration of substrates was much higher than the concentration of products.

Indene was also oxidized by P450 CYP116B3. Only one product, 1-indenol, (RT, 7.89 min; $M^+ = 132$) was identified by GC-MS (Figure 5.26). Indole has a similar chemical structure to indene, but interestingly, no oxidation activity towards indole by P450 CYP116B3 was detected.

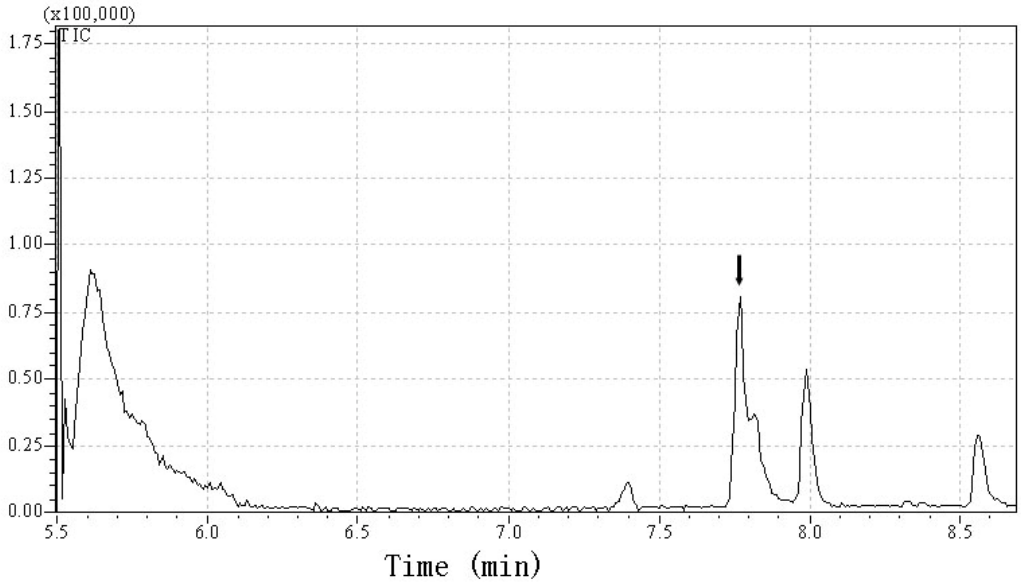

Figure 5.26: Indene was hydroxylated to 1-idenol by cytochrome P450 CYP116B3. The arrow indicates 1-indenol in the GC-MS chromatogram. All other peaks are impurities.

In order to confirm the selectivity of the new P450 monooxygenase, two other alkyl aromatic hydrocarbons were tested - *m*-xylene and ethyl benzene (Table 5.5 and Figure 5.27). Since activity towards these compounds was very low, whole-cell oxidation using recombinant *E. coli* cells producing the P450 enzyme was performed in order to produce a sufficient amount of the hydroxylated products. After the biotransformation overnight, only 3-methylbenzyl alcohol (RT, 6.78 min; $M^+ = 122$) was detected as a product of *m*-xylene oxidation. The peaks were integrated and the distribution of substrate/product was 87.1% to 12.9%, calculated from the peak areas. In case of ethyl benzene, however, two different products were identified: 1-phenylethyl alcohol (RT, 5.92 min; $M^+ = 122$) and 2-phenylethyl alcohol (RT, 6.62 min; $M^+ = 122$). The peaks were integrated and distribution of substrate/product was 86.5% to 7.5% to 6%, calculated by peak areas. No product was detected when *E. coli* cells with empty vector were used. Like toluene, both substrates were hydroxylated exclusively at the alkyl chain and no ring oxidation was observed. Such selectivity towards the alkyl chain

of aromatics can be applied in technical processes, since chemical oxidation is complicated and requires harsh reaction conditions.

No hydroxylation activity towards α-pinene, terpeneol, ethoxyresorufin was detected. Interestingly, no hydroxylation activity towards cyclohexane was detected with the purified cytochrome P450 CYP116B3 enzyme. One may speculate that another P450 monooxygenase present in *Rhodococcus ruber* is responsible for the hydroxylation of cyclohexane.

Table 5.5: Summary of substrates, products, and product formation rates using the P450 monooxygenase from *Rhodococcus ruber* DSM 44319.

Substrate	Product	Product formation rate (nmol x nmol P450 $^{-1}$ x min^{-1})
7-Ethoxycoumarin	7-Hydroxycoumarin	0.917 ± 0.05
Acenaphthene	1-Acenaphthenol	0.079 ± 0.01
Fluorene	9-Fluorenol	0.04 ± 0.005
Naphthalene	1-Naphthol	0.106 ± 0.01
Indene	1-Indenol	n. d.
Toluene	Benzyl alcohol	0.301 ± 0.01
Ethyl benzene	1-Phenylethyl alcohol	n. d.
	2-Phenylethyl alcohol	n. d.
m-Xylene	3-Methylbenzyl alcohol	n. d.

n. d. - not determined

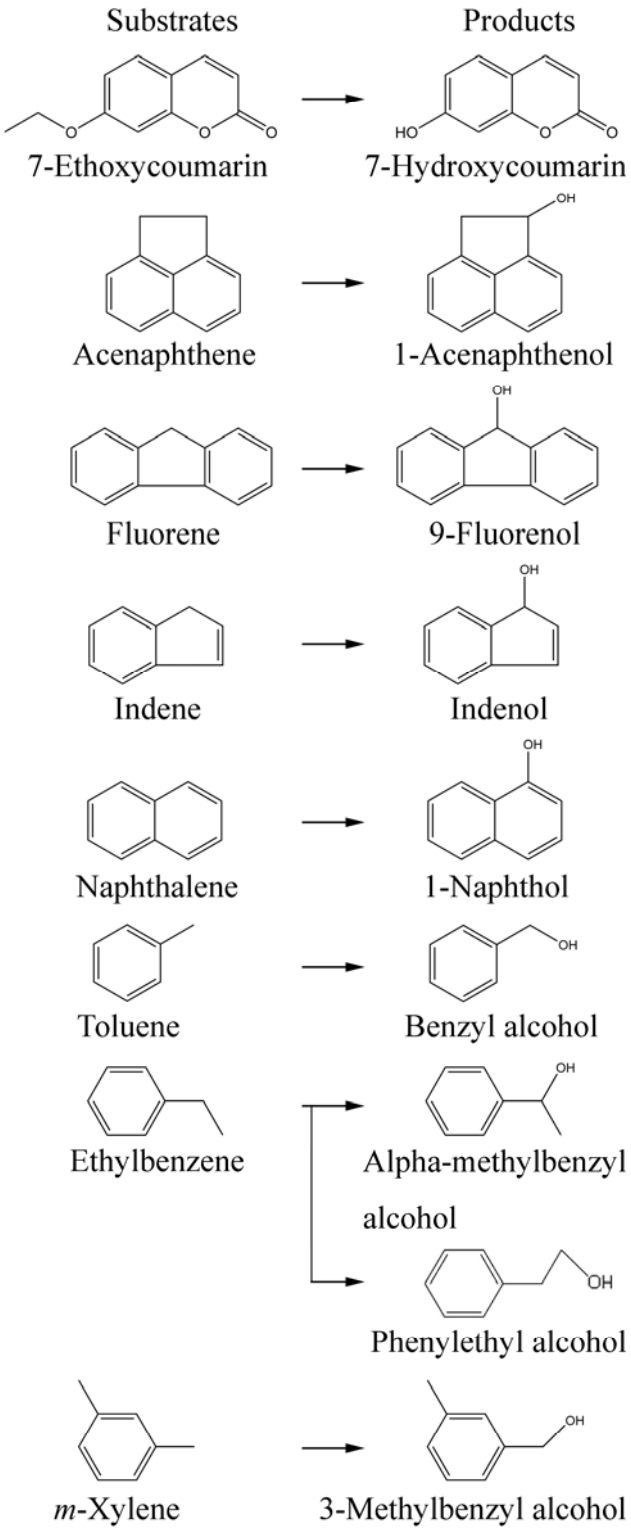

Figure. 3.27: Reactions catalyzed by P450 CYP116B3 from *Rhodococcus ruber* DSM 44319.

5.6 Protein engineering of CYP116B3

5.6.1 Rational design and Site directed mutagenesis

To improve activity and change selectivity of the new P450, we attempted to modify the heme domain of CYP116B3. The first strategy that we applied was rational design. A structure was modeled based on the sequence alignment and crystal structures of P450terp (CYP108A1) (PDB ID, 1CPT) (Boddupalli, Hasemann et al. 1992), P450epoK (CYP167A1) (PDB ID, 1PKF) (Nagano, Li et al. 2003) and P450eryF (CYP107A1) (PDB ID1Z8O) (Nagano, Cupp-Vickery et al. 2005) (Figure 3.28). Sequence identity levels between heme domain of CYP116B3 and three other P450 monooxygenases, P450terp, P450epoK and P450eryF were 29%, 28% and 28%, respectively. Although protein sequence identity was relatively low, the conserved sequences could be identified in the multiple sequence alignment, which was done using ClustalX. The high quality of the sequence alignment ensured the quality of the homology model. After analysis of the homology model, the amino acids, which are located in the substrate-binding pocket, were identified. We assumed, that a geometrical alternation in the binding pocket may affect substrate specificity. If the volume of the binding pocket is reduced, smaller substrate might be stabilized in the binding pocket and more specific oxidation might occur. To diminish the binding pocket, the alanine at position 109, which was located at the ceiling of the binding pocket, was substituted by a bulky amino acid, phenylalanine. The site directed mutagenesis was carried out by using Quick-Change PCR as described in Materials and Methods. The mutant A109F was expressed in *E. coli* BL21(DE3) and characterized. Only the activity towards 7-ethoxycoumarin and PAHs increased nearly 3-fold compared to wild-type enzyme. However, no hydroxylation activity towards smaller substrates, for example, α-pinene or cyclohexane, was observed.

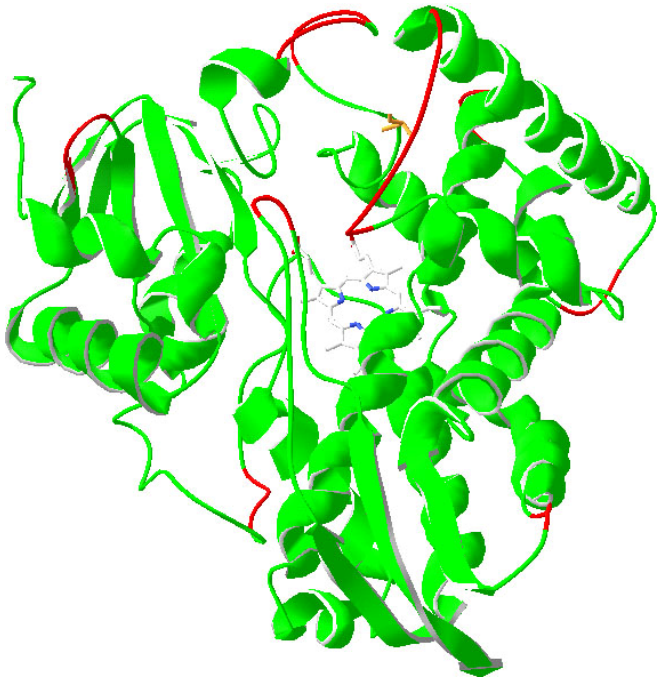

Figure 5.28: Computer model of CYP116B3 heme domain. The helix and sheets are shown in green. The loops are in red. The alanine residue at 109 is highlighted in orange.

5.6.2 Directed evolution

The results obtained from rational design show that our understanding of this P450 is still very limited. As directed evolution does not require any structural information for construction of a random mutant library, we decided to apply it for improvement of P450 CYP116B3. A random mutant library was generated by error-prone PCR.

5.6.2.1 High-throughput assay

Fist s high-throughput assay had to be developed to identify mutants with improved activity from a large number of variants. As mentioned before, both, substrate and product of the 7-ethoxycoumarin dealkylation are fluorescencent. The concentration of product 7-hydroxycoumarin can be directly measured by fluorescence spectroscopy; therefore, 7-ethoxycoumarin is a suitable test substance for a high-throughput screening. To avoid the tedious cell lysis step and the high costs of the co-factor NADPH, conversion of 7-ethoxycoumarin to 7-hydroxycoumain was carried out with whole *E. coli* cells. As the *E. coli* Novablue(DE3) strain contains T7 RNA polymerase and has high transformation efficiency, it was chosen as the expression host in the directed evolution experiments.

To determine whether dealkylation activity could be measured with whole *E. coli* cells, a test with wild-type CYP116B3 was carried out in microtiter plates. A plasmid (pET28a(+)) containing the wild-type gene of CYP116B3 was transformed into the *E. coli* Novablue(DE3) cells. 5 ml preculture were inoculated with a fresh colony from agar-plate and grown overnight. 4 ml of the preculture were used for inoculation of 400 ml of LB medium in a shake-flask and it was cultured at 30 °C until OD reached 0.8. Protein expression was induced by addition of 0.25 mM IPTG and was carried out overnight at 25 °C. Afterwards 600 µl of culture (contains about 1.2 mg cell wet weight) was taken out and centrifuged. Supernatant was discarded, the cells were resuspended in 180 µl (or in 160 µl for treatment with polymyxin B) potassium buffer (50 mM, pH 7.5) and transferred into a fresh microtiter plate. As in several publications permeabilized *E. coli* cells have been demonstrated to convert organic compounds without cell disruption (Schwaneberg, Otey et al. 2001), a permeabilizer polymyxin B was also tested. 20 µl polymyxin B (0.6 mM, dissolved in potassium buffer, 50 mM, pH 7.5) was added to 160 µl cell suspension. After addition of 20 µl 7-ethoxycoumarin (25 mM, dissolved in dimethyl sulfoxide, DMSO), increase in

fluorescence of 7-hydroxycoumarin was measured immediately. As a negative control *E. coli* cells containing the empty plasmid pET28a(+) have been used. A positive control was performed using cell lysate produced from the equivalent cell mass (about 1.2 mg cells). There was no significant difference in dealkylation activity towards 7-ethoxycoumarin between whole cells with or without polymeyxin B and cell lysate (Table 5.6). Cell membrane might be a diffusion barrier for substrate uptake, therefore DMSO that has been found to perturb the phospholipid bilayers, was used as a co-solvent (Anchordoguy, Carpenter et al. 1992). DMSO may facilitate the uptake of hydrophobic compounds by *E. coli* cells. However, polymyxin B permeabilized cells had no effect on P450 activity in this case. The result reveals that dealkylation activity could be measured with whole *E. coli* cells and no permeabilizer is required.

Table 5.6: The activity of whole cells with or without polymyxin B and cell lysate was compared in a total reaction volume of 200 µl. No activity was detectable in the negative control.

	Buffer (50 mM KPi, pH 7.5)	NADPH (1 mg/ml)	7-Ethoxy-coumarin (25 mM)	Polymyxin B (0.6 mM)	Product formation rate (fluorescence unit/min)
1.2 mg cells (CWW)	180 µl	-	20 µl	-	8.5
1.2 mg cells (CWW)	160 µl	-	20 µl	20 µl	7.9
10 µl cell lysate (from 1.2 mg cells, CWW)	150 µl	20 µl	20 µl	-	7.4

CWW = cell wet weight

5.6.2.2 Assay validation in microtiter plates

The main prerequisite of each assay suitable for screening of mutant libraries is it's miniaturization in microtiter plates. The most important question in our case was whether differences in activities of CYP116B mutants towards 7-ethoxycounmarin can be

distinguished when enzymes are expressed in a microtiter plate scale. To answer the question, the wild-type enzyme and the A109F mutant with a three-fold higher activity toward 7-ethoxycoumarin, were cultivated in a microtiter plate. 48 wells of a 96-well microtiter plate contained the wild type enzyme and the residual 48 wells contained the mutant A109F. The expression was carried out in a 96-deep-well microtiter plate containing 600 µl LB medium as described in "Material and Methods". From 48 repetitions of the mutant A109F, most samples showed activity in the range between 1 U/min and 2 U/min; 7 samples displayed higher activity of more than 3 U/min. The activity of most samples of the wild-type enzyme lies below 1 U/min, and only one sample displayed activity of 2 U/min (sample 8). Such a difference in activity of the same enzyme, produced in different wells is due to different protein expression levels. In this experiment the mutant A109F clearly exhibited higher activity towards 7-ethoxycoumarin than the wild-type enzyme (Figure 5.29). Hence, a mutant with higher activity towards 7-ethoxycoumarin could be identified using this assay. However, since some false positive may occur during primary screening, a secondary screening is necessary to confirm hits from the primary screening.

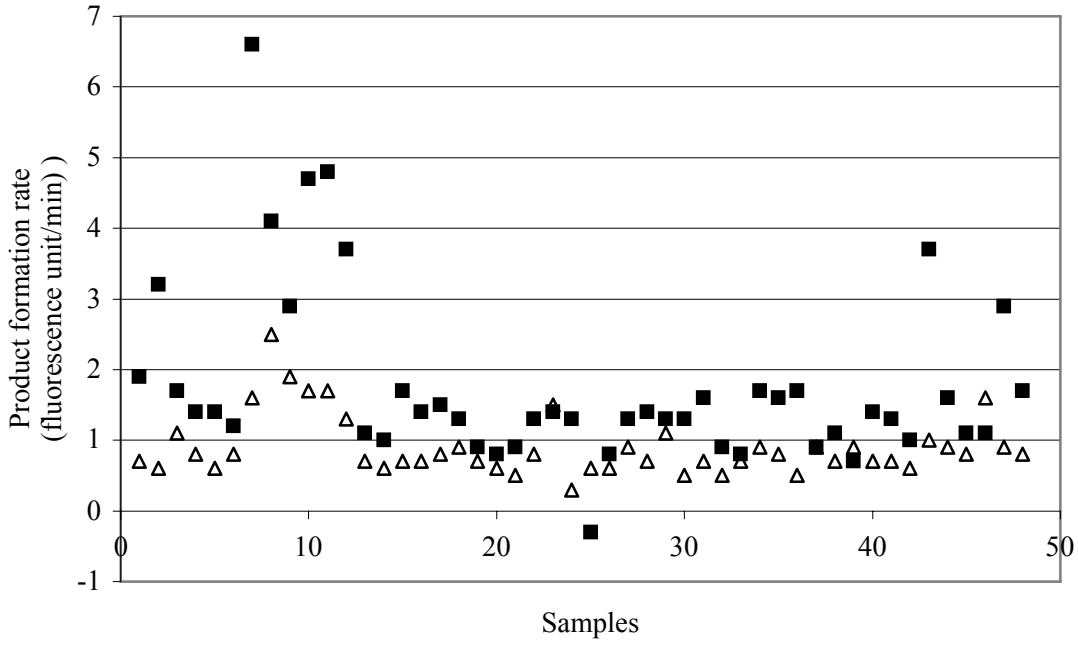

Figure 5.29: *O*-dealkylation of 7-ethoxycoumarin by whole cells was carried out in a microtiter plate. Squares and open triangles indicate the mutant A109F and wild-type CYP116B3, respectively.

5.6.2.3 Screening of the mutant library

The mutant library was created by error-prone PCR using a standard protocol as described in the "Materials and Methods" section. The mutant A109F was used as a template for the first round of error-prone PCR. The primary screening was performed in microtiter plates. The results were statistically analyzed using Microsoft-Excel 2003 and revealed an inactivation rate of the monooxygenase due to mutagenesis of approximately 35%. The first round of error-prone PCR mutagenesis served as an example. Twelve variants were selected, which exhibited obviously higher activity than the average activity of all active variants (Figure 5.30). The secondary screening was performed with twelve positive variants from the primary screening, which were cultivated on a larger scale for further detailed characterization. The variants were produced in a 500 ml shake-flask containing 100 ml LB medium, supplemented with corresponding antibiotics. After cell lysis, concentration of cytochrome P450 enzyme was determined from the CO-difference spectrum. The specific activity towards 7-ethoxycoumarin was measured using cell lysate and shown in Figure 5.31. Only three out of the twelve variants exhibited increased activity. Other variants were false positives. Variant 7F08 was chosen for sequencing. The sequence revealed that adenine was substituted by cytosine at position 799 which resulted in substitution of isoleucine by leucine at position 267 of protein sequence. Variant 7F08 was used as the template for the second round of error-prone PCR.

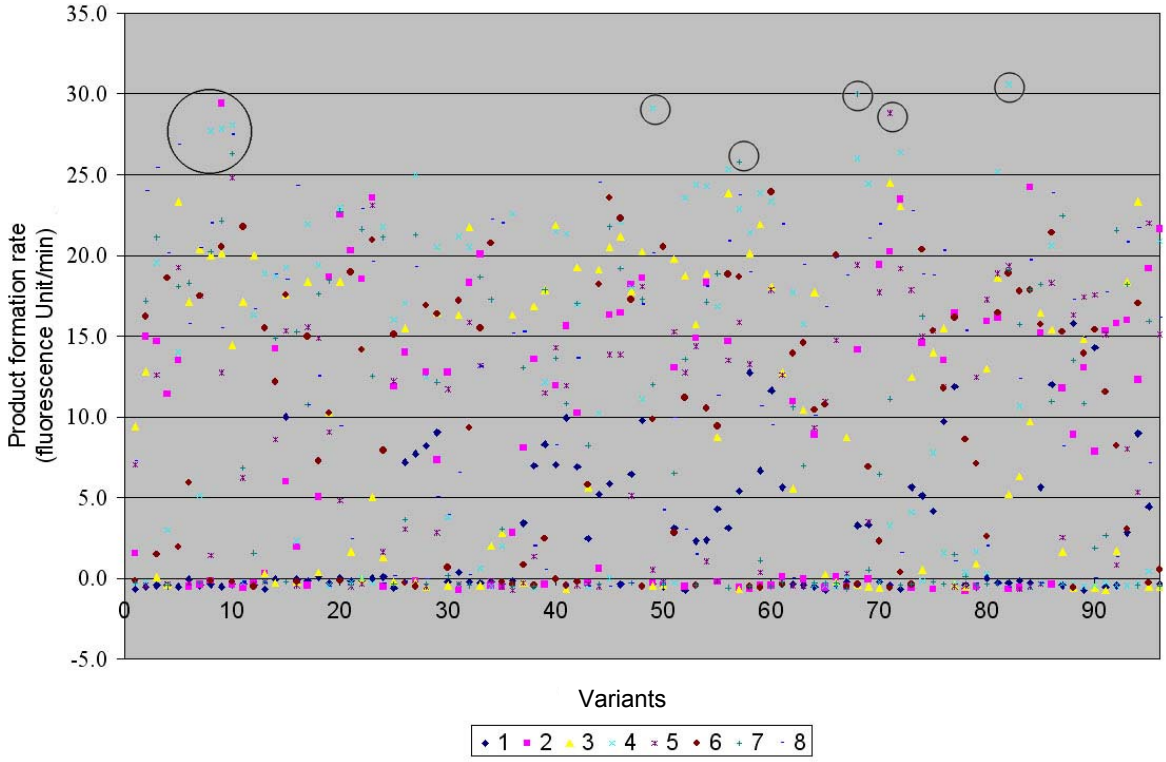

Figure 5.30: First primary screening. Eight 96-well microtiter plates were screened. The samples from different plates are indicated in different colors. The activity of most variants was below 25 U/min. The variants on the bottom of the figure were possibly inactivated due to mutagenesis. The circles indicate the selected variants for the secondary screening.

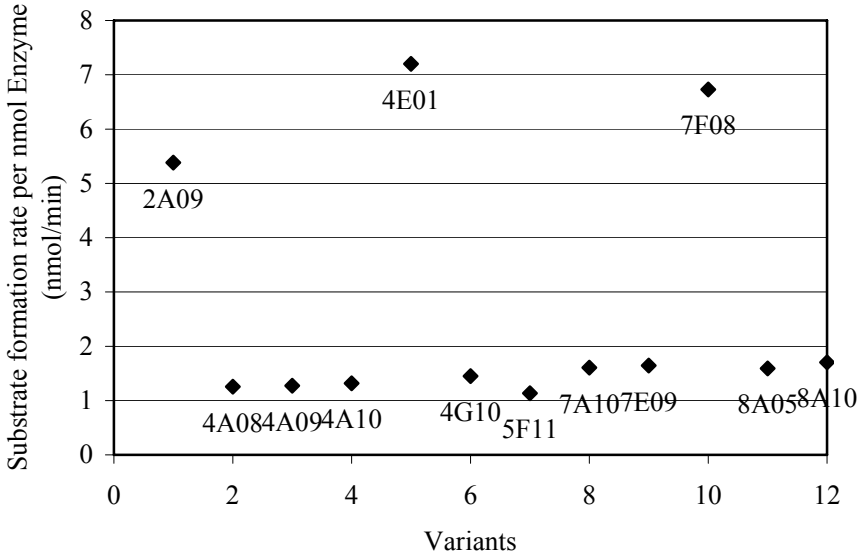

Figure 5.31: Twelve best variants were chosen for secondary screening. The variants were named by the plate number and the position in the microtiter plate.

5.6.2.4 Characterization of the random mutants

After four rounds of error-prone PCR, where the most active variant served as template for the next round of random mutagenesis, the activity towards 7-ethoxycoumarin was increased from 2 nmol to 223 nmol per nmol enzyme per minute (by variant 70A08). Totally, about 7800 variants were screened. The stepwise directed evolution always added a new mutation in the protein after each step of error-prone PCR. Each step of directed evolution exhibited an additive effect on the previous mutant (Table 5.7). Finally, the two best mutants, 70A08 (A86T/T91S/A109F/I179F/I267L) and 74H10 (T91S/A109L/I179F/I267L) were isolated from the fourth error-prone PCR library. To investigate the alteration of the substrate spectrum of the two mutants, a wide range of potential substrates was tested, including known substrates of wild-type CYP116B3. However, 70A08 and 74H10 did not show increased activity towards any substrates other than 7-ethoxycoumarin.

Table 5.7: Best mutants after four rounds of error-prone PCR

Round of error-prone PCR	Variant	Nucleotide substitution	Amino acid substitution	Product formation rate (nmol x nmol P450 $^{-1}$ x min^{-1})
1	07F08	799 A → C	I267L	7
2	30A06	256 G → A	A86T	30
3	44G08	535 C → T	L179F	100
4	70A08	271 A → T	T91S	223
	74H10	325 T → C	A109L	141

6. Discussion

6.1. Biotransformation

Bacterial P450 enzymes are in most cases not membrane-associated, water soluble and exhibit relatively high stability (Werck-Reichhart and Feyereisen 2000). Several bacterial strains were screened for monooxygenase activity. During the screening, two *Rhodoccocus* strains exhibited oxidation activity towards cyclohexane. Due to the complexity of microorganisms, biotransformation in whole-cell leads usually to overoxidation of alcohol to ketone or carboxylic acid (Adam, Lukacs et al. 2000). In this study the sole product, cyclohexanone, was detected after the biotransformation of cyclohexane using resting cells. However, only very low oxidation activity towards cyclohexane was observed. Probably the enzyme activity was very low, or the expression level of the corresponding enzyme was very low. We supposed that any monooxygenase in *Rhodoccocus* hydroxylated cyclohexane to cyclohexanol and the product was later oxidized to cyclohexanone. Therefore we decided to isolate the genes encoding the corresponding monooxygenases from both active *Rhodoccocus* strains for heterologous expression in *E. coli* and further investigation.

6.2. Gene isolation

Since the genomes of the two strains were not sequenced, we applied a PCR-based homology search to identify the corresponding P450 genes (Danielson, MacIntyre et al. 1997; Hyun, Kim et al. 1998). In PCR-based homology search is an essential first step to investigate related P450 monooxygenases, whose gene sequences are not published. In spite of a high GC-content of DNA from *Rhodococcus* strains, four degenerate primers with only 12.5% of degeneracy could be designed. These primers are based on sequence alignment between CYP116 from *Rhodoccocus* sp. NI86/21 and P450RhF from *Rhodococcus* sp. NCIMB 9784 and correspond to the regions of highest similarity. Because adenine and thymine have weaker hydrogen bonding than guanine and cytosine, they were chosen at the 3'-end of primers to reduce non-specific binding during PCR.

Two out of four primer combinations gave one specific PCR product derived from *Rhodococcus ruber*, but not from *Rhodococcus erhytropolis*. Sequence analysis and comparison with authentic P450 sequences revealed that it was a P450-like gene fragment.

There are many different methods to identify and isolate the entire gene if a gene fragment has already been amplified. Using "directed sequencing of genomic DNA", a large template DNA up to 4.6 Mb can be directly sequenced using flanking primers and BigDye terminators (Heiner, Hunkapiller et al. 1998). Although the signal intensity from genomic template is only 10 – 20% compared to the data from plasmid template, however, the use of BigDye terminators permitted direct sequencing on genomic templates, because the baseline noise in sequencing is reduced by increased brightness and improved resolution of BigDye terminators. Using a "long distance genome walking" method (Min and Powell 1998), a ssDNA is amplified in a primary PCR using a primer P1 which is located closely to the 3'-end of the known region. Subsequently, a "tailing"-reaction is carried out using Terminal Desoxyribonucleotide Transferase to attach a poly-dC-tail to the ssDNA. The ssDNA is used as template in a secondary PCR using a primer P2 which is located downstream of the primer P1, and an adapter-primer which consists of poly-dG and an *Eco*RI restriction site. Finally, this product of the secondary PCR can be further cloned and sequenced. However, because of low signal intensity and signal-noise ratio of obtained sequences, both methods, "directed sequencing of genomic DNA" and "long distance genome-walking" were not considered in this study. The classic method, the creation and screening of genomic DNA libraries provides usually good results. However, it is tedious and expensive. Therefore, we preferred a genomic walking method using PCR, described recently by Mishra et al. (Mishra, Singla-Pareek et al. 2002). The complicated and time-consuming probing of a genomic DNA library and the subsequent sequencing of positive clones could be avoided. To increase the quality and quantity of the obtained products, two step PCR was used. The primary PCR amplified a new gene fragment, which was an extension from the known gene fragment. However, most PCR products were non-specific as together with specific primers also completely degenerate walking primers were used. Because the specific primer was biotinylated, non-specific products could be easily eliminated by using streptavidin-coated beads. The secondary PCR with nested primers was run to amplify highly pure DNA of new gene fragment using a second specific primer, which should be located inside the new gene fragment. The method "directed sequencing of genomic DNA" was used 5 times stepwise, until the entire gene was completed. In each case only up to 400 base pairs could be obtained.

A novel cytochrome P450 monooxygenase with new features has been identified by the sequence-based screening approach. In comparison to classical enzyme activity screening, this approach facilitates the isolation, cloning, characterization, and expression of new cytochrome P450 genes from genetically closely related organisms without difficult steps of protein purification and sequencing. This approach can also be used in further study to discover more new enzymes with potential for industrial applications. However, as P450 genes generally share less than 20% sequence identity; this method is restricted to only closely related strains.

6.3. Cloning, expression and purification

This new P450 gene from *Rhodococcus ruber* DSM 44319 was named by the P450 nomenclature committee *CYP116B3.* A pET expression system was chosen for the overexpression of cytochrome P450 CYP116B3. The pET28a(+) contains the IPTG inducible strong T7-promotor, which facilitated the expression of sufficient amount of active enzyme. An optimal expression level of cytochrome P450 CYP116B3 was observed at low shaking frequency (between 110 and 150 rpm at 25 mm shaking diameter). Probably, the oxygen concentration in the medium could have influence on the expression. The high oxygen concentration in the medium may lead to the oxidative inactivation of cytochrome P450 (Karuzina and Archakov 1994b).

The gene has a 69% GC-content, however a possible limited expression level by secondary structure of mRNA due to high GC-content was not observed. The yield of active CYP116B3 enzyme was comparable to those for CYP102A2 in the same system (Budde, Maurer et al. 2004). Ishida et al. reported that introduction of a leader open reading frame upstream of the initiation codon could relieve the inhibitory effect (Ishida and Oshima 1994). In our case the pET28a(+) vector carries an N-terminal His-Tag®/thrombin/T7-Tag® which may play the roll of the leading opening reading frame upstream of the target P450 gene.

Attachment of a hexa histidine tag (His-Tag) to a target gene during cloning is a common strategy facilitating easier identification and purification of recombinant proteins expressed in *E. coli* (Hengen 1995). The P450 CYP116B3 protein was fused to an N-terminal His-tag that

allows one step purification by immobilized metal affinity chromatography. N-terminal His-tag has usually little or no effect on the protein function (Mason, He et al. 2002). No inhibition of the P450 CYP116B3 activity was observed in this study, either. The his-tagged protein was selectively bound on Ni-Sepharose column and subsequently eluted by highly concentrated imidazole. Because imidazole reversibly inhibits cytochrome P450 activity, it was removed by dialysis using a Vivaspin membrane. The purity of the protein was 85% and the recovery was 73% after one step purification.

6.4. Protein characterization

Thin layer chromatography detected only FMN within the reductase domain, as the sequence alignment had predicted. FMN-containing reductases are usually involved in the electron transport from NADH to dioxygenases (Correll, Batie et al. 1992), not to monooxygenases. Most known bacterial fusion P450 monooxygenases such as CYP102A1, CYP120A2 and CYP102A3 (Narhi and Fulco 1986; Gustafsson, Roitel et al. 2004), as well as the eukaryotic membrane-bound CYP505 from *Fusarium oxysporum* (Nakayama, Takemae et al. 1996) contain FAD/FMN-containing reductases of the cytochrome P450 reductase (CPR) type (belonging to class II). Thus, these results as well as a high sequence similarity to P450RhF from *Rhodococcus* sp. NCIMB 9784 indicate that the new natural fusion protein P450 CYP116B3 from *Rhodococcus ruber* DSM 44319 belongs to class IV of cytochrome P450 systems postulated by Roberts *et al*. (Roberts, Celik et al. 2003). It consists of a heme protein, an FMN-containing reductase, and a ferredoxin. Measurement of the reduction of cytochrome c showed strong preference for NADPH over NADH like in case of P450RhF.

All fusion P450 monooxygenase systems are self-sufficient, as they do not require any additional proteins for electron transfer from NAD(P)H to the heme iron. A single peptide represents an easier component for industrial applications than a multi-component protein complex in terms of high-performance biotransformation or enzyme bioreactor. The self-sufficient P450 monooxygenases are therefore good candidates for industrial applications.

6.5. Substrate spectra

P450RhF from *Rhodococcus* sp. NCIMB 9784 has only been reported to catalyze dealkylation of 7-ethoxycoumarin (Roberts, Celik et al. 2003). The P450 monooxygenase CYP116B3 from *Rhodococcus ruber* DSM 44319 was found to hydroxylate a broad range of aromatic hydrocarbons. All oxidized substrates contain at least one benzene ring (Figure 3.29). We suggest that the benzene ring is required for the recognition by the P450 CYP116B3 monooxygenase. Nevertheless, benzene itself is not a suitable substrate for this enzyme. Benzene may be excessively small for the enzymes binding pocket and could not be positioned close enough to the heme group. Interestingly, toluene was oxidized only at the methyl group and not on the benzene ring. Other two alkyl aromatics, *m*-xylene and ethyl benzene, were hydroxylated exclusively at the alkyl side chain.

Several bacterial P450s have been reported to catalyze hydroxylation of aromatic hydrocarbons in the nature. Bacterial P450-catalyzed aromatic hydroxylation occurs in both biosynthetic and biodegradative pathways. $P450_{StiL}$ from *Stigmatella aurantiaca* is found in the biosynthesis of stigmatellin A, which is an aromatic electron-transport inhibitor containing a chromone ring with a polyketide side chain. $P450_{StiL}$ introduces a phenolic hydroxyl group into chromone ring (Gaitatzis, Silakowski et al. 2002). $P450_{TxtC}$ from *Streptomyces acidiscabies* is able to catalyze multiple oxidations of a single substrate. $P450_{TxtC}$ catalyzes hydroxylation at the α-carbon of the phenylalanine residue of thaxtomin D and then at the β-position of the aromatic ring to give thaxtomin A (Healy, Krasnoff et al. 2002). Bacterial P450-catalyzed hydroxylation of aromatic hydrocarbons has also been implicated in biodegradation. $P450_{TdtD}$ from *Pseudomonas diterpeniphilia* A19-6a has been implicated in the biodegradation of resin acids (Morgan and Wyndham 2002).

Engineered P450s have been shown to catalyze hydroxylation of aromatic hydrocarbons. The mutants of P450cam demonstrated activity towards polycyclic aromatics (Harford-Cross, Carmichael et al. 2000) and alkyl benzene compounds (Bell, Harford-Cross et al. 2001). As well as P450cam, P450 BM-3 has been engineered to hydroxylate polycyclic aromatic hydrocarbons (PAH) (Li, Ogawa et al. 2001a). The PAHs are ubiquitous industrial harmful pollutants (Shuttleworth and Cerniglia 1995). Hydroxylation of PAHs by P450 monooxygenases enables first step in their biodegradation (Li, Ogawa et al. 2001a). Although the oxidation activity towards all substrates by the P450 CYP116B3 monooxygenase is quite

low it is probably an evidence of its natural function, On the one hand, this property opens up new possibilities for optimization of bioremediation processes. On the other hand, the ability of P450 CYP116B3 to hydroxylate the alkyl chain of aromatics makes this enzyme a useful biocatalyst for synthesis.

The closely related strain *Rhodococcus* sp. NCIMB 9784 was reported to contain two different P450 monooxygenases with different substrate specificities. P450camr catalyzes 6-*endo*-hydroxylation of (*1R*)-(+)-camphor (Grogan, Roberts et al. 2002). However, this reaction was not observed with P450RhF monooxygenase (Roberts, Grogan et al. 2002). This may also explain in the case of *Rhodococcus ruber* DSM 44319 why the cyclohexane oxidation observed during *in vivo* biotransformation could not be reproduced by the purified P450 CYP116B3 enzyme.

6.6. Rational design and Site-directed mutagenesis

The natural substrate of CYP116B3 is still unknown. The turnover rate of 7-ethoxycoumarin to 7-hydroxycoumarin by wild-type P450 CYP116B3 is very low (about 1 nmol of substrate converted per 1 nmol of P450 CYP116B3 per minute). However, all components (heme protein, reductase and ferredoxin within a peptide chain) could support a fast electron transfer. Activities towards other substrates were also quite low.

In order to increase enzyme activity, on the one hand and to alter the substrate specificity, on the other hand, we changed volume of the binding pocket to fit smaller substrates (Xu, Bell et al. 2005). The alanine residue at position 109 was substituted by a bulky amino acid phenylalanine to promote binding and oxidation of cyclohexane or alpha-pinene. Unfortunately, after the substitution no oxidation activity towards smaller substrates was detected. Only an increased activity towards 7-ethoxycoumarin and PAHs was observed. A possible explanation is, a newly formed π-π staking of phenyl rings between phenylalanine and PAH probably stabilizes substrates in the active site of the mutant stronger than in wild-type (Williams, Cosme et al. 2003). The interaction between the substrate and activated oxygen was benefited, resulting in increased oxidation rate of substrate. However, this modification did not affect product distribution.

It is too difficult to make accurate prediction of the substitution without enough structural information of P450 CYP116B3. As mentioned in the introduction, the advantage of directed evolution is that no structural information is required; therefore we decided to apply directed enzyme evolution of P450 CYP116B3 to improve its activity.

6.7. Assay development

One of the most critical steps in directed evolution is to find a simple, economical, and high-throughput screening system to measure enzyme activity directly on a microtiter plate scale. There are many strategies to screen a mutant library of P450 monooxygenases. One possibility is an NADPH oxidation assay, based on the principle that reduced NAD(P)H absorb at 340 nm and oxidized $NAD(P)^+$ not, so NADPH oxidation (consumption) can be monitored at a wavelength of 340 nm. However, NADPH consumption and product formation are not always correlated, because the reducing equivalent from NADPH can be diverted in forming reduced oxygen. P450 monooxygenase oxidation activity can not be accurately measured by this method. More accurate and precise methods are usually based on direct measurements of substrate consumption or product formation. Especially well appropriate for HTS is color reaction. As mentioned in the Introduction, the pNCA assay is a colorimetrical assay using modified substrate for screening of hydroxylation activity towards fatty acids. Another colorimetric assay based on Gibbs' reagent can be used in high-throughput screening of aromatic hydroxylation. Gibbs' reagent (2,6-dichloroquinone-4-chloroimide) reacts with phenols yields blue color compound, which is proportional to the concentration of phenols (Quintana, Didion et al. 1997). The ADH (alcohol dehydrogenase) assay is based on a cascade of two enzymes to improve regioselective hydroxylation of alkane by P450 enzyme (Lentz, Feenstra et al. 2006). Hydroxylation of alkane yields alcohol, which is further oxidized by ADH, the reduction of co-factor is coupled to concentration of alcohol and can be measured in a spectrophotometer. However, ADH assay requires a second enzyme with high substrate specificity.

As the natural substrate of CYP116B3 is unknown, we used a test substrate 7-ethoxycoumarin to obtain a higher oxidation activity for potential application and to find out important positions in the active site of the enzyme (so-called "hot spots") for further site-direct mutagenesis. 7-Ethoxycoumarin is metabolized by many mammalian cytochrome P450

enzymes active in xenobiotics metabolism and has been used as a prototypic substrate to monitor P450 activity in liver microsomes of mammals (Yamazaki, Inoue et al. 1996). Moreover, the product of dealkylation of 7-Ethoxycoumarin, 7-hydroxycoumarin is fluorescent, and can be directly measured via fluorescence spectroscopy.

In this study, a whole-cell assay was preferred for a cost-effective high-throughput screening system. Cell membrane is a diffusion barrier for hydrophobic compounds. However, in the present of DMSO, the cells did not show any rate-limiting effect in comparison to permeabilized cells or free enzyme. Commonly, glucose is required for regeneration of NADPH in whole-cell assay. However, the reaction duration in this screening was very short, only 15 minutes, so an addition of glucose was not necessary. A disadvantage of a whole-cell assay is that enzyme activity could not be normalized by enzyme concentration. False positives and false negatives could not be distinguished from very good or very bad results. Therefore, the false positives should be removed in a secondary activity screening, wherein false negatives will be sorted out already after the preliminary assay.

6.8. Directed evolution of P450 CYP116B3

Microtiter plates offer a powerful tool for handling a large number of parallel reactions under nearly the same condition. The advantage of the microtiter plate is that the handling can be automated using robotic pipetting system. For expression of P450 CYP116B3 in microtiter plates, several conditions were tested. Square shaped deep-well microtiter plates facilitated an optimal cells growth because of sufficient oxygen supply compared with round shaped deep-wells (Hermann, Lehmann et al. 2003) Expression of P450 CYP116B3 in LB-medium showed better results than in TB-medium. An addition of a heme precursor, aminolevulinic (Schwaneberg, Otey et al. 2001), did not affect the expression level of P450 CYP116B3.

The oxidation activity of the wild-type P450 CYP116B3 is much lower than expected for a self-sufficient fusion protein. After four rounds of directed evolution, the dealkylation activity of the best mutant P450 CYP116B3 A86T T91S A109F L179F I267L (70A08) (Figure 4.1) towards 7-ethoxycoumarin was increased up to 222.5 nmol substrate nmol $P450^{-1}$ $minute^{-1}$, while the wild-type P450 CYP116B3 has only a turnover rate 0.917 per minute towards 7-ethoxycoumarin.

Except position 109, all mutations are located outside the binding pocket. Positions 86 and 91 are near to the substrate entry channel. At position 86 the hydrophobic alanine residue was substituted by a hydrophilic amino acid threonine. In position 91 threonine residue was substituted by a serine, which is similar to threonine but smaller. Substitutions in these two positions may force the substrate to enter the binding pocket. The side chains of positions 267 and 179 are sterically close to each other. However, a direct contribution of the mutated positions 267 and 179 to the dealkylation activity towards 7-ethoxycoumarin can not be explained. Interestingly, in the fourth round of directed evolution one of the two mutations had a substitution of phenylalanine for leucine at position 109, which was identified as a "hot spot" in the homology model. This position alanine was previously substituted by phenylalanine in the rational design approach. Both substitutions Ala109Leu or Ala109Phe resulted in increased dealkylation activity towards 7-ethoxycoumarin. Position 109 may be one of the key positions that determine activity of this enzyme.

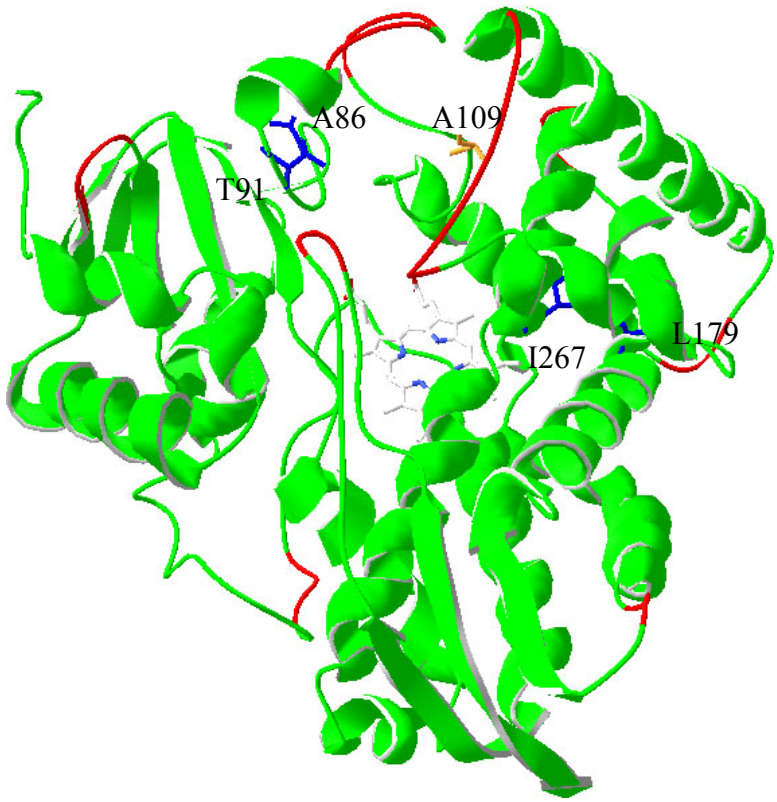

Figure 4.1: Computer model of CYP116B3 heme domain shows the mutated positions. The helix and sheets are shown in green. The loops are in red. The position 109 is highlighted in orange. Other positions, where substitution occurred, are highlighted in blue.

7. References

Adam, W., Z. Lukacs, et al. (2000). "Biocatalytic asymmetric hydroxylation of hydrocarbons with the topsoil-microorganism Bacillus megaterium." J Org Chem **65**(3): 878-82.

Akanuma, S., A. Yamagishi, et al. (1999). "Further improvement of the thermal stability of a partially stabilized Bacillus subtilis 3-isopropylmalate dehydrogenase variant by random and site-directed mutagenesis." Eur J Biochem **260**(2): 499-504.

Altschul, S. F. (1991). "Amino acid substitution matrices from an information theoretic perspective." J Mol Biol **219**(3): 555-65.

Anchordoguy, T. J., J. F. Carpenter, et al. (1992). "Temperature-dependent perturbation of phospholipid bilayers by dimethylsulfoxide." Biochim Biophys Acta **1104**(1): 117-22.

Arnold, F. H. (1998). "Design by directed evolution." Accounts of chemical research **31**(3): 125-131.

Atkins, W. M. and S. G. Sligar (1988). "The roles of active site hydrogen bonding in cytochrome P-450cam as revealed by site-directed mutagenesis." J Biol Chem **263**(35): 18842-9.

Bell, K. S., J. C. Philp, et al. (1998). "The genus Rhodococcus." J Appl Microbiol **85**(2): 195-210.

Bell, S. G., C. F. Harford-Cross, et al. (2001). "Engineering the CYP101 system for in vivo oxidation of unnatural substrates." Protein Eng **14**(10): 797-802.

Birnboim, H. C. and J. Doly (1979). "A rapid alkaline extraction procedure for screening recombinant plasmid DNA." Nucleic Acids Res **7**(6): 1513-23.

Boddupalli, S. S., R. W. Estabrook, et al. (1990). "Fatty acid monooxygenation by cytochrome P-450BM-3." J Biol Chem **265**(8): 4233-9.

Boddupalli, S. S., C. A. Hasemann, et al. (1992). "Crystallization and preliminary x-ray diffraction analysis of P450terp and the hemoprotein domain of P450BM-3, enzymes belonging to two distinct classes of the cytochrome P450 superfamily." Proc Natl Acad Sci U S A **89**(12): 5567-71.

Bornscheuer, U. T. (2002). "Methods to increase enantioselectivity of lipases and esterases." Curr Opin Biotechnol **13**(6): 543-7.

Bornscheuer, U. T. and M. Pohl (2001). "Improved biocatalysts by directed evolution and rational protein design." Curr Opin Chem Biol **5**(2): 137-43.

Budde, M., S. C. Maurer, et al. (2004). "Cloning, expression and characterisation of CYP102A2, a self-sufficient P450 monooxygenase from Bacillus subtilis." Appl Microbiol Biotechnol **66**(2): 180-6.

Cadwell, R. C. and G. F. Joyce (1992). "Randomization of genes by PCR mutagenesis." PCR Methods Appl **2**(1): 28-33.

Chen, R. (1999). "A general strategy for enzyme engineering." Trends Biotechnol **17**(9): 344-5.

Coia, G., A. Ayres, et al. (1997). "Use of mutator cells as a means for increasing production levels of a recombinant antibody directed against Hepatitis B." Gene **201**(1-2): 203-9.

Correll, C. C., C. J. Batie, et al. (1992). "Phthalate dioxygenase reductase: a modular structure for electron transfer from pyridine nucleotides to [2Fe-2S]." Science **258**(5088): 1604-10.

Cryle, M. J., Stok J. E. and De Voss J. J. (2003). "Reactions catalyzed by bacterial cytochromes P450." Aust J Chem **56**: 749-762.

Danielson, P. B., R. J. MacIntyre, et al. (1997). "Molecular cloning of a family of xenobiotic-inducible drosophilid cytochrome P450s: Evidence for involvement in host-plant allelochemical resistance." Proceedings of the National Academy of Sciences **94**: 10797-10802.

De Mot, R. and A. H. Parret (2002). "A novel class of self-sufficient cytochrome P450 monooxygenases in prokaryotes." Trends Microbiol **10**(11): 502-8.

Duetz, W. A., J. B. van Beilen, et al. (2001). "Using proteins in their natural environment: potential and limitations of microbial whole-cell hydroxylations in applied biocatalysis." Curr Opin Biotechnol **12**(4): 419-25.

Eigen, M. and W. C. Gardiner (1984). "Evolutionary molecular engineering based on RNA replication." Pure Appl. Chem **56**: 967¨C978.

Farinas, E. T., U. Schwaneberg, et al. (2000). "Directed Evolution of a Cytochrome P450 Monooxygenase for Alkane Oxidation." order **100**: 15.

Fersht, A. (1999). Structure and Mechanism in Protein Science. New York, W. H. Freeman.

Feyereisen, R. (1999). "Insect P450 enzymes." Annu Rev Entomol **44**: 507-33.

Finnerty, W. R. (1992). "The biology and genetics of the genus Rhodococcus." Annu Rev Microbiol **46**: 193-218.

Fruetel, J. A., R. L. Mackman, et al. (1994). "Relationship of active site topology to substrate specificity for cytochrome P450terp (CYP108)." J Biol Chem **269**(46): 28815-21.

Gaitatzis, N., B. Silakowski, et al. (2002). "The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase." J Biol Chem **277**(15): 13082-90.

Gliszczynska, A. and A. Koziolowa (1998). "Chromatographic determination of flavin derivatives in baker's yeast." J Chromatogr A **822**(1): 59-66.

Goodfellow, M. (1989). "Genus Rhodococcus." Bergey¡¯s Manual of Systematic Bacteriology **4**: 2362-2371.

Goodfellow, M., G. Alderson, et al. (1998). "Rhodococcal systematics: problems and developments." Antonie Van Leeuwenhoek **74**(1-3): 3-20.

Gotoh, O. (1992). "Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences." J Biol Chem **267**(1): 83-90.

Graham, S. E. and J. A. Peterson (1999). "How similar are P450s and what can their differences teach us?" Arch Biochem Biophys **369**(1): 24-9.

Gray, K. A., O. S. Pogrebinsky, et al. (1996). "Molecular mechanisms of biocatalytic desulfurization of fossil fuels." Nat Biotechnol **14**(13): 1705-9.

Greener, A., M. Callahan, et al. (1997). "An efficient random mutagenesis technique using an E. coli mutator strain." Mol Biotechnol **7**(2): 189-95.

Grogan, G., G. A. Roberts, et al. (2002). "P450(camr), a cytochrome P450 catalysing the stereospecific 6- endo-hydroxylation of (1 R)-(+)-camphor." Appl Microbiol Biotechnol **59**(4-5): 449-54.

Groves, J. T. (1985). "Key Elements of the Chemistry of Cytochrome P-450: The Oxygen Rebound Mechanism." Journal of Chemical Education **62**(11): 928-31.

Guengerich, F. P. (2001b). "Uncommon P450-catalyzed reactions." Curr Drug Metab **2**(2): 93-115.

Gustafsson, M. C., O. Roitel, et al. (2004). "Expression, purification, and characterization of Bacillus subtilis cytochromes P450 CYP102A2 and CYP102A3: flavocytochrome homologues of P450 BM3 from Bacillus megaterium." Biochemistry **43**(18): 5474-87.

Harford-Cross, C. F., A. B. Carmichael, et al. (2000). "Protein engineering of cytochrome p450(cam) (CYP101) for the oxidation of polycyclic aromatic hydrocarbons." Protein Eng **13**(2): 121-8.

Hasler, J. A. (1999). "Pharmacogenetics of cytochromes P450." Mol Aspects Med **20**(1-2): 12-24, 25-137.

Healy, F. G., S. B. Krasnoff, et al. (2002). "Involvement of a cytochrome P450 monooxygenase in thaxtomin A biosynthesis by Streptomyces acidiscabies." J Bacteriol **184**(7): 2019-29.

Hedegaard, J. and I. C. Gunsalus (1965). "Mixed function oxidation. IV. An induced methylene hydroxylase in camphor oxidation." J Biol Chem **240**(10): 4038-43.

Heiner, C. R., K. L. Hunkapiller, et al. (1998). "Sequencing multimegabase-template DNA with BigDye terminator chemistry." Genome Res **8**(5): 557-61.

Hengen, P. (1995). "Purification of His-Tag fusion proteins from Escherichia coli." Trends Biochem Sci **20**(7): 285-6.

Hermann, R., M. Lehmann, et al. (2003). "Characterization of gas-liquid mass transfer phenomena in microtiter plates." Biotechnology and Bioengineering **81**(2): 178-186.

Honkakoski, P. and M. Negishi (2000). "Regulation of cytochrome P450 (CYP) genes by nuclear receptors." Biochem J **347**(Pt 2): 321-37.

Hyun, C. G., J. M. Kim, et al. (1998). "An efficient approach for cloning P450 hydroxylase genes from actinomycetes." Journal of Microbiology and Biotechnology **8**(3): 295-299.

Ishida, M. and T. Oshima (1994). "Overexpression of genes of an extreme thermophile Thermus thermophilus, in Escherichia coli cells." J Bacteriol **176**(9): 2767-70.

Joo, H., Z. Lin, et al. (1999). "Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation." Nature **399**(6737): 670-3.

Karuzina, II and A. I. Archakov (1994b). "The oxidative inactivation of cytochrome P450 in monooxygenase reactions." Free Radic Biol Med **16**(1): 73-97.

Kitazume, T., N. Takaya, et al. (2000). "Fusarium oxysporum fatty-acid subterminal hydroxylase (CYP505) is a membrane-bound eukaryotic counterpart of Bacillus megaterium cytochrome P450BM3." J Biol Chem **275**(50): 39734-40.

Klingenberg, M. (1958). "Pigments of rat liver microsomes." Arch Biochem Biophys **75**(2): 376-86.

Kunst, F., N. Ogasawara, et al. (1997). "The complete genome sequence of the gram-positive bacterium Bacillus subtilis." Nature **390**(6657): 249-56.

Kurtzman, A. L., S. Govindarajan, et al. (2001). "Advances in directed protein evolution by recursive genetic recombination: applications to therapeutic proteins." Curr Opin Biotechnol **12**(4): 361-70.

Lambeth, J. D., D. W. Seybert, et al. (1982). "Steroidogenic electron transport in adrenal cortex mitochondria." Mol Cell Biochem **45**(1): 13-31.

Lechevalier, H. A. (1989). "Nocardioform actinomycetes." Bergey's manual of systematic bacteriology **4**: 2348-2404.

Lee, S. Y. (1996). "Bacterial polyhydroxyalkanoates." Biotechnology and bioengineering **49**(1): 1-14.

Lentz, O., A. Feenstra, et al. (2006). "Altering the regioselectivity of cytochrome P450 CYP102A3 of Bacillus subtilis by using a new versatile assay system." Chembiochem **7**(2): 345-50.

Lentz, O., Q. S. Li, et al. (2001). "Modification of the fatty acid specificity of cytochrome P450 BM-3 from Bacillus megaterium by directed evolution: a validated assay." Journal of Molecular Catalysis B: Enzymatic **15**(4): 123-133.

Leung, D. W., E. Chen, et al. (1989). "A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction." Technique **1**(1): 11¨C15.

Li, Q. S., J. Ogawa, et al. (2001a). "Engineering cytochrome P450 BM-3 for oxidation of polycyclic aromatic hydrocarbons." Appl Environ Microbiol **67**(12): 5735-9.

Li, Q. S., U. Schwaneberg, et al. (2001). "Rational evolution of a medium chain-specific cytochrome P-450 BM-3 variant." Biochim Biophys Acta **1545**(1-2): 114-21.

Liu, L., R. D. Schmid, et al. (2006). "Cloning, expression, and characterization of a self-sufficient cytochrome P450 monooxygenase from Rhodococcus ruber DSM 44319." Appl Microbiol Biotechnol.

Mansuy, D. (1994). "Cytochromes P-450 and Model Systems - Great Diversity of Catalyzed-Reactions." Pure and Applied Chemistry **66**(4): 737-744.

Marchler-Bauer, A., J. B. Anderson, et al. (2005). "CDD: a Conserved Domain Database for protein classification." Nucleic Acids Res **33**(Database issue): D192-6.

Martinez, C. A. and J. D. Stewart (2000). "Cytochrome P450s: Potential Catalysts for Asymmetric Olefin Epoxidations." Current Organic Chemistry **4**: 263-282.

Mason, A. B., Q. Y. He, et al. (2002). "Differential effect of a his tag at the N- and C-termini: functional studies with recombinant human serum transferrin." Biochemistry **41**(30): 9448-54.

Min, G. S. and J. R. Powell (1998). "Long-distance genome walking using the long and accurate polymerase chain reaction." Biotechniques **24**(3): 398-400.

Mishra, R. N., S. L. Singla-Pareek, et al. (2002). "Directional genome walking using PCR." Biotechniques **33**(4): 830-2, 834.

Mitani, F. (1979). "Cytochrome P450 in adrenocortical mitochondria." Mol Cell Biochem **24**(1): 21-43.

Morgan, C. A. and R. C. Wyndham (2002). "Characterization of tdt genes for the degradation of tricyclic diterpenes by Pseudomonas diterpeniphila A19-6a." Can J Microbiol **48**(1): 49-59.

Murai, N., F. Skoog, et al. (1980). "Relationships between cytokinin production, presence of plasmids, and fasciation caused by strains of Corynebacterium fascians." Proc Natl Acad Sci U S A **77**(1): 619-623.

Nagano, S., J. R. Cupp-Vickery, et al. (2005). "Crystal structures of the ferrous dioxygen complex of wild-type cytochrome P450eryF and its mutants, A245S and A245T: investigation of the proton transfer system in P450eryF." J Biol Chem **280**(23): 22102-7.

Nagano, S., H. Li, et al. (2003). "Crystal structures of epothilone D-bound, epothilone B-bound, and substrate-free forms of cytochrome P450epoK." J Biol Chem **278**(45): 44886-93.

Nagy, I., G. Schoofs, et al. (1995). "Degradation of the thiocarbamate herbicide EPTC (S-ethyl dipropylcarbamothioate) and biosafening by Rhodococcus sp. strain NI86/21 involve an inducible cytochrome P-450 system and aldehyde dehydrogenase." J Bacteriol **177**(3): 676-87.

Nakayama, N., A. Takemae, et al. (1996). "Cytochrome P450foxy, a catalytically self-sufficient fatty acid hydroxylase of the fungus Fusarium oxysporum." J Biochem (Tokyo) **119**(3): 435-40.

Narhi, L. O. and A. J. Fulco (1986). "Characterization of a catalytically self-sufficient 119,000-dalton cytochrome P-450 monooxygenase induced by barbiturates in Bacillus megaterium." J Biol Chem **261**(16): 7160-9.

Nelson, D. R., L. Koymans, et al. (1996). "P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature." Pharmacogenetics **6**(1): 1-42.

Nickerson, D. P., C. F. Harford-Cross, et al. (1997). "The catalytic activity of cytochrome P450cam towards styrene oxidation is increased by site-specific mutagenesis." FEBS Lett **405**(2): 153-6.

Nordblom, G. D., R. E. White, et al. (1976). "Studies on hydroperoxide-dependent substrate hydroxylation by purified liver microsomal cytochrome P-450." Arch Biochem Biophys **175**(2): 524-33.

O'Keefe, D. P., S. M. Lau, et al. (1993). "Efficiency and substrate specificity of Streptomyces griseolus cytochromes P-450SUI and P-450SU2 mono-oxygenase reactions." Biochem Soc Trans **21**(4): 1073-7.

Ogliaro, F., N. Harris, et al. (2000). "A model "rebound" mechanism of hydroxylation by cytochrome P450: Stepwise and effectively concerted pathways, and their reactivity patterns." Journal of the American Chemical Society **122**(37): 8977-8989.

Olsen, M., B. Iverson, et al. (2000). "High-throughput screening of enzyme libraries." Curr Opin Biotechnol **11**(4): 331-7.

Omura, T. and R. Sato (1964a). "The Carbon Monoxide-Binding Pigment of Liver Microsomes. I. Evidence for Its Hemoprotein Nature." J Biol Chem **239**: 2370-8.

Omura, T. and R. Sato (1964b). "The Carbon Monoxide-Binding Pigment of Liver Microsomes. Ii. Solubilization, Purification, and Properties." J Biol Chem **239**: 2379-85.

Peterson, J. A. and S. E. Graham (1998). "A close family resemblance: the importance of structure in understanding cytochromes P450." Structure **6**(9): 1079-85.

Pieper, U. and A. Steinbuchel (1992). "Identification, cloning and sequence analysis of the poly(3-hydroxyalkanoic acid) synthase gene of the gram-positive bacterium Rhodococcus ruber." FEMS Microbiol Lett **75**(1): 73-9.

Poulos, T. L., B. C. Finzel, et al. (1985). "The 2.6-A crystal structure of Pseudomonas putida cytochrome P-450." J Biol Chem **260**(30): 16122-30.

Poulos, T. L., B. C. Finzel, et al. (1987). "High-resolution crystal structure of cytochrome P450cam." J Mol Biol **195**(3): 687-700.

Quintana, M. G., C. Didion, et al. (1997). "Colorimetric method for a rapid detection of oxygenated aromatic biotransformation products." Biotechnology Techniques **11**(8): 585-587.

Reetz, M. T. and K. E. Jaeger (1999). "Superior biocatalysts by directed evolution." Top. Curr. Chem **200**: 31-57.

Roberts, G. A., A. Celik, et al. (2003). "A self-sufficient cytochrome p450 with a primary structural organization that includes a flavin domain and a [2Fe-2S] redox center." J Biol Chem **278**(49): 48914-20.

Roberts, G. A., G. Grogan, et al. (2002). "Identification of a new class of cytochrome P450 from a Rhodococcus sp." J Bacteriol **184**(14): 3898-908.

Ropp, J. D., I. C. Gunsalus, et al. (1993). "Cloning and expression of a member of a new cytochrome P-450 family: cytochrome P-450lin (CYP111) from Pseudomonas incognita." J Bacteriol **175**(18): 6028-37.

Schlichting, I., J. Berendzen, et al. (2000). "The catalytic pathway of cytochrome p450cam at atomic resolution." Science **287**(5458): 1615-22.

Schmid, A., J. S. Dordick, et al. (2001). "Industrial biocatalysis today and tomorrow." Nature **409**(6817): 258-68.

Schuler, M. A. and D. Werck-Reichhart (2003). "Functional genomics of P450s." Annu Rev Plant Biol **54**: 629-67.

Schwaneberg, U., C. Otey, et al. (2001). "Cost-effective whole-cell assay for laboratory evolution of hydroxylases in Escherichia coli." J Biomol Screen **6**(2): 111-7.

Schwaneberg, U., C. Schmidt-Dannert, et al. (1999). "A continuous spectrophotometric assay for P450 BM-3, a fatty acid hydroxylating enzyme, and its mutant F87A." Anal Biochem **269**(2): 359-66.

Schwede, T., J. Kopp, et al. (2003). "SWISS-MODEL: An automated protein homology-modeling server." Nucleic Acids Res **31**(13): 3381-5.

Scott, M. A., B. S. Graham, et al. (1995). "Rhodococcus equi--an increasingly recognized opportunistic pathogen. Report of 12 cases and review of 65 cases in the literature." Am J Clin Pathol **103**(5): 649-55.

Shuttleworth, K. L. and C. E. Cerniglia (1995). "Environmental aspects of PAH biodegradation." Appl Biochem Biotechnol **54**(1-3): 291-302.

Slusarczyk, H., S. Felber, et al. (2000). "Stabilization of NAD-dependent formate dehydrogenase from Candida boidinii by site-directed mutagenesis of cysteine residues." Eur J Biochem **267**(5): 1280-9.

Smith, G. C., D. G. Tew, et al. (1994). "Dissection of NADPH-cytochrome P450 oxidoreductase into distinct functional domains." Proc Natl Acad Sci U S A **91**(18): 8710-4.

Sono, M., M. P. Roach, et al. (1996). "Heme-Containing Oxygenases." Chem Rev **96**(7): 2841-2888.

Spee, J. H., W. M. de Vos, et al. (1993). "Efficient random mutagenesis method with adjustable mutation frequency by use of PCR and dITP." Nucleic Acids Res **21**(3): 777-8.

Spiegelman, S., I. Haruna, et al. (1965). "The synthesis of a self-propagating and infectious nucleic acid with a purified enzyme." Proc Natl Acad Sci U S A **54**(3): 919-27.

Stemmer, W. P. (1994a). "DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution." Proc Natl Acad Sci U S A **91**(22): 10747-51.

Stemmer, W. P. (1994b). "Rapid evolution of a protein in vitro by DNA shuffling." Nature **370**(6488): 389-91.

Steyer, F. and K. Sundmacher (2004). "VLE and LLE data for the system cyclohexane plus cyclohexene plus water plus cycloblexanol." Journal of Chemical and Engineering Data **49**(6): 1675-1681.

Tyson, C. A., J. D. Lipscomb, et al. (1972). "The role of putidaredoxin and P450 cam in methylene hydroxylation." J Biol Chem **247**(18): 5777-84.

van Beilen, J. B., W. A. Duetz, et al. (2003). "Practical issues in the application of oxygenases." Trends Biotechnol **21**(4): 170-7.

van Beilen, J. B., M. G. Wubbolts, et al. (1994). "Genetics of alkane oxidation by Pseudomonas oleovorans." Biodegradation **5**(3-4): 161-74.

Vermilion, J. L., D. P. Ballou, et al. (1981). "Separate roles for FMN and FAD in catalysis by liver microsomal NADPH-cytochrome P-450 reductase." J Biol Chem **256**(1): 266-77.

Walton, A. Z. and J. D. Stewart (2004). "Understanding and improving NADPH-dependent reactions by nongrowing Escherichia coli cells." Biotechnol Prog **20**(2): 403-11.

Werck-Reichhart, D. and R. Feyereisen (2000). "Cytochromes P450: a success story." Genome Biol **1**(6): REVIEWS3003.

White, R. E. (1991). "The involvement of free radicals in the mechanisms of monooxygenases." Pharmacol Ther **49**(1-2): 21-42.

Williams, P. A., J. Cosme, et al. (2003). "Crystal structure of human cytochrome P450 2C9 with bound warfarin." Nature **424**(6947): 464-8.

Wolfe, M. D. and J. D. Lipscomb (2003). "Hydrogen peroxide-coupled cis-diol formation catalyzed by naphthalene 1,2-dioxygenase." J Biol Chem **278**(2): 829-35.

Wong, T. S., F. H. Arnold, et al. (2004). "Laboratory evolution of cytochrome P450BM-3 monooxygenase for organic cosolvents." Biotechnol. Bioeng **85**: 351-358.

Xu, F., S. G. Bell, et al. (2005). "The Heme Monooxygenase Cytochrome P450 (cam) Can Be Engineered to Oxidize Ethane to Ethanol." Angew Chem Int Ed Engl.

Yamada, H. and M. Kobayashi (1996). "Nitrile hydratase and its application to industrial production of acrylamide." Biosci Biotechnol Biochem **60**(9): 1391-400.

Yamazaki, H., K. Inoue, et al. (1996). "7-Ethoxycoumarin O-deethylation catalyzed by cytochromes P450 1A2 and 2E1 in human liver microsomes." Biochem Pharmacol **51**(3): 313-9.

Zhao, H. and F. H. Arnold (1997). "Combinatorial protein design: strategies for screening protein libraries." Curr Opin Struct Biol **7**(4): 480-5.

8. Curriculum vitae

Personal details

Family Name:	Liu
	First name:Luo
Nationality:	Chinese (People's Republic of China)
Address:	Institute of Technical Biochemistry
	Allmandring 31
	D-70569, Stuttgart, Germany
Telephone:	+49 711 685 69984
Mobile:	+49 179 324 8971
Email:	luo.liu@itb.uni-stuttgart.de

Education

Title of Ph.D. Thesis at Institute of Technical Biochemistry **February 2003 –August 2007**

Stuttgart University, Germany

Working on the topic of biocatalysis using P450 monooxygenase.

"Cloning, characterization and engineering of a self-sufficient cytochrome P450 monooxygenase from *Rhodococcus ruber* DSM 44319"

Title of Master Thesis at Institute of Technical Biochemistry **May 2002 – February 2003**

Stuttgart University, Germany

"Using molecular-dynamic simulation to simulate the stereo-selectivity of P450 monooxygenase"

Curriculum "Technical Biology" **March 1999 – May 2002**

Master of Science

Stuttgart University, Germany

Curriculum "Virology & Molecular Biology" **September 1990 – July 1994**

Bachelor of Science

Wuhan University, China

Work experience

Associate Engineer **August 1994 – August 1998**

Beijing Pharmaceutical Factory, China

"Responsible for chemical production and quality control of para-aminosalicylic acid"

Publications

- Liu L, Schmid RD and Urlacher VB (2006) Cloning, expression, and characterization of a self-sufficient cytochrome P450 monooxygenase from *Rhodococcus ruber* DSM 44319. Appl Microbiol Biotechnol, April 2006. DOI: 10.1007/s00253-006-0355-0
- Engineering of cytochrome P450 CYP116B3 using directed evolution (In preparation)