

**The cytochrome P450 complement of the myxobacterium
Sorangium cellulosum So ce56 and characterization of two
members, CYP109D1 and CYP260A1**

**Dissertation
zur Erlangung des Grades
des Doktors der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultät III
Chemie, Pharmazie, Bio- und Werkstoffwissenschaften
der Universität des Saarlandes**

von

**Yogan Khatri
(M. Sc. in Medical Microbiology)**

**Saarbrücken
2009**

INDEX

PUBLICATIONS RESULTING FROM THIS WORK

DECLARATION

ABBREVIATIONS

ABSTRACT

Abstract (English version)

Abstract (German version)

SUMMARY

Summary (English version)

Zusammenfassung (German version)

LIST OF FIGURES

LISTS OF TABLES

Table of content	Page nos
------------------	----------

GENERAL INTRODUCTION

1. 1. Cytochrome P450

1. 1. 1.	General aspects	1
1. 1. 2.	Heme Structure in P450	2
1. 1. 3.	Structure of cytochrome P450	4
1. 1. 4.	Reactions and mechanism	5
1. 1. 5.	Electron transfer in cytochrome P450 systems	7
1. 1. 6.	Substrate binding properties of cytochrome P450	9
1. 1. 7.	Reactions catalyzed by P450	10

1. 2. Myxobacteria

1. 2. 1.	Taxonomy	11
1. 2. 2.	Characteristic features of myxobacteria	13
1. 2. 2. 1	Gliding motility	13
1. 2. 2. 2.	Fruiting body formation	13
1. 2. 3.	Application of myxobacterial secondary metabolites	14
1. 2. 4.	<i>Sorangium cellulosum</i> So ce56	15

CHAPTER I

The cytochrome P450 complement (CYPome) of *Sorangium cellulosum* So ce56

1. Introduction	16
2. Objectives	17

3. Materials and methods

3. 1. Materials

3. 1. 1.	Bacterial Strains	18
3. 1. 2.	Proteins	18
3. 1. 3.	Oligonucleotides	18
3. 1. 4.	Vectors	20

3. 2. Methods

3. 2. 1. Bioinformatic study

3. 2. 1. 1.	Bioinformatic analysis of P450s of <i>So ce56</i>	20
3. 2. 1. 2.	Phylogenetic analysis	21
3. 2. 1. 3.	Other bioinformatics studies related to P450s and neighboring ORF	21

3. 2. 2. Molecular biology methods

3. 2. 2. 1.	Isolation of genomic DNA	21
3. 2. 2. 2.	Plasmid isolation, purification and determination of the nucleic acid concentration	22
3. 2. 2. 3.	Digestion of DNA with endonucleases	22
3. 2. 2. 4.	Ligation of DNA	22
3. 2. 2. 5.	Agarose gel electrophoresis and extraction of DNA from agarose gel	23
3. 2. 2. 6.	Polymerase chain reaction and molecular cloning	23
3. 2. 2. 7.	Site-directed mutagenesis	26

3. 2. 3. Microbiological Methods

3. 2. 3. 1.	Cultivation and storage of bacteria	27
3. 2. 3. 2.	Preparation of competent cells	27
3. 2. 3. 2. 1.	Preparation of electro-competent cells	27
3. 2. 3. 2. 2.	Preparation of chemical competent cells	27
3. 2. 3. 3.	Transformation of <i>E. coli</i>	28
3. 2. 3. 3. 1.	Heat shock transformation	28
3. 2. 3. 3. 2.	Electroporation	28

3. 2. 4. Protein expression and purification

3. 2. 4. 1.	Heterologous gene expression of His ₆ -cytochrome P450	28
-------------	---	----

3. 2. 4. 2.	Cell lysis by sonication	29
3. 2. 4. 3.	Protein purification	29
3. 2. 4. 4.	Determination of the protein concentration	30
3. 2. 4. 5.	SDS (Sodium Dodecylsulfate) Polyacrylamide Gel Electrophoresis	30
3. 2. 5. Spectrophotometric characterization		
3. 2. 5. 1.	Spectrophotometric characterization of P450s of So ce56	30
3. 2. 5. 2.	Investigation of electron transfer partners	30
3. 2. 6. Biophysical characterization		
3. 2. 6. 1.	Circular dichroism (CD) measurement	31
3. 2. 6. 2.	Electron Paramagnetic Resonance (EPR) measurement	31
3. 2. 7. Mass measurement		33
3. 2. 8. Screening of compound library		33
4. Results		
4. 1. Bioinformatic analysis		
4. 1. 1.	Cytochrome P450 complement of <i>Sorangium cellulosum</i> So ce56 (CYPome of So ce56)	35
4. 1. 2.	Phylogenetic relationship of So ce56 CYPome	37
4. 1. 3.	The relationship of P450s of So ce56	38
4. 2. Genome mining of cytochrome P450		
4. 2. 1.	P450 potentially involved in secondary metabolite formation	41
4. 2. 2.	P450 with potential physiological role	42
4. 2. 3.	P450 clustered with regulatory elements	43
4. 2. 4.	P450 clustered with carbohydrate related regulatory elements	45
4. 2. 5.	P450 clustered with a terpene cyclase	45
4. 2. 6.	P450 with unique cluster	46
4. 2. 7.	P450 clustered with hypothetical protein	47
4. 3. Cloning, expression and purification of So ce56 cytochrome P450		
4. 3. 1.	Amplification and cloning of So ce56 CYPome	48

4. 3. 2.	Heterologous expression of So ce56 CYPome	48
4. 4.	Spectrophotometric characterization	50
4. 5.	Biophysical characterization	
4. 5. 1.	Circular Dichroism (CD) measurement of So ce56 cytochrome P450	55
4. 5. 2.	Electronic paramagnetic resonance (EPR) spectroscopy	57
4. 6.	Mass determination and conformation of active form of P450	60
4. 7.	Determination of electron donors	62
4. 8.	Screening of compound library	63
5.	Discussion	
5. 1.	Bioinformatic analysis	65
5. 2.	Genome mining of cytochrome P450	66
5. 3.	Expression, purification and characterization of cytochrome P450	68
6.	Outlook	72
7.	Literature review	73-85

CHAPTER II

A. CYP109D1 from *Sorangium cellulosum* So ce56: A fatty acid hydroxylase

A. 1.	Introduction	86
A. 2.	Objective	87
A. 3.	Materials and methods	
A. 3. 1.	Materials	88
A. 3. 2.	Methods	88
A. 3. 3.	Spin state shift and substrate dissociation constant determination	88
A. 3. 4.	<i>In vitro</i> assay of fatty acid and derivatization	88
A. 3. 5.	Fatty acid hydroxylation and GC/MS analysis	89
A. 4.	Results	
A. 4. 1.	Search for a fatty acid hydroxylating P450 of So ce56	91
A. 4. 2.	Spin state shift and substrate binding	93
A. 4. 3.	Characterization of different members of CYP109	

	family of So ce56	94
A. 4. 4.	Fatty acid hydroxylation and characterization	95
A. 3. 4. 1.	Hydroxylation of saturated fatty acid by CYP109D1	96
A. 4. 5.	Comparison of hydroxylation patterns of saturated fatty acids obtained by other CYP109 families of So ce56	102
A. 4. 6.	Hydroxylation of other fatty acids by CYP109D1	
A. 4. 6. 1.	Hydroxylation of unsaturated fatty acid by CYP109D1	103
A. 4. 6. 2.	Hydroxylation of branched fatty acid by CYP109D1	105
A. 5. Discussion		108
B. CYP109D1 from <i>Sorangium cellulosum</i> So ce56: A terpenoid epoxidase/hydroxylase		
B. 1. Introduction		
B. 1. 1.	Terpenoid compound	112
B. 1. 1. 1.	Geraniol and nerol	113
B. 1. 1. 2.	Limonene	113
B. 1. 1. 3.	Ionone	114
B. 1. 1. 4.	Role of cytochrome P450 in preparation of ionone derivatives	116
B. 2. Objective		116
B. 3. Material and methods		
B. 3. 1.	Materials	117
B. 3. 2.	Synthesized compound	117
B. 3. 3.	Methods	117
B. 3. 3. 1.	<i>In vitro</i> assay	118
B. 3. 3. 2.	Gas chromatography coupled with mass spectrometry (GC-MS) analysis	118
B. 3. 3. 3.	Thin layer chromatography (TLC)	119
B. 3. 3. 4.	Isolation of 3-hydroxy- α -ionone from reaction mixture	119
B. 3. 3. 5.	Computational methods	119
B. 4. Results		
B. 4. 1.	Expression and purification of recombinant CYP109D1	121
B. 4. 2.	Conversion of geraniol by CYP109D1	121
B. 4. 3.	Conversion of nerol by CYP109D1	122
B. 4. 4.	Conversion of limonene by CYP109D1	123

B. 4. 5.	Conversion of ionone by CYP109D1	
B. 4. 5. 1.	Conversion of α -ionone	125
B. 3. 5. 2.	Identification of the product of α -ionone oxidation	125
B. 4. 6.	Conversion of β -ionone	126
B. 4. 7	Homology modeling and docking of CYP109D1 with α -ionone and β -ionone	127
B. 5. Discussion		129
B. 6. Literature review		133-147

CHAPTER III

CYP260A1 from *Sorangium cellulosum* So ce56

1. Introduction		148
2. Objectives		149
3. Materials and methods		
3. 1. Materials		150
3. 2. Methods		
3. 2. 1.	Screening of a compound library	150
3. 2. 2.	Analytical methods	
3. 2. 2. 1.	<i>In vitro</i> steroid assay	150
3. 2. 2. 2.	<i>In vitro</i> nootkatone assay	150
3. 2. 2. 3.	High performance liquid chromatography (HPLC) of steroids and nootkatone	150
3. 2. 2. 4.	GC-MS analysis of nootkatone	151
4. Results		
4. 1.	Screening and validation of substrate/ligand	153
4. 2.	Selection of structural analogues of the Type I binding ligands	153
4. 3.	Binding of steroids to CYP260A1	155
4. 4.	Conversion of steroids by CYP260A1	156
4. 4. 1	Conversion of progesterone by CYP260A1	157
4. 4. 2.	Conversion of DOC by CYP260A1	159
4. 4. 2. 1.	Time dependent DOC conversion by CYP260A1	160
4. 5.	Conversion of nootkatone by CYP260A1	

4. 5. 1.	Selection of nootkatone as a substrate	163
4. 5. 2.	Binding of nootkatone to CYP260A1	164
4. 5. 3.	Conversion of nootkatone	165
4. 5. 4.	GC-MS of nootkatone conversion	166
5. Discussion		
5. 1.	Substrate screening and conversion of analogues compound.	167
5. 2.	Microbial conversion of progesterone and DOC	167
5. 3.	Triterpenoids and their biosynthesis in myxobacteria	172
5. 4.	Biotechnological application of CYP260A1	173
6. Outlook		174
7. Literature review		175-181
Acknowledgment		
Appendix		

Publications resulting from this work

A. Manuscripts

1. Kerstin Maria Ewen, Frank Hannemann, **Yogan Khatri**, Olena Perlova, Reinhard Kappl, Daniel Krug, Jürgen Hüttermann, Rolf Müller and Rita Bernhardt (2009). Genome mining in *Sorangium cellulosum* So ce56 - identification and characterization of the homologous electron transfer proteins of a myxobacterial cytochrome P450. *J. Biol. Chem.*, 284, 28590-8.
2. **Yogan Khatri**, Marco Girhard, Anna Romankiewicz, Frank Hannemann, Vlada B. Urlacher, Michael C. Hutter and Rita Bernhardt (2009). Regioselective hydroxylation of sesquiterpenoids by CYP109D1 from *Sorangium cellulosum* So ce56 (submitted to *Org. Biomol. Chem.*)
3. Marco Girhard, Tobias Klaus, **Yogan Khatri**, Rita Bernhardt and Vlada B. Urlacher (2009). Characterization of the versatile monooxygenase CYP109B1 from *Bacillus subtilis*. (under peer review in *Appl. Microbiol. Biotech.*)
4. **Yogan Khatri**, Frank Hannemann, Kerstin M. Ewen, Olena Perlova, Alexander O. Brachmann, Dominik Pistorius, Rolf Müller and Rita Bernhardt (2009). The cytochrome P450 complement (CYPome) of *Sorangium cellulosum* So ce56 and identification of CYP109D1 as a new fatty acid hydroxylase (resubmitting to *Chem. Biol.*)
5. **Yogan Khatri**, Frank Hannemann, Marco Girhard, Vlada B. Urlacher and Rita Bernhardt (2009). Molecular cloning, expression and purification of P450 enzymes from the myxobacterium *Sorangium cellulosum* So ce56 and characterization of three fatty acid hydroxylases genes (in preparation)

B. Patent application

Yogan Khatri¹, Rita Bernhardt¹, Frank Hannemann¹, Marco Girhard² and Vlada B. Urlacher². New biocatalyst for the epoxidation and hydroxylation of monoterpenes and sesquiterpenes (submitted).

¹Department of Biochemistry, Saarland University, D-66041 Saarbrücken, Germany

²Institute of Technical Biochemistry, Universitaet Stuttgart, Allmandring 31, 70569 Stuttgart, Germany.

C. Presentaion

- 1. Yogan Khatri**, Kerstin Ewen, Frank Hannemann, Olena Perlova; Rolf Müller and Rita Bernhardt (2008), The CYPome of myxobacteria (*Sorangium cellulosum*). Poster presentation - *In 9th International Symposium on Cytochrome P450 Biodiversity and Biotechnology*, June 8-12, Nice, France.

DECLARATION

I here by declare that this thesis is my original work and I have prepared it independently, except where specific reference is made to other sources. It has not been submitted in part either in German or in other country for any other degree.

Yogan Khatri

Place/date:

Signature:

ABBREVIATIONS

18-OH-B	18-hydroxy-corticosterone
ACTH	adrenocorticotrophic hormone
ACP	acyl carrier protein
AdR	adrenodoxin reductase
Adx	adrenodoxin
Aldo	aldosterone
Amp	ampicillin
APS	ammonium persulfate
ATP	adenosine triphosphate
CD	circular dichroism
CO	carbon monoxide
COD	carbon monoxide difference spectra
CYP	cytochrome P450
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxyribonucleotide triphosphate
DOC	deoxycorticosterone
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
EPR	electron paramagnetic resonance
ESI-TOF	electron-spray ionization-time of flight
Fad	flavine adenine dinucleotide
Fdx	ferredoxin
Fdr	ferredoxin reductase
FMN	flavine mononucleotide
GC-MS	gas chromatography - mass spectrometry
HPLC	high performance liquid chromatography
IPTG	isopropyl- β -D-thiogalactoside
Kb	kilobase
kDa	kilodalton

K _D	dissociation constant
Mbp	mega base pair
NADPH	nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance
OD	optical density
P450	cytochrome P450
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
Pfu	<i>Pyrococcus furiosus</i>
PKA	protein kinase K
PMSF	phenylmethylsulfonyl fluoride
RT	room temperature
SDS	sodium dodecyl sulfate
Taq	<i>Thermus aquaticus</i>
TEMED	<i>N,N,N',N'</i> -tetramethylethylenediamine
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet
vis	visible

Standard abbreviations for amino acids

A	Ala	Alanine
L	Leu	Leucine
R	Arg	Arginine
K	Lys	Lysine
N	Asn	Asparagine
M	Met	Methionine
D	Asp	Aspartic acid
F	Phe	Phenylalanine
C	Cys	Cysteine
P	Pro	Proline
Q	Gln	Glutamine
S	Ser	Serine
E	Glu	Glutamic acid

T	Thr	Threonine
G	Gly	Glycine
V	Val	Valine
H	His	Histidine
W	Trp	Tryptophan
I	Ile	Isoleucine
Y	Tyr	Tyrosine

Units

Current strength	Ampere	A
Concentration	Molar	M
	Micromolar	μM
Length	Meter	M
	Centimeter	cm
Mass	Gram	G
	Micro gram	μg
Molecular weight	Dalton	Da
	Kilodalton	kDa
Tension	Volt	V
Electricity	Watt	W
Temperature	degree Celsius	$^{\circ}\text{C}$
Volume	Liter	L
	Mililiter	ml
	Microliter	μl
Wave length	Nanometer	nm
Time	Second(s)	s
	Minute(s)	min
	Hour(s)	hr

ABSTRACT

This is a very first systematic study of cytochrome P450s from *Sorangium cellulosum* So ce56. In chapter I, the open reading frames (ORFs) encoding twenty-one cytochrome P450s with nine novel families were identified, and a phylogenetic tree and a physical map were constructed. All P450 ORFs were cloned, expressed in *Escherichia coli* in a soluble form, purified and characterized. The heterologous redox partners for nine of the P450s were identified. A compound library (~17,000 ligands) was screened with four P450s, and potential substrates and inhibitors were identified. In chapter II, CYP109D1 was shown as a new fatty acid hydroxylase. The GC-MS analysis of TMS-derivatized products of fatty acids showed sub-terminal ω -hydroxylation without any terminal (ω) product. In this chapter, CYP109D1 was also shown as a new biocatalyst for the epoxidation and hydroxylation of monoterpenes like geraniol, nerol and limonene. Furthermore, the selective hydroxylation of sesquiterpenoids, α -ionone and β -ionone to 3-hydroxy- α -ionone and 4-hydroxy- β -ionone, confirmed by ^1H NMR and ^{13}C NMR, are also shown. The regioselective hydroxylation of the sesquiterpenes was also explained by docking models. In chapter III, steroids (progesterone and DOC) were identified as analogues of potential substrates found by the screening of the compound library with CYP260A1. HPLC analysis of the *in vitro* products demonstrated the hydroxylation of both the steroids. Moreover, nootkatone, a sesquiterpene, was also shown to be hydroxylated and/or epoxidated by this novel P450.

ZUSAMMENFASSUNG

Diese Arbeit ist die bislang erste systematische Untersuchung der Cytochrome P450 aus *Sorangium cellulosum* So ce56. In Kapitel I wurden die Offenen Leserahmen (ORF) für 21 Cytochrome P450 mit neun neuen Familien identifiziert; ein phylogenetischer Baum und eine physikalische Genkarte wurden erstellt. Alle P450 ORFs wurden kloniert, in löslicher Form in *Escherichia coli* exprimiert, gereinigt und charakterisiert. Die heterologen Redoxpartner für neun der P450 wurden identifiziert. Eine Bibliothek (~17.000 Liganden) wurde mit vier P450 gescreent und potentielle Substrate und Inhibitoren wurden identifiziert. Kapitel II beschreibt CYP109D1 als neue Fettsäurehydroxylase. GC-MS Analyse der TMS-derivatisierten Produkte der Fettsäuren zeigte ω -Hydroxylierung, aber kein terminales (ω) Produkt. In diesem Kapitel wurde CYP109D1 als ein neuer Biokatalysator für die Epoxidierung und Hydroxylierung von Monoterpenen gezeigt. Zudem wurden die Sesquiterpenoide α -ionon und β -ionon ebenso selektiv zu 3-Hydroxy- α -ionon bzw. 4-Hydroxy- β -ionon hydroxyliert, was mittels ^1H NMR und ^{13}C NMR Analysen bestätigt werden konnte. Die regioselective

Hydroxylierung der Sesquiterpene konnten auch durch Docking-Modelle erklärt werden. In Kapitel III wurden die Steroide Progesteron und DOC als physiologisch signifikante Analoga der Substrat-Hits für CYP260A1 identifiziert. HPLC Analyse der *in vitro* Produkte zeigte die Hydroxylierung beider Steroide. Darüber hinaus wurde auch das Sesquiterpen Nootkaton von diesem neuen Cytochrom wurde hydroxyliert und/oder epoxidiert.

SUMMARY

The cytochrome P450 superfamily plays a critical role in the bioactivation and detoxification of a wide variety of xenobiotics, in the biosynthesis of endogenous compounds and in the production of secondary metabolites. It also catalyzes late-stage stereo- and regiospecific oxidations when associated with polyketide biosynthetic gene clusters in bacteria. The presence of huge numbers of P450s in the genome of *Sorangium cellulosum* So ce56 high-light the potential role of P450s during the complex physiology of this microorganism.

This work is a very first systematic study of cytochrome P450s done so far in myxobacteria. The open reading frames (ORFs) encoding twenty-one different cytochromes P450, eight ferredoxins and two reductases have been disclosed in the genome of *Sorangium cellulosum* So ce56.

The chapter I deals with the bioinformatics analysis, cloning, expression, purification and characterization of the *Sorangium cellulosum* So ce56 cytochrome P450 complement (CYPome). The bioinformatic analysis of the P450s revealed nine novel P450 families (CYP259, CYP260, CYP261, CYP262, CYP263, CYP264, CYP265, CYP266, and CYP267). A phylogenetic tree and a physical map of the *Sorangium cellulosum* So ce56 CYPome was constructed. The genomic mining of the *Sorangium cellulosum* So ce56 CYPome disclosed some important P450 features which could be useful for the prediction of potential substrates. We were able to find out that two of the P450s, CYP263A1 and CYP265A1, were clustered with a polyketidesynthase (PKS) and nonribosomal peptide synthetase (NRPS), respectively. Some P450s (CYP109C2, CYP110H1, CYP110J1, CYP260A1, CYP262A1, CYP262B1, and CYP261A1) are associated with regulatory elements. CYP264B1 was clustered with a terpene cyclase. Some of the other P450s (CYP124E1, CYP262A1 and CYP266A1) were clustered with neighbouring ORFs with potential physiological functions. In a unique cluster, two P450s (CYP109D1 and CYP259A1) were in close proximity and separated by only one ORF. The genomic mining also disclosed that none of the ferredoxins (Fdx) or ferredoxin reductases (FdR) was in an adjacent or neighbouring position with that of P450s. The ORFs of all the P450s of the CYPome were cloned, expressed in *Escherichia coli* in a soluble form and purified in an active form with a typical CO-difference peak at 448 or 450 nm. The characteristic apo- and holo-masses of some P450s were estimated and the biophysical properties (CD and EPR spectra) were studied for highly expressed P450s and for the representatives of new P450 families (CYP109C1, CYP109C2, CYP109D1, CYP260A1, CYP264A1, CYP264B1, and

CYP266A1). All measurements showed the typical nature of a P450 family member. Moreover, the absence of homologous redox partners was overcome for nine of the P450s (CYP109C1, CYP109C2, CYP109D1, CYP110J1, CYP124E1, CYP260A1, CYP264A1, CYP264B1, and CYP267B1) by heterologous electron donor partners from bovine adrenals, adrenodoxin (Adx) and adrenodoxin reductase (AdR). Furthermore, for four of the P450s (CYP109C1, CYP109C2, CYP260A1, and CYP266A1), potential substrates and inhibitors were screened with a compound library consisting of ~17,000 substances.

In the first part of Chapter II, considering the importance of fatty acids during the complex life cycle of myxobacteria, several potential fatty acid hydroxylases from *S. coelicolor* were identified. For the first time, a novel P450 enzyme, CYP109D1 was found to be a new fatty acid hydroxylase. The GC-MS analysis of the TMS-derivatized products of the fatty acids converted by CYP109D1 demonstrated the sub-terminal ω -hydroxylation without any terminal (ω) product for saturated fatty acids. The hydroxylation pattern of saturated fatty acid like, capric acid (C10) showed ω -1 and ω -2; lauric acid (C12) showed ω -1 to ω -3; tridecanoic acid (C13) and myristic acid (C14) showed ω -1 to ω -5, and palmitic acid (C16) showed ω -1 to ω -6 hydroxylation. Furthermore, a unsaturated fatty acid, oleic acid (18:1cis9), was also tested and showed selective hydroxylation at ω -6 and ω -7 positions. The branched fatty acids, 13-methyl myristic acid and 15-methyl myristic acid showed ω -2 to ω -6, and ω -2 to ω -5 hydroxylation, respectively. The second part of chapter II illustrates the role of CYP109D1 as a new biocatalyst for the epoxidation and hydroxylation of monoterpenes and sesquiterpenes. CYP109D1 was able to catalyze the epoxidation of linear compounds like geraniol and nerol. The geraniol was converted into 2,3-epoxygeraniol, 6,7-epoxygeraniol and 2,3-,6,7-diepoxygeraniol. In case of nerol, two epoxidated products, 2,3-epoxynerol and 6,7-epoxynerol, were obtained. The epoxidated products of both the linear terpenoids were confirmed with the synthesized product as standards. Moreover, the conversion of the cyclic monoterpene, limonene, into hydroxylated and epoxidated products has also been shown. CYP109D1 also converted the sesquiterpenoids, α -ionone and β -ionone into the selectively hydroxylated products of 3-hydroxy- α -ionone and 4-hydroxy- β -ionone, respectively, which was confirmed by ^1H NMR and ^{13}C NMR. The regioselective hydroxylation of the sesquiterpenes was also studied by docking studies using a computer model of CYP109D1 which gave a rational explanation for the observed selectivity and activities. All these data demonstrate not only the important role of CYP109D1 during terpenoid oxidation but also a potential biotechnological significance of this P450 for the production of selectively

epoxidated and hydroxylated terpenoid derivatives with higher market values than the original substrates.

The chapter III deals with the functional role of an additional novel member of the bacterial P450 family, CYP260A1. All the genes present in the close proximity of CYP260A1 in the genome code for hypothetical proteins. The individual homologue search done for each of the neighboring genes of CYP260A1 has no precise information for putative functions. Consequently, the potential putative substrates of CYP260A1 with respect to the genomic cluster could not be predicted. To overcome this, a compound library consisting of ~17,000 substances was screened with CYP260A1 and potential substrates as well as inhibitors were identified. Structural analogues of the Type I hits of the screening results were utilized and steroids as well as nootkatone were selected as potential substrates. The heterologous redox partners (Adx and AdR) were used for substrate (progesterone, DOC and nootkatone) conversion and the products were analysed in HPLC/GC-MS. CYP260A1 disclosed as a steroid hydroxylase and also sesquiterpene converter. CYP260A1 was able to hydroxylate progesterone and DOC. Progesterone was converted into more than three products, whereas DOC produced two major products during 1 hr of reaction time. One of the products has the same retention time as aldosterone and the other one was not identified. On successive increment of the incubation time of the *in vitro* assay, only the major product with the similar retention time of aldosterone was obtained. Since a steroid biosynthesis pathway in myxobacteria was already shown before, the steroid hydroxylation with CYP260A1 could have some physiological significance. Furthermore, CYP260A1 was also able to convert nootkatone, a sesquiterpene, into hydroxylated and/or epoxidated products. This gives new insights into the importance of CYP260A1 for steroid and terpenoid conversion having pharmaceutical and biotechnological applications.

ZUSAMMENFASSUNG

Die Cytochrom P450 Superfamilie spielt eine bedeutende Rolle in der Bioaktivierung und im Abbau einer großen Bandbreite von Xenobiotika, in der Biosynthese endogener Verbindungen und in der Produktion von sekundären Metaboliten. Darüber hinaus katalysieren sie, wenn sie in Bakterien mit Genklustern der Polyketidbiosynthese assoziiert sind, stereo- und regiospezifische Oxidationen. Die Anwesenheit einer großer Anzahl von P450 im Genom des *So ce56* unterstreicht die wichtige Rolle innerhalb der seiner komplexen Physiologie dieses Mikroorganismus.

Diese Arbeit repräsentiert die allererste systematische Untersuchung der Cytochrome P450 in Myxobakterien. Die offenen Leserahmen (OLRs) für 21 verschiedene Cytochrome P450, acht Ferredoxine und zwei Reduktasen wurden im Genom des *Sorangium cellulosum* *So ce56* identifiziert.

Das erste Kapitel dieser Arbeit befasst sich mit der bioinformatischen Analyse, Klonierung, Expression, Reinigung und Charakterisierung des *So ce56* Cytochrom P450 Komplements (CYPome). Die bioinformatische Analyse der P450 brachte neun neue P450 Familien (CYP259, CYP260, CYP261, CYP262, CYP263, CYP264, CYP265, CYP266 und CYP267) zum Vorschein. Davon ausgehend wurden ein phylogenetischer Baum und eine physikalische Genekarte der P450 erstellt. Die Gensuche des *So ce56* CYPome lieferte einige wichtige P450 Eigenschaften und könnte nützlich für die Substratvorhersage sein. Wir konnten zeigen, dass zwei der P450, CYP263A1 und CYP265A1, mit einer Polyketidsynthase (PKS) beziehungsweise mit einer nicht-ribosomalen Peptidsynthetase (NRPS) verbunden waren. Einige Cytochrome P450 (CYP109C2, CYP110H1, CYP110J1, CYP260A1, CYP262A1, CYP262B1, und CYP261A1) waren mit regulatorischen Elementen assoziiert. CYP264B1 war in einem Cluster mit einer Terpenzyklase angeordnet. Einige andere CYPome (CYP124E1, CYP262A1, und CYP266A1) waren mit den benachbarten OLRs, die potentielle physiologische Funktionen besitzen, verbunden. In einem einzigartigen Cluster wurden die zwei CYPome (CYP109D1 und CYP259A1) identifiziert. Diese lagen in enger Nachbarschaft und wurden nur durch einen einzigen OLR voneinander getrennt. Die Gensuche zeigte auch, dass keine der *So ce56* Ferredoxine (Fdx) oder Ferredoxin Reduktasen (FdR) benachbart oder näheren Umgebung eines der P450 war. Die OLR aller P450 wurden kloniert, erfolgreich in löslicher Form in *Escherichia coli* exprimiert und in aktiver Form mit einem typischen Peak bei 448/450 nm gereinigt. Die charakteristische Apo- und Holo-Massen der P450s wurden

bestimmt und die biophysikalische Eigenschaften (mittels CD und EPR Messungen) für einige sehr gut exprimierte Cytochrome oder für Repräsentanten der neuen P450 Familien (CYP109C1, CYP109C2, CYP109D1, CYP260A1, CYP264A1, CYP264B1 und CYP266A1) untersucht. Alle Messungen zeigten die typischen Merkmale eines Mitgliedes der P450 Familie. In Abwesenheit von homologen Redoxpartnern wurden die heterologen Elektronendonoren, bovines Adrenodoxin (Adx) und Adrenodoxin Reduktase (AdR) für neun der P450 (CYP109C1, CYP109C2, CYP109D1, CYP110J1, CYP124E1, CYP260A1, CYP264A1, CYP264B1 und CYP267B1) verwendet. Unterdessen wurde für vier dieser P450 (CYP109C1, CYP109C2, CYP260A1 und CYP266A1) mittels einer Ligandenbibliothek mit 17.000 Substanzen nach potentiellen Substraten und Inhibitoren gesucht. Dieses Kapitel befasst sich also mit den Grundlagen, Redoxpartnern und Substraten für die funktionale Untersuchung für einige dieser neuen Cytochrome P450.

Da Fettsäuren während des komplexen Lebenszyklus der Myxobakterien von großer Bedeutung sind, beschäftigt sich der erste Teil von Kapitel II, mit der Identifizierung mehrerer potentieller Fettsäurehydroxylasen aus *S. coelicolor*. Darüber hinaus wurde ein Mitglied der CYP109D Familie, CYP109D1, zum ersten Mal als eine neue Fettsäurehydroxylase beschrieben. Die GC-MS Analyse der TMS-derivatisierten Produkte der durch CYP109D1 umgewandelten Fettsäuren zeigte ω -Hydroxylierungen ohne terminale (ω) Produkte der gesättigten Fettsäuren. Das Hydroxylierungsmuster gesättigter Fettsäuren wie Caprylsäure (C10) zeigte ω -1 und ω -2; Laurinsäure (C12) zeigte ω -1 bis ω -3; Tridecansäure (C13) und Myristinsäure (C14) zeigten ω -1 bis ω -5, und Palmitinsäure (C16) zeigte ω -1 bis ω -6 Hydroxylierung. Auch die ungesättigte Fettsäure Ölsäure (18:1cis9) wurde getestet und zeigte selektive Hydroxylierungen an den Positionen ω -6 und ω -7. Die verzweigten Fettsäuren 13-Methylmyristinsäure und 15-Methylmyristinsäure zeigten ω -2 bis ω -6 beziehungsweise ω -2 bis ω -5 Hydroxylierungen.

Der zweite Teil von Kapitel II zeigt die Rolle des CYP109D1 als einen neuen Biokatalysator für die Epoxidierung und Hydroxylierung von Monoterpenen und Sesquiterpenen. CYP109D1 war in der Lage, die Epoxidierung linearer Verbindungen, wie z.B. Geraniol und Nerol, zu katalysieren. Geraniol wurde in 2,3-Epoxygeraniol, 6,7-Epoxygeraniol und 2,3- 6,7-Diepoxygeraniol umgewandelt. Im Falle des Nerols wurden zwei epoxidierte Produkte, 2,3-Epoxynerol und 6,7-Epoxynerol, erhalten. Die epoxidierten Produkte der beiden linearen Terpenoiden wurden mittels eines synthetisierten Produktes als Standard bestätigt. Darüber

hinaus konnte das zyklische Monoterpen Limonen in hydroxylierte und epoxidierte Produkte umgewandelt werden. Zudem hydroxylierte CYP109D1 die Sesquiterpenoide α -Ionon und β -Ionon ebenso selektiv zu 3-Hydroxy- α -ionon bzw. 4-Hydroxy- β -ionon, was durch ^1H NMR und ^{13}C NMR Analyse bestätigt werden konnte. Die regioselektive Hydroxylierung der Sesquiterpene wurde auch mittels Dockingstudien an einem Computermodell des CYP109D1 untersucht, welche eine logische Erklärung für die beobachteten Selektivitäten und Aktivitäten ergab. Alle diese Analysen zeigten nicht nur die bedeutende Rolle des CYP109D1 während der Terpenoidoxidation, sondern auch eine mögliche biotechnologische Bedeutung dieses P450 für die Produktion von selektiv epoxidierten und hydroxylierten Terpenoidderivaten mit höherem Marktwert als den Ausgangssubstraten.

Kapitel III dieser Arbeit behandelt die funktionale Rolle der neuen bakteriellen P450 Familie CYP260A1. Alle individuelle OLRs, die in nächster Umgebung zum CYP260A1 lagen, kodieren für hypothetische Proteine. Um den jeweiligen OLRs mögliche Funktionen zuzuordnen, wurde eine individuelle Homologiesuche durchgeführt. Diese lieferte aber keine genauen Informationen über putative Funktionen und somit waren alle OLRs hypothetische Proteine. Demzufolge konnten keine potentiellen, putativen Substrate für CYP260A1 vorausgesagt werden. Daher wurde eine Bibliothek mit ~17,000 Liganden mit CYP260A1 gescreent und sowohl potentielle Substrate als auch Inhibitoren identifiziert. Die strukturelle Analoga der Typ I Hits des Screenings und sowohl Steroide als auch Nootkaton wurden als potentielle Substrate ausgesucht. Die heterologen Redoxpartner (Adx und AdR) wurden zur Substratumwandlung von Progesteron, DOC und Nootkaton verwendet und die Produkte mittels HPLC/GC-MS analysiert. Es konnte zum ersten Mal gezeigt werden, dass CYP260A1 nicht nur als eine Steroidhydroxylase, sondern auch als ein Sesquiterpen-Umwandler fungiert. CYP260A1 hydroxylierte sowohl Progesteron als auch DOC. Die Hydroxylierung von Progesteron führte zu mehr als drei Produkten, die Hydroxylierung von DOC während einer Stunde Reaktionszeit führte zu zwei Hauptprodukten. Eines dieser Produkte hatte die gleiche Retentionszeit wie Aldosteron, während das andere nicht identifiziert wurde. Mit steigender Inkubationszeit des *in vitro* Assays wurde nur noch das Hauptprodukt, welches die gleiche Retentionszeit wie Aldosteron besaß, erhalten. Da ein Steroidbiosyntheseweg in Myxobakterien bereits bekannt ist, könnte die Steroidhydroxylierung mittels CYP260A1 von physiologischer Signifikanz sein. Zusätzlich war CYP260A1 auch in der Lage, Nootkaton, ein Sesquiterpen, in ein hydroxyliertes und/oder epoxidiertes Produkt umzuwandeln. Dies erlaubt neue Einblicke in die Bedeutung von CYP260A1 für die Umwandlung einiger anderer

Steroid- und Terpenoidverbindungen, die bereits pharmazeutisch und biotechnologisch angewendet werden.

LIST OF FIGURES

GENERAL INTRODUCTION

- Figure 1: Cytochrome P450 catalyzed hydroxylation reaction.
- Figure 2: Schematic representation of the structure of heme *b* (ferriprotoporphyrin).
- Figure 3: Crystal structure of CYP102A1 monooxygenase domain.
- Figure 4: Catalytic cycle of P450 monooxygenases.
- Figure 5: Rebound mechanism of the P450 catalysed hydroxylation.
- Figure 6: Schematic organization of different cytochrome P450 systems.
- Figure 7: The UV-difference spectra obtained by the addition of a substrate and inhibitor to cytochrome P450.
- Figure 8: Taxonomy of myxobacteria.
- Figure 9: Diagram of the life cycle of myxobacteria.
- Figure 10: Fruiting bodies of different myxobacterial species.
- Figure 11: Percentage distribution of the production of secondary metabolites from different myxobacterial strains.

CHAPTER I

- Figure I. 1: The representative diagram of micro titer plate that used for screening.
- Figure I. 2: Physical map of the *Sorangium cellulosum* So ce56 chromosome and distribution of cytochrome P450, ferredoxin (Fdx), and ferredoxin reductase (FdR) genes.
- Figure I. 3: Phylogenetic tree of *Sorangium cellulosum* So ce56 cytochrome P450.
- Figure I. 4: The genomic organization of CYP263A1.
- Figure I. 5: The genomic organization of CYP265A1.
- Figure I. 6: The genomic organization of P450s clustered with putative functional domain.
- Figure I. 7: The genomic organization of P450s clustered with regulatory elements.
- Figure I. 8: The genomic organization of P450s clustered with carbohydrate related regulatory elements.

- Figure I. 9: The genomic organization of P450s clustered with a terpene cyclase.
- Figure I. 10: A. The unique cluster of CYP109D1 and CYP259A1. B. The unique cluster of CYP267B1. C. The unique cluster of CYP264A1.
- Figure I. 11: SDS-PAGE of CYP109D1.
- Figure I. 12: SDS-PAGE of purified P450s of *So ce56*.
- Figure I. 13: Representative UV-visible characterization of *So ce56* P450s.
- Figure I. 14: The CO-difference spectras of *Sorangium cellulosum So ce56* P450.
- Figure I. 15: CD spectra of P450s in far-UV region.
- Figure I. 16: CD spectra of P450s in near-UV region.
- Figure I. 17: X-band EPR spectra of the CYP109 family.
- Figure I. 18: X-band EPR spectrum cytochrome P450 of *So ce56*.
- Figure I. 19: Determination of electron transfer partners for CYP109D1 and CYP264A1.

CHAPTER II

CHAPTER II.A

- Figure II. A. 1. The radial view of best maximum-parsimony tree obtained by PHYLIP analysis. .
- Figure II. A. 2: Spectral shift induced by binding lauric acid to CYP109D1.
- Figure II. A. 3: Spectral shift induced by binding myristic acid to CYP109D1
- Figure II. A. 4: Spectral shift induced by binding palmitic acid to CYP109D1
- Figure II. A. 5: GC chromatogram of subterminal ω -hydroxylation of saturated fatty acids by CYP109D1.
- Figure II. A. 6. 1: Mass spectra of the hydroxylated products of capric acid by CYP109D1.
- Figure II. A. 6. 2: Mass spectra of the hydroxylated products of lauric acid by CYP109D1.
- Figure II. A. 6. 3: Mass spectra of the hydroxylated products of tridecanoic by CYP109D1.
- Figure II. A. 6. 4: Mass spectra of the hydroxylated products of myristic acid by CYP109D1.
- Figure II. A. 6. 5: Mass spectra of the hydroxylated products of palmitic acid by CYP109D1.
- Figure II. A. 7: GC chromatogram of oleic acid of hydroxylation by CYP109D1.

- Figure II. A. 8: Mass spectra of the hydroxylated products of oleic acid by CYP109D1.
- Figure II. A. 9: Mass spectra of the hydroxylated products of 13-methyl myristic acid by CYP109D1.
- Figure II. A. 10: Mass spectra of the hydroxylated products of 15-methyl palmitic acid by CYP109D1.

CHAPTER II.B

- Figure II. B. 1: The structure of linear monoterpenes.
- Figure II. B. 2: Microbial biotransformation pathways for limonene.
- Figure II. B. 3: Possible conversion of β -ionone.
- Figure II. B. 4: The GC diagram of geraniol conversions.
- Figure II. B. 5: Mass spectra of geraniol obtained from the NIST arbitrary library.
- Figure II. B. 6: The GC diagram of nerol conversions.
- Figure II. B. 7: Mass spectra of nerol obtained from the NIST arbitrary library.
- Figure II. B. 8: The GC diagram of (*R*)-(+)-limonene conversions.
- Figure II. B. 9: Mass spectra of (*R*)-(+)-limonene obtained from the NIST arbitrary library.
- Figure II. B. 10: The GC diagram of α -ionone conversions.
- Figure II. B. 11: Mass spectra of α -ionone obtained from the NIST arbitrary library.
- Figure II. B. 12: The GC diagram of β -ionone conversions.
- Figure II. B. 13: Mass spectra of β -ionone obtained from the NIST arbitrary library.
- Figure II. B. 14: The obtained alignment and secondary structure of CYP109D1 (target) and the template (pdb entry 1z8p).
- Figure II. B. 15: Obtained docking positions in the homology model of CYP109D1.

CHAPTER III

- Figure III. 1: Type I binding spectra of CYP260A1 with progesterone.
- Figure III. 2: Type I binding spectra of CYP260A1 with DOC.
- Figure III. 3: The hydroxylated and dihydroxylated standards progesterone peaks.
- Figure III. 4: HPLC chromatogram of progesterone conversion by CYP260A1.
- Figure III. 5: HPLC chromatogram of CYP260A1 dependent progesterone conversion and comparison with the standard hydroxylated and dihydroxylated products.

- Figure III. 6: HPLC separation of aldosterone, corticosterone and DOC standards.
- Figure III. 7: HPLC chromatogram of the DOC conversion by CYP260A1.
- Figure III. 8: HPLC chromatogram of time dependent DOC conversion by CYP260A1.
- Figure III. 9: HPLC chromatogram of CYP260A1 dependent DOC conversion after 1 hr 30 min of incubation.
- Figure III. 10: HPLC chromatogram of CYP260A1 dependent DOC conversion after 2 hr of incubation
- Figure III. 11: The unrooted tree of So ce56 CYPome with CYP226 of *Burkholderia xenovorans* LB400, CYPDitQ of *Pseudomonas abietaniphila* BKME-9 and CYPtdtD of *Pseudomonas diterpeniphila*.
- Figure III. 12: Type I binding spectra of CYP260A1 with nootkatone.
- Figure III. 13: HPLC chromatogram of nootkatone conversion by CYP260A1.
- Figure III. 14: HPLC chromatogram of time dependent conversion of nootkatone by CYP260A1.
- Figure III. 15: MS data of the product of nootkatone conversion.
- Figure III. 16: Conversion of DOC to aldosterone.
- Figure III. 17: Steroid biosynthesis in myxobacteria.

LIST OF TABLE

GENERAL INTRODUCTION

Table 1: Chemical reactions catalyzed by cytochrome P450.

CHAPTER I

Table I. 1: *E. coli* strains and their genotypes.

Table I. 2: List of the oligonucleotides.

Table I. 3: PCR programmes used for the amplification of So ce56 CYPome from the genomic DNA of *Sorangium cellulosum* So ce56.

Table I. 4: Standard PCR reaction mix for genomic DNA amplification.

Table I. 5: PCR reaction mix for site directed mutagenesis.

Table I. 6: Parameters for sonication of cell pellet.

Table I. 7: Conserved domains of P450s in *Sorangium cellulosum* So ce56 (So ce56 CYPome).

Table I. 8: Cytochrome P450s in *Sorangium cellulosum* So ce56 with their closest homologues.

Table I. 9: The yield of expression and peak maxima of So ce56 cytochromes P450.

Table I. 10: Table illustrating the pattern of Soret maxima, α and β peak during oxidation, reduction and CO complex formation for all 21 *Sorangium cellulosum* So ce56.

Table I. 11: EPR spectral values of representative P450s of So ce56.

Table I. 12: Comparison of theoretical and experimental apo-protein mass and relative spectra of different cytochrome P450.

Table I. 13: Comparison of theoretical and experimental holo-protein mass and relative spectra of different cytochrome P450.

Table I. 14: Total numbers of Type I and Type II hits obtained from the screening of a compound library for CYP109C1, CYP109C2, CYP109D1, CYP260A1 and CYP266A1.

Table I. 15: The frequency of rare codons in *Sorangium cellulosum* So ce56 P450s.

CHAPTER II

CHAPTER II. A

- Table II. A. 1: The standard GC-MS parameters for fatty acid analysis.
- Table II. A. 2: The column oven program for fatty acid conversion.
- Table II. A. 3: The percentage similarity among the different members of CYP109 families
- Table II. A. 4: The characteristic derivatized units (sizes) for the silylated hydroxylation products from the different fatty acids that used in this work.
- Table II. A. 5: The product profile of CYP109 subfamilies (CYP109C1, CYP109C2 and CYP109D1) of *So ce56*.
- Table II. A. 6: Product profile of branched fatty acids with CYP109D1

CHAPTER III

- Table III. 1: HPLC gradient used for the analysis of progesterone conversion.
- Table III. 2: GC-MS analysis and identification of *in vitro* products of nootkatone
- Table III. 3: Structure analogues of selected Type I binding hits of screening.
- Table III. 4: Possible hydroxylation, epoxidation or oxidation of nootkatone by CYP260A1.
- Table III. 5: The summary of literature data on progesterone conversion in (a) bacteria and (b) fungus
- Table III. 6: The summary of literature data on DOC conversion in (a) bacteria and (b) fungus.

THE CYTOCHROME P450 COMPLEMENT OF THE MYXOBACTERIUM *Sorangium cellulosum* So ce56 AND CHARACTERIZATION OF TWO MEMBERS, CYP109D1 AND CYP260A1

GENERAL INTRODUCTION

The cytochromes P450 are available in all forms of life from prokaryotes to eukaryotes. The investigation of genome sequences has greatly facilitated the identification, classification and nomenclature of cytochrome P450 enzymes from different organisms. In plants, the P450s play a pivotal role for the production of secondary metabolites (Schuler and Werk, 2003). In yeast and fungi, a key P450 (CYP51) is essential for the production of sterols for membrane fluidity and integrity (Kalb *et al.*, 1987). In insects, P450s are required for hormone synthesis (Helvig *et al.*, 2004). Among 57 P450s in human, 15 P450s are involved in drug metabolism, 14 primarily involved in the steroid metabolism, 6 involved in vitamin metabolism and 9 P450s involved in the metabolism of fatty acids and eicosanoids, whereas the remaining 13 human P450s are considered as orphans and their biological functions remain to be firmly established (Guengerich *et al.*, 2005). In contrast, there is no cytochrome P450s in *E. coli*. Previously, it was presumed that bacteria would have less numbers of P450 but several genome sequence projects revealed the existence of higher numbers of P450s in some bacteria. Most recently, it was shown that the P450s are also available in virus and one of the P450 encoding genes, YP_143162, from *Acanthamoeba polyphaga* was heterologously expressed in *Escherichia coli* (Lamb *et al.*, 2009).

1. 1. CYTOCHROMES P450

Cytochrome P450 proteins comprise one of the largest super-families of enzymes (Werck-Reichhart and Feyereisen, 2000). They were first discovered in the 1950's–60's with simultaneous discovery of atmospheric oxygen incorporating into substrate molecules during metabolism catalyzed by metallo-proteins, which eventually lead to understanding the cytochromes having a part in oxidative metabolism (Estabrook, 2003).

1. 1. General aspects

The cytochrome P450s are a superfamily of heme *b*-containing monooxygenase enzymes, containing protoporphyrin IX as their prosthetic heme factor. P450 monooxygenases are 'mixed function oxygenases', constituting the terminal components of electron transport chains where both substrate and NAD(P)H are oxidized. Irrespective of the substrate, the

general function of a P450 is the same - the activation of molecular oxygen and subsequent insertion of a single atom of molecular oxygen into an organic molecule (Figure 1).

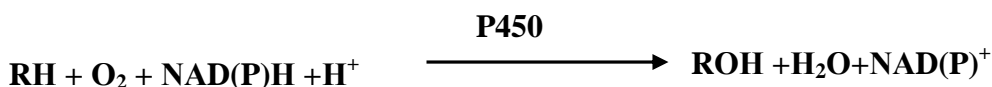


Figure 1: Cytochrome P450 catalyzed hydroxylation reaction. RH represents an oxidisable substrate and ROH the corresponding hydroxylated product.

The first experimental evidence relating to cytochromes P450 was provided in 1955 (Axelrod, 1955; Brodie *et al.*, 1955) when an enzyme system oxidizing xenobiotic compounds was identified in the endoplasmic reticulum of the liver. In 1958, a carbon monoxide binding pigment with an absorption maximum at 450 nm was detected in liver microsomes (Garfinkle, 1957; Klingenberg, 1958). The microsomal carbon monoxide-binding pigment was later demonstrated to contain iron-protoporphyrin IX (Omura and Sato, 1962; 1964), which was named cytochrome P450 after the primary characteristic feature in its absorption spectrum on complex formation with CO. The absorption peak of reduced P450 is still used for the estimation of the P450 content of a probe (Bernhardt, 2006). In cytochrome P450, *P* stands for pigment and *450* refers the maximum absorbance at 450 nm. *Cytochrome* stands for hemoprotein. The proteins of this family are commonly referred as CYPs, shorthand for cytochrome P450s.

1. 2. Heme Structure in P450

The heme structure found in P450s is a tetrapyrrole heme *b* (protoporphyrin IX) (Figure 2), which is a highly symmetrical and conjugated system, coordinated to an iron atom. The highly conjugated ring system leads to a strongly colored chromophore. The surrounding tetrapyrrole structure and axial ligands modulate the chemical properties of the heme iron. Heme electron transfer involves a single electron reversible oxidation/reduction of the iron between its 2⁺ and 3⁺ oxidation states, although the P450 catalytic cycle also involves other oxidation states of heme iron in the short-lived intermediates responsible for substrate oxidation (4⁺ and/or 5⁺). The properties of heme redox centers can be altered dramatically by changing the axial ligands. In P450s, four porphyrin central nitrogens provide the equatorial planar ligands for the heme iron, and the fifth ligand is usually a conserved cysteine sulphhydryl bond (the 'proximal' distal ligand, typically considered to be in the deprotonated thiolate form). A 6th (displaceable) water ligand is generally found opposite the cystein(at)e as

the distal ligand. This ligand makes way for binding of oxygen in the catalytic cycle. The heme macrocycle is typically bound non-covalently to the protein. However, the studies of the mammalian CYP4 family of P450s indicate that covalent linkage of a peripheral heme methyl group via a glutamate residue is feasible (LeBrun *et al.*, 2002a; 2002b).

Changes in electronic spectral properties occur during catalytic cycles of heme proteins and on interactions with various ligands. Spectral perturbations of the P450 heme result from the ability of various ligands to perturb the spin equilibrium or alter the coordination state of the proximal cysteinate-ligated heme iron. Substrates of P450s frequently shift the equilibrium in favor of the high-spin form, which occurs by displacement of a water molecule at the 6th ligand (distal axial ligand). Different spectral properties of the Soret band occur in the low-spin and high-spin forms, with maxima close to 420 nm and 390 nm, respectively (Munro *et al.*, 2000; Loew and Harris, 2000). The characteristic Type I binding spectra are represented by a decrease in an intensity of the Soret absorption peak at 420 nm coupled with a concomitant increase in the band intensity of 390 nm absorption (Gibson and Skett, 1994). The Type II binding spectrum is characterized by a decrease in absorption at around 390-405 nm with a concomitant increase at about 425-435 nm (Schenkman *et al.*, 1981). For several P450s (particularly certain well characterized bacterial isoforms), the change in spin-state facilitates the next step in their enzymatic cycle, one-electron reduction of heme iron from the redox partner.

Alterations in spin-state are an integral part of the characteristic and functional properties of the P450s. The signature electronic properties of heme sites are primarily due to π - π^* electron transitions within the protoporphyrin IX cofactor and the S \rightarrow Fe ligand-to-metal charge transfer bands. Also, normally associated with substrate-induced shift in spin-state are changes in the mid-point reduction potential of the ferric heme iron. Electronic re-distribution that occurs in the low- to high-spin transition in the iron *d*-orbitals and changes in overall energy levels of the orbitals leads to a more positive heme iron reduction potential in the high-spin state (Sharrock *et al.*, 1973).

The spectroscopic properties (i.e. UV-visible and other methods) of P450s provide a means to study and characterize a number of properties associated with these diverse biological catalysts. Accordingly, substrate dissociation constants (K_D) can be determined by the extent of the spin-state shifts induced during titration with a substrate and by using optical

spectroscopy (Daff *et al.*, 1997; Loew and Harris, 2000). Binding of inhibitory ligands (e.g. azoles) can also be analyzed by similar methods.

1. 1. 3. Structure of cytochrome P450

Cytochromes P450 share a common overall fold and topology despite less than 20% sequence identity across the gene super family (Hasemann, 1995). The typical crystal structure from the CYP102A is shown in Figure 3 (www.rcsb.org/pdb/results). The conserved P450 structural core is formed by a four-helix bundle composed of three parallel helices labeled D, L, and I and one antiparallel helix E (Presnell *et al.*, 1989). The prosthetic heme group is confined between the distal I-helix and proximal L-helix and bound to the adjacent Cys-heme-ligand loop containing the P450 signature amino acid sequences FxxGx(H/R)xCxG. The absolutely conserved cysteine in the proximal side is the ‘fifth’ ligand to the heme iron. This sulfur ligand is a thiolate (Dawson *et al.*, 1982) and is the origin of the characteristic name- giving 450 nm Soret absorbance observed for the ferrous CO-complex. Typically, the proximal ‘Cys’ forms two hydrogen bonds with neighboring backbone amides. A further interaction with a side chain is observed in some P450, for example, with a ‘Gln’ in CYP152A1 or a ‘Trp’ in nitric oxide synthase (Chen *et al.*, 2004, Ost *et al.*, 2001). The long I-helix forms a wall of the heme pocket and contains the signature amino acid sequences (A/G)Gx(E/D)T which is centered at a kink in the middle of the helix. The highly conserved ‘Thr’ preceded by an acidic residue is positioned in the active site and believed to be involved in catalysis (Imai *et al.*, 1989, Kimata *et al.*, 1995, Vidakovic *et al.*, 1998).

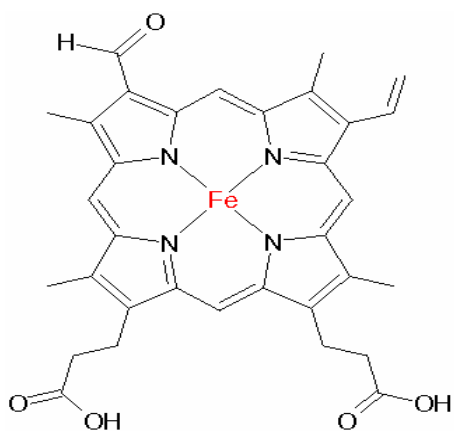


Figure 2: Schematic representation of the structure of heme *b* (ferriprotoporphyrin IX).

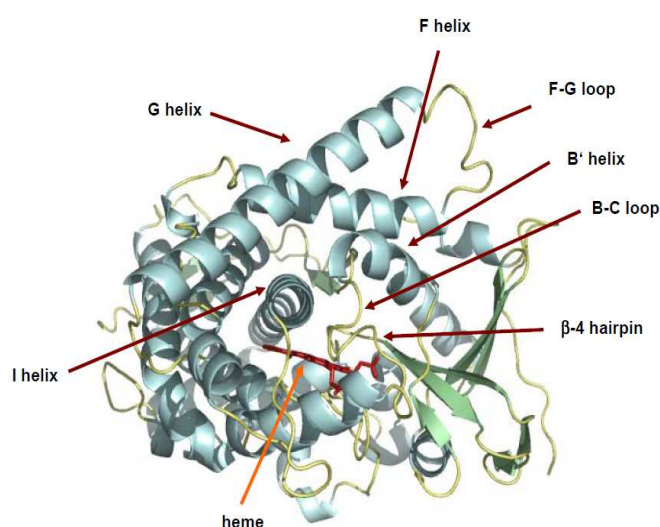


Figure 3: Crystal structure of CYP102A1 monooxygenase domain (www.rcsb.org/pdb/results).

1. 1. 4. Reactions and Mechanism

Based on crystal structures with and without substrate, ESR and NMR studies and other physicochemical methods a reaction mechanism for P450 monooxygenases was proposed. The actual reaction cycle shown in Figure 4 is nowadays commonly accepted (Denisov *et al.*, 2005).

(1) In the substrate free, inactive state the ferric heme iron Fe (III) is six times coordinated and in the low-spin state. The equatorial positions of the octahedral iron complex are coordinated with the nitrogens of protoporphyrin IX, while the cysteinate and a water molecule bind the axial positions.

(2) The water at the distal site of the heme is displaced upon substrate binding. The substrate is not taking the place of the water but binds mostly by hydrophobic interaction in the vicinity of the heme. This results in the transition to the high-spin Fe (III) complex, where the iron is only five times coordinated and in a pronounced “out of plane” structure. While some amino acid side chains show distinctive conformational changes, substrate binding induces only marginal rearrangement of the protein backbone (Haines *et al.*, 2001). More important than structural differences is the increase of the redox potential from around 300 to -170 mV. This is an important regulation mechanism, as after substrate binding the electron transfer guided by the potential gradient from NAD(P)H through the electron transfer chain to the heme iron is possible. The one electron transfer results in a high spin ferrous iron (Fe II) (3).

(4) Due to its four unpaired electrons and its out of plane structure this complex is predestined to bind triplet oxygen with very high affinity, resulting in “in plane” coordination. This is the last stable (at low temperature) intermediate of this cycle.

(5) The next step is again a one electron reduction to a very unstable postulated peroxo-iron (III) complex.

(6) Due to their high reactivity the following intermediates can not be spectroscopically characterized but it can be assumed that after incorporation of one proton a hydroperoxo-ferric intermediate is formed while through incorporation of the second proton the oxygen-oxygen bond is cleaved under release of water.

(7) The emerging reactive complex is often in literature termed “compound I” which most likely switches back to the initial state through a rebound mechanism (Figure I. 5).

(8) The reaction of this iron-oxo complex with the substrate is in general formulated as a radical reaction. However carbo-cations as intermediates or direct addition of the substrate (to) (7) under insertion of oxygen into a C-H bond are also discussed. During the rebound mechanism the carbon centered intermediate should oscillate through the binding plane under

loss of the sterical information. Due to the often high enantioselectivity of the P450 catalysed oxidation the recombination of this radical state has to be extremely fast. The resulting alcohol (Figure 5) is probably shortly bound to the oxygen of the Fe (III) - porphyrin system before its dissociation. Next to the rebound mechanism heterolysis or addition of R-H as reaction mechanism can not be completely excluded.

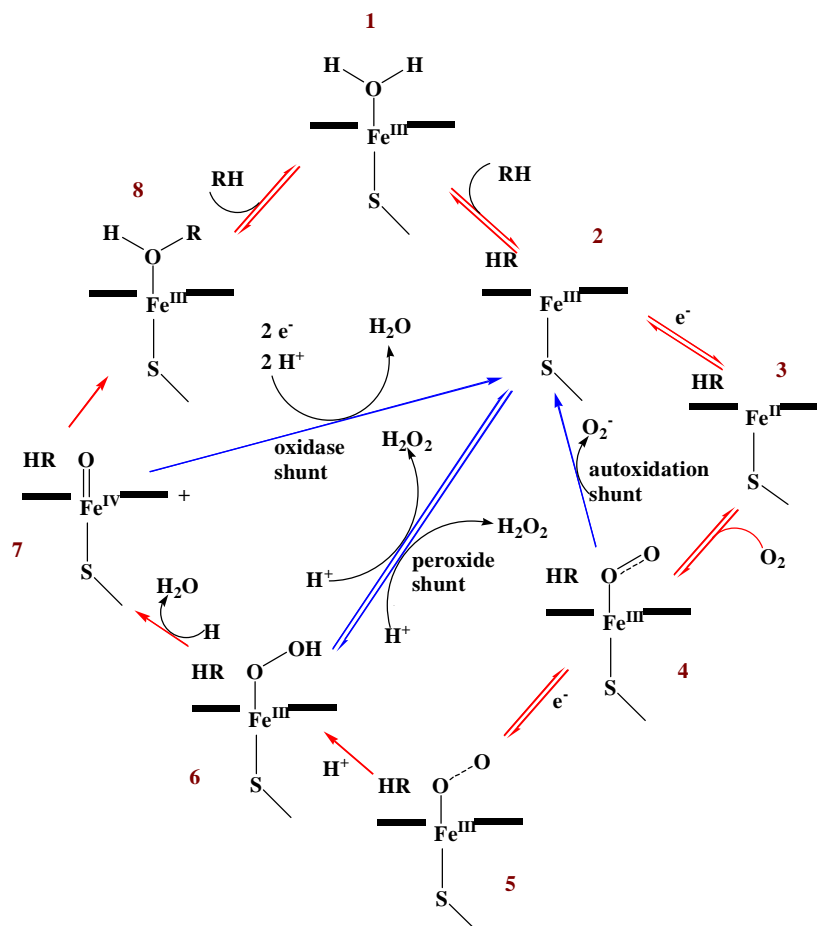


Figure 4: Catalytic cycle of P450 monooxygenases. The numbers shown in the figure are as described above.

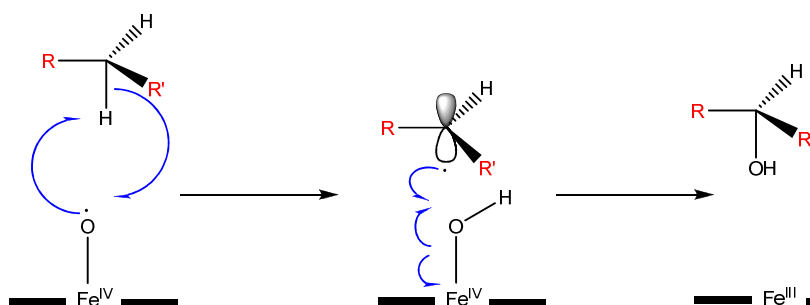


Figure 5: Rebound mechanism of the P450 catalyzed hydroxylation.

At least three more ways for dissociation of the activated oxygen from the heme iron or for substrate hydroxylation apart from the already discussed one exist. These reactions are called *shunt pathways*. The most thoroughly investigated reaction is the *peroxide shunt*. In the presence of powerful oxidants like periodates, organic peroxides and hydrogen peroxide the peroxo-complex (6) can directly be formed from the substrate bound high spin complex (2). This reaction is interesting for synthetic processes because it enables product formation without the use of NAD(P)H. In *auto-oxidation shunt* complex (2) is formed by dissociation of the superoxide from complex (4). On the other hand complex (2) can be formed through the *oxidase shunt* where oxygen is reduced under consumption of two molecules of NAD(P)H to water. These two reactions as well as the reverse *peroxide shunt* are often consolidated as the uncoupling reactions because they are unproductive concerning substrate oxidation but nevertheless consume NAD(P)H.

1. 1. 5. Electron transfer in cytochrome P450 systems

P450 monooxygenases (mixed function oxidases) incorporate one atom of molecular oxygen into the substrate and one atom is reduced to water. For this process electron equivalents are needed (Figure 4).

Regarding the electron transfer systems of P450, Chapman *et al.* defined four classes (<http://www.chem.edu.ac/champman/p450.html>). This classification is further elaborated by Hannemann *et al.* (2007), into 10 classes, the details of which are illustrated in Figure 6. The class I P450 systems [NAD(P)H→FdR→Fdx→P450] are present in most bacterial cytochrome P450 systems along with the mitochondrial P450 systems from eukaryotes. The class II P450 systems are abundantly present in the endoplasmic reticulum with two integral proteins-cytochrome P450 and the NADPH-cytochrome P450 reductase (CRP) containing the prosthetic groups FAD and FMN. The class III P450 systems also contains a three components P450 system [NAD(P)H→FdR→Fldx→P450] as in *Citrobacter braakii* where the natural immediate electron donor of the cytochrome P450 (P450cin) is flavodoxin (cinodoxin). The class IV P450 system [pyruvate, CoA→OFOR→Fdx→P450] is present in *Sulfolobus solfataricus*, the extreme acidothermophilic archaeon, where the potential redox partners of CYP119, a ferredoxin and its corresponding partners, 2-oxo-acid: ferredoxin oxidoreductase (OFOR) from *S. tokodaii* strain 7 and also from *S. solfataricus* itself, are found to be an efficient redox partner. The class V P450 system is a novel class of P450 found in *Methylococcus capsulatus* having two separate protein components - a putative NAD(P)H-

dependent reductase and a cytochrome P450-ferredoxin-fusion protein [NADH→FdR→(Fldx-P450)]. The class VI P450 system, present in *Rhodococcus rhodochrous* strain 11Y, has a putative NAD(P)H dependent flavoprotein reductase and a flavodoxin-P450 fusion protein [NAD(P)H→FdR→(Fldx-P450)]. In class VII P450 systems the cytochrome P450 is C-terminally fused to a reductase domain, which is usually not associate with P450 systems - a phthalate dioxygenase reductase domain [NADH→(PFOR-P450)] and is present in *Rhodococcus* sp. The class VIII P450 system is catalytically self sufficient as monooxygenases in which the P450 is fused to its eukaryotic-like diflavin reductase partner, cytochrome P450 reductase (CPR), in a single polypeptide chain. In class IX P450 systems the nitric oxide reductase is involved which was identified in *Fusarium oxysporum*. The class X P450 systems include all independent cytochromes P450 which do not need NAD(P)H and redox partners like allene oxide synthase, fatty acid hydro-peroxide lyase, divinyl ether synthase, prostacyclin synthase and thromboxane synthase.

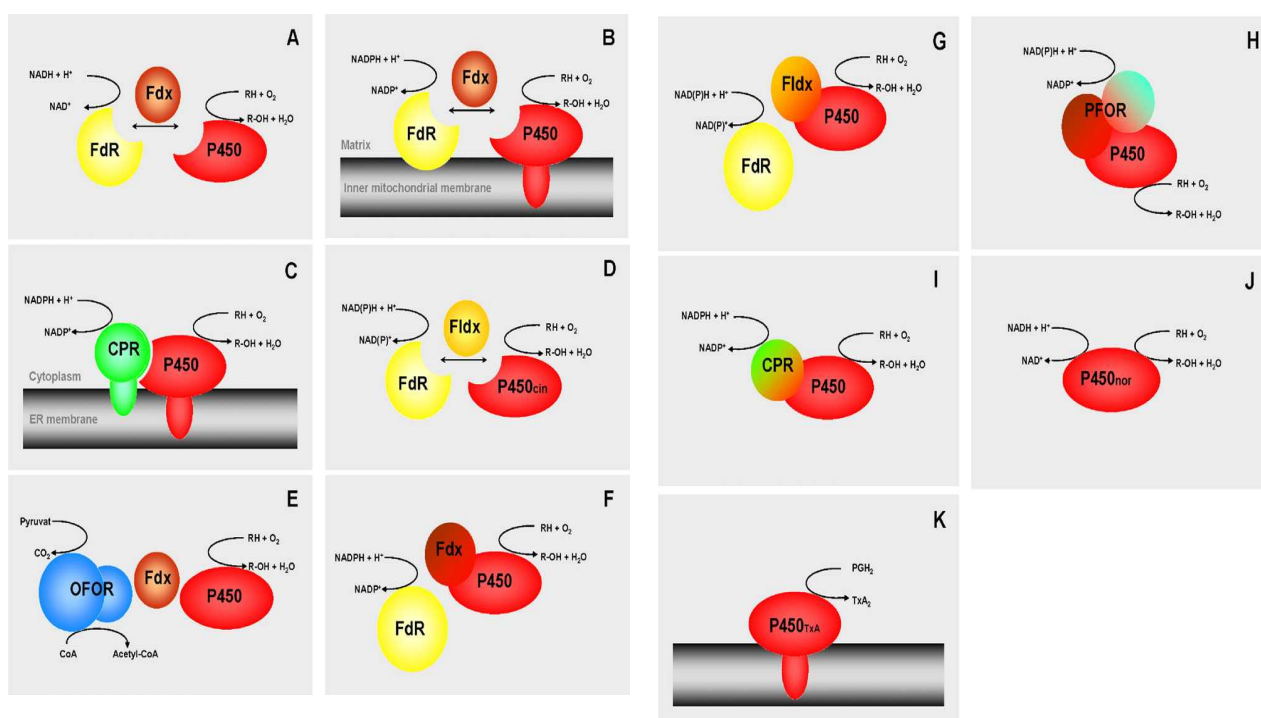


Figure 6: Schematic organization of different cytochrome P450 systems. (A) Class I, bacterial system; (B) class I, mitochondrial system; (C) class II microsomal system; (D) class III, bacterial system; example P450cin; (E) class IV, bacterial thermophilic system; (F) class V, bacterial [Fdx]-[P450] fusion system; (G) class VI, bacterial [Fldx]-[P450] fusion system; (H) class VII, bacterial [PFOR]-[P450] fusion system; (I) class VIII, bacterial [CPR]-[P450] fusion system; (J) class IX, soluble eukaryotic P450nor; (K) independent eukaryotic system, example P450TxA.

1. 1. 6. Substrate binding properties of cytochrome P450

The heme prosthetic group of the cytochrome P450 often serves as a useful chromophore during substrate and inhibitor binding studies. The oxidation of P450 bound substrate occurs at the heme, as a result certain aspects of substrate binding can be monitored readily with a spectrophotometer. Most commonly, ligand titration experiments are carried out to determine a spectral dissociation constant (K_D) for P450 substrates or inhibitors.

Three specific types of substrate binding spectra have been identified, namely Type I, Type II and reverse Type I, the latter being sometimes referred as modified Type II (Schenkman *et al.*, 1981). The characteristic Type I binding spectra (Figure 7) are represented by a decrease in the intensity of the Soret absorption peak at 420 nm coupled with a concomitant increase in the band intensity of 390 nm absorption (Gibson and Skett, 1994). Such spectrum is an indication of the substrate's influence on the P450 heme iron spin-state equilibrium, because low-spin P450 absorbs at around 418 nm (though the range varies from 416-420 nm) whereas the high-spin form gives rise to the 390 nm absorption (though the range varies from 385-394 nm). As a result, a Type I spectrum shows that the binding of substrate within the P450 heme locus brings about a shift in the Fe(III) spin equilibrium from low-spin to high-spin (Lewis, 2001). In most of the cases, changes in the magnitude of the Type I spectra indicate both the extent and the rate of metabolism, especially as there can be tight coupling between spin and redox equilibria in both the microsomal and bacterial P450 systems (Sligar and Gunsalus, 1976).

Compounds which tend to act as P450 inhibitors via ligation of the heme iron generally show Type II spectra (Figure 7). The structures of such molecules usually possess atoms containing non-bonded electrons, such as a nitrogen lone pair and there is also an accessibility in the structure of the ligand so that it may bind freely at the heme locus. It appears that the pK_a of the nitrogen is a factor in heme binding affinity of amines but the bond angle between the amine nitrogen and heme plane is also an important factor in those cases where steric hindrance can occur. The Type II binding spectrum is characterized by a decrease in absorption at around 390-405 nm with a concomitant increase at about 425-435 nm (Schenkman *et al.*, 1981). As a result, there is a shift in the absorption maxima corresponding to low- and high-spin Fe (III) towards longer wavelengths coupled with a shift in spin-state equilibrium from high- to low-spin iron.

The reverse Type I spectrum is a ‘mirror image’ of the Type I spectral change, with an increase in the 420 nm band and a decrease in absorbance at 390 nm. This is similar to a Type II change, except that there is no shift in the location of the maxima. The exact nature of this interaction is not known and it may be that the ligand binds to a different site on the enzyme.

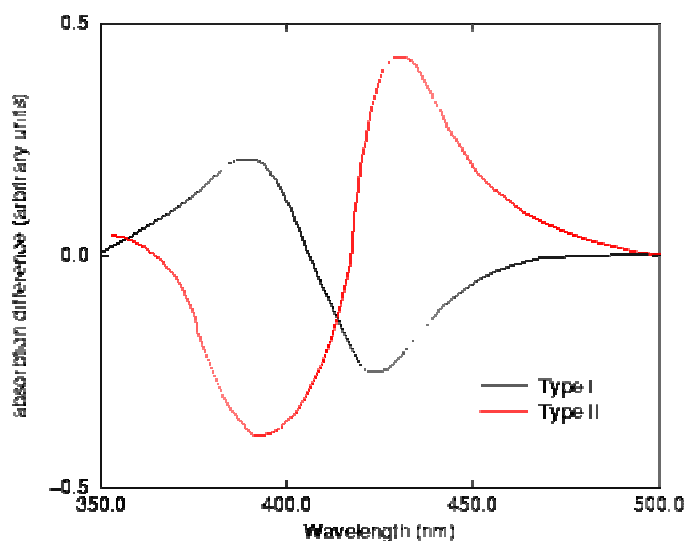


Figure 7: The UV-difference spectra obtained by the addition of a substrate or inhibitor to cytochrome P450. The black and red lines represent Type I and Type II difference spectra, respectively.

1. 1. 7. Reactions catalyzed by P450

Despite cytochrome P450 being considered mainly as oxygenases, during their long period of evolution, 3.5 billion years, the P450 family has become very versatile and has a broad field of chemical activity (Rendic and di Carlo, 1997). Among P450s there are isomerases, reductases, dehydrases, NO-synthases and in addition to oxidation, they contribute to various other reactions listed in Table 1 (Guengerich, 2001; Danielson, 2003; Bernhardt, 2006). However, exceptions can be found among the listed reactions.

Table 1: Chemical reactions catalyzed by cytochromes P450 (Guengerich, 2001; Danielson, 2003; Bernhardt, 2006).

Reactions	
NO-reduction	Alkene epoxidation
SO ₂ -reduction	Epoxide reduction
Hydroperoxide reduction	Isomerizations
1-electron oxidation	Oxidative C-C bond cleavage
Aldehyde deformylation	Alkyne oxygenation
Coupling reactions	Arene epoxidation
Ring formation, expansion	N-, S-, O-dealkylation
Hydrolysis	N-hydroxylation
1,2-shifts	N-oxidation
Mechanism-based heme inactivation	S-oxidation
Mechanism-based protein modification	Oxidative deamination
Reductive dehalogenation	Oxidative dehalogenation
Dehydratations	Dehydrogenation
Hydrocarbon hydroxylation	Steroid hydroxylation
Aromatic hydroxylation	

1. 2. Myxobacteria

The myxobacteria are a group of gram-negative, mainly soil-dwelling bacteria belonging to the delta subdivision of proteobacteria and having a high G + C content (~70%) in their genomic DNA (McCudry, 1989; Reichenbach, 1999). Myxobacteria are usually isolated from soil, decaying plant material, animal feces and tree bark, and produce extracellular hydrolytic enzymes that are able to degrade proteins, nucleic acids and some polysaccharides.

1. 2. 1. Taxonomy

In 1809, the first myxobacterium, *Polyangium vitellinum* (today *Kofteria flava*), was discovered by the German botanist H. F. Link (Reichenbach and Dworkin, 1992). Due to its fruiting bodies, the bacteria were termed as “gasteromycete” (fungi). In 1857, the British mycologist M. J. Berkeley discovered two additional species, *Stigmatella aurantiaca* and *Chondromyces crocatus*, which were classified as hyphomycetes (fungi imperfecti) (Berkeley, 1857). The genera name “myxobacteria” was formed by the American botanist Roland Thaxter in 1892 who was the first scientist to describe their life cycle (Thaxter, 1892).

Contemporary, phylogenetic analyses, by comparison of the 16S rRNA reveal that myxobacteria belong to the delta branch of the proteobacteria (Ludwig *et al.*, 1983). Based on morphological and physiological features and after several renaming, the order *Myxococcales* (myxobacteria) can be divided into three subgroups *Cystobacterineae*, *Sorangineae*, and *Nannocystineae* (Reichenbach, 2004; Shimkets *et al.*, 2005), five families, 17 genera and about 50 species (Figure 8).

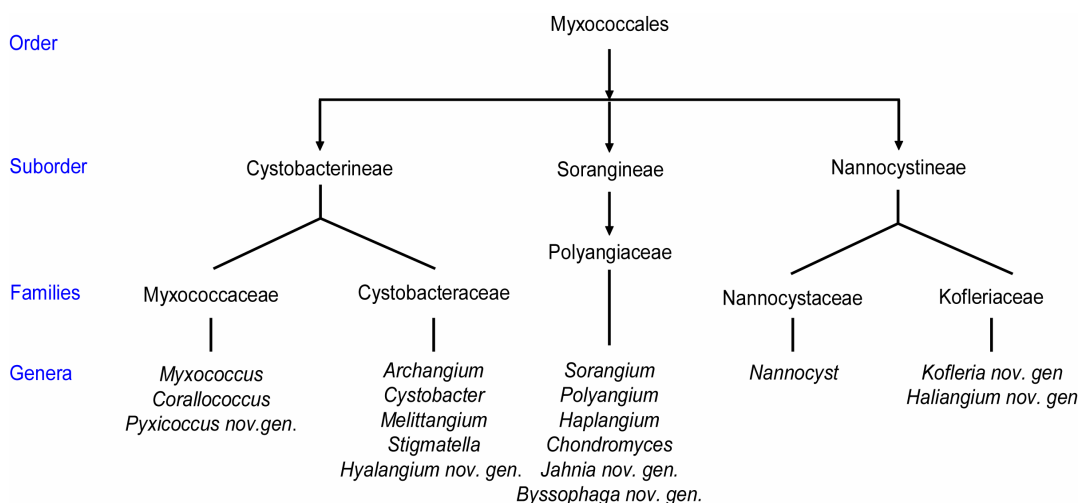


Figure 8: Taxonomy of myxobacteria (Reichenbach, 2004; Shimkets *et al.*, 2005).

Myxococcus xanthus, *Corallococcus* sp., *Archangium* sp. and *Stigmatella aurea* are typical members of the first subgroup and mainly obtain nutritional substrates by proteolytic or bacteriolytic activities. They feed upon other bacteria, utilizing the protein and lipid fraction as carbon and energy sources. The vegetative cells are slender and have tapering ends up to 1 µm in diameter and up to 20 µm in length. *Myxococcus xanthus* is the most extensively studied model organism among myxobacteria, and has a genome size of 9.14 Mb (Goldman *et al.*, 2006).

The *Sorangineae* like *Sorangium* (*Polyangium*) *cellulosum* and the *Byssophaga* genera are cellulose degraders and use inorganic nitrogen compounds while growing on cellulose and glucose (Reichenbach and Dworkin, 1992; Reichenbach, 2004). The vegetative cells of this subgroup are mainly cylindrical rods with rounded ends up to 1 µm wide and 10 µm long. In contrast, other members of the *Sorangineae* subgroup like *Haploangium* and *Chondromyces* and the subgroup *Nannocystineae* show in feeding experiments no cellulolytic activity.

1. 2. 2. Characteristic features of myxobacteria

The myxobacteria have two characteristics that distinguish them from other bacteria. At first, myxobacterial cells move by gliding motility in which cells move in the direction of their long axis on agar surfaces without the use of flagella. This gliding motility is utilized to accomplish swarming under vegetative conditions, during which cells of myxobacteria secrete hydrolytic enzymes that lyse other cells and convert their proteins into amino acids which are then metabolized. Secondly, myxobacteria can undergo complex cellular morphogenesis to form aggregates called fruiting bodies in which dormant myxospores are generated in response to starvation. The fruiting bodies can be very simple in shape and structure but also may occur as extremely complex designs.

1. 2. 2. 1. Gliding motility

One of the key elements in the development of *M. xanthus* is its gliding motility. The motility by gliding over a solid surface plays an important role in the developmental cycle of myxobacteria (Reichenbach and Dworkin 1992). In *M. xanthus*, wild-type gliding motility and cell swarming is the result of two separate systems. The S-, or social system, is responsible for the motility of groups of cells and is equivalent to twitching motility as seen in *Neisseria* and *Pseudomonas* (Connolly *et al.*, 1999; Wall and Kaiser, 1999; Merz *et al.*, 2000; Beatson *et al.*, 2002). Twitching motility is mediated by the extension and retraction of type IV pili. The A-, or adventurous system, is responsible for the motility of single cells and is related to gliding motility. Secretion-mediated gliding displayed by some cyanobacteria is due to the expulsion of a polyelectrolyte gel through membrane pores and is proposed to provide the thrust for A- type gliding (McBride, 2001; Thomasson *et al.*, 2002; Wolgemuth *et al.*, 2002). Many genes in *M. xanthus* have been shown to effect A- and S- type gliding separately, and a few have been shown to effect both systems.

1. 2. 2. 2. Fruiting body formation

The life cycle of myxobacteria consists of a vegetative growth cycle and a developmental cycle (Figure 9). The developmental cycle of myxobacteria, including fruiting body formation and sporulation, is triggered by the nutritional and physical changes of the environment. During fruiting body formation, myxobacteria sense the depletion of nutrient, move rhythmically (rippling), aggregate together on a solid surface, construct multicellular fruiting bodies (ca. 10^5 – 10^7 cells/fruiting body) and convert the vegetative cells into stress resistant myxospores inside the fruiting body (Dworkin, 1985). These fruiting bodies are resistant to

several stress conditions such as desiccation, sonication and UV radiation. Additionally, sporulation can be induced independently from starvation by different chemical agents such as glycerol, DMSO, indole and its derivatives (Dworkin, 1994; Gerth *et al.*, 1994; O'Connor and Zusman, 1997). Development of myxobacteria fruiting body requires intercellular communication. So far, at least five extracellular signals (*Asg*, *Bsg*, *Csg*, *Dsg*, and *Esg*) and a general starvation signal, (p)ppGpp, have been detected in *M. xanthus* (Downard *et al.*, 1993; Dworkin, 1996; Hagen *et al.*, 1978; LaRossa *et al.*, 1983). The morphology of fruiting bodies varies between different myxobacterial species. Whereas the *Stigmatella* and *Chondromyces* sp. form sophisticated multiple tree-like sporangioles, *Sorangium* strains produce simple knobs consisting of slime and myxospores Figure 10.

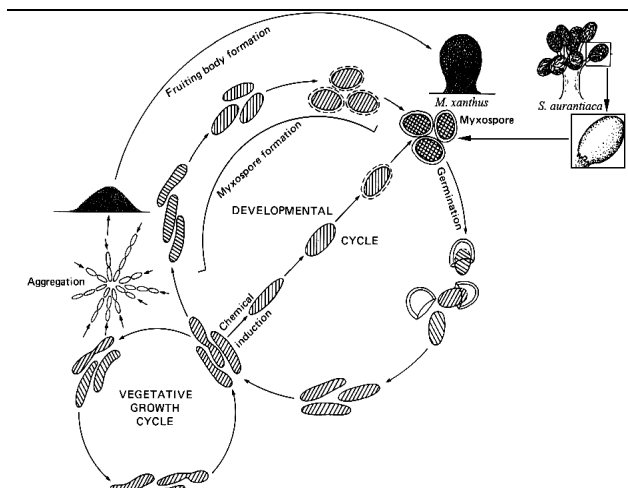


Figure 9: Diagram of the life cycle of myxobacteria (Dworkin, 1995). The fruiting body of *M. xanthus*, in which myxospores are embedded in the slime mound, and the fruiting body of *S. aurantiaca*, which consists of a stalk and sporangioles, are illustrated.

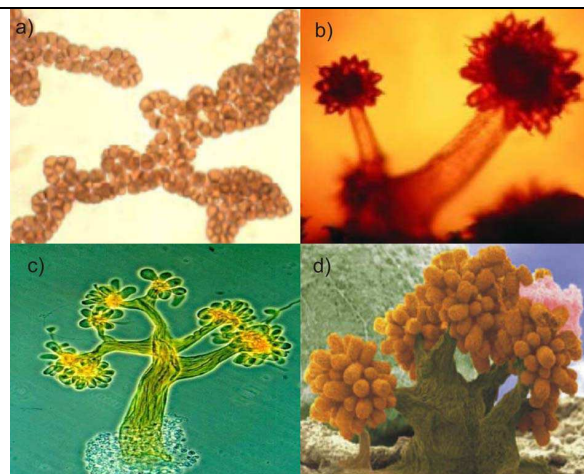


Figure 10: Fruiting bodies of different myxobacterial species. a). *Sorangium cellulosum* So ce56 (Gerth *et al.*, 2003); b) *Chondromyces apiculatus* (Reichenbach and Dworkin, 1992); c and d) *Chondromyces crocatus* (Reichenbach and Dworkin, 1992).

1. 2. 3. Application of myxobacterial secondary metabolites

The secondary metabolites that are produced by myxobacteria display a wide diversity of structures and biological activities. Myxobacteria are known to be prolific producers of interesting and novel bioactive substances applied in biotechnology and pharmacology (Reichenbach and Höfle 1993, 1999, Gerth *et al.*, 2003; Bode and Müller, 2006). About 7500 different myxobacteria have been isolated and many of them were screened for secondary metabolites. The literature shows that nearly 500 derivatives from 100 core structures were found with several novel core structures (Gerth *et al.*, 2003). Approximately, 50% of secondary metabolites are synthesized by different strains of *Sorangium cellulosum* (Figure

11). Different strains of *Sorangium cellulosum* produce several novel antimicrobial macrolides- thuggacins (Irschik *et al.*, 2007), phoxalone (Guo *et al.*, 2008), a new sesquiterpene- sorangiadenosine (Ahn *et al.*, 2008) and a new free-radical scavenger-soraphinol C (Li *et al.*, 2008). Along with this, *Sorangium cellulosum* also synthesizes a novel class of antineoplastic agents- the epothilones and their analogs (Mulzer, 2008).

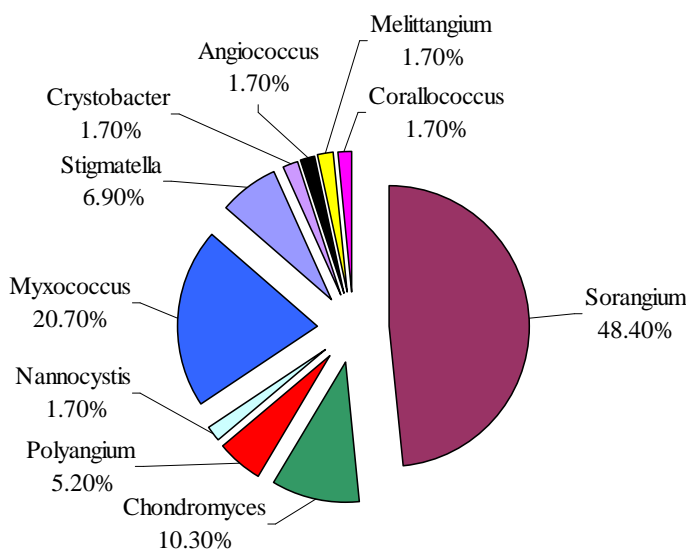


Figure 11: Percentage distribution of the production of secondary metabolites from different myxobacterial strains (replotted from Gerth *et al.*, 2003).

1. 2. 4. *Sorangium cellulosum* So ce56

The ~13.1 Mb genome of the model strain *Sorangium cellulosum* So ce56 is the largest yet discovered in bacteria (Schneiker, 2007). Varieties of biosynthetic gene clusters involved in natural product formation as well as numerous complex and unusual biosynthetic processes within *Sorangium cellulosum* have been identified (Wenzel and Müller, 2007). The selected strain So ce56 produces the natural secondary metabolites chivosazol, etnangien and myxochelin (Schneiker *et al.*, 2007; Wenzel and Müller, 2007; Menche *et al.*, 2008). The genomic sequence also revealed several additional regions of the chromosome encoding polyketide synthase (PKS) and/or nonribosomal polyketide synthase (NRPS) proteins, whose products have not been isolated till date (Schneiker *et al.*, 2007). *Sorangium* strains provide no or less information concerning major biological processes like the gliding system, fruiting body formation and cell-to-cell interaction.

CHAPTER I

The cytochrome P450 complement (CYPome) of *Sorangium cellulosum* So ce56

1. Introduction

Cytochrome P450 monooxygenase systems are found throughout the biological kingdoms from eubacteria, archaeobacteria to mammals. This gene superfamily plays a critical role in the bioactivation and detoxification of a wide variety of xenobiotics, in the biosynthesis of endogenous compounds and in the production of secondary metabolites (Gillam *et al.*, 2007; Bernhardt, 1995; 2006). Cytochromes P450 also catalyze late-stage stereo- and regiospecific oxidations when associated with polyketide biosynthetic gene clusters in bacteria (Graziani *et al.*, 1998; Zhao *et al.*, 2008; Buntin *et al.*, 2008).

It has been observed that cytochrome P450 sequence information for various organisms is increasing during the last several years, especially as a result of various genome projects. Unfortunately, the functional study of the corresponding proteins is lagging far behind, even for the most heavily studied human forms, of which the complete substrate range is still unclear. Till date 110 P450 families are listed in animals, 95 in plants, 205 in bacteria, 10 in Archaea, 61 in protists and 310 families in fungi (<http://drnelson.utmem.edu/p450stats.Feb2008.html>).

The investigation of P450s in *Sorangium cellulosum* So ce56 was started in parallel with the genome sequencing project. The recently completed genome sequence project of myxobacterium, *Sorangium cellulosum* So ce56 revealed 21 open reading frames (ORFs) of cytochromes P450 (Schneiker *et al.*, 2007). As this was the very first attempt to study the cytochrome P450s in So ce56, at the beginning of this work only the sequence of the P450 ORFs were in hand. The elucidation of the putative function of P450s can only be hypothesized after detailed theoretical study of all possible aspects.

2. Objectives

1. To perform bioinformatics analysis of the P450s of *So ce56*.
2. To clone, express and purify all 21 the P450s.
3. To study the biophysical and biochemical characterization of selected P450s.
4. To identify the heterologous redox partners for the selected P450s.
5. To screen a compound library consisting of ~17,000 substances as potential substrates for highly expressed P450s.

3. MATERIALS AND METHODS

3. 1. Materials

The instruments, chemicals, enzymes, kits, reagents and other accessories used in this work were of high grade. They are illustrated in Appendix I.

3. 1. 1. Bacterial strains

In this work several strains of *E. coli* were used for cloning and expression purposes. They are illustrated in Table I. 1.

Table I. 1: *E. coli* strains and their genotypes.

Strain	Genotype	Company
BL21 (DE3)	<i>E. coli</i> B F <i>dcm ompT hsd S</i> ($r_B^- m_B^-$)	Novagen
BL21	<i>E. coli</i> B F <i>dcm ompT hsd S</i> ($r_B^- m_B^-$) <i>gal</i>	Novagen
TOP10F'	<i>F'</i> { <i>lacIq</i> , <i>Tn10(TetR)</i> } <i>mcrA</i> _(<i>mrr</i> - <i>hsdRMS-mcrBC</i>) ϕ 80 <i>lacZ_M15_lacX74</i> <i>recA1 araD139</i> _(<i>ara-leu</i>)7697 <i>galU galK</i> <i>rpsL (StrR) endA1 nupG</i>	Invitrogen
JM109	<i>endA1</i> , <i>gyrA1</i> , <i>hsdR17</i> (<i>rk-mk+</i>), <i>mrcB+</i> , <i>recA1</i> , <i>relA1</i> , <i>supE44</i> , <i>thi-1</i> , _(<i>lac-proAB</i>), <i>F'</i> [<i>traD36, proAB, lacIqZ_M15</i>	Yanisch-Perron <i>et al.</i> (1985)

3. 1. 2. Proteins

Purification of bovine AdR and bovine Adx (4-108) (Uhlmann *et al.*, 1994) was carried out by Wolfgang Reinle. AdR was purified according to a procedure published by Sagara *et al.* (Sagara *et al.*, 1993). Adx (4-108) were purified following a protocol developed in our group (Uhlmann *et al.*, 1994). Spinach ferredoxin and ferredoxin reductase were purchased from Sigma Aldrich Company.

3. 1. 3. Oligonucleotides

The designed oligonucleotides to amplify the P450 ORFs from the genomic DNA of *So ce56* were purchased from MWG Biotech. The list of the primers is illustrated in Table I. 2.

Table I. 2: List of the oligonucleotides.

P450s of So ce56	Oigonucleotide sequences
CYP109C1	ccaatcat atgaaccttttcccggacgagat -3' C7_NdeI_For cccttctagactagtgatggtgatggtgatgccccgggcacggcgt C7_His.XbaI_Rev
CYP109C2	ccaatcatatgaacctgt tctcggaaga gatgc C6_NdeI_For ccaatctagatcagtgatggtgatggtgatgccccggggcaccgcc C6_His.XbaI_Rev
CYP109D1	ccaatcatatggaaaccgagaccgccccgagccc C8_NdeI_For ccaatctagatcagtgatggtgatggtgatggcggtggcgcggct C8_His.XbaI
CYP110H1	ccaatcatatgagccagttgccccgaggcccgc C12_NdeI_For cccttctagatcaatggtgatggtgatggtggcgcgccgtaaaag C12_His.XbaI_Rev
CYP110J1	cccatcatatgggaaatc tgccgccgg gc -5' C11_NdeI_For ccaatctagatcagtgatggtgatggtgatgggagaagccatggg C11_His.XbaI_Rev
CYP117B1	gggatcatatgaccaccg ccctccccgc t C9_NdeI_For gccttctagatcagtgatggtgatggtgatggccgctgaatcggac C9_His.XbaI_Rev
CYP124E1	ccaacatgatgaccctc aaaccgtcgatatcc C3_NdeI_For agaattctagt atg gtg atg gtg atg cctggcggggcgcgcgcaccgg C3_His.EcoRI_Rev
CYP259A1	ccaatcatatgctctccc tgggtcgaac cgacca C10_NdeI_For ccgatctagatcagtgatggtgatggtgatggcaccggcgaaaggc C10_His.XbaI_Rev
CYP260A1	ccaacatatggcagacgagaaagcggcggagcgcg C17_NdeI_For ccaaaagctttcagtgatggtgatggtgatggtgcaggagcacaggcag C17_His.HindIII_Rev
CYP260B1	ccaacatatgctgcctcgtaaaaatctattttt C16_NdeI_For ccaaaagcttttagtgatggtgatggtgatgcacgagcagcggcggaa C16_His.HindIII_Rev
CYP261A1	ccaatcatatggagaccgccagagcagaagtgc C15_NdeI_For ccaaaagctttcagtgatggtgatggtgatggcggtggcgggggggaa C15_His.HindIII_Rev
CYP261B1	ccaatcatatg aaggggg gcatgaacgc cacgcccc C21_NdeI_For ccaaaagctttcagtgatggtgatggtgatgcgcgctcgcgcgcttct C21_His.HindIII_Rev
CYP262A1	ccaatcatatggtggggctcctcccggcgtggcg C14_NdeI_For ccaaaagctttcaatggtgatggtgatggtgcacggcatggagcgtcat C14_His.HindIII_Rev
CYP262B1	ccaatcatatg cagccag atcttctctc cggcacac C20_NdeI_For ccaaaagctttcagtgatggtgatggtgatgccttcccggcgcggctc C20_His.HindIII_Rev
CYP263A1	ccaatcatatgagtccacgctcgaacggcagcc C13_NdeI_For ccaataagctttcagtgatggtgatggtgatggccccggcgctgtacg C13_His.HindIII_Rev
CYP264A1	ccaatcatatgagcgagcgcgtggacatcatgac C5_NdeI_For ccaaaagctttcagtgatggtgatggtgatggcccggcctgaagcggag C5_His.HindIII_Rev

CYP264B1	ccaatcatatgactcgacttaatctgttcgcc C4_NdeI_For ccaataagctttcagtgatggtgatggtgatgccagatgccctggacga C4_His.HindIII_Rev
Mutagenic primers	catcgctgatgaccacatggtcttgatcctcg C4_Mut_For cgaggatcaagaccatgtgggtcatcagcgatg C4_Mut_Rev
CYP265A1	ccaacatatgctgactcctcacacgctgcagccgg C18_NdeI_For ccaaaagctttcagtgatggtgatggtgatggcgccggcgccacgcggca C18_His.HindIII_Rev
CYP266A1	ccttcatatggagacgcaagtggccgcgacgc C2_NdeI_For ccaagaatttcagtgatggtgatggtgatggaatgatacagggagcgtctga C2_His.EcoRI_Rev
CYP267A1	ccttcatatgaactgccccgacgcaccgaagcc C1_NdeI_For ccaagaatttcagtgatggtgatggtgatgggagggcgagctctgcgggccta C1_His.EcoRI_Rev
CYP267B1	ccaatcatatggtcgacc aggacgcttt tccggagc C19_NdeI_For ccaaaagctttcaatggtgatggtgatggtggaacaccacaggcaccg C19_His.HindIII_Rev
CYP266B1P	ccaacatatggccaagctcccgcacactgcacagcc C22_NdeI_For ccaaaagctttcagtgatggtgatggtgatggcgccagcggcgcccg C22_His.HindIII_Rev

3. 1. 4. Vectors

In this work the cloning vector PCR-Topo Blunt end (Invitrogen), and the cloning vector pCWori+ and pET17b (Invitrogen) were used. The expression vector pCWori+ was provided by Michael R. Waterman, Vanderbilt University School of Medicine, Nashville, USA. The details of these vectors are illustrated in Appendix V.

3. 2. Methods

3. 2. 1 Bioinformatics study

3. 2. 1. 1. Bioinformatics analysis of P450s of *So ce56*

The bioinformatics study of the P450s of *So ce56* was done in collaboration with Dr. Olena Perlova and Prof. Dr. Rolf Müller, Pharmaceutical Biotechnology, Saarland University, during genome sequencing project of *Sorangium cellulosum So ce56*.

Cytochrome P450s encoding open reading frames (ORFs) of *Sorangium cellulosum So ce56* were initially analysed in the genomic database based on the heme-binding domain signature, FxxGx(H/R)xCxG as query sequence. The ORFs possessing such motifs were further analyzed for the presence of the highly conserved threonine in the putative I-helix and the conserved EXXR motif in the K-helix (Nelson, 1996; Denisov *et al.*, 2005). Amino acid sequences were deduced from the genes performed by ExPASy Translate tool

(<http://www.expasy.org/tools/dna.html>), and P450 family and subfamily assignments were made by Dr. David Nelson according to the P450 nomenclature (<http://drnelson.utm.edu/CytochromeP450.html>). Briefly, P450s showing more than 40% identity was placed in the same family and P450ss exhibiting more than 55% identity were categorized in the same subfamily. New families were assigned for the P450s showing less than 40% identity to proteins from other organisms.

3. 2. 1. 2. Phylogenetic analysis

Translated ORFs of respective P450 genes of the *Sorangium cellulosum* So ce56 were aligned using Clustal W2 (Thompson, 1994) and analyzed by the Neighbour Joining Algorithm with boot strap analysis of 1000 times replicating to resample the data and displayed through Tree View (Page, 1996).

3. 2. 1. 3. Other bioinformatics studies related to P450s and neighboring ORF

The percentage similarity and identity of the P450s of So ce56 were studied by pairwise sequence alignment using EMBOSS Pairwise Alignment Algorithms. The closeness of So ce56 P450s with other bacterial strains was identified by using David Nelson's P450s blast server (<http://132.192.64.52/blast/P450.html>). Comparison and analysis of the relatedness of each P450 gene of So ce56 and its neighboring genes were also studied by global NCBI protein BLAST and pfam (<http://pfam.sanger.ac.uk/>). The acquisition of several P450s genes/proteins was accessed from EMBL/NCBI gene data bank or Protein Data Bank (<http://www.rcsb.org/pdb/home>).

3. 2. 2. Molecular biology methods

All standard DNA techniques were performed according to the standard methods as described previously (Sambrook and Russell, 2001), unless indicated.

3. 2. 2. 1. Isolation of genomic DNA

Genomic DNA from *S. cellulosum* So ce56 was isolated using the Puregene[®] Genomic DNA Purification Kit (Gentra, Minneapolis, USA) as per the manufacturer's instructions. The genomic DNA was provided from Prof. Dr. Rolf Müller, Institute for Pharmaceutical Biotechnology, Saarland University. All the P450s containing plasmid DNAs were purified using either the NucleoSpin Plasmid QuickPure Kit or the Nucleobond Plasmid Purification AX 100 Kit (Macherey-Nagel, Düren, Germany).

3. 2. 2. 2. Plasmid isolation, purification and determination of the nucleic acid concentration

Plasmid purification was performed using a commercially available kit from Macherey-Nagel (Nucleobond[®] maxi or midi plasmid preparation kits) according to the manufacturer's instructions. After purification, the plasmid DNA concentration was determined spectroscopically by measuring the absorption at 260 nm. The principle of this method is based on the absorption ability of UV light by the ring structure of purines and pyrimidines in the DNA or RNA where 1 AU₂₆₀ corresponds to a dsDNA concentration of 50 µg/ml. (Hagemann, 1990). The concentration of nucleic acids was calculated by following the formula: $C \text{ mg/ml} = \text{OD}_{260} * \epsilon * f$ where C is the concentration in mg/ml, f = dilution factor and $\epsilon = 50$ (double strand DNA) or 40 (single strand DNA and RNA).

3. 2. 2. 3. Digestion of DNA with endonucleases

Restriction endonucleases enzymes were purchased from Roche Diagnostics (Basel, Switzerland), NEB (New England Biolabs, Beverly, MA, USA), and from Promega (Madison, WI, USA). Digestion of the DNA was carried out according to the manufacturer's instruction. Usually, 0.1-10 µg DNA were digested by 1-10 U of enzymes. Enzyme specific buffers provided with the enzymes were used to achieve the best results. Restriction digestions were performed in 20-50 µl volume. The incubation time and temperature were depended on the types of enzymes used, typically 1-4 hours to overnight at 37°C. The digested DNA was analyzed by agarose gel electrophoresis.

3. 2. 2. 4. Ligation of DNA

For the ligation purpose, new bonds between phosphate residues located at the 5' termini of double-stranded DNA and adjacent 3'-hydroxyl moieties have to be formed in the digested and purified DNA stretches. DNA ligase enzyme can catalyzes these phosphodiester bonds between adjacent 5'-phosphate and 3'-hydroxyl residues.

In this work, Fast-Link[™] DNA Ligation Kit (Invitrogen) was used and the manufacturer's instruction was followed with some modification. The ligation was done using the thermal cycler (16°C for 15 min, 20°C for 14 min and 70°C for 15 min) at different molar ratios of linker to insert (Sambrook and Russell, 2001).

3. 2. 2. 5. Agarose gel electrophoresis and extraction of DNA from agarose gel

Agarose gel electrophoresis is an effective method to separate DNA fragments for analytical and preparative purposes by means of size (base pair) differentiation. The negative charges of DNA molecules help to migrate towards the anode in an electrical field. The migration depends on the size (base pair), the conformation of the DNA (either relaxed or super coiled) and also on the pore size of the gel. To visualize the DNA, ethidium bromide, a fluorescent dye, is often used. It intercalates between stacked base pairs and upon ultraviolet illumination makes DNA visible.

Typically, for the DNA fragments between 1000 and 8000 bp length 1% melted agarose solution in 1 x TAE/TARE (Appendix II) was poured into the mould after addition of 0.01% ethidium bromide. The setted gel was transferred into the electrophoresis tank and was covered with 1 x TAE/TARE buffer. DNA samples (5-50 µl) with DNA loading buffer (5:1) were loaded on the gel. The Smart Ladder (Eurogentec Liège, Belgium) was used as reference DNA. Electrophoresis was performed at 80 V for 45 min. The gel was examined by ultraviolet light and documented. The extraction of DNA from agarose gel was done by using NucleoSpin (Macherey-Nagel) according to manufacturer's instruction.

3. 2. 2. 6. Polymerase chain reaction and molecular cloning

Polymerase chain reaction (PCR) is the most common technique use in molecular cloning and analysis of DNA. It is used to amplify large copy numbers of a specific region of DNA using DNA polymerase which catalyzes the polymerization of deoxyribonucleotides alongside a DNA strand. PCR requires several basic components like DNA template, two oligonucleotide primers, DNA polymerase, dNTPs and buffer solution. The PCR usually consists of a series of 20 to 40 cycles involving the denaturation of template, the annealing of primer, and the extension of the annealed primers by DNA polymerase.

In this work, the PCR reactions were carried out in the thermal cycler PT-100, MJ Research Inc. Each of the respective P450s from the high GC (71.38%) rich genomic DNA of *Sorangium cellulosum* So ce56 was amplified by using forward and reverse primers as described in Table I. 3. Direct or indirect cloning was done to design expression constructs in pCWori⁺. The insert possesses a *Nde* I site encoding the initiator codon ATG defining the 5' end of the open reading frame and a hexa-histidine tag followed by a new stop codon and either restriction site for *Eco*RI, *Hind*III or *Xba*I. In two P450s, CYP266A1 and CYP262A1,

the originl initiating codon GTG in the genomic open reading frame (ORF) was replaced with ATG within the *NdeI* site.

In case of CYP109C2, CYP109C1, CYP117B1, CYP259A1 and CYP110J1, the P450 fragments were subcloned in pCR4 Blunt-Topo vector according to manufacturer's instructions. All those cloned P450s have *NdeI* and *XbaI* site. But the *XbaI* site of those P450s inserts was methylated during transformation in normal *E. coli* and the restriction site was blocked. It was deactivated by transforming the plasmids into dam-negative *E. coli* GM1609 and was subsequently digested and inserted into the *NdeI* and *XbaI* sites of expression vector (pCWori⁺). The typical reaction cocktail is illustrated in Table I. 4.

Table I. 3: PCR programmes used for the amplification of the P450s from the genomic DNA of *So ce56*.

P450s of Soce56	GC %/ Tm value (°C)	Thermal cycle*	P450s of Soce56	GC %/ Tm value (°C)	Thermal cycle*
CYP109C1	45.2/66.8 60.9/>75	2	CYP261B1	61.1/>75 57.1/>75	4
CYP109C2	45.5/68.2 60.9/>75	2	CYP262A1	68.6/>75 51.0/>75	3
CYP109D1	61.8/>75 58.7/>75	3	CYP262B1	52.8/72.9 59.2/>75	3/2
CYP110H1	63.6/>75 56.5/>75	4	CYP263A1	58.8/74.3 57.1/>75	3a
CYP110J1	63.3/73.6 52.2/>72	3	CYP264A1	52.9/71.9 57.1/>75	3
CYP117B1	65.5/73.7 54.3/>75	1	CYP264B1	45.5/68.2 51/>75	5
CYP124E1	50/69.5 63.5/>75	3	Site directed mutagenesis		6
CYP259A1	55.9/73.1 56.5/>75	1	CYP265A1	60/>75 61.2/>75	3
CYP260A1	62.9/>75 53.1/>75	3/1	CYP266A1	62.5/74.6 49.1/>75	2
CYP260B1	34.3/64.8 52.1/>75	3a/2	CYP267A1	59,4/73.3 57.4/>75	1
CYP261A1	57.1/74.2 59.2/>75	3a	CYP267B1	55.6/74 53.1/>75	3/2
			CYP266B1P	64.9/>75 63.3/>75	3/2

***Thermal cycle:**

- 1:** Initial denaturation 98°C for 2 min, denaturation 98°C for 10 s, annealing 50°C for 35 s, elongation 72°C for 2 min for 35 cycle, final elongation 72°C for 10 min with phusion polymerase, DMSO and GC buffer.
- 2:** Initial denaturation 98°C for 2 min, denaturation 98°C for 15 s, annealing 55°C for 30 s, elongation 72°C for 1:30 min for 35 cycle, final elongation 72°C for 10 min with phusion polymerase, DMSO and GC buffer.
- 3:** Two step PCR: Initial denaturation 98°C for 5 min, denaturation 98°C for 15 s, annealing and elongation 72°C for 40 s for 30 cycles, final elongation 72°C for 10 m with phusion polymerase, DMSO and GC buffer.
- 3a:** Two step PCR: Initial denaturation 98°C for 5 min, denaturation 98°C for 15 s, annealing and elongation 72°C for 40 s for 30 cycles, final elongation 72°C for 10 min with phusion polymerase and HF buffer.
- 4:** Modified Two step PCR: Initial denaturation 98°C for 2 min, denaturation 98°C for 10 s, annealing and elongation 72°C for 55 s for 30 cycles, final elongation 72°C for 10 min with phusion polymerase, DMSO and GC buffer.
- 5:** Initial denaturation 98°C for 2 min, denaturation 98°C for 15 s, annealing 55°C for 30 s, elongation 72°C for 1:30 min for 35 cycles, final elongation 72°C for 10 min with phusion polymerase and HF buffer.
- 6:** Site directed mutagenesis: Initial denaturation 95°C for 5 min, denaturation 95°C for 30 s, annealing 55°C for 1 min, elongation 70°C for 9 min for 16 cycles, final elongation 72°C for 10 min with *Pfu* polymerase.

Table I. 4: Standard PCR reaction mix for genomic DNA amplification

Components	Volume
Genomic DNA (100-400 ng)	X µl
5X buffer (GC/HF)	10 µl
DMSO (only for GC buffer)	1.5 µl
dNTPs (200 µM each)	2 µl
Primer forward (0.5 µM)	1 µl
Primer reverse (0.5 µM)	1 µl
Water (ddH ₂ O)	Up to 50 µl

3. 2. 2. 7. Site-directed mutagenesis

Site-directed mutagenesis was used to make point mutations, and delete or insert single or multiple amino acids. In this method, *Pfu* DNA polymerase replicated both plasmid strands with high fidelity. The procedure was performed using a recombinant vector with an inserted gene of interest and two oligonucleotide primers containing the desired mutation. The primers were extended during temperature cycling by *Pfu* DNA polymerase. This incorporation of primers generated a mutant plasmid containing staggered nicks. The treatment of the product with *Dpn* I resulted in the digestion of the parental DNA template.

In this work site-directed mutagenesis was carried out for CYP264B1. In CYP264B1, the start codon (*Nde*I) was also present inside the ORF at 737 base pair towards 3' end. The *Nde*I site in the ORF was removed by a point mutation maintaining the original amino acid sequence. A point mutation was done with Site-Directed Mutagenesis Kit (Stratagene Ltd, Cambridge, UK) according to the manufacturer's instruction. For this, firstly, CYP264B1 encoding ORF was amplified from the genomic DNA of *So ce56* with primers (Table I. 2) in a reaction cocktail (Table I. 4) and cloned into pCR4 Blunt-Topo vector according to manufacturer's instruction. The positive clone of CYP264B1 was confirmed by automatic DNA sequencing at MWG Biotech. The positive clone of pCR4 Blunt-Topo_CYP264B1 was mutated my site directed mutagenesis with mutagenic primers (Table I. 2) in a reaction cocktail (Table I. 5) using programme 'Thermal cycle 6' (Table I. 4). The PCR product was digested with *Dpn*I for 4 hours at 37°C and 1µl was used for electro-transformation of competent *E. coli* TOP10F'. The positive mutated-clone of CYP264B1 in pCR4 Blunt-Topo vector was selected by restriction digestion and was confirmed by automatic DNA sequencing in MWG Biotech. Thus obtained positive mutated-insert was cut with *Nde*I and *Hind*III, and was finally cloned into expression vector pET17b at same restriction sites.

Table I. 5: PCR reaction mix for site directed mutagenesis

Components	Volume	Components	Volume
plasmid DNA (50 ng)	X µl	Primer 2 (0.2 µM)	1 µl
10 x <i>Pfu</i> reaction buffer	5 µl	<i>Pfu</i> DNA polymerase (2.5 u/µl)	0.5 µl
dNTPs (100 µM each)	2 µl	Water (ddH ₂ O)	Up to 50
Primer 1 (0.2 µM)	1 µl		µl

3. 2. 3. Microbiological Methods

3. 2. 3. 1. Cultivation and storage of bacteria

Unless otherwise stated, all the microbiological procedures for molecular cloning were carried out according to the standard methods as described previously (Sambrook and Russell, 2001).

3. 2. 3. 2. Preparation of competent cells

The process by which the foreign DNA is introduced into bacteria is known as bacterial transformation. When the bacteria grow under selective conditions (like antibiotic stress), the plasmid DNA remains stable in the cells, because of its resistance pattern to the selective antibiotic. The cells should be competent enough to incorporate the extra-chromosomal DNA. Depending upon the experimental purpose, chemical or electro-competent cells can be applied. In this work, both types of the competent cells were used.

3. 2. 3. 2. 1. Preparation of electro-competent cells

E. coli TOP10F' competent cell was used for the cloning purpose. The overnight seed culture of 4 ml was prepared and was inoculated in 400 ml of LB-medium (Merck). The culture was grown at 37°C at 180 rpm, till the absorbance reached to 0.4-0.5 at 600 nm. The culture flask was incubated on ice for 30 min and was centrifuged at 4000 rpm for 15 min at 4°C. The bacterial pellet was dissolved in 400 ml of ice cold sterile distilled water and was centrifuged as before. The pellet was again resuspended in 200 ml of ice cold sterile distilled water and was centrifuged as before. Thus obtained bacterial cell pellet was resuspended in 20 ml of ice cold sterile 10% glycerol and centrifuged as before. The pellet was finally dissolved in 2 ml of ice cold sterile 10% glycerol and was divided into aliquots of 50 µl in microfuge tubes. The tubes were immediately frozen by liquid nitrogen and stored at -80°C. The cell count was 3×10^{10} cells/ml.

3. 2. 3. 2. 2. Preparation of chemical competent cells

In this work, the competent cells were used for the protein expression and plasmid preparation purpose. The chemically competent cells were made from *E. coli* Top 10, BL21, BL21 (DE3), JM109 and Rosetta cells according to the need. Overnight cultures of each strain of *E. coli* cells were inoculated in 3 ml of nutrient broth (Merck) containing 10 mM Mg²⁺ either from 2M MgSO₄ or 2M MgCl₂ and were grown at 37°C for 12-16 hours at 180 rpm. The culture was 100 fold diluted in 50 ml main culture volume of LB medium containing 10 mM Mg²⁺ and was incubated at 37°C on a shaker at 150 rpm. The cell density was monitored once per

hour until the optical density (A_{600}) reached 0.3-0.5. The cell culture was harvested by centrifugation at 4000 rpm for 15 min at 4°C. The pellet was suspended in 14 ml TFP I buffer (Appendix II) and incubated on ice for 1 hour. After a further centrifugation, the cells were resuspended in 2 ml TFP II buffer (Appendix II). Aliquots of 100 μ l volume of the competent cells were transferred into a chilled sterile microfuge tube. The cells were frozen with liquid nitrogen and were stored at -70°C.

3. 2. 3. 3. Transformation of *E. coli*

3. 2. 3. 3. 1. Heat shock transformation

Heat shock transformation was applied to introduce the plasmid DNA into cells (Sambrook and Russell, 2001). The chemically competent cells were defrosted slowly on ice for 10 min and 1 μ l (20-100 ng) plasmid DNA was added. It was incubated for 30 min on ice. The heat shock was performed at 42°C for 90 sec followed by incubation on ice for 5 min. Then 250 μ l SOC medium (Appendix IV) was added to the cells and was incubated at 37°C for 1 h at 180 rpm. 100 μ l of the culture were plated on nutrient agar (Merck) plates containing appropriate antibiotic. The plates were incubated at 37°C overnight.

3. 2. 3. 3. 2. Electroporation

Higher transformation efficiency is generally required for cloning purpose so that the competent cells can have higher chance to incorporate the foreign plasmid. In this method, plasmid DNA was introduced into bacterial cells by dielectric breakdown of cell membranes. Electro-competent cells were thawed on ice and mixed with 3-5 μ l of dialyzed ligation mixture. The suspension (~50 μ l) was transferred to a chilled 0.1 cm electroporation cuvette and was jerked to remove air space. The cuvette was placed in Easyject Prima electroporator (Equibio) and an electric pulse of 18 kV/cm was applied. Cells were immediately resuspended in 400 μ l of SOC medium, incubated at 37°C for 1 hour and plated on nutrient agar plates containing appropriate antibiotics.

3. 2. 4. Protein expression and purification

3. 2. 4. 1. Heterologous gene expression of His₆-cytochrome P450

E. coli strains BL21, BL21 (DE3), JM109 and Rosetta and were used as heterologous hosts for the expression of P450s from *So ce56*. A single colony was inoculated into 50 ml Terrific Broth (TB) (Appendix IV) containing ampicillin (100 μ g per ml) and was grown overnight at 37°C. The main culture was inoculated with 1:100 dilution of the overnight culture in TB with

the same concentration of ampicillin and incubated at 37°C in a baffled flask at 90-120 rpm. When the optical density (A_{600}) of the culture reached 0.8 to 1.0, protein expression was induced with 0.8 mM isopropyl- β -D-thiogalactopyranoside (IPTG). In the mean time, 0.5 mM δ -aminolevulinic acid (δ -ALA) and ampicillin (50 μ g/ml) were added and cells were further grown at lower temperature (25°C to 30°C) for 36 to 48 hours in a New Brunswick Inova Shaker, model Excella E25. Expression of the P450s was monitored by analyzing the cell extracts by SDS-polyacrylamide gel electrophoresis.

3. 2. 4. 2. Cell lysis by sonication

The expressed cell pellet of P450 was harvested by centrifugation at 4°C. It was resuspended in 50 ml (per liter cell culture) of lysis buffer A (Appendix III). Phenylmethylsulphonyl fluoride (PMSF) was added to a final concentration of 1 mM and the suspension was sonicated in an ice bath. Cell debris was removed by centrifugation at $6 \times 10^4 \times g$ for 30 min at 4°C. The soluble fraction was used to record carbon monoxide-reduced difference (CO-difference or COD) spectra (Omura and Sato, 1964) to monitor the presence of P450. The parameters of the sonication are shown in Table 1.6.

Table 1. 6. Parameteres for sonication of cell pellet

Sonotrode type:	1	Pulse duration	30 sec
Oscillation amplitude	25 μ m	Medium temperature	10°C
Operating time	30 min	Stirrer speed	200 rpm
Pulse-duty factor	1	Pre-stir time	0 min

3. 2. 4. 3. Protein purification

Protein purification was carried out at 4°C using a purification unit (Äkta Prime, *Amersham Pharmacia* biotech) equipped with a flow cell UV/vis monitor and gradient mixer. The cytosolic fraction was applied to an immobilized metal ion affinity chromatography column (TALON, Clontech) that had been equilibrated with buffer B (Appendix III). The column was washed with 5 column volumes of buffer B. The tagged protein was eluted with buffer C (Appendix III). The fractions were collected and concentrated by ultra filtration (Centriprep 30, Millipore) to less than 2 ml. The concentrated sample was finally passed over a Superdex-75 column. The column was washed with 5 column volumes of buffer D (Appendix III). Fractions of highest purity ($A_{416}/A_{280} > 1.5$) were retained, pooled and concentrated by ultra filtration (Centriprep 30, Millipore).

3. 2. 4. 4. Determination of the protein concentration

Protein concentration was measured by using BC (Bicin-Choninic acid) Assay protein quantitation kit (Uptima Interchim, Montlucon, France) according to the manufacturer's instructions.

3. 2. 4. 5. SDS (Sodium Dodecylsulfate) polyacrylamide gel electrophoresis

The discontinuous gel electrophoresis was used to monitor the status of protein expression and the purity of the P450s obtained during successive steps of heterologous protein expression and purification. Proteins are separated according to their charge and molecular mass. Laemmli discontinuous gel electrophoresis method (Laemmli, 1970) was followed. In this work, for several different P450s, a 10 or 12 or 15 % separating gel solution (Appendix III) with 5% stacking gel solution (Appendix III) were used. The sample buffer composition was as described by Sambrook and Russell, 2001. The gels were stored in soaked papers with water at 4°C until use.

A portion of the protein sample was mixed with 2X SDS loading buffer (1:1 v/v) and heated at 100°C for 5 min. The samples were applied to the slots (wells) of the stacking gel and were run for 15 min at 120 V and then separated on a separating gel at 120 V until the bromophenol blue front, the tracking dye, reached to the bottom of the gel. The reference molecular weight marker, pre-stained broad range protein marker, from NEB (New England Biolabs, Beverly, MA, USA) was used. The gel was stained with Coomassie staining solution for 30 min at RT on shaker and then incubated in destaining solution until the protein bands become clearly visible. The gels were then slowly dried on a gel dryer (Model 583 gel dryer, BioRad) for storage.

3. 2. 5. Spectrophotometric characterization

3. 2. 5. 1. Spectrophotometric characterization of P450s of So ce56

All absorbance spectra were measured using a UV-2101PC, SHIMADZU spectrophotometer at room temperature. The carbon monoxide difference spectra of purified P450s were analyzed by the spectral method of Omura and Sato (Omura and Sato, 1964). Briefly, the solution of P450 enzyme (2.5 µM) in 50 mM potassium phosphate buffer (pH 7.5) was reduced with few grains of sodium dithionite. It was split into two cuvettes and a base line was recorded between 400 and 500 nm. The sample cuvette was bubbled gently with carbon monoxide for 30 seconds and a spectrum was recorded.

3. 2. 5. 2. Investigation of electron transfer partners

The native electron transfer partners of the *Sorangium cellulosum* So ce56 are not well known so far. The heterologous partners, spinach ferredoxin (Fdx)/ferredoxin reductase (FdR) and mammalian adrenodoxin (Adx)/adrenodoxin reductase (AdR) were used for this purpose. In this work, Adx (4-108) (Uhlmann *et al.*, 1994) was used. The competency of Fdx/FdR or Adx/AdR to reduce So ce56 P450s was determined by recording the formation of peak maxima at 450 nm during COD measurement (Lei *et al.*, 2004). Individual P450s with ratios of P450: Fdx/Adx: Fdr/AdR of 1:10:3 (200 pmol: 2 nmol: 600 pmol) were incubated with buffer E (Appendix III). After this NADPH was added to a final concentration of 1 mM. The spectrum of NADPH-reduced samples was recorded after bubbling with CO in the sample cuvette and was compared with the spectra of samples reduced with sodium-dithionite for each cytochrome P450.

3. 2. 6. Biophysical characterization

3. 2. 6. 1. Circular dichroism

Circular dichroism (CD) is the differential absorption of the left- and right-circularly polarized components of plane-polarized electromagnetic radiation. The CD spectroscopic effect is observed when linear polarized light passes through a chiral (asymmetric) environment which absorbs the right and left circularly polarized vector components to different extents (Holtzhauser, 1996). When the beams recombine after passing through the sample, elliptically polarized light is produced. Dichroism is then expressed as the ellipticity in degrees. The chromophore may be either intrinsically chiral or placed in an asymmetric environment.

In this work CD spectra were recorded at 25°C on a JASCO J715 spectrophotometer. Triplicate spectra of each were recorded and averaged. Far-UV spectra (195-260) were recorded with 5 µM P450 in 10 mM potassium phosphate buffer, pH 7.5. The near-UV spectra (300 to 450 nm) were recorded with 20 µM P450 in 10 mM potassium phosphate buffer, pH 7.5

3. 2. 6. 2. Electron Paramagnetic Resonance (EPR) measurements

An unpaired electron present in an atom has particular importance which will generate a net magnetic moment due to the spin and orbital magnetic moments of its unpaired electrons. The spin magnetic moments of unpaired electrons are randomly orientated in the absence of a

magnetic field, but the application of a field aligns the magnetic spin parallel or anti-parallel to the magnetic field orientation. In electron paramagnetic resonance (EPR) spectroscopy, the presence of an external magnetic field allows detection of unpaired electrons carrying an electron-spin magnetic moment. In a magnetic field each orbital energy level will be split into several energy levels and the transitions between them are associated with a net absorption of energy from the applied electromagnetic radiation (Hoff, 1989). This microwave absorption signal is also dependent on the orientation of the molecule bearing the unpaired electron with respect to the magnetic field, on the mobility of that molecule and on the presence of other unpaired electrons in the vicinity. The resulting resonance between the electron and a microwave field forms the basis of EPR, which is detected by a spectrometer to generate a signal. The frequency for EPR is calibrated in the dimensionless number called the g-factor which is defined in the following equation: $\nu = (g\beta B)/h$ where, ν = operating frequency of the spectrometer (constant), g = the g-factor, β = electron Bohr magnetron (9.273×10^{-23} Tesla), B = magnetic field (Tesla) at relevant point in EPR spectrum and h = Planck's constant. In the equation, $g\beta B$ describes interaction of electron spin with the magnetic field, characterized by the intrinsic g-factor (Zeeman interaction). The g-factor itself defines the centre of the resonance.

In this work, the EPR spectra were recorded on a Bruker ESP300 spectrometer (X-band, 9.5 GHz) equipped with a continuous flow helium cryostat ESR 900 and an ITC 4 temperature controller (Oxford Instruments) to allow measurements down to 5 K. The microwave frequency was measured by a HP 5350B frequency counter. Other parameters of measurements are given in the Figure legends in result. EPR measurements of oxidized P450s and the subsequent data analysis were done together with Dr. Reinhard Kappl, at the Institute of Biophysics, Universitätsklinikum, Universität des Saarlandes, Homburg.

Prior to the measurement, the oxidized form of So ce56 P450 samples of 100-150 μ M in 10 mM potassium phosphate buffer (pH 7.4) were transferred in EPR quartz tubes (Wilmad) and frozen in liquid nitrogen. The spectral parameters (g-tensor, line width) of the heme-centers were obtained by accumulation of the spectra with the programs Simfonia or Xsophe (Bruker). Standard measurement parameters used in this work were: modulation frequency- 100 kHz, microwave strength-10 mW, modulation amplitude-10 G and time constant- 0.3 sec.

3. 2. 7. Mass determination

In this work, seven of the P450s (CYP109C1, CYP109C2, CYP109D1, CYP260A1, CYP264A1, CYP264B1 and CYP266A1) were selected to determine the mass for both the apo- and holo-form. The mass of the P450s was measured in collaboration with Aurélie Mème, Laboratoire de Dynamique et Structure Moléculaire par Spectrométrie de Masse, Institut de Chimie, Starsbourg, France.

Prior to mass analysis, the samples were desalted on Vivaspin 500 (10 kDa) centrifugal concentrators (Millipore, Bedford MA, USA) in 20 mM ammonium acetate. Eight dilution/concentration steps were performed at 4°C and 10,000 rpm. Ammonium acetate enables native structures of proteins to be preserved and is compatible with ESI-MS analyses.

Mass spectrometric studies were performed in the positive ion mode using an ESI-TOF (Time-of-flight) mass spectrometer (microTOF, Bruker Daltonics, Bremen, Germany). Calibration of the instrument was performed using protonated horse myoglobin (Sigma) diluted to 2 pmol/μl in a 1:1 water-acetonitrile mixture (v/v) acidified with 1% formic acid. Purity and homogeneity of the samples were estimated by mass analysis in denaturing conditions: samples were diluted to 5 μM in a 1:1 water-acetonitrile mixture (v/v) acidified with 1% formic acid. In these conditions the non-covalent are disrupted, which allows the measurement of the molecular weight of the protein with a good precision. In non-denaturing conditions, the mass measurement of native cytochrome P450 was performed in 20 mM ammonium acetate to preserve its native conformation in solution. Samples were diluted to 10 μM. Great care was taken on instrumental parameters. Molecular species were assumed to be represented by series of peaks corresponding to multiply protonated ions.

3. 2. 8. Screening of compound library

In this work, four of the P450s (CYP109C1, CYP109C2, CYP260A1 and CYP266A1) were selected and a library consisting of ~17,000 substances was screened at the Forschungsinstitut für Molekulare Pharmakologie in Berlin, in cooperation with Prof. Oschkinat/Dr. von Kries.

The screening was done in 384 wells micro titer plate (MTPs) (3702 Corning) by an automated pipette robot cyclone ALH 3000 Workstation. This was equipped with plate-photometer (Tecan-Safire) and an automated dispenser (EasySypense2). During screening, for each plate 12 ml of P450s solution of final concentration 2.4 μM was used. The stocks of the

P450s (CYP109C1, CYP109C2, CYP260A1 and CYP266A1) were diluted in assay buffer (AB) (50 mM Tris-HCl, pH 7.5, 250 mM NaCl, 20% glycerin, 0.05% Tween-20 and 1 mM DTT). At each time only one of the P450s was used for screening. During screening, firstly 20 μ l of 1X AB was added with the dispenser. As illustrated in Figure I. 1, row 23 and 24 were used as controls. During screening, 0.5 μ M of each compound from the ligand library were added in each well of the titer plate with Cyclone ALH 3000 Workstation. Wavelength of 350-450 nm with an interval of 10 nm was fixed with Tecan-Safire plate-photometer and the absorbance of the substances were recorded. P450 solution of 30 μ l per well (1.5 μ M final concentration) was added, centrifuged and the mixture was incubated for 15 min at room temperature in the dark. The measured values were transported in excel data-sheet and were plotted as absorbance vs wavelength to observe Type I or Type II binding spectra. During the measurements, the background due to cytochrome P450 was subtracted from the value of the substance spectrum. For the validation, the hits of the final concentration of 200 μ M, 150 μ M, 100 μ M, 50 μ M and 16 μ M were prepared in 1X AB and 30 μ l per well of P450 (1.5 μ M final concentration) were added and plotted as before.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
A	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	RF	d
B	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	RF	d
C	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	RF	d
D	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	RF	d
E	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	RF	d
F	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	RF	d
G	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	RF	d
H	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	RF	d
I	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	P450	d
J	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	P450	d
K	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	P450	d
L	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	P450	d
M	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	P450	d
N	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	P450	d
O	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	P450	d
P	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	P450	d

Figure I. 1: The representative diagram of the 384 wells Micro Titer Plate (MTPs) that used for screening. The representative diagram shows that 23 and 24 rows were fixed for control. The references and controls are shown as: RF = Reference, d = DMSO as a Blank, s = replicated substance form ligand library and P450 = individual P450 solution of 1.5 μ M final concentration.

4. Results

4. 1. Bioinformatics analysis

4. 1. 1. Cytochrome P450 complement of *So ce56* (CYPome of *So ce56*)

The genome sequence project of *So ce56* revealed 21 putative cytochrome P450s. These P450s are distributed all over the genome which is illustrated in the chromosomal physical maps in Figure I. 2. The genomic abundance of P450 is 0.15% of all coding sequences, which is relatively low because of the large genomic size in comparison with 0.2% and 0.4% as described for *Streptomyces coelicolor* A(3)2 and *Streptomyces avermitilis*, respectively (Lamb *et al.*, 2003). All 21 P450s of *So ce56* possess a conserved threonine in the I-helix, and glutamic acid and arginine in the EXXR motif of the K-helix. The heme-binding signature (FxxGxxxCxG) is absolutely conserved in 19 P450s. In CYP267A1 and CYP265A1, the aromatic amino acid phenylalanine in the heme domain is replaced by the neutral aliphatic amino acid leucine. The amino acid residues conserved in the I-helix, K-helix and the heme-binding motif of individual P450 are illustrated in Table I. 7 with their respective amino acid positions. The details of the amino acids alignment of all 21 cytochrome P450s are illustrated in Appendix VIII. A.

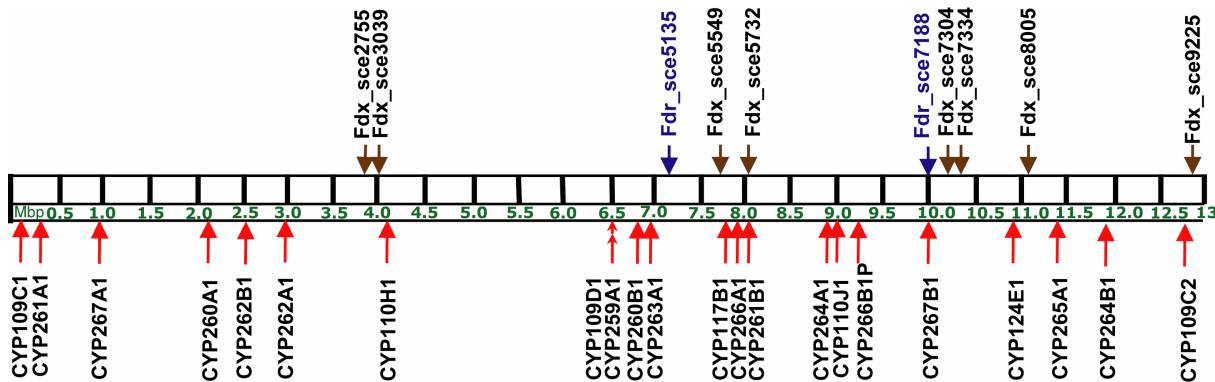


Figure I. 2: Physical map of the *Sorangium cellulosum* *So ce56* chromosome and distribution of cytochrome P450, ferredoxin (Fdx), and ferredoxin reductase (FdR) genes. All 21 *So ce56* P450s (red arrow) and one pseudo-cytochrome (CYP266B1P) with respective P450s annotation are assigned in the genomic location of the 13,033,779 base pairs genome. The distribution of 8 Fdx (brown arrow) and 2 FdR (in blue with blue arrow) are represented with respect to the gene number as annotated in NCBI.

Table I. 7: Conserved domains in *Sorangium cellulosum* So ce56 genome (So ce56 CYPome). Residues conserved in the I-helix (Thr), the K-helix (Glu and Arg), and the heme-binding motif (Phe, Gly and Cys) are highlighted in color. The unusual amino acid Leucine (L) is in italics-blue in heme-domain of CYP267A1 and CYP265A1. The numbers on the superscript of the starting and end of each domain represent the position of amino acids in P450 sequences.

P450 of So ce56	Gene name ²	I-helix	K- helix	Heme binding motif
CYP109C1	sce0122	²²⁶ AGTETA ²³¹	²⁶⁴ EEVLR ²⁶⁸	³²⁹ FGHG I HFCIG ³³⁸
CYP109C2	sce8913	²²⁶ AGHETT ²³¹	²⁶⁴ EEVLR ²⁸⁸	³²⁹ FGHG I HFCIG ³³⁸
CYP109D1	sce4633	²⁴¹ AGNDTT ²⁴⁶	²⁷⁹ EETLR ²⁸³	³⁴⁴ FGQ G THFCIG ³⁵³
CYP110H1	sce3065	²⁵¹ AGHETT ²⁵⁶	³⁰⁴ KEALR ³⁰⁸	³⁷⁵ FGGG H RR C IG ³⁸⁴
CYP110J1	sce6424	²⁵⁵ AGHETT ²⁶⁰	³⁰⁸ LEALR ³¹²	³⁷⁹ FGGG A RR C LG ³⁸⁸
CYP117B1	sce5528	²⁷² AGHETS ²⁷⁷	³²⁶ REALR ³³⁰	³⁹⁹ FGGG A H F CLG ⁴⁰⁸
CYP124E1	sce7867	²³⁹ AGHDTT ²⁴⁴	²⁷⁷ EEMLR ²⁸¹	³⁴² FG I G P H V CLG ³⁵¹
CYP259A1	sce4635	²⁵⁰ AGHETS ²⁵⁵	³⁰⁶ REVLR ³¹⁰	³⁷⁹ FGGG P H F CLG ³⁸⁸
CYP260A1	sce1588	²⁷⁹ GGYETT ²⁸⁴	³¹⁷ EEGMR ³²¹	³³¹ FGGG A H F CVG ³⁴⁰
CYP260B1	sce4806	²²⁸ GSYETT ²³³	²⁶⁶ EESTR ²⁷⁰	³³¹ FGGG L H Y CVG ³⁴⁰
CYP261A1	sce0200	²⁸⁷ AGEDTT ²⁹²	³⁴⁴ QETLR ³⁴⁸	⁴²¹ FG S G P R V CPG ⁴³⁰
CYP261B1	sce5725	²⁹⁹ AGEDTT ³⁰⁴	³⁵⁵ QETLR ³⁵⁹	⁴³¹ FGGG P R V CPG ⁴⁴⁰
CYP262A1	sce2191	²⁴⁷ AGNETT ²⁵²	³⁰² QEAMR ³⁰⁶	³⁷⁶ FGGG P R Q CIG ³⁸⁵
CYP262B1	sce1860	²⁷⁹ AGYETT ²⁸⁴	³³⁴ QEALR ³³⁸	⁴⁰⁸ FG I G Q R Q CVG ⁴¹⁷
CYP263A1	sce4885	²⁶³ VGHETS ²⁶⁸	³¹⁸ NECLR ³²²	³⁹² FG A G Q R I CLG ⁴⁰¹
CYP264A1	sce6323	²²⁹ GGLETT ²³⁴	²⁶⁷ EEMMR ²⁷¹	³³¹ FGHG A H F CLG ³⁴⁰
CYP264B1	sce8551	²³⁵ AGLETS ²⁴⁰	²⁷³ EEV M R ²⁷⁷	³³⁸ FGHG A H F CLG ³⁴⁷
CYP265A1	sce8224	²³⁴ GGLD T T ²³⁹	²⁷² EEILR ²⁷⁶	³³⁷ L G A G P H H CLG ³⁴⁶
CYP266A1	sce5624	²⁴⁹ AGFET V ²⁵⁴	²⁸⁷ DEMLR ²⁹¹	³⁵³ FGGG H H F CVG ³⁶²
CYP267A1	sce0675	²⁶³ GGHETT ²⁶⁸	³⁰¹ EELLR ³⁰⁵	³⁶⁶ L GL G V H Y C AG ³⁷⁵
CYP267B1	sce7167	²⁴³ AGHETT ²⁴⁸	²⁸¹ EEALR ²⁸⁵	³⁴⁷ FGGG I H F CLG ³⁵⁶

¹ CYP names as annotated at website <http://drnelson.utmem.edu/CytochromeP450.html>

² Gene name as annotated in NCBI

4. 1. 2. Phylogenetic relationship of *So ce56* CYPome

The huge genome of *S. cellulosum* So ce56 contains nine new families of cytochrome P450s. This includes CYP259 (subfamily CYP259A1), CYP260 (subfamilies CYP260A1 and CYP260B1), CYP261 (subfamilies CYP261A1 and CYP261B1), CYP262 (sub-families CYP262A1 and CYP262B1), CYP263 (subfamily CYP263A1), CYP264 (subfamilies CYP264A1 and CYP264B1), CYP265 (subfamily CYP265A1), CYP266 (subfamily CYP266A1) and CYP267 (subfamilies CYP267A1 and CYP267B1). The remaining seven P450s sequences belong to the existing P450 families. Three of the P450s sequences (CYP109C1, CYP109C2 and CYP109D1), belong to the existing CYP109 family. Two enzymes (CYP110H1 and CYP110J1) are the members of the CYP110 family. CYP117B1 and CYP124E1 belong to the already existing CYP117 and CYP124 families, respectively. The phylogenetic relationship of all P450s in *So ce56* shows the relatedness of the different P450 families (Figure I. 3). CYP266A1, a new member of CYP166 family, was found to be distantly related to the rest of the *So ce56* P450s.

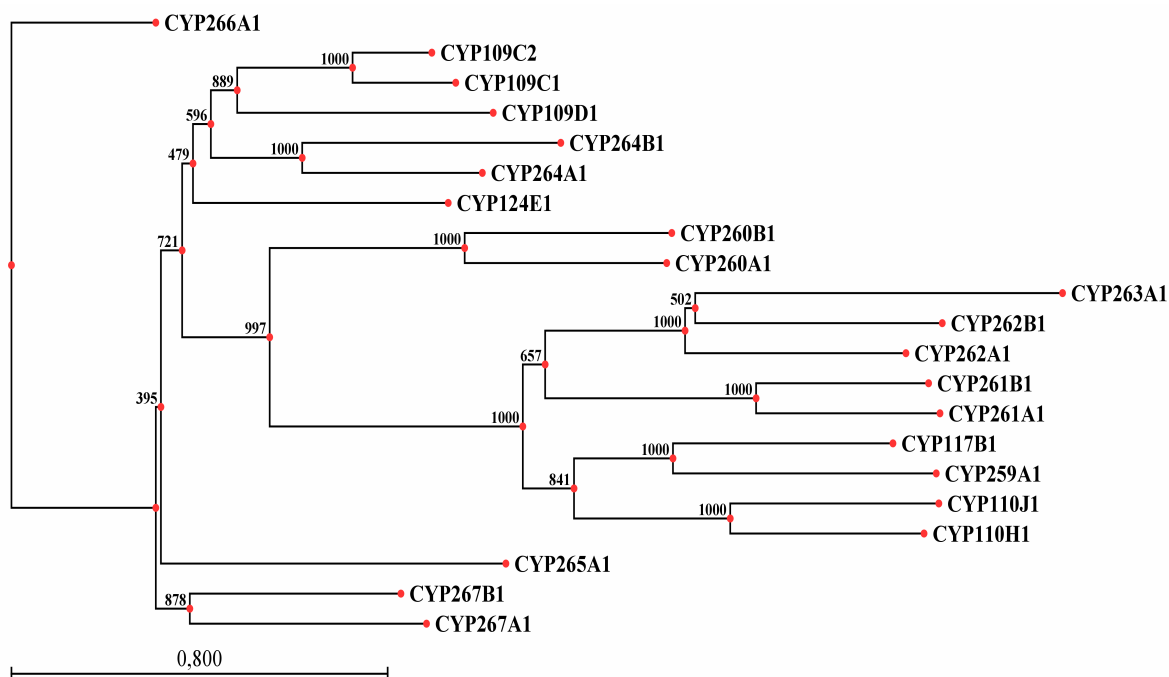


Figure I. 3. Phylogenetic tree of *Sorangium cellulosum* So ce56 cytochrome P450. The alignment was done with 10 gap setting and 1 gap extension having slow (very accurate) alignment input in CLC Workbench. The tree was constructed by using Neighbor Joining Algorithm with performing boot strap analysis of 1000 times replicate. Bootstrap values are shown at branch nodes. The bar in the left corner indicates 0.8 amino acid substitutions per amino acid for the branch length.

4. 1. 3. The relationship of P450s of So ce56

The relatedness of P450s of So ce56 was studied and their similarities were compared with its own P450s members. Moreover, to study the phylogenetic relationship each of the P450s of So ce56 was compared with other bacterial P450s in the database of P450s. For this, the amino acid sequences of each of the P450s were used as a query for the local BLAST search of the GenBank using a recent non-redundant protein data base from NCBI and the closest homologues for the different myxobacterial P450s were identified and illustrated in Table I. 8.

Among 21 P450s, only four families (CYP109, CYP110, CYP117 and CYP124) are belonging to the assigned P450 families. CYP109 family in So ce56 has 3 members (CYP109C1, CYP109C2 and CYP109D1). CYP109C1 has 68.1% sequence identity with CYP109C2 and 39.4% with CYP109D1 whereas CYP109C2 shows 38.8% sequence identity with CYP109D1 of Se ce56. These three members almost have 40% sequence identity with the members of CYP195A4 of *Stigmatella aurantiaca*, CYP113A1 of *Saccharopolyspora erythraea*, CYP107A1 of *Streptomyces avermitilis* and 30% identity with CYP105D5 of *S. coelicolor* A(3)2. (Table I. 8). This relatedness assumed the potential scope of CYP109D1 for xenobiotic and secondary metabolism as well as fatty acids hydroxylation as these homologues was performing such activities (www.drnelson.utm.edu/cytochrome_P450.html). CYP110H1 and CYP110J1 of So ce56 are the members of a larger family having nearly 40% identity with CYP110E1 and 44% identity with CYP110E4 of *Nostoc* sp and *Gloeobacter violaceus*, respectively, and the relatedness indicate the probability of having alkane affinity fatty acid hydroxylation (Torres *et al.*, 2005). CYP117B1 of So ce56 shares 41% sequence identity with CYP117A1 of *Bradyrhizobium japonicum* and CYP117A3 of *Mesorhizobium loti*, respectively, which suggested a probable role for terpenoid synthesis/conversion (Tully *et al.*, 1998). CYP124E1 of So ce56 has 44% and 40% sequence identity with CYP124A1 of *Mycobacterium tuberculosis* and CYP124 of *Rhodococcus* sp (strain RHA1), respectively, but both of them were uncharacterized P450s and their experimental evidences were not know till now. As a result, the putative function of CYP124A1 was not speculated.

The newly assigned P450 families (CYP259, CYP260, CYP261, CYP262 and CYP265) have 30-39% sequence identity with several other bacterial P450 (Table I. 8). However, the experimental data for them were not existed and their functions were not predicted. Likely, the other new family (CYP264, CYP266 and CYP267) of So ce 56 exhibits 36-40% sequence

identity with CYP107 of *Streptomyces* sp. and other bacterial P450s (Table I. 8) which suggest the homologue to have a potential role for xenobiotic and secondary metabolism (Haydock *et al.*, 1991; Jungmann *et al.*, 2005).

Table I. 8: Cytochrome P450s in *Sorangium cellulosum* So ce56 with their closest homologues. Pairwise sequence alignment was done to find the identity, similarity and the gaps of respective P450s.

CYP ^a names	Match in database ^b				
	Speices	CYP ^a name	Gene ID ^c	% Identity ^d	Gaps %
CYP109C1	<i>Solibacter usitatus</i>	-	Q023S4	52	1
	<i>Bacillus subtilis</i>	CYP109B1	<i>yjiB</i>	43	0
	<i>Saccharopolyspora erythraea</i>	CYP107B1	M83110	41	8
CYP109C2	<i>Bacillus subtilis</i>	CYP109B1	<i>yjiB</i>	43	0
	<i>Saccharopolyspora erythraea</i>	CYP113A1	<i>eryK</i>	40	3
	<i>Bradyrhizobium japonicum</i>	CYP107AA1	BAC51802	40	5
CYP109D1	<i>Saccharopolyspora erythraea</i>	CYP107B1	CAL99581	41	1
	<i>Bacillus subtilis</i>	CYP109A1	NP_389103	39	7
	<i>Streptomyces venezuelae</i>	CYP107L1	AAC64105	37	5
CYP110H1	<i>Nostoc</i> sp.	CYP110E1	BAB76532	40	2
	<i>Gloeobacter violaceus</i>	CYP110E4	AP006578	44	2
	<i>Myxococcus xanthus</i>	CYP209A1	CAB40542	36	7
CYP110J1	<i>Gloeobacter violaceus</i>	CYP110E5	AP006578	42	4
	<i>Nodularia spumigena</i>	CYP110D2	EAW43774.1	39	3
	<i>Nostoc punctiforme</i>	-	ZP_00109203	40	3
CYP117B1	<i>Mesorhizobium loti</i>	CYP117A3	NP_106891	41	0
	<i>Bradyrhizobium japonicum</i>	CYP117A1	BJ-4	41	0
	<i>Burkholderia rhizoxina</i>	-	CAL69892	40	0
CYP124E1	<i>Rhodococcus</i> sp	CYP125	ABG96453	45	0
	<i>Mycobacterium tuberculosis</i>	CYP124A1	Rv2266	44	3
	<i>Stigmatella aurantiaca</i>	-	EAU6400	42	1
CYP259A1 [*]	<i>Mesorhizobium loti</i>	CYP117A3	NP_106891	36	1
	<i>Bradyrhizobium japonicum</i>	CYP117A1	BJ-4 U12678	34	1
	<i>Pseudomonas fluorescens</i>	-	AAY92264.1	33	1
CYP260A1 [*]	<i>Saccharopolyspora erythraea</i>	CYP107B1	M83110	35	5
	<i>Micromonospora griseorubida</i>	CYP107E1	JP95	30	4
	<i>Streptomyces</i> sp.	CYP105D9	AF509565	33	3

CYP260B1*	<i>Streptomyces tubercidicus</i>	CYP105S2	AAT45294	31	8
	<i>Rhodococcus</i> sp.	CYP105	ABG94409	32	4
	<i>Pseudomonas aeruginosa</i>	CYP107S1	ABJ12582	32	8
CYP261A1*	<i>Burkholderia phytofirmans</i>	-	EAV04662	44	0
	<i>Streptomyces avermitilis</i>	CYP184A1	<i>cyp21</i>	29	5
	<i>Magnetospirillum magnetotacticum</i>	CYP173B1	AAP01002719	30	6
CYP261B1*	<i>Burkholderia phytofirmans</i>	-	EAV04662	44	1
	<i>Streptomyces avermitilis</i>	CYP184A1	<i>cyp21</i>	31	5
	<i>Streptomyces avermitilis</i>	CYP102B2	<i>cyp27</i>	30	4
CYP262A1*	<i>Myxococcus xanthus</i>	-	ABF92999	37	1
	<i>Bacillus halodurans</i>	CYP197A1	BAB04298	35	4
	<i>Chloroflexus aurantiacus</i>	-	ABY34735	41	1
CYP262B1*	<i>Chloroflexus aurantiacus</i>	CYP205A1	ABY34735	40	2
	<i>Streptomyces carzinostaticus</i>	CYP208A2	AY117439	35	4
	<i>Streptomyces avermitilis</i>	CYP184A1	<i>cyp21</i>	35	4
CYP263A1*	<i>Burkholderia thailandensis</i>	CYP184A1	BTH_II2347.	34	0
	<i>Streptomyces avermitilis</i>	CYP208A1	<i>cyp21</i>	35	5
	<i>Streptomyces globisporus</i>	-	AAL06697	35	3
CYP264A1*	<i>Stigmatella aurantiaca</i>	CYP107B1	EAU64026	58	1
	<i>Streptomyces</i> sp.	CYP107AE1	BD133545	37	5
	<i>Streptomyces avermitilis</i>	CYP107X1	<i>cyp23</i>	37	4
CYP264B1*	<i>Arthrobacter</i> sp.	-	ABK03359	39	2
	<i>Streptomyces avermitilis</i>	CYP147B1	<i>cyp3</i>	33	3
	<i>Bacillus licheniformis</i>	-	<i>yjiB</i>	32	2
CYP265A1*	<i>Saccharopolyspora erythraea</i>	CYP105P1	<i>cypA</i>	40	1
	<i>Streptomyces avermitilis</i>	CYP151A1	<i>pteC</i>	38	3
	<i>Mycobacterium smegmatis</i>	-	AF102510	37	1
CYP266A1*	<i>Bradyrhizobium japonicum</i>	CYP107AA1	BAC51802	40	1
	<i>Streptomyces avermitilis</i>	CYP147B1	<i>cyp3</i>	37	4

	<i>Bacillus subtilis</i>	CYP107H1	BioI	36	3
CYP267A1*	<i>Roseiflexus</i> sp.	-	ABQ89181	45	1
	<i>Streptomyces avermitilis</i>	CYP107P2	<i>cyp20</i>	41	3
	<i>Streptomyces venezuelae</i>	CYP107L1	<i>pikC</i>	41	6
CYP267B1*	<i>Stigmatella aurantiaca</i>	CYP107B1	EAU69748	51	0
	<i>Streptomyces venezuelae</i>	CYP107L1	AF087022	44	3
	<i>Bacillus amyloliquefaciens</i>	CYP107K1	CAG23962	42	2

Note: * = New members of P450 families found in So ce56 ; a = P450 names as annotated at the website: <http://drnelson.utmem.edu/CytochromeP450.html>; b = database search at NCBI; c = gene ID deposition in EMBL; d = the percentage identity represents the number of exact matches over the closed P450 length.

4. 2. Genome mining of cytochrome P450

Sequence alignment, homology modeling and spin-shift analysis are the existing theoretical and practical approaches for the prediction of potential substrates for several orphan cytochrome P450s. These approaches generally abolish the natural preponderance of corresponding neighboring genes of P450 in a genome. The study of proximal genes is more beneficial for those P450s which are in a cluster or in an operon with redox partners or with secondary metabolizing genes. The study of genomic organization of corresponding P450s may help to hypothesize some potential natural redox partners and/or potential natural/analogue substrates. Though various endogenous roles of P450s in several polyketide synthetase (PKS) gene cluster during secondary metabolite formation are already verified in some other strain of *Sorangium cellulosum* (Julien *et al.*, 2000; Tang *et al.*, 2000, Buntin *et al.*, 2008), the 13.1 Mbp genome of *Sorangium cellulosum* So ce56 has not been investigated so far. As this is the very first study of P450s in So ce56, a detailed genomic organization for each of the P450s with respect to the neighboring genes was studied. The genomic foot prints in accordance with the genome database of So ce56 with the direction of transcription (with arrow head of gene) were studied.

4. 2. 1. P450s potentially involved in secondary metabolite formation

It was observed that the genomic organization of P450s in So ce56 is different in comparison with other secondary metabolite producing strains like *Streptomyces* sp. In So ce56, though 21 P450s are detected, none of them are directly organized in a cluster with ferredoxins or ferredoxin reductases. However, two of the P450s, CYP263A1 and CYP265A1, are found in

a cluster with putative PKS and NRPS modules, respectively. The PKS module (sce4888) is detected downstream of CYP263A1 (sce4885). It is speculated that CYP263A1 potentially be involved in secondary metabolite biosynthesis (Figure I. 4).



Figure I. 4: The genomic organization of CYP263A1. CYP263A1 (brown) and the PKS gene (sce4888 in green) are shown. The respective neighboring genes are - *xynB4* (sce4884): xylan 1,4-beta-xylosidase, *sce4885*: CYP263A1, *fabG9* (sce4886): 3-oxoacyl-(acyl-carrier-protein) reductase, *sce4887*: putative phosphopantetheinyl transferase, *sce4888*: predicted PKS modules, *sce4889*: MarR family transcriptional regulator protein, *sce4890*: hypothetical protein, *sce4891*: protein kinase.

Likely, CYP265A1 is found close to the putative NRPS module (sce8219) upstream in the genome which could have a potential role during secondary metabolite formation (Figure I. 5). The neighboring genes of CYP265A1 (sce8224) also has several remarkable genes in upstream [a putative FkbH-protein (sce8221) which supposed to be involved in the methoxymalonyl-ACP biosynthesis] and downstream [a SAM-dependent methyl transferase (sce8225)].

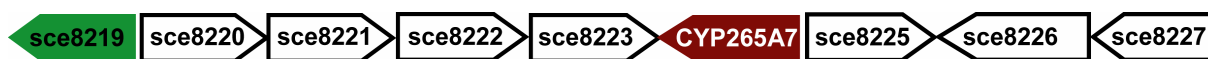


Figure I. 5: The genomic organization of CYP265A1. CYP265A1 (brown) and the NRPS gene (sce8219 in green) are shown. The respective neighboring genes are - *sce8219*: AMP-dependent synthetase and ligase which function as non-ribosomal peptide synthetase (NRPS) modules and related proteins, *sce8220*: hypothetical protein, *sce8221*: putative FkbH-protein, *sce8222*: predicted acyl carrier protein, *sce8223*: hypothetical protein, *sce8224*: CYP265A1, *sce8225*: SAM-dependent methyl transferase, *sce8226*: acyl-CoA-ligase/synthetase, *sce8227*: putative acyl carrier protein.

4. 2. 2. P450s with potential physiological role

In the genomic organization of *So ce56*, three of the P450s (CYP124E1, CYP262A1 and CYP266A1) are clustered with other neighboring genes having some putative physiological functions. The P450s present in this cluster speculated to have some role to produce important derivatives having some physiological significance during the bacterial lifecycle in an ecological niche. Concerning this, a gene *sce7866* in the adjacent upstream of CYP124E1 (sce7867) has a putative function of protocatechuate catabolism. Furthermore, CYP124E1 also has a putative phenylacetic acid (PA) degradation protein encoding gene (*sce7863*) in upstream (Figure I. 6. A). Similarly, in the upstream of CYP262A1 (sce2191), a putative

hydrolase encoding gene (*xysA*, sce2188) (Figure I. 6. B) is present. In CYP266A1 (sce5624), a gene encoding ubiquinone/menaquinone biosynthesis protein (*ubiE4*, sce5625) is present directly at an adjacent downstream. It was noticed that, CYP266A1 is the only P450 in So ce56 which has a NADH-flavin oxidoreductase (ferredoxin reductase, sce5629) within 4 genes downstream (Figure I. 6. C).

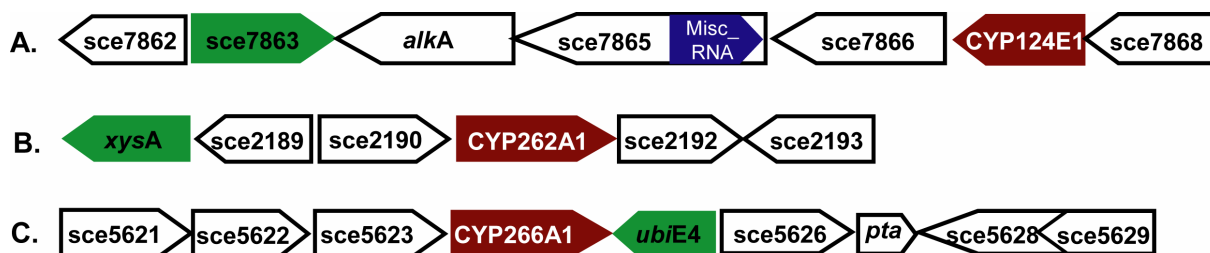


Figure I. 6: The genomic organization of P450s clustered with putative functional domain. The genes of P450s and the physiologically significant domains are shown in brown and green, respectively. **A. Neighboring genes of CYP124E1.** The respective neighboring genes are - sce7862: sigma-54 dependent DNA-binding response regulator, sce7863: predicted penylacetic acid (PA) degradation protein in bacteria, *alkA* (sce7864): putative DNA-3-methyladenine glycoside, sce7865: hypothetical protein, sce7866: predicted protocatechuate catabolism protein, sce7868: hypothetical protein. The miscellaneous RNA (Misc_RNA) overlapped on sce7865 is shown in blue. **B. Neighboring genes of CYP262A1.** The respective neighboring genes are - *xysA*: endo-1-4-beta xylanase, sce2189: hypothetical protein, sce2190: putative hydrolase, sce2191: CYP262A1, sce2192: hypothetical protein, sce2193: hypothetical protein. **C. Neighboring genes of CYP266A1.** The respective neighboring genes are - sce5621: predicted iron-sulphur (Fe-S) oxidoreductase, sce5622: putative methyl transferases, sce5623: SAM dependent methyl transferases, sce5624: CYP266A1, *ubiE4* (sce5625): ubiquinone/menaquinone biosynthesis protein, sce5626: transcriptional regulators, *pta* (sce5627): putative acyl transferases, sce5628: hypothetical protein, sce5629: NADH-flavin oxidoreductase. The ORFs of sce5628 and sce5629 are shown overlapped.

4. 2. 3. P450 clustered with regulatory elements

It was found that some of the P450s (CYP109C2, CYP110H1, CYP110J1, CYP260A1, CYP261A1 and CYP262B1) of So ce56 are found in a cluster with regulatory elements. CYP109C2 (sce8913) is clustered with the transcriptional regulator AraC family protein encoding genes (sce8912) adjacent upstream and protein kinase (sce 8916) downstream (Figure I. 7. A). CYP110H1 (sce3065) is clustered with two components hybrid system kinase (sce 3063) as a regulator towards 2 genes upstream. In an adjacent downstream of CYP110H1 a conserved hypothetical protein encoding gene (sce3066) having identity with the DUF899 thioredoxin family protein superfamily has been identified (Figure I. 7. B). The cluster of CYP110J1 (sce6424) has two regulators, transcriptional regulator (sce6423) upstream and protein kinase (sce6427) downstream. (Figure I. 7. C). The cluster of CYP260A1 (sce1588)

has a protein kinase regulator (sce1587) in an adjacent upstream along with a plethora of hypothetical protein encoding genes. (Figure I. 7. D). Likely, CYP261A1 (sce0200) clustered with two component sensor histidine kinase regulator (sce 0195) upstream. This P450 has several hypothetical protein (HP) encoding genes in both upstream and downstream (Figure I. 7. E). Likely, CYP262B1 (sce1860) has a regulator encoding gene (sce1859) directly at an adjacent upstream position (Figure 7. F). It has been shown that these all kinds of regulators are generally the eukaryotic-like-regulators (ELR) which was recently studied in *So ce56* (Perez *et al*, 2008) The real significance of these regulators with respective to P450s has not been studied so far.

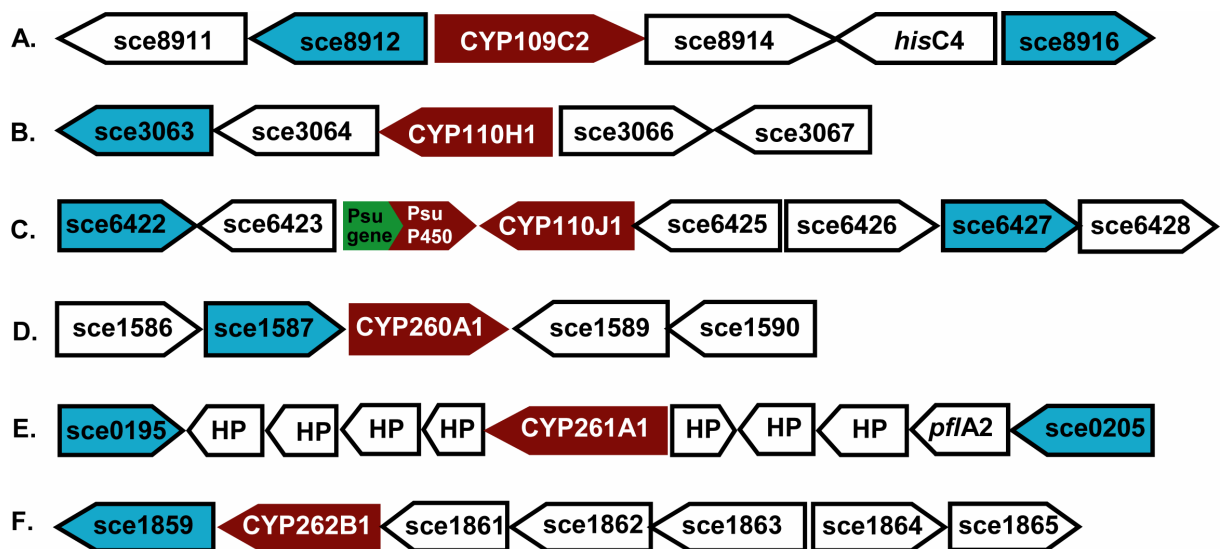


Figure I. 7: The genomic organization of P450s clustered with regulatory elements. The neighboring genes of CYP109C2, CYP110H1, CYP110J1, CYP260A1, CYP261A1 and CYP262B1 are illustrated. The genes of P450s and the regulators are shown in brown and blue, respectively. **A. Neighboring genes of CYP109C2.** The respective neighboring genes are - sce8911: probable class-V aminotransferase, sce8912: transcription regulator, AraC family protein, sce8913: CYP109C2, sce8914: hypothetical protein, *hisC4* (sce8915): histidinol-phosphate/aromatic aminotransaminase protein, sce8916: protein kinase. **B. Neighboring genes of CYP110H1.** The respective neighboring genes are - sce3063: two component hybrid system and regulator, sce3064: alpha-amylase catalytic region, sce3065: CYP110H1, sce3066: hypothetical protein, sce3067: hypothetical protein. **C. Neighboring genes of CYP110J1.** The respective neighboring genes are - sce6422: ArsR family transcriptional regulator, sce6423: transcriptional regulator MarR family protein., sce6424: CYP110J1, sce6425: hypothetical protein, sce6426: hypothetical protein, sce6427: protein kinase, sce6428: hypothetical protein. The pseudo-genes and the pseudo-P450 are shown in an adjacent downstream of CYP110J1 in green and brown, respectively. **D. Neighboring genes of CYP260A1.** The respective neighboring genes are - sce1586: hypothetical protein, sce1587: protein kinase, sce1588: CYP260A1, sce1589: hypothetical protein, sce1590: hypothetical protein. **E. Neighboring genes of CYP261A1.** The respective neighboring genes are - sce0195: two component sensor histidine kinase, *pflA2*: pyruvate formate lyase-activating enzyme, sce0205: two component sensor histidine kinase. The genes, represented

by HP, in the upstream and downstream are all hypothetical proteins. **F. Neighboring genes of CYP262B1.** The respective neighboring genes are - sce1859: protein kinase with two component sensor domain, sce1860: CYP262B1, sce1861: hypothetical protein, sce1862: periplasmic surface binding protein, sce1863: putative ABC type permease protein, sce1864: hypothetical protein, sce1865: ABC transporter ATP-binding protein.

4. 2. 4. P450s clustered with carbohydrate related regulatory elements

Two of the P450s, CYP109C1 and CYP267A1, are identified in a cluster of carbohydrate related regulatory elements. CYP109C1 (sce0122) is clustered with a predicted universal stress protein (*uspA*, sce0123) in an adjacent downstream position and a beta-xylosidase (*yagH*, sce0125), the third gene downstream (Figure I. 8. A). Likely, CYP267A1 (sce0675) is clustered with a polyphosphate-glucose phospho transferase (*ppgK*, sce0674) adjacent upstream and a succinyl glutamic semialdehyde dehydrogenase encoding gene (sce0676) in adjacent downstream position (Figure I. 8. B). The miscellaneous RNA present in the downstream of both of these P450s, make them more complex to predict their physiological significance.

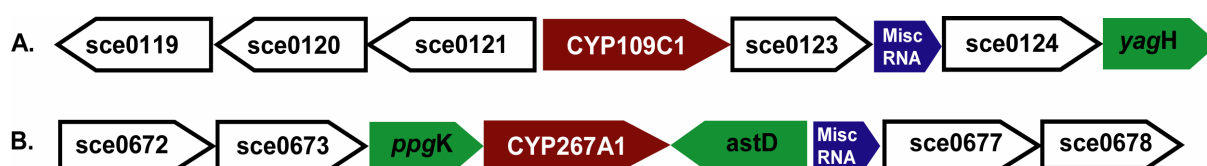


Figure I. 8: The genomic organization of P450s clustered with carbohydrate related regulatory elements. The cytochrome P450s (CYP109C1 and CYP267A1) and the carbohydrate related genes are shown in brown and green, respectively. The miscellaneous RNA (Misc_RNA) is shown in blue. **A. Neighboring genes of CYP109C1.** The respective neighboring genes are - sce0119: putative cell surface protein, sce0120: putative *ompA* family protein, sce0121: hypothetical protein, sce0122: CYP109C1, sce0123: predicted universal stress protein (*uspA*), sce0124: putative sodium solute transport protein, *yagH* (sce0125): beta-xylosidase protein. **B. Neighboring genes of CYP267A1.** The respective neighboring genes are - sce0672: hypothetical protein, sce0673: putative pectinacetyltransferase, *ppgK* (sce0674): phosphate-glucose phosphotransferase, sce0675: CYP267A1, *astD* (sce0676): succinyl glutamic semialdehyde dehydrogenase, sce0677: putative phosphoesterase, sce0678: hypothetical protein.

4. 2. 5. P450 clustered with a terpene cyclase

In So ce 56, only one P450, CYP264B1, is clustered with terpene cyclase gene. In the genome, CYP264A1 (sce8551) has a terpene cyclase gene, *geoA* (sce8552) adjacent downstream. It is observed that CYP264B1 and *geoA* are only separated by 63 bp (Figure I. 9) which could possibly be arranged in an operon. CYP264B1 might have a potential role during biosynthesis of putative cyclase.



Figure I. 9: The genomic organization of P450 clustered with a terpene cyclase. CYP264B1, terpene cyclase gene and regulators are shown in brown, green and blue, respectively. The neighboring genes of CYP264B1 are - sce8548: ABC Transporter/ATP-binding protein, sce8549: putative short-chain dehydrogenase, sce8550: transcriptional regulator, TetR family protein, sce 8551: CYP264B1, *geoA* (sce8552): putative terpene cyclase (geosmin), sce8553: protein kinase.

4. 2. 6. P450s with unique clusters

CYP109D1 (sce4633) and CYP259A1 (sce4635) are found in a gene cluster of *So ce56* genome. The P450 genes are in an anti-directional orientation in the genome and separated by a single gene, sce4634, encoding anti-anti sigma regulatory factor (Figure I. 10. A). The adjacent upstream gene (sce4632) of CYP109D1 belongs to the beta-CA superfamily protein and codes carbonate hydratases (CA) which are the zinc containing enzyme that catalyses the reversible hydration of carbon dioxide. This gene was found as a conserved gene in *Myxococcus xanthus* DK1622 (73% identity), *Streptococcus* sp N1 (52% identity) and in several *Acinetobacter* sp ADP1 (52% identity). The rest of the entire upstream gene clusters in the vicinity of this P450 codes for hypothetical proteins. The adjacent downstream gene (sce4636) of CYP259A1 (sce4635) has a conserved region of squalene cyclase (i.e. 343-416 amino acids out of 511 amino acid sequence). This gene is also conserved in *Rhizobium* sp (40% identity), *Mesorhizobium* sp (39% identity), and several other actinomycetes like *Mycobacterium tuberculosis* CDC1551 (30% identity) as a putative cyclase



Figure I. 10. A: The unique cluster of CYP109D1 and CYP259A1. The P450s (CYP109D1 and CYP259A1) and the regulatory gene are shown in brown and green, respectively. The neighboring genes are - sce4630: hypothetical protein, sce4631: hypothetical protein, sce4632: hypothetical protein, sce4633 (*cypA2*): CYP109D1, sce4634: anti-anti sigma regulatory factor, sce4635: CYP259A1, sce4636: hypothetical protein.

CYP267B1 also lies in a unique cluster. The neighboring region of CYP267B1 (sce7167) has a DNA repair exonuclease family protein (sce7166) in upstream and a hypothetical protein encoding gene (sce7168) in downstream position (Figure I. 10. B). A unique cluster of tymovirus analogue gene (sce7164), called as tymovirus 45/70 Kd protein, is present in the upstream of CYP267B1 in which a hypothetical protein encoding gene (sce7165) is

overlapped. Interestingly, several other genes encoding α -gluconotransferase (*malQ4*, *sce7169*), α -amylase family protein (*sce7170*) and ABC transporter ATP family protein (*sce7171*) were also present in downstream position of CYP267B1.



Figure I. 10. B: The unique cluster of CYP267B1. CYP267B1 (brown) and the overlapped gene of *sce7165* (green) on *sce7164* are shown. The neighboring genes of CYP267B1 are - *sce7165*: hypothetical protein, *sce7164*: tymovirus 45/70Kd protein, *sce7166*: DNA repair exonuclease family protein, *sce7167*: CYP267B1, *sce7168*: hypothetical protein, *malQ4* (*sce7169*): 4- α gluconotransferase, *sce7170*: alpha amylase family protein, *sce7171*: ABC transporter ATP binding protein.

CYP264A1 (*cypA5*) is also associated with other unique P450 clusters. CYP264A1 (*sce6323*) has a sigma-54 dependent transcriptional regulator (*sce6324*) in an adjacent downstream with a very unique signal domain XylR-N/v4r. This domain functions as an activator of aromatic catabolism (Jelsbak *et al.*, 2005). The putative function of this P450 was not hypothesized as all the remaining genes in upstream and downstream are coding for hypothetical proteins (Figure I. 10. C).



Figure I. 10. C: The unique clusters of CYP264A1 (*cypA5*). CYP264A1 (brown), the regulator (blue) and the structural (Str) RNA (deep blue) are shown. The neighboring genes of CYP264A1 are - *sce6320*: hypothetical protein, *sce6321*: ani-anti sigma factor protein, *sce6322*: hypothetical protein, *cypA5*: CYP264A1, *sce6324*: sigma-54 dependent transcriptional regulator, *sce6325*: hypothetical protein, *sce6326*: hypothetical protein.

4. 2. 7. P450s clustered only with hypothetical protein

The rest of the other P450s of So ce56 (CYP117B1, CYP260B1 and CYP261B1) were only clustered with hypothetical protein encoding genes. The adjacent and the neighboring genes of these P450s do not have any putative conserved domain to predict their potential function as a result, the putative functions of these P450s was not hypothesized.

4. 3. Cloning, expression and purification of So ce56 cytochrome P450

4. 3. 1. Amplification and cloning of the So ce56 CYPome

Amplification of all 21 cytochrome P450s from the GC (71.38%) rich genomic DNA of *Sorangium cellulosum* So ce56 was achieved by using several different combinations of reaction cocktail and thermal programs which is illustrated in Table I. 3-5. All P450 expression constructs were created in the pCWori⁺ vector; except for CYP264B1 the expression vector pET17b was used. The expression constructs were preliminary sorted by colony PCR, restriction digestion, and the full frame of each ORF was confirmed by sequencing.

4. 3. 2. Heterologous expression of cytochrome P450

The yields of the heterologous expression of the various P450s of So ce56 with respect to time, temperature and shaking speed were not alike. Time course studies were done to determine the optimum heterologous expression for individual P450s, considering pH of the culture medium (pH 6.8-7.2) as an indicator, followed by the visibly observed confluent reddish brown color of the cell pellet. Depending on the P450, the expression time was between 24 to 36 hours. Each of the expression trails for the respective P450 was monitored by whole-cell SDS-PAGE. The crude cell supernatant was also used to measure the COD for the confirmation of active form of P450. All P450s were heterologously expressed in a soluble-form. As a model example, the whole cell SDS-PAGE for CYP109D1 is illustrated in Figure I. 11. All 21 P450s from So ce56 were successfully purified in an active form by affinity chromatography and gel-chromatography. The representative SDS-PAGE of purified So ce56 P450s is shown in Figure I. 12. The concentrations of purified P450s were variable and the maximum amount was up to 1400 nmol/L for CYP266A1. The details of the P450 expression and characterization are illustrated in Table I. 9.

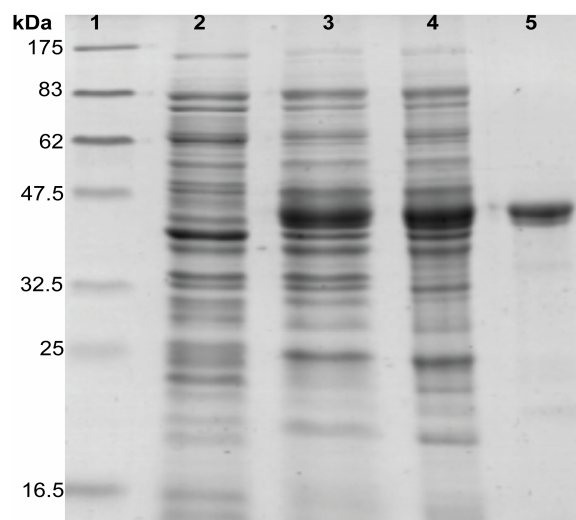


Figure I. 11. SDS-PAGE of CYP109D1. The expression and purification of CYP109D1 during heterologous expression in *E. coli* in soluble form are shown. Lane 1- Prestained marker (BioLab), Lane 2- cell lysate before induction (-IPTG), Lane 3- cell lysate after 36 hours of expression, Lane 4- cell free supernatant after sonication, Lane 5- purified CYP109D1.

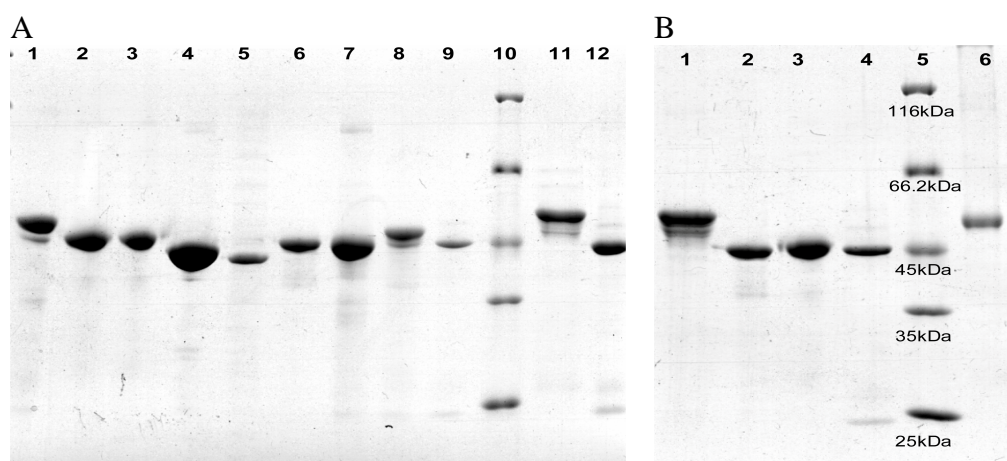


Figure I. 12: SDS-PAGE of purified So ce56 P450s. The designated lanes denote the individual purified P450s. A. Lane 1- CYP260A1, lane 2- CYP266A1, lane 3- CYP124E1, lane 4- CYP265A1, lane 5- CYP264A1, lane 6- CYP109C1, lane 7- CYP109C2, lane 8- CYP259A1, lane 9- CYP109D1, lane 10- protein MW marker (pEQLab), lane 11- CYP263A1, lane 12- CYP267B1, B. lane 1- CYP110J1, lane 2- CYP260B1, lane 3- CYP264B1, lane 4- 267A1, lane 5- protein MW marker (pEQLab), lane 6- CYP117B1.

Table I. 9. Expression yield and characterization of So ce56 cytochromes P450.

Individual cytochromes P450 were purified by affinity chromatography followed by size exclusion chromatography and the expression yields of the purified P450s were tabulated. The absorption maxima of each of the P450s are also shown.

P450s of So ce56	number of amino acids	P450 conc. nmol/L	Absorption maxima
CYP109C1	405	1333	450
CYP109C2	397	1300	448
CYP109D1	408	205	450
CYP110H1	444	5	450
CYP110J1	477	60	448
CYP117B1	459	5	448
CYP124E1	402	15	450
CYP259A1	442	8	450
CYP260A1	444	850	448
CYP260B1	392	100	448
CYP261A1	482	5	448
CYP261B1	493	2	450
CYP262A1	435	25	448
CYP262B1	470	2	450
CYP263A1	493	900	450
CYP264A1	390	440	448
CYP264B1	412	4	450
CYP265A1	398	21	450
CYP266A1	409	1400	448
CYP267A1	429	23	450
CYP267B1	405	70	450

4. 4. Spectrophotometric characterization

UV-visible absorption spectroscopy provides the primary technique for the recognition and characterization of P450 enzymes. The oxidized form of purified P450s of So ce56 followed typical spectral properties of the P450 enzyme class. The major Soret (γ) band of all 21 P450s coincided at 415-422 nm and the smaller α band at 532-537 nm and β band at 563-574 nm.

The reduced spectra, of all the P450s, showed diminished absorption maxima at Soret band with or without a slight shift of the peak and gave a single peak at 537-556 instead of the smaller α and β bands (Table 1. 10). The representative spectra of oxidized, reduced and the CO-complexed spectra for CYP109D1 are shown in Figure I. 13. A. The carbon monoxide bound form gave a typical peak at 448 or 450 depending on the individual P450s (Figure I. 13. B). Some of the P450s, CYP259A1, CYP260A1, CYP261A1, CYP264B1 and CYP265A1, gave a dominating P420 shoulder which was increased during time interval. The representative spectra for CYP264B1 are shown in Figure I. 13. C. All other CO difference spectras are illustrated in Figure I. 14.

Table I. 10: Optical characteristics of *So ce56* cytochromes P450. Soret maxima, α and β peaks during oxidation and reduction as well as CO complexed peak for all 21 *Sorangium cellulosum* *So ce56* are shown.

So ce56 P450s	Oxidized form			Reduced form		CO- difference (COD) Peak (nm)
	Soret (nm)	α (nm)	β (nm)	Soret (nm)	single peak at α and β regions (nm)	
CYP109C1	418	535	563	417	540	448
CYP109C2	417	533	566	417	535	448
CYP109D1	420	535	565	417	546	450
CYP110H1	418	534	566	418	554	450
CYP110J1	418	532	565	416	538	448
CYP117B1	416	534	568	416	552	448
CYP124E1	422	537	574	422	537	448
CYP259A1	420	535	565	418	550	452
CYP260A1	416	534	566	412	546	448
CYP260B1	418	535	560	416	548	448
CYP261A1	420	536	565	420	550	448
CYP261B1	420	534	570	420	DA*	450
CYP262A1	417	534	566	417	556	448
CYP262B1	418	534	568	416	552	450
CYP263A1	418	535	560	420	DA*	450
CYP264A1	415	535	566	415	DA*	450
CYP264B1	417	532	568	417	DA*	450
CYP265A1	419	535	568	415	552	448
CYP266A1	415	535	567	407	540	448
CYP267A1	418	535	570	417	548	448
CYP267B1	420	534	565	416	548	450

Note: DA*: Single peak at α and β regions was not obtained. Same absorption peaks as that of oxidized form at α and β bands were observed with diminished absorbance (DA).

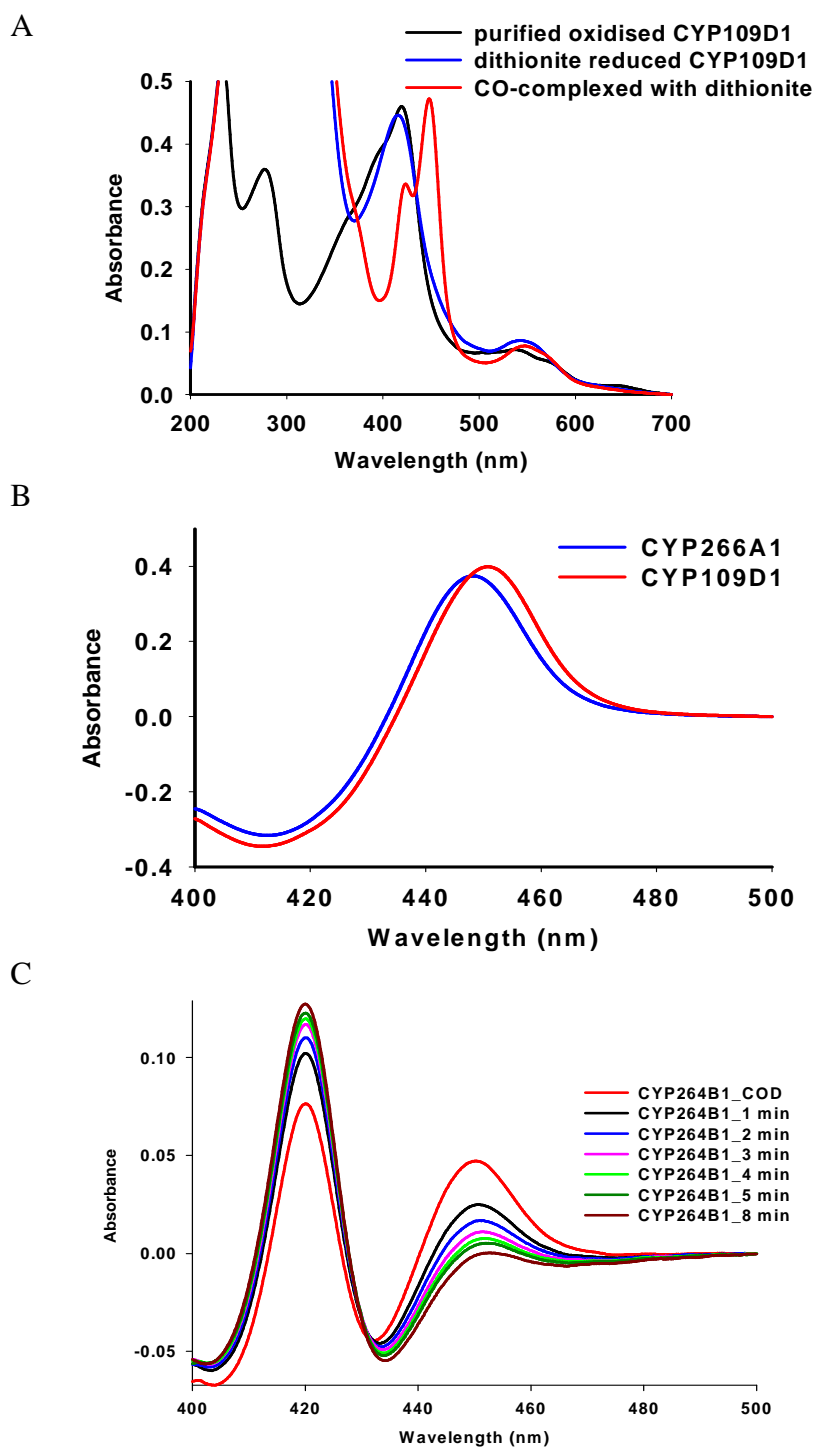
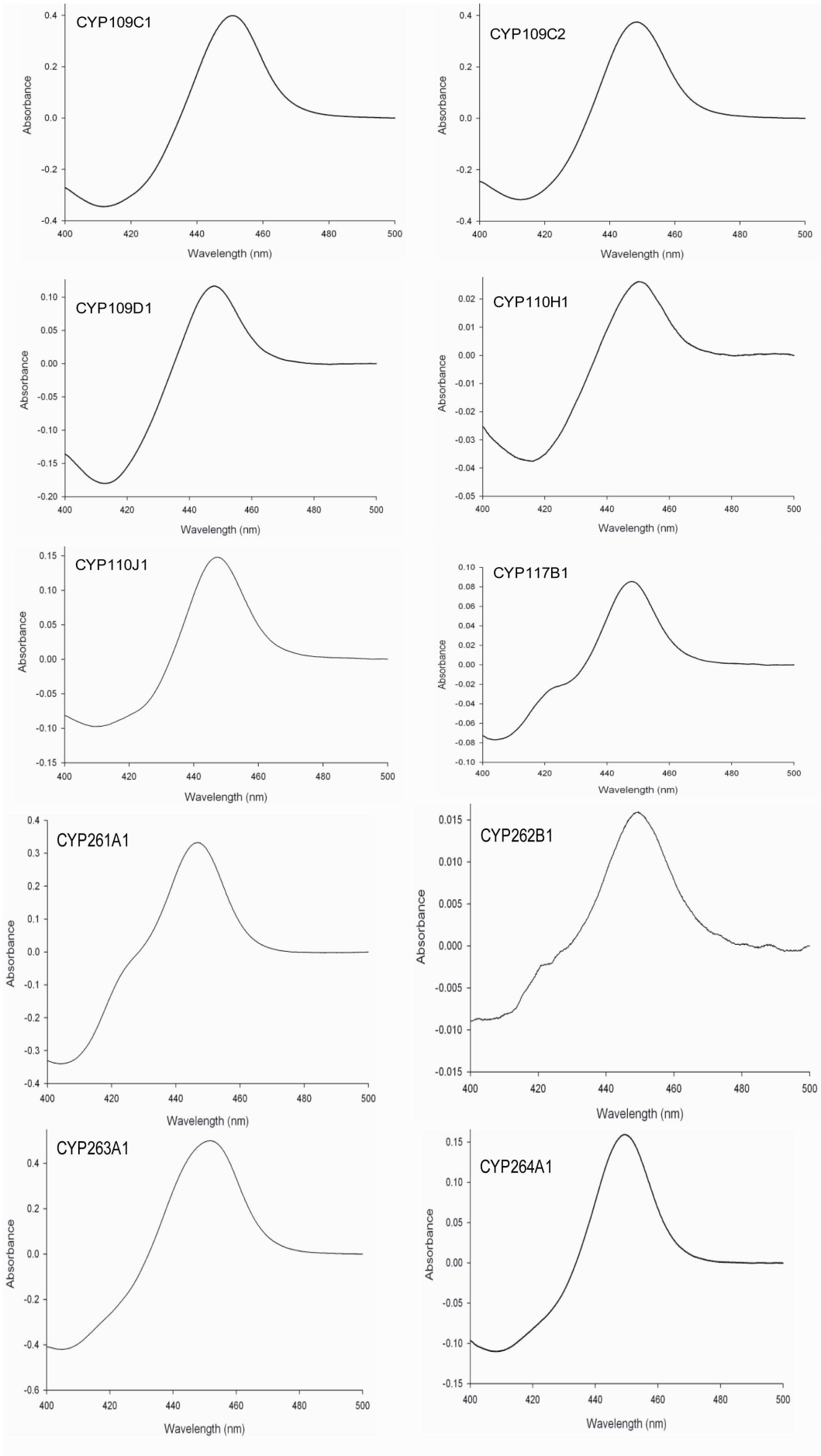


Figure I. 13: Representative UV-visible characterization of So ce56 P450s. A. Typical UV-visible spectra of purified CYP109D1. The oxidized (black line), dithionite reduced (blue line) and reduced CO-complexed from (red line) of CYP109D1 are illustrated. **B. The typical COD spectra of CYP109D1 and CYP266A1.** CYP109D1 (red line) and CYP266A1 (blue line) have peak maxima at 448 and 450 nm, respectively. **C. CO-difference spectra of CYP264B1.** The unstable P450 (CYP264B1), in absence of substrate, is shown with multiple spectra. The time dependent increment of the peak at 420 nm and a collapsing peak at 450 nm are shown. P420 form of CYP264B1 during the measurement of CO-difference spectrum (red line) is shown to be increased after 1 min (black line), 2 min (blue line), 3 min (pink line), 4 min (light green) and 5 min (dark green line). The P450 peak was almost collapsed at 8 min (brown line).



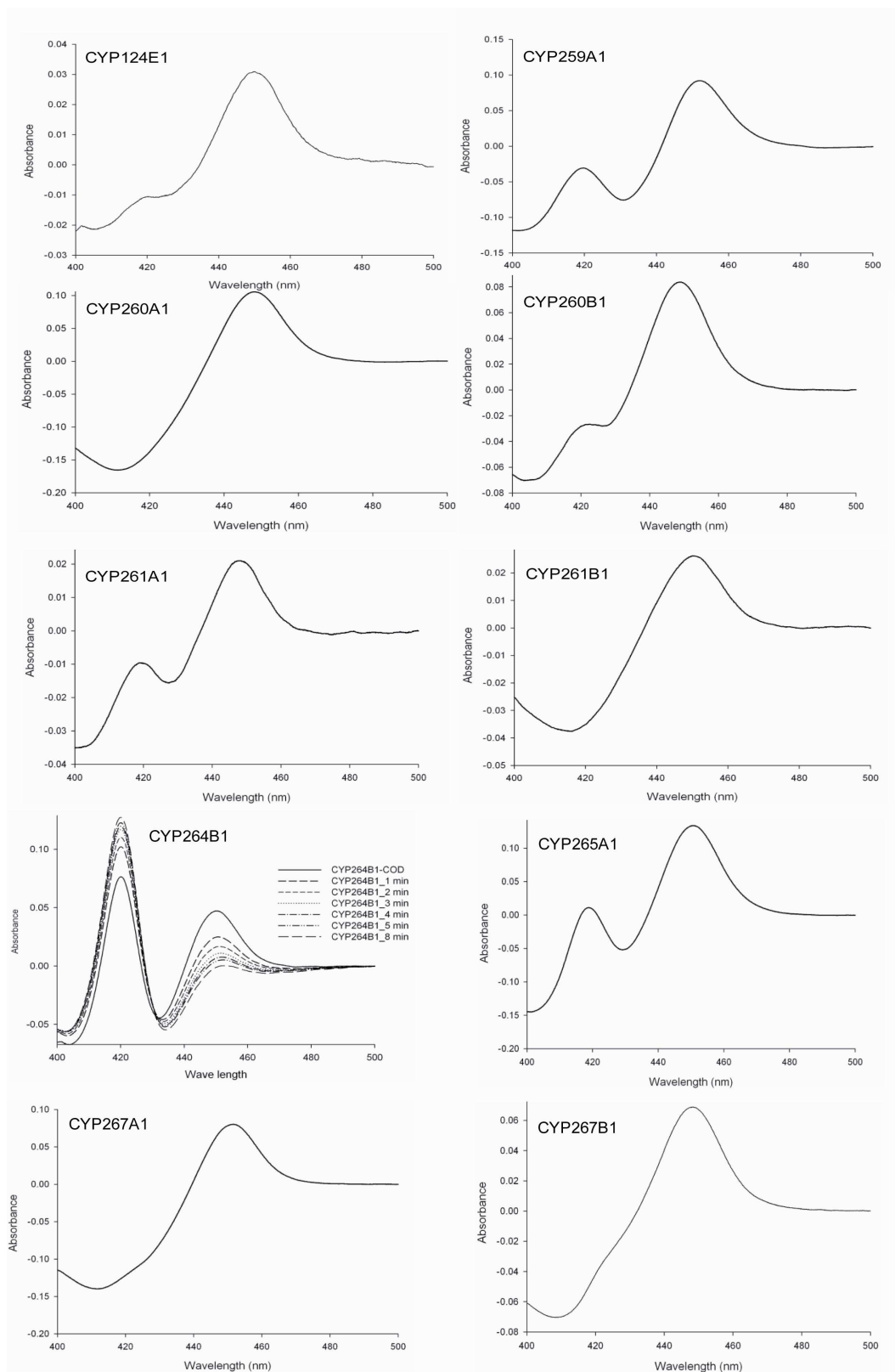


Figure. I. 14. CO-difference spectra of *Sorangium cellulosum* So ce56 P450s. Respective P450s are indicated in each spectrum. The unstable P450 (CYP264B1) is represented with multiple spectra with time dependent increment of the peak at 420 nm and a collapsing peak at 450 nm

4. 5. Biophysical characterization

Seven of the highly expressed, stable and the representative of the novel P450 enzymes (CYP109C1, CYP109C2, CYP109D1, CYP260A1, CYP264A1, CYP264B1 and CYP266A1) were selected for detail biophysical characterization. Circular Dichroism (CD) and Electron paramagnetic resonance (EPR) spectroscopy were recorded for them.

4. 5. 1. Circular Dichroism (CD) of So ce56 cytochrome P450

In the CD measurement, the spectra in the far-UV region (typically from 260 nm to 195 nm) and near-UV/visible region (typically from 300 nm to 500 nm) were measured. The representative far-UV spectra of CYP109C1, CYP109C2, CYP109D1, CYP260A1, CYP264A1, CYP264B1 and CYP266A1 are illustrated in Figure I. 15. The far-UV CD spectra for these P450s virtually look similar with a negative dichroic double band with minima at 208 and 220 nm. This indicates the typical α -helical protein characteristics and shows no significant change in the secondary structure. The trough at 208 nm corresponds to a weak but broad $n \rightarrow \pi^*$ transition (Munro *et al.*, 1999 and Kelly *et al.*, 2005).

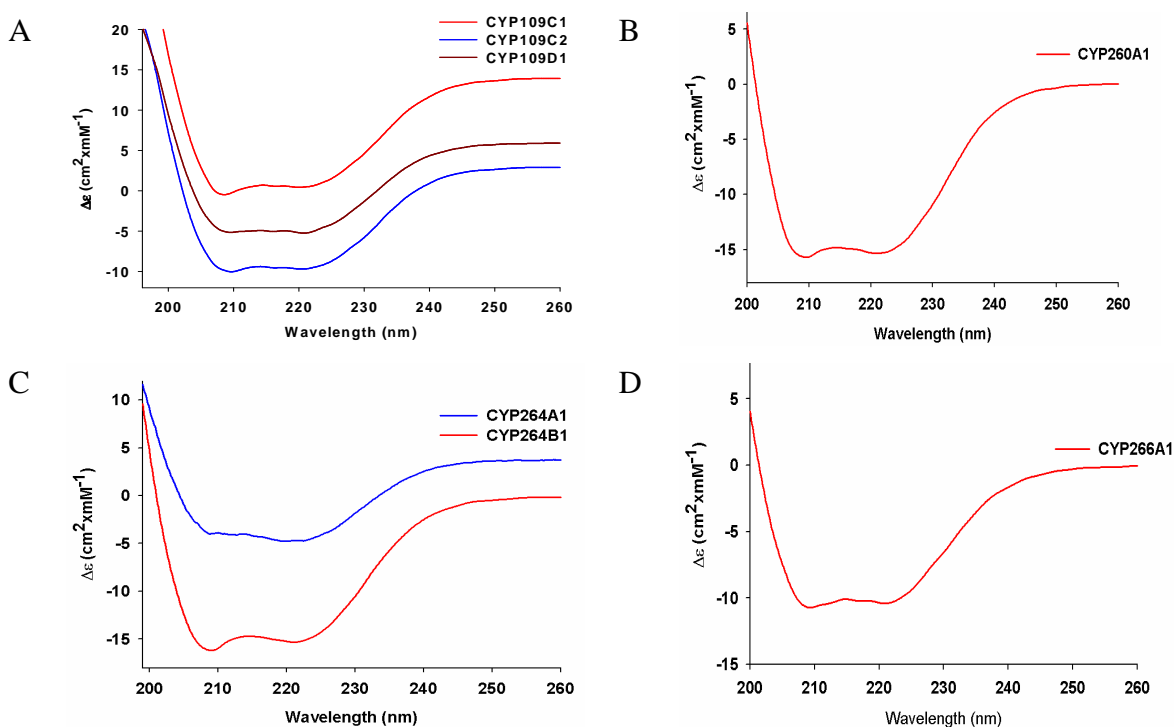
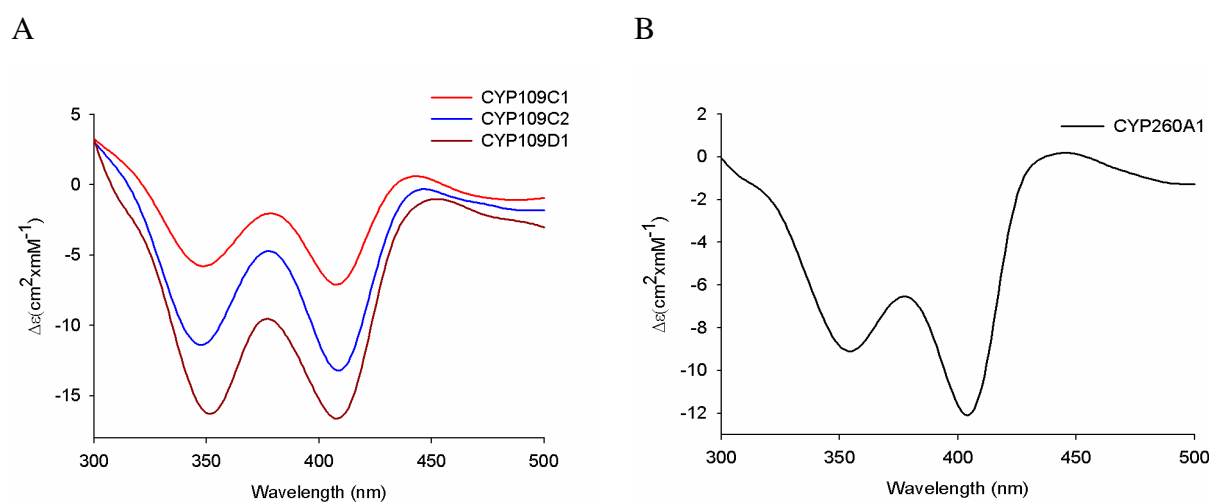


Figure I. 15: CD spectra of P450s in far-UV region. CD spectra recorded in the far-UV region are shown. The individual P450 concentration was 5 μ M in 10 mM potassium phosphate buffer, pH 7.4. The spectra were recorded in the absence of substrate at 20°C. A. The comparative CD spectra of CYP109C1 (red line), CYP109C2 (blue line) and CYP109D1 (brown line). B. CD spectra of CYP260A1. C. The comparative CD spectra of CYP264A1 (blue line) and CYP264B1 (red line). D. CD spectra of CYP266A1.

The near-UV/visible CD spectra provide information on the heme cofactor environment as well as on the tertiary structure of the protein. The spectrum in the near-UV region arises predominantly from the aromatic amino acids (phenylalanine, tryptophan and tyrosine) along with possible contributions from disulfide bonds. Aromatic CD spectral intensities are affected by protein flexibility, with the more highly mobile side-chains having lower CD intensities. In addition, interactions between aromatic side-chains in close proximity contribute to stronger signal intensity. Although the isolation of a heme chromophore exhibits no optical activity, because of the high symmetry, when bound in an asymmetric protein environment optical activity in the visible region may arise (Myer and Pande, 1978).

The representative CD spectra for near-UV region of typical P450s (CYP109C1, CYP109C2, CYP109D1, CYP260A1, CYP264A1, CYP264B1 and CYP266A1) of *So ce56* are shown in Figure I. 16. All tested P450s have displayed two large signals of negative peaks both at the delta and the Soret band. The near-UV CD spectra of the CYP109 family (CYP109C1, CYP109C2 and CYP109D1) showed similar delta band (peak at 350 nm), Soret band (peak at 408 nm) and environment of the heme group (350-450 nm). Both of the new P450 families of *So ce56*, CYP260A1 and CYP266A1, were characterized with distinct transitions at 350 and 403 nm, delta and Soret bands, respectively. In the case of CYP264A1 and CYP264B1, the delta bands centered at 348 and 350 nm, respectively, and the Soret bands were at 408 and 420 nm, respectively. It was found that the environment of the heme group (350-450 nm) of CYP264A1 in resting ferric form is slightly different with respect to CYP264B1.



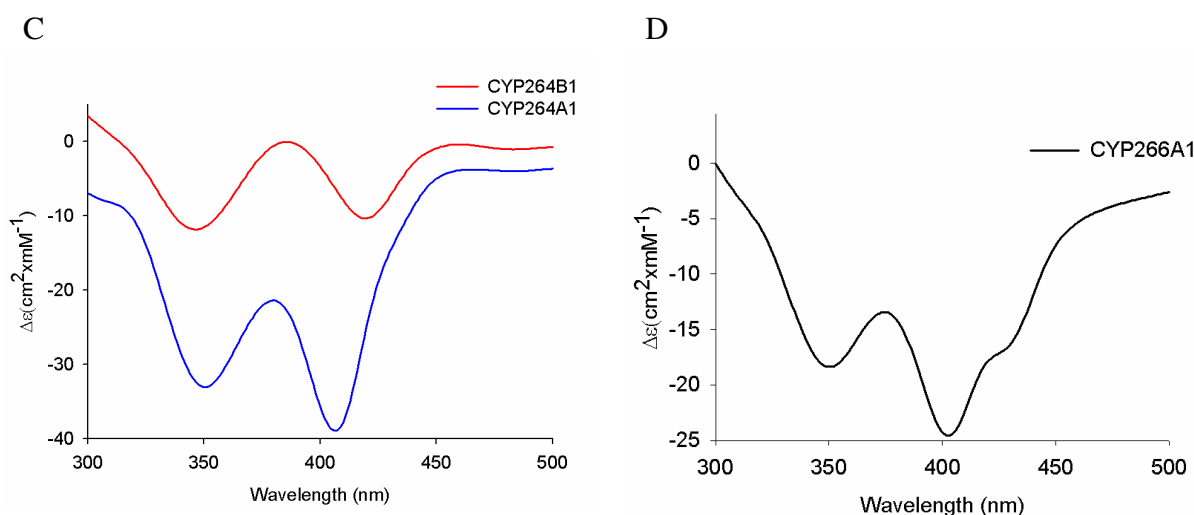


Figure I. 16: CD spectra of P450s in near-UV region. CD spectra recorded in the near-UV region are shown. The individual P450 concentration was 20 μM in 10 mM potassium phosphate buffer, pH 7.4. The spectra were recorded in absence of substrate at 20°C. A. The comparative CD spectra of CYP109C1 (red line), CYP109C2 (blue line) and CYP109D1 (brown line). B. CD spectra of CYP260A1. C. The comparative CD spectra of CYP264A1 (blue line) and CYP264B1 (red line). D. CD spectra of CYP266A1.

4. 5. 2. Electron paramagnetic resonance (EPR) spectroscopy

The EPR measurements of the cytochrome P450 are of particular interest. The EPR signal cannot be detected for the low-spin ferrous form, as there are no unpaired electrons. Therefore, the low-spin cytochrome P450 must be in the oxidized ferric state to exhibit an EPR spectrum. However, in the high-spin form, the unpaired electrons exist in both the reduced (ferrous) and oxidized (ferric) states, and can generate EPR signals in both states. Discrete EPR signals are produced for different spin-states and conformations of heme. Thus, obtained g -values differ depending on the ligation and environment of the heme iron (Hoff *et al.*, 1989; Mansuy and Renaud, 1995). The two most common types of spectra for iron-containing cytochromes are axial high spin and rhombic low spin. Axial high-spin systems have characteristic g -values of ~ 2 and ~ 8 , whereas rhombic low-spin systems have three distinct g -values between 1.0 and 3.8.

The X-band EPR spectra of the substrate free oxidized forms of the members of CYP109, CYP260, CYP264 and CYP266 of *So ce56* were examined. The major signals generated in the EPR spectrum of those P450s constitute a rhombic trio of g -tensor elements g_z , g_y , g_x (Figure I. 17-18). The occurrence of the three major peaks indicates a nuclear spin state of $\frac{1}{2}$ ($2I + 1$). The resulting g -values of g_z , g_y and g_x (Table I. 11) illustrated the typical low-spin ferric (Fe^{III}) heme (Mansuy and Renaud, 1995).

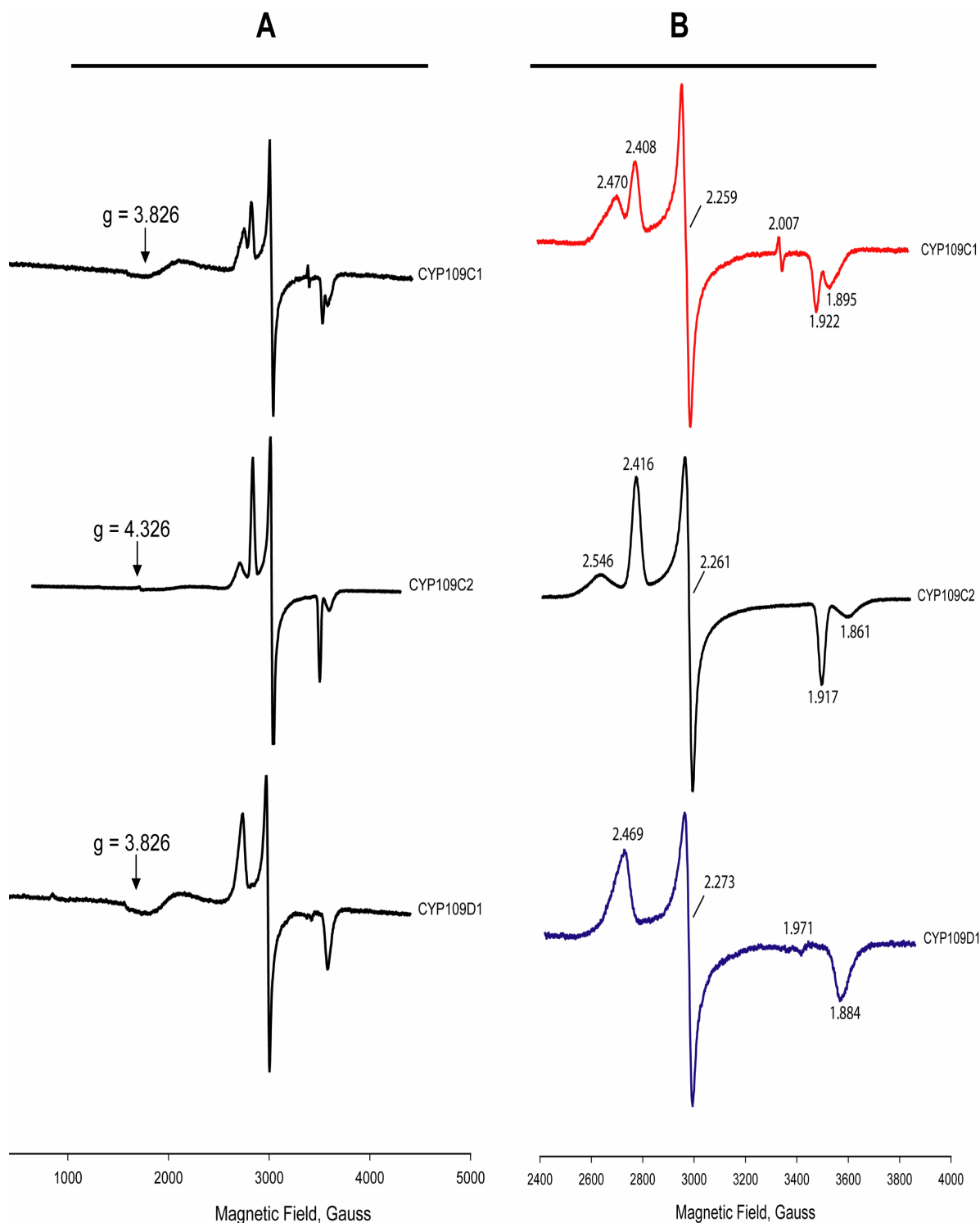


Figure I. 17: X-band EPR spectra of the CYP109 family (CYP109C1, CYP109C2 and CYP109D1). The individual P450 concentration was 150 μ M in 10 mM potassium phosphate buffer, pH 7.4. Substrate free P450s were used and the spectra at broad and narrow ranged were scanned. The g-values are indicated in each spectrum. On the panel A (the left side), the EPR spectra (broad range) for the CYP109 family are shown. It was measured with a field central of 2400 G and a sweep width of 4000 G. On the panel B (the right side), the expanded narrow range of the EPR spectra of panel A with a field central of 3140 G and a sweep width of 1440 G are shown for better resolution. In panel B, CYP109C1 (red line), CYP109C2 (black line) and CYP109D1 (blue line) are shown.

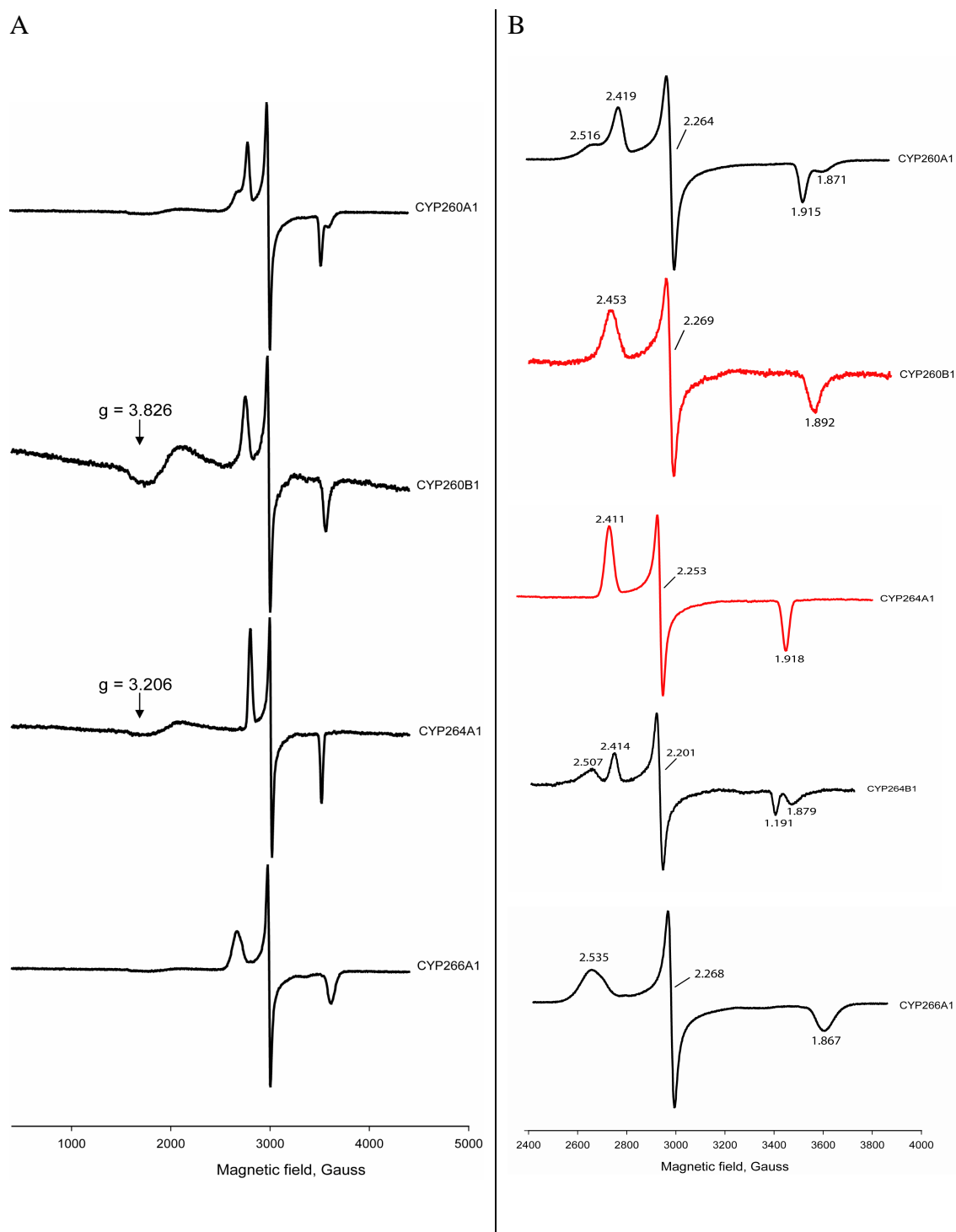


Figure I. 18: X-band EPR spectra of cytochrome P450s (CYP260A1, CYP260B1, CYP264A1, CYP264B1 and CYP266A1) of *So ce56*. The individual P450 concentration was 150 μ M in 10 mM potassium phosphate buffer, pH 7.4. Substrate free P450s were used and the spectra of broad and narrow range were scanned. The g-values are indicated in each spectrum. The respective P450s are shown in each spectrum. A. The EPR spectra (broad range) were measured with a field central of 2400 G and a sweep width of 4000 G. B. The EPR spectra (narrow range) of A with a field central of 3140 G and a sweep width of 1440 G are shown for better resolution. In panel B, CYP260A1 (black line), CYP260B1 (red line), CYP264A1 (red line), CYP264B1 (black line) and CYP266A1 (black line) are shown with respective P450 names.

Table I. 11: EPR spectral values of representative P450s of *So ce56*.

P450s of <i>So ce56</i>	g_z	g_y	g_x
CYP109C1	2.470/2.408	2.259	1.922
CYP109C2	2.546/2.411	2.261	1.917/1.864
CYP109D1	2.469	2.270	1.897
CYP260A1	2.516/2.419	2.264	1.191/1.871
CYP260B1	2.453	2.269	1.892
CYP264A1	2.411	2.250	1.918
CYP264B1	2.507/2.411	2.201	1.911/1.879
CYP266A1	2.535	2.268	1.867

Firstly, a broad range EPR spectrum was recorded to check the presence of any high spin form. After this the narrow range was scanned for better resolution for g_z , g_y and g_x . In this study, the signals at $g = 3.826$ and 3.206 were observed for CYP109C1, CYP109D1, CYP260B1 and CYP264A1 (Figure 27 and 28. A). This heterogeneity was supposed due to the purity of the sample (McLean *et al.*, 2002a). In case of CYP109C2 minor splitting of the g -value at $g = 4.325$ was observed and indicates some heterogeneity of the heme species and could be due to the trace amounts of high spin ferric heme and adventitious Fe^{III} or may even indicate a very low degree or rhombic splitting of the $S = 5/2$ high ferric state (Lipscomb, 1980; Dawson, 1983). In this study only the oxidized form of substrate free EPR was recorded.

The heterogeneity of low-spin heme was observed for CYP109C1, CYP109C2, CYP260A1 and CYP264B1 at g_z and g_x (Table I. 11). The signal shifts at g_z were also reported in the 4-phenylimidazole-bound form of P450cam (2.45/2.5, 2.25. 1.89) (Uchida *et al.*, 2004). The signals at g_z 2.43/2.41 and g_x 1.92 indicate a proportion of the sample retains a distal water ligand (Girvan *et al.*, 2004; Lawson *et al.*, 2004). In previous studies, some heterogeneity of the low-spin heme g_x signal has also been obtained in P450cam with signals at both $g = 1.970$ and 1.191 (Green *et al.*, 2001; Lawson *et al.*, 2004).

4. 6. Mass determination and conformation of active form of P450 of *So ce56*

The presence of the target band of corresponding P450s in correlation with the theoretical molecular weight in SDS-PAGE confirmed the successful expression. Sometimes the molecular weight obtained from SDS-PAGE was not identical with the theoretical values.

Seven of the P450s (CYP109C1, CYP109C2, CYP109D1, CYP260A1, CYP264A1, CYP264B1 and CYP266A1) were chosen to determine the masses using ESI-TOF for both the apo- and holo-forms. The experimental and theoretical masses are illustrated in Table I. 12-13 and corresponding spectras are shown in Appendix VI. A. The theoretical and the experimental masses of apo- and holo-forms as well as the measured cofactor were nearly identical.

Table I. 12: Comparison of theoretical and experimental apo-protein mass and relative spectra of different cytochrome P450s.

P450 of So ce56	Apo-protein theoretical MW (Da)*	Apo-protein measured mass (Da)	Spectrum (Appendix VI. A)
CYP109C1	45802.64	45799.54 ± 0.44	O7380AM
CYP109C2	44924.15	44923.43 ± 0.45	O7379AM
CYP109D1	45213.81	45213.84 ± 0.52	O7382AM
CYP260A1	49746.75	49616.03 ± 0.65	O7392AM
CYP264A1	43337.64	43205.66 ± 0.43	O7377AM
CYP264B1	47414.64	47281.95 ± 0.38	O7374AM
CYP266A1	46673.87	46672.66 ± 0.45	O7373AM

Note: * The calculated theoretical apo-protein molecular weight includes the hexa-histidine tag attached on the C-terminal of the P450s during cloning procedure.

Table I. 13: Comparison of theoretical and experimental masses of holo-protein.

P450 of So ce56	Holo-protein theoretical MW (Da)*	Holo -protein measured mass (Da)	Cofactor theoretical MW (Da)	Cofactor measured mass (Da)	Spectrum (Appendix VI. B)
CYP109C1	46419.13	46409.77 ± 0.98	616.49	610.23	O7412AM
CYP109C2	45540.64	45532.59 ± 0.43	616.49	609.16	O7410AM
CYP109D1	45830.30	45823.14 ± 0.47	616.49	609.30	O7415AM
CYP260A1	50363.24	50223.46 ± 0.58	616.49	607.43	O7418AM
CYP264A1	43954.12	43815.30 ± 0.63	616.49	609.64	O7407AM
CYP264B1	48031.13	47891.57 ± 0.32	616.49	609.62	O7405AM
CYP266A1	47290.36	47282.04 ± 0.37	616.49	609.38	O7402AM

Note: * The calculated theoretical holo-protein molecular weight includes the hexa-histidine tag attached on the C-terminal of the P450s during cloning procedure.

4. 7. Determination of electron donors

The genomic information of *So ce56* has shown that none of the putative genes encoding the ferredoxins and ferredoxin reductases are in an adjacent position. As a result, in a plethora of several ferredoxins and ferredoxin reductases, it was difficult to predict particular electron donor partners for specific P450s for distinct physiological purpose. But it has been shown that bacterial P450s are known to use a variety of electron transfer mechanisms (Hannemann *et al.*, 2007). Mostly in bacterial systems, electrons move from NAD(P)H to flavoproteins and the P450s accept electrons either from flavoproteins or ferredoxin (Fdx) proteins (which receive the electrons from ferredoxin reductase (Fdr) (Katagiri *et al.*, 1968; Narhi *et al.*, 1986).

To determine the redox partners, at first CO-difference spectra with the reduction of dithionite was recorded. Thus obtained peak maximum at 450 nm was compared with the NADPH reduced CO-complexed peak obtained by the reduction of dithionite in the presence of redox partners. Considering this, the combination of commercial spinach Fdx and FdR as heterologous electron transfer partners, a strategy that is often used in bacterial systems, has been followed with myxobacterial P450s. None of the tested P450s (CYP109C1, CYP109C2, CYP109D1, CYP110J1, CYP124E1, CYP260A1, CYP264A1, CYP264B1, CYP266A1 and CYP267B1) was able to get reduced with the spinach system. The other combination of heterologous electron transfer partners of bovine adrenal adrenodoxin (Adx) and adrenodoxin reductase (AdR) were also used in the ratio of P450s:Adx:AdR of 1:10:3. Among those tested P450s, except CYP266A1, all others were able to reduce by NADPH and gave a peak maximum at 450 nm during complexed with carbon monoxide in the presence of redox partners. During this, for CYP264A1 the dithionite reduced CO-difference peak was fully recovered with the reduction of NADPH within short interval of time. Likely, CYP109D1 was also able to recover almost 75% of the dithionite reduced peak (Figure I. 19) when NADPH was used as reducing agent in presence of Adx and AdR.

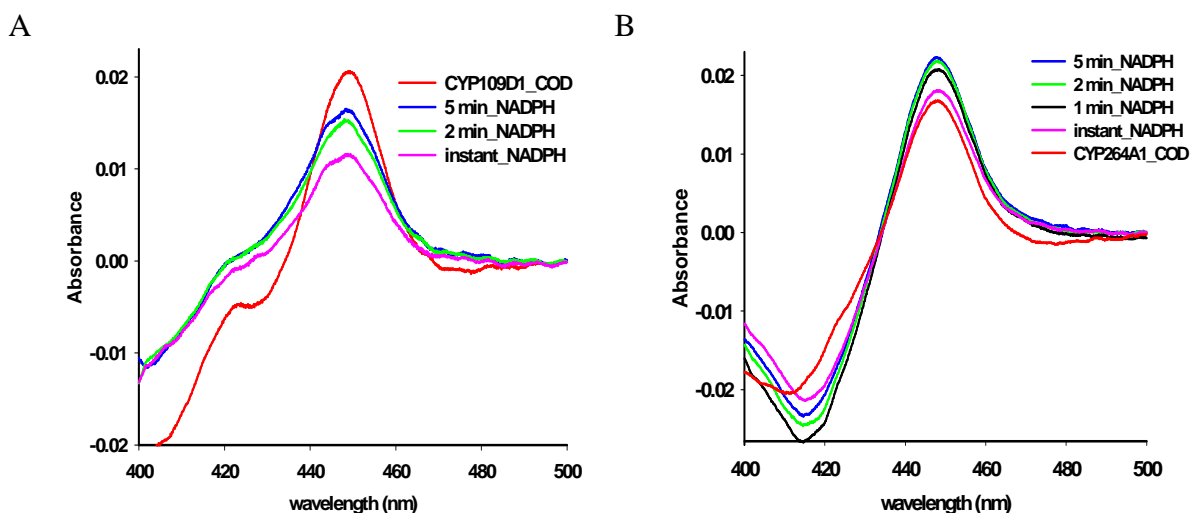


Figure I. 19: Determination of electron transfer partners for CYP109D1 and CYP264A1. The dithionite reduced CO-difference (COD) spectrum (red line) was compared with the NADPH reduced CO-difference spectrum of instant (pink line), 1 min (black line), 2 min (green line), 5 min (blue line) for A. CYP109D1 and B. CYP264A1. The NADPH reduced CO-difference spectra were recorded in a 1 ml mixture of P450:Adx:AdR of 1:10:3 (200 pmol:2 nmol:600 pmol) in 10 mM potassium phosphate buffer, pH 7.4 containing 20% glycerol.

4. 8. Screening of a compound library

The potential physiological and biotechnological applications of P450 cannot be illustrated unless a potential substrate becomes known. In general, the identification of potential substrates for novel P450s is not easy. To explore the substrate range, four of the P450s (CYP109C1, CYP109C2, CYP260A1 and CYP266A1) were selected because of their higher expression yield and stability (Table I. 10). A compound library consisting of 17,000 substances was screened for those P450s and the potential substrates and inhibitors were identified. Each of the primary screening results was validated as described in the ‘Method 3.2.8’. According to the validated results of the positive hits, Type I and Type II binding ligands were identified. The numbers of the hits for CYP109C1, CYP109C2, CYP260A1 and CYP266A1 are tabulated in Table I. 14 and their structures with respective Chem-Div numbers are illustrated in Appendix VII. During the screening it was found that some of the Type I and Type II hits were common among the screened P450s and this is also illustrated in Appendix VII.

Table I. 14: Total numbers of Type I and Type II hits obtained from the screening of a compound-library for CYP109C1, CYP109C2, CYP260A1 and CYP266A1. The purified P450s of 900 ml of 3 μ M each was required to screen a compound-library consisting of ~17,000 substances. During primary screening, individual P450 of 1.5 μ M and each of the compounds of 0.5 μ M was used, and screened in 384 wells micro-titer plate by an automated pipette robot cyclone ALH 3000 Workstation as described in Methods I. 3. 2. 8. The wavelength of 350-450 nm with Tecan-Safire plate-photometer and the absorbance of the substances were recorded. The primary hits were further validated by 200 μ M, 150 μ M, 100 μ M, 50 μ M and 16 μ M of individual compounds from the library with 1.5 μ M final concentration of P450. The number represents the counts of the validated hits for the P450s.

P450 of So ce56	Numbers of Type I hits	Numbers of Type II hits	Total hits	Remarks (For detail structure and Chem-Div nos)
CYP109C1	16	9	25	Appendix VII. A
CYP109C2	8	18	26	Appendix VII. B
CYP260A1	17	20	37	Appendix VII. C
CYP266A1	16	21	37	Appendix VII. D

5. Discussion

Myxobacteria attract many researchers because of their pharmaceutical importance and complex physiology. They exhibit a multitude of possibilities for the production of bioactive substances. It was shown that about 100 novel basic compounds and 500 different variants have been isolated from these organisms (Reichenbach, 2001, 1999; Gerth *et al.*, 2003). The genus *Sorangium* alone can produce nearly 50% of the metabolites isolated from myxobacteria (Gerth *et al.*, 2003). Different strains of *Sorangium cellulosum* produce several novel antimicrobial macrolides, the thuggacins (Irschik *et al.*, 2007), phoxalone (Guo *et al.*, 2008) and other components like a new sesquiterpene, sorangiodenosine (Ahn *et al.*, 2008), a new free-radical scavenger, soraphinol C (Li *et al.*, 2008), and a novel class of antineoplastic agents - the epothilones and their analogs (Mulzer, 2008). The selected strain So ce56 produces the natural secondary metabolites chivosazol, etnangien and myxochelin (Schneiker *et al.*, 2007; Wenzel and Müller, 2007; Menche *et al.*, 2008). The genomic sequence of So ce56 also revealed several additional regions encoding for polyketide synthase (PKS) and/or non ribosomal polyketide synthase (NRPS) proteins, whose products have not been isolated till date (Schneiker *et al.*, 2007). The complex life cycle of So ce56 is mediated by an extensive regulatory network, including enhancer binding proteins, two component regulatory systems, extra cytoplasmic function family protein sigma factors, and serine/threonine/tyrosine protein kinases (Perez *et al.*, 2008).

5. 1. Bioinformatic analysis

The genome sequence project of the *Sorangium cellulosum* So ce 56 revealed 21 P450s with amino acids ranges from 390 to 493 in length (Table I. 10). This added 9 novel cytochrome P450 families (CYP259, CYP260, CYP261, CYP262, CYP263, CYP264, CYP265, CYP266 and CYP267) into the P450 database. Furthermore, 7 new members of existing P450 families of CYP109 (CYP109C1, CYP109C2 and CYP109D1), CYP110 (CYP110H1 and CYP110J1), CYP117 (CYP117B1) and CYP124 (CYP124E1) were also disclosed. The presence of such higher numbers of P450s in the So ce56 genome underlines the potential significance during its lifecycle.

The phylogenetic relationship of P450s of So ce56 shows 5 distinct P450 branches (Figure I. 3). The cluster 1: CYP266A1, cluster 2: CYP267 (CYP267A1 and CYP267B1), cluster 3: CYP265A1, cluster 4: CYP109 (CYP109C1, CYP109C2 and CYP109D1), CYP124E1 and CYP264 (CYP264A1 and CYP264B1), and cluster 5: CYP110 (CYP110H1 and CYP110J1),

CYP117B1, CYP259A1, CYP260 (CYP260A1 and CYP260B1), CYP261 (CYP261A1 and CYP261B1), CYP262 (CYP262A1 and CYP262B1), and CYP263A1. Most interestingly, the cluster-1, CYP266A1, represents the most distant branch having 30% identity with human cytochrome P450 CYP5A1 (Thromboxane synthase), 30% identity with fungal P450 CYP55A3 (*Cylindrocarpon tonkinense* D78512) and 36% with bacterial CYP107H1 of *Bacillus subtilis* BioI US1868. The functional role of this P450 could be interesting for further studies.

All 21 P450s of *So ce56* contain conserved K- helices (having glutamine and arginine), I- helices (having threonine) and the heme binding motif. A variant of the heme binding motif was observed in CYP267A1 and CYP265A1, where leucine is present instead of phenylalanine (Table I. 7). The presence of leucine in the heme domain of CYP267A1 and CYP265A1 could be the topic for further investigations of heme signature variants if the substrates and redox partners are known.

The bioinformatic analysis of *So ce56* also revealed 8 ferredoxin genes and 2 reductases. In the *So ce56* genome neither ferredoxin nor ferredoxin reductase genes were organized in a cluster with P450. It was shown that one P450 from *S. coelicolor* A(3)2 and five P450s from *S. avermitilis* were found to be in a cluster with a ferredoxin (Parajuli *et al.*, 2004). Likewise, it has been shown that none of the ferredoxin reductase was directly clustered with P450 in *S. avermitilis* and *S. coelicolor* A(3)2, however CYP147B1 and CYP105Q1 from *S. avermitilis* were associated with ferredoxin followed by ferredoxin reductase gene (Parajuli *et al.*, 2004). In comparison with *So ce56*, the fully sequenced 9 Mb genome of *Streptomyces avermitilis* has 33 P450s with 9 Fdx and 6 Fdr genes (Tully *et al.*, 1998). Likely, *Streptomyces coelicolor* A(3)2 has 18 P450s with 6 Fdx genes and 4 Fdr reductases in a 8.7 Mb chromosome (Lamb *et al.*, 2002). Only one natural electron transfer chain was determined in *S. coelicolor* A(3)2, where the genes coding for ferredoxin reductase FDR1, ferredoxin Fdx4 and CYP105D5 were located in close distance in the genome and showed a functional role during fatty acid hydroxylation (Chun *et al.*, 2007). The presence of 21 P450s, 8 Fdx and 2 Fdr in *So ce56* symbolizes the complex and ambiguous electron transfer pathways.

5. 2. Genome mining of cytochrome P450

As this is a very first study of P450s in *So ce56*, at the beginning of this work only gene sequences were available. Though various endogenous roles of P450s in several polyketide

synthetase (PKS) gene cluster during secondary metabolite formation were already verified in other strain of *Sorangium cellulosum* (Tang *et al.*, 2000; Julien *et al.*, 2000; Buntin *et al.*, 2008), the 13.1 Mbp genome of *Sorangium cellulosum* So ce56 was a new candidate. At the beginning, neither there was an idea of potential substrates for any of the P450s nor were the homologous or heterologous redox partners known. As a result, the study of the genomic organization of individual P450 at the beginning was the only way to hypothesize the potential redox partners or substrates. So, studies on the neighboring genes of the respective P450s were done to find possible genomic informations.

In So ce56, all 21 P450s are distributed all over the genomic DNA (Figure I. 2). Concerning the distribution of genes (including P450s) in So ce56, Schneiker *et al.*, 2007, studied the simultaneous analysis of the genome for GC content, dinucleotide frequency and codon usage and revealed a relatively uniform pattern of distribution of genes over the first two-thirds of the chromosome (i.e. up to 8.5 Mbp), but nine anomalous sections are present in the last third of the genome where the majority of the insertion sequences (16 of 18) is present (Schneiker *et al.*, 2007). It was found that, among all the P450s 76% of them were within that 8.5 Mb genome and the rest (24%) were on the last third of the genome, where the insertion sequences are presents. It was also found that the three new P450 families, CYP264 (CYP264A1 and CYP264B1), CYP265A1 and CYP267B1, along with three new members CYP109C2, CYP110J1 and CYP124E1, are present in the remaining last-third of the genome (Figure I. 12). Considering this, as 88.88% of the genes in this region was predicted to be transferred horizontally (Schneiker *et al.*, 2007), the P450s found in this region could have evolutionary significance. Interestingly, most of the P450s lying in this region were either completely new-families or members. Thus no experimental data for these P450s are available though being members of old families. It is very remarkable that even being bacterial P450s, all the tested P450s (CYP264A1, CYP110J1, CYP267B1, CYP124E1, CYP264B1 and CYP109C2 (Figure I. 2) present in this region were able to get electrons from bovine Adx and AdR.

The genome mining of individual P450s of So ce56 disclosed several characteristic properties on the genomic level and also provided some information for putative substrates (at least for some novel P450s). Two of the P450s of So ce56, CYP263A1 and CYP265A1, have a distinct PKS/NPRS module but the gene products have not been isolated till date. The similar theoretical consideration of PKS/NRPS module in the genomic informations is also available

for several *Streptomyces* and *Mycobacterium* sp. (Parajuli *et al.*, 2004; Zhao and Waterman, 2007). An understanding of the molecular basis for the biosynthesis of many complex polyketide antibiotics has only been gained over the past decade and has provided an explanation for the basis of the extensive structural diversity (Khosla *et al.*, 2001). Moreover, within *Streptomyces* sp., it has been shown that P450s are involved in many polyketide biosynthetic pathways and catalyzed the stereo- and regio-specific oxidative tailoring of antibiotics like novobiocin from *S. spheroides* with CYP163A1 (Chen and Walsh, 2001), doxorubicin from *S. peucitius* with CYP129A2 (Dickens *et al.*, 1997), amphotericin B from *S. nodosus* with CYP105H4 and CYP161A3 (Caffrey *et al.*, 2001), complestatin from *S. lavendulae* with CYP165E1 and CYP165B5 (Chiu *et al.*, 2001), avermectin from *S. avermitilis* with CYP171A1 (Ikeda *et al.*, 1999), PKSIII of *S. coelicolor* with CYP158A2 (Austin *et al.*, 2004), and so on. In this regard, the potential roles of CYP263A1 and CYP265A1 from So ce56 could be further research topics.

The genome mining of So ce56 also disclosed a putative terpene cyclase gene adjacent to CYP264B1 in the genome. This ORF resembles the genomic organization of CYP170A1 of *S. coelicolor* A(3)2 where the coupled action of this CYP170A1 with epi-isozizaene synthetase for the biosynthesis of albaflavenone has been described (Zhao *et al.*, 2008). A similar cluster was also reported in *S. avermitilis* where the gene SAV2999 encoding CYP183A1 has been described to catalyse the conversion of pentalenene(3) to pentalen-13-al(7) by stepwise allylic oxidation via pentalen-13-ol(6) (Quaderer *et al.*, 2006). This could also be a new aspect for further research.

The genomic mining of the So ce56 P450s also disclosed the presence of several complex regulators and other hypothetical protein encoding genes in a cluster. Unfortunately, no experimental data are available for the P450s.

5. 3. Expression, purification and characterization of cytochrome P450

All 21 cytochrome P450s of So ce56 were successfully cloned and expressed in *E. coli* in a soluble form and the highest yield of expression (after purification) was up to 1400 nmol/L for CYP266A1 whereas six of the P450s (CYP264B1, CYP110H1, CYP263A1, CYP261A1, CYP262B1 and CYP261B) have yield of <5 nmol/L (Table I. 9). Concerning this, the careful analysis of the gene inferred the following three reasons for differential expression levels. Firstly, the gene products could be toxic or a metabolic burden for the host cell. An inducible

expression system could avoid this problem. However, most of the commonly used promoters in *E. coli* are not tightly regulated (Wilms *et al.*, 2001). In this study, the expression vector pCWori⁺ was used which has *Lac*, *Tac* and *UV5* promoters which are of bacterial origin where the basal level leakage was obvious during protein expression. As the pET system has T7 promoter of viral origin, such basal level leakage could be controlled. It was observed that one of the P450s, CYP109D1, of *So ce56* when recloned in pET17b and co-expressed with the molecular chaperones pGro12_GroES/GroEL in host *E. coli* BL21(DE3) showed an expression level was 3-4 fold higher than the expression with pCWori⁺ in *E. coli* BL21. In addition, it was observed that the basal level leakage during expression could be somehow controlled by inducing the bacterial culture at high cell density.

Secondly, the high GC content could be responsible for the poor expression in *E. coli* but the effect was not uniform for all the *So ce56* P450s. The higher GC ratios present in the P450s of *So ce56* reflect the differences in the codon usage during heterologous expression. An artificial gene modifying such codon usage and adapting to *E. coli* could be synthesized for those P450s with low expression level (Khorana *et al.*, 1979). Though different approaches were implemented to synthesize the artificial genes, all of them were relatively expensive and time consuming. So during this work, this idea was not followed to improve the expression level of poorly expressed P450s. Alternatively, rare codon optimizing expression strains like *E. coli* Rosetta was used. These strains are engineered to contain extra copies of genes that can encode the tRNA that most frequently limit translation of heterologous protein expression in *E. coli*. In this study, the attempt of using such host for some of the low expressed P450s of *So ce56* somehow increased the expression level but it was still not high enough. Interestingly, the study of heterologous expression of 18 different P450s from the GC rich micro-organism *Streptomyces coelicolor* A(3)2 has shown that the P450s expression levels were also not uniform and varied from less than 50 to 3000 nmol/L culture (Lamb *et al.*, 2002).

Lastly, the GC-rich regions are also known for maintaining a strong secondary structure which could influence for the sequencing or amplification and even the heterologous expression of the corresponding genes. The GC-rich P450s sequences of *So ce56* may be the other limiting factor for the over-expression, as it might form more strong mRNA secondary structures. The translation initiation sequence is of crucial important for the expression of GC-rich genomes and is thus a possible target for optimization of the expression (Tang *et al.*,

1999). A practical method for the over-expression of GC-rich genes in *E. coli* by modification of the initial sequence was also published (Ishida *et al.*, 1997). Regarding this, in the genome two of the P450s (CYP262A1 and CYP267A1) of *So ce56* have GTG as a start codon which was converted to ATG during the cloning process where the expression yield of pure protein was fairly high (125 nmol/L) for CYP262A1 and even higher (1400 nmol/L) for CYP266A1.

Moreover, in this study, the frequency of rare-codons in the P450s of *So ce56* has also been analyzed. During the examination, a large number of rare-used codons in *E. coli*, particularly arginine (CGG and CGA), proline (CCC and CCT) and glycine (GGA and GGG) were observed (Table I. 15). The effect of rare-codons in all P450s was not uniform, irrespective of which the expression level in general was high.

Table I. 15: The frequency of rare codons in *Sorangium cellulosum* *So ce56* P450s. The rare-codon includes Arginine (Arg) , Proline (Pro) and Glycine (Gly).

P450s of <i>So ce56</i>	Gene*	Rare codons			Expression level (nmol/L cell culture)
		Arg (AGG, AGA, CGA)	Pro (CCC)	Gly (GGA, GGG)	
CYP109C1	sce0122	5	10	9	1333
CYP109C2	sce8913	7	6	5	1300
CYP109D1	sce4663	4	12	7	205
CYP110H1	sce3065	8	18	7	5
CYP110J1	sce6424	10	12	14	60
CYP117B1	sce5528	2	16	6	5
CYP124E1	sce7868	5	16	9	15
CYP259A1	sce4635	5	12	14	8
CYP260A1	sce1588	8	9	8	850
CYP260B1	sce4806	6	2	13	100
CYP261A1	sce0200	5	9	9	5
CYP261B1	sce5725	4	11	11	2
CYP262A1	sce2191	6	10	8	25
CYP262B1	sce1860	6	9	17	2
CYP263A1	sce4885	3	16	9	900
CYP264A1	sce6323	7	6	5	440
CYP264B1	sce8551	9	8	9	4
CYP265A1	sce8224	8	8	10	21
CYP266A1	sce5624	5	8	8	1400
CYP267A1	sce0675	1	11	3	23
CYP267B1	sce7167	1	10	9	70

* gene ID deposition in NCBI

Furthermore, several published results regarding the expression and purification of bacterial P450s have shown different expression conditions. The expression was shown to be optimized by different induction times, temperatures, shaking-speeds and aerations. As a result, there were no common expression parameters that can be equally followed for all the P450s of *So ce56*. However, it was found that the expression was always better during induction at higher OD of 0.9-1.5. This study also showed that the low shaking speed in a baffled flask with higher ratio of culture volume/flask generally yielded a better expression level. Regarding the optimum temperature for the expression, decreasing of the expression temperature from 37°C to 25-28°C was the best condition for a good expression of all the P450s of *So ce56*.

All the P450s of *So ce56* were purified by Ni-NTA affinity chromatography at 4°C. The protein was further purified by size exclusion chromatography during which the imidazole present after affinity chromatography was also removed. During this study, the effect of imidazole was not analyzed for any of the P450s. The purity of all the P450s was monitored by SDS-PAGE and the active form of the P450 was confirmed by COD measurements. Moreover, to illustrate the holo and apo-masses, seven of the P450s (CYP109C1, CYP109C2, CYP109D1, CYP260A1, CYP264A1, CYP264B1 and CYP266A1) were selected. The coincidence of the experimental holo-mass and the co-factor with that of theoretical values confirmed the active form (Table I. 12-13).

The spectrophotometric characterization of all the P450s showed a typical nature of P450s with peak maxima at 450 or 448 nm (Table I. 10) of the CO-complexed reduced P450. In some of the P450s, the P420 shoulder was more pronounced and the addition of the putative substrates helped to recover the P450 peak with minor or no shoulder at 420 nm, which represents the instability of P450s in an absence of the substrate.

In this study, the absence of redox partners for *So ce56* P450s has been overcome by using the mammalian heterologous electron partners adrenodoxin reductase (AdR) (Sagara *et al.*, 1993) and the soluble (2Fe-2S) cluster-containing ferredoxin, adrenodoxin (Adx₄₋₁₀₈) (Uhlmann *et al.*, 1994). They were able to transfer electrons to CYP109C1, CYP109C2, CYP109D1, CYP110J1, CYP124E1, CYP260A1, CYP264A1, CYP264B1, and CYP267B1. Likely, our group has already shown that the mammalian adrenodoxin and adrenodoxin reductase were able to couple with the bacterial CYP106A2 from *Bacillus megaterium* (Virus *et al.*, 2006;

Hannemann *et al.*, 2006; Lisurek *et al.*, 2008; Virus *et al.*, 2008) during the conversion of steroids.

6. Outlook

The myxobacterial strain *Sorangium cellulosum* So ce56 has the largest bacterial genome sequenced till date. Twenty-one P450s encoding ORFs, disclosing nine novel P450 families, were identified and a phylogenetic tree as well as a physical map of the So ce56 chromosome have been deduced. All P450s of So ce56 have been identified, cloned, expressed in *E. coli* and the corresponding soluble proteins were purified in an active form with a yield up to 1400 nmol/L for CYP266A1. The characteristic ferrous-carbon monoxide [Fe(II)CO] complex having a typical peak maximum at 448 nm and/or 450 nm confirmed the active existence of So ce56 P450s. The apo- and holo- mass of selected P450s (CYP109C1, CYP109C2, CYP109D1, CYP260A1, CYP264A1, CYP264B1 and CYP266A1) were not only in agreement with theoretical masses but also confirm their active form. The absence of homologous redox partners was overcome by using mammalian Adx and AdR for nine of the P450s (CYP109C1, CYP109C2, CYP109D1, CYP110J1, CYP124E1, CYP260A1, CYP264A1, CYP264B1, and CYP267B1). Screening of a compound-library of potential substrates disclosed positive hits for CYP109C1, CYP109C2, CYP260A1 and CYP266A1 and opened the way for *in vitro* substrate conversion and subsequent product identification studies.

7. Literature review

1. Ahn, J.W., Jang, K.H., Chung, S.C., Oh, K.B., and Shin, J. (2008). Sorangiadenosine, a new sesquiterpene adenoside from the myxobacterium *Sorangium cellulosum*. *Org. Lett.* *10*, 1167-116.
2. Austin, M.A., Izumikawa, M., Bowman, M.E., Udvary, D.W., Ferrer, J.L., Moore, B.S., and Nol, J.P. (2004). Crystal structure of a bacterial type III polyketide synthase and enzymatic control of reactive polyketide intermediate. *J. Biol. Chem.*, *279*, 45162-45174.
3. Axelrod, J. (1955). The Enzymatic Demethylation of Ephedrine. *J. Pharmacol. Exp. Ther.* *114*, 430-438.
4. Beatson, S.A., Whitchurch, C.B., Sargent, J.L., Levesque, R.C. and Mattick, J.S. (2002). Differential regulation of twitching motility and elastase production by Vfr in *Pseudomonas aeruginosa*. *J. Bacteriol.* *184*, 3605-13.
5. Berkeley, M.J. (1857). Introduction to Cryptogamic Botany (on *Stigmatella* and *Chondromyces*), pp.313-315. H. Bailliere Publishers, London.
6. Bernhardt, R. (1995). Cytochrome P450: structure, function and generation of reactive oxygen species. *Rev. Physiol. Biochem. Pharmacol.* *127*, 1-80.
7. Bernhardt, R. (2006). Cytochromes P450 as versatile biocatalysts. *J. Biotechnol.* *124*, 128-145.
8. Bode, H.B. and Müller, R. (2006). Analysis of myxobacterial secondary metabolism goes molecular. *J. Ind. Microbiol. Biotechnol.* *33*, 577-588.
9. Brodie, B.B., Axelrod, J., Cooper, J.R., Gaudette, L., Ladu, B.N., Mitoma, C., and Udenfriend, S. (1955). Detoxication of Drugs and Other Foreign Compounds by Liver Microsomes. *Science*, *121*, 603-604.
10. Buntin, K., Rachid, S., Scharfe, M., Blöcker, H., Weissman, J.K., and Müller, R. (2008). Production of the antifungal isochromanone ajudazols A and B in *Chondromyces crocatus* Cm c5: Biosynthetic machinery and cytochrome P450 modifications. *Angew. Chem. Int. Ed. Engl.* *47*, 4595-4599.
11. Caffrey, P., Lynch, S., Flood, E., Finnan, S., and Oliynyk, M. (2001). Amphotericin biosynthesis in *Streptomyces nodosus*: deductions from analysis of polyketide synthase and late genes. *Chem Biol.* *8*, 713-723.
12. Chen, Z., Ost, T.W., and Schelvis, J.P. (2004). Phe393 mutants of cytochrome P450 BM3 with modified heme redox potentials have. *Biochemistry.* *43*, 1798-1808.

13. Chen, H., and Walsh, C.T. (2001). Coumarin formation in novobiocin biosynthesis: beta-hydroxylation of the aminoacyl enzyme tyrosyl-S-NovH by a cytochrome P450 NovI. *Chem. Biol.* 8, 301-312.
14. Chiu, H.T., Hubbard, B.K., Shah, A.N., Eide, J., Fredenburg, R.A., Walsh, C.T., and Khosla, C. (2001). Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster. *Proc. Natl. Acad. Sci. U. S. A.* 98, 8548-8553.
15. Chun, Y.J., Shimada, T., Sanchez-Ponce, R., Martin, M.V., Lei, L., Zhao, B., Kelly, S.L., Waterman, M.R., Lamb, D.C., and Guengerich, F.P. (2007). Electron transport pathway for a *Streptomyces* Cytochrome P450: Cytochrome P450 105D5-catalyzed fatty acid hydroxylation in *Streptomyces coelicolor* A3(2). *J. Biol. Chem.* 282, 17486-17500.
16. Connolly, L., Yura, T. and Gross, C.A. (1999). Autoregulation of the heat shock response in Procaryotes. In Bukau, B. (ed.) *Molecular chaperones and folding catalysts*. Harwood academic publishers, Amsterdam, pp. 13-34.
17. Daff, S.N., Chapman, S.K., Turner, K.L., Holt, R.A., Govindaraj. S., Poulos, T.L., Munro A.W. (1997). Redox control of the catalytic cycle of flavocytochrome P450 BM-3. *Biochemistry*, 36, 13816-13823.
18. Danielson, P.B., (2003). The cytochrome P450 superfamily: Biochemistry, evolution and drug metabolism in humans. *Curr. Drug. Metab.* 3, 561-597.
19. Dawson, J.H., Andersson, L.A., and Sono, M. (1982). Spectroscopic Investigations of Ferric Cytochrome P-450-Cam Ligand Complexes - Identification of the Ligand Trans to Cysteinate in the Native Enzyme. *J. Biol. Chem.* 257, 3606-3617.
20. Denisov, I.G., Makris, T.M., Sligar, S.G., and Schlichting, I. (2005). Structure and chemistry of cytochrome P450. *Chem. Rev.* 105, 2253-2278.
21. Dickens, M.L., Priestley, N.D., and Strohl, W.R. (1997). In vivo and in vitro bioconversion of epsilon-rhodomyacinone glycoside to doxorubicin: functions of DauP, DauK, and DoxA. *J. Bacteriol.* 179, 2641-2650.
22. Downard, J., Ramaswamy, S.V. and Kil, K.S. (1993). Identification of *esg*, a genetic locus involved in cell-cell signaling during *Myxococcus xanthus* development. *J. Bacteriol.* 175, 7762-7770.
23. Dworkin, M. (1985). *Developmental biology of the bacteria*. The Benjamin/Cummings Publishing Company, Inc, California.

24. Dworkin, M. (1994). Increases in the intracellular concentration of glycerol during development in *Myxococcus xanthus* S. Courtney Frasch [corrected and republished in FEMS Microbiol. Lett. 122, 321-5
25. Dworkin, M. (1996). Recent advances in the social and developmental biology of the myxobacteria. Microbiol. Rev, 60, 70-102.
26. Estabrook, R. (2003). A passion for P450s (remembrances of the early history of research on cytochrome P450). Drug Metab. Dispos. 31,1461-73.
27. Freemont, P.S., Hanson, I.M., Trowsdale, J. (1991). A novel cysteine-rich sequence motif. Cell. 64, 483–484.
28. Garfinkel, D. (1957). Isolation and Properties of Cytochrome-B5 from Pig Liver. Arch. of Biochem. Biophys.71, 111-120.
29. Gerth, K., Bedorf, N., Irschik, H., Hofle, G. and Reichenbach, H. (1994). The soraphens: a family of novel antifungal compounds from *Sorangium cellulosum* (Myxobacteria). I. Soraphen A1 alpha: fermentation, isolation, biological properties. J. Antibiot. Tokyo, 47, 23-31.
30. Gerth, K., Pradella, S., Perlova, O., Beyer, S. and Müller, R. (2003). Myxobacteria: proficient producers of novel natural products with various biological activities-past and future biotechnological aspects with the focus on the genus *Sorangium*. J. Biotechnol. 106, 233-253.
31. Gibson, G.G and Skett, P. (1994). Introduction to drug metabolism. London: Blackie Academic and Professional.
32. Gigon, P.L., Gram, T.E., and Gillette, J.R. (1969) Studies on the rate of reduction of hepatic microsomal cytochrome P-450 by reduced nicotinamide adenine dinucleotide phosphate: effect of drug substrates, Mol. Pharmacol. 5, 109–122.
33. Gillam, E.M.J., and Hunter, D.J.B. (2007). Chemical defense and exploitation. Biotransformation of xenobiotics by cytochrome P450 enzymes. In: The Ubiquitous Roles of Cytochrome P450 Proteins: Metal Ions In Life Sciences, A. Sigel, ed. (John Wiley and Sons, Ltd), 3, pp. 477-560.
34. Girvan, H.M., Marshall, K.R., Lawson, R.J., Leys, D., Joyce, M.G., Clarkson, J., Smith, W.E., Cheesman, M.R., and Munro, A.W. (2004). Flavocytochrome P450 BM3 Mutant A264E Undergoes Substrate-dependent Formation of a Novel Heme Iron Ligand Set. J. Biol. Chem. 279, 23274-23286
35. Goldman, B. S., Nierman, W. C., Kaiser, D., Slater, S. C., Durkin, A. S., Eisen, J. A., Ronning, C. M.,Barbazuk, W. B., Blanchard M., Field C., Halling, C., Hinkle, G.,

- Iartchuk, O., Kim, H.S., Mackenzie, C., Madupu, R., Miller, N., Shvartsbeyn, A., Sullivan, S. A., Vaudin, M., Wiegand R. and Kaplan H. B. (2006). Evolution of sensory complexity recorded in a myxobacterial genome. *Proc. Natl. Acad. Sci. U.S.A.* *103*, 15200-15205.
36. Graziani, E.I., Cane, D.E., Betlach, M.C., Kealey, J.T., and McDaniel, R. (1998). Macrolide biosynthesis: A single cytochrome P450, PicK, is responsible for the hydroxylations that generate methymycin, neomethymycin, and picromycin in *Streptomyces venezuelae*. *Bioorg. Med. Chem. Lett.* *8*, 3117-3120.
 37. Green, A., Rivers, S., Cheesman, M., Reid, G., Quaroni, L., Macdonald, I., Chapman, S., and Munro, A. (2001). Expression, purification and characterization of cytochrome P450 BioI: a novel P450 involved in biotin synthesis in *Bacillus subtilis*. *J. Biol. Inorg. Chem.* *6*, 523-533.
 38. Guengerich, F.P. (2001). Uncommon P450-catalyzed Reactions. *Curr. Drug Metab.* *2*, 93-115.
 39. Guengerich, F.P., Wu, Z.L., and Bartleson, C.J. (2005). Function of human cytochrome P450s: characterization of the orphans. *Biochem. Biophys. Res. Commun.* *338*, 465-469.
 40. Guo, W.J. and Tao, W.Y., (2008). Phoxalone, a novel macrolide from *Sorangium cellulosum*: structure identification and its anti-tumor bioactivity in vitro. *Biotechnol. Lett.* *30*, 349-356.
 41. Hagemann (1990). *Gentechnologische Arbeitsmethoden*.
 42. Hagen, D.C., Bretscher, A.P. and Kaiser, D. (1978). Synergism between morphogenetic mutants of *Myxococcus xanthus*. *Dev. Biol.* *64*, 284-296.
 43. Haines, D.C. and Tomchick, D.R., (2001). Pivotal role of water in the mechanism of P450BM-3. *Biochemistry*, *40*, 13456-13465.
 44. Hannemann, F., Bichet, A., Ewen, K.M., and Bernhardt, R. (2007). Cytochrome P450 systems-biological variations of electron transport chains. *Biochem. Biophys. Acta. Gen. Sub.* *1770*, 330-344.
 45. Hannemann, F., Virus, C., and Bernhardt, R. (2006). Design of an Escherichia coli system for whole cell mediated steroid synthesis and molecular evolution of steroid hydroxylases. *J. Biotechnol.* *124*, 172-181.
 46. Hasemann, C.A., Kurumbail, R.G., Boddupalli, S.S., Peterson, J.A. and Deisenhofer, J. (1995). Structure and function of cytochromes P450: a comparative analysis of three. *Structure*, *3*, 41-62.

47. Haydock, S.F., Dowson, J.A., Dhillon, N., Roberts, G.A., Cortes, J., and Leadlay, P.F., (1991). Cloning and sequence analysis of genes involved in erythromycin biosynthesis in *Saccharopolyspora erythraea*: sequence similarities between EryG and a family of S-adenosylmethionine-dependent methyltransferases. *Mol. Gen. Genet.* *230*, 120-128.
48. Helvig, C., Koener, J.F., Unnithan, G.C., and Feyereisen, R. (2004). CYP15A1, the cytochrome P450 that catalyzes epoxidation of methyl farnesoate to juvenile hormone III in cockroach *Corpora allata*. *Proc. Natl. Acad. Sci. U. S. A.* *101*, 4024-4029.
49. Hoff, A.J. (Ed.) (1989) *Advanced EPR in Application in Biology and Biochemistry*, Elsevier Science, Amsterdam.
50. Ikeda, H., Nonomiya, T., Usami, M., Ohta, T., and Omura, S. (1999). Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*. *Proc. Natl. Acad. Sci. U. S. A.* *96*, 9509-9514.
51. Imai, M., Shimada, H., Watanabe, Y., Matsushima-Hibiya, Y., Makino, R., Koga, H., Horiuchi, T. and Ishimura, Y. (1989). Uncoupling of the cytochrome P-450cam monooxygenase reaction by a single. *Proc. Natl. Acad. Sci. U. S. A.* *86*, 7823-7827.
52. Irschik, H., Reichenbach, H., Höfle, G., and Jansen, R. (2007). The thuggacins, novel antibacterial macrolides from *Sorangium cellulosum* acting against selected Gram-positive bacteria. *J. Antibiot.* *60*, 733-738.
53. Ishida, M., Yoshida, M., and Oshima, T. (1997). Highly efficient production of enzymes of an extreme thermophile, *Thermus thermophilus*: A practical method to overexpress GC-rich genes in *Escherichia coli*. *Extremophiles.* *1*, 157-162.
54. Jelsbak, L., Givskov, M., and Kaiser, D. (2005). Enhancer-binding proteins with a forkhead-associated domain and the σ^{54} regulon in *Myxococcus xanthus* fruiting body development. *Proceedings of the Proc. Natl. Acad. Sci. U. S. A.* *102*, 3010-3015.
55. Julien, B., Shah, S., Ziermann, R., Goldman, R., Katz, L., and Khosla, C. (2000). Isolation and characterization of the epothilone biosynthetic gene cluster from *Sorangium cellulosum*. *Gene* *249*, 153-160.
56. Jungmann, V., Molnar, I., Hammer, P.E., Hill, D.S., Zirkle, R., Buckel, T.G., Buckel, D., Ligon, J.M., and Pachlatko, J.P. (2005). Biocatalytic conversion of avermectin to 4-oxo-avermectin: characterization of biocatalytically active bacterial strains and of cytochrome p450 monooxygenase enzymes and their genes. *Appl. Environ. Microbiol.* *71*, 6968-6976.

57. Kalb, V.F., Woods, C.W., Turi, T.G., Dey, C.R., Sutter, T.R., and Loper, J.C. (1987). Primary structure of the P450 lanosterol demethylase gene from *Saccharomyces cerevisiae*. *DNA*. 6, 529-537.
58. Karuzina, I.I, and Archakov, A.I. (1994). The oxidative inactivation of cytochrome P450 in monooxygenase reactions. *Free Radic. Biol. Med.* 16, 73-97.
59. Katagiri, M., Ganguli, B.N., and Gunsalus, I.C. (1968). A Soluble Cytochrome P-450 Functional in Methylene Hydroxylation. *J. Biol. Chem.* 243, 3543-3546.
60. Kelly, S.M., Jess, T.J. and Price, N.C. (2005) How to study proteins by circular dichroism. *Biochim. Biophys. Prot. Proteom.* 1751, 119-139.
61. Khorana, H.G. (1979). Total synthesis of a gene. *Science*. 203, 614-625.
62. Kimata, Y., Shimada, H., Hirose, T. and Ishimura, Y., (1995). Role of Thr-252 in cytochrome P450cam: a study with unnatural amino acid. *Biochem. Biophys. Res. Commun.* 208, 96-102.
63. Klingenberg, M. (1958). Pigments of Rat Liver Microsomes. *Arch. Biochem. Biophys.* 75, 376-386.
64. Koegl, M., Hoppe, T., Schlenker, S., Ulrich, H.D., Mayer, T.U., Jentsch, S. (1999). A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell*. 96, 635-644.
65. Kosla, C.H. and Harbury, P.B. (2001). Modular enzymes. *Nature*. 409,247-252.
66. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 227, 680-685.
67. Lamb, D.C., Ikeda, H., Nelson, D.R., Ishikawa, J., Skaug, T., Jackson, C., Omura, S., Waterman, M.R., and Kelly, S.L. (2003). Cytochrome P450 complement (CYPome) of the avermectin-producer *Streptomyces avermitilis* and comparison to that of *Streptomyces coelicolor* A3(2). *Biochem. Biophys. Res. Commun.* 307, 610-619.
68. Lamb, D.C., Skaug, T., Song, H.L., Jackson, C.J., Podust, L.M., Waterman, M.R., Kell, D.B., Kelly, D.E., and Kelly, S.L. (2002) The cytochrome P450 complement (CYPome) of *Streptomyces coelicolor* A3(2). *J. Biol. Chem.* 277, 24000-24005.
69. Lamb, D.C., Lei, L., Warrilow, A.G., Lepesheva, G.I., Mullins, J.G., Waterman, M.R. and Kelly, S.L. (2009). The first virally encoded cytochrome P450. *J. Virol.* 83, 8266-9.
70. LaRossa, R., Kuner, J., Hagen, D., Manoil, C. and Kaiser, D. (1983). Developmental cell interactions of *Myxococcus xanthus*: analysis of mutants. *J. Bacteriol*, 153, 1394-1404.

71. Lawson R.J., Leys, D., Sutcliffe, M.J., Kemp, C.A., Cheesman, M.R., Smith S.J., Clarkson, J. Smith, W.E., Haq, I., Perkins, J.B., and Munro, A.W. (2004). Thermodynamic and Biophysical Characterization of Cytochrome P450 BioI from *Bacillus subtilis*. *Biochemistry*, *43*, 12410-12426.
72. LeBrun, L.A., Hoch,U. and de Montellano, P. R. O. (2002a). Autocatalytic mechanism and consequences of covalent heme attachment in the cytochrome P4504A family. *J. Biol. Chem.* *277*. 12755-12761.
73. LeBrun, L.A., Xu, F.Y., Kroetz, D.L., and de Montellano, P.R.O. (2002b). Covalent attachment of the heme prosthetic group in the CYP4F cytochrome p450 family. *Biochemistry*. *41*, 5931-5937.
74. Lewis, D.F.V. (2001). On the recognition of mammalian microsomal cytochrome P450 substrates and their characteristics - Towards the prediction of human P450 substrate specificity and metabolism. *Biochemical Pharmacology*. *60*, 293-306.
75. Lisurek, M., Simgen, B., Antes, I. and Bernhardt, R. (2008). Theoretical and experimental evaluation of a CYP106A2 low homology model and production of mutants with changed activity and selectivity of hydroxylation. *Chembiochem*. *9*, 1439–1449
76. Lewis, D.F.V. (1996). *Cytochrome P450 – Structure function and Mechanism* (Taylor and Francis, Basingstoke).
77. Lewis, D.F.V., Ito, Y. and Goldfarb, P.S. (2006). Structural modeling of the human drug-metabolizing cytochromes P450. *Curr. Med. Chem.* *13*, 2645-52.
78. Li, X., Yu, T. K., Kwak, J. H., Son, B. Y., Seo, Y., Zee, O. P., and Ahn, J. W. (2008). Soraphinol C, a new free-radical scavenger from *Sorangium cellulosum*. *J. Microbiol. Biotechnol.* *18*, 520-522.
79. Liu, Y., Moenne-Loccoz, P., Loehr, T.M., and Ortiz de Montellano, P.R. (1997). Heme oxygenase-1, intermediates in verdoheme formation and the requirement for reduction equivalents. *J. Biol. Chem.* *272*, 6909-6917.
80. Loew, G.H., and Harris, D.L. (2000). Role of the heme active site and protein environment in structure, spectra and function of the cytochrome P450s. *Chemical Reviews*. *100*, 407-419.
81. Lorick, K.L, Jensen, J.P., Fang, S., Ong, A.M., Hatakeyama, S., Weissman, A.M. (1999). RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proc. Natl. Acad. Sci. U. S. A.* *96*, 11364–11369.

82. Ludwig, W., Schleifer, K.H., Reichenbach, H., Stackebrandt, E. (1983). A phylogenetic analysis of the myxobacteria *Myxococcus fulvus*, *Stigmatella aurantiaca*, *Cystobacter fuscus*, *Sorangium cellulosum* and *Nannocystis exedens*. Arch. Microbiol. 135, 58-52.
83. Mansuy, D., and Renaud, J. (1995) Heme-thiolate proteins different from cytochromes P450 catalysing monooxygenations, in Cytochrome P450: Structure, mechanism and specificity (Ortiz de Montellano P.R., ed), Plenum press, New York. pp. 537-574.
84. McBride, M.J. (2001). Bacterial gliding motility: multiple mechanisms for cell movement over surfaces. Annu. Rev. Microbiol. 55, 49-75.
85. McCudry, H. (1989). Fruiting Gliding Bacteria. In: Holt, J. (Ed.), Bergey's Manual of Systematic Bacteriology. Williams & Wilkins, Baltimore, pp. 2139-2143.
86. McLean, K.J., Dunford, A.J., Neeli, R., Driscoll, M.D., and Munro, A.W. (2007). Structure, function and drug targeting in Mycobacterium tuberculosis cytochrome P450 systems. Arch. Biochem. Biophys. 464, 228-240.
87. Menche, D., Arikian, F., Perlova, O., Horstmann, N., Ahlbrecht, W., Wenzel, S.C., Jansen, R., Irschik, H., and Müller, R. (2008). Stereochemical determination and complex biosynthetic assembly of ethangien, a highly potent RNA-polymerase inhibitor from the myxobacterium *Sorangium cellulosum*. J. Am. Chem. Soc. (In press).
88. Merz, A.J., So, M. and Sheetz, M.P. (2000). Pilus retraction powers bacterial twitching motility. Nature. 407, 98-102.
89. Mulzer, H.J. (2008). The epothilones- an outstanding family of anti-tumor agents: From soil to the clinic. (Springer-Verlag, Austria).
90. Munro, A.W., Noble, M.A., Ost, T.W.B., Green, A.J., McLean, K.J., Robledo, L., Miles C.S., Murdoch, J., and Chapman, S.K. (2000) Flavocytochrome P450 BM-3 – substrate selectivity and electron transfer in a model cytochrome P450, Subcellular Biochemistry 35: Enzyme catalysed electron and radical transfer (ed. Holzenburg and Scrutton), Plenum publishers, New York.
91. Myer, Y. P., and Pande, A. J. (1978). In The Porphyrins (Dolphin, D., ed.), Academic Press, New York, Vol. 3, pp. 271–322.
92. Narhi, L.O., and Fulco, A.J. (1986). Characterization of a catalytically self-sufficient 119,000-dalton cytochrome P-450 monooxygenase induced by barbiturates in *Bacillus megaterium*. J. Biol. Chem. 261, 7160-7169.

93. Nelson, D.R., Koymans, L., Kamataki, T., Stegeman, J.J., Feyereisen, R., Waxman, D.J., Waterman, M.R., Gotoh, O., Coon, M.J., Estabrook, R.W., Gunsalus, I.C., Nebert, D.W. (1996). P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* 6, 1-42.
94. Noble, M.A., Miles, C.S., Chapman, S.K., lysek, D.A., MacKay, A.C., Reid, G.A., Hanzlik, R.P., Munro, A.W. (1999). Roles of key active-site residues in flavocytochrome P450 BM3. *Biochem. J.* 339, 371.
95. O'Connor, K.A. and Zusman, D.R. (1997). Starvation-independent sporulation in *Myxococcus xanthus* involves the pathway for beta-lactamase induction and provides a mechanism for competitive cell survival. *Mol. Microbiol.* 24, 839-850.
96. Omura, T., and Sato, R. (1962). A New Cytochrome in Liver Microsomes. *J. Biol. Chem.* 237, 1375-1376.
97. Omura, T., and Sato, R. (1964). Carbon Monoxide-Binding Pigment of Liver Microsomes.I. Evidence for Its Hemoprotein Nature. *J. Biol. Chem.* 239, 2370-2378.
98. Ost, T.W., Munro, A.W., Mowat, C.G., Taylor, P.R., Pesseguiro, A., Fulco, A.J., Cho, A.K., Cheesman, M.A., Walkinshaw, M.D. and Chapman, S.K. (2001). Structural and spectroscopic analysis of the F393H mutant of flavocytochrome P450. *Biochemistry* 40, 13430-13438.
99. Page, R.D.M. (1996). Tree View: An application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12, 357-358.
100. Parajuli, N., Basnet, D.B., Chan Lee, H., Sohng, J.K., and Liou, K. (2004). Genome analyses of *Streptomyces peucetius* ATCC 27952 for the identification and comparison of cytochrome P450 complement with other *Streptomyces*. *Arch. Biochem. Biophys.* 425, 233-241.
101. Perez, J., Castaneda-Garcia, A., Jenke-Kodama, H., Müller, R., and Munoz-Dorado, J. (2008). Eukaryotic-like protein kinases in the prokaryotes and the myxobacterial kinome. *Proc. Natl. Acad. Sci.* (In press).
102. Presnell, S.R. and Cohen, F.E. (1989). Topological distribution of four-alpha-helix bundles. *Proc. Natl. Acad. Sci. U. S. A.* 86, 6592-6596.
103. Quaderer, R., Omura, S., Ikeda, H., and Cane, D.E. (2006). Pentalenolactone Biosynthesis. Molecular Cloning and Assignment of Biochemical Function to PtlI, a Cytochrome P450 of *Streptomyces avermitilis*. *J. Am. Chem. Soc.* 128, 13036-13037.
104. Reichenbach, H. (1999). The ecology of the myxobacteria. *Environ. Microbiol.* 115-21.

105. Reichenbach, H. (2001). Myxobacteria, producers of novel bioactive substances. *J. Ind. Microbiol. Biotechnol.* 27, 149-156.
106. Reichenbach, H. (2004). The Myxococcales. In: Garrity, G.M. (Ed.), *Bergey`s Manual of Systematic Bacteriology*. Springer-Verlag, New York, pp. 1059-1143.
107. Reichenbach, H. and Dworkin M. (1992). The myxobacteria “The Prokaryotes” 2nd Edition Vol .IV, Chapter 188.
108. Reichenbach, H. and Dworkin, M. (1992). The order *Myxobacterales*. In Starr, M.P., Trüper, H.G., Balows, A. and Schlegel, H.G. (eds.), *The prokaryotes*. Springer-Verlag, Berlin, pp. 328-355.
109. Reichenbach, H. and G. Höfle (1999). Myxobacteria as producers of secondary metabolites. In *Drug Discovery from Nature*, pp. 149-179. Ed: S. Grabley & R. Thiercke. Berlin, Heidelberg: Springer.
110. Reichenbach, H. and Höfle, G. (1993). Biologically active secondary metabolites from myxobacteria. *Biotech. Adv.* 11, 219-277.
111. Rendic, S. and Carlo, Di. (1997). Human Cytochrome P450 Enzymes: a status report summarizing their reactions, substrates, inducers and inhibitors. *Drug Metab. Rev.* 29, 413-580.
112. Sagara, Y., Wada, A., Takata, Y., Waterman, M.R., Sekimizu, K. and Horiuchi, T. (1993). Direct expression of adrenodoxin reductase in *Escherichia coli* and the functional characterization. *Biol. Pharm. Bull.* 16, 627-630.
113. Sambrook, J. and Russell, D.W. (2001). *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
114. Schenkman, J.B., Sligar, S.G., and Cinti, D.L. (1981). Substrate interaction with cytochrome P-450. *Pharmacol. Ther.* 12, 43-71.
115. Schneiker, S., Perlova, O., Kaiser, O., Gerth, K., Alici, A., Altmeyer, M.O., Bartels, D., Bekel, T., Beyer, S., Bode, E., Bode, H.B., Bolten, C.J., Choudhuri, J.V., Doss, S., Elnakady, Y.A., Frank, B., Gaigalat, L., Goesmann, A., Groeger, C., Gross, F., Jelsbak, L., Jelsbak, L., Kalinowski, J., Kegler, C., Knauber, T., Konietzny, S., Kopp, M., Krause, L., Krug, D., Linke, B., Mahmud, T., Martinez-Arias, R., McHardy, A.C., Merai, M., Meyer, F., Mormann, S., Munoz-Dorado, J., Perez, J., Pradella, S., Rachid, S., Raddatz, G., Rosenau, F., Ruckert, C., Sasse, F., Scharfe, M., Schuster, S.C., Suen, G., Treuner-Lange, A., Velicer, G.J., Vorholter, F.-J., Weissman, K.J., Welch, R.D., Wenzel, S.C., Whitworth, D.E., Wilhelm, S., Wittmann, C., Blocker, H., Puhler, A.,

- and Muller, R. (2007). Complete genome sequence of the myxobacterium *Sorangium cellulosum*. *Nat. Biotech.* 25, 1281-1289.
116. Schuler, M.A., and Werck-Reichhart, D. (2003). Functional genomics of P450s. *Annu. Rev. Plant Biol.* 54, 629-667.
 117. Schenkman, J.B., Sligar, S.G., and Cinti, D.L. (1981). Substrate interaction with cytochrome P-450. *Pharmacol. Ther.* 12, 43-71.
 118. Sharrock, M., Munck, E., Debrunner, P.G., Marshall, V., Lipscomb, J.D. and Gunsalus, I.C. (1973). Mossbauer studies of cytochrome P-450 cam. *Biochemistry.* 12, 258-265.
 119. Shimkets, L.J., Dworkin, M. and Reichenbach, H. (2005). The myxobacteria. In: *The Prokaryotes*, third ed. Springer-Verlag, New York, NY, release 3.19.
 120. Sligar, S.G. and Gunsalus, I.C. (1976) A thermodynamic model of regulation: modulation of redox equilibria in camphor monooxygenase. *Proceedings of the National Academy of Sciences of the United States of America.* 73, 1078-1082.
 121. Tang, L., Shah, S., Chung, L., Carney, J., Katz, L., Khosla, C., and Julien, B. (2000). Cloning and Heterologous Expression of the Epothilone Gene Cluster. *Science.* 287, 640-642.
 122. Tang, S.C., (1976). Letter: Oxidized cytochrome P-450. Magnetic circular dichroism evidence for. *J. Am. Chem. Soc.* 98, 3707-3708.
 123. Tang, W., and Tseng, H. (1999). A GC-rich sequence within the 5' untranslated region of human basoquin mRNA inhibits its translation. *Gene.* 237, 35-44.
 124. Thaxter, R. (1892). On the Myxobacteriaceae, a new order of Schizomycetes. *Bot. Gaz.* 17, 389-406.
 125. Thomasson, B., Link, J., Stassinopoulos, A.G., Burke, N., Plamann, L. and Hartzell, P.L. (2002). MglA, a small GTPase, interacts with a tyrosine kinase to control type IV pili-mediated motility and development of *Myxococcus xanthus*. *Mol. Microbiol.* 46, 1399-413.
 126. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* 22, 4673-4680.
 127. Torres, S., Fjetland, C., and Lammers, P. (2005). Alkane-induced expression, substrate binding profile, and immunolocalization of a cytochrome P450 encoded on the *nifD* excision element of *Anabaena* 7120. *BMC Microbiol.* 5, 16.

128. Tully, R.E., van Berkum, P., Lovins, K.W., and Keister, D.L. (1998). Identification and sequencing of a cytochrome P450 gene cluster from *Bradyrhizobium japonicum*. *Biochem. Biophys. Acta.* *1398*, 243-255.
129. Uchida, E., Kagawa, N., Sakaki, T., Urushino, N., Sawada, N., Kamakura, M., Ohta, M., Kato, S., and Inouye, K. (2004). Purification and characterization of mouse CYP27B1 overproduced by an *Escherichia coli* system coexpressing molecular chaperonins GroEL/ES. *Biochem. Biophys. Res. Commun.* *15*, 323, 505-11.
130. Uhlmann, H., Kraft, R., and Bernhardt, R. (1994). C-terminal region of adrenodoxin affects its structural integrity and determines differences in its electron transfer function to cytochrome P-450. *J. Biol. Chem.* *269*, 22557-22564.
131. Vidakovic, M., Sligar, S.G., Li, H. and Poulos (1998). Understanding the role of the essential Asp251 in cytochrome p450cam using. *Biochemistry*, *37*, 9211-9219.
132. Virus, C., Lisurek, M., Simgen, B., Hannemann, F. and Bernhardt, R. (2006). Function and engineering of the 15beta-hydroxylase CYP106A2. *Biochem. Soc. Trans.* *34*, 1215–1218
133. Virus, C. and Bernhardt, R. (2008). Molecular Evolution of a Steroid Hydroxylating Cytochrome P450 Using a Versatile Steroid Detection System for Screening. *Lipids.* *43*, 1133–1141
134. Wall, D. and Kaiser, D. (1999). Type IV pili and cell motility. *Mol. Microbiol.* *32*, 1-10.
135. Wenzel, C.S., and Müller, R. (2007). Myxobacterial natural product assembly lines: fascinating examples of curious biochemistry. *Nat. Prod. Rep.* *24*, 1211-1224.
136. Werck-Reichhart, D. and Feyereisen, R. (2006). Cytochromes P450: a success story. *J. Biotechnol.* *124*, 128-45.
137. Wilms, B., Hauck, A., Reuss, M., Sylđatk, C., Mattes, R., Siemann, M., and Altenbuchner, J. (2001). High-cell-density fermentation for production of L-N-carbamoylase using an expression system based on the *Escherichia coli* rhaBAD promoter. *Biotechnol. Bioeng.* *73*, 95-103.
138. Wolgemuth, C., Hoiczuk, E., Kaiser, D. and Oster, G. (2002). How myxobacteria glide? *Curr. Biol.* *12*, 369-77.
139. Zhao, B., and Waterman, M.R. (2007). Novel properties of P450s in *Streptomyces coelicolor*. *Drug Metabol. Rev.* *39*, 343-352.

140. Zhao, B., Lin, X., Lei, L., Lamb, D.C., Kelly, S.L., Waterman, M.R., and Cane, D.E. (2008). Biosynthesis of the sesquiterpene antibiotic albaflavenone in *Streptomyces coelicolor* A3(2). *J. Biol. Chem.* 283, 8183-8189.

CHAPTER II

A. CYP109D1 from *Sorangium cellulosum* So ce56: A fatty acids hydroxylase

A. 1. INTRODUCTION

Myxobacteria have very complex fatty acids profiles and up to 40 different compounds belonging to all known classes of fatty acids were identified (Ware and Dworkin *et al.*, 1973; Bode *et al.*, 2005, Kaneda, 1991). It was shown that odd numbered, iso- or in some case, anteiso-branched fatty acids were predominant in myxobacteria and in some other gliding bacteria (Fautz *et al.*, 1979). It was also presumed that the unexpected abundance of fatty acids play an important role in the complex life cycle of myxobacteria which culminates in the formation of fruiting bodies containing heat and desiccation resistant myxospores (Shimkets, 1999; Moraleda-Munoz and Shimkets, 2007; Kaiser, 2003). It was previously demonstrated that the straight chain fatty acids are dispensable in *Myxococcus xanthus* for vegetative growth and fruiting body formation (Bode *et al.*, 2006). Moreover, fatty acids were identified to play a role in chemotaxis and it was also illustrated that phosphatidylethanolamine (PE) containing fatty acid, 16:1 ω 5(Δ^{11}), elicits a chemotactic response (Curtis *et al.*, 2006).

Members of cytochrome P450 families have been shown to play a pivotal role in fatty acid hydroxylation and conversion. Several publications have shown that mammalian, (Fer *et al.*, 2008; Hardwick, 2008; Holmes *et al.*, 2004; Hashizume *et al.*, 2002), plant (Imaishi *et al.*, 2007; Sangeewa *et al.*, 2007; Morant *et al.*, 2007; Coon, 2005), fungal (Kim *et al.*, 2007; Katayume *et al.*, 2008; Eschenfeld, *et al.*, 2003) and bacterial (Li *et al.*, 2000; Chun *et al.*, 2007; Green *et al.*, 2001; Girhard *et al.*, 2007; Cryle *et al.*, 2006; Budde *et al.*, 2006) cytochrome P450s were able to do fatty acid hydroxylation. Among them, P450BM-3 (CYP102) has one of the highest turnover numbers for a monooxygenation reaction (Gustafsson *et al.*, 2001) but the exact physiological role is still unclear. The hydroxy fatty acids have also gained industrial interests because of their special properties such as higher viscosity and reactivity compared with normal fatty acids (Bagby and Calson, 1989). These special chemical properties enable hydroxy fatty acids to have a wide range of industrial applications including the production of resins, waxes, nylons, plastics, lubricants, cosmetics and additives in coatings and paintings. In addition, hydroxy fatty acids were reported to contain antimicrobial activities against plant pathogenic fungi and some bacteria (Bajpai *et al.*, 2004; Hou and Forman, 2000; Kato *et al.*, 1984; Shin *et al.*, 2004). Recently, much effort

has been focused on the microbial production of hydroxy fatty acids from various fatty acid substrates.

A. 2. Objectives of the Chapter II. A

1. To search for fatty acid hydroxylating P450s of *So ce56*.
2. To show the functional role of the CYP109 family as fatty acid hydroxylases.
 - a) To analyze the position of hydroxylation in the fatty acid products
 - b) To compare fatty acid hydroxylation patterns of CYP109 family members (CYP109C1, CYP109C2 and CYP109D1) of *So ce56*.

A. 3. MATERIALS AND METHODS

A.3.1. Materials

The instruments, chemicals, enzymes, kits, reagents and other accessories used in this work were of high grade. They are illustrated in Appendix I.

The included fatty acids in this study were even-carbon-saturated fatty acid [capric acid (decanoic acid) ($C_{10}H_{20}O_2$), lauric acid (dodecanoic acid) ($C_{12}H_{24}O_2$), myristic acid (tetradecanoic acid) ($C_{14}H_{28}O_2$), and palmitic acid (hexadecanoic acid) ($C_{16}H_{32}O_2$)], odd-carbon-saturated fatty acid [tridecanoic acid ($C_{13}H_{22}O_2$)], mono-unsaturated fatty acid [Oleic acid (cis-9-octadecanoic acid) (18:1 cis-9)] and branched fatty acids (13-methyl myristic acid and 15-methyl palmitic acid).

A. 3. 2. Methods

The bioinformatics studies, protein expression, purification and spectrophotometric characterizations are as illustrated in Chapter I.

A. 3. 3. Spin state shift and substrate dissociation constant determination

The spin-state shifts upon substrate binding were assayed at room temperature under aerobic condition using an UV-Vis scanning photometer (UV-2101PC, Shimadzu, Japan) equipped with two tandem cuvettes (Noble *et al.*, 1999). One chamber of each cuvette contained 5 μ M CYP109D1 in buffer E (Appendix III) and the second chamber contained buffer alone. The fatty acids dissolved in DMSO were added in small aliquots (< 5 μ l) from an appropriate stock into the CYP109D1 containing chamber of the sample cuvette. Meantime, equal amounts of substrate were added into the buffer containing chamber of the reference cuvette. The spectral changes between 350 and 500 nm were recorded. When saturation of CYP109D1 with a substrate was achieved, the dissociation constant (K_D) was calculated by fitting a hyperbolic curve of the peak-to-trough difference to substrate concentration by using Sigma Plot program.

A. 3. 4. *In vitro* assay of fatty acid and derivatization

The *in vitro* reactions were performed in a final volume of 0.5 ml in buffer E (Appendix III) containing 200 μ M fatty acids as a substrate and a NADPH-regenerating system, consisting of glucose-6-phosphate (5mM), glucose-6-phosphate dehydrogenase (1 unit) and magnesium chloride (1 mM). The combinations of reconstituted redox partners P450: Adx: AdR or P450:

Fdx: FdR of 1: 10: 1 (1 μ M: 10 μ M: 1 μ M) were used for assay purpose. During incubation at 30°C on a thermo shaker, NADPH was added to 500 μ M final concentration and the mixture was incubated for 20 min. The reaction was stopped with an equal volume of diethyl ether. The product was extracted with diethyl ether for two times and was vacuum dried. This product was dissolved in 35 μ l of derivatizing solution (N,O-bis(trimethylsilyl)trifluoroacetamide + 1% trimethylchlorosilan) in a GC-vile, incubated at 70°C for 30 min prior to GC-MS analysis and was analyzed as below.

A. 3. 5. Fatty acid hydroxylation and GC-MS analysis

All the GC-MS experiments were performed in Institute of Technical Biochemistry, Universitaet Stuttgart in collaboration with Dr. Vlada B. Urlacher.

GC-MS analysis was performed on Shimadzu GCMS-QP2010 (column: FS-Supreme-5, length: 30 m, internal diameter: 0.25 mm, film thickness: 0.25 μ m) using helium as a carrier gas. The mass spectra were collected using electron ionization. The column oven was programmed as in Table II. 1. The programs that were used for fatty acid analysis are shown in Table II. 2.

Table II. A. 1: The standard GC-MS parameters for fatty acids analysis.

Standard GC parameters		Standard MS parameters	
Injection mode	split	Ion source temperature	200°C
Flow control mode	Linear velocity	Interface temperature	285°C
Purge flow	3 ml/min	Start m/z	50
		End m/z	450

Table II. A. 2 (1-4): The column oven program for fatty acid analysis.

1. Column oven program for decanoic, dodecanoic and tridecanoic acid				2. Column oven program for tetradecanoic and pentadecanoic acid			
Rate (°C/min)	Final (°C)	Temp (min)	Hold (min)	Rate (°C/min)	Final (°C)	Temp (min)	Hold (min)
-----	165		1	-----	190		1
6	215		0	7	245		0
30	275		1	30	275		1

3. Column oven program for hexadecanoic and heptadecanoic acid			4. Column oven program for oleic acid		
Rate (°C/min)	Final Temp (°C)	Hold (min)	Rate (°C/min)	Final Temp (°C)	Hold (min)
-----	190	1	-----	190	1
6	255	0	8	300	1
30	285	1			

The mono-hydroxylated fatty acids were identified by their characteristic fragmentation pattern. The stereochemistry of the products was not identified.

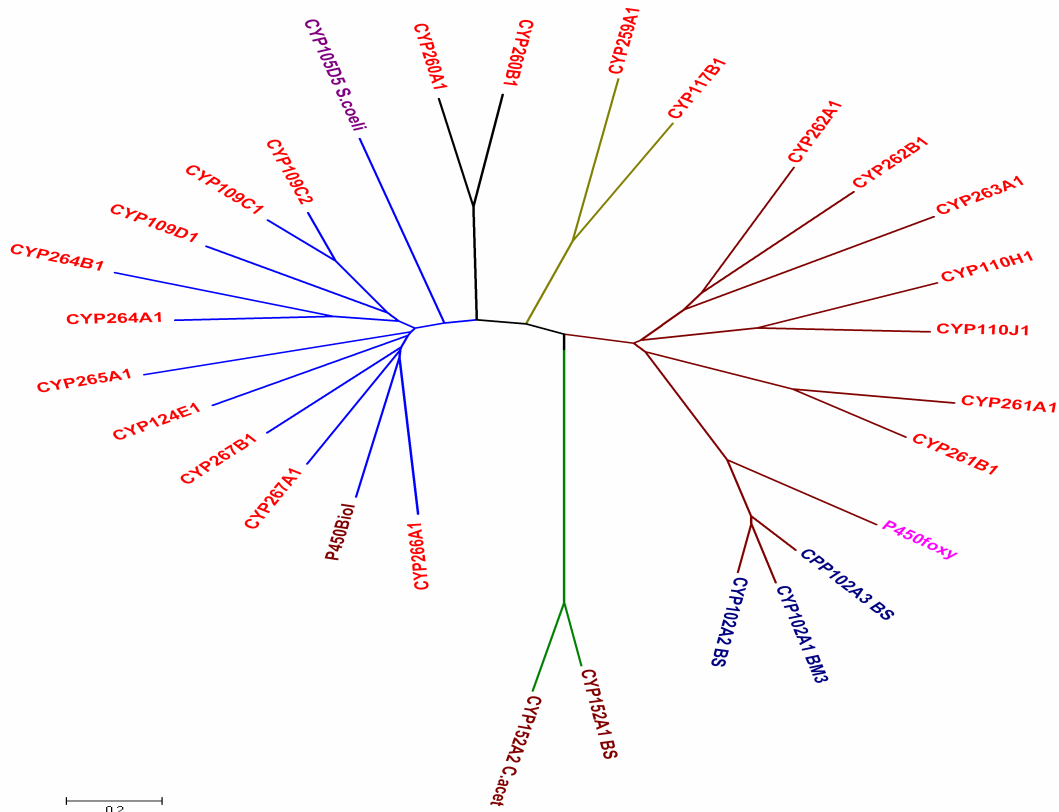
A. 4. RESULTS

A. 4. 1. Search for a fatty acid hydroxylating P450 of So ce56

The inherent role of the lipids during the complex life cycle of myxobacteria strains (Schneiker, *et al.*, 2007; Ware and Dworkin, 1973), guided us to presume fatty acids as potential substrates for some of the cytochromes P450s. The amino acid sequences of fatty acid hydroxylating P450s [CYP102A1 of *Bacillus megaterium*, CYP102A2, CYP102A3 and CYP152A1 of *Bacillus subtilis*, P450foxy from *Fusarium oxysporum*, CYP152A2 of *Clostridium acetobutylicum*, P450BioI of *Bacillus subtilis* and CYP105D5 of *Streptomyces coelicolor* A(3)2] were retrieved from NCBI. The P450 amino acid sequences of So ce56 were aligned (Appendix VII. A-B) and a unrooted phylogenetic tree was constructed. It was found that CYP109 (CYP109C1, CYP109C2, CYP109D1), CYP124E1, CYP264 (CYP264A1 and CYP264B1), CYP265A1, CYP267 (CYP267A1 and CYP267B1) of So ce56 were clustered with CYP109D5 of *S. coelicolor* A(3)2 and P450BioI of *B. subtilis*. Likely, CYP110 (CYP110H1 and CYP110J1), CYP261 (CYP261A1 and CYP261B1), CYP262 (CYP262A1 and CYP262B1) and CYP263A1 are clustered with CYP102A1 of *B. megaterium*, and CYP102A2 and CYP102A3 of *Bacillus subtilis* (Figure II. A. 1. a).

In order to highlight the significance of fatty acid hydroxylating P450s in So ce56, CYP109 family member (CYP109C1, CYP109C2 and CYP109D1) was selected for further studies. All these P450 members were found to have relatively high expression yield (Table I. 9, Chapter I). Moreover, the heterologous redox partners, Adx and AdR, were already identified as an efficient electron transfer partners (Chapter I) which is the bottle neck for the reconstitution of the monooxygenase activity. Furthermore, these P450s were found to have a very high homology with the enzymes P450BioI of *Bacillus subtilis* and CYP105D5 of *Streptomyces coelicolor* A(3)2 (Figure II. A. 1. B), the well characterized fatty acid hydroxylases for which the fatty acid hydroxylation patterns were already shown (Green *et al.*, 2001 and Chun *et al.*, 2007). The individual pairwise alignment of CYP109D1 showed 30.6% identity (47.8% similarity) with P450BioI of *Bacillus subtilis* and 31.2% identity (44.2% similarity) with CYP105D5 of *Streptomyces coelicolor* A(3)2. As a result, several fatty acids were selected for the subsequent conversion study.

A.



B.

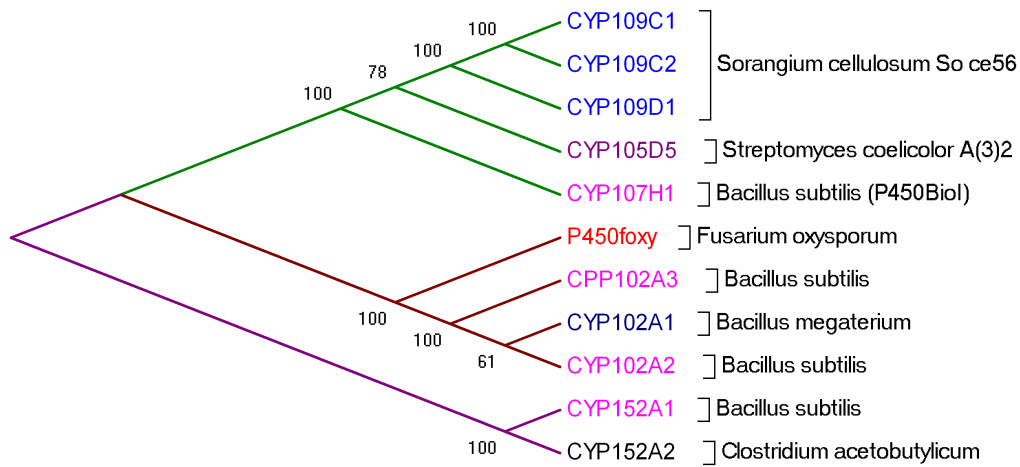


Figure II. A. 1. The radial view of best maximum-parsimony tree obtained by PHYLIP analysis. **A.** The amino acid sequences of fatty acid hydroxylating CYPs (CYP102A1, CYP102A2, CYP102A3, P450foxy, CYP152A1, CYP152A2, P450BioI and CYP105D5) were retrieved from EMBL database. The blue branches represent the cluster of So ce56 P450s related with P450BioI (CYP107H1) from *Bacillus subtilis* and CYP105D1 of *S. coelicolor* A(3)2. The brown lines represent the cluster of So ce56 with CYP102 family from *Bacillus sp* and P450foxy from *Fusarium oxysporum*. CYP260 family members (black branch), CYP259A1 and CYP117B1 (yellowish green) of So ce56 are not clustered with other fatty acid hydroxylases. **B.** The relatedness of CYP109 (CYP109C1, CYP109C2 and CYPD1) family member of So ce56 with fatty acid hydroxylases CYP105D5 and P450BioI (cluster with green branches). P450s from *Sorangium cellulosum* So ce56, *Streptomyces coelicolor* A(3)2, *Bacillus subtilis*, *Fusarium oxysporum*, *Bacillus megaterium* and *Clostridium acetobutylicum* are shown in blue, violet, magenta, red, dark blue and black, respectively.

A. 4. 2. Spin state shift and substrate binding

The differential spectroscopy study was firstly performed for CYP109D1 with long chain saturated fatty acids like lauric acid, myristic acid and palmitic acid. These fatty acids showed perturbation of the heme spectra of CYP109D1 with a shift of the resting Soret band from 420 nm towards a new position at ~390 nm, typical of the high spin form (Figure II. A. 2-4). The apparent dissociation constant (K_D) and the maximum absorbance shift (A_{\max}) obtained at the saturation of related fatty acids were calculated for CYP109D1. The calculated K_D and ΔA_{\max} were $1.84 (\pm 0.164) \mu\text{M}$ and $0.27 (\pm 0.004)$ respectively for lauric acid, with a calculated regression (R) of 0.99 (Figure II. A. 2). In case of myristic acid, the K_D was $2.02 (\pm 0.118) \mu\text{M}$ and ΔA_{\max} was $0.41 (\pm 0.005)$ with $R = 0.99$ (Figure II. A. 3). Likely, K_D and ΔA_{\max} for palmitic acid were $1.43 (\pm 0.110) \mu\text{M}$ and $0.38 (\pm 0.006)$, respectively, with $R = 0.99$ (Figure II. A. 4). The very low K_D values demonstrate the tight binding of these fatty acids with CYP109D1.

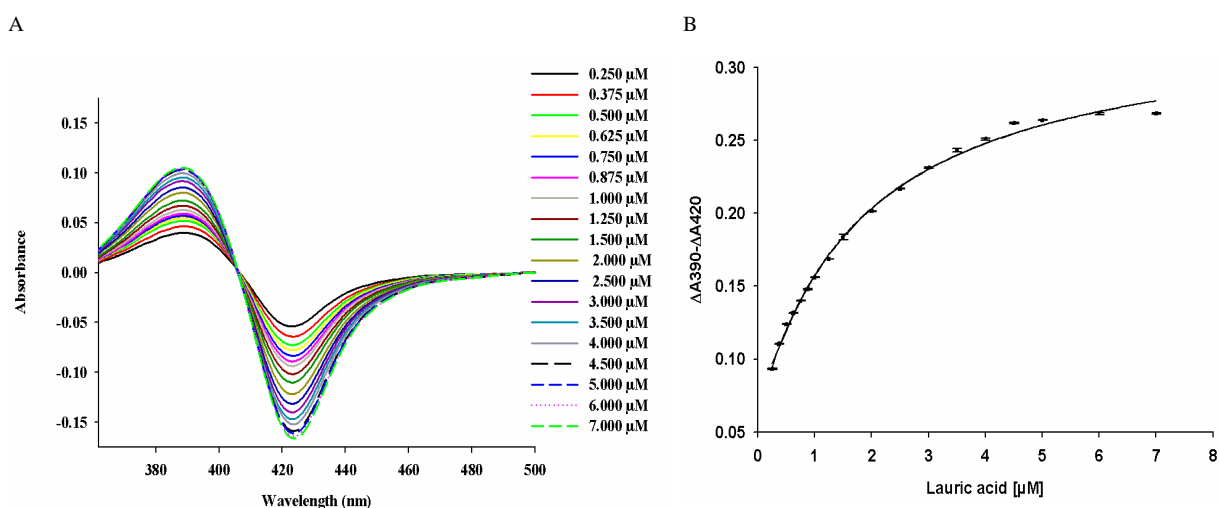


Figure II. A. 2: A. Spectral shift induced by binding of lauric acid to CYP109D1. Varying concentrations of lauric acid, dissolved in DMSO, were titrated in small (μl) aliquots to the solution of CYP109D1 ($5 \mu\text{M}$) in 10mM potassium phosphate buffer, pH 7.4 containing 20% glycerol, with the starting spectrum subtracted from the subsequent traces, so that the final concentrations were $0.250 \mu\text{M} - 7.000 \mu\text{M}$ in paired-tandem cuvettes. **B. The plot of absorbance change vs. concentration of the lauric acid.** The points were plotted to hyperbolic fit and the values for $K_D = 1.84 (\pm 0.164) \mu\text{M}$ and $\Delta A_{\max} = 0.27 (\pm 0.004)$ were obtained with regression (R) = 0.99 and standard error of estimate of 0.003.

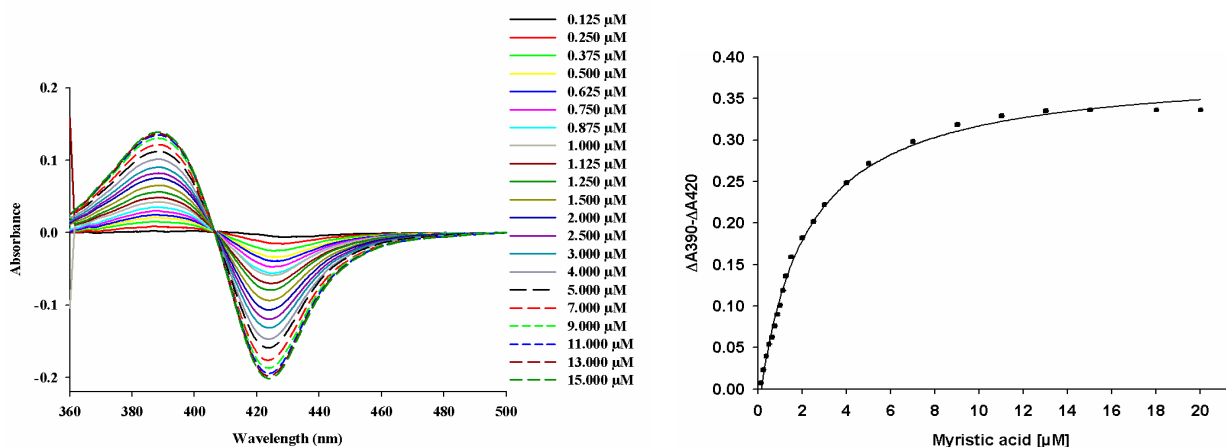


Figure II. A. 3: A. Spectral shift induced by binding of myristic acid to CYP109D1. Varying concentrations of myristic acid, dissolved in DMSO, were titrated in small (μl) aliquots to the solution of CYP109D1 ($5 \mu\text{M}$) in 10mM potassium phosphate buffer, pH 7.4 containing 20% glycerol, with the starting spectrum subtracted from the subsequent traces, so that the final concentrations were $0.125 \mu\text{M}$ - $15.000 \mu\text{M}$ in paired-tandem cuvettes. **B. The plot of absorbance change vs. concentration of the myristic acid.** The points were plotted to hyperbolic fit and the values for $K_D = 2.02 (\pm 0.118) \mu\text{M}$ and $\Delta A_{\text{max}} = 0.41 (\pm 0.005)$ were obtained with regression (R) = 0.99 and standard error of estimate of 0.006.

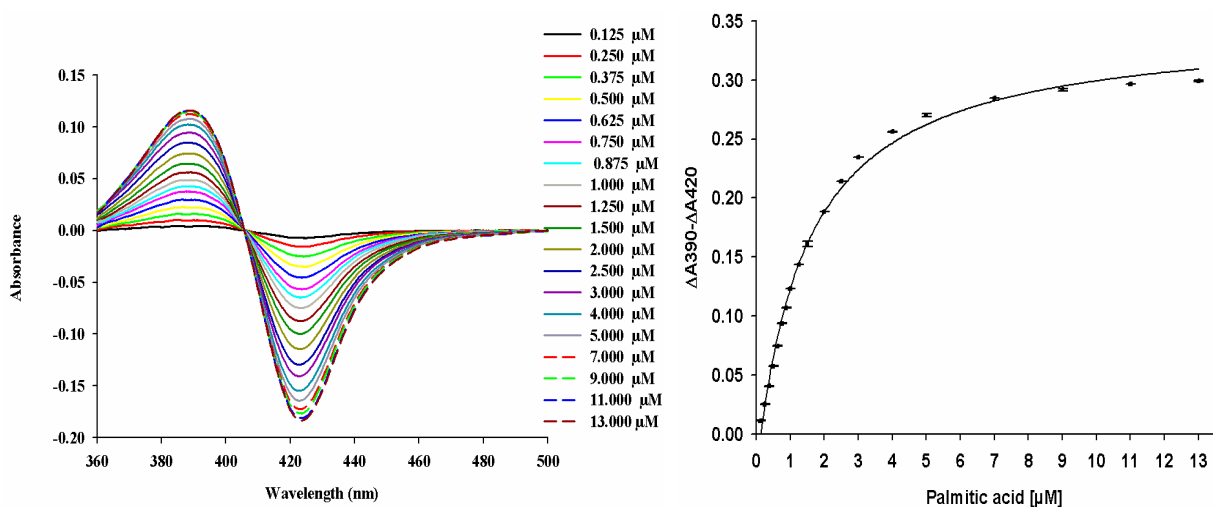


Figure II. A. 4: A. Spectral shift induced by binding of palmitic acid to CYP109D1. Varying concentrations of palmitic acid, dissolved in DMSO, were titrated in small (μl) aliquots to the solution of CYP109D1 ($5 \mu\text{M}$) in 10mM potassium phosphate buffer, pH 7.4 containing 20% glycerol, with the starting spectrum subtracted from the subsequent traces, so that the final concentrations were $0.125 \mu\text{M}$ - $13.000 \mu\text{M}$ in paired-tandem cuvettes. **B. The plot of absorbance change vs. concentration of the palmitic acid.** The points were plotted to hyperbolic fit and the values for $K_D = 1.43 (\pm 0.110) \mu\text{M}$ and $\Delta A_{\text{max}} = 0.38 (\pm 0.006)$ were obtained with regression (R) = 0.99 and standard error of estimate of 0.007.

A. 4. 3. Characterization of different members of CYP109 family of *So ce56*

Two sequences of CYP109, P27632 and AF015825 from *Bacillus subtilis* were reported long before (Ahn and Wake, 1991 and Kunst, *et al.*, 1997) but their substrates have not been

identified so far. As this work was the very first to illustrate the novel properties of a CYP109 family member as a fatty acid hydroxylase, a detailed study was necessary. The percentage similarities among all CYP109 members were calculated and are illustrated in Table II. A. 3. All the subfamilies of CYP109 from *Bacillus subtilis* and *So ce56* showed more than 50% similarity except for CYP109D1 of *So ce56* and CYP109B (AF015825) of *B. subtilis*.

Table II. A. 3: The percentage similarities among the different members of the CYP109 families. The values were obtained from the pair wise sequence alignment and the values are expressed in percentage similarity of amino acid sequences.

		<i>Bacillus subtilis</i>		<i>Sorangium cellulosum</i> (<i>So ce56</i>)		
		CYP109A (P27632*)	CYP109B (AF015825*)	CYP109C1	CYP109C2	CYP109D1
<i>Bacillus subtilis</i>	CYP109A (P27632*)	-	61.6	53.3	55.9	52.1
	CYP109B (AF015825*)	61.6	-	55.5	58.3	46.6
<i>So ce56</i>	CYP109C1	53.3	51.5	-	79.3	51.2
	CYP109C2	55.9	58.3	79.3	-	52.3
	CYP109D1	52.2	46.6	51.2	52.3	-

NOTE: * gene number in EMBL database

A. 4. 4. Fatty acid hydroxylation and characterization

The *in vitro* conversion of fatty acids catalyzed by CYP109D1 were performed in a reconstitution system with heterologous redox partners, Adx-AdR, and were measured by GC-MS after derivatization assuming that all the derivatized products were extracted with an equal efficiency. The substrate and products were identified by their characteristic fragmentation patterns. The possible derivatized masses for the major fragments were calculated by using ChemDraw Ultra 11.0 and were confirmed with an authentic library and the archived MS spectra of trimethylsilyl-(TMS) methyl-hydroxy fatty acids on <http://www.lipidlibrary.co.uk>. An overview of the characteristic fragments for the products of the investigated saturated fatty acid is illustrated in Table II. A. 4.

Table II. A. 4: The characteristic derivatized units (sizes) for the silylated hydroxylation products from the different fatty acids used in this work. The sub-terminal ω -hydroxylations of saturated, unsaturated and branched fatty acids are shown.

Substrate		ω -0	ω -1	ω -2	ω -3	ω -4	ω -5	ω -6	ω -7
Saturated fatty acid	Capric acid (C10)	-	117, 317	131 303	-	-	-	-	-
	Lauric acid (C12)	-	117, 345	131, 331	145, 317	-	-	-	-
	Tridecanoic acid (C13)	-	117, 359	131, 345	145, 331	159, 317	173, 303	-	--
	Myristic acid (C14)	-	117, 373	131, 359	145, 345	159, 331	173, 317	-	-
	Palmitic acid (C16)	-	117, 401	131, 387	145, 373	159, 359	173, 345	187, 331	-
Unsaturated fatty acid	Oleic acid (18:1 cis9)	-	-	-	-	-	-	187, 357	201, 343
Branched fatty acid	13-Methyl myristic acid	-	-	145, 359	159 345	173, 331	-	-	-
	15-Methyl palmitic acid	-	-	145, 387	159, 373	173, 359	187, 345	201, 331	-
		-	-						
		-	-						

A. 4. 4. 1. Hydroxylation of saturated fatty acid by CYP109D1

Five saturated fatty acids of carbon-length C10 to C16 were tested in this work, including capric acid (C10), lauric acid (C12), tridecanoic acid (C13), myristic acid (C14) and palmitic acid (C16).

As shown in Table II. A. 4 CYP109D1 was able to convert all the tested saturated fatty acids. They were hydroxylated at sub-terminal positions. The terminal hydroxylation was not obtained. Capric acid was hydroxylated at ω -1 (RT 8.11 min) and ω -2 (7.91 min) positions. Lauric acid gave 3 hydroxylated products at ω -1 (RT 9.86 min), ω -2 (RT 9.67 min) and ω -3 (RT 9.30 min). Tridecanoic acid was hydroxylated at ω -1 (RT 15.16 min), ω -2 (RT 15.04 min), ω -3 (RT 14.79 min), ω -4 (RT 14.65 min) and ω -5 (RT 14.55). Myristic acid was also

converted into 5 hydroxylated products of ω -1 (RT 9.36 min), ω -2 (RT 9.20 min), ω -3 (RT 8.88 min), ω -4 (RT 8.68 min) and to ω -5 (RT 8.57). Similarly, palmitic acid was converted to 7 hydroxylated products at ω -1 (RT 11.61 min), ω -2 (RT 11.41 min), ω -3 (RT 14.02 min), ω -4 (RT 10.84 min), ω -5 (RT 10.67) and ω -6 (RT 10.57) (Figure II. A. 5). The characteristic fragmentation patterns of the silylated hydroxylation products of the investigated saturated fatty acids were assigned in each of the mass spectra in (Figure II. A. 6. 1-5). During product analysis, mainly two major fragments have been considered and the derivatization is shown in the inset of each mass spectra.

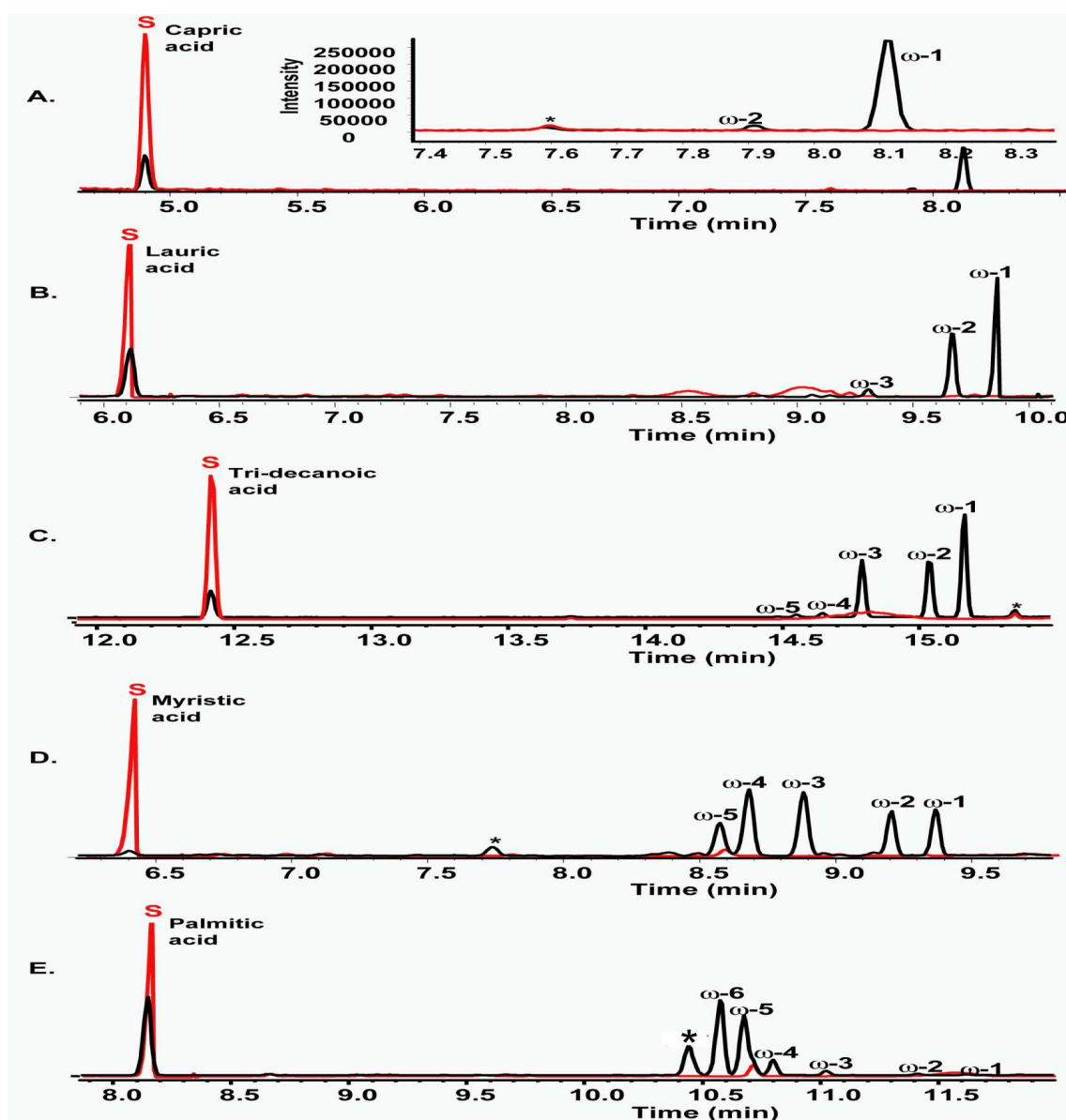


Figure II. A. 5: GC chromatogram of subterminal ω -hydroxylation of saturated fatty acids by CYP109D1. A-E represents the hydroxylation of capric acid, lauric acid, tridecanoic acid, myristic acid and palmitic acid, respectively. The red and black lines represent the negative and the product, respectively. Substrate (S) and ω -hydroxylation are shown. The ‘*’ represents non specific product.

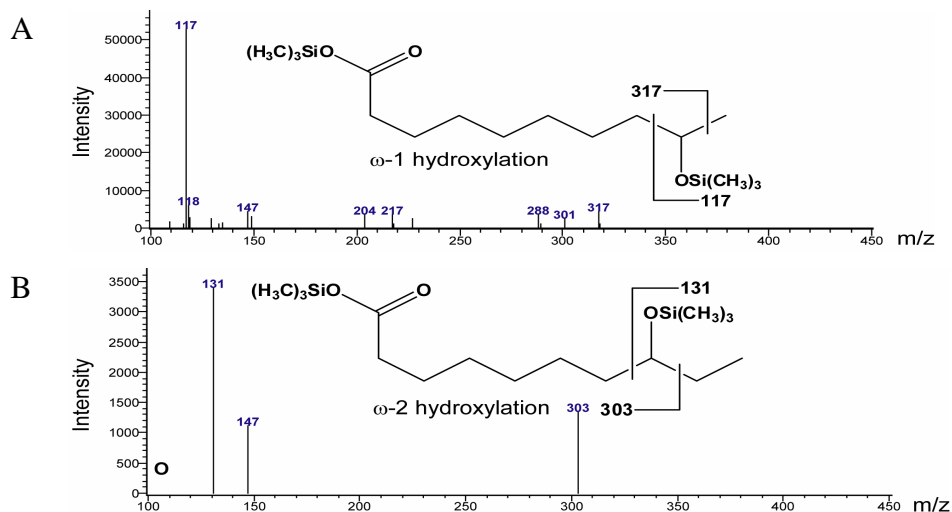


Figure II. A. 6. 1: Mass spectra of the hydroxylated products of capric acid by CYP109D1. The MS of A. ω -1 and B. ω -2 hydroxylated products of capric acid are shown. In the inset the derivatization pattern of each of the ω -products of capric acid are shown.

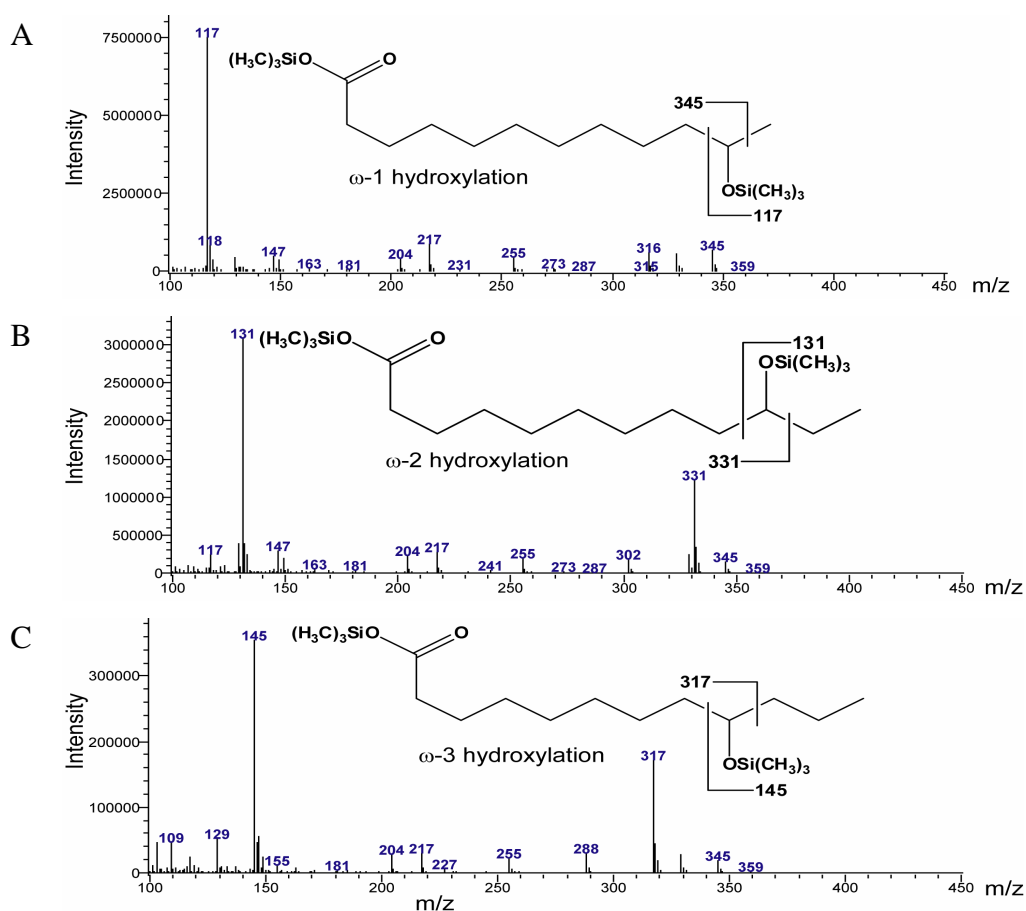


Figure II. A. 6. 2: Mass spectra of the hydroxylated products of lauric acid by CYP109D1. The MS of A. ω -1, B. ω -2 and C. ω -3 hydroxylated products of lauric acid are shown. In the inset the derivatization pattern of each of the ω -products of lauric acid are shown.

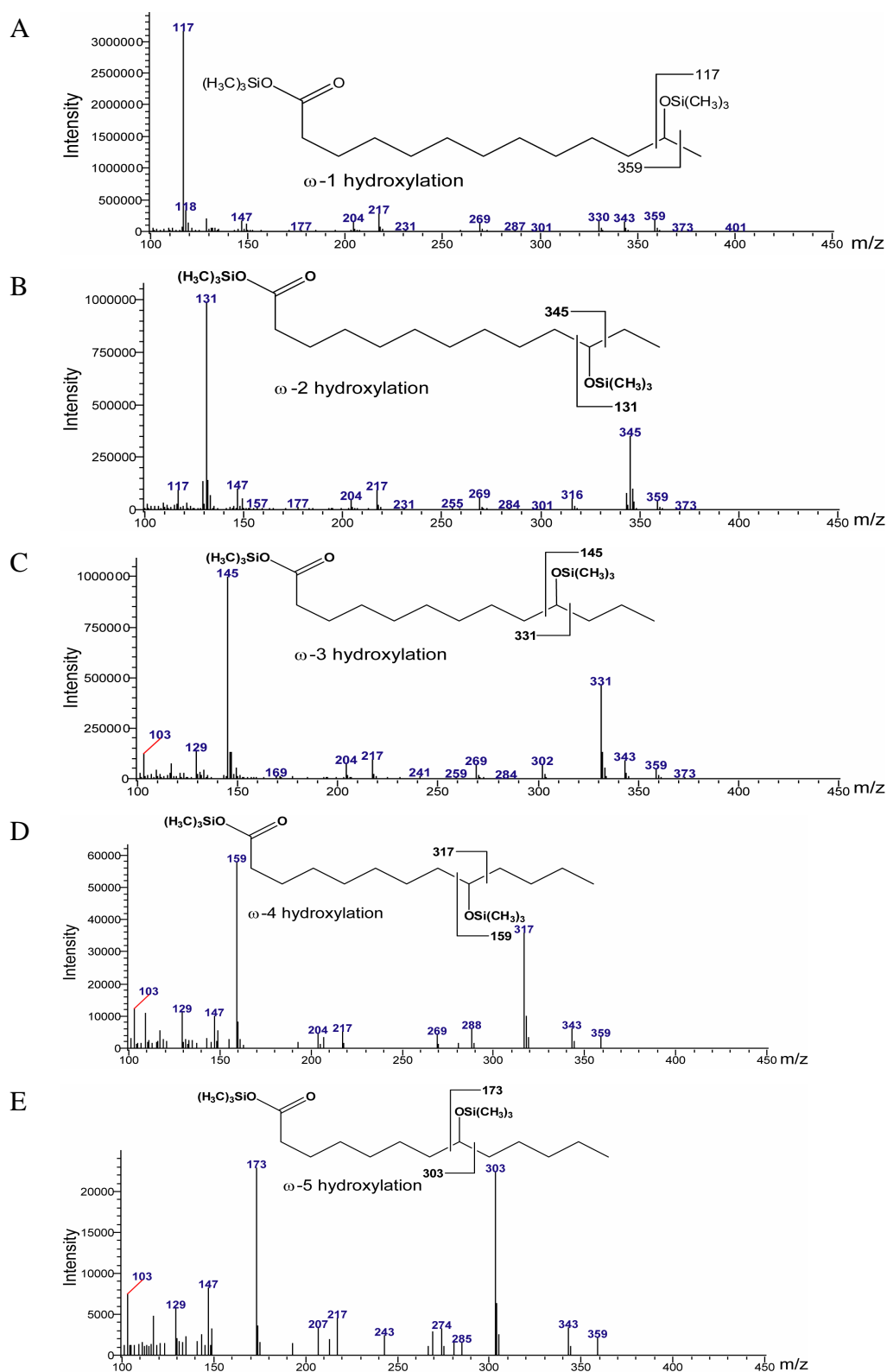


Figure II. A. 6. 3: Mass spectra of the hydroxylated products of tridecanoic acid by CYP109D1. The MS of A. ω -1, B. ω -2, C. ω -3, D. ω -4 and E. ω -5 hydroxylated products of tridecanoic acid are shown. In the inset the derivatization pattern of each of the ω - products of tridecanoic acid are shown.

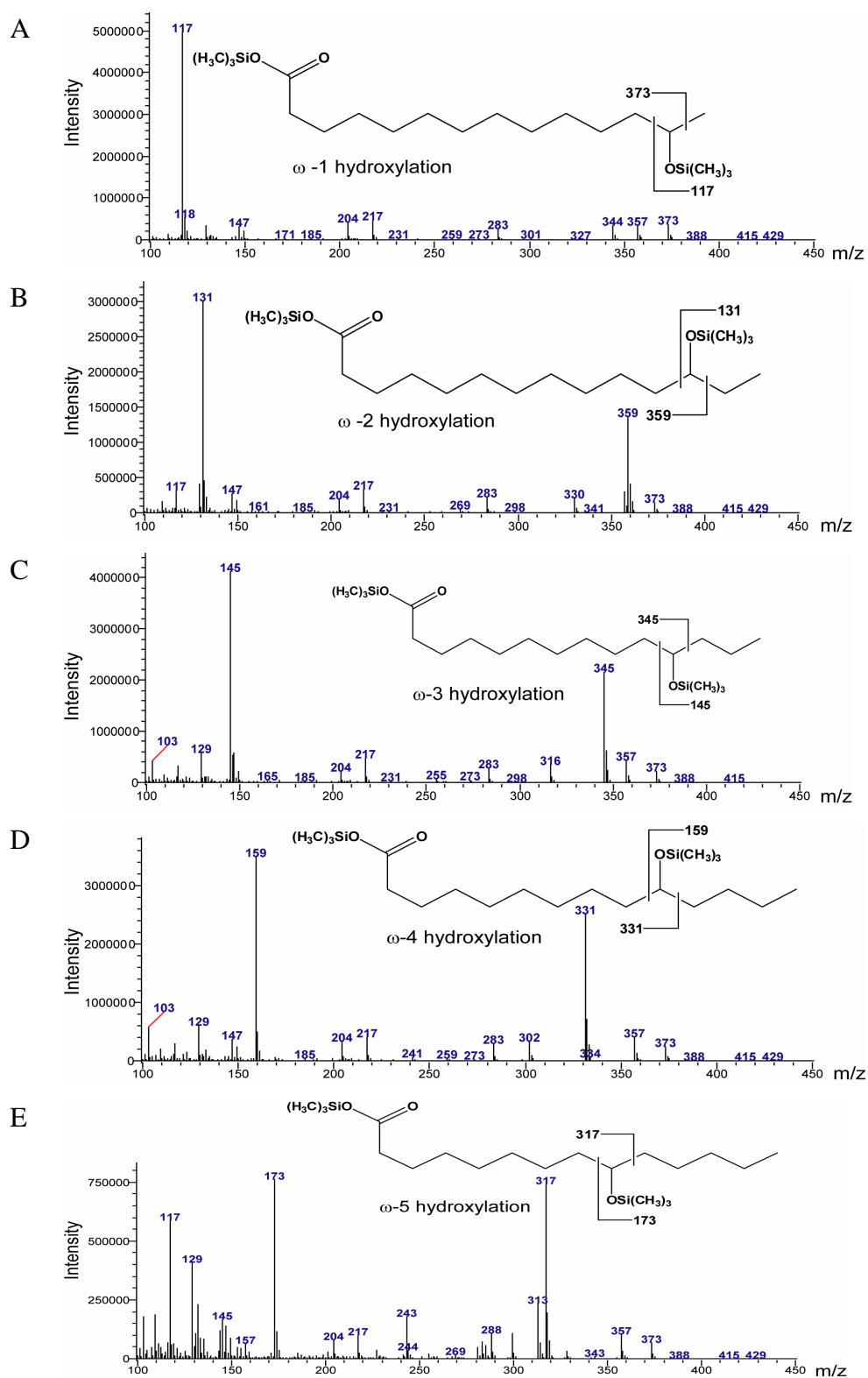


Figure II. A. 6. 4: Mass spectra of the hydroxylated products of myristic acid by CYP109D1. The MS of A. ω-1, B. ω-2, C. ω-3, D. ω-4 and E. ω-5 hydroxylated products of myristic acid are shown. In the inset the derivatization pattern of each of the ω-products of myristic acid are shown.

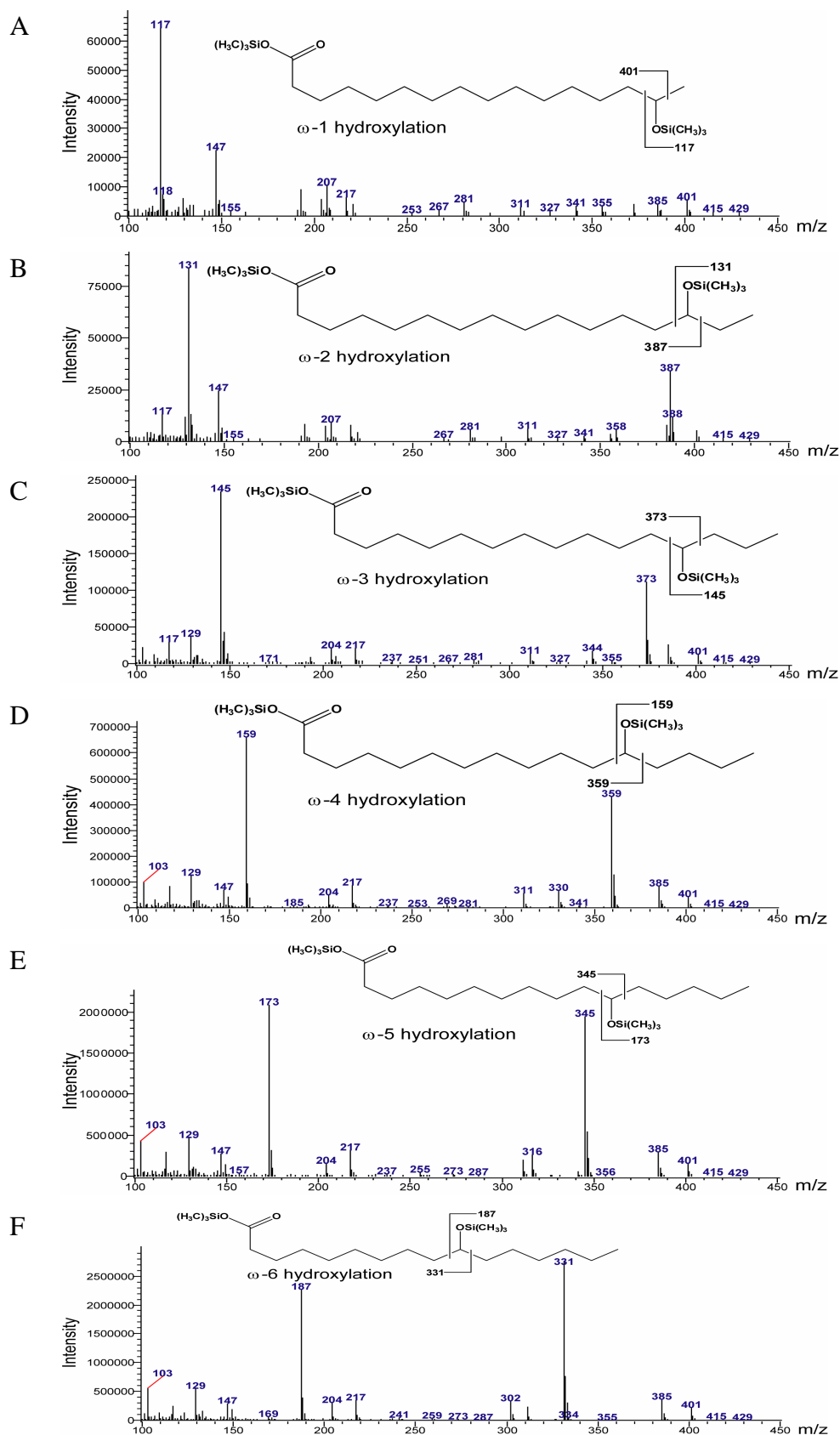


Figure II. A. 6. 5: Mass spectra of the hydroxylated products of palmitic acid by CYP109D1. The MS of A. ω-1, B. ω-2, C. ω-3, D. ω-4, E. ω-5 and F. ω-6 hydroxylated products of palmitic acid are shown. In the inset the derivatization pattern of each of the ω-products of palmitic acid are shown.

A. 4. 5. Comparison of hydroxylation patterns of saturated fatty acids obtained by other CYP109 members of *So ce56*

The other two members of the CYP109 family, CYP109C1 and CYP109C2, were also used for saturated fatty acid hydroxylation. Both P450s (CYP109C1 and CYP109C2) were able to get electrons from the heterologous redox proteins, Adx and AdR (Chapter I). The spin shift experiment with corresponding saturated fatty acids for CYP109C1 and CYP109C2 was not done. Since CYP109C1 and CYP109C2 were clustered with CYP109D1 (Figure II. A. 1), fatty acids were selected as putative substrates and their conversion was studied directly. The GC-MS analysis was done and was compared with the hydroxylation pattern of CYP109D1. It was found that CYP109C1 was not able to convert any of the tested fatty acids. CYP109C2 has converted all the tested saturated fatty acids into their corresponding hydroxylated products, except for capric acid (C10) (Table II. A. 5). It was found that CYP109C2 mediated hydroxylation reactions were less efficient but more selective than those of CYP109D1. CYP109C2 was able to convert lauric acid and tridecanoic acid at ω -1 to ω -3, myristic acid at ω -1 to ω -2 and palmitic acid only at ω -1 position. The details of hydroxylated products, and the corresponding retention times are demonstrated in Table II. A. 5. CYP109C2 catalyzed reaction showed that the sub terminal ω -1 hydroxylated products (area percent) were always dominating (Table II. A. 5). It was found that the hydroxylation pattern by CYP109C1 was more selective towards the sub-terminal end for the saturated fatty acids. It seemed that the hydroxylation pattern of CYP109C2 is not depending on the chain length.

However, the hydroxylating reaction of CYP109D1 was much more efficient than that of CYP109C2 (Table II. A. 5). It was found that, CYP109D1 has converted ~57% of capric acid into ω -1 and ω -2 hydroxylated products. Likely, 72% of the lauric acid was converted to ω -1 to ω -3, 88% tridecanoic acid to ω -1 to ω -4, 98% myristic acid to ω -1 to ω -4, and 71% of palmitic acid to ω -1 to ω -5. It was noticed that the hydroxylation pattern of CYP109D1 depends on the chain length. In moderate short chain fatty acids like capric acid, lauric acid and tridecanoic acid, the sub-terminal ω -1 hydroxylation product is predominating and subsequently decreasing towards higher ω -position with the product pattern of ω -1 > ω -2 > ω -3 > ω -4 hydroxylated product (Figure II. A. 5 and Table II.5). In case of long chain fatty acids like myristic acid and palmitic acid the sub-terminal hydroxylations were not consistent rather the higher ω -position of hydroxylations was more predominant. In case of myristic acid, 98% of the substrate was converted into ω -1 to ω -5 hydroxylated products where ω -3 and ω -4 has 27 and 26% product predominance, respectively. Likely, for palmitic acid, 72%

of the substrate was converted into ω -1 to ω -7 hydroxylated products where 25% of the product was predominant in ω -5 and 27% in ω -6 position (Table II. A. 5). These results showed that CYP109D1 was much more efficient for saturated fatty acid hydroxylation than CYP109C2.

A. 4. 6. Hydroxylation of other fatty acids by CYP109D1

As CYP109C1 was not able to hydroxylate any of the tested saturated fatty acids and CYP109C2 has very low activity, only CYP109D1 was used to test the hydroxylation of unsaturated and branched fatty acids.

A. 4. 6. 1. Hydroxylation of unsaturated fatty acid by CYP109D1

The unsaturated fatty acid oleic acid was used in this work. CYP109D1 was able to convert the oleic acid (RT 13.67 min) into 2 hydroxylated product of ω -7 (RT 15.34 min) and ω -6 (RT 15.64 min) (reconstitution with Adx and AdR) (Figure A. 7). During conversion, 56% of the oleic acid (200 μ M) was converted into ω -7 and ω -6 products, where the ω -7 product was much predominated (50%). The characteristic fragmentation patterns of the oleic acid and its products are illustrated in Figure. A. 8.

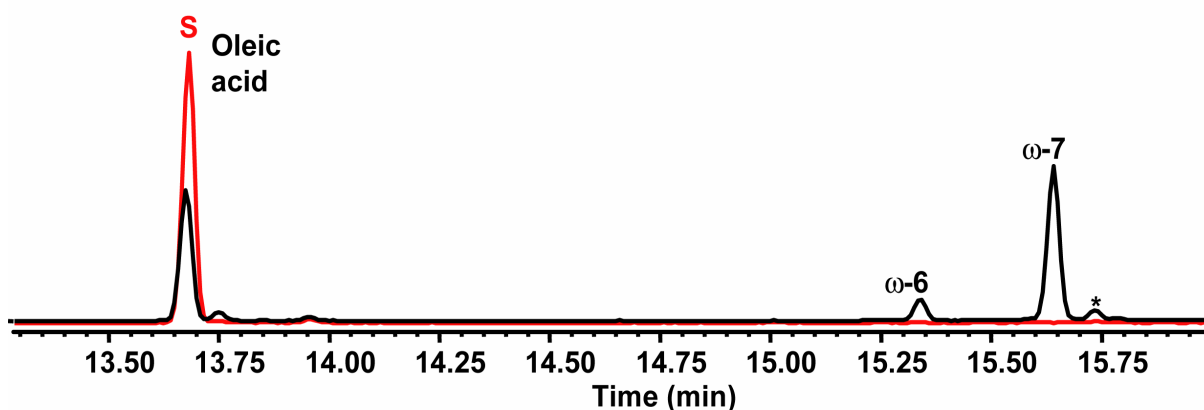


Figure II. A. 7: GC chromatogram of oleic acid hydroxylation by CYP109D1. The red and black line represents the negative control and conversion, respectively. 'S' represents the oleic acid and the '*' represents the impurity. The ω -6 and ω -7 hydroxylated product peaks are shown.

Table II. A. 5: The product profiles of saturated fatty acids by the members of CYP109 (CYP109C1, CYP109C2 and CYP109D1) of *So ce56*. The comparison of hydroxylation position, respective retention time (RT) and the area percentage of the substrate and the sub terminal ω -hydroxylated products for CYP109D1, CYP109C1 and CYP109C2 are illustrated. CYP109C1 has not shown any hydroxylation activity.

P450s of <i>So ce56</i>	Substrate RT/Area %	ω -7	ω -6	ω -5	ω -4	ω -3	ω -2	ω -1	Total Conversion %
		Retention time (RT) in min / Area %							
Capric acid									
CYP109C1	-	-	-	-	-	-	-	-	NC*
CYP109C2	-	-	-	-	-	-	-	-	NC*
CYP109D1	4.89/ 43.03	-	-	-	-	-	7.91/ 2.80	8.11/ 54.17	56.97
Lauric acid									
CYP109C1	-	-	-	-	-	-	-	-	NC*
CYP109C2	6.12/ 95.97	-	-	-	-	9.30/ 0.24	9.67/ 1.46	9.86/ 2.33	4.03
CYP109D1	6.12/ 27.74	-	-	-	-	9.30/ 3.93	9.67/ 31.28	9.86/ 37.05	72.26
Tridecanoic acid									
CYP109C1	-	-	-	-	-	-	-	-	NC*
CYP109C2	12.42/ 91.33	-	-	-	-	14.79/ 0.93	15.04/ 2.31	15.16/ 5.43	8.67
CYP109D1	12.42/ 11.66	-	-	14.55/ 1.07	14.65/ 1.71	14.79/ 22.19	15.04/ 23.74	15.16/ 39.63	88.34
Myristic acid									
CYP109C1	-	-	-	-	-	-	-	-	NC*
CYP109C2	6.40/ 84.92	-	-	-	-	-	9.20/ 3.11	9.36/ 11.97	15.08
CYP109D1	6.40/ 1.80	-	-	8.57/ 12.18	8.68/ 25.61	8.88/ 26.58	9.20/ 15.53	9.36/ 16.95	98.2
Palmitic acid									
CYP109C1	-	-	-	-	-	-	-	-	NC*
CYP109C2	8.14/ 88.00	-	-	-	-	-	-	10.70/ 12.00	12.00
CYP109D1	8.14/ 28.56	10.44/ 11.97	10.57/ 27.60	10.67/ 24.77	10.80/ 4.93	11.02/ 1.44	11.41/ 0.55	11.61/ 0.21	71.44

Note: NC* - No conversion

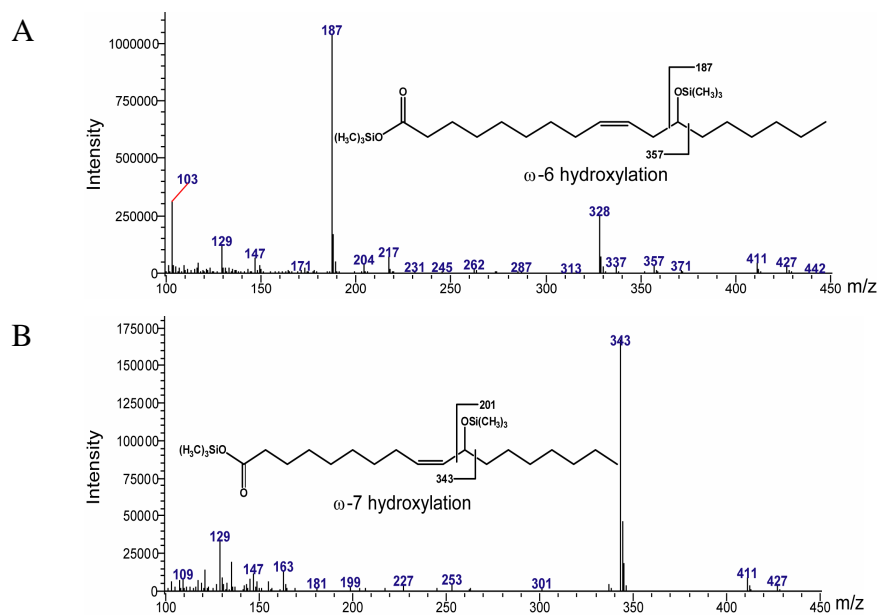


Figure II. A. 8: Mass spectra of the hydroxylated products of oleic acid by CYP109D1. The MS of A. ω -6 and B. ω -7 hydroxylated products of oleic acid are shown. In the inset the derivatization pattern of each of the ω -products of oleic acid are shown.

A. 4. 6. 2. Hydroxylation of branched fatty acid by CYP109D1

In this work, the branched fatty acids, 13-methyl myristic acid and 15-methyl palmitic acid were used as substrates for CYP109D1 after reconstitution with Adx and AdR as redox partners. CYP109D1 was able to convert 13-methyl myristic acid at ω -2 (RT 12.95 min), ω -3 (RT 12.47min) and ω -4 (RT 13.37min) (Table II. A. 6). In case of 15-methyl palmitic acid, the hydroxylation also occurred at sub-terminal positions at ω -2 (RT 15.10 min), ω -3 (RT 14.63 min), ω -4 (RT 14.52) and ω -5 (RT 14.48) and ω -6 (14.45 min). The retention time and the area percentage of the substrate and the respective hydroxylation products are illustrated in Table II. A. 6. The characteristic fragmentation patterns are illustrated in Figure II. A. 9-10. The total conversion of both the fatty acids was below 15%. The terminal hydroxylation was not detected.

Table II. A. 6: Product profile of branched fatty acids with CYP109D1. Hydroxylation positions with respective retention time and area percentage of the substrate and product for CYP109D1 are illustrated.

Substrate	ω -6	ω -5	ω -4	ω -3	ω -2	ω -1	Total Conversion %
	Retention time (RT) in min/ Area %						
13-methyl myristic acid							
10.12/ 98.00	-	-	12.37/ 0.59	12.47/ 0.49	12.95/ 0.87	-	2.00
15-methyl palmitic acid							
12.4/ 84.90	14.45 2.01	14.48/ 8.18	14.52/ 0.81	14.63/ 0.98	15.10/ 0.75	-	15.1

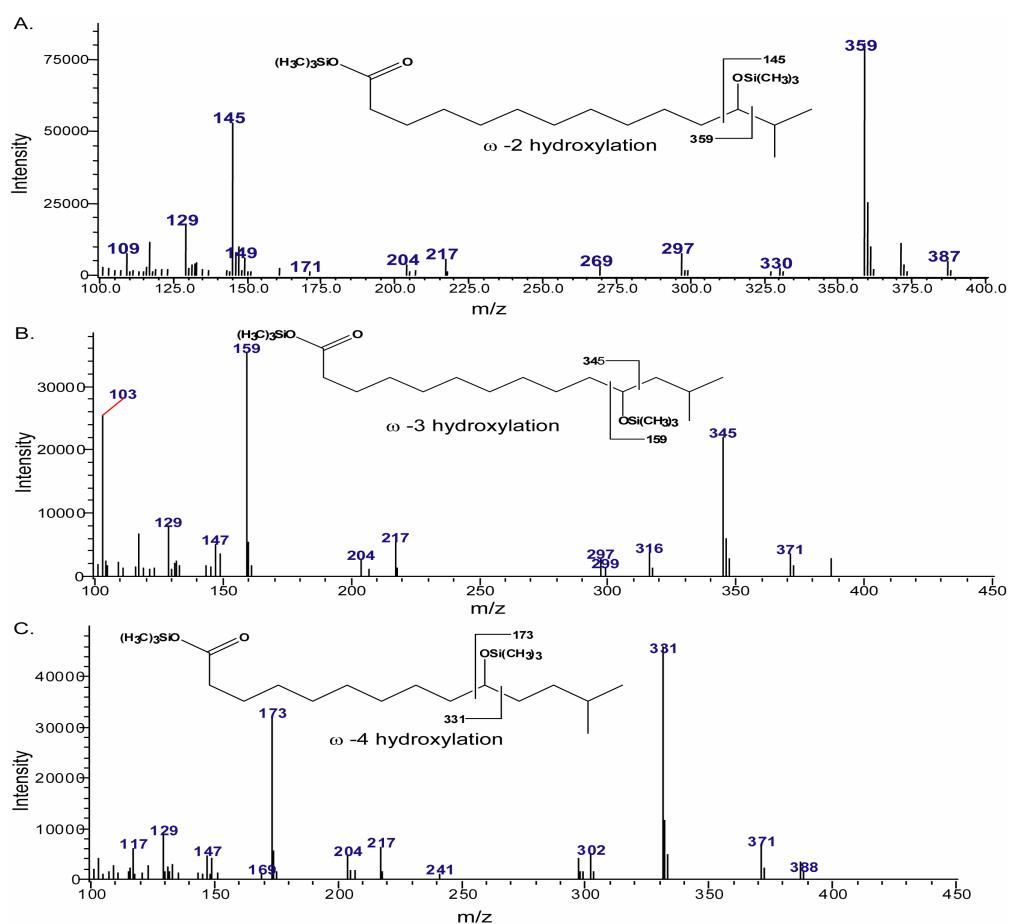


Figure II. A. 9: Mass spectra of the hydroxylated products of 13-methyl myristic acid by CYP109D1. The MS of A. ω -2, B. ω -3 and C. ω -4 hydroxylated products of 13-methyl myristic acid are shown. In the inset the derivatization patterns of each of the ω -products of 13-methyl myristic acid are shown.

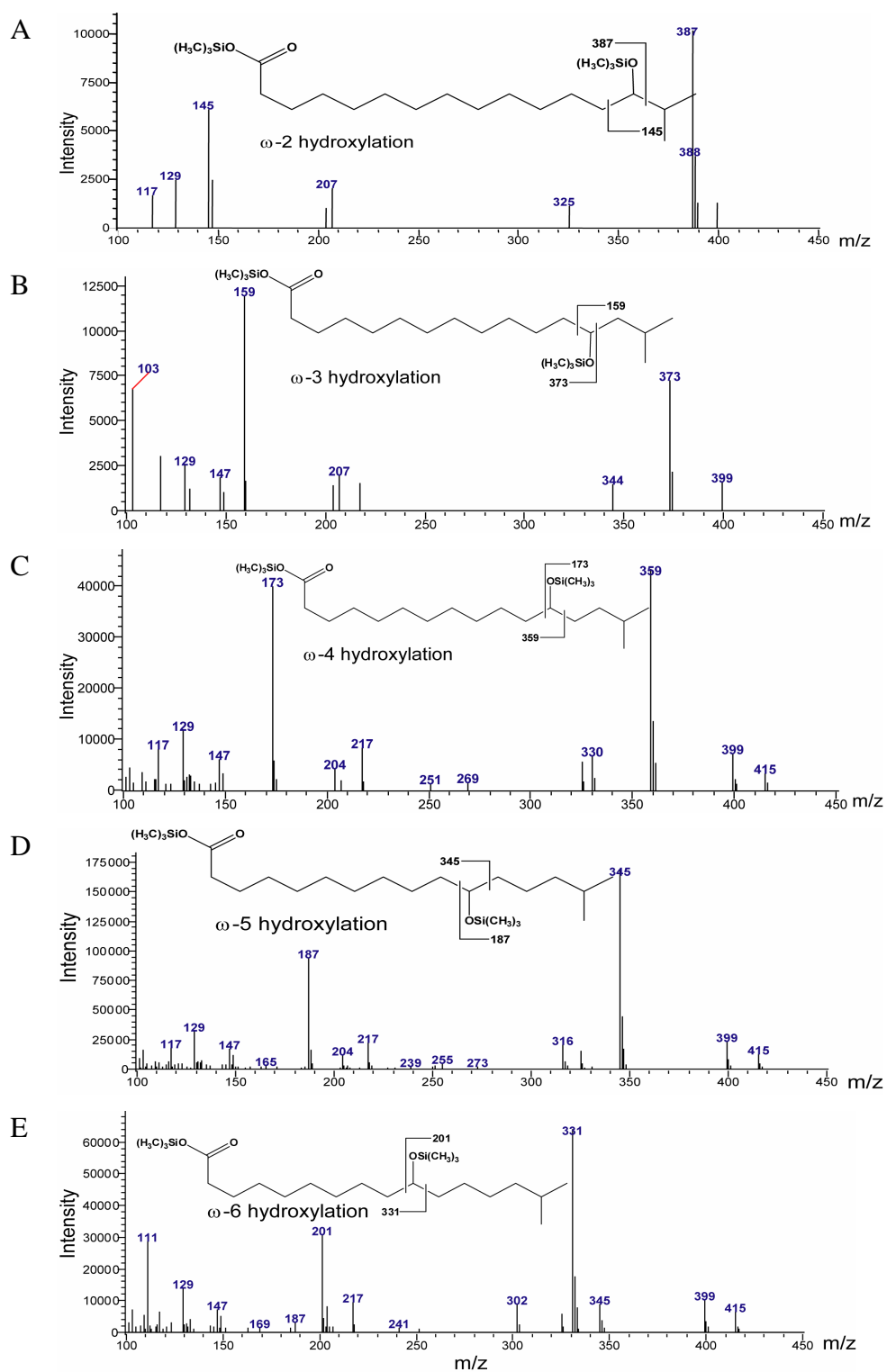


Figure II. A. 10: Mass spectra of the hydroxylated products of 15-methyl palmitic acid by CYP109D1. The MS of A. ω-2, B. ω-3, C. ω-4, D. ω-5 and E. ω-6 of hydroxylated products of 15-methyl palmitic acid are shown. In the inset the derivatization pattern of each of the ω-products of 15-methyl palmitic acid are shown.

A. 5. Discussion

Fatty acids occur widely in nature as metabolic intermediates or secondary metabolites. In bacteria, the straight chain fatty acids lauric acid, myristic acid, palmitic acid, stearic acid, hexadecanoic acid, octadecanoic acid, cyclopropanoic acid, 10-methyl hexadecanoic acid and 2- or 3-hydroxy fatty acids are commonly found (Lechevalier, 1977; Siristova *et al.*, 2008). The occurrence and functional role of hydroxylated long chain fatty acids in animals, plants and fungi are well established (Dyk *et al.*, 1994; Lösel, 1988). Many of these compounds have potential physiological activities and play an important role in the cell signaling process. A large number of fungi that produce hydroxylated long chain fatty acids are plant pathogens. However, much less is known about the function of hydroxy fatty acids in bacteria. It was shown that *Nocardia cholesterolicum* and *Flavobacterium* DS5 convert oleic acid to 10-hydroxy stearic acid and linoleic acid to 10-hydroxy-12 (Z)-octadecanoic acid which act as bioactive compounds (Hou, 2008). In *Mycobacterium tuberculosis* and *Mycobacterium leprae* 3-hydroxylated long chain fatty acids play a crucial role in the envelope architecture (Rafidinarivo *et al.*, 2008). Though the hydroxylated fatty acids govern some vital functions in bacteria, the real mechanisms of the production were not clear. However, it has been shown that the terminal and sub-terminal hydroxylation of long chain fatty acids are generally ascribed to the action of cytochromes P450 (Boddupalli *et al.*, 1990; Li *et al.*, 2000) and thus obtained fatty acid derivatives speculated to have some physiological role.

It has been shown that the fatty acids and their derivatives are of prime concern in the life cycle of myxobacteria (Shimkets, 1999; Curtis *et al.*, 2006; Rosenbluh and Dworkin, 1996; Gerth *et al.*, 2003; Guo and Tao, 2008; Ahn *et al.*, 2008). Considering the physiological significance of lipids in myxobacteria, fatty acids were assumed as putative substrates for some of the P450s of *So ce56*. Though several homologues of fatty acid hydroxylases were identified, only the enzymes of the CYP109 family (CYP109C1, CYP109C2 and CYP109D1) were selected for further studies in this dissertation because of their higher expression level (Chapter I) and identified redox partners (Chapter I). The putative substrates, fatty acids (lauric acid, myristic acid and palmitic acid), were firstly used to characterize the binding spectra with CYP109D1, the most distant P450 among three members of CYP109 family (CYP109C1, CYP109C2 and CYP109D1) (Figure II. A. I) and showed a tight binding. Beside this, CYP109D1 also hydroxylated capric acid into ω -1 and ω -2, lauric acid to ω -1 to ω -3, tridecanoic acid to ω -1 to ω -5, myristic acid to ω -1 to ω -5, and palmitic acid to ω -1 to ω -6. Further on, the other two CYP109 members (CYP109C1 and CYP109C2) of *So ce56* were

used to study the *in vitro* conversion of saturated fatty acids. Under identical experimental condition, CYP109C1 was not able to convert any of the tested saturated fatty acids. However, the sub-terminal hydroxylated products of saturated fatty acids were obtained in case of CYP109C2 although the area percentage of the products was much lower than that with CYP109D1. It could be that either the predicted substrates (fatty acids) are not good substrates for CYP109C2 or that the redox partners might not be efficient enough with this P450. Our data show that the three P450s (CYP109C1, CYP109C2 and CYP109D1), despite being same family members and homologues to each other, were not able to perform the same reactivity under identical reaction conditions. An application of theoretical studies such as docking experiments should be the next stage in gaining a better understanding of the effect of fatty acids on CYP109C1, CYP109C2 and CYP109D1. An investigation on the hydroxylation of saturated fatty acid was also done for CYP105D5 from *S. coelicolor* A(3)2 where ω -1, ω -2, ω -3 and ω -4 hydroxylation of lauric acid and ω -1 hydroxylation of oleic acid (Chun *et al.*, 2007) were described. Likewise, P450BioI (CYP107H1) from *Bacillus subtilis* was also shown to catalyze the ω -hydroxylation of myristic acid (Green *et al.*, 2001). In both the cases the distinct physiological role is not clear. Similarly, even for the well studied fatty acid hydroxylase P450 BM-3 (CYP102A1) the physiological role is currently unclear. However, Wolf's group speculated that the role of P450BM-3 is to protect *Bacillus megaterium* by removing toxic fatty acids from the environment (Gustafsson *et al.*, 2001). Likewise, the hydroxylation of fatty acids by CYP102A1 was also shown (Capdevila *et al.*, 1996) but the physiological role again was not clear. In this study, though the clear physiological role of the products of hydroxylated saturated fatty acids during the life cycle of *So ce56* is not clear, it is speculated that the hydroxylated products obtained by P450s could successively be converted into related alcohols and carboxylic acids which could have important role during normal homeostasis or during other biosynthetic process.

In our study, the hydroxylation of fatty acid by CYP109D1 was further extended by using mono-unsaturated fatty acid and branched-fatty acid. CYP109D1 was able to convert oleic acid (the monounsaturated omega-9-fatty acid) into the ω -6 hydroxylated product (12-hydroxy-9-octadecenoic acid) also called as ricinoleic acid which has various physiological and biotechnological significances. It is reported that 12-hydroxyoctadecanoic, ricinoleic and other fatty acids can enhance the production of greenish-yellow pigment (identified as phenazine-1-carboxylic acid) in *Pseudomonas* species (Keudell *et al.*, 2000). As *Sorangium cellulosum* *So ce56* are also yellowish-orange colored, the importance of such hydroxylated

product by P450 could be a matter of further studies. It was shown that the ricinoleic acid can also undergo many familiar organic reactions to form useful derivatives like epoxidated, hydrogenated and hydroxylated forms (Park *et al.*, 2004). So it is assumed that CYP109D1 possibly could have an important role to produce some physiologically important derivatives from unsaturated fatty acids during the lifecycle of *So ce56*. Moreover, ricinoleic acid could have some biotechnological applications. The presence of hydroxyl- and carboxyl-groups as well as Δ -9 unsaturation make ricinoleic acid beneficial as an important precursor for the synthesis of useful bio-materials like biopolymers (Ebata *et al.*, 2007). It was also reported that ricinoleic acid can be enzymatically polymerized to yield a higher molecular weight polyricinoleate, which showed an excellent low temperature flexibility and elasticity to a thermosetting elastomer (Ebata *et al.*, 2007).

Branched chain fatty acids also have special importance in microbes. Monomethyl branched-chain fatty acids (mmBCFAs) in *iso*- or in an *anteiso*- configuration are found in various bacteria including cold tolerant and thermophilic species (Merkel and Perry, 1977; Annous *et al.*, 1997; Batrakov *et al.*, 2000; Jahnke *et al.*, 2001, Groth *et al.*, 2002, Nichols *et al.*, 2002, Ferreira, *et al.*, 1997). In most of the cases, mmBCFs play a contributing role to the membrane function, regulating fluidity (Rilfors *et al.*, 1978; Suutari and Laakso, 1994; Cropp *et al.*, 2000; Jones *et al.*, 2002) and protein permeability (van de Vossenberg, 1999). The branched fatty acids are also responsible for the pathogenic role in some microbes, e.g. mycocerosic acid present in the cell wall of the pathogenic mycobacterium is responsible for the growth and pathogenesis (Gago, *et al.*, 2006). The role of branched-chain fatty acids during the tolerance of pH stress in *Listeria monocytogenes* was also studied and it was shown that a higher proportion of branched fatty acid is produced during alkaline condition for the adaptability (Giotis *et al.*, 2007). Along with this, it has been shown that even for the growth and development of the eukaryotic cell *Caenorhabditis elegans*, mono-methyl branched fatty acids became essential (Kniazeva *et al.*, 2004). Concerning this, in several myxobacterial studies the pronounced role of branched chain fatty acid has been documented during gliding motility, chemotaxis and fruiting body formation. The fatty acids present in the total hydrolysates of several gliding bacteria (*Myxococcus fulvus*, *Stigmatella aurantiaca*, *Cytophaga johnsonae*, *Cytophaga* sp. strain Samoa and *Flexibacter elegans*) showed that 13-methyl-tetradecanoic acid, 15-methyl-hexadecanoic acid, hexadecanoic acid, and hexadecenoic acid, 2- and 3-hydroxy fatty acids comprised up to 50% of the total fatty acids. It was shown that 2-hydroxy-15-methyl hexadecanoic acid was the characteristic fatty acid for

the myxobacteria in contrast; 2-hydroxy-13-methyl-tetradecanoic acid, 3-hydroxy-13-methyl-tetradecanoic acid, and 3-hydroxy-15-methyl-hexadecanoic acid were dominant in the *Cytophaga* and *Flexibacter* group (Fautz, *et al.*, 1979). Similarly, *M. xanthus* fatty acids (including the branched-chain fatty acids) have been observed to be important for a variety of the effects on developing cell-stages and showed that the branched-chain fatty acids were released during the cell to cell signaling (Downard and Toal, 1995). It has also been shown that phosphatidylethanolamine (PE) containing 16:1 omega-5C is likely to play a role in developmental chemo-attraction in *Myxococcus xanthus* (Kearns *et al.*, 2001). Likewise, it was also shown that the iso-branched ether lipids are specific markers that are observed during the developmental sporulation in the *M. xanthus* (Ring *et al.*, 2006). In order to highlight this, two of the branched fatty acid, 13-methylmyristic acid and 15-methyl palmitic acid, were also tested. CYP109D1 was able to hydroxylate both fatty acids at sub-terminal position. Terminal hydroxylation was not observed. It is assumed that the hydroxylated products of the branched chain fatty acid and other fatty acids might have an important role during gliding motility, cell signaling and fruiting body formation as well as energy storage and dissemination.

CHAPTER II

B. CYP109D1 from *Sorangium cellulosum* So ce56: A terpenoid epoxidase/hydroxylase

B. 1. INTRODUCTION

B. 1. 1. Terpenoid compounds

Terpenes and terpenoids are the most diverse family of natural products synthesized notably from plants serving a range of important physiological and societal functions (Roberts, 2007). Over 40,000 different terpenoids have been isolated mainly from plant but also from animal and microbial species (Rohdich, *et al.*, 2005; Withers and Keasling, 2007). The basic molecular formulae of terpenes are the multiples of $(C_5H_8)_n$ where 'n' is the number of linked isoprene units. Terpenes could be hemiterpenes (with one isoprene unit e.g. prenol, isovaleric acid), monoterpenes (with two isoprene units, $C_{10}H_{16}$, e.g. geraniol, nerol, limonene), sesquiterpenes (with three isoprene units, $C_{15}H_{24}$, e.g. farnesenes, farnesol, nootkatone etc), diterpenes (with four isoprene units, $C_{20}H_{32}$, e.g. cafestol, cembrene, taxadiene), sesterterpene (with 25 carbons and five isoprene units e.g. geranylarnesol), triterpenes (with six isoprene units, $C_{30}H_{48}$, e.g. several precursors of steroids), tetraterpenes (with eight isoprene units, $C_{40}H_{64}$, e.g. lycopene, carotenes etc) and polyterpenes (with long chains of many isoprene units e.g. rubber).

Monoterpenes, in particular, represent a cheap and abundantly available pool of chiral substances that can be transformed into valuable bioactive compounds which are used in flavor, fragrance, and pharmaceutical industries or as building blocks in organic chemistry (de Carvalho and da Fonseca, 2006). Though the oxygenated products of terpenes are highly valuable compounds, their selective oxidation under mild conditions is still a great challenge. The regio- and stereo- selective functionalization of terpenes, especially terpene hydrocarbons, is known to be a difficult task for organic chemistry, and the few approaches described in the literature show several disadvantages with respect to reaction complexity and specificity (Newhall, 1958; Bhawal *et al.*, 2001; Stainslaw *et al.*, 2002; Duetz *et al.*, 2003). In such cases, the application of P450s monooxygenase could be useful. It is already shown that cytochrome P450 monooxygenases are capable of doing selective oxidation of different terpenes and terpenoids, often with a high activity (Peterson *et al.*, 1992; Hawkes *et al.*, 2002).

B. 1. 1. 1. Geraniol and nerol

The acyclic monoterpenes studied in this work were geraniol and nerol which are monoterpenoids and alcohols (Figure II. B. 1). Geraniol is abundantly present in the essential oils of rose, palmarosa and citronella, and widely used in fragrance chemistry. Geraniol is also a pheromone of certain species of bees, being secreted by the scent glands of worker bees to signal the location of nectar-bearing flowers and the entrances to their hives. Geraniol also finds an application as insect repellants or deterrents (<http://chemicaland21.com/specialtychem/perchem/NEROL.htm>). Moreover, it has been shown that the geraniol and its derivatives are present in 76% of the investigated deodorants on the European market, in 33% of the cosmetic based natural ingredients (Rastogi *et al.*, 1998) and in 41% domestic and household products (Rastogi *et al.*, 2001). Nerol is the cis-isomer of geraniol. They are also one of the compounds in the fragrance mix (FM) that is used for clinical screening of contact allergy among consecutive dermatitis patients which is responsible for only 5% of positive patch test reactions to the individual components of FM (Schnuch *et al.*, 2004) and is considered to be a weak allergen.

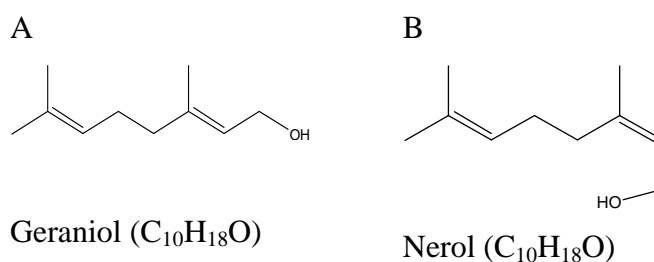


Figure II. B. 1: Structure of linear monoterpenes. A. geraniol B. nerol

B. 1. 1. 2. Limonene

Limonene (4-isopropenyl-1-methylcyclohexene), a monocyclic monoterpene, is the most widespread terpene in the world and is formed by more than 300 plant species (Burdock, 1995). (4*R*)-(+)-Limonene is the most widespread form. (4*R*)-Limonene is the major constituent of citrus peel essential oils, in which it is usually found at concentrations between 90 and 96% (Nonino, 1997). However, several plants form a mixture of both enantiomers, while others produce only (4*S*)-(-)-limonene (Burdock, 1995). Different studies have been undertaken to evaluate possibilities of biotransformation of limonene with an aim to produce more valuable natural flavor compounds (Cadwallader, *et al.*, 1989; Chang and Oriel, 1994; Dhere and Dhavlikar, 1970, Noma *et al.*, 1992, van Dyk, 1998, Wust and Croteau, 2002). Five different microbial transformation pathways for limonene have been proposed (van der Werf *et al.*, 2000) and are illustrated in Figure II. B. 2.

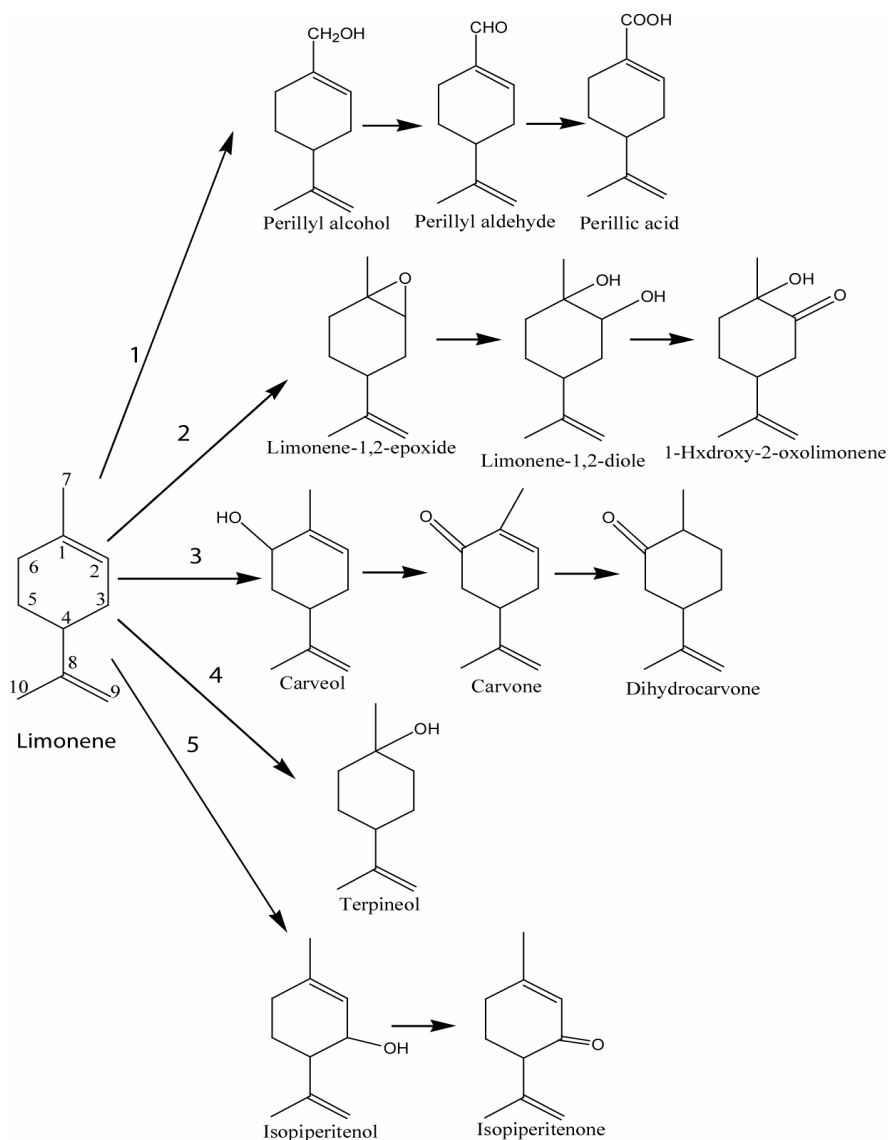


Figure II. B. 2: Microbial biotransformation pathways for limonene.

B. 1. 1. 3. Ionone

Ionones are sesquiterpenoids, naturally occurring unsaturated hydrocarbons whose carbon skeletons are composed exclusively of isoprene C₅ units [CH₂=C(CH₃)-CH=CH₂], consisting of five carbon atoms attached to eight hydrogen atoms (C₅H₈). Ionone can be synthesized from citral and acetone with calcium oxide as a basic heterogeneous catalysis and serves as an example of an aldol condensation followed by rearrangement reactions (Noda *et al.*, 1998; Russell and Kenyo, 1955).

Ionones (α -, β -, and γ) and their hydroxyl forms are essential intermediates for the synthesis of a number of carotenoids (Eschenmoser, *et al.*, 1981; Oritani and Yamashita, 1984) and can

be used in the industrial synthesis of zeaxanthin and cantaxanthin. Ionone derivatives with a trimethylcyclohexane building block constitute substantial aroma components of floral scents (Eugster and Maeki-Fischer, 1991; Winterhalter and Rouseff, 2002), and thus attracted an attention of flavor and fragrance industry. The α -ionone shows a more floral odour than beta. It is also used in vitamin A (retinol) production for cosmetics and toiletries. α -ionone and its derivatives have been described as effective male attractants for *B. latifrons* (Flath, *et al.*, 1994; Ishida, *et al.*, 2008). They form important components of insect lures (Donaldson, *et al.*, 1990; McQuate and Peck, 2001) which can favor insect pollination.

On the other hand, the ionone derivatives, e.g. 3-hydroxy- β -ionone, could provide valuable intermediates for the chemoenzymatic synthesis of carotenoids, e.g. for astaxanthin and zeaxanthin (Loeber, *et al.*, 1971). However, β -ionone has been analyzed for colon cancer chemoprevention and treatment, and a β -ionone induced cell growth inhibition and apoptosis in human colon cancer HCT116 cell line was shown (Janakiram, *et al.*, 2008). It has also been reported that the β -ionone and several 3-hydroxy esters of β -ionone have antifungal activity (Wyatt and Kue, 1992). Hydroxy- β -ionones are also versatile compounds for the synthesis of phytohormones like abscisic acid (Oritani and Yamashita, 1994). The possible products from the β -ionone are illustrated in Figure II. 7.

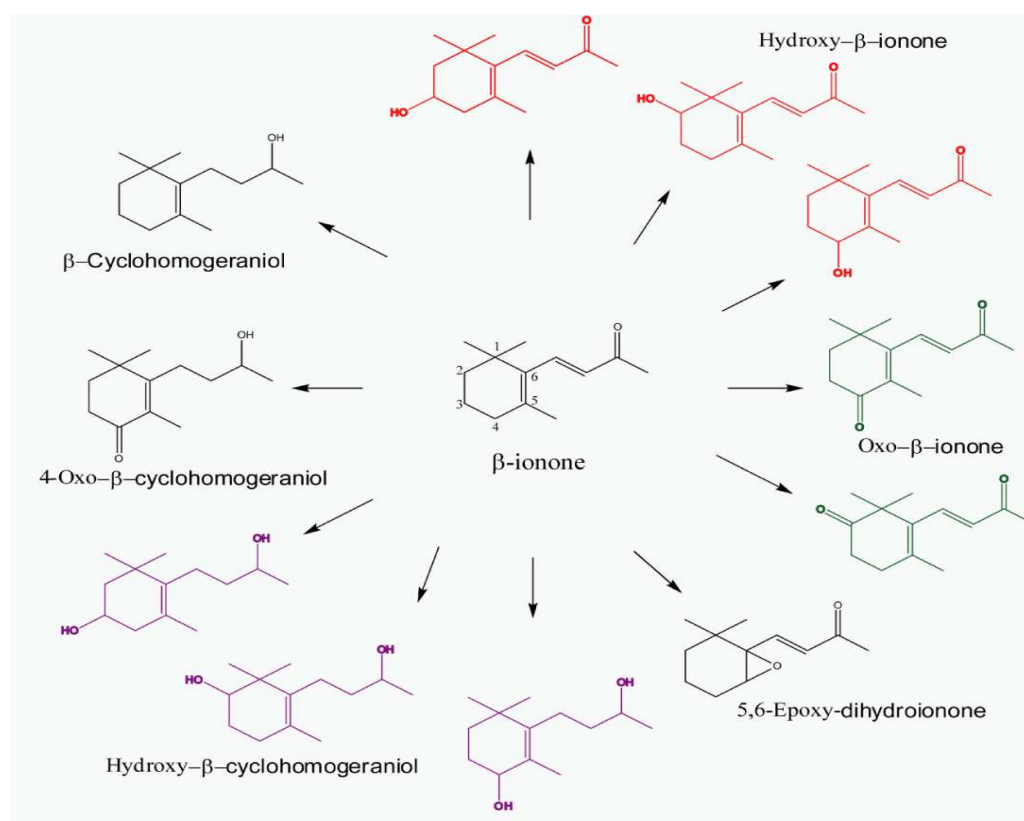


Figure II. B. 3: Possible conversion of β -ionone. Similar compounds are represented by the same color.

B. 1. 1. 3. 1 Role of cytochrome P450 in preparation of ionone derivatives

A few methods of preparing hydroxy derivatives of ionones like the chemical optical resolution of α -ionone (Haag, *et al.*, 1980) or on selective enzymatic hydrolysis of the 4-hydroxy- β -ionone have been published (Oritani and Yamashita, 1994; Kakeya, *et al.*, 1991). Microbial transformation of the α - and β -ionone has been reported for several fungal strains, especially *Aspergillus* sp (Lorroche *et al.*, 1995; Mikam *et al.*, 1981; Yamazaki *et al.*, 1988). Similarly, the conversion of β -ionone into 4-hydroxy- β -ionone and the conversion of α -ionone into 4-hydroxy- α -ionone by the several *Streptomyces* sp was studied (Lutz-Wahl, *et al.*, 1998). The biotransformation of the α -ionone and β -ionone catalyzed by P450s derived from actinomycetes (Celik, *et al.*, 2005) and the engineered cytochrome P450BM-3 (Urlacher, *et al.*, 2006).

B. 2. Objective of Chapter II. B.

1. To find terpenes and terpenoids as substrates for CYP109D1
2. To show CYP109D1 as a biocatalyst for epoxidation/hydroxylation of linear, cyclic and sesquiterpenoids.
 - a. To analyze epoxidation of linear and cyclic mono terpenoids.
 - b. To analyze the hydroxylation of sesquiterpenoids.
 - c. To illustrate the regioselectivity of hydroxylation by docking.

B. 3. MATERIAL AND METHODS

B. 3 1. Materials

The bacterial stains, other proteins, oligonucleotides and the vectors were as illustrated in Chapter I.

B. 3. 2. Synthesized compound

The synthesized products of geraniol *E*-3,7-dimethyl-6,7-epoxy-2-octen-1-ol and *E*-3,7-dimethyl-2,3-epoxy-6-octen-1-ol, and the products of nerol *Z*-3,7-dimethyl-2,3-epoxy-6-octen-1-ol and *Z*-3,7-dimethyl-6,7-epoxy-2-octen-1-ol were used as authentic standard. Their synthesis procedures were published elsewhere (Watanabe *et al.*, 2007). These synthesized products were available in Institute of Technical Biochemistry, Universitaet Stuttgart.

B. 3. 3. Methods

The bioinformatics studies, protein expression, purification and spectrophotometric analysis were as illustrated in Chapter I, except that the yield of protein expression was increased by coexpression of chaperons. CYP109D1 was recloned from pCWori (pCWori_CYP109D1) into expression vector pET17b and coexpressed with molecular chaperones GroES/GroEL as described previously (Nishihara *et al.*, 1998; Zöllner *et al.*, 2008). The expression host, *E. coli* BL21(DE3) was transformed with two expression plasmids, pET17b_CYP109D1, and pGro12_GroES/GroEL and cultured overnight in TB medium containing 100µg/ml ampicillin and 50µg/ml kanamycin. The overnight culture was hundred fold diluted in 200 ml of the TB medium in 2 liter baffled flask containing 100 µg/ml ampicillin and 50 µg/ml kanamycin and grown at 37°C and 95 rpm unless the optical density at 600 nm reached 1.2. Upon culturing the heme synthesis was supported by the addition of 0.8 mM of δ-ALA and the expression was induced with 1 mM IPTG. For the chaperone expression 4 mg/ml final concentration of arabinose was added. At the same time 50 µg/ml of the ampicillin was also added to overcome ampicillin degradation at high temperatures or its deactivation by beta-lactamases. The cultures were grown at 28°C for additional 24-36 hours at 95 rpm. The cultures were harvested by centrifugation at 4500 rpm and the pellet was stored at -20°C until purification of the protein. CYP109D1 was purified and the active form of P450 was confirmed by CO-difference spectrum (COD) measurements as described in Chapter I.

B. 3. 3. 1. *In vitro* assay

The substrates used in this work were acyclic (geraniol and nerol) and cyclic (limonene) monoterpenes as well as sesquiterpenoids (α -ionone and β -ionone). Reactions were performed in a final volume of 0.5 ml of buffer E (Appendix II) with 200 μ M substrate and a NADPH-regenerating system, consisting of glucose-6-phosphate (5mM), glucose-6-phosphate dehydrogenase (1 unit) and magnesium chloride (1 mM). The reconstitution with the redox partners Adx and AdR was performed with the ratios of CYP109D1: Adx: AdR of 1: 10: 1 (1 μ M: 10 μ M: 1 μ M) was used. During incubation at 30°C on a thermo shaker, NADPH was added to 500 μ M final concentration and the mixture was incubated for 30 min. The reaction was stopped and extracted twice with an equal volume of ethyl acetate. The organic phase was pooled and the volume was reduced to 200 μ l with the flow of nitrogen gas. The samples were run on either thin layer Silica Gel 60 F254 plates and/or directly analyzed using gas chromatography coupled with mass spectrometry as mentioned below.

B. 3. 3. 2. Gas chromatography coupled with mass spectrometry (GC-MS) analysis

All the GC-MS measurements were carried out at Institute of Technical Biochemistry, Universitaet Stuttgart. GC-MS was performed on a Shimadzu GC/MS QP2010 (Kyoto, Japan) equipped with an FS Supreme-5 column (0.25 mm x 30 m, 0.25 μ m, CS-Chromatography Service, Langerwehe, Germany). The following program was used for the analysis of the substrates and products.

Geraniol and nerol: The column temperature was controlled at 120°C for 3 min. The temperature was then raised to 165°C at the rate of 5°C min⁻¹, and then to 280°C at 30°C min⁻¹. It was kept at 280°C for 1 min. The temperatures of injector and the interface were fixed at 285°C respectively.

Limonene: The column temperature was controlled at 80°C for 2 min. The temperature was then raised to 180°C at the rate of 10°C min⁻¹, and then to 250°C at 30°C min⁻¹. It was kept at 250°C for 1 min. The temperatures of injector and the interface were fixed at 250°C.

α - and β -ionone: The column temperature was controlled at 150°C for 1 min. The temperature was then raised to 250°C at 20°C min⁻¹. It was kept at 250°C for 5 min. The temperatures of injector and the interface were fixed at 250°C respectively.

B. 3. 3. 3. Thin layer chromatography (TLC)

Thin layer chromatography (TLC) was used to monitor the conversion of α - and β -ionone and to inspect the substrate and/or product fractions obtained during the silica column chromatography. Samples/products were spotted onto TLC aluminum sheets (5 x 10 cm², silica gel layer thickness, 0.2 mm; Silica Gel 60 F254; Merck), and the plates were developed in solvent tank containing n-hexane:ethyl acetate (3:2). Products were visualized firstly at UV light at 254 nm and then confirmed by spraying with a vanillin solution (2.5% (w/v) vanillin in ethanol:sulfuric acid (H₂SO₄) (95:5) and subsequent heating at 110°C for 2 min.

B. 3. 3. 4. Isolation of 3-hydroxy- α -ionone from the reaction mixture

For the identification of the products of α -ionone oxidation, the *in vitro* reaction described above was scaled up 120 times. The organic phase was dried using vacuum; the obtained product was completely dissolved in the mixture n-hexane:ethyl acetate (3:2) and purified on the 150 ml silica gel column. The complete separation of remained α -ionone and the product was monitored by TLC. The purified product fractions were pulled and dried in the rotary vacuum device (SpeedVac Concentrator 5301, Eppendorf) and 10 mg was used for NMR analysis (Avance 500 (Bruker Biospin GmbH, Rheinstetten, Germany). The structure was confirmed by ¹H- and ¹³C NMR and IR analysis.

B. 3. 3. 5. Computational methods

The computational work was done in collaboration with PD Dr. Michael C. Hutter, Center for Bioinformatics, Saarland University.

P450eryF (CYP107A1) from *Saccharopolyspora erythrae* (accession code Q00441) shows a sequence identity of 31.6% to CYP109D1 of *Sorangium cellulosum* So ce56. The corresponding crystallographic structure (pdb entry 1Z8P) was chosen as structural template. The alignment generated with CLUSTAL W (version 2) (Larkin *et al*, 2007) was used as input for SWISS-MODEL (Arnold *et al*, 2007; Schwede *et al*, 2003; Guex *et al*, 1997). The coordinates of the heme-porphyrin atoms from the template structure were added subsequently to the obtained homology model. The vicinity around the porphyrin was structurally relaxed by force field optimization using the AMBER3 set of parameters as implemented in HYPERCHEM (version 6.02, 1999). The carboxylate groups of the heme-form ionic hydrogen-bonds with the side chains of Arg293, Arg104, His100, and His349. Compounds for docking were generated manually and energetically optimized using the

MM+ force field as implemented in HYPERCHEM AUTODOCK (version 4.00) was applied for docking of α - and β -ionone into the homology model of CYP109D1 (Huey *et al*, 2007; Morris *et al*, 1998). The Windows version 1.5.2 of Autodock Tools was used to compute Gasteiger-Marsili charges for the enzyme and the ligands (Sanner *et al*, 1999). A partial charge of +0.400 *e* was assigned manually to the heme-iron, which corresponds to Fe(II), that was compensated by adjusting the partial charges of the legating nitrogen atoms to -0.348 *e*. Flexible bonds of the ligands were assigned automatically and verified by manual inspection. A cubic grid box (50 x 50 x 50 points with a grid spacing of 0.375 Å) was centered above the heme-iron at the putative position of the ferryl oxygen. For each of the ligands 100 docking runs were carried out applying the Lamarckian genetic algorithm using default parameter settings, except for the mutation rate that was increased to 0.05.

B. 4. RESULTS

B. 4. 1. Expression and purification of recombinant CYP109D1

Previously, CYP109D1 was expressed in *E. coli* BL21 using a bacterial expression vector pCWori and the expression level was only 205 nmol.L⁻¹ culture media (Chapter I). The co-expression of pET17b_CYP109D1 with the molecular chaperons GroES/GroEL resulted the four fold increased in the expression level with up to 900 nmol/L soluble P450 enzyme. CO-difference spectra (COD) (Omura and Sato, 1962) demonstrate peak maximum at 450 nm which confirmed the active form of the enzyme.

B. 4. 2. Conversion of geraniol by CYP109D1

The homologous redox partners for CYP109D1 from *So ce56* are not known so far. Therefore, the heterologous redox partners adrenodoxin (Adx) and adrenodoxin reductase (AdR) were used for *in vitro* conversion. GC-MS analyses demonstrated that CYP109D1 was able to convert geraniol into epoxidated products. The substrate geraniol has a retention time (RT) of 6.37 min and the product peaks I, II and III were obtained at RT of 8.15, 8.72 and 8.93 min, respectively (Figure II. B. 4).

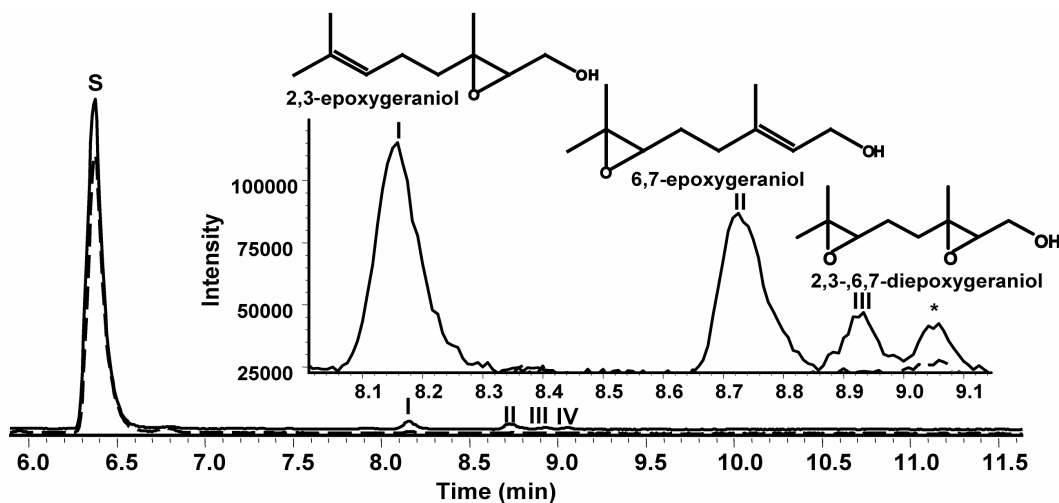


Figure II. B. 4: The GC diagram of geraniol conversion. The solid and dashed lines represent conversion and negative control (geraniol only), respectively. 'S' is the substrate and the inset highlights the products. Peak I, peak II and peak III are identified as 2,3-epoxy-, 6,7-epoxy and 2,3-,6,7-diepoxygeraniol, respectively. The '*' represents a non-specific peak.

The mass spectra of the peak I and peak II corresponded to the authentic 2,3-epoxy- and 6,7-epoxygeraniol, respectively. Peak III was correlated with the diepoxidated geraniol at 2,3- and 6,7-C-C double bonds, respectively (Figure II. B. 4). All the mass spectra were identical with the mass spectra of the authentic epoxides of geraniol (Figure II. B. 5).

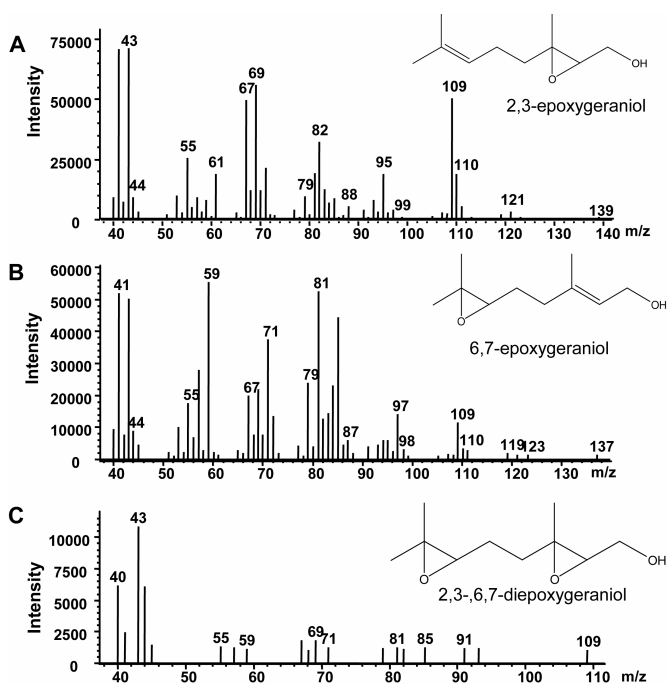


Figure B. 5: Mass spectra of geraniol obtained from the NIST arbitrary library. The MS of this arbitrary was determined from authentic reference compounds. A. The MS of peak I is shown to be 2,3-epoxygeraniol. B. The MS of peak II is shown to be 6,7-epoxygeraniol C. The MS of peak III is shown to be 2,3-, 6,7-epoxygeraniol.

B. 4. 3. Conversion of nerol by CYP109D1

CYP109D1 was also able to convert nerol into two distinct products, peak I and peak II, showing retention times of 7.90 and 8.10 min, respectively (Figure II. B. 6).

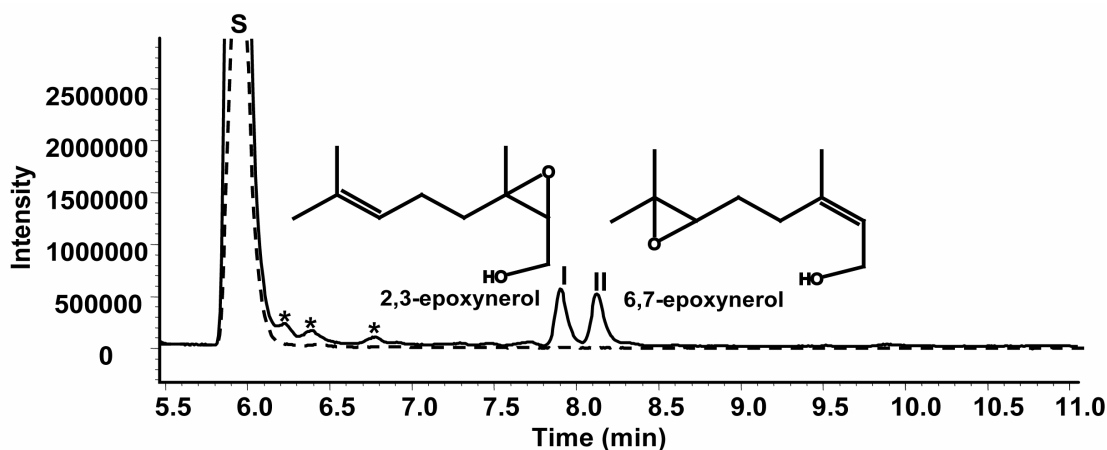


Figure II. B. 6: GC diagrams of nerol conversions. The solid and dashed lines represent conversion and negative control (nerol only), respectively. 'S' is the substrate. Peak I and peak II are identified as 2,3-epoxy and 6,7-epoxynerol, respectively. The '*' represents non-specific peaks.

The mass spectra of the peak I and peak II corresponded to the 2,3-epoxy- and 6,7-epoxynerol, respectively (Figure II. B. 6). The product mass spectra were also compared to

the mass spectra of the authentic epoxidated products of nerol and revealed identical peaks (Figure II. B. 7).

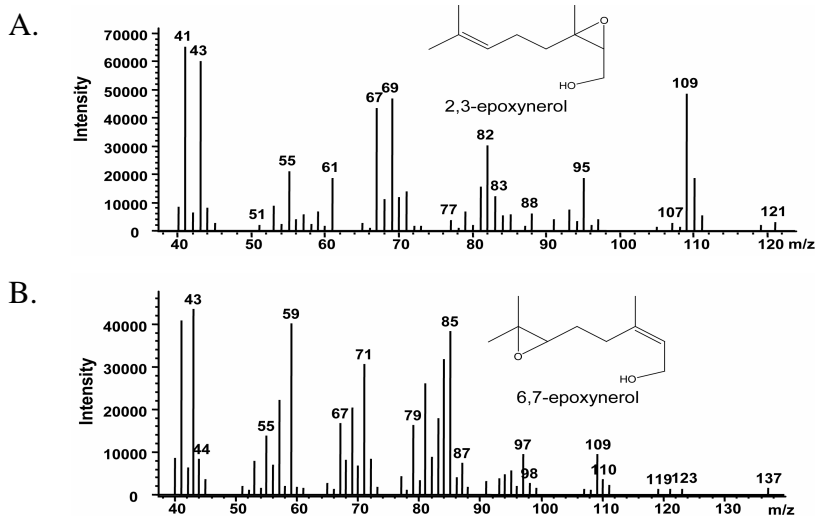


Figure II. B. 7: Mass spectra of nerol obtained from the NIST arbitrary library. The MS are the same as that for the authentic products of the nerol. A. The MS of peak I is shown to be 2,3-epoxynerol. B. The MS of peak II is shown to be 6,7-epoxynerol

B. 4. 4. Conversion of limonene by CYP109D1

GC-MS analyses demonstrated that CYP109D1 was able to convert (*R*)-(+)-limonene into several products. The total conversion of the substrate was nearly 10%. Six product peaks I, II, III, IV, V and VI, at retention times of 6.78, 7.40, 7.46, 8.39, 8.45 and 8.71 min, respectively, were obtained (Figure II. B. 8).

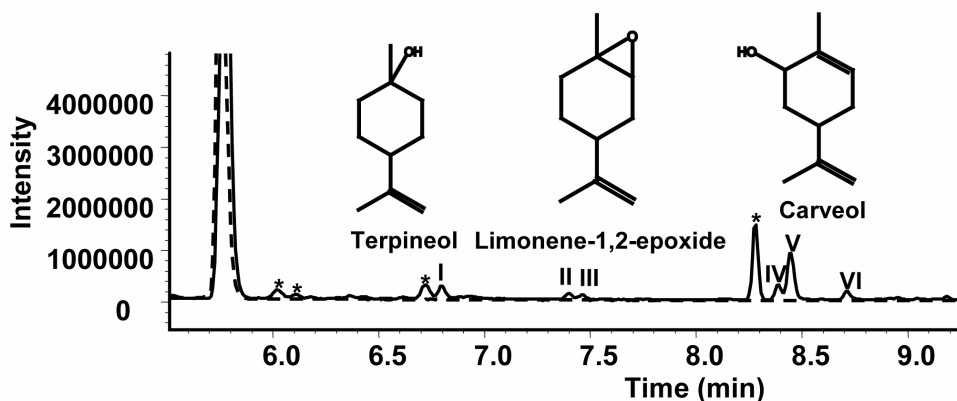


Figure II. B. 8: GC diagrams of (*R*)-(+)-limonene conversions. In the figure, the solid line represents the conversion and the dashed line is for negative control ((*R*)-(+)-limonene only). Peak I is identified as a terpeneol, peak II and peak III as *cis*- and/or *trans*- limonene-1,2-epoxide, and peak IV, peak V and peak VI as *cis*- and/or *trans*-carveol. The '*' represents non-specific peaks.

The products of (*R*)-(+)-limonene could be in *cis*- or *trans*-form. In the absence of authentic standards, the absolute configuration of the products could, however, not be identified. So, the product mass spectra were compared with data in the NIST arbitrary library. According to the found coincidence, the mass spectra of the peak I was matched with terpineol (Figure II. B. 9), peak II and peak III with *cis*- and/or *trans*- limonene-1,2-epoxide (Figure II. B. 9), peak IV, peak V and peak VI with *cis*- and *trans*-carveol (Figure II. B. 9). Though the configuration was not able to be allocated, it was clear that CYP109D1 was able to hydroxylate and epoxidate (*R*)-(+)-limonene at several positions.

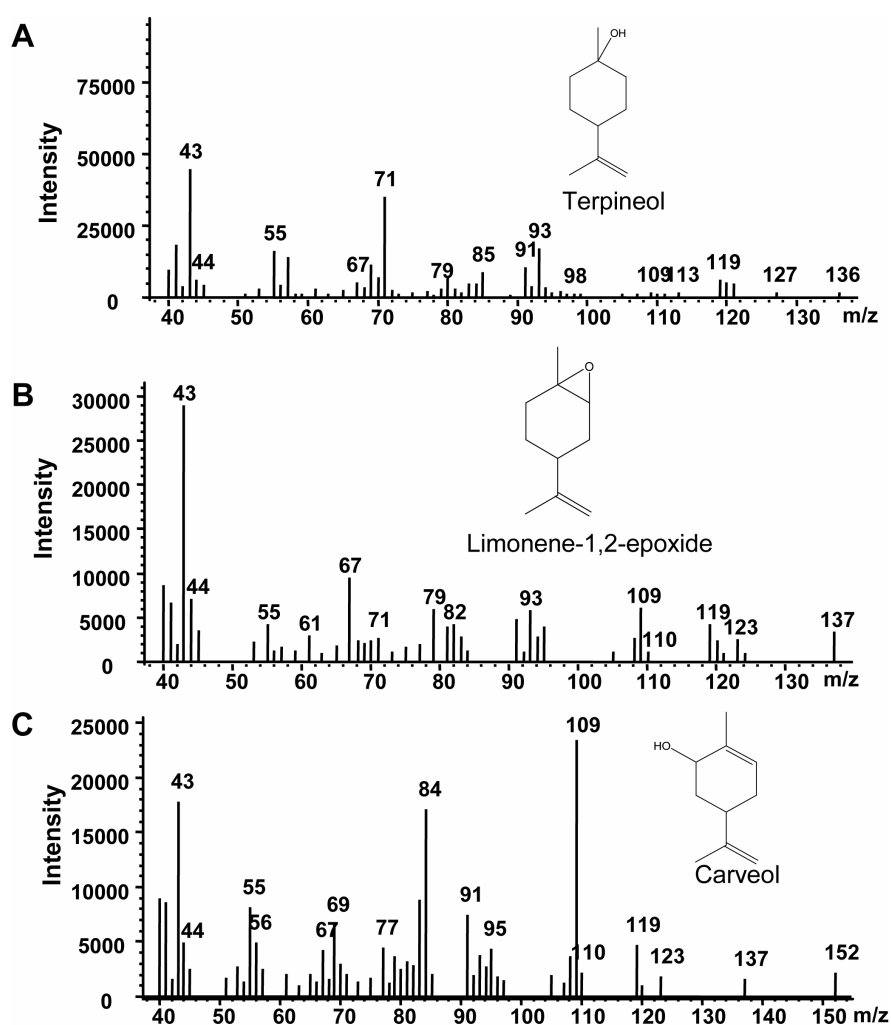


Figure II. B. 9: Mass spectra of (*R*)-(+)-limonene obtained from the NIST arbitrary library. A. The MS of the peak I is identified as a terpineol, B. peak II and peak III are identified as *cis*- and/or *trans*- limonene-1,2-epoxide. C. The peak IV, peak V and peak VI are identified as *cis*- and/or *trans*-carveol.

B. 4. 5. Conversion of ionone by CYP109D1

B. 4. 5. 1. Conversion of α -ionone

Besides hydroxylating acyclic and monocyclic monoterpenes, CYP09D1 was able to convert α -ionone into a single product. The product, peak I, was obtained at a retention time of 5.03 min (Figure II. B. 10) and the mass spectrum of the product is illustrated in Figure II. B. 11. A similarity search for the mass spectrum within the NIST arbitrary library did not give the precise expected hints on the product. Therefore, we purified and characterized this product.

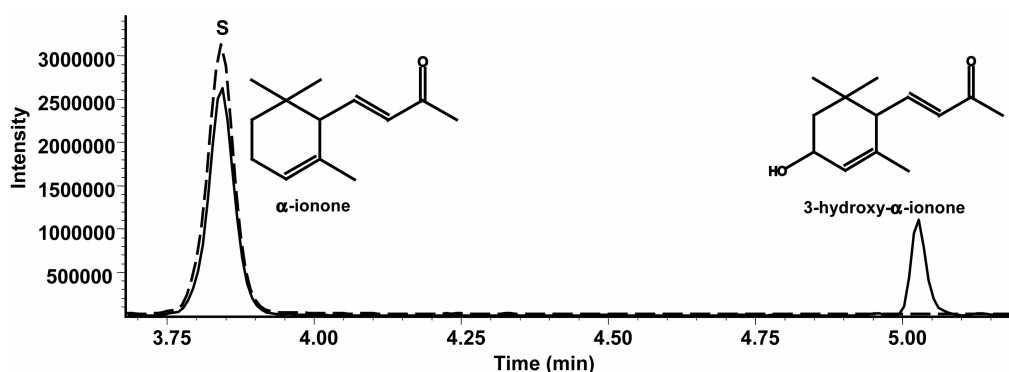


Figure II. B. 10: GC diagrams of α -ionone conversions. The solid line represents the conversion and the dashed line is for negative control (α -ionone only). The product peak is identified as 3-hydroxy- α -ionone and is illustrated in the inset.

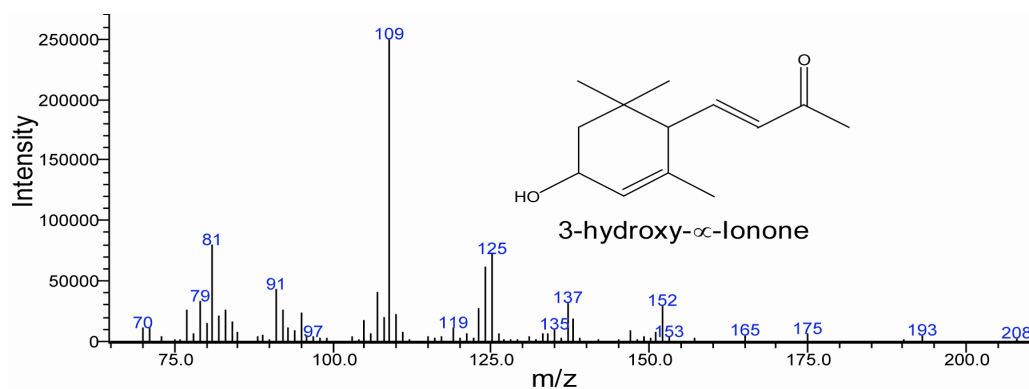


Figure II. B. 11: Mass spectra of 3-hydroxy- α -ionone obtained from the NIST arbitrary library.

B. 4. 5. 2. Identification of the product of α -ionone oxidation

A large scale conversion and purification on silica gel column was done to obtain enough amounts of the single product of α -ionone. After the silica gel preparation, 16 mg of the purified product were obtained. The identification of additional functional groups in the product molecule was done through IR spectroscopy, and was confirmed by ^1H and ^{13}C NMR. The product was found to be 3-hydroxy- α -ionone (Table II.B.4). The detailed results of the α -ionone products were as follows.

TLC: $R_f = 0.67$ (for substrate), 0.30 (for product), hexane- ethyl acetate (3:2)

GC-MS (EI). ($t_R = 5.030$ min.), (208, M^+ , 27%), (109, 100%), (81, 31.88%), (125, 28.65%), (91, 16.93%), (107, 16.05%), (137, 12.95%), (152, 11.78%)

IR. The IR-spectroscopy study of the product gave a characteristic IR absorption of O-H. The O-H stretch hydrogen bond of alcohol was appeared in the region of $3600-3200\text{ cm}^{-1}$ with $\nu_{\max}/\text{cm}^{-1}$ (FT). 3449

$^1\text{H NMR}$: δ_{H} (CDCl_3 ; 200 MHz), 6.47 (1 H, dd, $J = 15.7$ and 10 Hz); 6.03 (1 H; d, $J = 15.8$ Hz); 5.56 (1 H, br); 4.20 (1 H, br); 2.43 (1 H; d, $J = 10$ Hz); 2.19 (3 H, s); 1.77 (1 H, dd, $J = 13.6$ and 6.0 Hz); 1.51 (3 H, s); 1.33 (1 H; dd, $J = 13.6$ and 6.6 Hz); 0.96 (3 H, s); 0.82 (3 H, s). **$^{13}\text{C NMR}$ δ :** 33.35 (C-1), 43.83 (C-2), 65.48 (C-3), 125.83 (C-4), 135.20 (C-5), 54.30 (C-6), 147.06 (C-7), 133.62 (C-8), 198.30 (C-9), 27.20 (C-10), 29.70 (C-11), 24.21 (C-12), 22.66 (C-13).

B. 4. 6. Conversion of β -ionone

CYP109D1 was also able to convert β -ionone into a single product. The product, peak I, was obtained at a retention time of 5.77 min (Figure II. B. 12) and the mass spectrum of the product is illustrated in Figure II. B. 13. The similarity search of the mass spectra within the NIST mass spectra library showed that the MS-data was 83% similar to that of 4-hydroxy- β -ionone. Analysis of the authentic 4-hydroxy- β -ionone (Urlacher *et al.*, 2006) by GC-MS confirmed that the product was 4-hydroxy- β -ionone. The detailed results for the β -ionone product were as follows.

TLC. $R_f = 0.93$ (for substrate), 0.34 (for product), hexane-ethyl acetate (3:2)

GC-MS (EI). ($t_R = 5.199$ min.), (208, M^+ , 15%), (109, 100%), (123, 29.41%), (91, 22.90%), (208, 19.63%), (95, 19.00%), (137, 17.85%), (79, 15.78%), (107, 15.68%).

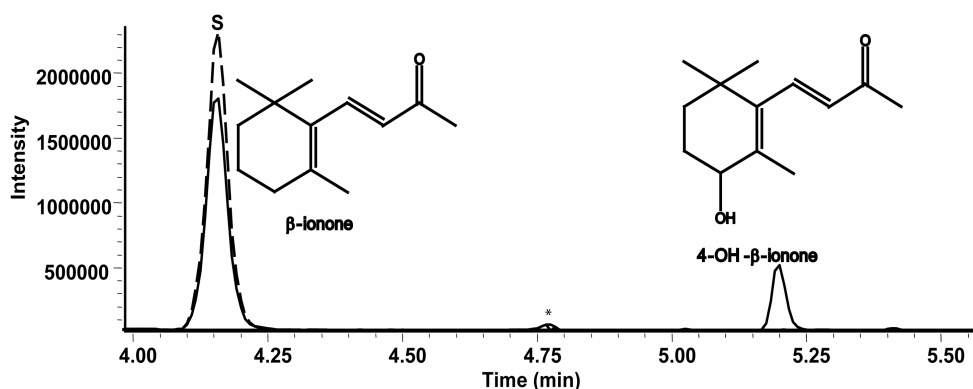


Figure II. B. 12: GC diagrams of β -ionone conversions. The solid line and the dashed line represent the conversion and negative control (β -ionone only), respectively. The product peak is identified as 4-hydroxy- β -ionone and is illustrated in the inset.

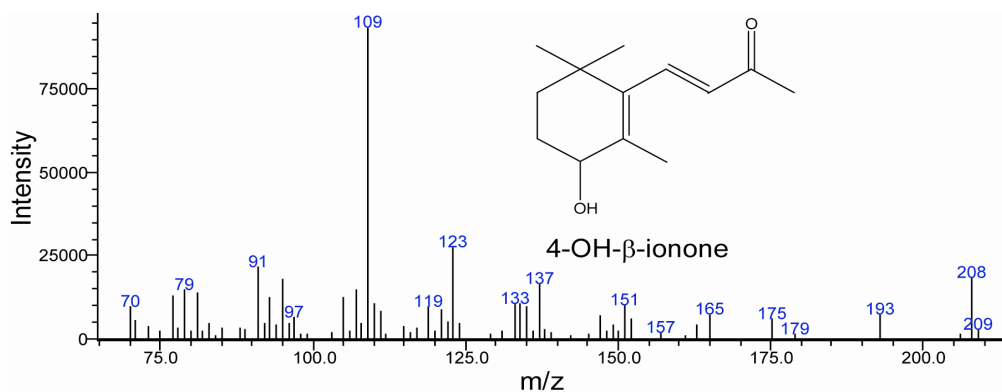


Figure II. B. 13: Mass spectra of β -ionone obtained from the NIST arbitrary library.

B. 4. 7. Homology modeling and docking of CYP109D1 with α - and β -ionone.

Unfortunately, the crystal structure of CYP109D1 is not known yet. To get more insight into the structural basis for the regiospecificity of hydroxylation of α - and β -ionone, computer modeling of CYP109D1 and its complexes with α - and β -ionone has been performed in cooperation with PD Dr. Michael C. Hutter, Center for Bioinformatics, Saarland University.

CYP109D1 from *So ce56* has the highest identity (31.6%) with P450eryF (CYP107A1) from *Saccharopolyspora erythraea* among the known crystal structures of bacterial cytochrome P450s that have been documented in protein data bank (PDB). The alignment is shown in Figure 14.

The corresponding crystallographic structure (pdb entry 1Z8P) was chosen as structural template. The carboxylate groups of the heme- moiety form ionic hydrogen-bonds with the side chains of Arg293, Arg104, His100, and His349. Both the *R* and *S*-isomer of α -ionone were found in binding poses that have the 3-position in suitable arrangement for hydroxylation (Figure 15). The distances towards the iron atom are 3.90 Å for the *R*-isomer and 2.96 Å for the *S*-isomer, respectively. Both the isomers show a hydrogen-bond (2.1 Å) between the carbonyl oxygen and the NE2-nitrogen atom of His 94. All other obtained binding poses exhibit the alternative sites for hydroxylation much more distant. This indicates structural implications as reason for the regioselective hydroxylation of α -ionone. Likewise, the closest distance found between the iron atom and the 4-position in β -ionone is 4.34 Å, whereas all other hydroxylation sites within the ring are again further away. The larger separation found for β -ionone agrees very well with the lower turn-over compared to β -ionone.

TARGET	7	PSFS	PEQIDLSAPS	VIADFYFAYR	ALRGRSPVLY	ARVPAGGAAG
1z8pA	2	ttvp	dles----ds	fhvdwyrtya	elretapvtp	vr-flgqdaw
TARGET				hhhhh	hhhhh	ssss
1z8pA				hhhhh	hhhhh	ss s ssss
TARGET	51	LGEPI-RAYA	LLRHAEVLA	LR-----D	PQTSSNVTD	KIRVLPRITL
1z8pA	41	lvtgydeaka	aledlrlssd	pkkkypgvev	efpaylgfpe	dvznyfatnm
TARGET		sss	hh hhh	sss		h hhhhhh
1z8pA		sss	hhhhhh	hhh	sss	h hhhhhh
TARGET	93	LHDDPPRTH	LRLVRSFT	PRIAELEPW	IGRLAASLE	ATGDGP-SDL
1z8pA	91	gtdcppthr	lrlvsgft	vrrveampr	veqitaellid	evgdsgvvd
TARGET		hhh	hhhhh	hhhhhhhhh	hhhhhhhhh	h sss
1z8pA		hhh	hhhhh	hhhh	hh	hhhhhhhhh
TARGET	142	MGAYAMPLM	MVIATLLGIP	AERYVQFRSW	SESVMSYSGI	PAERASRGK
1z8pA	141	vdrfahplpi	kvicellgvd	ekyrgefgrw	sseilvmdpe	raeqrgqaar
TARGET		hh	hhhhh	hhhhhhh	hhh	hhhh
1z8pA		hh	hhhhh	hhhhhhh	hhh	hhhh
TARGET	192	AMVDFFAEL	EARRAPSGD	LISALVEAEI	DG-ARLDTPE	AVGFCVGLLV
1z8pA	191	evnfildiv	errrtepgdd	llsalirvqd	dddgrlsade	ltsialvll
TARGET		hhhhhhhhh	hhhhh	hhhhhhh	hhh	hhhhhhhhh
1z8pA		hhhhhhhhh	hhhhh	hhhhhhh	hhh	hhhhhhhhh
TARGET	241	AGNDTTNLI	GNMAHLLSER	PELYRRAQD	RSLVGPPIIE	TLRHSSPVQR
1z8pA	241	agfessvsl	gigtlyllth	pdqialvrrd	psalpnavee	ilryiappet
TARGET		h	hhhhh	hhhhhhh	hhhhhhh	hhhh
1z8pA		h	hhhhh	hhhhhhh	hhhhhhh	hhhh
TARGET	291	LLRVTTFPVD	VSGVMIPAGH	LVDVVFGAAN	RDPVFEEDP	AFRLDRPFAE
1z8pA	291	ttrfaaeve	iggvaipqys	tvlvangaan	rdpkqfpdph	rfdvtrdtrg
TARGET		ssssss	ss s	sss	s	ssssshhh
1z8pA		ssssss	ss s	sss	s	ssssshhh
TARGET	341	HLAFGQGFH	CIGAALARME	ARIANALD	CYESITPGEA	-PFLRQTRAI
1z8pA	341	hlsfqqghf	cmgrplakle	gevalraifg	rfpalslgid	addvwrrsl
TARGET			hhhhhhh	hhhhhhhhh	h	sss
1z8pA			hhhhhhh	hhhhhhhhh	h	sss
TARGET	390	MPLGFESLPL	VLRR			
1z8pA	391	llrgidhlpv	rl dg-			
TARGET			sss	s		
1z8pA			sss	ss		

Figure 14: The obtained alignment and secondary structure of CYP109D1 (target) and the template (pdb entry 1z8p). In the figure, the α -helices are denoted by 'h' and the β -sheets are indicated by 's'.

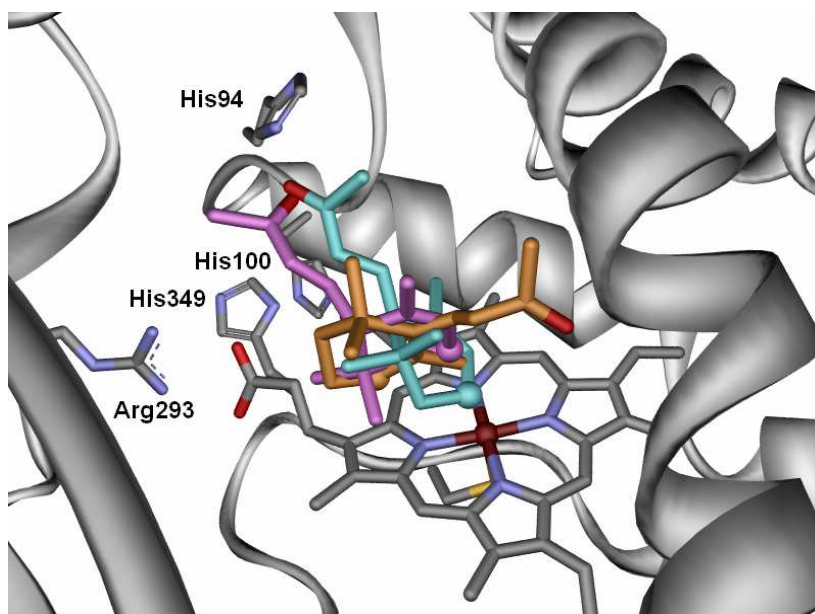


Figure 15: Obtained docking positions in the homology model of CYP109D1. α -ionone *R*-isomer is shown in pink, *S*-isomer in cyan, β -ionone in orange. The respective carbonyl oxygen atoms are shown in red. The positions where hydroxylation occurs are marked as balls.

B. 5. Discussion

Most of the terpenes and terpenoids are commercially valuable with activities usually depending on characteristic functional groups and the absolute stereochemical configurations. Several attempts have been undertaken to discover new biologically active terpenoids, having higher market value. However, the selective oxy-functionalization of these compounds and oxidation of the resulting alcohols to the corresponding carbonyls, which are the key reactions in the fine chemical industry, are still challenging to organic chemistry due to low specificities and the need for hazardous and expensive catalysts and reactants (Newhall, 1958; Lochynski *et al.*, 2002; Sheldon, 2000; Duetz *et al.*, 2003; Buehler and Schmid, 2004).

This requirement can only be met either by conventional isolation methods or by the enzymatic or microbial production of terpene and terpenoid derivatives from natural raw material, which has been the target of intensive research activities during the past decades (Schrader and Berger, 2001). Nowadays, biocatalysis offers promising synthetic alternatives due to the huge natural diversity of oxygenating and oxidizing enzymes acting as highly regio- and stereo-selective catalysts. The great majority of these biotransformations, including some terpene oxidations, were conducted with whole cells expressing cytochrome P450s which, however, resulted in rather low productivities and the formation of some other by-products (Schrader and Berger, 2001; de Carvalho and da Fonseca, 2006). Although such items usually show low activities, instability, and depend on the costly cofactors NAD(P)H and/or redox partners that are usually associated with P450 enzymes, they still provide a great technical potential because of their high regio- and stereo-selectivity (Duetz *et al.*, 2001, 2003; Urlacher and Schmid, 2002). In addition, recent methods like directed evolution can be used to improve the stability and efficiency of these proteins.

From a viewpoint of chemical synthesis, monoterpenes are particularly interesting as flavorings and fragrances or as pharmaceuticals (de Carvalho and da Fonseca, 2006). They can contain alkene functional groups with different degrees of substitution. They can also have other functionalities that may further alter the chemistry of the carbon-carbon double bond. These products are only possible through radical-reactions but are not easy and most of the time not selective to introduce desired epoxidation or chiral alcohol functions in those compounds. As those disadvantages can be overcome by the application of P450, we have cloned and expressed a cytochrome P450, CYP109D1. Using a reconstituted system with Adx and AdR as redox partners, we were able to demonstrate the conversion of monoterpenes as

well as sesquiterpenes. Geraniol and nerol, the oldest known organic compounds, were chosen for this work as acyclic monoterpenes. CYP109D1 was able to convert both the compounds into 2,3- and 6,7-epoxy- as well as into the 2,3-6,7-diepoxy form (only for geraniol). It was reported that cytochrome P450 BM-3 mutant (R47L/Y51F/F87V) was able to convert geraniol into two and nerol into three epoxidized products (Wong *et al.*, 2000; Watanabe *et al.*, 2007). It is notable that even the wild type CYP109D1 was able to give selective epoxidated products and the study of selective mutants of CYP109D1 might be useful to produce more regio- or stereo-selective products. These epoxy derivatives have several applications. Some terpene epoxides like 6,7-epoxygeraniol are the major constituent of juvenile hormones in several beetles (Bowers, 1969; Archelas, *et al.*, 1993), and 2,3-epoxygeraniol as well as the corresponding acetate are part of the pheromone bouquet of male neotropical orchid bee (Eltz *et al.*, 2006). Likewise, 2,3-epoxyneryl is used by acarid mites as sex pheromone (Mori *et al.*, 1995). The acyclic terpene-derived epoxides have also been used as valuable building blocks for the synthesis of the aggregation pheromone of the Colorado potato beetle (Tashiro and Mori, 2005), alpha-damascone (Bovolenta *et al.*, 2004; Fráter, *et al.*, 1998; Sefton *et al.*, 1989), 2,2-dialkylchromanols (Inoue *et al.*, 1998) or antibiotic terpene (+)-tuberine (Taber, *et al.*, 1994) and decalins (Corey and Sodeoka, 1991).

As a cyclic monoterpene, (*R*)-(+)-limonene was chosen for our investigations. The conversion of (*R*)-(+)-limonene to several products like terpineol, (*cis*- and/or *trans*-) limonene-1,2-epoxide, and *cis*- and/or *trans*- carveol, showed that CYP109D1 is not stereo- and regioselective in the case of this substrate. Thus, the production of mutants of CYP109D1 with higher regio- and stereoselectivity of hydroxylation/epoxidation of limonene to obtain more valuable derivatives could be the focus of further studies. There have been many reports concerning the biotransformation of limonene with the aim towards the production of more valuable natural flavor compounds (Cadwallader *et al.*, 1989; Haudenschild *et al.*, 2000; Rama and Bhattacharyya, 1977 and van Dyk, *et al.*, 1998). One enzyme system that transforms limonene to perillyl alcohol was found in *Bacillus stearothermophilus* BR388. However, this enzyme system was not sufficiently regiospecific and significant quantities of carveol, carvone, and terpineol were produced (Oriel *et al.*, 1997). Furthermore, it was shown that the cytochrome P450 in *Mycobacterium* sp. strain HXN-1500 was able to specifically hydroxylate (*L*)-limonene at position 7 to produce enantiopure (-)-perillyl alcohol with a specific activity of 3 U/g (dry weight) of cell (van Beilen *et al.*, 2005). Likewise, (*R*)-limonene has been shown to be oxidized by CYP102A7 to racemic mixtures of (*R*)-limonene-

1,2-epoxide, limonene-8,9-epoxide and carveol (Dietrich *et al.*, 2008). It has also been shown to be regio- and stereoselectively converted by *Xanthobacter sp.* C20 to form (4*R*,8*R*)-limonene-8,9-epoxide (van der Werf, *et al.*, 2000). Interestingly, cytochrome P450 enzymes in liver microsomes of some animals have been shown to oxidize limonene to several oxidation products such as 1,2- and 8,9-epoxides, carveol (a product by 6-hydroxylation), perillyl alcohol (a product by 7-hydroxylation) (Watabe *et al.*, 1980, 1981; Jager *et al.*, 1999; van Beilen, 2005, Shimada *et al.*, 2002). It has also been reported that limonene enantiomers were oxidized to carveols and perillyl alcohols by CYP2C11 in liver microsomes of untreated rats and by CYP2B1 in those of phenobarbital-treated rats (Miyazawa *et al.*, 2002). Similarly, other studies showed that (+) - and (-)-limonene enantiomers were oxidized to respective carveol (6-hydroxylation) and perillyl alcohol (7-hydroxylation) metabolites by CYP2C9 and CYP2C19 in human liver microsomes (Miyazawa, *et al.*, 2002).

The study of terpenoid hydroxylation by CYP109D1 was further extended by the use of sesquiterpenoids as substrates. Most interestingly, CYP109D1 was able to catalyze the selective hydroxylation of the sesquiterpenes α - and β -ionone which are valuable fragrance constituents. Due to their organoleptic properties and the distinctive fine violet and rose scents, manufacturers of perfumes, soaps, cosmetics and fine chemicals have special interest on them (Ohloff, 1994; Pybus and Sell, 1999). Moreover, ionones are also appreciated as important synthetic building blocks (Fra'ter *et al.*, 1998; Buchecker *et al.*, 1973; Baraldi *et al.*, 1986; Ubelhart *et al.*, 1986; Colombo *et al.*, 1992). Among this, 4-hydroxy- β -ionone is an important intermediate for the synthesis of the carotenoids (Enzell *et al.*, 1977; Eschenmoser *et al.*, 1981, Brenna *et al.*, 2002) and also of deoxyabscisic acid, a synthetic analogue of the phytohormone abscisic acid (Oritani and Yamashita, 1984; Tamura and Nagao, 1970).

This study has demonstrated that CYP109D1 was able to convert α -ionone into 3-hydroxy- α -ionone that was confirmed by ^1H and ^{13}C NMR. The bioconversion of α -ionone and β -ionone to their corresponding mono-hydroxylated derivatives along with other compounds have been examined using a recombinant *Escherichia coli* whole cell system expressing cytochrome P450 SU1 and SU2, or P450 SOY (Celik A, *et al.*, 2005). Likewise, P450BM-3 of *Bacillus megaterium* also showed the conversion of α -ionone into several mixtures of different hydroxylated and/or epoxidated product (Carmichael and Wong, 2001). Similarly, the triple P450 BM-3 mutant (F78Q L188Q A74G) produced a mixture of different hydroxylated products from α -ionone (Appel *et al.*, 2001). Along with this, the P450BM-3 mutant F87V

also produced four products from α -ionone (Urlacher *et al*, 2006). In this regard, we showed that CYP109D1, despite being wild type, was able to give a single hydroxylated product (3-hydroxy- α -ionone) from α -ionone. To the best of our knowledge, the single hydroxylated product of α -ionone alone has not been obtained so far using a P450-catalyzed reaction, even for the rationally designed mutant of P450BM-3, which produced a mixture of other products along with 3-hydroxylated- α -ionone. We further demonstrated that CYP109D1 converted β -ionone into 4-hydroxy- β -ionone exclusively, which was confirmed by comparison to an authentic standard (Watanabe *et al*, 2007). The hydroxylation of β -ionone into 4-hydroxy- β -ionone was also published for wild type P450 BM-3 and its improved mutants (e.g. A74E F87V P386S) (Urlacher *et al*, 2006).

To further elucidate the structural basis for the regioselective hydroxylation catalyzed by CYP109D1, we performed docking studies using a homology model of CYP109D1. According to the computed interaction energies, both the α - and β -ionone should bind with micro-molar affinity. The obtained docking positions show that both ionones in corresponding orientations enable the experimentally observed regioselective hydroxylation (Figure 5). Other binding positions that would allow alternative hydroxylation in the 4-position of α -ionone (both *R* and *S*-isomer) were not found. Likewise, no plausible docking positions allowing the hydroxylation in the 3-position of β -ionone were obtained. Moreover, as mentioned above, we found a higher amount of corresponding products for α -ionone compared to β -ionone under identical reaction conditions. This agrees very well with the docking results. The distance between the iron atom, the putative ferryl oxygen, and the site of hydroxylation is considerably larger for β -ionone.

Finally, our results demonstrated that CYP109D1 acts as selective epoxidase or hydroxylase during terpene and terpenoid conversions which generates more valuable compounds for fragrance- and flavor-industry.

6. Literature review

1. Ahn, J.W., Jang, K.H., Chung, S.C., Oh, K.B., and Shin, J. (2008). Sorangiadenosine, a new sesquiterpene adenoside from the myxobacterium *Sorangium cellulosum*. *Org. Lett.* *10*, 1167-116.
2. Annous, B.A., Becker, L.A., Bayles, D.O., Labeda, D.P. and Wilkinson, B.J. (1997). Critical role of anteiso-C15:0 fatty acid in the growth of *Listeria monocytogenes* at low temperatures. *Appl. Environ. Microbiol.* *63*, 3887–3894.
3. Appel, D., Lutz-Wahl, S., Fischer, P., Schwaneberg, U. and Schmid, R.D., (2001). A P450 BM-3 mutant hydroxylates alkanes, cycloalkanes, arenes and heteroarenes. *J. Biotechnol.* *88*, 167–171.
4. Archelas, A., Delbecque, J.P. and Furstoss, R. (1993). Microbiological transformations. 30. Enantioselective hydrolysis of racemic epoxides: the synthesis of enantiopure insect juvenile hormone analogs (Bower's compound). *Tetrahedron Asymmetry.* *4*, 2445–46.
5. Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics*, *22*, 195-201.
6. Atkins, W.M., and Sligar, S.G. (1988). The roles of active site hydrogen bonding in cytochrome P-450cam as revealed by site-directed mutagenesis. *J. Biol. Chem.* *263*, 18842-18849.
7. Bagby, M.O. and Calson, K.D. (1989). Chemical and biological conversion of soybean oil for industrial products. In: Cambie RC (ed) *Fats for the future*, Ellis Horwood, Chichester (UK), pp. 301–317.
8. Bajpai, V., Shin, S.Y., Kim, M.J., Kim, H.R. and Kang, S.C. (2004). Antifungal activity of bioconverted oil extract of linoleic acid and fractionated dilutions against phytopathogens *Rhizoctonia solani* and *Botrytis cinerea*. *Agric. Chem. Biotechnol.* *47*, 199–204.
9. Baraldi, P.G., Barco, A., Benetti, S., Pollini, G.P., Polo, E. and Simoni, D. (1986). The intramolecular nitrile oxide cycloaddition route to forskolin. *J. Chem. Soc. Chem. Commun.* 757-758.
10. Batrakov, S.G., Mosezhnyi, A.E., Ruzhitsky, A.O., Sheichenko, V.I. and Nikitin, D.I. (2000). The polar-lipid composition of the sphingolipid-producing bacterium *Flectobacillus major*. *Biochim. Biophys. Acta.* *1484*, 225–240.

11. Bernhardt, R. (1995). Cytochrome P450: structure, function and generation of reactive oxygen species. *Rev. Physiol. Biochem. Pharmacol.* *127*, 1-80.
12. Bhawal, B.M, Joshi, S.N., Krishnaswamy, D., Deshmukh, A.R. (2001). (+)-3-Carene, an efficient chiral pool for the diastereoselective synthesis of β -lactams. *J. Indian. Inst. Sci.* *81*, 265–276.
13. Boddupalli, S.S., Estabrook, R.W., and Peterson, A.J. (1990). Fatty acid monooxygenase by cytochrome P-450 BM-3. *J. Biol. Chem.* *265*, 4233-4239.
14. Bode, H.B., Dickschat, J.S., Kroppenstedt, R.M., Schulz, S., and Muller, R. (2005). Biosynthesis of iso-fatty acids in myxobacteria: iso-even fatty acids are derived by alpha-oxidation from iso-odd fatty acids. *J. Am. Chem. Soc.* *127*, 532-533.
15. Bode, H.B., Ring, M.W., Kaiser, D., David, A.C., Kroppenstedt, R.M., and Schwar, G. (2006). Straight-chain fatty acids are dispensable in the myxobacterium *Myxococcus xanthus* for vegetative growth and fruiting body formation. *J. Bacteriol.* *188*, 5632-5634.
16. Bovolenta, M., Castronovo, F., Vadala, A., Zanoni, G., and Vidari, G. (2004). A Simple and Efficient Highly Enantioselective Synthesis of α -Ionone and α -Damascone. *J. Org. Chem.* *69*, 8959-8962.
17. Bowers, E.S. (1969). Juvenile hormone: activity of aromatic terpenoid ethers. *Science.* *164*, 323-325.
18. Brenna, E., Fuganti, C., Serra, S. and Kraft, P. (2002). Optically Active Ionones and Derivatives: Preparation and Olfactory Properties. *Eur. J. Org. Chem.*, 967-978.
19. Buchecker, R., Egli, R., Regel-Wild, H., Tschärner, C., Eugster, C.H., Uhde, G., Ohloff, G. (1973). Absolute Konfiguration der enantiomeren α -cyclogeraniumsäuren, α -cyclogeraniale, α -Jonone, β -Jonone, α - und ϵ -Carotine. *Helvetica Chimica. Acta.* *56*, 2548-2563.
20. Budde, M., Morr, M., Schmid, R.D., and Urlacher, V.B. (2006). Selective hydroxylation of highly branched fatty acids and their derivatives by CYP102A1 from *Bacillus megaterium*. *Chembiochem.* *7*, 789-794.
21. Buehler, B. and Schmid, A. (2004). Process implementation aspects for biocatalytic hydrocarbon oxyfunctionalization. *J. Biotechnol.* *113*, 183–210.
22. Burdock, G. A. (1995). Fenaroli's handbook of flavour ingredients, 3rd ed., pp. 107. CRC Press, Boca Raton, Fla.
23. Cadwallader, K.R., Braddock, R.J., Parish, M.E. and Higgins, D.P. (1989). Bioconversion of (+)-limonene by *Pseudomonas gladioli*. *J. Food Sci.* *54*:1241-1245.

24. Capdevila, J.H., Wei, S., Helvig, C., Falck, J.R., Belosludtsev, Y., Truan, G., Graham-Lorence, S.E., and Peterson, J.A. (1996). The Highly Stereoselective Oxidation of Polyunsaturated Fatty Acids by Cytochrome P450BM-3. *J. Biol. Chem.* 271, 22663–22671.
25. Carmichael, A.B., and Wong, L.L. (2001). Protein engineering of *Bacillus megaterium* CYP102: the oxidation of polycyclic aromatic hydrocarbons. *Eur. J. Biochem.* 268, 3117–3125.
26. Celik, A., Flitsch, S.L., and Turner, N.J. (2005). Efficient terpene hydroxylation catalysts based upon P450 enzymes derived from actinomycetes. *Org Biomol Chem.* 3, 2930-2934.
27. Chang, H.C. and Oriel, P. (1994). Bioproduction of perillyl alcohol and related monoterpenes by isolates of *Bacillus stearothermophilus*. *J. Food Sci.* 59, 1241-1245.
28. Chun, Y.J., Shimada, T., Sanchez-Ponce, R., Martin, M.V., Lei, L., Zhao, B., Kelly, S.L., Waterman, M.R., Lamb, D.C., and Guengerich, F.P. (2007). Electron transport pathway for a *Streptomyces* Cytochrome P450: Cytochrome P450 105D5-catalyzed fatty acid hydroxylation in *Streptomyces coelicolor* A3(2). *J. Biol. Chem.* 282, 17486-17500.
29. Colombo, M.I., Zinczuk, J. and Ruveda, E. A. (1992). Synthetic routes to forskolin. *Tetrahedron.* 48, 963-1037.
30. Coon, M.J. (2005). Omega Oxygenases: Nonheme-iron enzymes and P450 cytochromes. *Biochem. Biophys. Res. Commun.* 338, 378-385.
31. Corey, E.J., and Sodeoka, M. (1991). An effective system for epoxide-initiated cation-olefin cyclization. *Tetrahedron Lett.* 32, 7005-7008.
32. Cropp, T.A., Smogowicz, A.A., Hafner, E.W., Denoya, C.D. and McArthur, H.A., (2000). Fatty-acid biosynthesis in a branched-chain alpha-keto acid dehydrogenase mutant of *Streptomyces avermitilis*. *Can. J. Microbiol.* 46, 506–514.
33. Crowell, P.L. and Gould, M.N. (1994). Chemoprevention of mammary carcinogenesis by hydroxylated derivatives of *d*-limonene. *Crit Rev Oncog*, 5, 1–22.
34. Cryle, M.J., Espinoza, R.D., Smith, S.J., Matovic, N.J., and De Voss, J.J. (2006). Are branched chain fatty acids the natural substrates for P450(BM3)? *Chem. Commun. (Camb)*. 2353-2355.
35. Curtis, P.D., Geyer, R., White, D.C., and Shimkets, L.J. (2006). Novel lipids in *Myxococcus xanthus* and their role in chemotaxis. *Environ. Microbiol.* 8, 1935-1949.

36. de Carvalho, C.C. and da Fonseca, M.M. (2006). Biotransformation of terpenes. *Biotechnol. Adv.* 24, 134–142
37. Dhere, S.G. and Dhavilkar, R.S. (1970). Microbial transformations of terpenoids: limonene. *Sci.Cult.* 7, 292.
38. Dietrich, M., Eiben, S., Asta, C., Do, T., Pleiss, J., and Urlacher, V. (2008). Cloning, expression and characterisation of CYP102A7, a self-sufficient P450 monooxygenase from *Bacillus licheniformis*. *Appl. Microbiol. Biotechnol.* 79, 931-940.
39. Donaldson, J.M.I., McGovern, T.P., Ladd, T.L.J. (1990). Floral attractants for the Cetoniinae and Rutelinae (Coleoptera: Scarabaeidae). *J. Econ. Entomol.* 83, 1298–1305.
40. Downard, J., and Toal, D. (1995). Branched-chain fatty acids: the case for a novel form of cell-cell signalling during *Myxococcus xanthus* development. *Mol Microbiol.* 16, 171-175.
41. Duetz, W.A., Bouwmeester, H., van Beilen, J.B., Witholt, B. (2003). Biotransformation of limonene by bacteria, fungi, yeast, and plants. *Appl. Microbiol. Biotechnol.* 61, 269–277.
42. Duetz, W.A., van Beilen, J.B., and Witholt, B. (2001). Using proteins in their natural environment: potential and limitations of microbial whole-cell hydroxylations in applied biocatalysis. *Curr. Opin. Biotechnol.* 12, 419-425.
43. Dyk, M.S., Kock, J.L.F., and Botha, A. (1994). Hydroxy long-chain fatty acids in fungi. *World J. Microb. Biot.* 10, 495-504.
44. Ebata, H., Toshima, K. and Matsumura, S. (2007). Lipase-catalyzed synthesis and curing of high-molecular weight polyricinoleate. *Macromol. Biosci.* 7, 798-803.
45. Ebata, H., Yasuda, M., Toshima, K., and Matsumura, S. (2008). Poly (ricinoleic acid) based novel thermosetting elastomer. *J. Oleo. Sci.* 57, 315-320.
46. Eltz, T., Ayasse, M., and Lunau, K. (2006). Species-Specific Antennal Responses to Tibial Fragrances by Male Orchid Bees. *J. Chem. Ecol.* 32, 71-79.
47. Enzell, C.R., Wahlberg, I. and Aasen, A.J.(1977). Isoprenoids and alkaloids of tobacco. *Fortschr. Chem. Org.Naturst.*, 34, 1.
48. Eschenfeldt, W.H., Zhang, Y., Samaha, H., Stols, L., Eirich, L.D., Wilson, C.R., and Donnelly, M.I. (2003). Transformation of fatty acids catalyzed by cytochrome P450 monooxygenase enzymes of *Candida tropicalis*. *Appl. Environ. Microbiol.* 69, 5992-5999.

49. Eschenmoser, W., Uevelhart, P., Eugster, C.H. (1981). Synthesis and structure of the enantiomeric 6-hydroxy- α -ionone and *cis*- and *trans*-5,6-dihydroxy-5,6-dihydro- β -ionone. *Helv. Chim. Acta.* *64*, 2681–2690.
50. Eugster, C.H., Maerki-Fischer, E. (1991). The chemistry of rose pigments. *Angew. Chem. Int. Ed. Engl.* *30*, 654–672.
51. Fautz, E., Rosenfelder, G., and Grotjahn, L. (1979). Iso-branched 2- and 3-hydroxy fatty acids as characteristic lipid constituents of some gliding bacteria. *J. Bacteriol.* *140*, 852-858.
52. Fer, M., Corcos, L., Dreano, Y., Plee-Gautier, E., Salaun, J.-P., Berthou, F., and Amet, Y. (2008). Cytochromes P450 from family 4 are the main omega hydroxylating enzymes in humans: CYP4F3B is the prominent player in PUFA metabolism. *J. Lipid Res.* *49*, 2379-2389.
53. Ferreira, A.C., Nobre, M.F., Rainey, F.A., Silva, M.T. and Wait, R. (1997). *Deinococcus geothermalis* sp. nov. and *Deinococcus murrayi* sp. nov., two extremely radiation-resistant and slightly thermophilic species from hot springs. *Int. J. Syst. Bacteriol.* *47*, 939–947.
54. Flath, R.A., Cunningham, R.T., LiqulDo, N.J., and McGovern, T.P. (1994). Alpha-Ionol as Attractant for Trapping *Bactrocera latifrons* (Diptera: Tephritidae). *Journal of Economic Entomology* *87*, 1470-1476.
55. Fra'ter, G., Bajgrowicz, J.A., Kraft, P. (1998). Fragrance chemistry. *Tetrahedron.* *54*, 7633-7703.
56. Gago, G., Kurth, D., Diacovich, L., Tsai, S.C., and Gramajo, H. (2006). Biochemical and structural characterization of an essential acyl coenzyme A carboxylase from *Mycobacterium tuberculosis*. *J Bacteriol.* *188*, 477-486.
57. Gerth, K., Pradella, S., Perlova, O., Beyer, S. and Müller, R. (2003). Myxobacteria: proficient producers of novel natural products with various biological activities-past and future biotechnological aspects with the focus on the genus *Sorangium*. *J. Biotechnol.* *106*, 233-253.
58. Giotis, E.S., McDowell, D.A., Blair, I.S., and Wilkinson, B.J. (2007). Role of branched-chain fatty acids in pH stress tolerance in *Listeria monocytogenes*. *Appl Environ Microbiol.* *73*, 997-1001. Epub 2006 Nov 1017.
59. Girhard, M., Schuster, S., Dietrich, M., Dürre, P., and Urlacher, V.B. (2007). Cytochrome P450 monooxygenase from *Clostridium acetobutylicum*: A new [alpha]-fatty acid hydroxylase. *Biochem.Biophys.Res. Commun.* *362*, 114-119.

60. Green, A., Rivers, S., Cheesman, M., Reid, G., Quaroni, L., Macdonald, I., Chapman, S., and Munro, A. (2001). Expression, purification and characterization of cytochrome P450 BioI: a novel P450 involved in biotin synthesis in *Bacillus subtilis*. *J. Biol. Inorg. Chem.* *6*, 523-533.
61. Groth, I., Schumann, P., Schutze, B., Augsten, K. and Stackebrandt, E. (2002). *Knoellia sinensis gen. nov., sp. nov.* and *Knoellia subterranea sp. nov.*, two novel actinobacteria isolated from a cave. *Int. J. Syst. Evol. Microbiol.* *52*, 77–84.
62. Guex, N., and Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling. *Electrophoresis*, *18*, 2714-2723.
63. Guo, W.J. and Tao, W.Y., (2008). Phoxalone, a novel macrolide from *Sorangium cellulosum*: structure identification and its anti-tumor bioactivity in vitro. *Biotechnol. Lett.* *30*, 349-356.
64. Gustafsson, M.C., Palmer, C.N., Wolf, C.R., and von Wachenfeldt, C. (2001). Fatty-acid-displaced transcriptional repressor, a conserved regulator of cytochrome P450 102 transcription in *Bacillus* species, *Arch. Microbiol.* *176*, 459-464.
65. Haag A, Eschenmoser W, Eugster CH (1980) Synthese von (-)-(R)-4-hydroxy- β -ionone and (-)-(5R,6S)-5-Hydroxy-4,5-dihydro- α -ionone aus (-)-(S)- α -ionone. *Helv Chim Acta* *63*:10–5
66. Hagvall, L., Baron, J.M., Borje, A., Weidolf, L., Merk, H., and Karlberg, A.T. (2008). Cytochrome P450-mediated activation of the fragrance compound geraniol forms potent contact allergens. *Toxicol. Appl. Pharmacol.* *233*, 308-313.
67. Hallahan, D.L., and West, J.M. (1995). Cytochrome P-450 in plant/insect interactions: geraniol 10-hydroxylase and the biosynthesis of iridoid monoterpenoids. *Drug Metabol. Drug Interact.* *12*, 369-382.
68. Hannemann, F., Virus, C., and Bernhardt, R. (2006). Design of an *Escherichia coli* system for whole cell mediated steroid synthesis and molecular evolution of steroid hydroxylases. *J. of Biotechnol.* *124*, 172-181.
69. Hardwick, J.P. (2008). Cytochrome P450 omega hydroxylase (CYP4) function in fatty acid metabolism and metabolic diseases. *Biochem. Pharmacol.* *75*, 2263-2275.
70. Hashizume, T., Imaoka, S., Mise, M., Terauchi, Y., Fujii, T., Miyazaki, H., Kamataki, T., and Funae, Y. (2002). Involvement of CYP2J2 and CYP4F12 in the Metabolism of Ebastine in Human Intestinal Microsomes. *J. Pharmacol. Exp. Ther.* *300*, 298-304.
71. Haudenschild, C., Schalk, M., Karp, F., and Croteau, R. (2000). Functional Expression of Regiospecific Cytochrome P450 Limonene Hydroxylases from Mint (*Mentha* spp.)

- in *Escherichia coli* and *Saccharomyces cerevisiae*. *Arch. Biochem. Biophys.* *379*, 127-136.
72. Hawkes, D.B., Adams, G.W., Burlingame, A.L., Ortiz de Montellano, P.R., and De Voss, J.J. (2002). Cytochrome P450(cin) (CYP176A), isolation, expression, and characterization. *J. Biol. Chem.*, *277*, 27725–27732.
 73. Holmes, V.E., Bruce, M., Shaw, P.N., Bell, D.R., Qi, F.M., and Barrett, D.A. (2004). A gas chromatography-mass spectrometry method for the measurement of fatty acid [omega] and [omega]-1 hydroxylation kinetics by CYP4A1 using an artificial membrane system. *Anal. Biochem.* *325*, 354-363.
 74. Hou, C.T. (2008). New bioactive fatty acids. *Asia Pac. J. Clin. Nutr.* *17*, 192-5.
 75. Hou, C.T. and Forman, R.J. (2000). Growth inhibition of plant pathogenic fungi by hydroxy fatty acids. *J. Ind. Microbiol. Biotechnol.* *24*, 275–276.
 76. Huey, R., Morris, G.M., Olson, A.J., and Goodsell, D.S. (2007). A Semiempirical Free Energy Force Field with Charge-Based Desolvation, *J. Comput. Chem.* *28*, 1145-1152.
 77. HYPERCHEM, version 6.02 (1999). Hypercube Inc., Gainsville, FL.
 78. Imaishi, H., and Petkova-Andonova, M. (2007). Molecular cloning of CYP76B9, a cytochrome P450 from *Petunia hybrida*, catalyzing the omega-hydroxylation of capric acid and lauric acid. *Biosci. Biotechnol. Biochem.* *71*, 104-113.
 79. Inoue, S., Asami, M., Honda, K., Shrestha, K.S, Takahashi, M. and Yoshino, T. (1998). *Synlett.* *679*.
 80. Ishida, T., Enomoto, H., and Nishida, R. (2008). New Attractants for Males of the Solanaceous Fruit Fly *Bactrocera latifrons*. *J. Chem. Ecol.* *34*, 1532-1535.
 81. Jager, W., Mayer, M., Platzer, P., Rezicek, G., Dietrich, H. and Buchbauer, G. (1999). Stereoselective metabolism of the monoterpene carvone by rat and human liver microsomes. *J. Pharm. Pharmacol.*, *52*, 191–197.
 82. Jahnke, L.L., Eder, W., Huber, R., Hope, J.M. and Hinrichs, K.U., (2001). Signature lipids and stable carbon isotope analyses of Octopus Spring hyperthermophilic communities compared with those of *Aquificales* representatives. *Appl. Environ. Microbiol.* *67*, 5179–5189.
 83. Janakiram, N.B., Cooma, I., Mohammed, A., Steele, V.E., and Rao, C.V. (2008). Beta-ionone inhibits colonic aberrant crypt foci formation in rats, suppresses cell growth, and induces retinoid X receptor-alpha in human colon cancer cells. *Mol. Cancer Ther.* *7*, 181-190.

84. Jones, S.L., Drouin, P., Wilkinson, B.J. and Morse, P.D. (2002). Correlation of long-range membrane order with temperature-dependent growth characteristics of parent and a cold-sensitive, branched-chain-fatty-acid-deficient mutant of *Listeria monocytogenes*. *Arch. Microbiol.* 177, 217–222.
85. Kaiser, D. (2003). Coupling cell movement to multicellular development in myxobacteria. *Nat. Rev. Micro.* 1, 45-54.
86. Kaneda, T. (1991). Iso- and anteiso-fatty acids in bacteria: biosynthesis, function and taxonomic significance. *Microbiol. Rev.* 55, 2, 288-302.
87. Kato, T., Yamaguchi, Y., Abe, N., Uyehara, T., Nakai, T., Yamanaka, S. and Harada, N. (1984). Unsaturated hydroxy fatty acids, the self-defensive substances in rice plant against rice blast disease. *Chem. Lett.* 25, 409–412.
88. Kearns, D.B., Venot, A., Bonner, P.J., Stevens, B., Boons, G.J., and Shimkets, L.J. (2001). Identification of a developmental chemoattractant in *Myxococcus xanthus* through metabolic engineering. *Proc. Natl. Acad. Sci. U. S. A.* 98, 13990-13994.
89. Keudell, K.C., Huang, J.K., Wen, L., Klopfenstein, W.E., Bagby, M.O., Lanser, A.C., Plattner, R.D., Peterson, R.E., and Weisleder, D. (2000). Fatty acids enhanced tubermycin production by *Pseudomonas* strain 2HS. *Microbios.* 102, 27-38.
90. Kim, D., Cryle, M.J., De Voss, J.J., and Ortiz de Montellano, P.R. (2007). Functional expression and characterization of cytochrome P450 52A21 from *Candida albicans*. *Arch. Biochem. Biophys.* 464, 213-220.
91. Kitazume, T., Yamazaki, Y., Matsuyama, S., Shoun, H., and Takaya, N. (2008). Production of hydroxy-fatty acid derivatives from waste oil by *Escherichia coli* cells producing fungal cytochrome P450foxy. *Appl. Microbiol. Biotechnol.* 79, 981-988.
92. Kniazeva, M., Crawford, Q.T., Seiber, M., Wang, C.Y., and Han, M. (2004). Monomethyl branched-chain fatty acids play an essential role in *Caenorhabditis elegans* development. *PLoS Biol.* 2, E257. Epub 2004 Aug 2031.
93. Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessières, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Codani, J.J., Connerton, I.F., Danchin, A., et al. (1997). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature*, 390, 249-56.
94. Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J. D.,

- Gibson, T. J., and Higgins D. G. (2007). CLUSTALW, version 2, *Bioinformatics*, 23, 2947-2948.
95. Larroche, C., Creuly, C. und Gros, J.-B. (1995). Fed-batch biotransformation of β -ionone by *Aspergillus niger*. *Appl. Microbiol. Biotechnol.* 43, 222-227.
 96. Lechevalier, M.P. (1977). Lipids in bacterial taxonomy - a taxonomist's view. *CRC Crit. Rev. Microbiol.* 5, 109–210.
 97. Li, Q.S, Schwaneberg, U., Fischer, P. and Smith R.D. (2000). Directed Evolution of the Fatty-Acid Hydroxylase P450 BM-3 into an Indole-Hydroxylating Catalyst. *Chemistry - A Eur. Jour.* 6, 1531-1536.
 98. Lochynski, S., Kowalska, K. and Wawrzenczyk, C. (2002). Synthesis and odour characteristics of new derivatives from the carane system. *Flavour Fragr. J.* 3, 181–186.
 99. Loeber, D.E., Russell, S.W., Toubé, T.P., Weedon, B.C.L. and Diment. J., (1971). Carotenoids and related compounds. Part XXVIII. Syntheses of zeaxanthin, β -cryptoxanthin, and zeinoxanthin (α -cryptoxanthin). *J. Chem. Soc. C1971.*, 404-408.
 100. Lösel, D. M. (1988). Fungal lipids. In *Microbial Lipids*, C., Ratledge, and S.G., Wilkinson ed. (San Diego:Academic Press), 1, pp. 699-806.
 101. Lutz-Wahl, S., Fischer, P., Schmidt-Dannert, C., Wohlleben, W., Hauer, B., and Schmid, R.D. (1998). Stereo- and Regioselective Hydroxylation of α -Ionone by *Streptomyces* Strains. *Appl. Environ. Microbiol.* 64, 3878-3881.
 102. McQuate, G.T., Peck, S.L. (2001). Enhancement of attraction of α -ionol to male *Bactrocera latifrons* (Diptera: Tephritidae) by addition of a synergist, cade oil. *J. Econ. Entomol.* 94,39–46.
 103. Merkel, G.J. and Perry, J.J. (1977). Effect of growth substrate on thermal death of thermophilic bacteria. *Appl. Environ. Microbiol.* 34, 626–629.
 104. Mikami, Y., Fukunaga, Y., Arita, M. und Kisaki, T. (1981). Microbial transformation of β -ionone and β -methylionone. *Appl. Environ. Microbiol.* 41, 610-617.
 105. Miyazawa, M., Shindo, M. and Shimada, T. (2001). Oxidation of 1,8-cineole, the monoterpene cyclic ether originated from *Eucalyptus polybractea*, by cytochrome P450 3A enzymes in rat and human liver microsomes. *Drug Metab. Dispos.* 29, 200–205.
 106. Miyazawa, M., Shindo, M. and Shimada, T. (2002). Sex differences in the metabolism of (+)- and (-)- limonene enantiomers to carveol and perillyl alcohol derivatives by cytochrome P450 enzymes in rat liver microsomes. *Chem. Res. Toxicol.* 15, 15-20.

107. Moraleda-Munoz, A., and Shimkets, L.J. (2007). Lipolytic enzymes in *Myxococcus xanthus*. *J. Bacteriol.* *189*, 3072-3080.
108. Morant, M., Jorgensen, K., Schaller, H., Pinot, F., Moller, B.L., Werck-Reichhart, D., and Bak, S. (2007). CYP703 Is an Ancient Cytochrome P450 in Land Plants Catalyzing in-Chain Hydroxylation of Lauric Acid to Provide Building Blocks for Sporopollenin Synthesis in Pollen. *Plant Cell.**19*, 1473-1487.
109. Mori, N., and Kuwahara, Y. (1995). Synthesis of (2R, 3R)-epoxyneral, a sex pheromone of the acarid mite, *Caloglyphus* sp. (Astigmata: Acaridae). *Tetrahedron Letters* *36*, 1477-1478.
110. Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., and Olson, A.J. (1998). Automated Docking Using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function. *J. Comput. Chem.* *19*, 1639-1662.
111. Newhall, W.F. (1958). Derivatives of (+)-Limonene. I. Esters of trans-p-Menthane-1,2-diol. *J. Org. Chem.* *23*, 1274-1276.
112. Nichols, D.S., Presser, K.A., Olley, J., Ross, T. and McMeekin, T.A. (2002) .Variation of branched-chain fatty acids marks the normal physiological range for growth in *Listeria monocytogenes*. *Appl. Environ. Microbiol.* *68*, 2809–2813.
113. Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., and Yura, T. (1998). Chaperone Coexpression Plasmids: Differential and Synergistic Roles of DnaK-DnaJ-GrpE and GroEL-GroES in Assisting Folding of an Allergen of Japanese Cedar Pollen, Cryj2, in *Escherichia coli*. *Appl. Environ. Microbiol.* *64*, 1694-1699.
114. Noble, M.A., Miles, C.S., Chapman, S.K., Lysek, D.A., MacKay, A.C., Reid, G.A., Hanzlik, R.P., and Munro, A.W. (1999). Roles of key active-site residues in flavocytochrome P450 BM3. *Biochem. J.* *339*, 371-379.
115. Noble, M.A., Miles, C.S., Chapman, S.K., Lysek, D.A., MacKay, A.C., Reid, G.A., Hanzlik, R.P., and Munro, A.W. (1999). Roles of key active-site residues in flavocytochrome P450 BM3. *Biochem. J.* *339*, 371-379.
116. Noda, C., Alt, G.P., Werneck, R.M., Henriques, C.A., and Monteiro, J.L.F. (1998). Aldol Condensation of Citral with Acetone on Basic Solid Catalysts. *Brazilian J. Chem. Eng.* *15*.
117. Noma, Y., Yamasaki, S. and Asakawa, Y. (1992). Biotransformation of limonene and related compounds by *Aspergillus cellulosa*. *Phytochemistry*, *31*, 2725-2727.
118. Ohloff, G. (1994). Scent and Fragrances: The Fascination of Fragrances and their Chemical Perspectives. Springer-Verlag, Berlin.

119. Oriel, P.J., Savithiry, S., and Chang, H.C. (1997). Process for the preparation of monoterpenes using bacterium containing recombinant DNA. U.S. patent 5,688,673.
120. Oritani, T. and Yamashita, K. (1994). Biological activity and structure of abscisic acid. *Kagaku to Seibutsu*, 22, 21-28.
121. Park, S., Jin, F. and Lee, J. (2004). Effect of biodegradable epoxidized castor oil on physicochemical and mechanical properties of epoxy resins. *Macromol. Chem. Phys.* 205, 2048-2054.
122. Peterson, J.A., Lu, J.Y., Geisselsoder, J., Grahamlorence, S., Carmona, C., Witney, F., and Lorence, M.C. (1992). Cytochrome P450terp–Isolation and purification of the protein and cloning and sequencing of its operon. *J. Biol. Chem.*, 267, 14193–14203.
123. Pybus, D.H. and Sell, C.S. (1999). *The Chemistry of Fragrances*. The Royal Society of Chemistry, London.
124. Rafidinarivo, E., Laneelle, M.A., Montrozier, H., Valero-Guillen, P., Astola, J., Luquin, M., Prome, J.C., and Daffe, M. (2008). Trafficking pathways of mycolic acids: structures, origin, mechanism of formation and storage form of mycobacteric acids. *J. Lipid Res.*, (in press).
125. Rama Devi, J. and Bhattacharyya, P. K. (1977). Microbiological transformations of terpenes. Part XXII. Fermentation of geraniol, nerol & limonene by a soil pseudomonad, *Pseudomonas incognita* (Linalool strain). *Indian J. Biochem. Biophys.* 14, 288-291.
126. Rastogi, C.S., Johansen, D.J. and Menne, T. (1996). Natural ingredients based cosmetics. *Contact Dermatitis*, 34, 423-426.
127. Rastogi, C.S., Heydorn, S., Johansen, D.J. and Basketter, A. D. (2001). Fragrance chemicals in domestic and occupational products. *Contact Dermatitis*, 45, 221–225.
128. Rilfors, L., Wieslander, A. and Stahl, S. (1978). Lipid and protein composition of membranes of *Bacillus megaterium* variants in the temperature range 5 to 70 degrees C. *J. Bacteriol.* 135, 1043–1052.
129. Ring, M.W., Schwar, G., Thiel, V., Dickschat, J.S., Kroppenstedt, R.M., Schulz, S., and Bode, H.B. (2006). Novel iso-branched ether lipids as specific markers of developmental sporulation in the myxobacterium *Myxococcus xanthus*. *J Biol Chem.* 281, 36691-36700.
130. Roberts, S.C. (2007). Production and engineering of terpenoids in plant cell culture. *Nat. Chem. Biol.* 3, 387-395.

131. Rohdich, F., Bacher, A., and Eisenreich, W. (2005). Isoprenoid biosynthetic pathways as anti-infective drug targets. *Biochem. Soc. Trans.* *33*, 785-791.
132. Rosenberg, E., and Dworkin, M. (1996). Autocides and a paracide, antibiotic TA, produced by *Myxococcus Xanthus*. *J. Ind. Microbiol.*, *17*, 424-431.
133. Russell, A. and Kenyon L.R. (1955). Pseudoionone. *Org. Synth.*, Coll. 3, 747.
134. Sangeewa G. Rupasinghe, H.D.M.A.S. (2007). Molecular definitions of fatty acid hydroxylases in *Arabidopsis thaliana*. *Proteins: Structure, Function, and Bioinformatics.* *68*, 279-293.
135. Sanner, M.F. (1999). Python: A Programming Language for Software Integration and Development. *J. Mol. Graph. Model.* *17*, 57-61
136. Schneiker, S., Perlova, O., Kaiser, O., Gerth, K., Alici, A., Altmeyer, M.O., Bartels, D., Bekel, T., Beyer, S., Bode, E., Bode, H.B., Bolten, C.J., Choudhuri, J.V., Doss, S., Elnakady, Y.A., Frank, B., Gaigalat, L., Goesmann, A., Groeger, C., Gross, F., Jelsbak, L., Jelsbak, L., Kalinowski, J., Kegler, C., Knauber, T., Konietzny, S., Kopp, M., Krause, L., Krug, D., Linke, B., Mahmud, T., Martinez-Arias, R., McHardy, A.C., Merai, M., Meyer, F., Mormann, S., Munoz-Dorado, J., Perez, J., Pradella, S., Rachid, S., Raddatz, G., Rosenau, F., Ruckert, C., Sasse, F., Scharfe, M., Schuster, S.C., Suen, G., Treuner-Lange, A., Velicer, G.J., Vorholter, F.-J., Weissman, K.J., Welch, R.D., Wenzel, S.C., Whitworth, D.E., Wilhelm, S., Wittmann, C., Blocker, H., Puhler, A., and Muller, R. (2007). Complete genome sequence of the myxobacterium *Sorangium cellulosum*. *Nat. Biotech.* *25*, 1281-1289.
137. Schnuch, A., Lessmann, H., Geier, J., Frosch, J.P. and Uter, W. (2004). Contact allergy to fragrances: frequencies of sensitization from 1996 to 2002. Results of the IVDK, *Contact Dermatitis*, **50**, 65–76.
138. Schrader, J. and Berger, R.G. (2001). Biotechnological production of terpenoid flavor and fragrance compounds. In: *Biotechnology, 2nd ed., Special Processes*. Rehm H-J, Reed G, Pühler A, Stadler P (eds), Wiley-VCH, Weinheim, New York, 373–422.
139. Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* *31*, 3381-3385.
140. Sefton, M.A., Skouroumounis, G.K., Massy-Westropp, R.A., Williams, P.J. (1989). Norisoprenoids in *Vitis vinifera* white wine grapes and the identification of a precursor of damascenone in these fruits. *Aust. J. Chem.* *42*, 2071–2084.
141. Sheldon, R.A. (2000). Atom efficiency and catalysis in organic synthesis. *Pure. Appl. Chem.* *72*, 1233–1246.

142. Shimada, T., Shindo, M., and Miyazawa, M. (2002). Species differences in the metabolism of (+)- and (-)-limonenes and their metabolites, carveols and carvones, by cytochrome P450 enzymes in liver microsomes of mice, rats, guinea pigs, rabbits, dogs, monkeys, and humans. *Drug Metab. Pharmacokinet.* *17*, 507-515.
143. Shimkets, L.J. (1999). Intercellular signaling during fruiting-body development of *Myxococcus xanthus*. *Annu. Rev. Microbiol.* *53*, 525-549.
144. Shin, S.Y, Kim, H.R., Kang, S.C. (2004). Antibacterial activity of various hydroxy fatty acids bioconverted by *Pseudomonas aeruginosa* PR3. *Agric. Chem. Biotechnol.* *47*, 205–208
145. Siristova, L., Melzoch, K., and Rezanka, T. (2008). Fatty acids, unusual glycolipids and DNA analyses of thermophilic bacteria isolated from hot springs. *Extremophiles* (in press).
146. Stanislaw, L., Kowaiska, K. and Wawrzynczyk, C. (2002). Synthesis and odour characteristics of new derivatives from the carane system. *Flavour and Fragrance Journal* *17*, 181-186.
147. Suutari, M. and Laakso, S. (1994). Microbial fatty acids and thermal adaptation. *Crit. Rev. Microbiol.* *20*, 285–328.
148. Taber, D.F., Bhamidipati, R.S., and Thomas, M.L. (1994). Cascade Cyclization: Synthesis of (+)-Tuberine. *The Journal of Organic Chemistry.* *59*, 3442-3444.
149. Tamura, S. and Nagao, M. (1970). *Syntheses and biological activities* of new plant growth inhibitors structurally related to abscisic acid. *Agric. Biol. Chem.*, *34*, 1393.
150. Tashiro, T., and Mori, K. (2005). Enzyme-assisted synthesis of (S)-1,3-dihydroxy-3,7-dimethyl-6-octen-2-one, the male-produced aggregation pheromone of the Colorado potato beetle, and its (R)-enantiomer. *Tetrahedron: Asymmetry* *16*, 1801-1806.
151. Teichmann, B., Linne, U., Hewald, S., Marahiel, M.A. and Bölker, M.(2007). A biosynthetic gene cluster for a secreted cellobiose lipid with antifungal activity from *Ustilago maydis*. *Mol. Microbiol.* *66*, 525-533.
152. Ubelhart, P., Baumeler, A., Haag, A., Prewo, R., Bieri, J.H. and Eugster, C.H. (1986). Optisch aktive 4,5-Epoxy-4,5-dihydro- α -ionone und Synthese der stereoisomeren 4,5:4',5'-Diepoxy-4,5,4',5'-tetrahydro- ϵ , ϵ -carotene und der sterische Verlauf ihrer Hydrolyse. *Helv. Chim. Acta.* *69*, 816-834.
153. Urlacher, V. and Schmid, R.D. (2002). Biotransformations using prokaryotic P450 monooxygenases. *Curr. Opin. Biotechnol.* *13*, 557–564.

154. Urlacher, V.B., Makhsumkhanov, A., and Schmid, R.D. (2006). Biotransformation of beta-ionone by engineered cytochrome P450 BM-3. *Appl. Microbiol. Biotechnol.* *70*, 53-59.
155. van Beilen, J.B., Holtackers, R., Luscher, D., Bauer, U., Witholt, B., and Duetz, W.A. (2005). Biocatalytic Production of Perillyl Alcohol from Limonene by Using a Novel *Mycobacterium* sp. Cytochrome P450 Alkane Hydroxylase Expressed in *Pseudomonas putida*. *Appl. Environ. Microbiol.* *71*, 1737-1744.
156. van de Vossenbergh, J.L., Driessen, A.J., da Costa, M.S. and Konings, W.N. (1999) Homeostasis of the membrane proton permeability in *Bacillus subtilis* grown at different temperatures. *Biochim. Biophys. Acta.* *1419*, 97–104.
157. van der Werf, M.J., Keijzer, P.M. and van der Schaft, P.H. (2000). *Xanthobacter* sp. C20 contains a novel bioconversion pathway for limonene. *J. Biotechnol.* *84*, 133–143.
158. van Dyk, M.S., van Rensburg E. and Moleleki, N. (1998). Hydroxylation of (+)limonene, (- α -pinene and (-) β -pinene by a *Hormonema* sp. *Biotechnol. Lett.* *20*, 431-436.
159. Ware, J.C., and Dworkin, M. (1973). Fatty acids of *Myxococcus xanthus*. *J. Bacteriol.* *115*, 253-261.
160. Watabe, T., Hiratsuka, A., Isobe, M. and Ozawa, N. (1980). Metabolism of *d*-limonene by hepatic microsomes to non-mutagenic epoxides toward *Salmonella typhimurium*. *Biochem. Pharmacol.*, *29*, 1068–1071.
161. Watabe, T., Hiratsuka, A., Ozawa, N. and Isobe, M. (1981). A comparative study on the metabolism of *d*-limonene and 4-vinylcyclohex-1-ene by hepatic microsomes. *Xenobiotica* *11*, 333–344.
162. Watanabe, Y., Laschat, S., Budde, M., Affolter, O., Shimada, Y., and Urlacher, V.B. (2007). Oxidation of acyclic monoterpenes by P450 BM-3 monooxygenase: influence of the substrate E/Z-isomerism on enzyme chemo- and regioselectivity. *Tetrahedron.* *63*, 9413-9422.
163. Werf van der, M.J., Swarts, H.J., and de Bont, J.A. (1999). *Rhodococcus erythropolis* DCL14 contains a novel degradation pathway for limonene. *Appl. Environ. Microbiol.* *65*, 2092-2102.
164. White, R.E., Groves, J.T. and McClusky, G.A. (1979). Electronic and steric factors in regioselective hydroxylation catalyzed by purified cytochrome P-450. *Acta. Biol. Med. Germ.* *38*, 475-482.

165. White, R.E., Miller, J.P., Favreau, L.V., and Bhattacharyya, A. (1986). Stereochemical dynamics of aliphatic hydroxylation by cytochrome P-450. *J. Am. Chem. Soc.*, *108*, 6024-6031.
166. Williams, R.J, Britton, G. Charlton, J.M, Goodwin, T.W. (1967). The stereospecific biosynthesis of phytoene and polyunsaturated carotenes. *Biochem J.* *104*, 767–777.
167. Winterhalter, P. and Rouseff, R.L. (2002). Carotenoid-derived aroma compounds. Washington DC, Am. Chem. Soc.
168. Withers, S.T., and Keasling, J.D. (2007). Biosynthesis and engineering of isoprenoid small molecules. *Appl. Microbiol. Biotechnol.* *73*, 980-990.
169. Wong, L.L., Bell, S.G., Carmichael, A.B. (2000). Process for oxidising terpenes. European Patent WO 00/31273, 2 June 2000.
170. Wust, M., and Croteau, R.B. (2002). Hydroxylation of Specifically Deuterated Limonene Enantiomers by Cytochrome P450 Limonene-6-Hydroxylase Reveals the Mechanism of Multiple Product Formation. *Biochemistry.* *41*, 1820-1827.
171. Wyatt, S.E. and Kue, J. (1992). The accumulation of β -ionone and 3-hydroxy esters of β -ionone in tobacco immunized by foliar inoculation with tobacco mosaic virus. *Phytopathology*, *82*, 580-582.
172. Yamazaki, Y., Hayashi, Y., Arita, M., Hieda, T. und Mikami, Y. (1988). Microbial conversion of α -ionone, α -methylionone, and α -isomethylionone. *Appl. Environ. Microbiol.* *54*, 2354-2360.
173. Zöllner, A., Kagawa, N., Waterman, M.R., Nonaka, Y., Takio, K., Shiro, Y., Hannemann, F., and Bernhardt, R. (2008). Purification and functional characterization of human 11beta hydroxylase expressed in *Escherichia coli*. *FEBS J.*, *275*, 799-810.

CHAPTER III

CYP260A1 from *Sorangium cellulosum* So ce56

1. INTRODUCTION

Genomic sequence of organisms has been determined at an accelerating rate and the genes encoding cytochrome P450 enzymes are also increasing. This large pool of the putative P450 genes has been considered as a potential resource for new versatile biocatalysts (Bernhardt, 2006). The genomic sequencing project of *Sorangium cellulosum* So ce56 also revealed several new P450 family members. Among them, CYP260A1 represents a novel bacterial P450 which could have potential function as a biocatalyst for the oxidative metabolism of endogenous and/or exogenous compounds. However, the identification of natural substrates for such novel P450 enzyme is not easy. Based on the sequence homologies with functionally characterized P450s enzymes, in some extent, it is possible to predict general functions of the P450 enzyme. Nevertheless, the sequence information of P450 alone is not sufficient enough to determine their catalytic functions or actual substrates.

Cytochrome P450 enzymes catalyze several reactions as described before in 'General Introduction'. Owing to the broad range of reactivities of the P450 enzymes, several *in vitro* screening techniques have been developed to access the potential substrates and inhibitors, including radiometric, fluorogenic, and rapid liquid chromatography/mass spectrometry (LC/MS)-based assays (Crespi *et al.*, 1997; Cohen *et al.*, 2003; Di *et al.*, 2007; Trubetskoy *et al.*, 2005; Furuya *et al.*, 2008; Sukumaran *et al.*, 2009; Rebe *et al.*, 2009). Among them, the most feasible and simple method is spin-shift where the heme prosthetic group of the cytochrome P450 serves as a useful chromophore during substrate and inhibitor binding studies. The oxidation of P450 bound substrate occurs at the heme, as a result certain aspects of substrate binding can be monitored readily with a spectrophotometer. Three specific types of substrate binding spectra (Type I, Type II and reverse Type I) have been identified (Schenkman *et al.*, 1981) as described before in 'General Introduction'. Generally, nitrogen-containing heme ligands, which bind more avidly than oxygen-containing ligands, are more often inhibitors than substrates (Gigon *et al.*, 1968). Many substrates (Type I ligands) displace the heme water ligand, shifting the iron spin equilibrium toward the high-spin form, whereas nitrogen heterocycles and anilines (Type II ligands) can replace the water ligand to stabilize the low-spin form. However, some Type II ligands can also act as substrates (Kunze *et al.*, 2006; Pearson *et al.*, 2006).

2. Objectives of Chapter III

1. To screen a compound library consisting of ~17,000 substances for CYP260A1.
2. To select structural analogues of the screening hits.
3. To determine and analyze substrates and their products.

3. MATERIALS AND METHODS

3. 1. Materials

The chemicals, reagents and other accessories used in this work were of high grade. They are illustrated in Appendix I.

3. 2. Methods

3. 2. 1. Screening of a compound library

The detailed procedure for the screening of a compound library was described in Chapter I.

3. 2. 2. Analytical methods

3. 2. 2. 1. *In vitro* steroid assay

The steroids, progesterone and DOC were used as substrates and were analyzed as described below.

Reactions were performed in a final volume of 0.5 ml of buffer E (Appendix III) containing 200 μ M steroid as a substrate and a NADPH-regenerating system, consisting of glucose-6-phosphate (5 mM), glucose-6-phosphate dehydrogenase (1 unit) and magnesium chloride (1 mM). The combinations of reconstituted redox partners CYP260A1: Adx: AdR of 1: 10: 1 (1 μ M: 10 μ M: 1 μ M) were used for assay purpose. During incubation at 30°C on a thermo shaker, NADPH was added to 500 μ M final concentration and the mixture was incubated for 15 min to 2 hr. The reactions were stopped with an equal volume of chloroform. The products were extracted with chloroform for two times. Thus obtained organic phases were vacuum dried, dissolved in 200 μ l of solvent acetonitrile:water (40:60) and were analyzed by HPLC as described below.

3. 2. 2. 2. *In vitro* nootkatone assay

The other substrate used in this work was nootkatone. The assay was done as described above with 200 μ M of nootkatone as a substrate. The samples were measured by HPLC or GC-MS as mentioned below.

3. 2. 2. 3. High performance liquid chromatography (HPLC) of steroids and nootkatone

HPLC is widely used to separate, identify and quantify the compounds. The sample to be separated is transported under pressure in the mobile phase through a column which holds the stationary phase. Separation is based on the differences of the solubility of substances in the

mobile and stationary phase. In this work, reversed phase HPLC (RP-HPLC) was used where the mobile phase was more polar than the stationary phase. Due to this, the more hydrophobic substance will be retained longer in the column.

After evaporation of the organic (chloroform) phase, the steroids (progesterone and DOC) and nootkatone were resuspended in acetonitrile and separated on a Jasco reversed phase HPLC system (Tokyo, Japan) composed of an auto sampler AS-2050 plus, pump PU-2080, gradient mixer LG-2080-02 and an UV-detector UV-2075. A reversed phase column (Nucleodur R100-5 C18ec, particle size 3 μm , length 125 mm, internal diameter 4 mm, Macherey-Nagel) or reversed phase column (150 Novapac C18, 3.9 id, 4 μm) was used to separate the steroids (progesterone and DOC) and nootkatone. Progesterone and nootkatone were monitored at 240 nm with retention time (RT) of 35 min whereas for the DOC same wavelength with RT of 10 min was used. In all cases, the column temperature was kept constant at 25°C with a peltier oven. The mobile phase was a mixture of acetonitrile:water (40:60) for both the substrates. 5 μl of the samples were injected for analysis. The peaks were identified by using the ChromPass software (V.1.7.403.1, Jasco). Pure steroids (>99%) were used as standards to identify the retention time (RT) of the peaks on HPLC. The separation of the progesterone was done by a gradient elution as mentioned in Table III.1.

Table III. 1: HPLC gradient used for the analysis of progesterone conversion. The gradient of phase A (40:60 acetonitrile/water) and phase B (100% acetonitrile) are shown.

Time (min)	0	10	15	18	22
Flow (ml/min)	1.0	1.0	1.0	1.0	1.0
A (%)	100	100	50	50	100
B (%)	0	0	50	50	0

3. 2. 2. 4. GC-MS analysis of nootkatone

The GC-MS measurement was performed in Institute of Technical Biochemistry, Universität Stuttgart. The analysis was performed on Shimadzu GCMS-QP2010 (column: FS-Supreme-5, length: 30 m, internal diameter: 0.25 mm, film thickness: 0.25 μm) using helium as a carrier gas. The mass spectra were collected using electron ionization. The column oven was programmed as in Table III. 2 for the measurement of nootkatone and its products.

Table III. 2. GC-MS analysis and identification of *in vitro* products of nootkatone

Column Oven Temp	150°C		
Injection Temp	250°C		
Program	Rate (°C/min)	Final Temperature (°C)	Hold (min)
	-----	150	4
	10	250	3
	50	300	1

4. Results

4. 1. Screening and validation of substrates/ligands

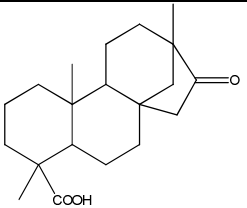
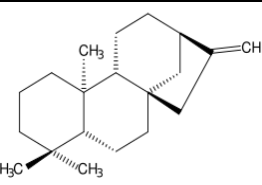
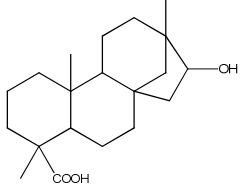
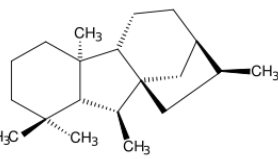
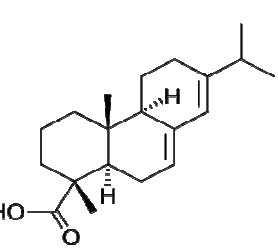
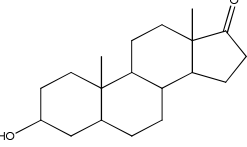
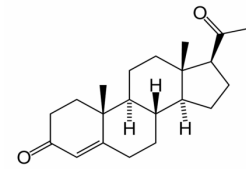
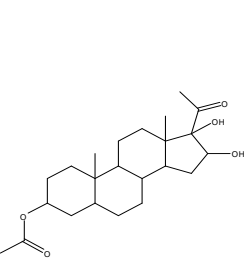
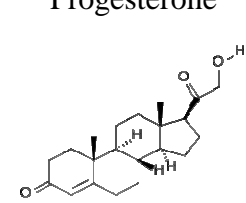
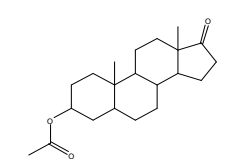
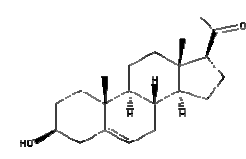
The myxobacterial cytochrome P450 CYP260A1 is one of the novel members of bacterial cytochrome P450s. The genomic information of this P450 (Chapter 1) was not able to give a precise hint for putative substrates. To solve this, a compound library consisting of ~17,000 substances was screened at the Forschungsinstitut für Molekulare Pharmakologie in Berlin, in cooperation with Prof. Oschkinat/Dr. von Kries.

During screening, binding spectra of potential ligands with CYP260A1 were observed and the potential substrates (having Type I binding) and the potential inhibitors (having Type II binding) were identified. For this, primary screening was done with 1.5 μ M of CYP260A1 and 0.5 μ M of ligand. Then, each of the primary screening hits was validated again with 1.5 μ M each of CYP260A1 and the corresponding ligand. After this, concentration dependent validation was done with fixed concentration of CYP260A1 (1.5 μ M) and respective ligands of 6.25 μ M, 12.5 μ M, 25 μ M, 50 μ M and 100 μ M. Altogether 33 positive hits (including both Type I and Type II ligands) were obtained. Among them, 15 ligands showed Type I binding and the rests 19 ligands showed Type II binding spectra. The detailed structures of the ligands with Chem-Div number are shown in Appendix VII (C). Validated representative Type I binding spectra for five of the ligands (Chem-Div nos: 25796, 25800, 25826, 26173 and 26267) are shown in Appendix VII (E).

4. 2. Selection of structural analogues of the Type I binding ligands

Though the screening of the compound-library identified Type I (potential substrates) and Type II (potential inhibitors) hits, the real physiological significance of those ligands is not known. As this work was more focused on the substrate identification, the best Type I validated hits (Chem-Div nos: 25796, 25800, 25826, 26173 and 26267) displaying absorption maxima at 390 were chosen and searched for analogous compounds. This is necessary, since the Chem. Div. library contains rather complicated compounds with limited accessibility. On the basis of the chemical structure of those ligands, the structural analogous compounds having distinct physiological functions were searched manually. The ligands 25796 and 25800 have the ring structure similar to that of *ent*-Kaurene, gibberellic acid and abietic acid (Table III. 3). Likely, the ligands 25826, 26173 and 26267 have an analogue ring-structure of steroids like progesterone, DOC and pregnenolone.

Table III. 3: Structure analogues of selected Type I binding hits of screening.

Type I binding ligand (Structure and Chem-Div nos)	Structure analogues of Type I ligand	Remark
 <p>25796_N062-0002</p>	 <p><i>ent</i>-Kauren</p>	<p><i>ent</i>-Kaurene is the intermediate product for gibberellin biosynthesis pathway</p>
 <p>25800_N062-006</p>	 <p><i>ent</i>-Gibberellin</p>	<p>Gibberellin is plant growth hormone</p>
	 <p>Abietic acid</p>	<p>Abietic acid (resin acid) is used in lacquers, varnishes and soaps and for the analysis of resins and the preparation of metal resinates.</p>
 <p>25826_N045-0002</p>	 <p>Progesterone</p>	<p>Progesterone, DOC, pregnenolone are the important steroid in human and mammalian</p>
 <p>26173_1563-0004</p>	 <p>11- deoxycorticosterone (DOC)</p>	
 <p>26267_N050-0003</p>	 <p>Pregnenolone</p>	

4. 3. Binding of steroids to CYP260A1

In this work, to study the substrate binding and its conversion, progesterone and DOC (the analogue structure of 25826 and/or 26173 and /or 26267) were selected. The selected steroids were tested by difference spectroscopy. Both steroids were able to induce a high spin shift (Type I binding) representing the displacement of the water molecule bound to the sixth position of the heme iron (Figure III. 1-2).

They were further used to determine the apparent dissociation constant (K_D) and the maximum absorbance shift (A_{max}) by titration with increasing substrate concentrations. The K_D and A_{max} for progesterone were $7.28 (\pm 2.67) \mu\text{M}$ and $2.29 (\pm 0.22)$, respectively, with a correlation coefficient of 0.99 (Figure III. 1). For DOC, the calculated K_D was $8.49 (\pm 2.24) \mu\text{M}$ and A_{max} was $1.90 (\pm 0.14)$ (with a correlation coefficient of 0.99) (Figure III. 2).

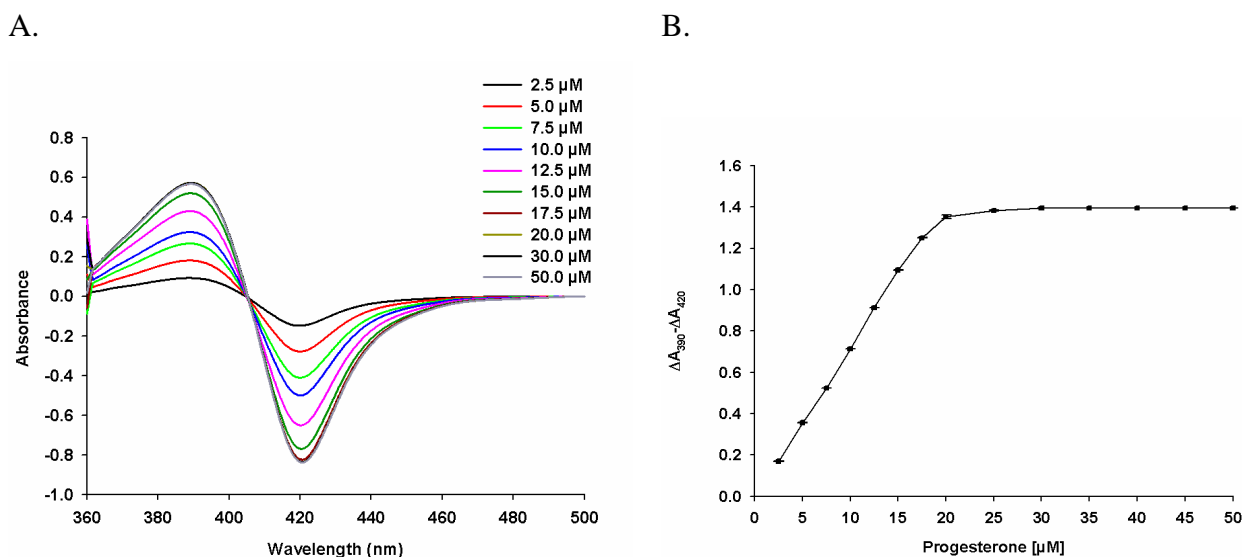


Figure III. 1: Type I binding spectra of CYP260A1 with progesterone. Varying concentrations of progesterone (dissolved in DMSO) were added to a solution of CYP260A1 (10 μM) in 10 mM potassium phosphate buffer (pH 7.4) containing 20% glycerol (v/v). **A. Difference spectra of CYP260A1 with progesterone.** The difference spectra with the starting spectrum subtracted from the subsequent traces. The concentrations of added progesterone were 2.5 (black line), 5.0 (red line), 7.5 (bright green line), 10.0 (blue line), 12.5 (pink line), 15.0 (dark green line), 17.5 (brown line), 20.0 (light green line), 30.0 (blue line), and 50.0 (grey line) μM . **B. Plot of absorbance change vs. concentration of progesterone (from part A).** The individual titration was performed for three times. The mean values and their standard deviations of the difference between absorption maxima and minima were calculated. The points were fitted to a quadratic equation (in Sigma Plot) and the calculated K_D and A_{max} with a correlation coefficient of 0.99 were $7.28 (\pm 2.67) \mu\text{M}$ and $2.29 (\pm 0.22)$, respectively. Each point represents the mean of three successive measurements.

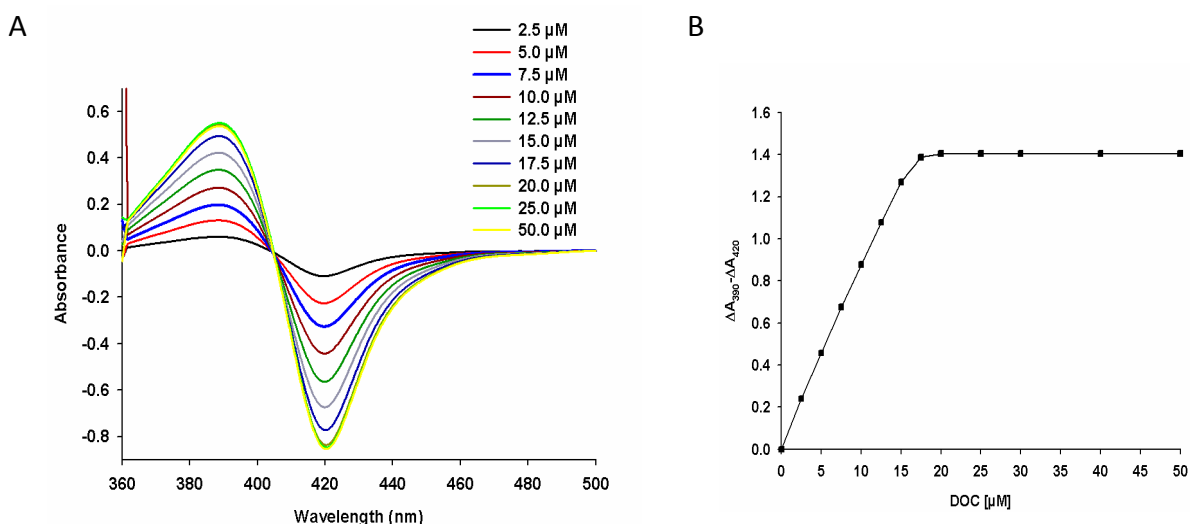


Figure III. 2: Type I binding spectra of CYP260A1 with DOC. Varying concentrations of DOC (dissolved in DMSO) were added to a solution of CYP260A1 (10 μM) in 10 mM potassium phosphate buffer (pH 7.4) containing 20% glycerol (v/v) **A. Difference spectra of CYP260A1 with DOC.** Difference spectra, with the starting spectrum subtracted from the subsequent traces. The concentrations of added DOC were 2.5 (black line), 5.0 (red line), 7.5 (blue line), 10.0 (brown line), 12.5 (dark green line), 15.0 (grey line), 17.5 (dark blue line), 20.0 (light green line), 25.0 (green line) and 50.0 (yellow line) μM . **B. Plot of absorbance change vs. concentration of DOC (from part A).** The individual titration was performed for three times. The mean values and their standard deviations of the difference between absorption maxima and minima were calculated. The points were fitted to a quadratic equation (in Sigma Plot) and the calculated K_D and A_{max} with a correlation coefficient of 0.99 were $8.49 (\pm 2.24) \mu\text{M}$ and $1.90 (\pm 0.14)$, respectively.

4. 4. Conversion of steroids by CYP260A1

The *in vitro* conversion of progesterone and DOC, the analogues of the Type I ligand hits of the screening, with CYP260A1 was analyzed by using HPLC. In this study, this technique was not used for the quantification purpose but, rather used to test the ability of the novel P450 CYP260A1 to hydroxylate (convert) the steroids (progesterone and DOC). The product peaks of progesterone and DOC were assigned with respect to the retention time (RT) of several standards of hydroxylated products of the progesterone which were run at the same time. The standard hydroxylated products of progesterone that are used in this work were 11α -, 6β -, 15β -, $11\alpha,6\beta$ -, 9α and $15\beta,11\alpha$ - hydroxylated progesterone. The applied HPLC condition was able to separate them (Figure III. 3). During this work, the RT for the standard hydroxylated progesterone of $15\beta,11\alpha$ -, $11\alpha,6\beta$ -, 11α -, and 9α - were 1.47, 1.64, 4.14 and 4.79 min, respectively. The applied method was also suitable enough to separate the respective substrates as well as the expected products (Figure III. 4-5).

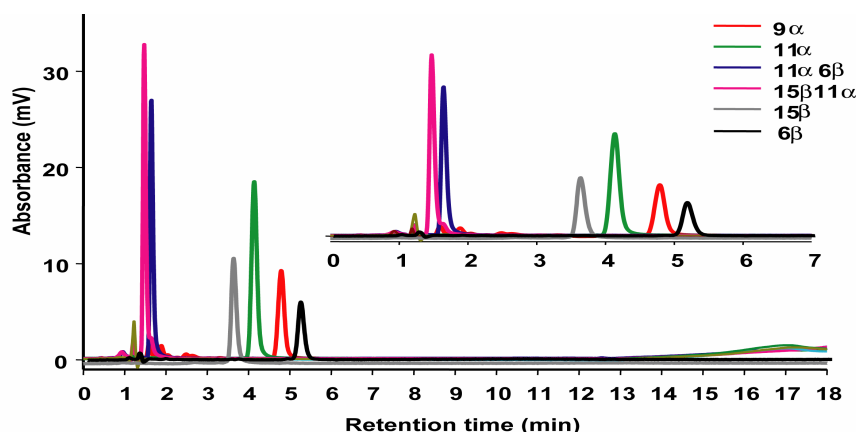


Figure III. 3: The hydroxylated and dihydroxylated standards progesterone peaks. The chromatograms were recorded at 240 nm after HPLC separation with a gradient solvent system consisting of phase A (acetonitrile/water 40:60) and phase B (100% acetonitrile) at a flow rate of 1 ml/min on a C18 reverse phase column. The applied method was able to separate all the hydroxylated and dihydroxylated standards of progesterone products of 9 α - (red line), 11 α - (green line), 11 α , 6 β - (blue line), 15 β , 11 α - (pink line), 15 β - (grey line) and 6 β - (black line).

4. 4. 1. Conversion of progesterone by CYP260A1

In presence of the heterologous redox partners, Adx and AdR, CYP260A1 was able to convert progesterone into several products. During HPLC analysis, progesterone and its products were eluted from the column within less than 20 min (Figure III. 4). The substrate peak was obtained at 16.08 min. During 1 hr incubation of the *in vitro* assay, one major product (peak II at 2.66 min), four minor products (peaks I, III, IV and V at 2.29, 4.67, 5.65 and 8.23 min, respectively) and five very minor products (peak a, b, c, d and e at 1.17, 1.58, 1.85, 3.32 and 4.09, respectively) were obtained (Figure III. 4).

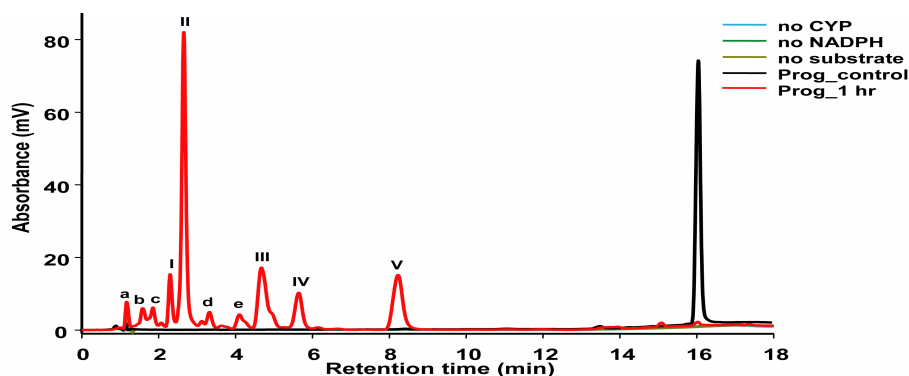


Figure III. 4: HPLC chromatogram of progesterone conversion by CYP260A1. The conversion (red line) and control, progesterone only, (black line) are shown. Several controls like the absence of CYP260A1 (blue line), absence of NADPH (dark green line) and absence of substrate (light green line) are also shown. The major products are represented by roman letter (I-V) and the minor products are represented by alphabets (a-e).

It was found that the product peak III (RT 4.67 min) has similar retention time as 9 α -OH-progesterone (RT 4.70 min) whereas the peak 'e' (RT 4.09 min) showed similar movement with 11 α -OH-progesterone (RT 4.10 min). The major product peak II (RT 2.66 min) can not be identified using the available standards as illustrated in Figure III. 5.

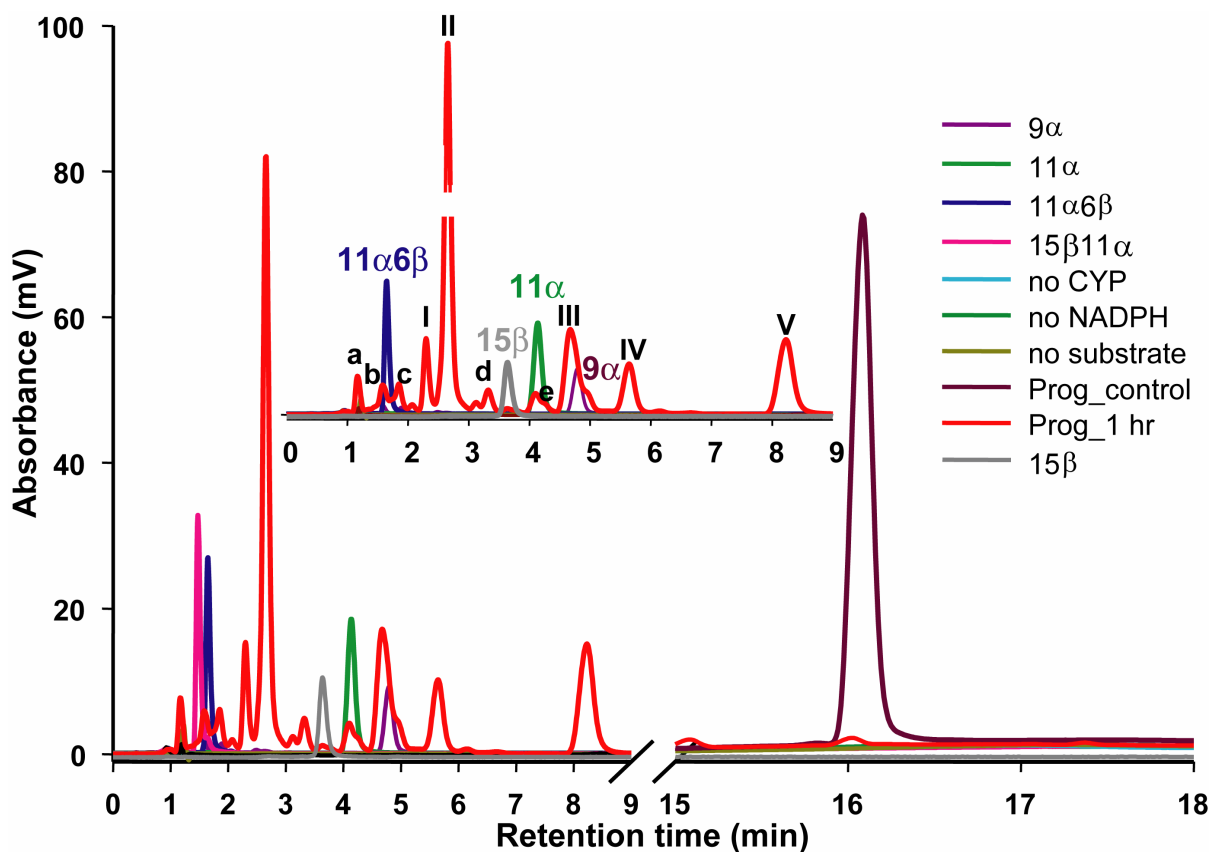


Figure III. 5: HPLC chromatogram of CYP260A1 dependent progesterone conversion and comparison with the hydroxylated and dihydroxylated standard products. The chromatograms of control, progesterone only, (brown line) and the conversion (red line) are shown. The chromatogram peaks of the hydroxylated and dihydroxylated standard progesterone of 9 α - (voilet line), 11 α - (green line), 11 α 6 β - (dark blue line), 15 β 11 α - (pink line) and 15 β -OH-progesterone (grey line) are overlaid with the conversion chromatogram. The inset illustrates the magnified view of the overlaid products parts. The product peak 'e' and peak III are shown to have similar RT as standard 11 α - (green line) and 9 α -OH-progesterone (violet line), respectively. The negative controls, with absence of both CYP109D1 and NADPH (blue line), absence of NADPH (dark green line) and absence of substrate (light green line) during assay are also shown.

4. 4. 2. Conversion of DOC by CYP260A1

CYP260A1 was also able to convert DOC with the reconstituted heterologous redox partners. The product profiles varied depending on the incubation time of the *in vitro* assay. Aldosterone and corticosterone were used as standards to compare the RT of the DOC products. The applied HPLC conditions were able to separate the standards (aldosterone and corticosterone) and the substrate DOC within less than 10 min (Figure III. 6). The substrate peak, DOC, has a RT of 7.40 min. the standards, aldosterone and corticosterone, have RT values of 1.81 and 3.27 min, respectively.

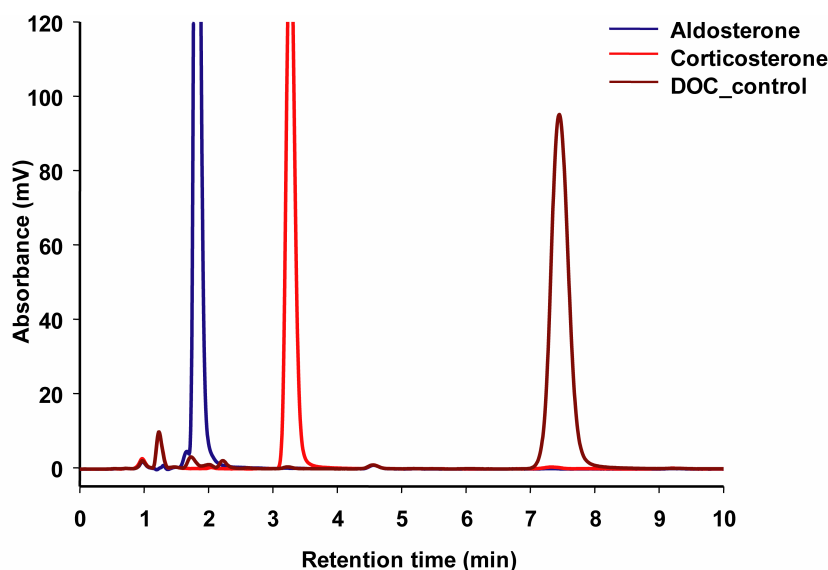


Figure III. 6: HPLC separation of aldosterone, corticosterone and DOC standards. The chromatograms were recorded with an isocratic solvent system consisting of acetonitrile/water (40:60) at a flow rate of 1 ml/min on a C18 reverse phase column. The aldosterone (blue line), corticosterone (red line) and DOC (brown line) are shown.

During 1 hr incubation of the *in vitro* assay for DOC conversion, one major product (peak II at 2.42 min), one minor product (peak I at 1.80 min), and some very minor products (peak a, b, c, d, e and f at 2.03, 2.12, 2.28, 3.54, 5.00 and 5.20 min, respectively) were obtained (Figure III. 7). The major product (peak II) could not be identified whereas the peak I showed similar movement (RT 1.80 min) as aldosterone (RT 1.81 min).

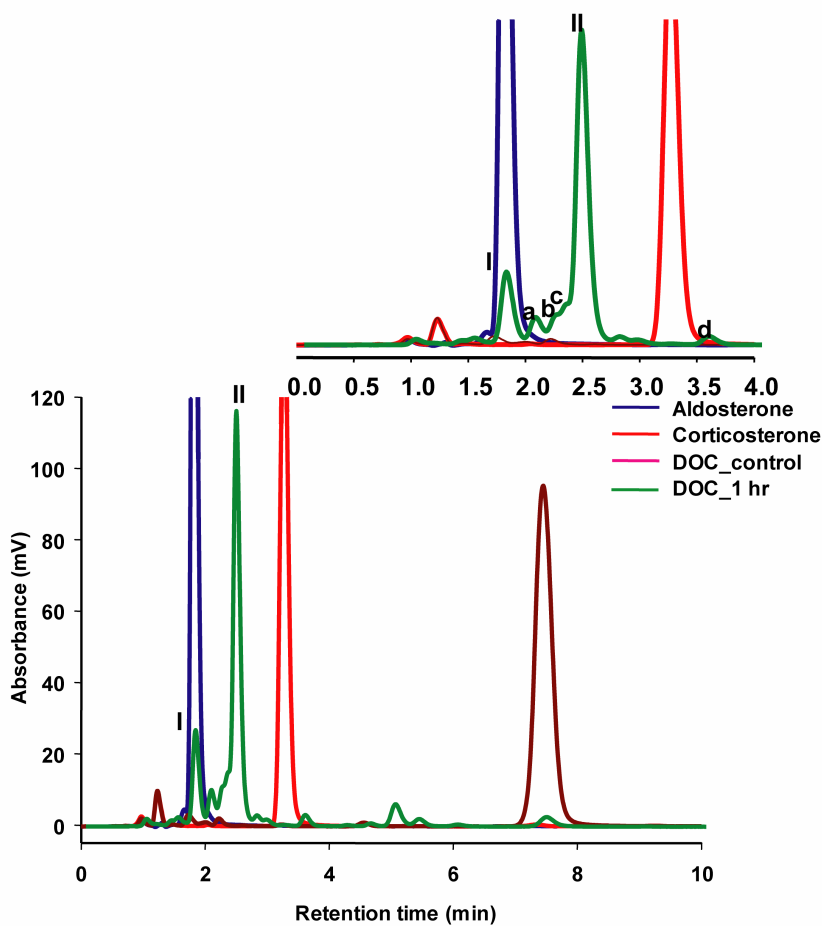


Figure III. 7: HPLC chromatogram of DOC conversion by CYP260A1. The substrate DOC (brown line) and its conversion products (green line) are shown. Aldosterone (blue line) and corticosterone (red line) are used as standard. The inset illustrates the magnified view of the product peaks. The products peak I and II as well as peaks a-f are illustrated.

4. 4. 2. 1. Time dependence of DOC conversion by CYP260A1

Time dependent DOC hydroxylation by CYP260A1 was also carried out. Each of the independent *in vitro* assays was stopped after 6, 15, 30, 45 and 60 min and was measured by HPLC. During successive increment of the incubation time from 5 min to 60 min of the *in vitro* assay, the substrate, DOC, peak was gradually diminished and after 60 min it was almost vanished (Figure III. 8). During this time, the abundance of the product peaks has increased, but the product patterns were remaining the same.

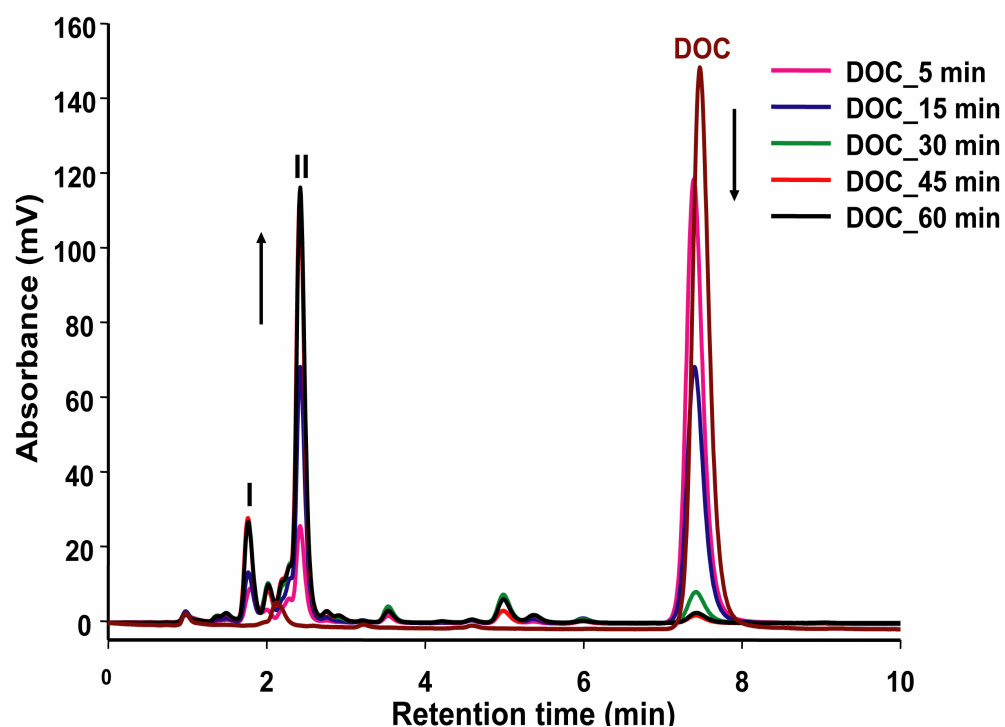


Figure III. 8: HPLC chromatogram of time dependent DOC conversion by CYP260A1. Time dependent *in vitro* conversion of DOC was carried out and was measured by HPLC. The product obtained after 5 (pink line), 15 (blue line), 30 (green line), 45 (red line) and 60 min (black line) incubation of the assay are shown. The control substrate, DOC only (brown line) is also shown. The minor and major products, peak I and II are also illustrated. The down-arrow denotes the decrement of the DOC and the up-arrow denotes the increment of the product abundance during time interval.

The product patterns of the DOC were changed when the incubation time of the *in vitro* assay was increased to more than 1 hr, during which the main product peak II (RT 2.42 min) was gradually shifted to the peak I (RT 1.80 min) (Figure III. 9). After 2 hr of incubation, a single major product peak I (RT 1.80 min) was obtained and the former major peak II (RT 2.42 min) obtained at 1 hr of incubation was almost lost. On comparison with the standard, it has been found that the movement of the peak I has almost similar RT with that of standard aldosterone (RT 1.81 min) (Figure III. 10), which could be a poly or dihydroxylation from the peak II to I.

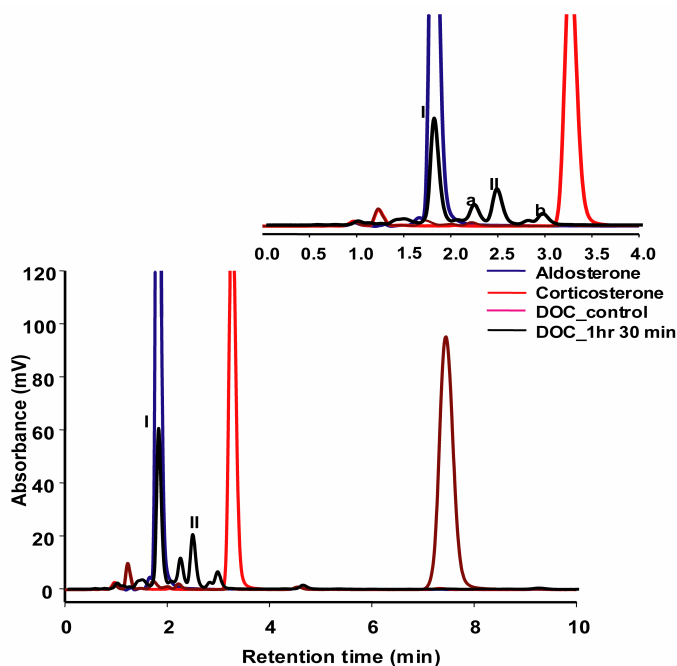


Figure III. 9: HPLC chromatogram of CYP260A1 dependent DOC conversion after 1 hr 30 min of incubation. The substrate DOC (brown line) and its conversion (black line) are shown. Aldosterone (blue line) and corticosterone (red line) are used as standards. The inset illustrates the magnified view of the product peaks. The products peak I and II as well as peaks a-b are shown.

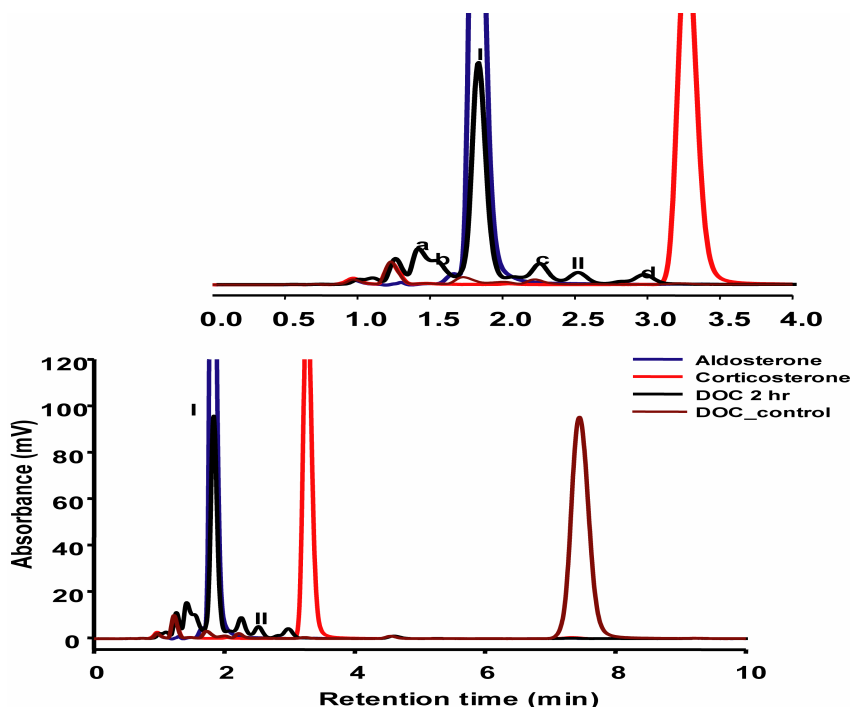


Figure III. 10: HPLC chromatogram of CYP260A1 dependent DOC conversion after 2 hr of incubation. The substrate DOC (brown line) and its conversion (black line) are shown. The aldosterone (blue line) and corticosterone (red line) are used as standards. The inset illustrates the magnified view of the product peaks. The products peak I and II as well as peaks a-d are shown.

4. 5. Conversion of nootkatone by CYP260A1

4. 5. 1. Selection of nootkatone as a substrate

The screening hits also guided that some of the sesquiterpenoid could also be a potential substrate for CYP260A1. The sesquiterpenes related compounds have several important functions in bacterial life cycle, and their alcohol, aldehyde and ketone derivatives are biologically active or precursors to other biological metabolites (Fraga, 2004). Due to this, all 21 P450s of the So ce56 were aligned (Appendix VIII. C) and a unrooted phylogenetic tree was constructed with the well characterized P450, like P450_{ditQ} from *Pseudomonas abietaniphila* BKME-9 and P450_{tdtD} from *Pseudomonas diterpeniphila* which are involved in the metabolism of terpenoids (Smith *et al*, 2004), and CYP226 from *Burkholderia xenovorans* LB400 involved in diterpenes catabolism (Smith *et al*, 2007) (Figure III. 11). It was found that CYP260A1 shows 40.5% similarity (24.7% identity) to P450_{ditQ} , 42% similarity (25.8% identity) to tdtD and 39.1% similarity (23.4% identity) to CYP226 from *B. xenovorans* LB400. In these reference, one of the sesquiterpenes, nootkatone, was selected as a potential substrate for CYP260A1.

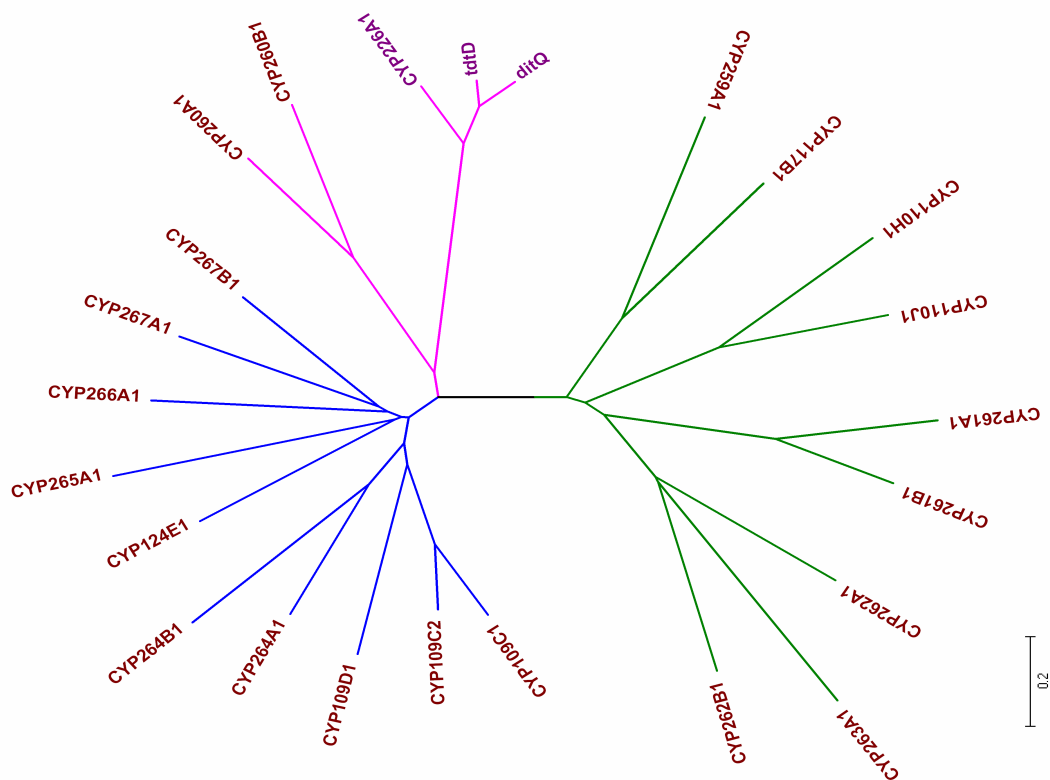


Figure III. 11: The unrooted tree of So ce56 CYPome with CYP226 of *Burkholderia xenovorans* LB400, CYP_{ditQ} of *Pseudomonas abietaniphila* BKME-9 and CYP_{tdtD} of *Pseudomonas diterpeniphila*. Only CYP260 families of So ce56 are in same clan with that of CYP226, CYP_{ditQ} and CYP_{tdtD} (cluster of pink branches). The bar on the right corner indicates 0.2 amino acid substitution per amino acid for the branch length.

4. 5. 2. Binding of nootkatone to CYP260A1

Nootkatone (NK) was able to induce a high spin shift. This was further used to determine the apparent dissociation constant (K_D) and the maximum absorbance shift (A_{\max}) obtained at the saturation with the nootkatone. The K_D and A_{\max} were calculated to be $7.87 (\pm 3.16) \mu\text{M}$ and $2.27 (\pm 0.19) \mu\text{M}$, respectively with the correlation coefficient of 0.99 (Figure III. 12). This showed that the nootkatone binds firmly with CYP260A1 and could acts as a good substrate.

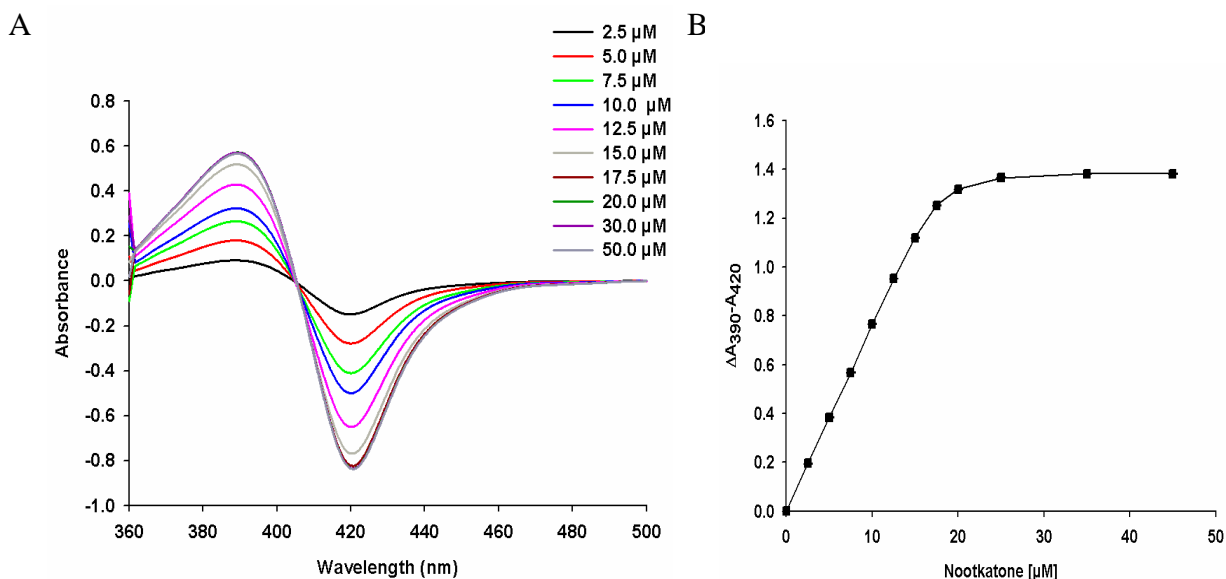


Figure III. 12: Type I binding spectra of CYP260A1 with nootkatone. Varying concentrations of nootkatone (dissolved in DMSO) were added to a solution of CYP260A1 ($10 \mu\text{M}$) in 10 mM potassium phosphate buffer ($\text{pH } 7.4$) containing 20% glycerol (v/v). The plot of the absorbance change vs. concentration of the substrate was fitted to a quadratic equation (in Sigma Plot), and apparent dissociation constant (K_D) and the maximum absorbance shift (A_{\max}) obtained at the saturation with the nootkatone were calculated. **A. Difference spectra of CYP260A1 with nootkatone.** Difference spectra, with the starting spectrum subtracted from the subsequent traces. The concentrations of nootkatone added were 2.5 (black line), 5.0 (red line), 7.5 (green line), 10.0 (blue line), 12.5 (pink line), 15.0 (grey line), 17.5 (brown line), 20.0 (dark green line), 30.0 (purple line) and 50.0 (dark grey) μM . **B. Plot of absorbance change vs. concentration of nootkatone (from part A).** Each of the titration was performed for three times. The mean values and their standard deviations of the difference between absorption maxima and minima were calculated. The points were fitted to a quadratic equation and yielded K_D and A_{\max} were $7.87 (\pm 3.16) \mu\text{M}$ and $2.27 (\pm 0.19)$, respectively, with a correlation coefficient of 0.99.

4. 5. 3. Conversion of nootkatone

CYP260A1 can convert nootkatone (NK) into two major and some minor products after reconstitution with heterologous redox partners, Adx and AdR. During HPLC analysis, the substrate and the products were eluted within 30 min from the column. The substrate peak was obtained at 29.35 min. During 1 hr incubation of the *in vitro* assay, two major products (peak III and peaks IV at 1.77 min and 4.70 min, respectively) and two minor products (peak I and II at 1.22 and 1.59 min, respectively) were observed (Figure III. 13). Time dependent *in vitro* assays were done from 15 min to 2 hr but the product patterns were not changed (Figure III. 14).

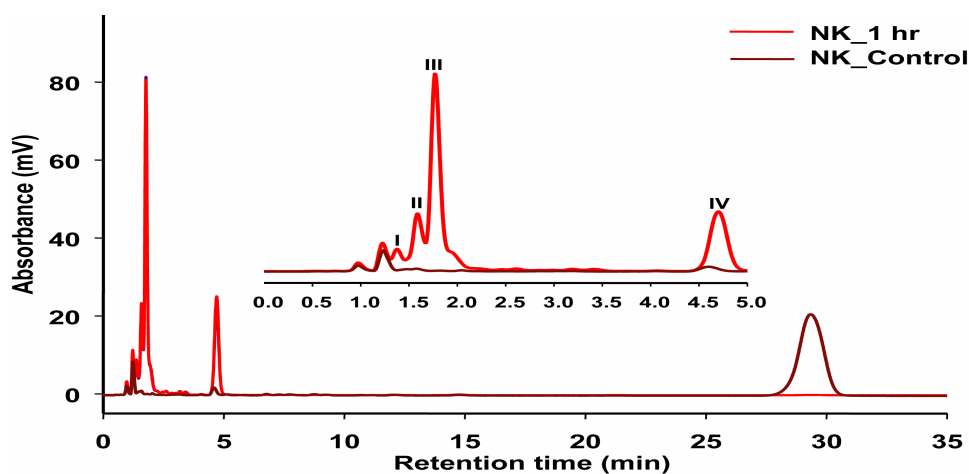


Figure III. 13: HPLC chromatogram of nootkatone conversion by CYP260A1. The chromatograms were recorded at 240 nm after HPLC separation with an isocratic system consisting of acetonitrile/water 40:60 at a flow rate of 1 ml/min on a C18 reverse phase column. The applied method was able to separate the products (red line) and the substrate, nootkatone, (brown line). The inset illustrates the magnified view of the product peaks I-IV.

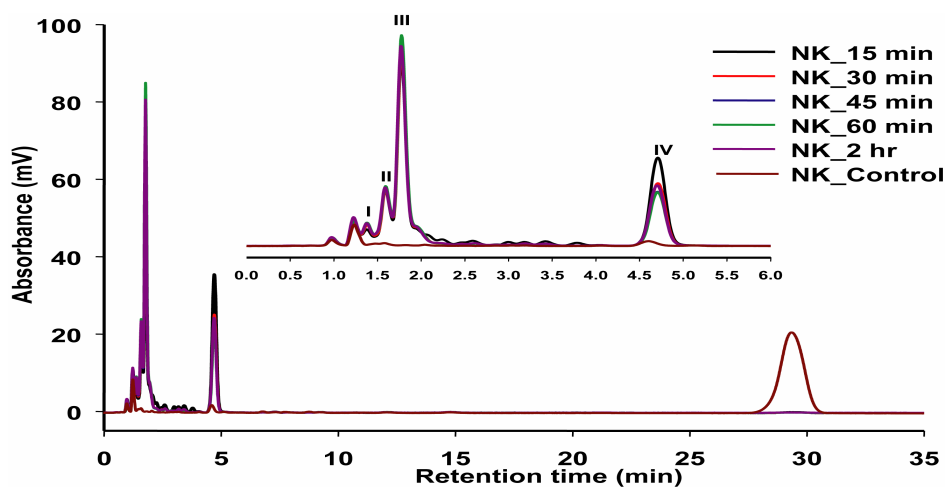


Figure III. 14: HPLC chromatogram of time dependent conversion of nootkatone by CYP260A1. The substrate nootkatone (brown line) and its conversion at 15 (black line), 30 (red line), 45 (blue line), 60 min (green line) and 2 hr (purple line) are shown. The insets illustrates the magnified view of the product peaks I-IV.

4. 5. 4. GC-MS of nootkatone conversion

During HPLC measurement, due to absence of the authentic standards of nootkatone derivatives, the product peaks were not assigned. Therefore, the conversion was also measured by GC-MS and mass spectra were obtained. The MS of peak III at 1.77 min and peak IV at 4.70 min were measured and their mass spectra were illustrated in Figure III.18 and Table III. 4. The total mass of the peak III was 234 which suggested the possible hydroxylation or epoxidation. The product peak IV has the mass value 232 corresponding to a possible oxygenated product. It was further found that peak I (RT 1.22 min) was nonspecific and peak II (1.59 min) was the regioisomer of peak III (Figure III.15).

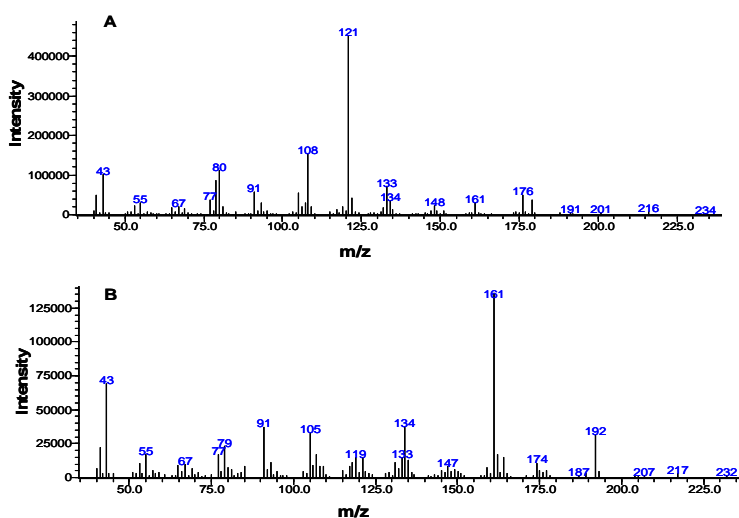


Figure III. 15: MS data of the product of nootkatone conversion. MS data of A. Peak III and B. Peak IV are shown.

Table III. 4: Possible hydroxylation, epoxidation or oxidation of nootkatone by CYP260A1. The possible hydroxylation and epoxidation sites for nootkatone with respect to the total mass value are shown.

<p>Illustration of nootkatone structure</p>	<p>The possible total mass value (m/z) of 234 could be due to hydroxylation or epoxidation</p> <ol style="list-style-type: none"> hydroxylation could be at position 3, 6, 7, 8 or 9 (the potential hydroxylation position at 15 would be m/z of 220.31) epoxidation at 1 and 10 or 13 and 14 (the double epoxidation would reveal m/z of 250.33). Possible total mass value (m/z) of 232 could be due to oxidation at 3, 6, 8 or 9.
---	---

5. Discussion

CYP260A1 is one of the novel P450s family members found in *Sorangium cellulosum* So ce56. Though the heterologous proteins Adx and AdR were identified as electron transfer partners (Chapter I), the substrates of this P450 were not known. Three approaches were implemented to explore potential substrates. Firstly, the genomic informations of the adjacent and neighboring genes of CYP260A1 were analyzed but all the proximal genes encoded for hypothetical proteins (Chapter I). As a result, the genomic informations were not sufficient enough to predict a putative substrate. Secondly, a compound-library consisting of 17,000 substances was screened and potential hits were identified. As a last approach, CYP260A1 was aligned to find the closest homologues of well characterized sesquiterpenoid converting P450s to get a hint on a potential substrate.

5. 1. Substrate screening and conversion of analogues compound.

Type I and Type II substrate binding modes have been extensively studied demonstrating spectral changes in the heme–Soret absorption band (Schenkman, *et al.* 1967; Kumaki *et al.*, 1978; Roberts *et al.*, 2005). The screening of the compound library carried out during this study enabled us to find potential substrates (Type I binding ligands) and inhibitors (Type II binding ligands). As we are focused on finding some potential substrates, only Type I hits were considered. Unfortunately, most of the hits have a complex structure and were not natural substances. Nevertheless, the chemical structure of the screening hits enabled us to search for analogous compounds having some potential physiological or biotechnological application. The structure analogues of the Type I binding ligands were searched (Table III. 3), and steroids (progesterone and DOC) were selected. CYP260A1 was able to hydroxylate both the steroids. Concerning the steroid hydroxylating bacterial cytochrome P450s, only very few bacteria has been described so far, among which 15 β -hydroxylase (CYP106A2) from *B. megaterium* ATCC 13368 has been extensively studied (Lisurek, *et al.*, 2004; Kang, *et al.*, 2004; Virus and Bernhardt, 2008). Several other steroid hydroxylases like CYP163A2 from *Streptomyces roseochromogenes* (Berrie *et al.*, 1999, 2001), CYP105D1 of *Streptomyces griseus* (Taylor *et al.*, 1999) and *Arthrobacter simplex* (Kominek *et al.*, 1993) have also been described.

5. 2. Microbial conversion of progesterone and 11-deoxycorticosterone (DOC)

Progesterone and 11-deoxycorticosterone (DOC) are important steroids having direct pharmaceutical implications. DOC is a precursor molecule for the production of aldosterone.

The major pathway for the aldosterone production is in the adrenal glomerulosa zone of the adrenal gland in mammals. It is produced from progesterone via 21-hydroxylation, 11-hydroxylation as well as 18-hydroxylation and 18-oxidation (Figure III. 16) (Smith *et al.*, 2008).

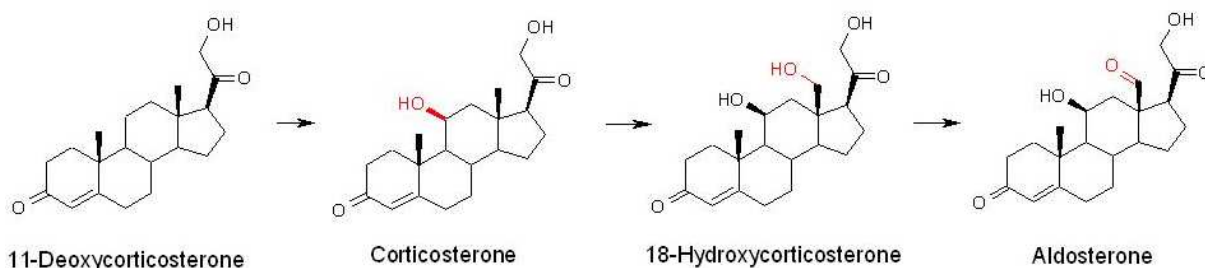


Figure III. 16: Conversion of DOC to aldosterone.

There are several published reports concerning the microbial conversion of progesterone and DOC. Most of the progesterone converting enzymes are from filamentous fungus. P450 mediated conversions of progesterone are shown in *Bacillus* sp., *Mycobacterium* sp. and *Streptomyces* sp.. The details of progesterone and DOC conversion in bacteria and fungus are illustrated in Table III. 5-6.

Table III. 5: Summary of literature data on progesterone conversion in (a) bacteria and (b) fungus.

a. Bacterial progesterone conversion

Organism/enzyme	Product	Reference
<i>Bacillus megaterium</i> ATCC 13368, 15 β -Hydroxylase (CYP106A2)	15 β -OH-Progesterone	Berg <i>et al.</i> , 1976
	11 α -OH-Progesterone	Lisurek <i>et al.</i> , 2004
	9 α -OH-Progesterone	Kang <i>et al.</i> , 2004
	6 β -OH-Progesterone	
<i>Bacillus thermoglucosidasium</i> Strain 12060 cytochrome P450	6 β -OH-Progesterone	Sideso <i>et al.</i> , 1998
	6 α -OH-Progesterone	
	Androstenedion	
<i>Bacillus stearothermophilus</i> cytochrome P450	20 α -OH-Progesterone	Al-Awadi <i>et al.</i> , 2001, 2002
	6 β -OH-Progesterone	
	6 α -OH-Progesterone	
	9,10-Seco-20 α -OH-Progesteron	
	6 β ,20 α -DiOH-Progesterone	

	6 α ,20 α -DiOH-Progesterone 5 α -Pregnen-3,6,20-trione 6-Dehydroprogesterone	
<i>Caldariella acidophila</i>	4-Pregnen-3,6,20-trione 5 α -Pregnan-3,6,20-trione 20 α -OH-4-pregnen-3-one 6 β -OH-Progesterone 6 α -OH-Progesterone	De Rosa <i>et al.</i> , 1981
<i>Mycobacterium smegmatis</i>	17 α -OH-Progesteron	Jenkins <i>et al.</i> , 2004
<i>Streptomyces roseochromogenes</i> strain 10984 Progesterone-16 α -Hydroxylase	16 α -OH-Progesterone 2 β ,16 α -DiOH-Progesterone	Berri <i>et al.</i> , 1999, 2001
<i>Pseudomonas</i> sp.	Androsta-1,4-dien-3,17-dione 17 β -Hydroxyandrosta-1,4-diene-3- on	Dhar and Samanta., 1993
<i>Rhodococcus rhodochrous</i> Baeyer-Villager-Monooxygenase	Testosteronacetate	Miyamoto <i>et al.</i> , 1995

b. Fungal progesterone conversion

Organism/enzyme	Product	Reference
<i>Acremonium strictum</i> PTCC 5282	15 α -OH-Progesterone 7 β ,15 β -DiOH-Progesterone 6 β ,11 α -DiOH-Progesterone 11 α ,15 β -DiOH-Progesterone 6 β ,11 α ,17 α -TriOH-Progesterone 11 α ,15 β ,17 α -TriOH-Progesterone 7 β ,15 β ,17 α -TriOH-Progesterone	Yoshihama, 1993 Faramarzi <i>et al.</i> , 2003
<i>Aspergillus fumigatus</i>	11 α -OH-Progesterone 15 β -OH-Progesterone 7 β -OH-Progesterone 7 β ,15 β -DiOH-Progesterone 11 α ,15 β -DiOH-Progesterone	Smith <i>et al.</i> , 1994
<i>Apergillus ochraceus</i>	11 α -OH-Progesterone	Samanta <i>et al.</i> ,

cytochrome P450			1987
<i>Cochliobolus lunatus</i>	M118	11 β -OH-Progesterone	Vita <i>et al.</i> , 1994
cytochrome P450		11 β ,14 α -DiOH-Progesterone	
		7 α ,14 α -DiOH-Progesterone	
		11-keto-14 α -OH-Progesterone	
<i>Fusarium culmorum</i>		15 α -OH-Progesterone	Kolek <i>et al.</i> , 1998
		12 β ,15 α -DiOH-Progesterone	
<i>Mortierella isabellina</i>		14 α -OH-Progesterone	Holland <i>et al.</i> , 1992, 95
<i>Mucor piriformis</i>		14 α -OH-Progesterone	Madyastha and
14 α -Hydroxylase		5 β ,14 α -DiOH-Progesterone	Joseph, 1993, 1996
		6 β ,14 α -DiOH-Progesterone	
		7 α ,14 α -DiOH-Progesterone	
		7 β ,14 α -DiOH-Progesterone	
<i>Phycomyces blakesleeanus</i>		7 α -OH-Progesterone	Smith <i>et al.</i> , 1991
		15 β -OH-Progesterone	Ahmed <i>et al.</i> , 1995
		6 β -OH-Progesterone	
		14 α -OH-Progesterone	
		15 α -OH-Progesterone	
<i>Penicillium decumbens</i>		15 α -OH-Progesterone	Holland <i>et al.</i> , 1995
ATCC10436			
<i>Rhizopus nigricans</i>		11 α -OH-Progesterone	Kim and Kim, 1991
cytochrome P450		6 β ,11 α -DiOH-Progesterone	
		11 α -OH-5 α -pregnan-3,20-dione	
<i>Trichophyton mentagrophyton</i>		15 α -OH-Progesterone	Clemons <i>et al.</i> , 1989
		11 α -OH-Progesterone	
		1-Dehydroprogesterone	

Table III. 6: Summary of literature data on DOC conversion in (a) bacteria and (b) fungi.

a. Bacterial DOC conversion

Organism/enzyme	Product	Reference
<i>Achromobacter kashiwasakiensis</i>	6 β -OH-RSS (RSS:11-deoxycortisol)	Tsuda <i>et al.</i> , 1959
<i>Bacillus</i> sp.	6 β -OH-RSS 15 α -OH-RSS 14-OH-RSS	Mahato and Garai, 1986
<i>Bacillus megaterium</i> ATCC 13368; CYP106A2	15 β -OH-RSS 6 β -OH-RSS 11 α -OH-RSS	Virus and Bernhardt, 2008
<i>Mycobacterium smegmatis</i>	Prednisolon	Ghanem and Yusef, 1992 a and b.

b. Fungal DOC conversion

Organism/enzyme	Product	Reference
<i>Apergillus fumigatus</i>	17 α ,21-Dihydroxy-5 α -pregn-1-en-3,20-dione 17 α , 20 α ,21-Trihydroxy-5 α -pregn-1-en-3-dione 6 β , 17 α ,21-Trihydroxypregn-4-en-3,20-dione 15 β , 17 α ,21-Trihydroxypregn-5 α -pregnan-3,20-dione	Garai and Mahato, 1992
<i>Beauveria bassiana</i>	11 α -OH-RSS (Epihydrocortison)	Zeng and Zhang, 1992
<i>Cunninghamella elegans</i>	11 α -OH-RSS (Epihydrocortison) 11 β -OH-RSS (Cortisol)	Dlugonski <i>et al.</i> , 1991
<i>Curvularia lunata</i>	11 β -OH-RSS (Cortisol)	Zuidweg <i>et al.</i> , 1962
<i>Hormodendrum olivaceum</i>	15 α -OH-RSS	Bernstein <i>et al.</i> , 1960 Allen <i>et al.</i> , 1961
<i>Spicaria simplicissima</i>	15 β -OH-RSS	Bernstein <i>et al.</i> , 1960 Allen <i>et al.</i> , 1961

5. 3. Triterpenoids and their biosynthesis in myxobacteria

Triterpenes belong to a large group of compounds and are arranged in a four or five ring configuration of 30 carbons. They are assembled from a C₅ isoprene unit through the cytosolic mevalonate pathway to make a C₃₀ compound. It has been published that the triterpenoids function as an important pharmaceutical compounds and showed antimicrobial (Chen *et al.*, 2008; Yadava and Jharbade, 2008; Horiuchi *et al.*, 2007; Ovenden *et al.*, 2005), antiviral (Wu *et al.*, 2007) and cytotoxic (Zhang *et al.*, 2008; Sogno *et al.*, 2009) activities.

The triterpenoid lipids like steroids are synthesized in almost all eukaryotic cells and are used to control the fluidity and flexibility of their cell membranes (Ourisson and Nakatani, 1994). In mammals, steroid hormones are mainly being synthesized in the adrenal and sexual glands (Bernhardt, 2006) and play an important role as glucocorticoids, mineralocorticoids and sex hormones. Whereas, bacteria rarely synthesize such compounds but frequently replace them with a different class of triterpenoids, the pentacyclic hopanoids as building blocks in their membranes (Ourisson *et al.*, 1987; Kannenberg and Poralla, 1999). The biosynthesis of hopanoids and steroids starts from squalene, a linear precursor that is formed by condensation of six isopentenyl units. The biosynthetic steps giving rise to the cyclic products belong to the most complicated reactions in nature catalysed by one single protein, squalene-hopene cyclase (Shc), and different forms of (*S*)-2,3-oxidosqualene cyclases (Osc). The latter type of enzyme is responsible for the formation of the protosteryl cation, the first cyclization product of (*S*)-2,3-oxidosqualene (Figure III. 17) (Michal, 1999). A variety of further reactions leads to different cyclic triterpenes, which are the precursors of membrane sterols, such as the zoosterine cholesterol, the mycoesterine ergosterol (both derived from lanosterol) and the phytosterine sitosterol (derived from cycloartenol).

It has been shown that the steroids are found in other bacterial species like, *Methylococcus capsulatus* (Bird *et al.*, 1971) for physiological regulations. The myxobacteria like, *Nannocystis excedens* and *Polyangium* sp. have also been shown to produce steroids (Kohl *et al.*, 1983) that have rarely been found in prokaryotes. There are also several reports on the existence of steroids in other myxobacteria (Kohl *et al.*, 1983). Interestingly, Bode *et al.*, (2003), reported that myxobacteria *Stigmatella aurantiaca* also produce cycloartenol, the well-known first cyclization product of steroid biosynthesis in plants and algae. The physiological role of CYP260A1 in *Sorangium cellulosum* So ce56 is not known but the steroid

hydroxylase activity of CYP260A1 speculates to have some important role during its complex life cycle.

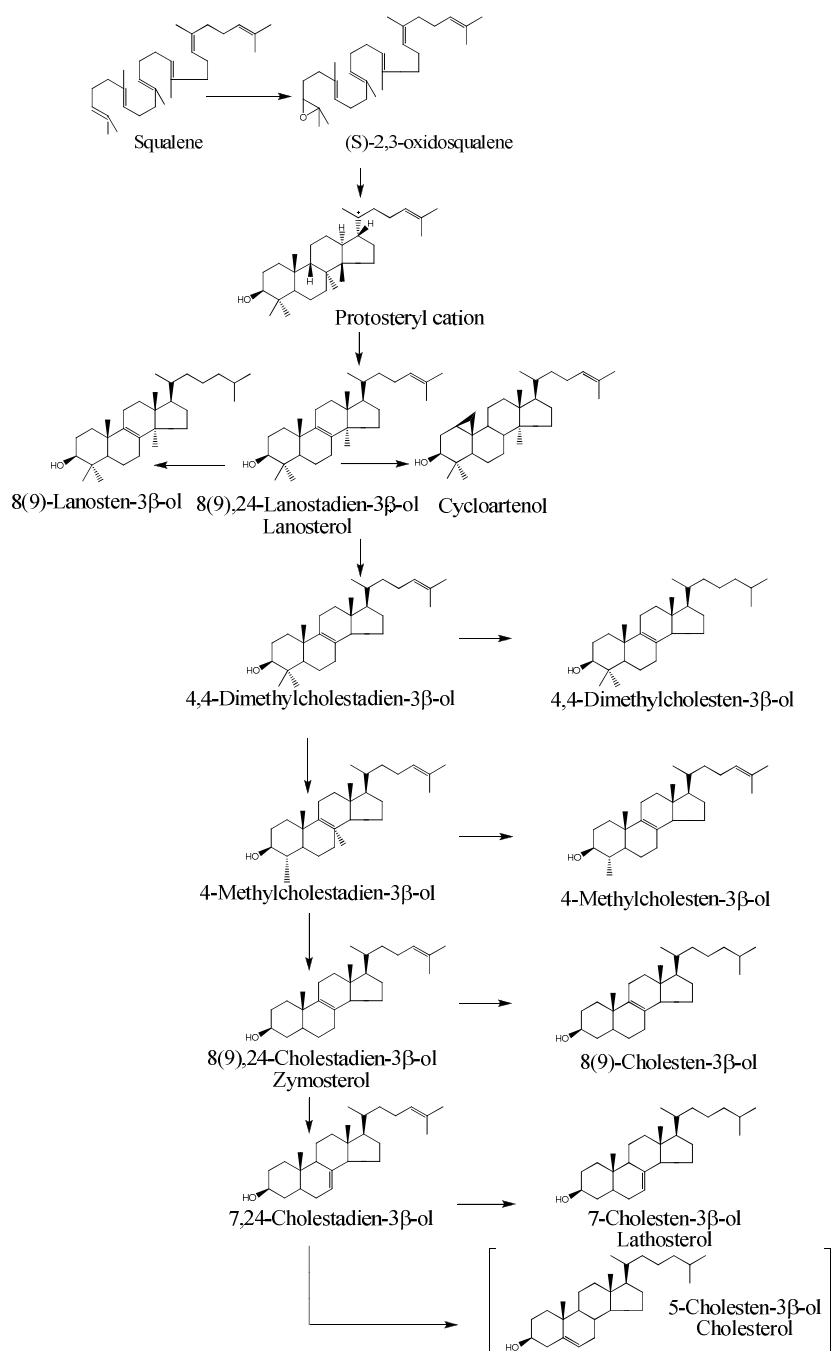


Figure III. 17: Steroid biosynthesis in myxobacteria. All steroids except for cholesterol (shown in brackets) have been isolated from different myxobacterial strains.

5. 4. Biotechnological application of CYP260A1

A variety of steroids are widely used as anti-inflammatory, diuretic, anabolic, contraceptive, antiandrogenic, progestational, and anticancer agents. So nowadays, the multi-step combinations of chemical and biochemical processes are used in the industrial production of

steroids (Farnandes *et al.*, 2003). The immediate shortfall of chemical synthesis are regarding the hazard, cost and time. Microorganisms or their isolated enzymes became a cheap and efficient alternative. The importance of microbial biotechnology in the production of steroid drugs and hormones was realized for the first time in 1952 when Murray and Peterson of Upjohn Company patented the process of 11 α -hydroxylation of progesterone by a *Rhizopus* sp (Murray and Peterson, 1952). The 11 α -, 11 β -, and 16 α -hydroxylations are now exclusively achieved in the steroid industry by microbial transformations. The other hydroxylations that seem to have potential for industrial exploitation are 9 α - and 14 α -hydroxylations. 14 α -hydroxyandrost-4-ene-3,17-dione is a useful intermediate for the preparation of hormones. 14 α -hydroxyandrost-4-ene-3,6,17-trione possess androgen activity and are useful inhibitors for breast cancer cells (Yoshioka *et al.*, 1994). The detailed study of the microbial biotransformations of steroids concerning the hydroxylation and C-C cleavage was reviewed before (Mahato and Garai, 1997; Donova *et al.*, 2005; Malaviya and Gomes, 2008).

Concerning this, as CYP260A1 of *So ce56* was hydroxylating the steroids, the potential biotechnological importance was more pronounced. Beside this, as the biosynthesis of steroid pathway in myxobacteria has already been shown, the steroid hydroxylation with CYP260A1 could have some important physiological significance too. Furthermore, CYP260A1 was also able to convert nootkatone into potentially hydroxylated and/or epoxidated products. This gives a new insight to study the conversion of several other sesquiterpenes, having some pharmaceutical and biotechnological applications. The importance of novel P450, CYP260A1, needs further exploration.

6. Outlook

The novel myxobacterial cytochrome P450 CYP260A1 from *Sorangium cellulosum* *So ce56* was found to be a potential steroid hydroxylase and sesquiterpene converter. These preliminary results give a new insight for the novel bacterial P450 for the conversion of several other steroids and terpenoids having some pharmaceutical and biotechnological applications. The physiological role of CYP260A1 during the life cycle of *So ce56* for steroid hydroxylation, the less predominated in bacterial system, could potentially be a new insight for further studies.

7. Literature review

1. Ahmed, F., Williams, R.A., and Smith, K.E. (1995). Microbial transformation of steroids-IX. Purification of progesterone hydroxylase cytochrome P-450 from *Phycomyces blakesleeanus*. J. Steroid Biochem. Mol. Biol. 52, 203-208.
2. Ahmed, F., Williams, R.A., and Smith, K.E. (1996). Cytochromes P-450 11 alpha-hydroxylase and C17-C20 lyase and a 1-ene-dehydrogenase transform steroids in *Nectria haematococca*. J. Steroid Biochem. Mol. Biol. 58, 337-349.
3. Al-Awadi, S., Afzal, M., and Oommen, S. (2001). Studies on *Bacillus stearothermophilus*. Part 1. Transformation of progesterone to a new metabolite 9,10-seco-4-pregnene-3,9,20-trione. J. Steroid Biochem. Molec. Biol. 78, 493-498.
4. Al-Awadi, S., Afzal, M., and Oommen, S. (2002). Studies on *Bacillus stearothermophilus*. Part II. Transformation of progesterone. J. Steroid Biochem. Molec. Biol. 82, 251-256.
5. Allen W.S. and Feldman L.I. (1961). Method of preparing 15-hydroxy pregnenes. US Patent 3, 010, 877.
6. Berg, A., Gustafsson, J.A., and Ingelman-Sundberg, M. (1976). Characterization of a cytochrome P-450-dependent steroid hydroxylase system present in *Bacillus megaterium*. J. Biol. Chem. 251, 2831-2838.
7. Bernhardt, R. (2006). Cytochromes P450 as versatile biocatalysts. J. Biotechnol. 124, 128-145.
8. Bernstein, S., Heller, M., Feldman, L.I., Allen, W.S., Blank, R.H. and Linden, C.E. (1960). C-15 substituted steroids. 15-alpha and 15- beta-hydroxy-reichstein's substance S and transformation products. J. Am. Chem. Soc. 82, 3685-3689.
9. Berrie, J.R., Williams, R.A.D., and Smith, K.E. (1999). Microbial transformations of steroids-XI. Progesterone transformation by *Streptomyces roseochromogenes*-purification and characterisation of the 16[alpha]-hydroxylase system. J. Steroid Biochem. Mol. Biol. 71, 153-165.
10. Berrie, J.R., Williams, R.A.D., and Smith, K.E. (2001). Microbial transformations of steroids-XII. Progesterone hydroxylation profiles are modulated by post-translational modification of an electron transfer protein in *Streptomyces roseochromogenes*. J. Steroid Biochem. Mol. Biol. 77, 87-96.
11. Bird, C., Lynch, J., Pirt, F., Reid, W., and Brooks, C. (1971). Steroids and squalene in *Methylococcus capsulatus* grown on methane. Nature. 230.

12. Bode, B.H., Zeggel, B., Silakowski, S., Wenzel, C.S., Reichenbach, H., and Müller, R. (2003). Steroid biosynthesis in prokaryotes: identification of myxobacterial steroid and cloning from the myxobacterium *Stigmatella aurantiaca*. *Mol. Microbiol.* *47*, 471-481.
13. Chen, J.J., Fei, D.Q., Chen, S.G., and Gao, K. (2008). Antimicrobial triterpenoids from *Vladimiria mulinensis*. *J. Nat. Prod.* *71*, 547-50.
14. Clemons, K.V., Stover, E.P., Schar, G., Stathis, P.A., Chan, K., Tokes, L., Stevens, D.A., and Feldman, D. (1989). Steroid metabolism as a mechanism of escape from progesterone-mediated growth inhibition in *Trichophyton mentagrophytes*. *J. Biol. Chem.* *264*, 11186-11192.
15. Cohen, L.H., Remley, M.J., Raunig, D., Vaz, A.D.N, (2003). In vitro drug interactions of cytochrome P450: an evaluation of fluorogenic to conventional substrates. *Drug Metab. Dispos.* *31*, 1005-1015.
16. Crespi, C.L., Miller, V.P., and Penman, B.W. (1997). Microtiter plate assays for inhibition of human, drug-metabolizing cytochromes P450. *Anal. Biochem.* *248*, 188-190.
17. De Rosa, M., Gambacorta, A., Sodano, G., and Trabucco, A. (1981). Transformation of progesterone by *Caldariella acidophil*, an extreme thermophilic bacterium. *CMLS*, *37*, 6, 541-542.
18. Dhar, A., and Samanta, T.B. (1993). Novel oxidative cleavage of C17-C20 bond in pregnane by a *Pseudomonas sp.* *J. Steroid Biochem. Mol. Biol.* *44*, 101-104.
19. Di, L., Kerns., E.H., Li, S.Q., and Carter, G.T. (2007). Comparison of cytochrome P450 inhibition assays for drug discovery using human liver microsomes with LC-MS, rhCYP450 isozymes with fluorescence, and double cocktail with LC-MS. *Int. J. Pharm.*, *335*, 1-11.
20. Dlugonski, J., Bartnicka, K., Zemelko, I., Chojecka, V., and Sedlaczek L. (1991). Determination of cytochrome P450 in *Cunninghamella elegans* intact protoplast and cell free preparations capable of steroid hydroxylation. *J. Basic. Microbiol.* *31*, 347-356.
21. Donova, M.V., Egorova, O.V., and Nikolayeva, V.M. (2005). Steroid 17[beta]-reduction by microorganisms-a review. *Process Biochem.* *40*, 2253-2262.
22. Faramarzi, M.A., Tabatabaei, Y.M., Amini, M., Zarrini, G., and Shafiee, A. (2003). Microbial hydroxylation of progesterone with *Acremonium strictum*. *FEMS Microbiol. Lett.* *222*, 183-186.

23. Fernandes, P., Cruz, A., Angelova, B., Pinheiro, H.M., and Cabral, J.M.S. (2003). Microbial conversion of steroid compounds: recent developments. *Enzyme Microb. Technol.* *32*, 688-705.
24. Fraga, B.M. (2004). Natural sesquiterpenoids. *Nat. Prod. Rep.* *21*, 669-693.
25. Furuya, T., Nishi, T., Shibata, D., Suzuki, H., Ohta, D., and Kino, K., (2008). Characterization of orphan monooxygenases by rapid substrate screening using FT-ICR mass spectrometry. *Chem. Biol.* *15*, 563-72.
26. Garai, S., and Mahato, S.B. (1992). Novel reduction and hydroxylation products formed by *Aspergillus fumigatus* from Reichstein's substance S. *Steroid.* *62*, 253-257.
27. Ghanem, K.M. and Yusef H.H. (1992b). Some nutritional requirements of a mixed culture transforming Reichstein's compound S into prednisolone. *Can. J. Microbiol.* *38*, 753-757.
28. Ghanem, K.M. and Yusef, H.H. (1992a). Transformation of Reichstein's compound S into prednisolone by a mixed culture under some physiological conditions. *Rev. Latinoam. Microbiol.* *34*, 107-114.
29. Gigon, P.L., Gram, T.E., and Gillette, J.R. (1969) Studies on the rate of reduction of hepatic microsomal cytochrome P-450 by reduced nicotinamide adenine dinucleotide phosphate: effect of drug substrates, *Mol. Pharmacol.* *5*, 109-122.
30. Holland, H.L., Nguyen, D.H., and Pearson, N.M. (1995). Biotransformation of corticosteroids by *Penicillium decumbens* ATCC 10436. *Steroids.* *60*, 646-649.
31. Holland, H.L., Poddar, S., and Tripet, B. (1992). Effect of cell immobilization and organic solvents on sulfoxidation and steroid hydroxylation by *Mortierella isabellina*. *J. Ind. Microbiol.* *10*, 195-197.
32. Horiuchi, K., Shiota, S., Hatano, T., Yoshida, T., Kuroda, T., and Tsuchiya, T. (2007). Antimicrobial activity of oleanolic acid from *Salvia officinalis* and related compounds on vancomycin-resistant enterococci (VRE). *Biol. Pharm. Bull.* *30*, 1147-9.
33. Jenkins, R.L., Wilson, E.M., Angus, R.A., Howell, W.M., Kirk, M., Moore, R., Nance, M., and Brown, A. (2004). Production of androgens by microbial transformation of progesterone in vitro: a model for androgen production in rivers. *Environ. Health Perspect.* *112*, 1508-1511.
34. Kang, M.J., Lisurek, M., Bernhardt, R. and Hartmann, R.W. (2004). Use of high-performance liquid chromatography/electrospray ionization collision-induced dissociation mass spectrometry for structural identification of monohydroxylated progesterones. *Rapid Commun. Mass Spectrom.* *18*, 2795-2800.

35. Kannenberg, E.L., and Poralla, K. (1999). Hopanoid biosynthesis and function in bacteria. *Naturwissenschaften*, 86, 168–176.
36. Kim, M.H., and Kim, M.N. (1991). Transformation pathway of the progesterone by *Rhizopus nigricans*. *Misaengmul Hakhoechi*, 29, 111-116.
37. Kohl, W., Gloe, A., and Reichenbach, H. (1983). Steroids from the myxobacterium *Nannocystis exedens*. *J. Gen. Microbiol.* 129, 1629–1635.
38. Kolek, T., and Swizdor, A. (1998). Biotransformation XLV. Transformations of 4-ene-3-oxo steroids in *Fusarium culmorum* culture. *J. Steroid Biochem. Mol. Biol.* 67, 63-69.
39. Kominek, L.A., Wolf, J.H. and Steiert, P.S. (1993). 1,2-dehydrogenation of steroidal 21-esters with *Arthrobacter simplex* or *Bacterium cyclooxydans*. U.S. Pat. 5, 225, 335.
40. Kumaki, K., Sato, M., Kon, H., Nebert, D. W. (1978). Correlation of type I, type II, and reverse type I difference spectra with absolute changes in spin state of hepatic microsomal cytochrome P-450 iron from five mammalian species. *J. Biol. Chem.* 253, 1048-1058.
41. Kunze, K.L, Nelson, W.L, Kharasch, E.D, Thummel, K.E., and Isoherranen, N. (2006) Stereochemical aspects of itraconazole metabolism in vitro and in vivo. *Drug Metab. Dispos.*, 34, 583–590.
42. Lisurek, M., Kang, M.J., Hartmann, R.W., and Bernhardt, R. (2004). Identification of monohydroxy progesterones produced by CYP106A2 using comparative HPLC and electrospray ionisation collision-induced dissociation mass spectrometry. *Biochem. Biophys. Res. Commun.* 319, 677-682.
43. Madyastha, K.M. (1996). Novel microbial transformations of steroids. *Adv. Exp. Med. Biol.* 405, 259-270.
44. Madyastha, K.M., and Joseph, T. (1993). Studies on the 14 alpha-hydroxylation of progesterone in *Mucor piriformis*. *J. Steroid Biochem. Mol. Biol.* 45, 563-569.
45. Mahato, S.B. and Garai, S. (1997). Advances in microbial steroid biotransformation. *Steroid*, 62, 332-345.
46. Malaviya, A., and Gomes, J. (2008). Androstenedione production by biotransformation of phytosterols. *Bioresour. Technol.* 99, 6725-6737.
47. Michal, G. (1999). *Biochemical Pathways*. Heidelberg, Spektrum Akad.-Verlag.
48. Miyamoto, M., Matsumoto, J., Iwaya, T., and Itagaki, E. (1995). Bacterial steroid monooxygenase catalyzing the Baeyer-Villiger oxidation of C21-ketosteroids from

- Rhodococcus rhodochrous*: the isolation and characterization. *Biochim. Biophys. Acta.* 1251, 115-124.
49. Murray, H.C., and Peterson D.H. (1952). Oxygenation of steroids by Mucorales fungi. U.S. Patent 2602769 (Upjohn Co., Kalamazoo, Michigan, USA).
 50. Ourisson, G., and Nakatani, Y. (1994). The terpenoid theory of the origin of cellular life: the evolution of terpenoids to cholesterol. *Chem. Biol.* 1, 11–23.
 51. Ourisson, G., Rohmer, M., and Poralla, K. (1987). Prokaryotic hopanoids and other polyterpenoid sterol surrogates. *Annu. Rev. Microbiol.* 41, 301–333.
 52. Oviden, S.P., Yu, J., Bernays, J., Wan, S.S., Christophidis, L.J., Sberna, G., Tait, R.M., Wildman, H.G., Lebellier, D., Platel, D., May, T.W., and Meurer-Grimes, B.M. (2005). Cerylimycins [corrected] A and B: antibacterial triterpenes from the new species *Tricholoma* sp. AU1. *J. Nat. Prod.* 68, 409-12.
 53. Pearson, J., Hill, J., Swank, J., Isoherranen, I., Kunze, K., and Atkins, W. (2006) Surface plasmon resonance analysis of antifungal azoles binding to CYP3A4 with kinetic resolution of multiple binding orientations. *Biochemistry*, 45, 6341–6353.
 54. Rabe, S.K., Spengler M., Erkelenz, M., Müller, J., Gandubert, V.J., Hayen, H., and Niemeyer, C.M., (2009). Screening for cytochrome p450 reactivity by harnessing catalase as reporter enzyme. *Chembiochem.* 10, 751-7.
 55. Roberts, A.G., Campbell, A.P., and Atkins, W.M. (2005). The thermodynamic landscape of testosterone binding to cytochrome P450 3A4: Ligand binding and spin state equilibria. *Biochemistry.* 44, 1353-1366.
 56. Samanta, T.B., and Ghosh, D.K. (1987). Characterization of progesterone 11 alpha-hydroxylase of *Aspergillus ochraceus*: a cytochrome P-450 linked monooxygenase. *J. Steroid. Biochem.* 28, 327-332.
 57. Schenkman, J.B., Remmer, H., and Estabrook, R.W. (1967). Spectral studies of drug interaction with hepatic microsomal cytochrome. *Mol. Pharmacol.* 3, 113-123.
 58. Schenkman, J.B., Sligar, S.G., and Cinti, D.L. (1981). Substrate interaction with cytochrome P-450. *Pharmacol. Ther.* 12, 43-71.
 59. Sideso, O., Williams, R.A.D., Welch, S.G., and Smith, K.E. (1998). Progesterone 6-hydroxylation is catalysed by cytochrome P-450 in the moderate thermophile *Bacillus thermoglucosidasius* strain 12060. *J. Steroid Biochem. Mol. Biol.* 67, 163-169.
 60. Smith, C., Marks, D.A., and Lieberman, M. (2008). Marks' Basic Medical Biochemistry: A Clinical Approach. Lippincott Williams & Wilkins.

61. Smith, J.D., Martin, J.J.V., and Mohn, W.W. (2004). A cytochrome P450 involved in the metabolism of abietane diterpenoids by *Pseudomonas abietaniphila* BKME-9. *J. Bacteriol.* *186*, 3631-3639.
62. Smith, J.D., Patrauchan A.M., Florizone, C., Eltis, D.L. and Mohn, W.W. (2007). Distinct roles for two CYP226 family cytochrome P450 in abietane diterpenoid catabolism by *Burkholderia xenovorans* LB400. *J. Bacteriol.* *190*, 1575-1583.
63. Smith, K.E., Ahmed, F., Williams, R.A.D., and Kelly, S.L. (1994). Microbial transformations of steroids--VIII. Transformation of progesterone by whole cells and microsomes of *Aspergillus fumigatus*. *J. Steroid Biochem. Mol. Biol.* *49*, 93-100.
64. Smith, K.E., Latif, S.A., and Kirk, D.N. (1991). Microbial transformation of steroids-VII. Hydroxylation of progesterone by extracts of *Phycomyces blakesleeanus*. *J. Steroid Biochem. Mol. Biol.* *38*, 249-256.
65. Sogno, I., Vannini, N., Lorusso, G., Cammarota, R., Noonan, D.M., Generoso, L., Sporn, M.B., and Albini, A. (2009). Anti-angiogenic activity of a novel class of chemopreventive compounds: oleanic acid terpenoids. *Recent Results Cancer Res.* *181*, 209-12.
66. Sukumaran, M.S., Potsaid, B., Lee, Y.L., Clark, S.D., and Dordick, S.J. (2009). Development of fluorescent-based, ultra high-throughput screening platform for nanoliter-scale cytochrome P450 microarray. *J. Biomol. Screen.* *14*, 668
67. Taylor, M., Lamb, D.C., Cannell, R., Dawson, M. and Kelly, S.L. (1999). Cytochrome P450105D1 (CYP105D1) from *Streptomyces griseus*: Heterologous Expression, Activity, and Activation. Effects of Multiple Xenobiotics. *Biochem. Biophys. Res. Commun.* *263*, 338-342
68. Trubetskoy, O.V., Gibson, J.R., and Marks, B.D. (2005). Highly miniaturized formats for in vitro drug metabolism assays using Vivid fluorescent substrates and recombinant human cytochrome P450 enzymes. *J. Biomol. Screen.* , *10*, 56-66.
69. Tsuda, K., Iizuka, H., Ohki, E., Sato, Y., Naito, A., and Hattori M. (1959). Hydroxylierung von Progesteron und Reichsteins Substanz S durch *Achromobacter kaishiwasakiensis* Nov. Sp. *J. Gen. Appl. Microbiol.* *5*, 7-12.
70. Virus, C., and Bernhardt, R. (2008). Molecular evolution of a steroid hydroxylating cytochrome P450 using a versatile steroid detection system for screening. *Lipids.* *43*, 1133-1141.

71. Vita, M., Smith, K., Rozman, D., and Komel, R. (1994). Progesterone metabolism by the filamentous fungus *Cochliobolus lunatus*. *J. Steroid Biochem. Mol. Biol.* *49*, 87-92.
72. Wu, Z.J., Ouyang, M.A., Wang, C.Z., and Zhang, Z.K. (2007). Six new triterpenoid saponins from the leaves of *Ilex oblonga* and their inhibitory activities against TMV replication. *Chem. Pharm. Bull.* *55*, 422-7.
73. Yadava, R.N., and Jharbade, J. (2008). New antibacterial triterpenoid saponin from *Lactuca scariola*. *Fitoterapia.* *79*, 245-9.
74. Yoshihama, M. (1993). Microbial hydroxylation of steroid hormones and their pharmaceutical applications. *Yukijirushi Nyugyo Kenkyusho Hokoku*, *99*, 1-70.
75. Yoshioka, H., Asada, H., and Fujita, S. (1994). Process for production of 6[3, 14 α -dihydroxy-4-androstene-3,17-dione. European Patent.
76. Zeng, B., and Zhang, B. (1992). The relationship between the growing characterization and steroid 11-alpha-hydroxylation of *Beauveria bassiana* AS69 during fermentation. *Weishengwuzue Zazhi*, *12*, 43-45.
77. Zhang, M., Wang, W.L., Fang, Y.C., Zhu, T.J., Gu, Q.Q., and Zhu, W.M. (2008). Cytotoxic alkaloids and antibiotic nordammarane triterpenoids from the marine-derived fungus *Aspergillus sydowi*. *J. Nat. Prod.* *71*, 985-9.
78. Zuidweg, M.H.J., van der Waard W.F., and De Flines, J (1962). Formation of hydrocortisone by hydroxylation of Reichstein's compound S with an enzyme preparation from *Curvularia lunata*. *Biochim. Biophys. Acta.* *58*, 131-133.

APPENDIX

APPENDIX I

A. INSTRUMENTS

Instrument	Characteristics	Company
FPLC	Äkta prime	
Electrophoresis unit		Bio-Rad
Autoclave	LVSA 50/70	Zirbus IMM 20
Balance	KB 600-2 SBA 62 SBC 31	Kern Scaltec Scaltec
Centrifuge	Pico Fuge HF-120 3K30 (rotors 19776, 12153) 2K15 (rotor 12148) Avanti J-20 Himac CP75 β (Ultra-centrifuse)	Stratgene Sigma Sigma Beckman Coulter Hitachi
CD-spectrometer		Jasco 715
Electroporator	EasyjecT Prima	Equibio
FPLC	Äkta prime	
Freezer	Revco Ultima II Forma Scientific	Thermo Scientific Thermo Scientific
Gel dryer	Model 583	Bio-Rad
HPLC system	PU-2080Plus Intelligent HPLC-Pump, AS-2050 Intelligent Sampler, DG-2080-53 3-Line Degasser, LG-2080-02 Ternary-Gradient Unit, LC-Net II/ADC, UV-2075Plus Intelligent UV/VIS-Detector	Jasco
Incubator		Heraeus
Incubation shaker	Multitron II Innova "4230"	Infors HT New Brunswick Scientific
Microwave oven		Siemens
Peltier thermal cycler	DNA Engine	Biorad
pH meter	766 Calimatic	Knick
Power supply unit	PowerPac 300 ECPS 3000/150	Bio-Rad Pharmacia
SDS gel electrophoresis unit	Mighty Small II	Hoefer
SpeedVac		UniEquip Univapo
Thermocycler		MJ Research
Thermomixer	5436	Eppendorf
Ultrasonic desintegrator	USD-30	Emich Ultraschall
UV/VIS Spectrophotometer	UV-2100 UV-2101 PC	Shimadzu
Vortex VF2	VF2	IKA RLabortechnik

B. CHEMICALS

All commercially obtained materials used in this work were of the finest quality available purchased from the companies listed below:

Amersham Pharmacia Biotech	Kodak
Bio-Rad Laboratories	Merck
BioTez GmbH	MWG Biotech GmbH
Biozym GmbH	New England Biolabs
Clontech Laboratories GmbH	Promega
Difco Laboratories	Qiagen
Dulbeco	Roche Molecular Biochemicals
Fluka AG	Serva GmbH & Co KG
Gibco/BRL GmbH	Sigma-Aldrich
Invitrogen BV	Stratagene

C. ENZYMES

Bioline

Stratagene Ltd, Cambridge, UK

New England Biolabs

Boehringer

CLONTECH Laboratories

MBI Fermentas

Promega Corporation

C. KITS

Purpose	Characteristics	Company
Plasmid DNA purification	NucleoBond AX 100	Macherey-Nagel
PCR clean-up Gel extraction	NucleoSpin Extract II	Macherey-Nagel
Protein quantification	BC Assay	Uptima

D. COLUMN MATERIALS FOR PROTIEN PURIFICATION

Affinity chromatography: Talon® Resin, Metal affinity resin, Clontech

Size exclusion chromatography: Superdex-75, GE Healthcare.

E. CONSUMABLES

Amersham Pharmacia Biotech

Fisher Scientific

Falcon

Machery-Nagel GmbH & Co KG

Millipore

Carl Roth GmbH

Qiagen GmbH

APPENDIX II

II. 1. BUFFERS AND COMMONLY USED SOLUTIONS

a. Buffer for chemical competent cell preparation

TFP buffer

A		B	
1M CH ₃ COOK	3 ml	1M MnCl ₂	5 ml
1M KCl	10 ml	autoclave separately	
1M CaCl ₂	1 ml		
Glycerol	12 ml		
Fill water up to 80 ml and set pH to 6.1			
Mix A and B and adjust the final volume of 100 ml. filter sterilize and stored at 4°C.			

II. 2. Common Buffers

a. Sodium phosphate buffer (50 mM, pH 7.4)

1M Na ₂ HPO ₄	46.6 ml
1M NaH ₂ PO ₄	3.4 ml add d/w up to 1000 ml

b. Potassium phosphate buffer (40 mM, pH 7.3)

1M K ₂ HPO ₄	40 ml
1M KH ₂ PO ₄	10 ml add d/w up to 1000 ml

c. Potassium phosphate buffer (10 mM, pH 7.4)

1M K ₂ HPO ₄	8.4 ml
1M KH ₂ PO ₄	1.6 ml add d/w up to 1000 ml

II.3. BUFFERS AND SOLUTIONS FOR AGAROSE GEL ELECTROPHORESIS

10 x TAE Buffer

Tris base	242.28 g
EDTA-Na	18.61 g
Acetic acid	60.05 ml
pH 8.5 ± 0.2 Add d/w up to 1000 ml	

10 x Loading buffer

Bromophenol blue	2.5 mg
Xylencyanole	2.5 mg
Glycerol	300 µl
add d/w up to 1 ml	

50 x TARE Buffer [Tris/Acetate/reduced EDTA, pH 8.0]

2M	Tris	24.1 g
Tris/Acetate,	Acetic acid	5.7 ml
pH 8.0	5 mM EDTA	1 ml
		Add d/w up to 100 ml

APPENDIX III

1. Buffers and solutions for polyacrylamide gel electrophoresis

<i>separating gel</i>	Stock solutions	Final acrylamide concentration		
		10%	12%	15%
	4 x LT	3.75 ml	3.75 ml	3.75 ml
	10% APS	75 μ l	75 μ l	75 μ l
	dest.H ₂ O	ad 15 ml	ad 15 ml	ad 15 ml
	30% AA/Bis	5.0 ml	6.0 ml	7.5 ml
	TEMED	7.5 μ l	7.5 μ l	10 μ l
<i>stacking gel</i>	Stock solutions	5% acrylamide concentration		
	4 x UT	2.5 ml		
	10% APS	50 μ l		
	Dest.H ₂ O	ad 10 ml		
	30% AA/Bis	1.6 ml		
	TEMED	5 μ l		
4 x Buffer for separating gel (LT)	1.5 M Tris/Cl, pH=8,8			
	0.4% SDS			
4 x Buffer for stacking gel (UT)	0.5 M Tris/Cl, pH=6,8			
	0.4% SDS			
2 x SDS loading buffer	125 mM Tris/Cl, pH=6.8			
	20% Glycerol			
	4% SDS			
β -mercaptoethanol	10%			
Bromphenol blue	0.004%			
Staining solution	0.1% Coomassie Brilliant Blue G-250			
	40% Methanol			
	10% Acetic acid			
Destaining solution	25% Methanol			
	10% Acetic acid			

2. Buffers and solutions for protein purification

A. PMSF stock - 100 mM PMSF in isopropanol

B. Cell lysis buffer (Buffer A)

50 mM Sodium phosphate buffer, pH 8.0
300 mM NaCl
20% Glycerin
10 mM imidazole

C. Buffer for affinity chromatography

i. Equilibrium buffer (Buffer B)

50 mM Sodium phosphate
buffer, pH 8.0
300 mM NaCl
20% Glycerin
20 mM imidazole

ii. Elution buffer (Buffer C)

50 mM Sodium phosphate buffer, pH
8.0
300 mM NaCl
20% Glycerin
150 mM imidazole

D. Buffer for size exclusion chromatography (Buffer D)

Potassium phosphate buffer (10 mM, pH 7.4) containing 20% glycerol.

E. Buffer E

10mM potassium phosphate buffer, pH 7.4 containing 20% glycerol and 3 mM DTT.

APPENDIX IV

1. Culture media

i. TB media

A	Bacto tryptone	12 g	B	KH ₂ PO ₄	2.31 g
	Bacto yeast extract	24 g		K ₂ HPO ₄	12.54 g
	Glycerin	4 ml		Add d/w up to 100 ml and autoclave it	
	Add d/w up to 900 ml and autoclave it				

Mix solution A and B for use

ii. SOC Medium

Trypton	2%
Yeast extract	0.5%
NaCl	10 mM
KCl	2.5 mM
MgCl ₂	10 mM

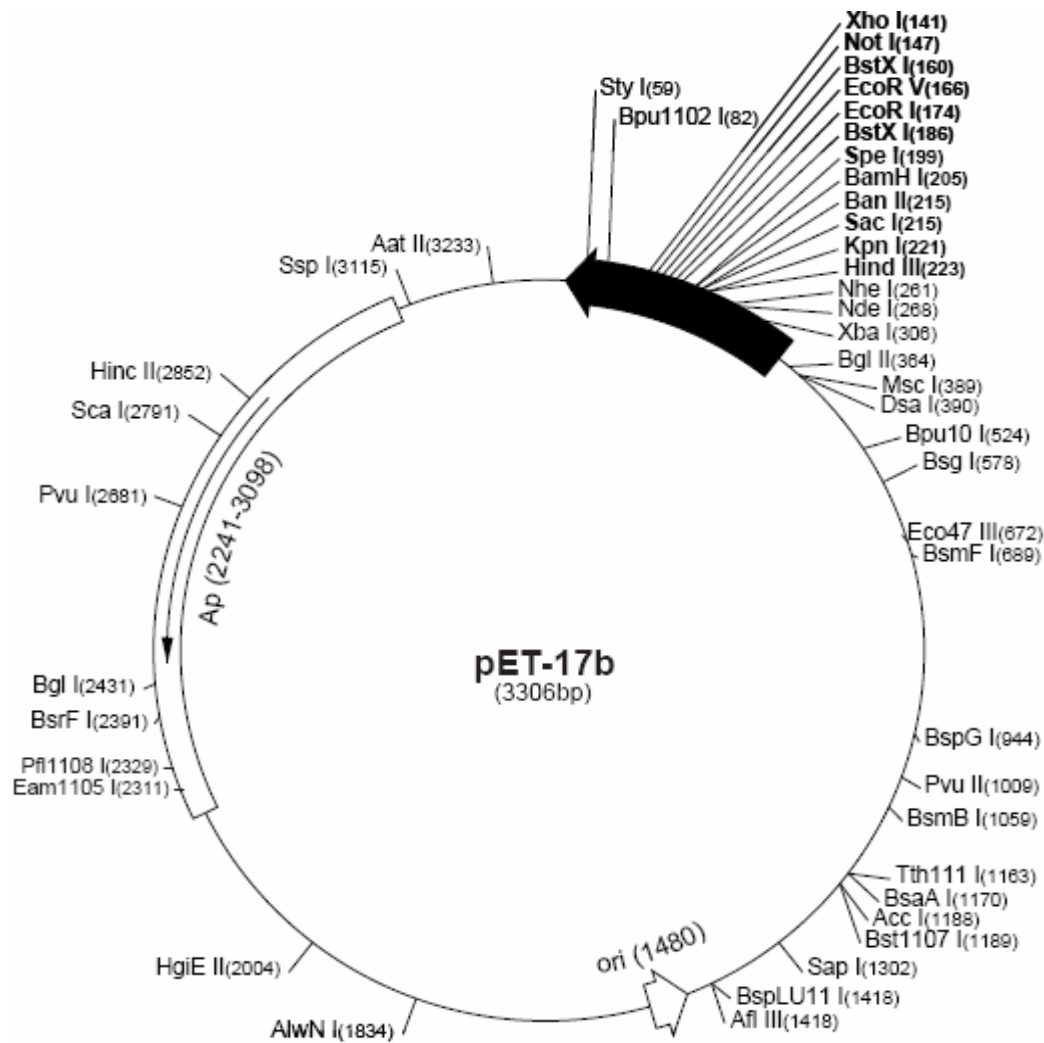
Adjust pH 7.0, autoclave, add 20mM glucose and filter sterilize by 0.22-micron filter.

2. ANTIBIOTICS AND OTHERS

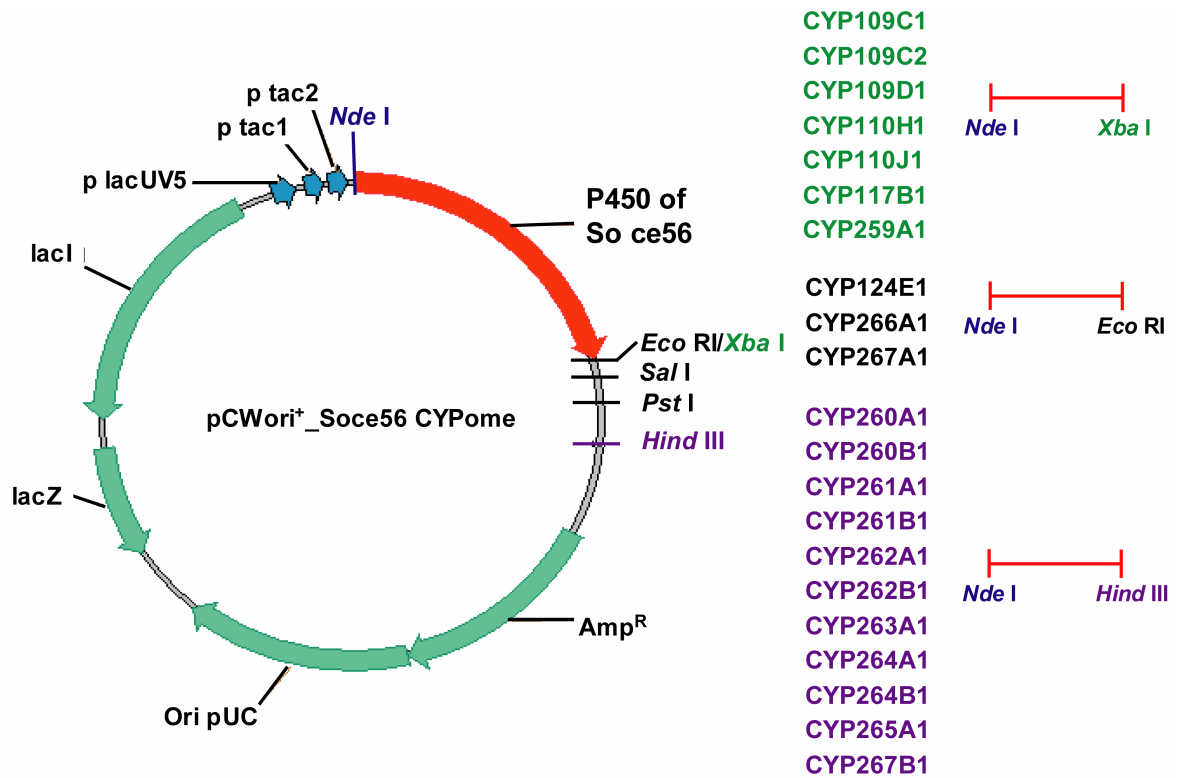
Antibiotics	Stock solution	Solvent	Working conc.
Ampicillin	100 mg/ml	water	100 µg/ml
Kanamycin	50 mg/ml	water	50 µg/ml
Chloramphenicol	50 mg/ml	ethanol	50 µg/ml

APPENDIX V

1. Vector map of pET17b



2. Vector map of pCWori⁺. The vector map with three promoters (lacUV5, tac1 and tac2) and multiple cloning site (*EcoRI*, *XbaI*, *SalI*, *PstI* and *HindIII*) are shown. The P450 inserts of *So ce56* having same restriction sites are shown in same color. Each P450 shown in green, black and purple has *XbaI*, *EcoRI* and *HindIII* sites, respectively, at C-terminal after hexahistidine tag..

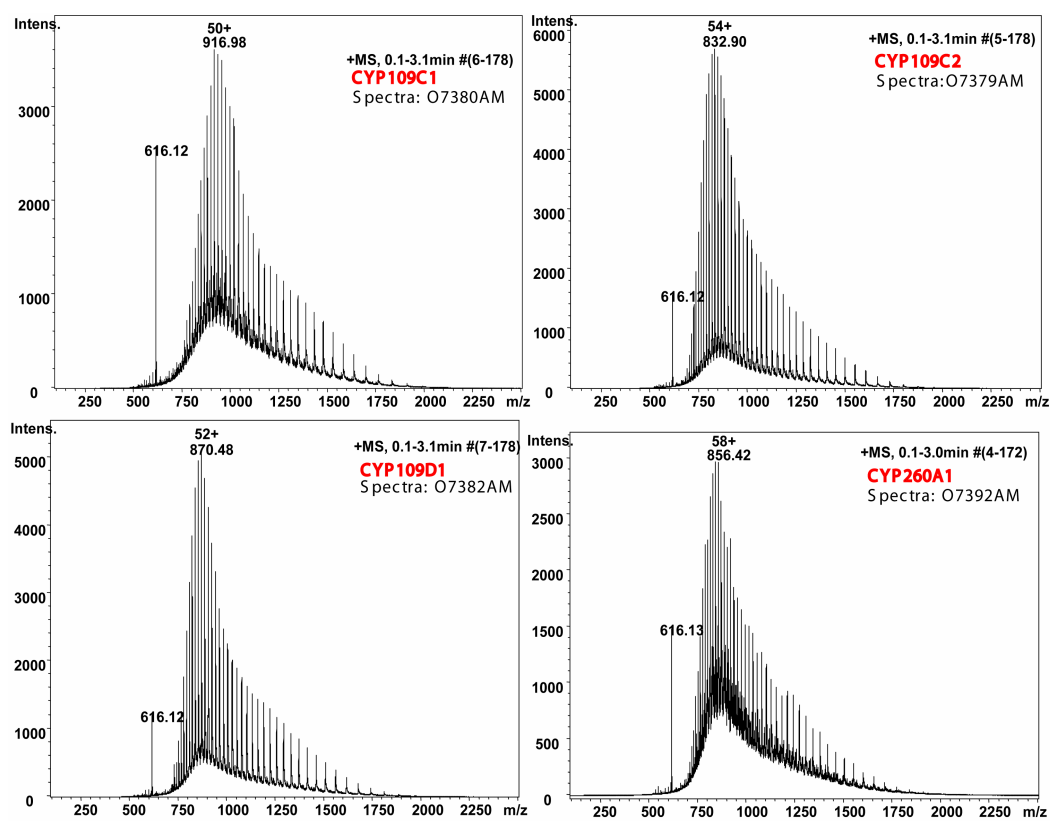


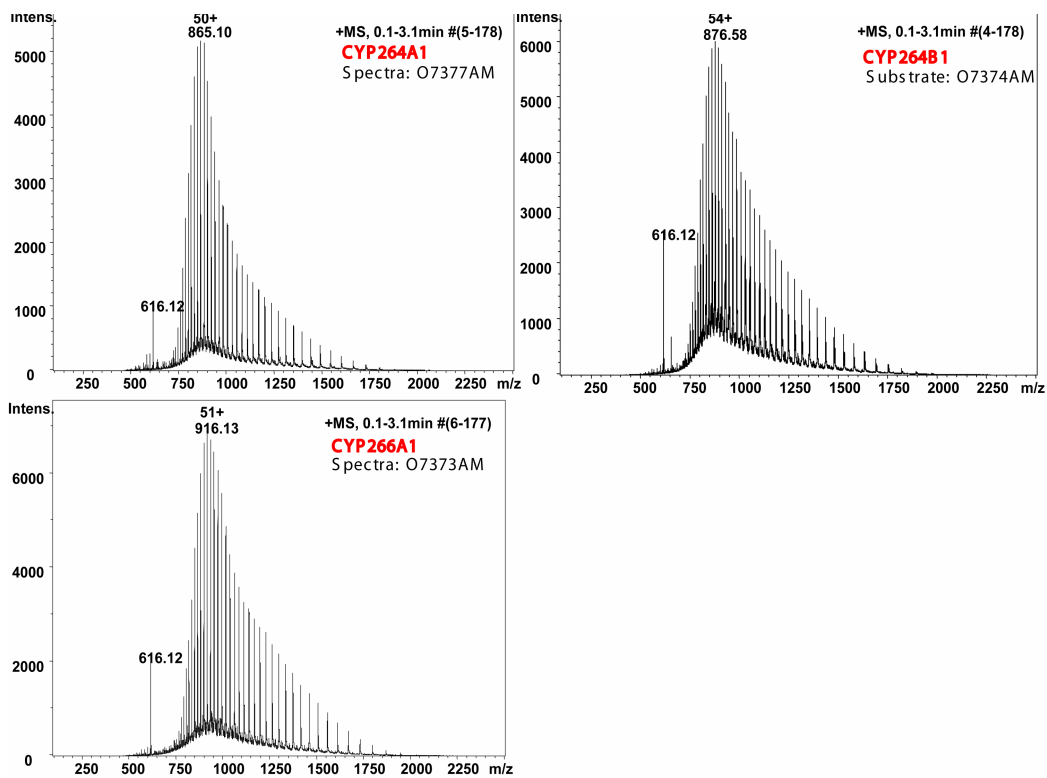
APPENDIX VI

Mass spectra of A. apo-mass, and B. holo-mass of CYP109C1, CYP109C2, CYP109D1, CYP260A1, CYP264A1, CYP264B1 and CYP266A1 of *So ce56* are shown. The mass of the P450s was measured in collaboration with Aurélie Mème, Laboratoire de Dynamique et Structure Moléculaire par Spectrométrie de Masse, Institut de Chimie, Starsbourg, France. Mass spectrometric studies were performed in the positive ion mode using an ESI-TOF (Time-of-flight) mass spectrometer (microTOF, Bruker Daltonics, Bremen, Germany).

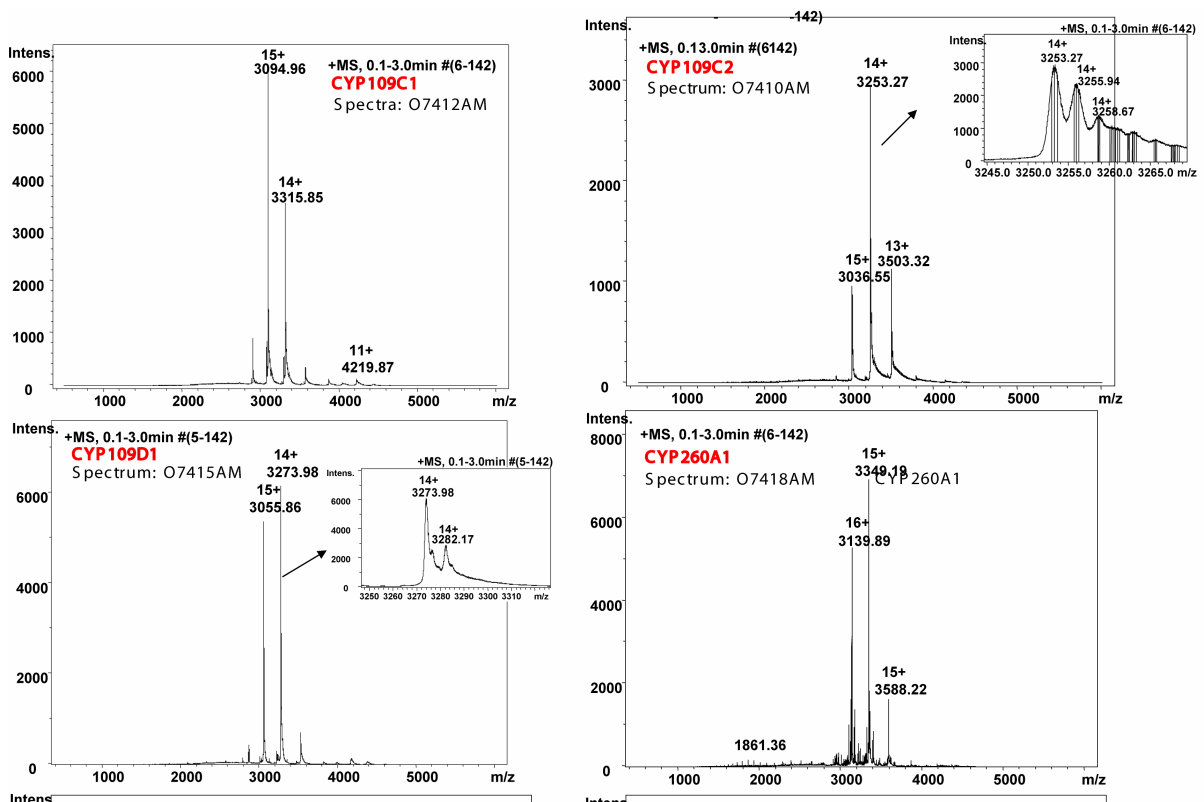
Each peak on the spectrum shows the corresponding P450 mass with different state of charge, for example, in the spectrum of CYP109C1 (O7380AM), the peak at 916.98 Da corresponding to the P450 with 50 charges (50 protons on the protein). As the abscissa of the spectrum is in m/z, the real mass of the protein is obtained by multiplying the mass 916.98 by 50. Software coupled with the system calculates the real mass for each peak according to its state of charge and calculates the average mass for all the peaks.

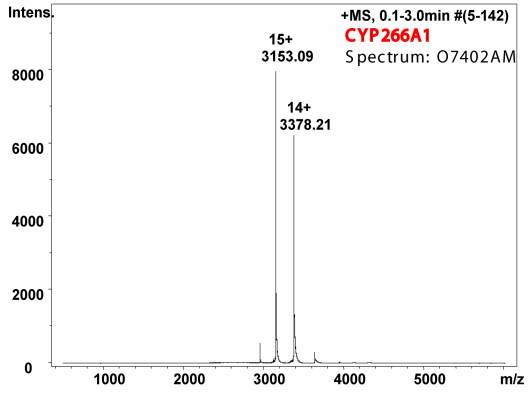
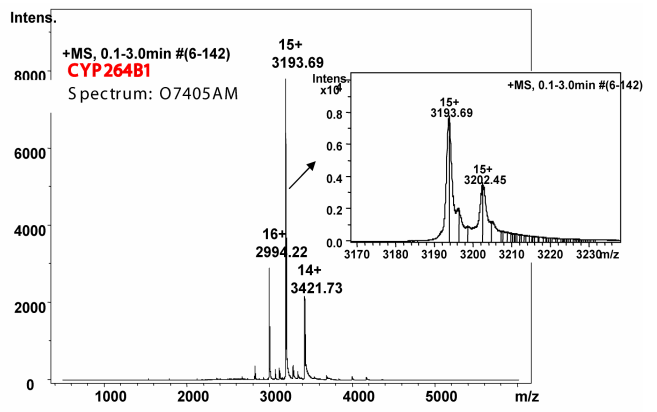
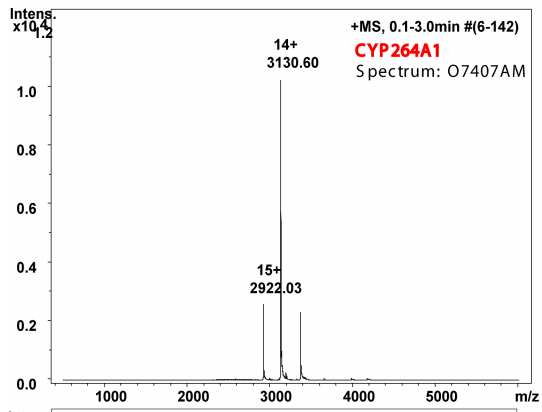
A. Mass spectra (apo- mass) of different P450s of *So ce56* (Table I. 14 in text)





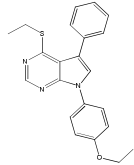
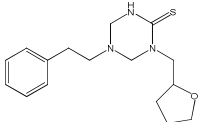
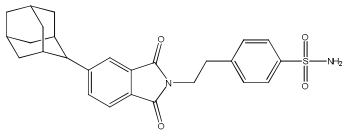
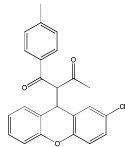
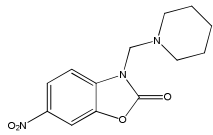
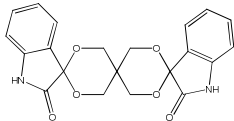
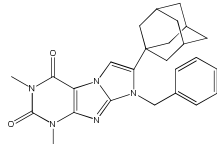
B. Mass spectra (holo- mass) of different P450s of So ce56 (Table I. 14 in text)

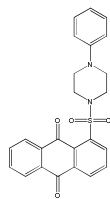




APPENDIX VII

A. SCREENING RESULTS FOR CYP109C1

CYP109C1 Type I hits	Compound code (Chem_Div nos)/ Remarks
	31649, C074-0114 (also for CYP260A1) Type I- Primary screening Type I- 1st validation Type II- Clear with concentration dependent validation
	32245, 5701-2503 Type I- Primary screening Type II- 1st validation Type I- Clear with concentration dependent validation
	33152, 7406-0023 (also for CYP260A1) Type I- Primary screening Not clear in 1st validation Type I- Clear with concentration dependent validation
	33160, 7746-0008 Type I- Primary screening Type I- Clear with concentration dependent validation
	33629, 8006-6674 Type II- Primary screening Not clear in 1st validation Type I- Clear with concentration dependent validation
	34062, 8012-8381 Type I- Primary screening Not clear in 1st validation Type I- Clear with concentration dependent validation
	34355, 8012-8982 Type I- Primary screening Not clear in 1st validation Type I- Clear with concentration dependent validation

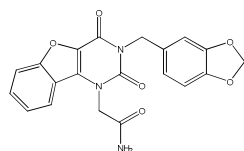


35049, 8016-2764

Type II- Primary screening

Not clear in 1st validation

Type I- Clear with concentration dependent validation

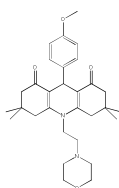


35443, E545-0590

Type I- Primary screening

Not clear in 1st validation

Type I- Clear with concentration dependent validation

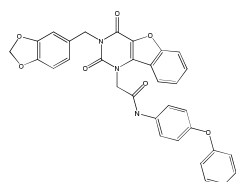


35444, C047-0116

Type II- Primary screening

Not clear in 1st validation

Type I- Clear with concentration dependent validation

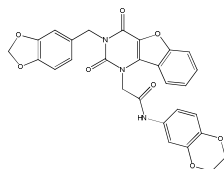


35447, E545-0637

Type I- Primary screening

Not clear in 1st validation

Type I- Clear with concentration dependent validation

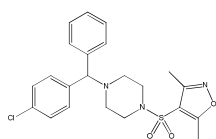


35449, E545-0638

Type I- Primary screening

Not clear in 1st validation

Type I- Clear with concentration dependent validation

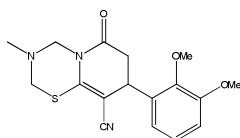


36880, C499-0591

Type I- Primary screening

Not clear in 1st validation

Type I- Clear with concentration dependent validation



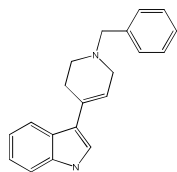
37517, C499-0591

(also for CYP109C2 and CYP266A1)

Type I- Primary screening

Not clear in 1st validation

Type I- Clear with concentration dependent validation

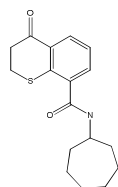


37518, C800-0101

Type II- Primary screening

Not clear in 1st validation

Type I- Clear with concentration dependent validation



38786, E213-1274

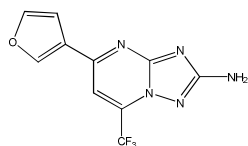
Type I- Primary screening

Not clear in 1st validation

Type I- Clear with concentration dependent validation

CYP109C1 Type II hits

Compound code/Remarks

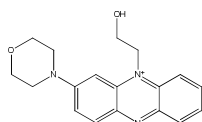


31305, 7395-3009

Type II- Primary screening

Type II- 1st validation

Not Clear with concentration dependent validation

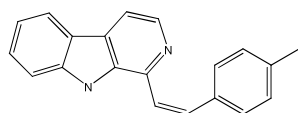


31632, 5451-0327

Type II- Primary screening

Type II- 1st validation

Not 100% Clear with concentration dependent validation

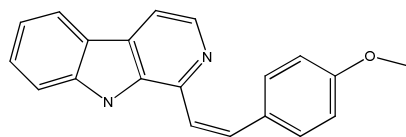


33182, 7462-0037

(also for CYP109C2, CYP260A1 and CYP266A1)

Type II- Primary screening

Type II- Clear with concentration dependent validation



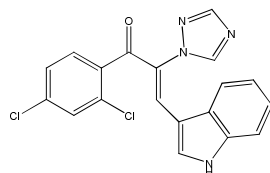
33202, 7462-0046

(also for CYP109C2, CYP260A1 and CYP266A1)

Type II- Primary screening

Not clear with 1st validation

Type II- Clear with concentration dependent validation

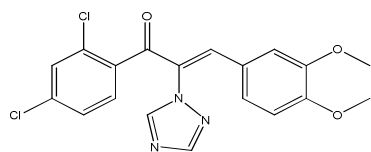


33463, D155-0026

(also for CYP109C2, CYP260A1 and CYP266A1)

Type II- Primary screening

Type II- Clear with concentration dependent validation

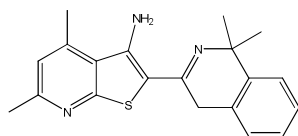


33469, D155-0050

(also for CYP109C2, CYP260A1 and CYP266A1)

Type II- Primary screening

Not clear with concentration dependent validation



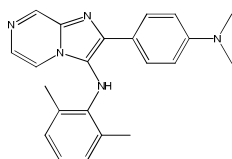
34218, 8013-4052

(also for CYP109C2, CYP260A1 and CYP266A1)

Type II- Primary screening

Not clear in 1st validation

Type II- Clear with concentration dependent validation



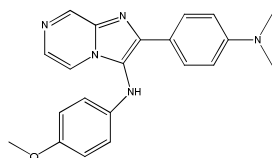
35863, C169-0140

(also for CYP260A1)

Type II- Primary screening

Not clear in 1st validation

Type II- Clear with concentration dependent validation



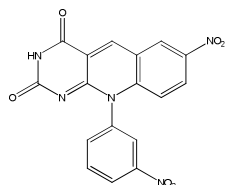
35933, C169-0144

(also for CYP109C2, CYP260A1 and CYP266A1)

Type II- Primary screening

Type II- Clear in 1st validation

Type II- Clear with concentration dependent validation



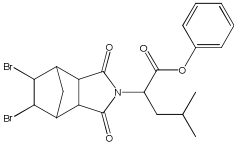
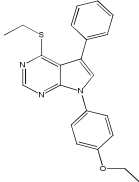
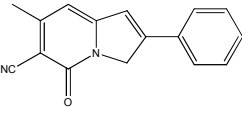
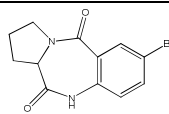
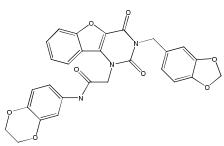
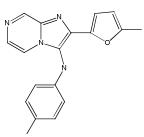
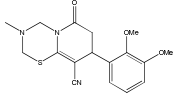
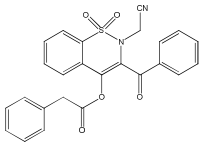
2371-0057

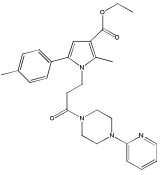
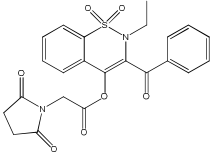
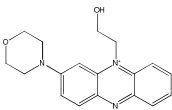
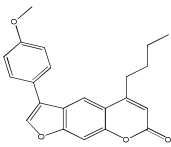
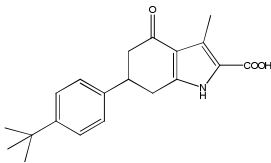
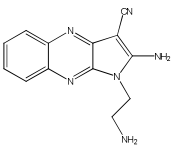
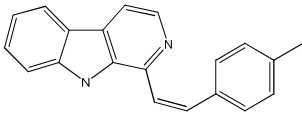
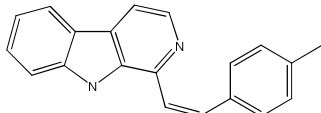
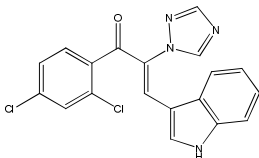
Type II- Primary screening

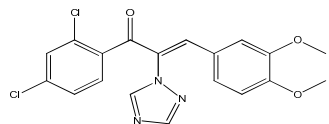
Type II- Clear in 1st validation

Type II- Clear with concentration dependent validation

B. SCREENING RESULTS FOR CYP109C2

CYP109C2 Type I hits	Compound code/Remarks
	31358, 5089-1245 (also for CYP266A1) Type I- Primary screening Not clear with concentration dependent validation
	31649, 5451-0327 (also for CYP266A1) Type I- Primary screening Type I- Clear with concentration dependent validation
	34168, 8013-0849 Type II- Primary screening Type II- in first validation Type I- Clear with concentration dependent validation
	35086, 8017-2885 Type I- Primary screening Type I- Clear with concentration dependent validation
	35449, E545-0638 Type I- Primary screening Not clear in first validation Type I- Clear with concentration dependent validation
	35862, C169-0253 Type I- Primary screening Type I- Clear with concentration dependent validation
	37517, C499-0591 (also for CYP266A1 and CYP260A1) Type I- Primary screening Type I- Clear with concentration dependent validation
CYP109 Type II hits	Compound code/Remarks
	31506, 5149-0173 Type II- Primary screening Not clear with concentration dependent validation

	<p>38219, C721-0718 Type II- Primary screening Type II- Clear with concentration dependent validation but not 100% sure</p>
	<p>31518, D097-0031 Type II- Primary screening Not clear with concentration dependent validation</p>
	<p>31632, 5451-0327 Type II- Primary screening Type II- Clear with concentration dependent validation</p>
	<p>31892, 5436-1027 Type II- Primary screening Type II- Clear with concentration dependent validation</p>
	<p>32628, C276-0392 Type II- Primary screening Not clear with concentration dependent validation</p>
	<p>32873, 6257-0356 Type II- Primary screening Type II- Clear with concentration dependent validation</p>
	<p>33182, 7462-0037 (also for CYP109C1 and CYP266A1) Type II- Primary screening Type II- Clear with concentration dependent validation</p>
	<p>33202, 7462-0046 (also for CYP109C1, CYP266A1 and CYP260A1) Type II- Primary screening Type II- Clear with concentration validation</p>
	<p>33463, D155-0026 (also for CYP109C1, CYP266A1 and CYP260A1) Type II- Primary screening Type II- Clear with concentration dependent validation</p>

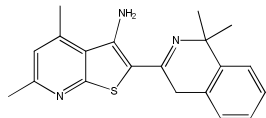


33469, D155-0050

(also for CYP109C1, CYP266A1 and CYP260A1)

Type II- Primary screening

Type II- Clear with concentration dep. Val.

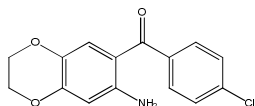


34218, 8013-4052

(also for CYP109C1, CYP266A1 and CYP260A1)

Type II- Primary screening

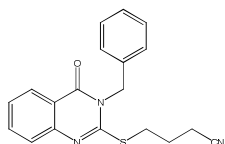
Type II- Clear with conc. Dep. Val.



34484, 8015-3948

Type II- Primary screening

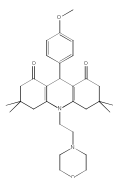
Type II- Clear with concentration dependent validation



35124, 8016-9750

Type II- Primary screening

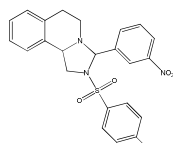
Type II- Clear with concentration dependent validation



35444, C047-0116

Type II- Primary screening

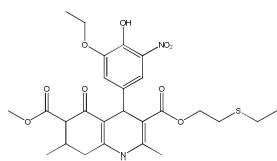
Type II- Clear with concentration dependent validation



35820, 8012-4720

Type II- Primary screening

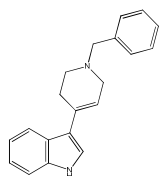
Type II- Clear with concentration dependent validation



35822, C119-0040

Type II- Primary screening

Type II- Clear with concentration dependent validation



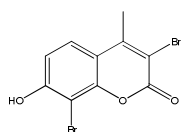
37518, C800-0101

(CYP260A1)

Type II- Primary screening

Type II- 1st validation

Type II- Clear with concentration dependent validation



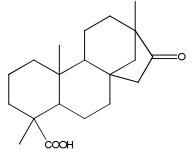
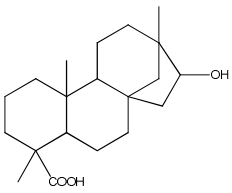
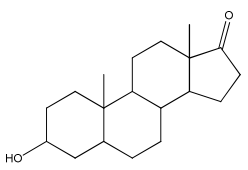
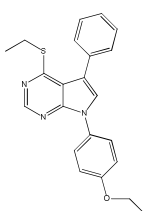
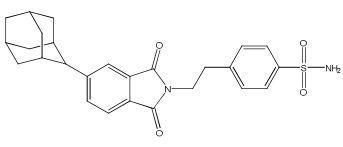
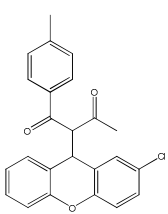
40120, K815-0059

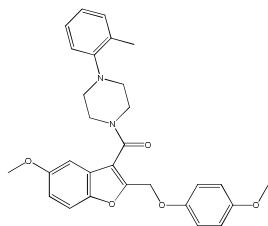
(also for CYP266A1)

Type II- Primary screening

Type II- Clear with concentration dependent validation

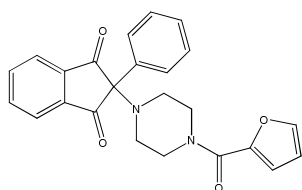
C. SCREENING RESULTS FOR CYP260A1

CYP260A1 Type I hits	Compound code/Remarks
	N062-0006 Type I- Primary screening Type I- Clear with validation
	N045-0002 Type I- Primary screening Type I- Clear with validation
	R092-0022 Type I- Primary screening Type I- Clear with validation
	31649, C074-0114 (also for CYP109C2 and CYP266A1) Type I- Primary screening Not clear with concentration dependent validation
	33152, 7406-0023 (also for CYP109C1) Type I- Primary screening Not clear in 1st validation
	33160, 7746-0008 Type I- Primary screening Not clear with concentration dependent validation



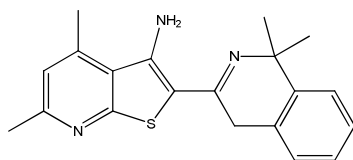
33780, 8016-0529

Type I- Primary screening
Not 100% clear with concentration
dependent validation for Type I



33883, E938-0001

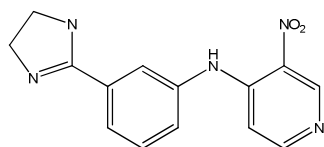
Type I- Primary screening
Not 100% clear with concentration
dependent validation for Type I



34218, 8013-4052

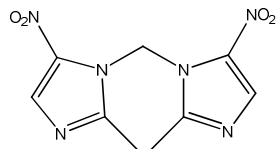
(also for CYP109C1, CYP109C2 and
CYP266A1)

Type II- Primary screening
Type I- Clear with concentration dependent
validation



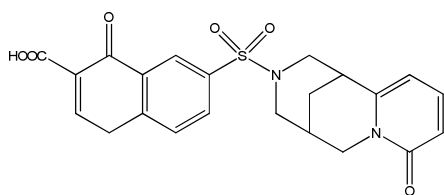
34286, 8013-6360

Type II- Primary screening
Type I- Clear with concentration dependent
validation



34723, 8015-0565

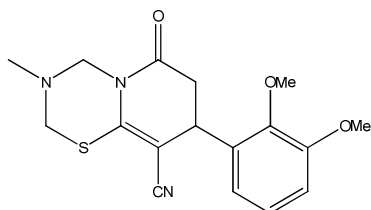
Type I- Primary screening
Type I- Clear with concentration dependent
validation



36260, K511-0020

Type I- Primary screening

Type I- Clear with concentration dependent validation



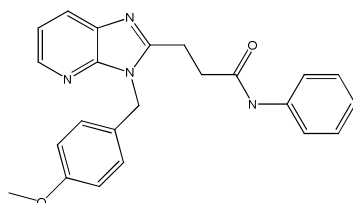
37517, C499-0591

(CYP266A1)

Type I- Primary screening

Not clear in 1st validation

Type I- Clear with concentration dependent validation but not 100% sure

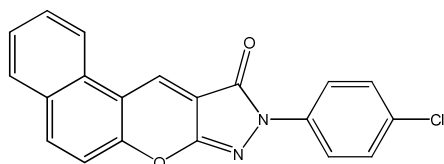


37633, C614-0577

Type II- Primary screening

Type II- 1st validation

Type I- Clear with concentration dependent validation but not 100% sure



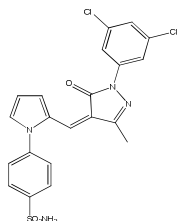
37880, C652-0093

Type I- Primary screening

Type I- Clear with concentration dependent validation but not 100% sure

CYP260A1 Type II hits

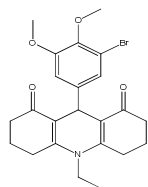
Compound code/Remarks



31132, 4588-0118

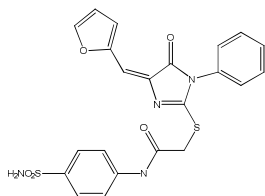
Type II- Primary screening

Type II- Clear with concentration dependent validation



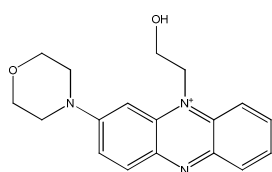
31359, 4984-3625

Type II- Primary screening
Not clear with concentration dependent validation



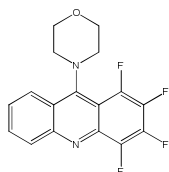
31472, 5183-0085

Type II- Primary screening
Type II- 1st validation
Type II- Clear with concentration dependent validation



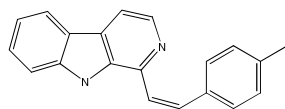
31632, 5451-0327

Type II- Primary screening
Type II- 1st validation
Type II- Clear with concentration dependent validation



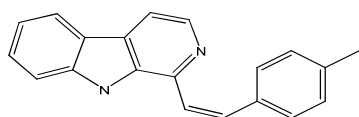
32990, 6550-0046

Type II- Primary screening
Type II- Clear with concentration dependent validation



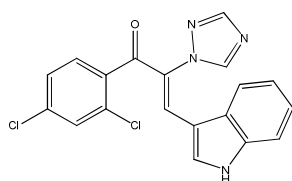
33182, 7462-0037

(also for CYP109C1, CYP109C2 and CYP266A1)
Type II- Primary screening
Type II- Clear with concentration dependent validation



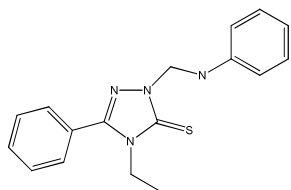
33202, 7462-0046

(also for CYP109C1, CYP109C2 and CYP266A1)
Type II- Primary screening
Type II- Clear with concentration dependent validation



33463, D155-0026

(also for CYP266A1)
Type II- Primary screening
Type II- Clear with concentration dependent validation



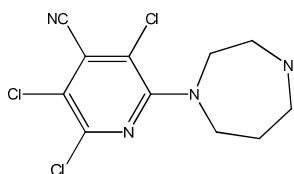
33469, 8008-4502

(also for CYP109C1, CYP109C2 and CYP266A1)

Type II- Primary screening

Type II- 1st validation

Type II- Clear with concentration dependent validation

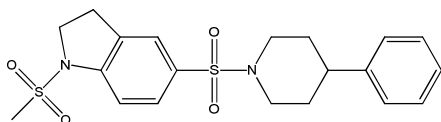


34103, 8012-7514

(also for CYP109C1, CYP109C2 and CYP266A1)

Type II- Primary screening

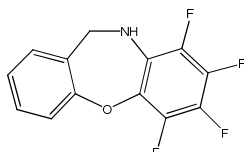
Not clear with concentration dependent validation



35081, 8016-3010

Type II- Primary screening

Type I- Clear with concentration dependent validation

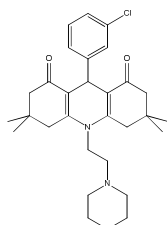


35108, 8017-4048

Type II- Primary screening

Not clear 1st validation

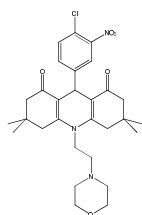
Type I- Clear with concentration dependent validation



35448, C047-1182

Type II- Primary screening

Type II- Clear with concentration dependent validation

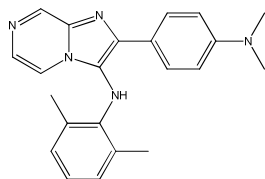


35450, C047-1207

(also for CYP266A1)

Type II- Primary screening

Type II- Clear with concentration dependent validation

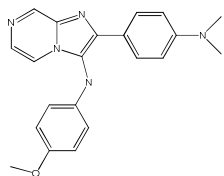


35863, C169-0140

(also for CYP109C1)

Type II- Primary screening

Type II- Clear with concentration dependent validation

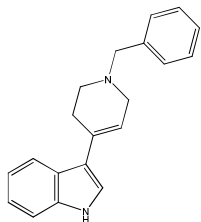


35933, C169-0144

Type II- Primary screening

Type II- Not clear 1st validation

Type II- with concentration dependent validation but not 100% sure

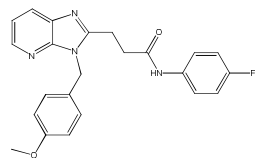


37518, C800-0101

(CYP109C1)

Type II- Primary screening

Type II- with concentration dependent validation but not 100% sure

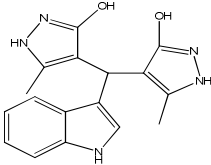
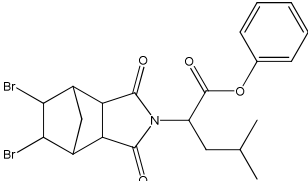
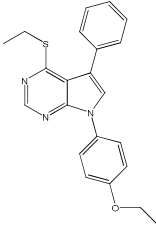
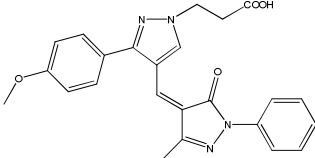
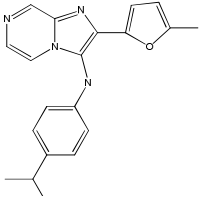
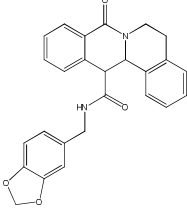


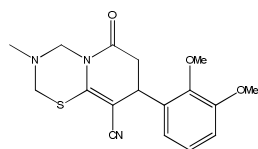
37635, C614-0578

Type II- Primary screening

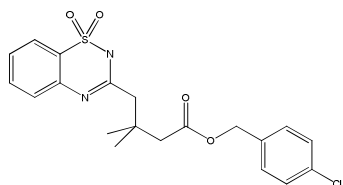
Type II- with concentration dependent validation but not 100% sure

D. SCREENING RESULTS FOR CYP266A1

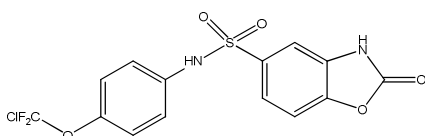
Screening hits for CYP266A1	
CYP266A1 Type I hits	Compound code/Remarks
	31021, 4373-3072 Type I- Primary screening Type I- Clear with concentration dependent validation
	31358, 5089-1245 (also for CYP109C2) Type I- Primary screening Not clear with concentration dependent validation
	31649, C074-0114 (CYP109C2 and CYP260A1) Type I- Primary screening Type I- Clear with concentration dependent validation
	31655, 5237-0515 Type I- Primary screening Not clear with concentration dependent validation
	35862, C169-0253 Type I- Primary screening Type II- Clear with concentration dependent validation
	36386, C240-0165 Type I- Primary screening Type II- Clear with concentration dependent validation



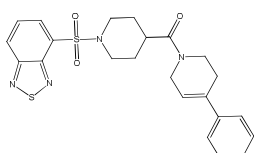
37517, C499-0591
(also for CYP109C2 and CYP260A1)
Type I- Primary screening
Not 100% clear with concentration dependent validation



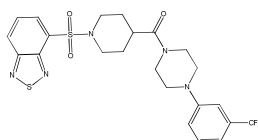
39394, K261-0593
Type I- Primary screening
Type I- Clear with concentration dependent validation



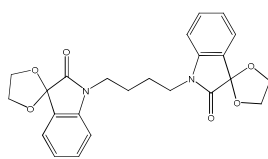
24650 – 8012-0674
Type I- Primary screening
Type I- Clear with concentration dependent validation



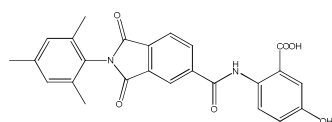
24025 – C361-2424
Type I- Primary screening
Type I- Clear with concentration dependent validation



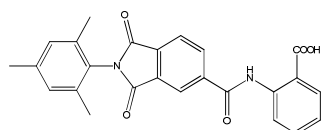
24027 – C361-2436
Type I- Primary screening
Type I- Clear with concentration dependent validation



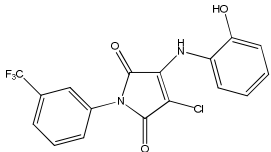
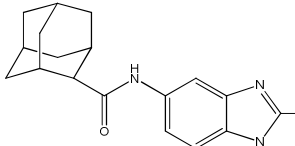
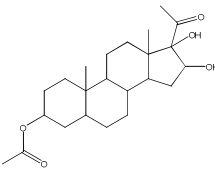
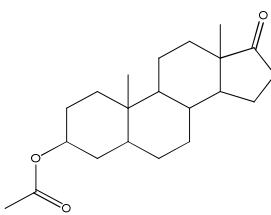
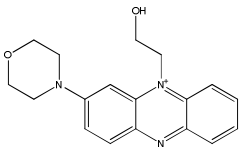
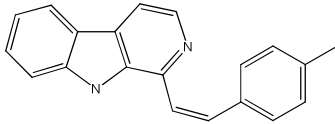
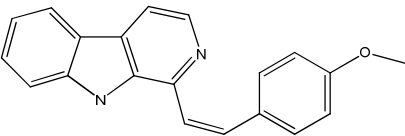
25722 – 8006-1508
Type I- Primary screening
Type I- Clear with concentration dependent validation

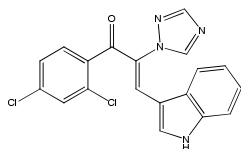


25979 – 4334-1827
Type I- Primary screening
Type I- Clear with concentration dependent validation



25943 – 2922-2951
Type I- Primary screening
Type I- Clear with concentration dependent validation

	<p>30373 – 3793-2107 Type I- Primary screening Type I- Clear with concentration dependent validation</p>
	<p>25101 – 3769-2011 Type I- Primary screening Type I- Clear with concentration dependent validation</p>
	<p>1563-0004 Type I- Primary screening Type I- Clear with concentration dependent validation</p>
	<p>N050-0003 Type I- Primary screening Type I- Clear with concentration dependent validation</p>
CYP266A1 Type II hits	Compound code/Remarks
	<p>31632, 5451-0327 Type II- Primary screening Type II- Clear with concentration dependent validation</p>
	<p>33182, 7462-0037 (also for CYP109C1, and CYP109C2) Type II- Primary screening Type II- Clear with concentration dependent validation</p>
	<p>33202, 7462-0046 (also for CYP109C1, CYP109C2 and CYP266A1) Type II- Primary screening Type II- Clear with concentration dependent validation</p>

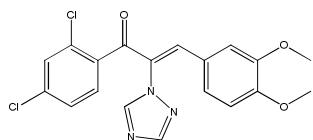


33463, D155-0026

(also for CYP109C1, CYP109C2 and CYP266A1)

Type II- Primary screening

Type II- Clear with concentration dependent validation

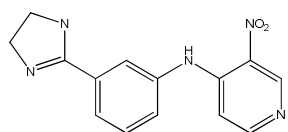


33469, D155-0050

(also for CYP109C1, CYP109C2 and CYP266A1)

Type II- Primary screening

Type II- Clear with concentration dependent validation

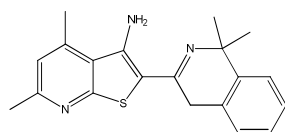


34286, 8013-6360

(also for CYP109C1, CYP109C2 and CYP266A1)

Type II- Primary screening

Type II- Clear with concentration dependent validation

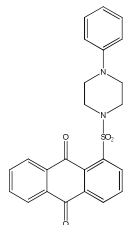


34218, 8013-4052

(also for CYP109C1, CYP109C2 and CYP266A1)

Type II- Primary screening

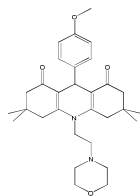
Type II- Clear with concentration dependent validation



35049, 8016-2764

Type II- Primary screening

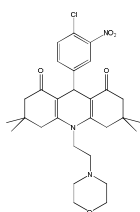
Type II- Clear with concentration dependent validation



35444, C047-0116

Type II- Primary screening

Type II- Clear with concentration dependent validation

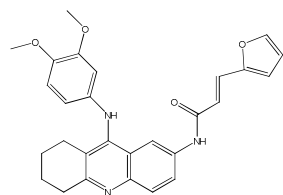


35450, C047-1207

(also for CYP266A1)

Type II- Primary screening

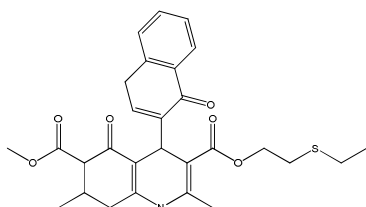
Type II- Clear with concentration dependent validation



35613, C066-3765

Type II- Primary screening

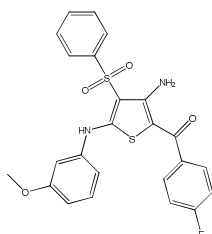
Type II- Clear with concentration dependent validation



35820, C119-0038

Type II- Primary screening

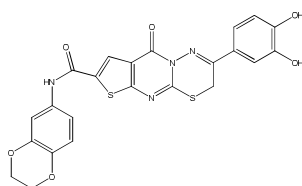
Type II- Clear with concentration dependent validation



36261, C200-3873

Type II- Primary screening

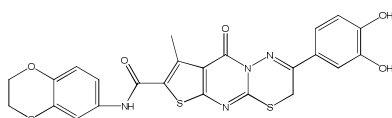
Type II- Clear with concentration dependent validation



36293, K310-0067

Type II- Primary screening

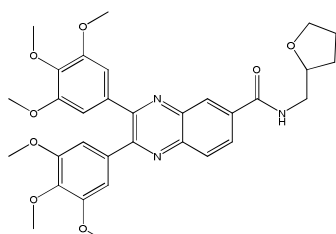
Type II- Clear with concentration dependent validation



36293, K310-0067

Type II- Primary screening

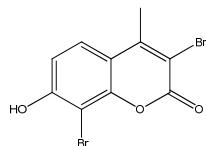
Type II- Clear with concentration dependent validation



39637, K216-1373

Type II- Primary screening

Type II- Clear with concentration dependent validation

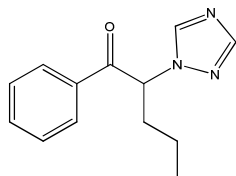


40120, K815-0059

(also for CYP109C2)

Type II- Primary screening

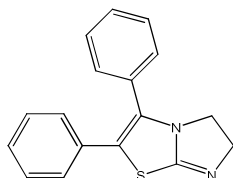
Type II- Clear with concentration dependent validation



25051 – 5509-0020

Type II- Primary screening

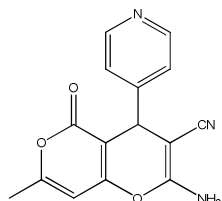
Type II- Clear with concentration dependent validation



29543 – 3442-0170

Type II- Primary screening

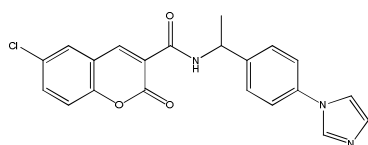
Type II- Clear with concentration dependent validation



24247 – 8003-6142

Type II- Primary screening

Type II- Clear with concentration dependent validation



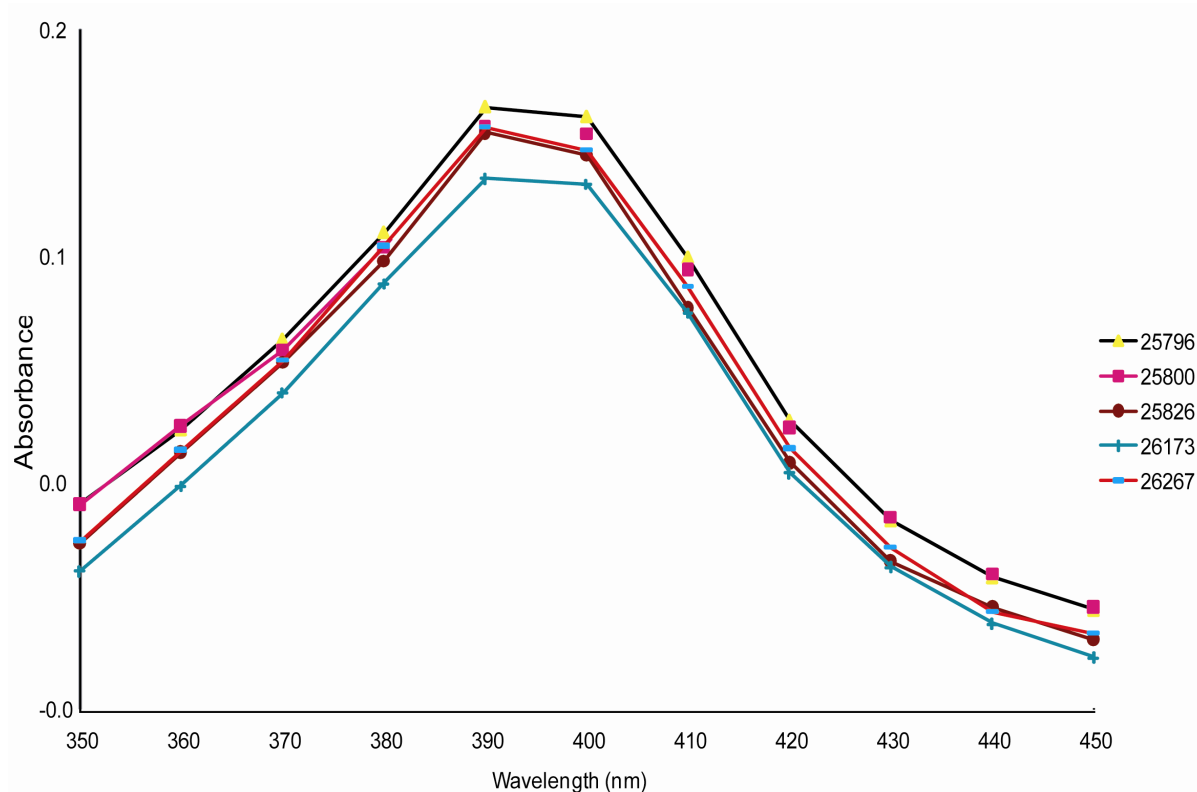
26844 – 8017-8439

Type II- Primary screening

Type II- Clear with concentration dependent validation

D. Representative Type I binding spectra for the hits of CYP260A1 during screening.

During screening only major shifts was considered. The represented Type I hits are shown with the peak maxima at 390 nm. The ligands 25796 (black line with yellow box), 25800 (pink line with pink box), 25826 (brown line with solid brown circle), 26173 (blue line with blue bar) and 26267 (red line with blue mark) are shown. The serial number assigned for each spectrum represents the Chem-Div library number. The screening was done in 384 wells micro titer plate in an automated pipette robot Cyclone ALH 3000 Workstation as described in 'Methods' (Chapter I. 3. 2. 8). The screening was done with 1.5 μM of CYP260A1 and 0.5 μM of ligands in 50 μl assay buffer (50 mM Tris-HCl, pH 7.5, 250 mM NaCl, 20% glycerin, 0.05% Tween-20 and 1 mM DTT). The wavelength of 350-450 nm with an interval of 10 nm was fixed with Tecan-Safire plate-photometer and the absorbance of the substances were recorded and analyzed.



	130	140	150	160	170	180
CYP261A1	FRRW	SPLQ	ELIDEMGFGG	VFSAEGDQ	WTRQRRLLVMS	AFTAGQLRES
CYP261B1	FRRV	RAIE	AVSREADMVG	LFSADGDA	WRKQRLIHP	TFHPRHVEGF
CYP102A1_BM3	DKNL	SQAL	KFVRDFAGDG	LFTSWTHEKN	WKAHNILLP	SFSQQAMKGY
CYP102A2_BS	DKSI	EGAL	EKVRAFSGDG	LFTSWTHEPN	WRKAHNILMP	TFSQRAMKDY
CPP102A3_BS	DKNL	GKGL	QKVRDFGGDG	LFTSWTHEPN	WQKAHRLLP	SFSQKAMKGY
P450foxy	KKTL	KSVL	SQVREGVHDG	LFTAFEDEN	WGKAHRLLVP	AFGPLSIRGM
CYP262A1	FGKE	GGMW	KPIGRLLGNG	LVTAGGDE	WLRNRRRMQP	LFSSRQLAGL
CYP262B1	YRKG	GAIW	ESVRSLLIGNG	LPTSEGDL	WLRQRRMIQP	EFHRDRLAAM
CYP263A1	YPK	GKRY	HELAPVLGNG	LVNSEGEL	WRRQRRHVQP	QFNHANTLKF
CYP110H1	FLPF	ATRA	IAPLVGEHSL	LMLSDEPH	RRERKLLTP	PFHGDRMRAY
CYP110J1	FDVW	GVQL	TEPVFGTSSV	VVTAGARH	RRDRRLLP	PFSAAGMRGY
CYP260A1	STR	MFQAGILN	GGLAAMQGD	HARMRRVYM	FELPRAVSQY	
CYP260B1	STR	TYDTGIMK	GALVTLGGEA	HTRMRRLFNA	VLSPRVISRY	
CYP109C1	SS	VAPPTGKAPD	WIVFSD	PRHTKLRSIVLR	AFTPRSIAGL	
CYP109C2	SS	VSPPGTRTAE	WLIFSD	PRHTKLRLITR	AFTPRVAVGL	
CYP109D1	SS	VTDKIRVLPR	ITLLHDDPR	HTHLRRLVSR	SFTPRRIAEI	
CYP264A1	SSQ	GFRAAWQPA	WVGHNPLAS	SILAMD	GPDHARLRGLVSR	AFGAPAIARI
CYP264B1	SSA	GLRMAEPP	YLRRQNPLSG	SMLLAD	PRHGQLRSISSR	AFTANMVSTL
CYP124E1	CS	GRGLFIED	LPPGDMRDN	PDVMIMDDPR	HARFRALVSK	GFTPRVIOQL
CYP265A1	NM	GTDVLTRO	GVTEGLVRRM	YEYEFLLSTEGAA	HARLRGLLKK	GFTPDVAVDAL
CYP267A1	AKRSSAFVTK	LPDEVRQRL	PLRRN	LA	SWALLLDPPE	HTRIRSLINK
CYP267B1	RD	FRKLPDE	VRRRYFPLSD	RRTTFMD	QHMLDADDPD	HTRLRRAIVQR
P450BioI	VR	ESSTKYQ	DLSHV	QN	QMLLFQNPQD	HRRLRTLASG
CYP266A1	ADRYVALADT	VPPEQKEMNS	YIVKS	LS	MFMLNVENPT	HRRLRNLNTR
CYP259A1	MDH	EH	IKGIVG		RSMAAQDGLL	HRHMRSAINP
CYP117B1	SSY	L	TEGGVGHFLG	QTLMAHDGAS	HRRLRTAMNG	PFSPKGLDAA
CYP152A1_BS	QRQN	ALP	KRVQKSLFGV	NAIQGMDGSA	HHRKMLFLS	LMTPPHQKRL
CYP152A2_C.acet	KRQG	AAP	KRVQETLLGE	NAIQTLDGES	HHRKMLFML	LTNQVQKRL
CYP105D5_S.coeli	TDRTR	DGF	PATSA	ARLAAVRERR	TALLGVDDPE	HRAQRMMVLP
Clustal Consensus						

	190	200	210	220	230	240
CYP261A1	HGTLSTITRR	LREWRASAA	RGAPVDARRD	LARYTVDVTT	AVAFGLDMNL	IERTDPLSD
CYP261B1	FPSIRDITGR	LRELWARAAD	ERAARDVLGD	LMRYSVDVTS	SVAFGRDLNT	LERGADALQE
CYP102A1_BM3	HAMMVDIAVQ	LVQKWERLNA	D-EHIEVPED	MTRLTLDITIG	LCCGFNYRFNS	FYR-----D
CYP102A2_BS	HEKMVDIAVQ	LIQKWARLNP	N-EAVDVPGD	MTRLTLDITIG	LCCGFNYRFNS	YYR-----E
CPP102A3_BS	HSMLLDIATQ	LIQKWSRLNP	N-EEIDVADD	MTRLTLDITIG	LCCGFNYRFNS	FYR-----D
P450foxy	FPEMHDIATQ	LCMKFARHGP	R-TPIDTSDN	FTRLALDTLA	LCAMDFRFYS	YYK-----E
CYP262A1	VDRMFDVVEG	DLPRLEERAR	AGAVDDMDKE	MMQLTQRVIL	ATMFG---VS	ITP-----R
CYP262B1	CDLMVQAIID	GMAGFGAAAA	AGRPVNAERE	LPHITMKVIL	NTMFG---SG	ITK-----E
CYP263A1	VPIIVHEMEA	VLRRWDAQPG	E-FERDINDD	MMDVTFGIAG	EAFFG---AA	LH-----A
CYP110H1	AAAMADTAAR	RLTEAAR---	APRAVAQEL	TQAISLDVII	RAVFGVEEPS	RTS-----
CYP110J1	GDAIAEISLD	VASRWRP---	GRSFSMLAA	TQAIALDVIV	RVVFGVRGEA	RVG-----
CYP260A1	EER---FVRP	ISEQVVDRLA	GKPRVDLLED	FAMELPRRVI	GELFGFPAEK	LH-----
CYP260B1	EEA---TVTP	VARRVVERLV	RKERAELFDD	FAISMPMGVT	SALFGLPBER	IA-----
CYP109C1	EP---RIRE	LSRDLLDPWI	ERGEMDLAAD	YAGPLPTMVI	AEMLGVPED	RA-----
CYP109C2	EP---RIRR	LSRELLDRTI	ESGQMDVAED	YAVPLPLLVI	AEMLGAPAAD	QP-----
CYP109D1	EP---WIGR	LAASLLEATG	D-GPSDLMGA	YAMPLPMVI	ATLLGLPAER	YV-----
CYP264A1	EQ---RARD	LCERLAGRLD	G--EVDFTAA	AAAPLPAFVI	SELLGLDHAL	EP-----
CYP264B1	EH---HMRS	MAVRLTDDL	HRRVVEFISE	FASRAQVSVL	AKLIGFDPGL	EG-----
CYP124E1	ES---HVRE	LVTRLIDDAC	ERGGCDFASD	IAGKLPLSVI	LEVIGVRED	QE-----
CYP265A1	RP---RMRE	RMHALLDEIG	DRTEFDVFTA	VAFERFPGQVI	IELLGLPVSF	EDP-----
CYP267A1	RS---RVET	LVNELLDAVA	PAGRMDVLRD	LGDLLPLLVI	GEVLGVP AED	RH-----
CYP267B1	RP---RIQE	IADGLIDAVI	DRRRMELIAD	FAPPLPTAVI	AELLGLPVED	RG-----
P450BioI	QP---YIIE	TVHLLDQVQ	GKKKMEVISD	FAPPLASFVI	ANILGVP EED	RE-----
CYP266A1	RP---SAHA	VVNELLDAVQ	PRGHMDVVAD	LAYPMPKFI	CGILGMPVED	MG-----
CYP259A1	EAA---AAIA	SATAAKIDAW	LQQESFAVHP	AMRDLTLDII	LRICGVP TGD	IR-----
CYP117B1	EVG---AIVA	TSVERKVRWS	LGRRDVQLLR	ETRELALEV M	FKITGV EDE	LP-----
CYP152A1_BS	AELMFEWKA	AVTRWEK---	ADEVVLE	EAKEILCRVA	CYWAGVPLKE	TEVK-----
CYP152A2_C.acet	AELTTEKWEA	SASKWH---	TKSIVLFN	EANEILCOVA	CHWAGVPLME	SDIK-----
CYP105D5_S.coeli	RPS---IRRI	VGERLDAMIA	QGPADLVTA	FALPVPSMVI	CALLGVPYAD	HEF-----
Clustal Consensus						

	250	260	270	280	290	300
CYP261A1	QLETVFASLN	RRVFAPFPYW	RHVKLPADRA	LDQALSEVRT	RMLDILRTR	AELDRDPARA
CYP261B1	HLQRVFLGLN	RRMLAPLPYW	RYLGLPADRA	YDRALVAVKR	LILDLVTAAR	REIEGGADRP
CYP102A1_BM3	QPHPFITSMV	RALDEAMNKL	QRANPDDPAY	DENK-RQFQE	DIKVMNDLVD	KIIADRKASG
CYP102A2_BS	TPHPFITSMV	RALDEAMHQM	QRLDVQDKLM	VRTK-RQFRY	DIQTMFSLVD	SIIAERRANG
CPP102A3_BS	SQHPFITSMV	RALKEAMNQS	KRLGLQDKMM	VTK-LQFQK	DIEVMNSLVD	RMIAERKANP
P450foxy	ELHPFIEAMG	DFLPESGNRN	RRPPFAPNFL	YRAANEKFG	DIALMKSVAD	EVVAARKASP
CYP262A1	EADSLGEVLL	VAIQALNARM	FLYFMPDRL-	-LPGERALRD	ATARIDEAIL	RLVRRER-RRS
CYP262B1	EADAVGGSRLR	YALDYMLLGA	ALRALPSWIP	-APGRRRFER	SAKAIDEHVF	RFAIQR-RAQ
CYP263A1	HTDTVREAFK	YALSIALKRM	YSLVNPPLSW	PLPSHLRFRR	AMDRVHAVID	GIIDGY-QRG
CYP110H1	AFARAVVAMT	DALTPALTFL	P--PLQRELG	GLGPYARFRR	RVGELDALFR	AQIERARAAE
CYP110J1	RTREAVLGLI	ESLSPSFMII	P--ALRRDFA	GFGPYARHKR	AARALDALLF	EEIR-ARRAE
CYP260A1	ETDERVRAML	RGLVRMHDP	AVAESQR---	-----	---AYGETL	GLITEVVERE
CYP260B1	ENDALIRKMI	RSVVMQDPV	VVAEGRS---	-----	---AHAAME	AQLREIAERE
CYP109C1	RLLRFSEVIV	NLSHAISGGE	EAAARAVSEHA	-----	---VVKEMK	VYLAGLIEAR
CYP109C2	HFKRWDAIL	DLSHTVSGSE	EAAARLDAFT	-----	---AVTAEMQ	AYLRGLVEQR
CYP109D1	QFRSWSSEVM	SYSG-IPAE	RASRGK---	-----	---AMV	DFFAAELEAR
CYP264A1	HFKRWDDLL	SVTPEP-ASA	EHAARVR---	-----	---ATIAELD	RYMADVIAAR
CYP264B1	HFKRWATDLV	IVGVIPPEDH	ARIAEVR---	-----	---RTIDEME	QYMLGLLASR
CYP124E1	QMLDWTTRFF	GASDPAYG--	VTPPEELN---	-----	---AVLHNMN	AYAHQLAEQR
CYP265A1	DFRRWCSDLG	YIVGLEVK-	KHLPRLE---	-----	---AAFGLR	GYIPTIAR
CYP267A1	RLKGWSNALS	GFLGAGPRTL	ETAGGAL---	-----	---SAVALE	DYFRGILAAE
CYP267B1	RFRRTWKILL	----APAKDR	EFVERAQ---	-----	---PVVEEFA	AYFRALADAR
P450BioI	QLKEWAASLI	QTIDFTRSRK	ALTEGNI---	-----	---MAVQAMA	-YFKELIQKR
CYP266A1	LIKQLSDDVS	VYIGSAGKAA	GCTPPAY---	-----	---HAIVEFS	KLFRPLVEAR
CYP259A1	AWSRAYRELF	LGVYPIPRGL	PGMPAYRSDR	-----	---ARAWIDD	RLRKLALAL
CYP117B1	EWRHYEDIM	LVILSLPFDV	PGSPRRGLR	-----	---ARAWLDE	RIGRILGVR
CYP152A1_BS	ERADDFIDMV	DAFGAVGPRH	WKGRARP---	-----	---RAEEWIE	VMIEDARAGL
CYP152A2_C.acet	NRAEDFSSMI	DSFGAVGPRH	WKGKKARN---	-----	---TIEAWIK	EIIENVRSGR
CYP105D5_s.coeli	-FEEQSRRL	RGPLPADTRD	AR-----	-----	---DRLE	AYLGELIDRK

Clustal Consensus

	310	320	330	340	350	360
CYP261A1	AAPRSILLEAL	LVARDADDPK	ARLSEQEVLG	NVLTILLAGE	DITADTMAWM	LHFMALRQDV
CYP261B1	ARPRTLEILM	IAARDGEDPR	ARLSDAEVYA	NVLTMLLAGE	DITANTLAWI	LYYLGSRQDV
CYP102A1_BM3	EQSD-DLLTH	MLNGKDPETG	EPIDDENIRY	QIITFLIAGH	ETTSGLLSFA	LYFLVKHPHV
CYP102A2_BS	DQDEKDLLAR	MLNVEDPETG	EKLDDENIRF	QIITFLIAGH	ETTSGLLSFA	TYFLKHPDK
CPP102A3_BS	DENIKDLSL	MLYAKDPVTG	ETLDDENIRY	QIITFLIAGH	ETTSGLLSFA	IYCLLTPEK
P450foxy	-SDRKDLLAA	MLNGVDPQTG	EKLSDENITN	QLITFLIAGH	ETTSGLTSFA	MYQLKHPDA
CYP262A1	KEERDDLLSL	LLLRAD-ESG	SGMDDRQLRD	ELVTMFIAGN	ETTATMTWL	FYLLDRNPGI
CYP262B1	PGRCGDLLSI	LLATVDAETG	EQMTNQLRD	EAVSMFLAGY	ETTSHALWA	LHLLVHPDL
CYP263A1	AGGEDNVIVR	LMNSVDPETG	AKMDRAQLRD	EIKTILMVGH	ETSSVTASWA	LYQLARHPV
CYP110H1	GD---DILSL	MVSARYDDGS	R-MSDQAFID	ELRTLFLFAGH	ETTALALAWA	LDHVHRHPGV
CYP110J1	GDASQDILGL	MMSARHDDGA	G-MSDVEIAD	QLRALLFAGH	ETTAMSLGWA	MYWLHREPAV
CYP260A1	SRDTSDELLG	EILRTLKA-E	HMDTIEASRQ	IVLSLILGQY	ETTSWLVAE	IHALLAHDT
CYP260B1	VAHPSDILLG	EIARAIVA-E	GLGGVEACEG	VVLTILLGSY	ETTSWMLANL	LVALLAHDA
CYP109C1	RRAPAEIDLT	RLVEAEVD-G	ERLDEGDILN	FFQFLLAAGT	ETATNLIDNA	ILCFLESPAE
CYP109C2	RAAPEDDILT	RLVEAEVD-G	ERLNEDEILG	FFQILLVAGH	ETTTNLIGNA	MLCFLENDE
CYP109D1	RRAPSGDLIS	ALVEAEID-G	ARLDTPEAVG	FCVGLLVAGN	DITTNLIGNM	AHLLSERDEL
CYP264A1	RRSPSDDIVS	ELARAG----	ELLDGREIID	LLVSILGGGL	ETTHFLGSS	MLLLAERDAE
CYP264B1	RRHLENDIVS	ELLRSRRD-D	DGTTDQDLVS	LLSILLVAGL	ETSTSLMTHM	VLILAQRQMW
CYP124E1	RKEPKDDMLS	LLMAAEVD-G	EKLISYTEFGG	FFNLLLTAGH	DITKNLISNG	MLALLEHPDQ
CYP265A1	RAEPRGDILS	TLAAAEQN-G	DRFSEELVC	QIVALLFGGL	DITRRTIGKA	FLVMLEHPDQ
CYP267A1	RQSPGNDILS	QLLILAEQ-G	MILGEQELLS	TCCMLLFGGH	ETTKNLISNG	LLALLLRSE
CYP267B1	RKAPRDDILS	GLLLAEQ-G	HKLSPAEELS	MVFLLLVAGH	ETTVHLIASG	MLLLSHPEAE
P450BioI	KRHPQDDMIS	MLLKGREK--	DKLTEEEAAS	TCILLIATAGH	ETTIVNLISNS	VLCLLQHPDQ
CYP266A1	RKEPKDDILS	SMVTTRVD-G	DSLSDDEVIA	NCILFLVAGF	ETVTNLICAG	TLALLEHPDQ
CYP259A1	GGGAAAGSIVE	MLSNARDDG	EPISDQGIVD	NLRIVLVAAGH	ETSASILAWI	VIVLSRRSEV
CYP117B1	ARGDAKGLLP	ALLTARDEQG	EPISDQDLVG	NLRILLLVAGH	ETSASVMAWC	VAHLAESPAV
CYP152A1_BS	LKTTSGTALH	EMAFHTQEDG	SQIDSRMAAI	ELINILRIP-I	VAISYFLVFS	ALALHEHPKY
CYP152A2_C.acet	IRAEESPLH	EIAFYIDVNG	QQMPAEMAAI	ELINILRIP-I	VAISTFITFS	ALALYEHSEY
CYP105D5_s.coeli	RRAPGEGLLD	DLVRRQAS-E	GATDREQLIA	FAVILLVAGH	ETTANMISLG	TYTLTLPGR

Clustal Consensus

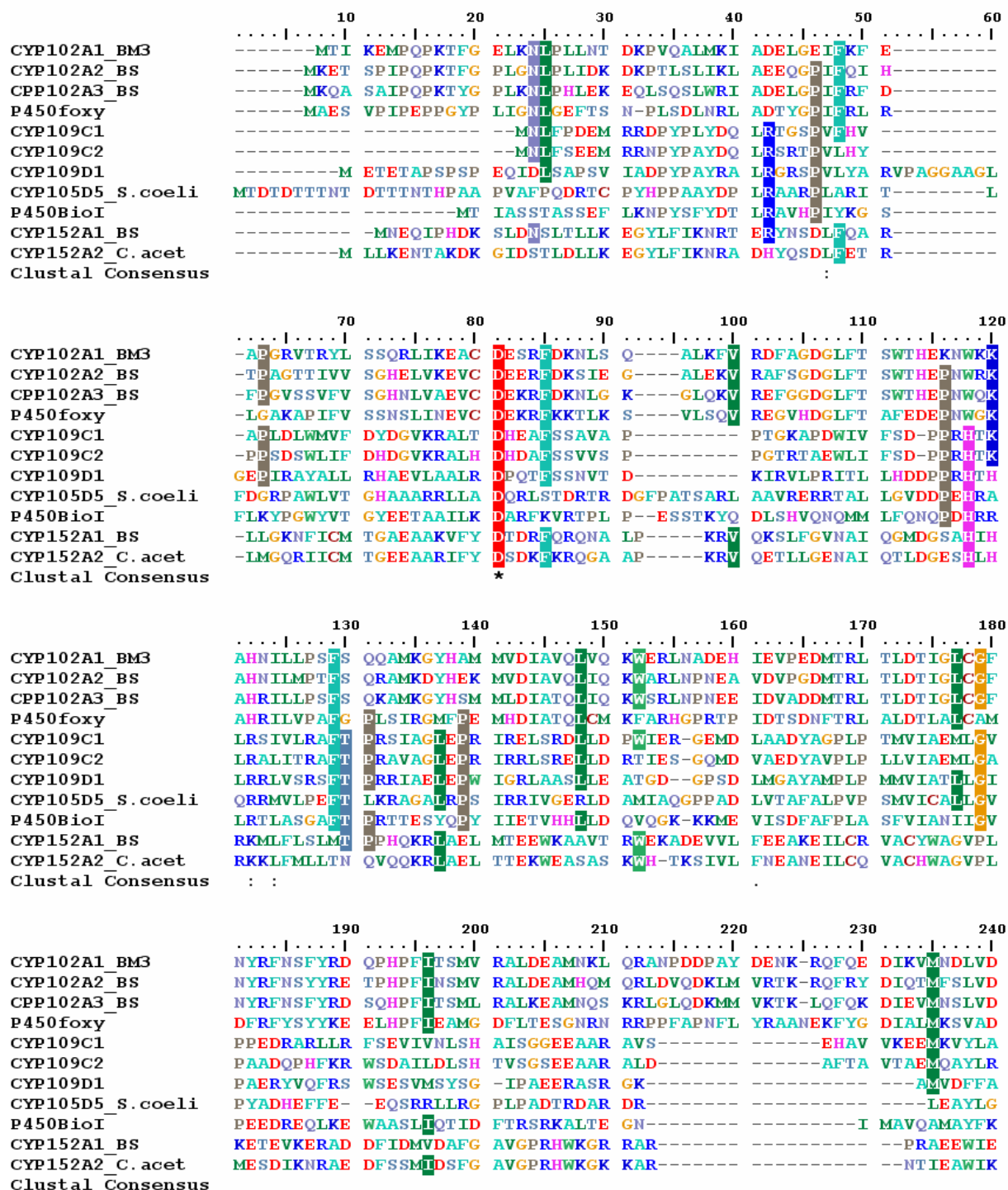
	370	380	390	400	410	420
CYP261A1	QARMRAEVDA	ELGDADLPES	PEHAQRRLRYL	GAVAQETLRL	KSAAPILFME	ACADTVVG-S
CYP261B1	QARAREEVDR	VLGGGLIP-T	LEQCKRLEYT	AAVTQETLRL	RSAAPLLFLE	ANHDTAVG-D
CYP102A1_BM3	LQKAAEEAAR	VLVD--PVPS	YKQVKQLKYV	GMVLNEALRL	WPTAPAFSLY	AKEDTVLGGG
CYP102A2_BS	LKKAYEEVDR	VLTD--AAPT	YKQVLELTYI	RMILNESLRL	WPTAPAFSLY	PKEDTVIGGK
CPP102A3_BS	LKKAQEEADR	VLTD--DTPE	YKIQQLKYI	RMVLNETLRL	YPTAPAFSLY	AKEDTVLGGG
P450foxy	YSKVQKEVDE	VVGR--GPVL	VEHLTKLPYI	SAVLRETLLR	NSPITAFGLE	AIDDTFLGGK
CYP262A1	ERKLRAEIEE	VVGD--RRPT	AADLSRMEYT	KMVIQEAARM	YPPSWLVPRT	VKEDDQIC-G
CYP262B1	LRALATEVDE	ALGD--RRPG	FADVPRLLPLA	LAVVQEAALRL	YPPAYWIPRT	AIEDDEID-G
CYP263A1	CARLTDEIDR	VLGG--RSPT	ARDIESMPYL	TMVFNECLRL	LPSPVFFILRS	PLEDDVIC-G
CYP110H1	LARLRDEIDA	LGPE---PD	PERLAALPYL	DAVCKEALRV	YPIVTESPRL	LAQPFRLG-E
CYP110J1	LARLLDELDT	LGPA---PE	ADALASLPYL	EAVCLEALRL	HPPVVDVARV	VKRPFRLK-G
CYP260A1	LARVRQDPS-	-----	-----LL	PAAIEEGMRW	CPSSFGVL-R	MVERDVRLLD
CYP260B1	MNQLRQQPS-	-----	-----LL	PQAIEESTRW	CSSAAGIV-R	FVEREATIGG
CYP109C1	LFRLRAAPE-	-----	-----LL	PSAIEEVLRLH	RSPLOMVF-R	ETRRAVEVHG
CYP109C2	LARLRRAPE-	-----	-----LL	PSAIEEVLRLY	RSPVQAMF-R	VTRRDVPMHG
CYP109D1	YRRAQQDRS-	-----	-----LV	GPITIEETLRH	SSPVQRLL-R	VTRRPVDSG
CYP264A1	LERLRASPQ-	-----	-----LI	PRFIEEMMRV	DGPTQSVR-R	LTTSDVALAG
CYP264B1	MDRLRAEPA-	-----	-----LI	PHFIEEVMRF	EAPVHATM-R	LTVTETELGG
CYP124E1	RRRLDDPS-	-----	-----LL	PTAIEEMLRF	TPPVYFVR-R	SAVRDIELHG
CYP265A1	RRLLQERP-	-----	-----LA	EPAAEEILRY	APGTRVAV-R	VATCDLELHG
CYP267A1	REALRATPS-	-----	-----LI	GPAVEELLRY	DSPVQWMS-R	VALDDIELDG
CYP267B1	RRRLDEDPG-	-----	-----LV	GSAVEEALRC	EGPAELSTIR	WSLEDIELEGG
P450BioI	LLKQRENPD-	-----	-----LI	GTAVEECLRY	ESPTQMTA-R	VASEDIDICG
CYP266A1	LELLKRDSR-	-----	-----LM	EGAIDEMLRV	YPPVNRTA-R	LCVEDIPLRG
CYP259A1	WSALCDEHAA	APDAP-VPMS	AREADRLPVA	NAIFREVLRLM	YTPVWFIF-R	TVTETITLSG
CYP117B1	WRALREEASS	APDLP-R--S	PADLRRFPFA	EAVFREALRL	HPPVPHDA-R	RAVADFELDG
CYP152A1_BS	KEWLRSGNSR	-----	-----ER	EMFVQEVRRY	YPPGPFPLG-A	LVKKDFVWNN
CYP152A2_C.acet	REKLQSKDIR	-----	-----YL	EMFTQEVRRY	YPPFAPFVG-A	RVRKDFLWNN
CYP105D5_s.coeli	LAELRADPA-	-----	-----LL	PGAVEELMRV	LSIADGLL-R	MATEDIHDVG
Clustal Consensus						

	430	440	450	460	470	480
CYP261A1	LRLAGTRMI	LLTRQIALKE	EHFH-DAASF	MEERWLTTPRP	AGAGKHDPRR	ALAFGSGP--
CYP261B1	VAVPAGAGVV	ALTRATGRKP	EYFG-DEESF	REERWLGGA	S-TLPHPRM	ALAFGGGP--
CYP102A1_BM3	YPLEKQDELM	VLIPQLHRDK	TWGGDDVEEF	REER-----	ENPSAIPQHA	FKPFGNGQ--
CYP102A2_BS	FPITNDRIS	VLIPQLHRDR	DAWGKDAEEF	REER-----	EHQDQVPHHA	YKPFNGNQ--
CPP102A3_BS	YPIISKQPVVT	VLIPKLHRDQ	NAWGPDAEDF	REER-----	EDPSSIPHHA	YKPFNGNQ--
P450foxy	YLVKKQEIIVT	ALLSRGHVDP	VVYGNADKDF	IPERMLDDEF	ARLNKEYPNC	WKPFGNGK--
CYP262A1	YVPVAGATVI	LSQVYMHHP	AFWE-APAEF	DEER---FTP	ERSASRPRYS	YMPFGGGP--
CYP262B1	FHIPAGTMVG	VMTYVLRHHP	DHWE-APARF	DEGR---FTP	EHARARHPLA	FIPFGIGQ--
CYP263A1	YKVAAGSTVA	IVPVWTHRHP	EFWS-DPLEF	KEER---FAN	HRSRADDKLA	FIPFGAGQ--
CYP110H1	HELLEGTGVA	PCILLVHHHP	ELYP-EPSRF	REER-----	FLERKFSPEF	YLPFGGCH--
CYP110J1	YTIPTAGEAIA	ASPLLLHGRE	DLYP-SPERF	RESR-----	FLDRKFTPFE	FIAFGGGA--
CYP260A1	QALSAGTVVC	LAGIAGNYDE	TAYP-SPEVY	DIDRK-----	-----PLPA	ANVFGGGA--
CYP260B1	ETLAAGTILY	LSLIARHYDE	EIYP-RPETF	DIHRR-----	-----PVGM	LN-FGGGL--
CYP109C1	QVTPACKLVL	PVIGSANRDP	HQFH-DPGRF	DIGR-----	-----DPNP	HVAFGHGI--
CYP109C2	QVTPAGKAVL	AMIGSANRDS	QQFR-DPDRF	DIAR-----	-----DPNP	HIAFGHGI--
CYP109D1	VMTIPAGHLVD	VVFGAANRDP	AVFE-EPDAF	RLDR-----	-----PPAE	HLAFGGGT--
CYP264A1	VTIPTAGSLVL	ALVGSANRDE	VRFV-DPDRF	DLHR-----	-----G-QP	SLTFGHGA--
CYP264B1	TRLPRAHAVVA	LLISSGLRDE	ARFQ-EPDRF	NEER-----	-----GDQA	NLAFGHGA--
CYP124E1	QRITAGDKVV	LWYVSANRDE	TVFS-DPHRF	DVGR-----	-----TPNE	HLSFGIGP--
CYP265A1	VPTRAGEMVI	AATGAANRDP	RVFE-APETF	DITR-----	-----RCAA	PLSLGAGP--
CYP267A1	VRIIPKDRAF	LVLGAANRDP	AQFP-DPDKL	DFRR-----	-----TDIR	HISLGLGV--
CYP267B1	ARVPAGEGVV	AGLLAANRDP	QHFP-DPDRF	DIGR-----	-----SPNR	HIGFGGCI--
P450BioI	VTIRQCEQVY	LLLGAANRDP	SIFT-NPDPV	DITR-----	-----SPNP	HLSFGHGH--
CYP266A1	KVIKKQIIVV	LMLGAGNRDP	SEYP-DPDRF	DIARE-----	-----NRSR	PLSFGGCH--
CYP259A1	RRIPRNTPLA	LSPAQFGYDP	SLFP-DPDRF	DLSRWR-----	-----DRSIPP	GGPLELAAFGGGP--
CYP117B1	RAVPAGTHVA	IPLVLLSRDP	ELYP-DPDSY	REERWV-----	-----GRKEALS	PLEIAQFGGGA--
CYP152A1_BS	CEFKKQTSVL	LDLYGTNHDP	RLWD-HPDEF	REER-----	-----FAERENL	FDMIPQGGGHAE
CYP152A2_C.acet	CEFKKEMVLV	LDIYGTNHDS	RIWQ-KPYEF	IEDR-----	-----FRSYKGNL	FDFIPQGGGDP
CYP105D5_s.coeli	QTIIRAGDGVV	FSTSVINRDE	SVYP-EPDAL	DWHR-----	-----PARH	HVAFGHGI--
Clustal Consensus						

	490	500	510	520	530	540
CYP261A1	--RVCPGRAL	SMVESAMVGA	MVARDFDVSL	VDPAR----	PVKELLGFTMK	PEGLFVRFPP
CYP261B1	--RVCPGRAM	ALLECAVTVG	MVRSFTVHL	ADHTA----	PVQERMAFAMQ	PEGLRIRFTK
CYP102A1_BM3	--RACIGQQF	ALHEATLVLG	MMIKHDFE	DHTNY----	E DHTNY----	PEGFVVKAKS
CYP102A2_BS	--RACIGMQF	ALHEATLVLG	MLIKYFTL-I	DHENY----	E LDIKQTLTK	PGDFHISVQS
CPP102A3_BS	--RACIGMQF	ALQEATMVLG	LVIKHFEL-I	NHTGY----	E LKIKEALTK	PDDFKITVKP
P450Foxy	--RACIGRPF	AWQESLLAMV	VIFQNFNTM	TDPNY----	A LEIKQTLTK	PDHFIYINATL
CYP262A1	--RQCTGNLF	SIMBAQIVIA	VITRRRLRMRL	VPGHP----	V SPQAVATLRP	RHGLKMTLHA
CYP262B1	--RQCVGKEF	ALMEGQLILA	RLLDQRYRISA	VPGR-----	T RLHFATTLRT	SGGVWLRLEP
CYP263A1	--RIGLGEFM	GQLEGKIMVA	MAIDRYRRL	VPGFD----	P RCRGFISLQP	VNGMRMLKR
CYP110H1	--RRCTGAAF	AMFEMKIVLG	VALSAAVEFRL	LDERP----	P RPVRRNLTL	PSGGVPLVL-
CYP110J1	--RRCLGAAF	AMYEMKVALG	TILGRYRLRL	ESQAP----	I HHVRRGLTMG	PSGDVAMILE
CYP260A1	--HFCVGGAPL	ARMEARVGLQ	ALLARFPGLR	AVPEERPSFM	YGAKDSVAHG	PDKLPVLLH-
CYP260B1	--HYCVGGAPL	ARMEARVGVV	LILERRFPALR	ADPTVQPTFS	TAPRGAAAF	PDQIPALLV-
CYP109C1	--HFCIGAAPL	ARLEARIALP	DILLQRLKGLR	RASDE----	PW EPRKAIIVHG	PARLPTRFEP
CYP109C2	--HFCIGAPL	SRLEARIALA	DLVRLGGLE	RASDA----	PW EPRRAIIVHG	PARLPTRFEP
CYP109D1	--HFCIGAAPL	ARMEARIALN	ALIDCYESIT	PGEAP----	PL RQTRAIMPLG	FESLPLVLR
CYP264A1	--HFCIGAAPL	ARMEAKVALP	VIVRRIGEVV	RAPG-----	EI PYNRTLTVRG	PVSLPLRFRP
CYP264B1	--HFCIGVFL	ARVQARIVLE	ELLRCHRIV	LRTD-----	RL EWQAALNTRS	PVALPIEVIP
CYP124E1	--HYCLGLVFL	ARLEARVAFE	ELLRRLPDLE	LAGPA----	V RLRSN-WSSG	VKSMPVRCAR
CYP265A1	--HFCIGAYL	ARAEIQEAVP	TILRRLEEA	LASPV----	V VRTAREGSPG	PRVIPLRCPV
CYP267A1	--HYCAGSAL	ARVEAQAAIS	TILRRFPDAE	LSPGP----	L TWRMNPGMRG	VTALPIELGP
CYP267B1	--HFCIGAML	ARTEAAIAFS	TILRRLLPRIE	LATSTR----	DI VWSEWPTIRG	PAAVPVVF--
P450BioI	--HYCLGSSL	ARLEAQIAIN	TILDRMPSLN	LADFE----	WRYRP----	---
CYP266A1	--HFCIGSHL	ARMEGEVALG	ALLDRMPNLR	LATQE----	V EWRGNSRFRG	LRALPVSF--
CYP259A1	--HFCIGYNV	ACVEALQLQF	ILVVALRRAG	LVPHLEGASP	KHVFYPTGHP	TAATRVAFR
CYP117B1	--HFCIGYHL	AWMEIVQLLV	ALGRELPASG	--PRLGGGFP	ASRYLPLVLRP	SGGTRVFRD
CYP152A1_BS	KGHRCPGEGE	TIEVMKASLD	ELVHQIEYDV	PEQLH----	YSLARMPSLP	ESGFVMSGIR
CYP152A2_C.acet	STRCPGEGE	TLEIMKTSLD	FISTKIDFTV	PDQDLS----	YLSKIPITLP	KSGFIIDNIN
CYP105D5_S.coeli	--HQCIGQNL	ARAELEIALE	SIFDRLLPTLR	LAAPAD----	EI PFKPGDTIQG	MLELPVAW--
Clustal Consensus	: * * :					

	550	560	570	580	590	600
CYP261A1	RHR-----					
CYP261B1	RASA-----					
CYP102A1_BM3	KK--IPLGGI	PSPSTEQSAK	KVRKKAENAH	-----NT	PILLVLYGSNM	GTAEGTARDL
CYP102A2_BS	RH--QEAIHA	DVQAAEKAAAP	DEQKEKTEAK	GASVIGLNNR	PILLVLYGSDT	GTAEGVAREL
CPP102A3_BS	RK--TAAINV	QRKEQADIIKA	ETKPKETKPK	-----HGT	PILLVLFGSNL	GTAEGIAGEL
P450Foxy	RHGMTPTELE	HVLANGGATS	SSTHNIKAAA	NLDAKAGSGK	PMAIFYGSNS	GTCEALANRL
CYP262A1	V-----					
CYP262B1	RPGR-----					
CYP263A1	REPSSAVRAA	PAGREEPAGA	EPVVTEPVAA	AMSAHGCPYS	ARA-----	
CYP110H1	----RARGRS	R-----	LYAAA--			
CYP110J1	GERERARAEG	RPAAVQGSPA	EPALRPAAPO	PAGRCPMGFS		
CYP260A1						
CYP260B1						
CYP109C1	GRRAAVHAAG	AHTPCPG--				
CYP109C2	GPRLGG----	APG----				
CYP109D1	SRATA-----					
CYP264A1	A-----					
CYP264B1	VSTTAARESP	VVQGIW----				
CYP124E1	PAR-----					
CYP265A1	RPR-----					
CYP267A1	QSSAS-----					
CYP267B1						
P450BioI						
CYP266A1						
CYP259A1	C-----					
CYP117B1						
CYP152A1_BS	RKS-----					
CYP152A2_C.acet	LKL-----					
CYP105D5_S.coeli						
Clustal Consensus						

B. The amino acid sequences of the fatty acid hydroxylating P450s [CYP102A1 from *Bacillus megaterium*; P450BioI (CYP107H1), CYP102A2, CYP102A3 and CYP152A1 from *Bacillus subtilis*, CYP152A from *Clostridium acetobutylicum*, CYP105D5 from *Streptomyces coelicolor* A(3)2 and P450foxy from *Fusarium oxysporum*] were accessed from EMBL database. They were aligned with CYP109 family of So ce56 in Clustwl W2. A tree was generated to find out the relatedness of CYP109 family of So ce56 with that of known fatty acids hydroxylase P450s.



CYP102A1_BM3	KIITADRKASG	EQSD-DLLTH	MLNGKDPETG	EPIDDENIRY	QIITFLIAGH	ETTSGLLSFA
CYP102A2_BS	SIIABRRANG	DQDEKOLLAR	MLNVEDPETG	EKLDDENIRF	QIITFLIAGH	ETTSGLLSFA
CPP102A3_BS	RMIABRRKAP	DENIKOLLSL	MLYAKDPVTG	EILDDENIRY	QIITFLIAGH	ETTSGLLSFA
P450foxy	EVVAARKASP	-SDRKDLLAA	MLNGVDPQTG	EKLSDENITN	QLITFLIAGH	ETTSGLLSFA
CYP109C1	GLIEARRRAP	---ABDLLTR	LVEA--EVDG	ERLIEGDILN	FFQFLLAAGT	ETATNLIDNA
CYP109C2	GLVEQRRAAP	---EDDLLTR	LVEA--EVDG	ERLNEDEILG	FFQLLLVAGH	ETTTNLIIGNA
CYP109D1	AELEARRRAP	---SGDLISA	LVEA--EIDG	ARLDTPEAVG	FCVGLLVAGN	DTTTNLIIGNM
CYP105D5_s.coeli	ELIDRKRRAAP	---GEGLLDD	LVRR--QASE	GATDREQLIA	FAVILLVAGH	ETTANMISLG
P450BioI	ELIQKRKRHP	---QQDMISM	LLKG--REK-	DKITEEEAAS	TCILLAIAGH	ETTIVNLISMS
CYP152A1_BS	VMIEDARAAGL	LKTTSGTALH	EMAFHTQEDG	SQLDSRMAAI	ELINILRIP-I	VAISYFLVFS
CYP152A2_C.acet	EILENVRSGR	IRAEESGPLH	EIAFYIDVNG	QQMPAEMAAI	ELINILRIP-I	VAISTFIFFS
Clustal Consensus	:	:	:	:	:	:

		310	320	330	340	350	360
CYP102A1_BM3	LYFTVKNPHV	LQKAAEEAAR	VLVDPVPSYK	QVKQLKYVGM	VLNEALRLWP	TAPAFSLYAK	
CYP102A2_BS	TYFLKHPDK	LKKAYEEVDR	VLTDAAPTYK	QVLELTYIRM	ILNESLRLWP	TAPAFSLYPK	
CPP102A3_BS	IYCLLTHPEK	LKKAQEEADR	VLTDDTPEYK	QIQQLKYIRM	VLNETLRLYP	TAPAFSLYAK	
P450foxy	MYQLLKNPEA	YSKVQKEVDE	VVGRGPVLVE	HLTKLPYISA	VLRETLRLNS	PITAFGLEAI	
CYP109C1	ILCFLESPAE	LFRIRAAPE-	-----	-----LLPS	AIEEVLRLRS	PLQMVFRETR	
CYP109C2	MLCFLENPDE	LARRAAPE-	-----	-----LLPS	AIEEVLRLRS	PVQAMFRVTR	
CYP109D1	AHLLESRPEL	YRRAQDRS-	-----	-----LVGP	LIIEETLRHS	PVQRLLRVTT	
CYP105D5_s.coeli	TYTLLEINPGR	LAELRADPA-	-----	-----LLPG	AVEELMRVLS	IADGLLRMAT	
P450BioI	VLCLLQHPAQ	LLKLRNPDP-	-----	-----LIGT	AVEECLRYES	PTQMTARVAS	
CYP152A1_BS	ALALHEHPKY	KEWLRSGNSR	-----	-----EREM	FVQEVRRYYP	FQPFGLALVK	
CYP152A2_C.acet	ALALYEHSEY	REKIQSKDIR	-----	-----YLEM	FTQEVRRYYP	FAPFVGARVR	
Clustal Consensus	:	:	:	:	.*.*	:	:

		370	380	390	400	410	420
CYP102A1_BM3	EDTVLGGEYP	LEKQDELMVL	IPQLHRDKTI	WGDDVEEFRP	ER-----	FEN	PSAIPQHAFK
CYP102A2_BS	EDTVLGGKFP	ITNDRISVL	IPQLHRDRDA	WGDAEEFRP	ER-----	FEH	QDQVPHHAYK
CPP102A3_BS	EDTVLGGEYP	ISKQQPVTVL	IPKLHRDQNA	WGDAEDFRP	ER-----	FED	PSSIPHHAYK
P450foxy	DDTVLGGKYL	VKKEIVTAL	LSRGHVDPVV	YGNDAKFIIP	ERMLDDEFAR		LNKEYPNCWK
CYP109C1	RAVEVHG-QV	IPAGKLVLPV	TGSANRDPHQ	FH-DPGRFDI	GR-----		DPNPHV
CYP109C2	RDVPMHG-QV	IPAGKAVLAM	TGSANRDSQQ	FR-DPDRFDI	AR-----		DPNPHI
CYP109D1	RPVDVSG-VM	IPAGHLVDVV	FGANRDPAV	FE-EPDAERL	DR-----		PAEHL
CYP105D5_s.coeli	EDIDVDG-QT	IRAGDGVVFS	TSVINRDES	YP-EPDADW	HR-----		PARHHV
P450BioI	EDIDICG-VT	IRQGEQVYLL	LGAANRDESI	FT-NPDVFDI	TR-----		SPNPHL
CYP152A1_BS	KDFVWNN-CE	FKKCTSVLLD	LYGTNHDPRL	WD-HPDEFRP	ER-----	FA	ERENLFDMI
CYP152A2_C.acet	KDFLWNN-CE	FKKEMLVLLD	LYGTNHDSRI	WQ-KPYEFTP	DR-----	FR	SYKGNLDFDI
Clustal Consensus	:	:	.*	:	:	.*	:

		430	440	450	460	470	480
CYP102A1_BM3	PFGNGQ----	RACIGQQFAL	HEATLVLGMM	LKHFDL-EDH	TNYELDIKET	LTLKPEGFVV	
CYP102A2_BS	PFGNGQ----	RACIGMQFAL	HEATLVLGMI	LKYFTL-IDH	ENYELDIKQT	LTLKPGDFHI	
CPP102A3_BS	PFGNGQ----	RACIGMQFAL	QBATMVLGLV	LKHFEI-INH	TGYELKIKEA	LTIKPDDFKI	
P450foxy	PFGNGK----	RACIGRPFAP	QESLLAMVVL	FQNFNFTMTD	PNYALEIKQT	LTIKPDHFI	
CYP109C1	AFGHCI----	HFCIGAALAR	LEARIALPDL	LCRLKG----	-----LRRASD	EPWEPK-AI	
CYP109C2	AFGHCI----	HFCIGAPLSR	LEARVALADL	VARLGG----	-----LERASD	APWEPK-AI	
CYP109D1	AFGOCT----	HFCIGAALAR	MEARIALNAL	LD CYES----	-----ITPGEA	PPLRQTR-AI	
CYP105D5_s.coeli	AFGFGI----	HQCIGQNLAR	AELIATLES	FDRLPT----	-----LRLAAP	ADEIPFKPGD	
P450BioI	SFGHGH----	HVCLGSSLAR	LEAQIAINTL	LQRMPS----	-----LNLADF	EWR--YRPLF	
CYP152A1_BS	PQGGCHAEKG	HRCBEGGITI	EVMKASLDFL	VHQIEYDVPE	QSLHYSLARM	PSLPESGIVM	
CYP152A2_C.acet	PQGGCDPSST	HRCBEGGITL	EIMKTSLDFL	STKIDFTVPD	QDLSYLSKI	PTLPKSGFTI	
Clustal Consensus	. * *	: * * :	:	:	:	:	:

```

CYP102A1_BM3      KAKSKK--IP LGGIPSPSTE QSAKKVRKKA ENAH----- --NTPLLLVLY GSNMGTAEGT
CYP102A2_BS       SVQSRH--QE AIHADVQAAE KAAPDEQKEK TEAKGASVIG LNNRP LLVLY GSDTGTAEGV
CPP102A3_BS       TVKPRK--TA AINVQRKEQA DIKAETKPKK TKPK----- -HGTP LLVLF GSNLGTAEGL
P450foxy          NATLRHGMP TELEHVLGN  GATSSSTHNI KAAANLDAKA GSGKPM AIFY GSNSGTCEAL
CYP109C1          IVHGPA----- ----- ----- ----- --RLP IRFEP GRRAAVHAAG
CYP109C2          HVHGPA----- ----- ----- ----- --RLP IRFAP GPRLGG-----
CYP109D1          MPLGFE----- ----- ----- ----- --SLP LVLRR SRATA-----
CYP105D5_s.coeli TIQGM----- ----- ----- ----- --ELP VAW-- -----
P450BioI          GFRAL----- ----- ----- ----- --ELP VTFE-----
CYP152A1_BS       SGIRRK----- ----- ----- ----- -----
CYP152A2_C.acet  DNINLKL----- ----- ----- ----- -----
Clustal Consensus

```

```

                    550      560      570      580      590      600
CYP102A1_BM3      ARDLADIAMS KGFAPQVA-T LD SHAGNLPR EGAVLIVTAS YNGHPPDNAK QFVDWLDQAS
CYP102A2_BS       ARELADTASL HGVRTKTA-P LND RIGKLPK EGAVVIVTSS YNGKPPSNAG QFVQWLQELK
CPP102A3_BS       AGELAAQGRQ MGFETAETA-P LDD YIGKLPK EGAVVIVTAS YNGAPPDNAA GFVEWLKLEL
P450foxy          ANRLASDAPS HGFSATTVGP LDQAKQNLPE DRP VVIVTAS YEQPPSNAA HFIKWMEDLD
CYP109C1          AHTPCPG----- ----- ----- ----- -----
CYP109C2          ----APG----- ----- ----- ----- -----
CYP109D1          -----
CYP105D5_s.coeli -----
P450BioI          -----
CYP152A1_BS       -----
CYP152A2_C.acet  -----
Clustal Consensus

```

```

                    610      620      630      640      650      660
CYP102A1_BM3      ADEVKGVRYG VFGCGDKNWA TTYQKVPAFI DETLAAKGAE NIADRGEADA SD-DFEGTYE
CYP102A2_BS       PGELEGVHYA VFGCGDHNWA STYQYVPRFI DEQLAEKGAT RFSARGEADV SG-DFEGQLD
CPP102A3_BS       EGQLKGVSYA VFGCGNRSWA STYQRIPRLI DDMKAKGAS RLTAIGEDA AD-DFESHRE
P450foxy          GNDMEKVSYA VFACGHHDWV ETFHRIPKLV DSTLEKRGGT RLVPMSADA ATSDMFSDFE
CYP109C1          -----
CYP109C2          -----
CYP109D1          -----
CYP105D5_s.coeli -----
P450BioI          -----
CYP152A1_BS       -----
CYP152A2_C.acet  -----
Clustal Consensus

```

```

                    670      680      690      700      710      720
CYP102A1_BM3      EWREH-MWSD VAAVFNLDI- ENSEDNKSTL SLQFVDSAAD MPLAKMHGAF STNVVASKEL
CYP102A2_BS       EWKKS-MWAD AIKAFGLELN ENADKERSTL SLQFVRGLGE SPLARSYEAS HASIAENREL
CPP102A3_BS       SWENR-FWKE TMDAFDINEI AQKED-RPSL SITFLSEATE TPVAKAYGAF EGIVLENREL
P450foxy          AWEDIVLWPG LKEKYKISDE ESGGQ----- KLLLEVSTP RKTSLRQDVE EALVVAEKTL
CYP109C1          -----
CYP109C2          -----
CYP109D1          -----
CYP105D5_s.coeli -----
P450BioI          -----
CYP152A1_BS       -----
CYP152A2_C.acet  -----
Clustal Consensus

```


	130	140	150	160	170	180	
tdtD	TSG-----	--FIAEGLGS	F-----	FDG	LILITAMDGEA	HKNIRSLLOP	VFMPETVNIR-
ditQ	TSG-----	--FIAEGLGA	F-----	FDG	LILITAMDGDA	HKNIRNLLQP	VFMPETVNIR-
CYP226A1	TSG-----	--FIAEGLGS	F-----	FDG	LILITGLDGEA	HRRARALLQP	VFLPEVVNIR-
CYP260A1	STR-----	--MFQAGI-	-----	LNG	-GDAAMQGDE	HARMRRVYNM	FFLPRAVSQ-
CYP260B1	STR-----	--TYDTGI-	-----	MKG	-ALVTLGGEA	HIRMRRLFNA	VLSPRVISR-
CYP124E1	CSGRGLFIED	-LPPGDMRD-	-----	NPD	-VMIMDDPPR	HARFRALVSK	GFTPRVIQR-
CYP265A1	NMGTDVLRQ	GVTEGLVRR-	-----	MYE	GFILLSTEGAA	HARLRGLLKK	GFTPDVAVDA-
CYP267A1	AKRSSAFVTK	LPDEVQRLE	PLRRN--	LA	SWALLDDPPE	HIRIRSLINK	AFVPRLVEG-
CYP267B1	RD-----	FRK	LPDEVRRRYF	PLSDRTTFMD	QHMLDADPPD	HIRLRRAIVQR	AFSPRRMEG-
CYP266A1	ADRYVALADT	VPPPEQKEMNS	YIVKS--	LS	MFMLNVENPT	HFRLRNLNFR	SFTPKSIAA-
CYP109C1	SS-----	-AVAPPTGKA	P-----	D	WIVFSD-PPR	HTKLRSLVLR	AFTPRSIAG-
CYP109C2	SS-----	-VVSPPGTRT	A-----	E	WLIFSD-PPR	HTKLRALITR	AFTPRAVAG-
CYP109D1	SS-----	-NVTDKIRVL	P-----	R	ITLHDDPPR	HTHLRRLVSR	SFTPRRTAE-
CYP264A1	SSQGFRAAWQ	PAWVG-HNPL	A-----	S	SILAMD-GPD	HARLRGLVSR	AFGPAIAR-
CYP264B1	SSAGLRMATE	PPYLRRQNPL	S-----	G	SMLAD-PPR	HGQLRSISSR	AFTANMVST-
CYP261A1	QRILRERPKE	FRRWSPLOEL	ID-----	EMGF	GGVFSAEQDQ	WTRQRRLVMS	AFTAGQLRES
CYP261B1	RDVLQRRPAV	FRRVRAIEAV	SR----	EADM	VGLFVSADGDA	WRKORQLIHP	TFHPRHVEGF
CYP262A1	QYVLHDNWRN	FGKEGGMWKP	IG----	RLLG	NGLVTAGGDE	WLRNRRRMQP	LFSSRQLAGL
CYP262B1	QHVLRLDNARN	YRKGGAIWES	VR----	SLIG	NGLPTSEGDL	WLRQRRLVQP	QFHRLDLAAM
CYP263A1	AHVLRNARN	YPK-GKRYHE	LA----	PVLG	WGLVNSEGEL	WRRQRHVIQP	QFNHANTLKF
CYP110H1	QDVMTPAPPET	FLPFATRAIA	P-----	LVGE	HSLLMLSDER	HRRERKLLTP	PFHGDRMRAY
CYP110J1	RAVYTDADPE	FDVWGVQLTE	P-----	VFGT	SSVVVTAGAR	HRRDRRLVLP	PFSAGAMRGY
CYP259A1	LDLLRAP--E	VVMDDPEH-	IK----	GIVG	RSMAAQDGPL	HRHMRSAINP	SFSPQGF---
CYP117B1	FSLFKSK--A	VDSSYLTEGG	VG----	HLFG	QTLMAHDCAS	HRRRLRTAMNG	PFSPKGLD--

Clustal Consensus

	190	200	210	220	230	240
tdtD	--WKETKIDR	VIREEYLRPM	VASKRADIME	FALYFPIRVI	YSLIG----	----FPEDRP
ditQ	--WKETKIDR	VIREEYIKPL	VAKSADLMQ	FGLYFPIRVI	YALIG----	----FPEDKP
CYP226A1	--WRESKMEP	IVRNEFIGPM	VPQRADLMH	FGLHFPRLTI	YSLIG----	----FADDRP
CYP260A1	--YEEFVVRP	ISEQVVDRLA	GKPRVDLLED	FAMELPRRVI	GELFG----	----FPAEKL
CYP260B1	--YEEATVTP	VARRVVERLV	RKERAELFDD	FAISMPMGVT	SALFG----	----LPEERI
CYP124E1	--LES-HVRE	LVTRLIDDAC	ERGGCDFASD	IAGKLPISVI	LEVIQ----	----VREDQD
CYP265A1	--LRP-RMRE	RMHALLDEIG	DRTEFDFVTA	VAERFPQQVI	IELLG----	----LPSFE
CYP267A1	--LRS-RVET	LVNELLDAVA	PAGRMDVLRD	LGDLLPLLLVI	GEVLG----	----VPAEDR
CYP267B1	--LRP-RIQE	IADGLIDAVI	DRRRMELIAD	FAPPLPTAVI	AELLG----	----LPEEDR
CYP266A1	--MRP-SAHA	VVNELLDAVQ	PRGHMDVVAD	LAYPMPIKEI	CGILG----	----MPVEDM
CYP109C1	--LEP-RIRE	LSRDLLDPWI	ERGEMDLAAD	YAGPLPTMVI	AEMLG----	----VPEEDR
CYP109C2	--LEP-RIRR	LSRELLDRTI	ESGQMDVAED	YAVPLPLLVI	AEMLG----	----APAADQ
CYP109D1	--LEP-WIGR	LAASLEATG	D-GPSDLMGA	YAMPLPMVI	ATLLG----	----IPAERY
CYP264A1	--IEQ-RARD	LCERLAGRLD	G--EVDFTIAA	AAAPLPFVVI	SELLG----	----LDHALE
CYP264B1	--LEH-HMRS	MAVRLTDDL	HRRVVEFISE	FASRAQVSVL	AKLIG----	----FDGGLG
CYP261A1	HGTLSTITRR	LRERWRASAA	RGAPVDARRD	LARYTVDDVT	AVAFGLDMNL	IERTDPLSD
CYP261B1	FPSIRDITGR	LRELWARAAD	ERAGMDVLGD	LMRYSVDVTS	SVAFGRDLNT	LERGADALQE
CYP262A1	VDRMFQVVEG	DLPRLLEEAR	AGAVVDMKE	MMQLTQRVIL	ATMFG--	VSI
CYP262B1	CDLMVQAIDD	GMAGFGAAAA	AGRFPVNAERE	LPHITMKVIL	NTMFG--	SGI
CYP263A1	VPITVEHMEA	VLRRWDAQPG	E-FERDINDD	MMDVTFGIAG	EAFFG--	AAL
CYP110H1	AAAMADTAAR	RLTEAAR---	-APRAVAQEL	TQAISLDVVI	RAVFG--	VEE
CYP110J1	GDAIAEISLD	VASRWRP---	-GRSFSMLAA	TQAIALDVIV	RVVFG--	VRG
CYP259A1	AKEAAAAIAS	AIAAKIDAWL	QQESFAVHPA	MRDLTLDIIL	RICGVP---	---
CYP117B1	AAEVGAIVAT	SVERKVRSWL	GRRDVQLLRE	TRELALEVME	KITGVE---	---

Clustal Consensus

	250	260	270	280	290	300
tdtD	EEIEQYAAWA	LAILAGPQVD	PEKAAAARGA	AMEAAQALYD	VVKVVVAQR	AEGATGDD--
ditQ	EEIEQYAAWA	LAILAGPQVD	PEKAAAARGA	AMHAAQALYD	VVKVVVAER	EAGAEQDD--
CYP226A1	EQVEQYAAWA	LAILAGPQVD	AQKAAIARKA	AMEAAEALYA	AIRSEVAVVR	AKGAEQED--
CYP260A1	HETDERVRAM	LRGLVRMH-D	PAAVAESQRA	YGETLGLITE	VVERESRD--	----TSDT--
CYP260B1	AENDALIRKM	IRSVVMPQ-D	PVVVAEGRSA	HAAMEAQLRE	IAEREVAH--	----PSDT--
CYP124E1	E-QMLDWTTR	FFG--ASDPA	YGVTPPEELN	--AVLHNMNA	YAHQLAEQRR	K--EPKDD--
CYP265A1	DDDFRRWCSD	IGY--IVGLE	VKKHLPRLE-	--AAFEGLRG	YIEPTIAARR	R--EPRGD--
CYP267A1	H-RLKGWSNA	LSGFLGAGRP	TLEIAGGAL-	--SAVAELED	YFRGVIAARR	Q--SPGND--
CYP267B1	G-RFRRWTKI	LL-----	APAK	DREFVERAQ-	--PVVEEFAA	YFRALADARR
CYP266A1	G-LIKQLSDD	VSUYIGSAGK	AAGCIPPAY-	--HAIVEFSK	LFRPLVEARR	K--EPKDD--
CYP109C1	A-RLLRFSEV	IVNLSHAISG	GEEAARAVSE	HAVVKEEMKV	YLAGLIEARR	R--APAEV--
CYP109C2	P-HFKRWSDA	ILLDSHTVSG	SEEAAARALDA	FTAVTAEMQA	YLRGLVEQRR	A--APEDD--
CYP109D1	V-QFRSWSES	VMSYSG-IPA	EERASRGK--	-----	AMVD	FFAAVELEARR
CYP264A1	P-HFKRWMD	LLSVTPEP-A	SAEHAARVR-	--ATIAELDR	YMADVIAARR	R--SPSDD--
CYP264B1	G-HFKRWATD	LVIVGVIPPE	DHARIAEVR-	--RTIDEMEQ	YMLGLLASRR	R--HLEND--
CYP261A1	QLETVFASLN	RRVFAPFPYV	RHVYKLPADRA	LDQALSEVRT	RMLDILRTR	AELDRDPARA
CYP261B1	HLQRFVFLGLN	RRMLAPLPYW	RYVGLPADRA	YDRALVAVKR	LILDLVADR	RIEIEGADRP
CYP262A1	VLLVAIQALN	ARMFYFMPD	RL--LPGERA	LRDAIARIDE	AILRLVRERR	RSKEERDD--
CYP262B1	SLRYALDYML	LGAAALRALPS	WLP-APGRRR	FERSAKAIDE	HVFRFIAQRR	AQPGRGGD--
CYP263A1	AFKYALSIAL	KRMYSLVNP	LSWLPSPHLR	FRRAMDRVHA	VIDGIIIDYQ	RGAGGEDN--
CYP110H1	VVAMDALTP	ALTFLLPQLP	ELGGLGPPYAR	FRRRVGELDA	LFRAQIETAR	AAPGD---
CYP110J1	VLGLIESLSP	SFMIIPALRR	DFAGFGPYAR	HKRAARALDA	LLFEEIR-AR	RAEGDASQ--
CYP259A1	IRAWSRAYRE	LFLGVYPIPG	RLPGMPAY-R	SDRARAWIDD	RIRKLLALAR	GGGAAGS---
CYP117B1	LPEWRHYDED	LMLVILSLPF	DVPGSPRR-R	GLRARAWLDE	RIGRILAGVR	ARGDAKG---

Clustal Consensus

	490	500	510	520	530	540
tdtD	IG---QFVAK	TEINCAINAI	LDLMPNIRLD	PDKPAPEIIG	---AQLRSPH	HVHVIWD---
ditQ	IG---QFVAK	TEINCAINAI	LDLLPNIRLD	PSKPAPQIVG	---AQLRSPH	EVHVIWD---
CYP226A1	IG---QFIAK	VELQVAVNAI	LDLLPNLRLD	PDRSPPKIVG	---AQLRSPD	AVHVVWD---
CYP260A1	VG---APLAR	MEARVGLQAL	LARFPGGLRAV	PEERPSFMYG	AKDSVAHGPD	KLPVLLH---
CYP260B1	VG---APLAR	MEARVGVSLI	LERFPPALRAD	PTVQPTFSTA	PRGAAAFSPD	QIPALLV---
CYP124E1	LG---LVILAR	LEARVAFEEL	LRRLPDLELA	GPA---VRL	RSN-WVSGVK	SMPVRCARPA
CYP265A1	LG---AYLAR	AELQEAVPIL	TRRLTEAALA	SPV---VVR	TAREGSPGPR	VIPLRCRVRP
CYP267A1	AG---SALAR	VEAQAAISTF	LRRFPPDAELS	PGP---LTW	RMNPGMRGVT	ALPIELGPQS
CYP267B1	LG---AMLAR	IEAATAFSTI	LRRLPRIELA	TSTR--DIVW	SEWPTIRGPA	AVPVVF----
CYP266A1	IG---SHLAR	MEGEVALGAL	LORMPNLRLA	TQE---VEW	RGNSRFRGLR	ALPVSF----
CYP109C1	IG---AALAR	LEARIALPDL	LORLKLRRRA	SDE---PWEP	RKAIIVHGPA	RLPIRFEPGR
CYP109C2	IG---APLSR	LEARVALADL	VARLGGLELA	SDA---PWEP	RRAIHVHGPA	RLPIRFAPGP
CYP109D1	IG---AALAR	MEARIALNAL	LDCEYESITPG	EAP---PLRQ	TRAIIMPLGFE	SLPLVLRRSR
CYP264A1	LG---AALAR	MEAKVALVVL	VERIIGEVTTRA	PG---EIPY	NRTLTVRGPV	SLPLRFAPA-
CYP264B1	LG---VFLAR	VQARIVLEEL	LRRCRIVLR	TD---RLEW	QAALNTRSPV	ALPIEIVPVS
CYP261A1	PG---RALSM	VESAMVGAMV	ARDFDVSLVD	PAR---PVKE	LLGFTMKPEG	-LFVRFPPRH
CYP261B1	PG---RAMAL	LECAVTGGMV	LRSFVHLAD	HTA---PVQE	RMAFAMQPEG	-LRIRFTKRA
CYP262A1	IG---NLFST	MEAQIVIAVL	LRRLRMRLVP	GHP---VSPQ	AVATLRPRHG	-LKMTLHAV-
CYP262B1	VG---KEFAL	MEGQLITARL	LQRYRISAVP	GRT---TRLH	FATTLRTSGG	-VWLRLERP
CYP263A1	LG---EFMGQ	LECKIMVAMA	LQRYRLRLLVP	GFD---PRCR	GFISLQPVNG	-MRMILKRRE
CYP110H1	IG---AAFAM	FEMKIVLGVA	LSAWEFRLLD	ERP---PRPV	RRNLTLSPSG	GVPLVL----
CYP110J1	LG---AAFAM	YEMKVALGTI	LSRYRLRLES	QAP---IHHV	RRGLTMGPSG	DVAMILEGER
CYP259A1	LGYNVACVEA	LQLQFILVRA	LRRAAGLVPHL	EGA---SPKH	VYFPTGHPFA	ATRVAFRRC-
CYP117B1	LGYHLAWMEI	VQFLVALGRE	LEPAGS--PRL	QGG---FPAS	RYLPLVLRPSG	GTRVRFDDG--
Clustal Consensus	*	.	:			

	550	560	570	580
tdtD	-----	-----	-----	-----
ditQ	-----	-----	-----	-----
CYP226A1	-----	-----	-----	-----
CYP260A1	-----	-----	-----	-----
CYP260B1	-----	-----	-----	-----
CYP124E1	R-----	-----	-----	-----
CYP265A1	R-----	-----	-----	-----
CYP267A1	SAS-----	-----	-----	-----
CYP267B1	-----	-----	-----	-----
CYP266A1	-----	-----	-----	-----
CYP109C1	RAAVHAAGAH	TPCPG---	-----	-----
CYP109C2	RLGG-----	--APG---	-----	-----
CYP109D1	ATA-----	-----	-----	-----
CYP264A1	-----	-----	-----	-----
CYP264B1	TTAARESPVV	QGIW----	-----	-----
CYP261A1	R-----	-----	-----	-----
CYP261B1	SA-----	-----	-----	-----
CYP262A1	-----	-----	-----	-----
CYP262B1	GR-----	-----	-----	-----
CYP263A1	PSSAVRAAPA	GREEPAGAEP	VVTEPVAAAM	SAHGCPYSAR A
CYP110H1	-RARGRSR--	LYAAA----	-----	-----
CYP110J1	ERARAEGRPA	AVQGSAPAEA	LRPAAPQPAG	RCPMGFS----
CYP259A1	-----	-----	-----	-----
CYP117B1	-----	-----	-----	-----
Clustal Consensus				

ACKNOWLEDGMENTS

My tenure at Department of Biochemistry, Saarland University has been wonderful and rewarding experience. I am forever indebted to these people for their contributions to my scientific carrier and personal growth over the years.

I wish to express my deepest gratitude to my supervisor Prof. Dr. habil. Rita Bernhardt for introducing me to the intriguing field of cytochrome P450. Her knowledge, understanding, encouragement and guidance have been of great value for me. I was not only having a very pleasant feeling during professional discussion concerning our research but also sharing my personal problems and sentiments too.

My respect and admiration for Prof. Dr. Rolf Müller is always there and become a source of inspiration. His cooperation and collaboration with myxobacterium project of *Sorangium cellulosum* was the beginning of my research carrier with this P450 world in Prof. Bernhardt's group.

I am thankful to the German Academic Exchange Service (Deutscher Akademischer Austausch Dienst) for awarding this full-funded PhD-fellowship and giving an opportunity to study in Germany. I express my special thanks to our secretary Ms. Gabi Schon for helping me in every administrative aspect.

I also wish to express my gratitude to Dr. Frank Hannemann for the valuable suggestions throughout my PhD duration. I had a wonderful time and many great remembrances with him. I am indebt to Dr. Norio Kagawa and much grateful for sharing his profound experiences of P450 expression during his stay in our group. I extend my gratitude to Dr. Vlada Urlacher, Institute of Technical Biochemistry, University of Stuttgart, for her research cooperation. I am grateful for her helps and supports during my stay there. I wish to thank Dr. Reinhard Kappl, Institute of Biophysics, Universitätsklinikum, Universität des Saarlandes, Homburg, for EPR measurement and Dr. Michael C. Hutter, Center for Bioinformatics, for modeling and docking experiments. I would like to thank Aurélie Môme, Laboratoire de Dynamique et Structure Moléculaire par Spectrométrie de Masse, Institut de Chimie, Starsbourg, France for the cooperation of mass measurements. My sincere thank also goes to Mr. Wolfgang Reinle for the purification of adrenodoxin and adrenodoxin reductase. My thanks also goes to Dr.

Michel Lisurek, Forschungsinstitut für Molekulare Pharmakologie, Berlin, for the cooperation of substrate screening.

I would like to thank Dipl. Biol. Britta Wilzewski and Dipl. Biol. Martin Wörner, not just for keeping work smoothly and helping out whatever needed, but also for helping me to keep my life in proper perspective. I am also thankful for the wonderful time that I shared with Dr. Anja Berwanger and Dipl. Chem. Berna Mersinli. I am very grateful for their help and sharing my personal feeling during my stay. I am also thankful to Nguyen family, Dr. Hoang and Mrs. Thoa, for sharing my feeling and helping me during my PhD work. I would also like to thank Dr. Andreas Bichet for suggestions and everlasting funny stuff during my stay that will always be in my memory. I am also thankful to Dipl. Biol. Kerstin Ewen for her help, cooperation and working together in same myxo-project. My thanks also go to Dr. Tarek Hakki, Dipl. Chem. Michael Ringle, Dipl. Biol. Michael Kleser, Dipl. Biol. Tin Ming Kwai, Dipl. Chem. Sabrina Bleif, M. Sc. in Biotech. Anna Hobler, M. Sc. Biol. Thuy Li Thy Bich, Dipl. Biol. Danella Rauf, Dipl. Biol. Bernd Lauer and Dipl. Biol. Simon Janocha and Dipl. Biol. Jens Neunzig for nice time and everlasting memories. I would also like to share my thanks to our technician Katharina Bompais and Dip. Biol. Antje Edien Plach. Large thanks also go to my various colleagues in other groups which have made my PhD enjoyable throughout these years.

My loving thank is always for my sweet wife Rajita for her encouragement, support and suggestions. Rajita! thank you very much, you make me able to withstand. I am indebt with my parents Hajuri and Keshav, and my beloved brothers Gagan and Raman. I am also thankful to my parents in law Raj Gopal and Pabita, and my dear brother in law Ritesh and Aashish. During this time, though I was far away from you, there was no distance in between. Sweet remembrance and everlasting memory was always in front and be ever! I love you a lot for your blessings and everlasting support!