

**Characterization of CYP264B1 and a terpene cyclase  
of a terpene biosynthesis gene cluster from the  
myxobacterium *Sorangium cellulosum* So ce56**

**Dissertation**

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

von

**Thuy Thi Bich Ly**

Saarbrücken

2011

Tag des Kolloquiums: 16.12.2011

Dekan: .....

Berichterstatter: Prof. Dr. Rita Berhardt

Prof. Dr. Rolf Müller

.....

Vorsitz: Prof. Dr. Uli Müller

Akad. Mitarbeiter: Dr. Gert-Wieland Kohring

---

**CONTENTS**

<b>CONTENTS</b> .....	<b>i</b>
<b>ABBREVIATIONS</b> .....	<b>v</b>
<b>AIM OF THE WORK</b> .....	<b>1</b>
<b>1. INTRODUCTION</b> .....	<b>3</b>
<b>1.1. Myxobacterium <i>Sorangium cellulosum</i> So ce56</b> .....	<b>3</b>
1.1.1. General information.....	3
1.1.2. A terpene biosynthesis gene cluster of So ce56.....	4
<b>1.2. Cytochromes P450</b> .....	<b>5</b>
1.2.1. General aspects.....	5
1.2.2. Catalytic mechanism.....	7
1.2.3. Electron transport systems.....	9
1.2.4. Structural features.....	10
1.2.5. Interactions of cytochromes P450 with their ligands.....	11
<b>1.3. Terpene cyclases and terpenoid biosynthesis</b> .....	<b>13</b>
1.3.1. Terpenoids.....	13
1.3.2. Terpenoid biosynthesis pathways.....	14
1.3.3. Terpene cyclases.....	16
1.3.3.1. <i>Catalytic mechanism</i> .....	17
1.3.3.2. <i>Structure features</i> .....	19
<b>2. MATERIALS AND METHODS</b> .....	<b>21</b>
<b>2.1. Materials</b> .....	<b>21</b>
2.1.1. Bacterial strains.....	21
2.1.2. Genomic DNA.....	21
2.1.3. Vectors.....	22
2.1.4. Primers.....	23
2.1.5. Proteins.....	23
2.1.6. Culture media.....	23

<b>2.2. Methods</b> .....	24
2.2.1. Plasmid construction.....	24
2.2.1.1. Construction of <i>pETG</i> for terpene cyclase <i>GeoA</i> expression.....	24
2.2.1.2. Construction of <i>pETC4AA</i> for <i>in vivo</i> conversion.....	24
2.2.1.3. Construction of <i>pETC4AAG</i> for terpene production.....	25
2.2.1.4. Construction of <i>pETGAA</i> for terpene production.....	25
2.2.1.5. Construction of <i>pAC-MvC4</i> for terpene production.....	26
2.2.2. Heterologous gene expression of CYP264B1 and <i>GeoA</i> .....	26
2.2.3. Purification of His <sub>6</sub> -CYP264B1 and His <sub>6</sub> - <i>GeoA</i> .....	26
2.2.4. Determination of P450s concentration.....	27
2.2.5. Determination of redox partners for CYP264B1.....	28
2.2.6. Substrate determination.....	28
2.2.6.1. Substrate determination for CYP264B1 and analysis of substrate binding	28
2.2.6.2. Substrate determination for <i>GeoA</i> .....	29
2.2.7. <i>In vitro</i> conversion.....	29
2.2.7.1. <i>In vitro</i> conversion with CYP264B1.....	29
2.2.7.2. <i>In vitro</i> conversion with <i>GeoA</i> .....	29
2.2.7.3. <i>In vitro</i> conversion with <i>GeoA</i> and CYP264B1.....	29
2.2.8. Kinetic analysis of CYP264B1.....	30
2.2.9. <i>In vivo</i> substrate conversion with CYP264B1.....	30
2.2.10. Terpenoid production in mevalonate pathway-engineered <i>E. coli</i> .....	31
2.2.10.1. Terpene production with <i>GeoA</i> .....	31
2.2.10.2. Terpene production with <i>GeoA</i> and CYP264B1.....	31
2.2.11. Product purification and identification.....	32
2.2.11.1. Silicagel column.....	32
2.2.11.2. Thin layer chromatography.....	32
2.2.11.3. High performance liquid chromatography.....	32
2.2.11.4. Gas chromatography-mass spectrometry.....	33
2.2.11.5. Nuclear magnetic resonance spectroscopy.....	33
2.2.12. Homology modeling and substrate docking into CYP264B1 active site....	33
<b>3. RESULTS</b> .....	<b>35</b>
<b>3.1. Expression and characterization of CYP264B1</b> .....	<b>35</b>

---

3.1.1. Protein expression and purification.....	35
3.1.2. Spectrophotometric characterization.....	36
3.1.3. Redox partners.....	37
3.1.4. Substrate screening.....	38
3.1.5. Substrate binding constant.....	40
3.1.6. Substrate conversion.....	42
3.1.6.1. Conversion of $\alpha$ -ionone.....	42
3.1.6.2. Conversion of $\beta$ -ionone.....	43
3.1.6.3. Conversion of (+)-nootkatone.....	44
3.1.6.4. Conversion of other sesquiterpenes.....	46
3.1.7. Enzyme kinetics.....	47
3.1.8. Homology modeling and substrate docking.....	49
3.1.9. <i>E. coli</i> whole cell conversion for CYP264B1 substrate.....	51
<b>3.2. Cloning, expression and characterization of GeoA.....</b>	<b>53</b>
3.2.1. Computer-based identification of GeoA.....	53
3.2.2. Cloning of the GeoA gene.....	55
3.2.3. Protein expression and purification.....	56
3.2.4. Substrate identification.....	57
3.2.5. Product identification.....	58
<b>3.3. Identification of products of the <i>TPSgc</i> encoding enzymes.....</b>	<b>60</b>
3.3.1. Product formation in the <i>in vitro</i> conversions.....	60
3.3.2. Comparison of sesquiterpene product X with valencene.....	63
3.3.3. Conversion of the products of GeoA by different CYPs of So ce56.....	64
<b>3.4. Investigation of terpenoid production in engineered <i>E. coli</i>.....</b>	<b>66</b>
3.4.1. <i>E. coli</i> strain and culture medium.....	66
3.4.2. Terpene production with GeoA.....	67
3.4.3. Terpene production with CYP264B1 and GeoA.....	69
<b>4. DISCUSSION.....</b>	<b>73</b>
<b>4.1. Characterization of CYP264B1.....</b>	<b>73</b>
<b>4.2. <i>E. coli</i> whole-cell biocatalyst for sesquiterpene conversion.....</b>	<b>78</b>
<b>4.3. Characterization of GeoA and identification of the <i>TPSgc</i>.....</b>	<b>80</b>

---

<b>4.4. Investigation of the terpene production in engineered <i>E. coli</i>.....</b>	<b>83</b>
<b>4.5. Outlook.....</b>	<b>88</b>
<b>5. SUMMARY / ZUSAMMENFASSUNG.....</b>	<b>90</b>
<b>6. REFERENCES.....</b>	<b>92</b>
<b>7. APPENDIX.....</b>	<b>110</b>
<b>STATEMENT / ERKLÄRUNG.....</b>	<b>118</b>
<b>ACKNOWLEDGMENT.....</b>	<b>119</b>

**ABBREVIATIONS**

AdR	Adrenodoxin reductase
Adx	Adrenodoxin
Amp	Ampicillin
APS	Ammonium persulfate
bp	Base pair
BSA	Bovine serum albumin
C	Celsius
CD	$\beta$ -hydropropyl-cyclodextrin
CDP-ME	Diphosphocytidyl-2-C-methylerythritol
Cm	Chloramphenicol
CO	Carbon monoxide
CoA	Coenzyme A
CYP	Cytochrome P450
Da	Dalton
DMAPP	Dimethylallyl diphosphate
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DPMVA	Mevalonate-5-pyrophosphate
<i>E. coli</i>	<i>Escherichia coli</i>
FAD	Flavine adenine dinucleotide
FMN	Flavine mononucleotide
FPP	Farnesyl pyrophosphate
GCMS	Gas chromatography - mass spectrometry
GGPP	Geranylgeranyl pyrophosphate
GPP	Geranyl pyrophosphate
h	Hour
HMBPP	4-hydroxy-3-methylbut-2-enyldiphosphate
HMG	3-hydroxy-3-methylglutaryl
HPLC	High performance liquid chromatography
IPP	Isopentenyl pyrophosphate
Kb	Kilobase

---

kDa	Kilodalton
$K_d$	Dissociation constant
Kan	Kanamycin
IPTG	Isopropyl- $\beta$ -D-thiogalactoside
L	Liter
M	Mevalonolactone
MecPP	2-C-methyl-D-erythritol
MEP	Mevalonate-independent pathway
min	Minute
MVA	Mevalonate-dependent pathway
NADPH	Nicotinamide adenine dinucleotide phosphate
nm	Nanometer
NMR	Nuclear magnetic resonance
OD	Optical density
P450	Cytochrome P450
PAGE	Polyacrylamid gel electrophoresis
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulfonyl fluoride
rpm	Round per minute
SDS	Sodium dodecylsulfate
TEMED	<i>N,N,N',N'</i> -tetramethylethylenediamine
TLC	Thin layer chromatography
TPSgc	Terpene biosynthesis gene cluster

## AIM OF THE WORK

Within the recent decades, increasing attention has been paid on the discovery and exploitation of natural products from plants, bacteria, and fungi. Natural products are often secondary metabolites, which have biological activities and pharmacologically interesting properties, and may act as antibiotics, insecticides, herbicides, antifungals, cytostatics, anticholesteremics, antiparasitics, or antitumor drugs (Reichenbach and Hofle 1993). Thus, a great number of secondary metabolites are widely commercially available as e.g. medicines, but also as flavor and fragrance ingredients, food additives, and agrochemicals (Vijaya Sree *et al.* 2010).

During the last 20 years, myxobacteria have been known to be one of the major sources of microbial secondary metabolites besides actinomycetes and fungi. Their ability to produce a variety of structurally unique compounds and metabolites with rare biological activities made them a promising source of novel pharmaceutically interesting compounds (Bode and Müller 2006; Wenzel and Müller 2007). More than 100 different basic compounds and approximately 500 structural variants have been isolated from these organisms (Reichenbach 2001; Gerth *et al.* 2003). About 7.500 different myxobacteria have been isolated, and numerous strains have been analyzed chemically (Gerth *et al.* 2003; Bode and Müller 2006; Wenzel and Müller 2007). Of these, members of the genus *Sorangium* produce nearly 50 % of the metabolites isolated from myxobacteria (Gerth *et al.* 2003; Schneiker *et al.* 2007). *Sorangium cellulosum* Soce 56, a myxobacterium with the largest genome yet discovered in bacteria, has been known as a producer of many novel low molecular weight natural products and valuable secondary metabolites such as chivosazol, etnangien and myxochelin (Schneiker *et al.* 2007; Wenzel and Müller 2007). In 2007, the genome of *Sorangium cellulosum* So ce56 was completely sequenced. Genomic data mining revealed 17 loci involved in secondary metabolism, including carotenoid and terpenoid biosynthesis (Schneiker *et al.* 2007). This opens up new possibilities to exploit secondary metabolites from this strain.

If a promising compound is found, it may be applied in a natural or chemically modified form (Reichenbach and Hofle 1993). The modification of natural products by changing their chemical structures, e.g. adding functional groups to enhance or alter biological properties, is very important to develop new compounds. For example, camptothecin – a

compound found in the stem wood of the Chinese ornamental tree *Camptotheca acuminata* can act as an interesting anticancer drug by inhibiting DNA topoisomerase I. However, its therapeutic activity is easily lost due to the intrinsic instabilities resulting from the rapid hydrolysis of the lactone ring in the body. Therefore, modification this compound to an analogue with adequately long biological life/activity in its active lactone form has been received a lot of attention from scientists (Srivastava *et al.* 2005). In chemical organic syntheses, regio- or stereo-selective oxidation is key reaction. However, these reactions are difficult to be achieved by chemical synthetic processes, in addition to cost and time inefficiency as well as hazardous impact (Curci *et al.* 2006; Bicas *et al.* 2009). In contrast, bio-catalysts, especially, cytochrome P450 mono-oxygenases, are considered as an alternative approach for regioselective modifications because of their ability to introduce atomic oxygen into allylic positions, double bonds or even into non-activated C–H bonds, and have thus a great potential for application in the production of natural molecules (van Beilen *et al.* 2003; Urlacher *et al.* 2004, Bernhardt 2006, Cankar 2011).

In the context of the So ce56 genome project, a cooperation between the Institute of Biochemistry and the Institute of Pharmaceutical Biotechnology at Saarland University was initiated to exploit the cytochromes P450 of this myxobacterial strain. During his PhD thesis, Dr. Yogan Khatri cloned and expressed all 21 P450s of So ce56 in *E. coli* and also characterized some of them, e.g CYP109D1 and CYP260A1 (Khatri PhD thesis 2009). Continuing a part of this project, my work is targeted at the characterization of CYP264B1, one of 21 P450s of So ce56. Since the gene coding for CYP264B1 is located in a cluster with gene coding for a terpene cyclase gene, characterization of CYP264B1 and the neighbouring terpene cyclase to elucidate the biological function of this cluster in the terpene biosynthetic pathway of myxobacteria So ce56 is very interesting for both basic research and industrial application.

By using modern molecular biology techniques and various chemical analysis methods, the presented work focuses on the characterization of two important enzymes, CYP264B1 and a terpene cyclase, of a gene cluster of *S. cellulosum* So ce56. The purposes of the work are: (i) understanding the characteristics of CYP264B1, including substrate specificity and catalytic functions, (ii) screening potential substrates of CYP264B1, (iii) characterizing the terpene cyclase to elucidate the role of the terpene biosynthesis gene cluster in the biosynthetic pathway of *S. cellulosum* So ce56.

## 1. INTRODUCTION

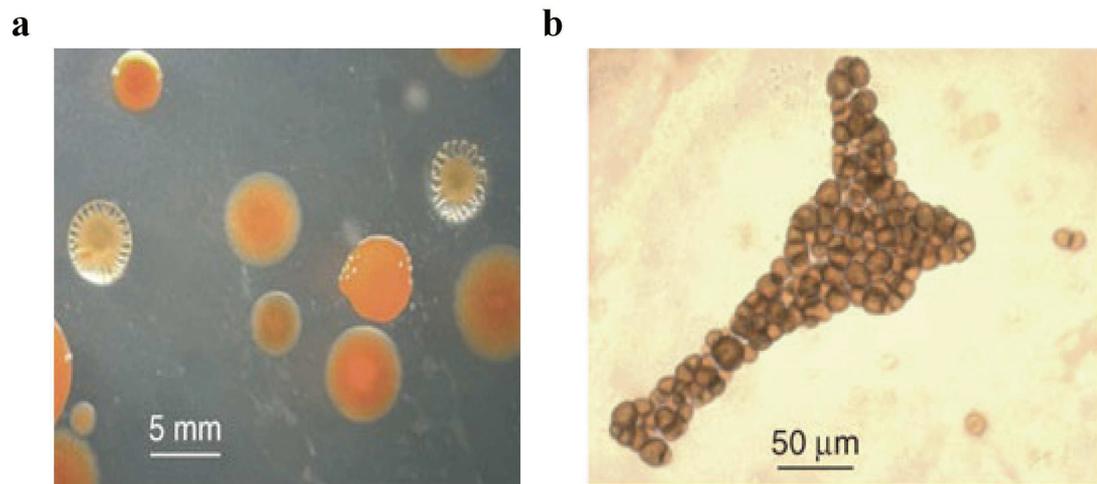
### 1.1. Myxobacterium *Sorangium cellulosum* So ce56

#### 1.1.1. General information

Myxobacteria are a large group of bacteria living predominantly in the soil. They do not have flagella and move on solid surfaces by gliding. Some physiological characteristics of So ce56 are shown in Figure 1.1. In comparison to other bacteria, genomes of myxobacteria are very large, about 9-10 million nucleotides. In 2007, the complete genome sequence of *Sorangium cellulosum* So ce56 was released, and it is the largest bacterial genome reported to date, with 13,033,799 base pairs. Genome analysis of So ce56 revealed a GC content of 71.38%, coding sequences of 86.45%, and 9,367 predicted protein-coding sequences (Schneiker *et al.* 2007).

Myxobacteria represent a very promising source for the discovery of novel classes of secondary metabolites. From these, the genus *Sorangium* is particularly valuable with approximately 50% of metabolites isolated from myxobacteria (Gerth *et al.* 2003; Schneiker *et al.* 2007). Genome sequencing has revealed that So ce56 possesses a high metabolic potential. This strain is known to produce natural products as chivosazole and myxochelin as well as the polyketide metabolite etnangien which is an inhibitor of bacterial and viral nucleic acid polymerases (Schneiker *et al.* 2007).

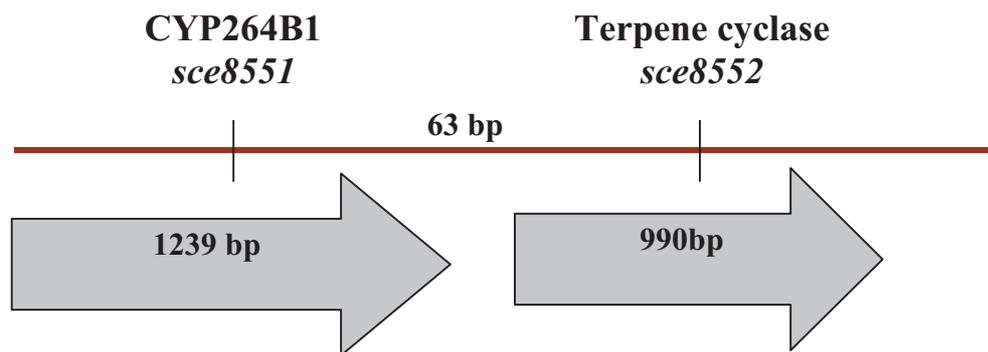
The So ce56 genome encodes a large number of biotechnologically important enzymes, including proteases, nitrilases, lipases, amidases, cellulases, and hydantoinases. In addition, numbers of genes encoding members of the short-chain dehydrogenase/reductase, S8 family peptidases and subtilisin-like enzymes were also found in the genome of this strain, providing a great potential for industrial production of chiral alcohols and additives used in laundry detergents (Schneiker *et al.* 2007). Especially, the So ce56 genome contains 21 putative cytochromes P450 which are known as versatile biocatalysts (Bernhardt 2006; Schneiker *et al.* 2007).



**Figure 1.1: Physiological characteristics of So ce56.** (a) *S. cellulosum* colonies, (d) Fruiting bodies (Schneiker *et al.* 2007)

### 1.1.2. A terpene biosynthesis gene cluster of So ce56

The genome of the myxobacterium So ce56 has three locus\_tags of terpene cyclases (sce1440, sce6369 and sce8552) which are potentially involved in the carotenoid and terpenoid biosynthesis (Schneiker *et al.* 2007). Locus sce1440 encodes a putative terpene cyclase composed of 771 amino acids located between locus sce1439 (encodes a major facilitator superfamily permease) and sce1441 (encodes a nucleotide-binding protein) with a distance of 528 and 21 bp, respectively. This putative terpene cyclase has substantial similarity to genes in other bacteria that are involved in the biosynthesis of geosmin, a sesquiterpene-derived natural product that gives soil its characteristic smell (Dickschat *et al.* 2005; Jiang *et al.* 2006; Schneiker *et al.* 2007). Locus sce6369 encodes a predicted pentalenene synthase composed of 315 amino acids with major domains for terpene cyclases and located between two hypothetical proteins with a distance of 614 and 59 bp. Locus sce8552 is a putative terpene cyclase (gene name: geoA). It is located at 63 bp downstream of the gene coding for a cytochrome P450 - CYP264B1 (gene name: CypA8) and 81 bp upstream of a protein kinase (NCBI). The distance of 63 base pairs between CYP264B1 and GeoA implies that they are likely arranged in an operon and together involved in the biosynthesis of terpenoids in So ce56 (Khatri 2011). CYP264B1 and the terpene cyclase GeoA are therefore regarded as a terpene biosynthesis gene cluster (*TPSgc*) of So ce56 (Figure 1.2).



**Figure 1.2: Organization of genes coding for CYP264B1 and terpene cyclase GeoA in the *TPSgc* of the myxobacterium *Sorangium cellulosum* So ce56**

## 1.2. Cytochromes P450

### 1.2.1. General aspects

Cytochromes P450 belong to the superfamily of heme-containing monooxygenase enzymes (Ruckpaul et al. 1989; Bernhardt 2006). They are often abbreviated as CYP or P450. Their name is based on their character as hemoproteins (cytochrome), pigment (P) and the typical absorption band at 450 nm of their reduced carbon-monoxide bound form (Omura and Sato 1964). This unusual spectral property is induced by a cysteine thiolate group (Ichikawa 1967; Murakami 1967), forming the fifth ligand of the heme iron (Hannemann et al. 2006). The ability of reduced, CO-bound P450 to produce an absorption peak at 450 nm is used for the estimation of the P450 concentration (Omura and Sato 1964). The members of the P450 superfamily are classified according to the recommendation of a nomenclature committee (Nelson *et al.* 1996), based on amino acid identity, phylogenetic criteria and gene organization (Werck-Reichhart and Feyereisen 2000). They are designated as CYP, followed by an Arabic number for the gene family (with more than 40% amino acid identity), a capital letter indicating the subfamily (with more than 55% amino acid identity), and another numeral for the individual gene, e.g. CYP106A2, CYP264B1, CYP260A1.

Cytochromes P450 are found in all domains of life, from eukaryotes to prokaryotes and even virus (Lamb *et al.* 2009). Plants have the highest number of P450s, human has 57 P450s, while *Escherichia coli* has none. In bacteria, P450s are involved in catabolism of compounds used as energy or carbon sources (Gunsalus *et al.* 1978; Schiffler *et al.*

2003), detoxification of xenobiotics (Taylor *et al.* 1999), metabolism of fatty acids and biosynthesis of antibiotics (Werck-Reichhart *et al.* 2000; Bernhardt 2006). In yeast and fungi, they are involved in the synthesis of membrane sterols and mycotoxins, detoxification of phytoalexins, and metabolism of lipid carbon sources (Kalb *et al.* 1987; Werck-Reichhart *et al.* 2000). In insects, P450s are required for juvenile hormone synthesis (Helvig *et al.* 2004), catabolism of ecdysteroids and detoxification of foreign chemicals (Feyereisen 1999). In plants, P450s play an important role in the synthesis of cutin and lignin barriers and of defense substances, UV protectants, flower pigments, plant hormones, defensive compounds (Kahn *et al.* 2000, Schuler and Werck-Reichhart 2003). In mammals, they take part in the production of signaling molecules, steroid hormones and retinoic acid (Hasler 1999), are involved in the biosynthesis of steroid hormones, vitamin D, fatty acids and drug metabolism (Guengerich *et al.* 2005).

P450s catalyze more than 20 different reactions summarized by Dawson and coworkers (Sono *et al.* 1996) (Table. 1.1) and some more unusual reactions were mentioned by Guengerich and coworkers (Guengerich *et al.* 2001). The most common reaction catalyzed by P450s is a monooxygenase reaction in which one atom of oxygen is inserted into an organic substrate (RH) while the other oxygen atom is reduced to water (Mansuy *et al.* 1998):



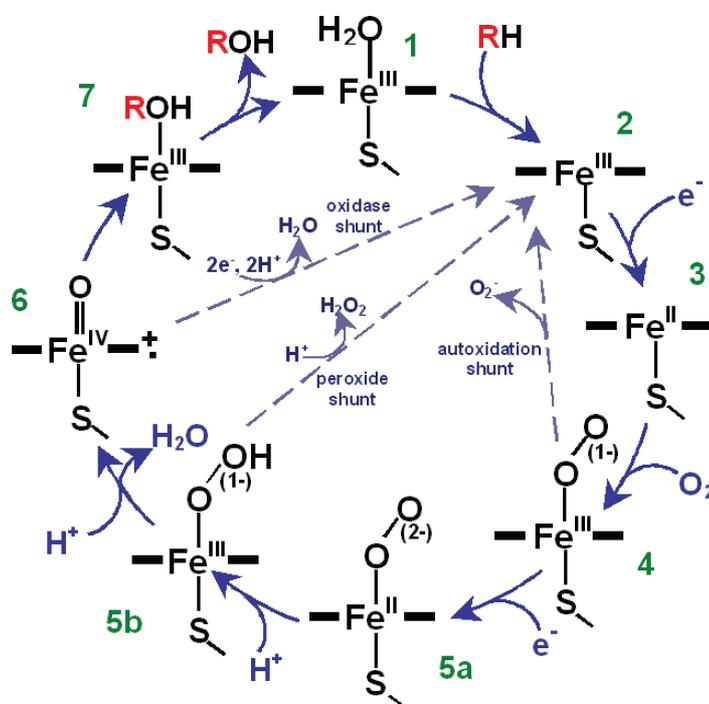
**Table 1.1: Reactions catalyzed by cytochromes P450, according to (Sono *et al.* 1996)**

Reactions	
Hydrocarbon hydroxylation	Oxidative dehalogenation
Alkene epoxidation	Alcohol and aldehyde oxidations
Alkyne oxygenation	Dehydrogenation
Arene epoxidation	Dehydratations
Aromatic hydroxylation	Reductive dehalogenation
N-Dealkylation	N-Oxide reduction
S- Dealkylation	Epoxide reduction
O- Dealkylation	Reductive $\beta$ -scission of alkyl peroxides
N-Hydroxylation	NO reduction
N-Oxidation	Isomerizations
S-Oxidation	Oxidative C-C bond cleavage
Oxidative deamination	

Due to their capability of performing various reactions on a variety of substrates, cytochromes P450 are considered as versatile biocatalysts and have great application potential (Bernhardt 2006; Urlacher and Eiben 2006; Urlacher and Girhard 2011). For example, a P450 monooxygenase has been used for the biotransformation of steroids to hydrocortisone with *Curvularia sp* at commercial scale (Petzoldt et al. <http://www.schering.de>); they were applied for the production of transgenic plants such as carnations and blue roses (Holton *et al.* 1993; Ogata *et al.* 2005) and for the production of vitamin D analogs (Sakaki *et al.* 1999). Nowadays, with the possibilities of genetic engineering techniques and various mutagenesis methods, the diversity of P450-catalyzed reactions becomes more and more available and applicable to biotechnological use (Bernhardt 2006).

### 1.2.2. Catalytic mechanism

A common catalytic cycle of the cytochromes P450 was proposed in 1968 (Gunsalus *et al.* 1975; Griffin 1979) and then updated by Denisov *et al.* (Denisov *et al.* 2005). This cycle includes substrate binding, reduction of the ferric heme, binding of molecular oxygen, activation of oxygen, insertion of an iron-bound oxygen into the substrate, and product release (Figure 1.3).



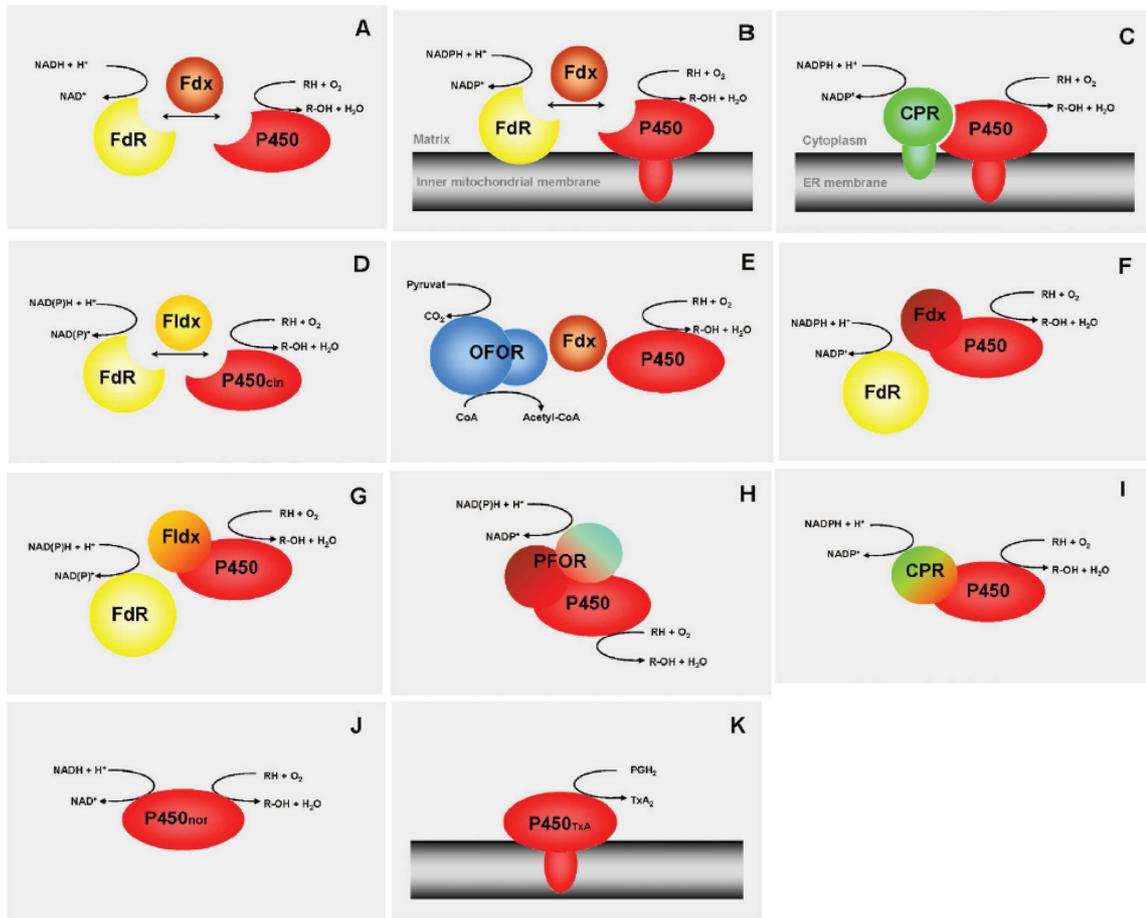
**Figure 1.3:** The catalytic cycle of cytochrome P450 (Denisov *et al.* 2005)

According to Gunsalus *et al.* (Gunsalus *et al.* 1975), Griffin (Griffin 1979) and Denisov *et al.* (Denisov *et al.* 2005), the catalytic cycle of the cytochromes P450 can be described as follows: **(1)** P450 in the substrate-free form: The ferric heme iron Fe (III) is in the low-spin state, and the sixth coordination position opposite the cysteinate is occupied by a water molecule as the distal ligand. **(2)** Substrate binding: Substrate binds to the enzyme active site, mostly by hydrophobic interaction in the vicinity of the heme and displaces the water molecule. This induces a low- to high-spin shift of the heme iron and an increase of the reduction potential of the enzyme heme, which favors the electron uptake. The five-coordinate iron atom is out of the porphyrin plane. **(3)** The first reduction:  $\text{Fe}^{3+}$  ion is reduced by an electron transferred from NAD(P)H via an electron transfer chain, results in a high-spin ferrous state Fe (II). **(4)** Oxygen binding: Molecular oxygen binds covalently to the distal axial coordination position of the heme iron, forming a superoxide complex  $\text{Fe}^{2+}\text{-OO}^-$  then undergoes a slow conversion to a more stable complex  $\text{Fe}^{3+}\text{-OO}^-$ , resulting in “in plane” coordination which is the last relatively stable intermediate in this cycle. **(5a)** Second reduction: One-electron reduction of this complex by insertion of the second electron to a ferric peroxo state  $\text{Fe}^{3+}\text{-OO}^{2-}$ , which is easily protonated to form the hydroperoxoferric  $\text{Fe}^{3+}\text{-OOH}$ -intermediate **(5b)**. **(6)** Oxygen cleavage: After the formation of an unstable transient  $\text{Fe-OOH}_2$  from the  $\text{Fe}^{3+}\text{-OOH}$ - complex, a heterolytic scission of the O-O bond occurs, following the release of the water molecule and the generation of the ferryl-oxene ( $\text{O} = \text{FeIV}$ ) activated oxygen intermediate. **(7)** Product formation and release: The active ferryl-oxene inserts an oxygen atom into the substrate forming a hydroxylated form of the substrate. The monooxygenated product can be released from the active site of the enzyme, which returns to its initial state. In addition, there are three abortive branch points, which can possibly occur under physiological conditions: (i) auto-oxidation of the oxy-ferrous enzyme (4) together with a release of a superoxide anion and return of the enzyme to its resting state (2), (ii) a peroxide shunt, where the coordinated peroxide or hydroperoxide anion (5a, b) dissociates from the iron forming hydrogen peroxide, thus completing the unproductive (in terms of substrate turnover) two electron reduction of oxygen, and (iii) an oxidase shunt at the stage of the ferryl-oxo intermediate (6) where oxygen is reduced to water instead of being introduced into the substrate (Denisov *et al.* 2005).

### 1.2.3. Electron transport systems

Cytochromes P450 are external monooxygenases, since they utilize electrons from an external donor to activate molecular oxygen for substrate conversion. For almost all P450s, the electrons cannot be transferred directly from NAD(P)H to the P450 but are mediated via an electron transfer chain (ETC). The components participating in an ETC are called redox partners. As the cytochrome P450 enzyme family is widespread in nature, their catalytic functions and the composition of their electron transfer chains are variable. Based on the types of redox proteins participating in electron transfer, Hannemann *et al.* (Hannemann *et al.* 2007) classified known P450s into 10 classes as shown in Figure 1.4 and summarized as follows:

Class I P450s includes most bacterial cytochromes as well as the mitochondrial P450s from eukaryotes. In this class, the electrons (reduction equivalents) are transferred from NAD(P)H to a FAD-containing reductase (FdR). Then, they are transferred to a ferredoxin (Fdx) which in turn reduces the P450s. In class II P450s, the ETC is composed of the cytochrome P450 and the NADPH cytochrome P450 reductase (CPR) located in the endoplasmic reticulum (ER). The CPR contains the prosthetic groups FAD and FMN which transfer electrons from NADPH to the P450. Class III P450s has been typically found in bacteria. The system consists of FdR, Flavodoxin (Fldx) containing FMN and the novel cytochrome P450cin (discovered from *Citrobacter braakii*). The Class IV P450s system includes Fdx and its corresponding partner, 2-oxoacid ferredoxin oxidoreductase (OFOR). Such a system has been found in *Sulfolobus solfataricus*. Class V P450s system contains putative NAD(P)H-dependent reductase (FdR) and a cytochrome P450-ferredoxin-fusion protein (Fdx-P450). The class VI P450s system is composed of a putative NAD(P)H-dependent flavoprotein reductase (FdR) and a flavodoxin-P450-fusion protein (Fldx-P450). Class VII P450s constitutes a fusion of the phthalate-family oxygenase with P450 (PFOR-P450). Class VIII P450s contains P450s fused to their eukaryotic-like diflavin reductase partner (CPR-P450). Class IX P450s comprises nitric oxide reductase (P450nor) using NADH as an electron donor. P450nor catalyzes dependent of NADH and independent of other electron transfer proteins. Class X P450s consists of independent redox partner P450s. This P450 group includes allene oxide synthase, fatty acid hydroperoxide lyase, divinyl ether synthase, prostacyclin synthase and thromboxane synthase (Hannemann *et al.* 2007).

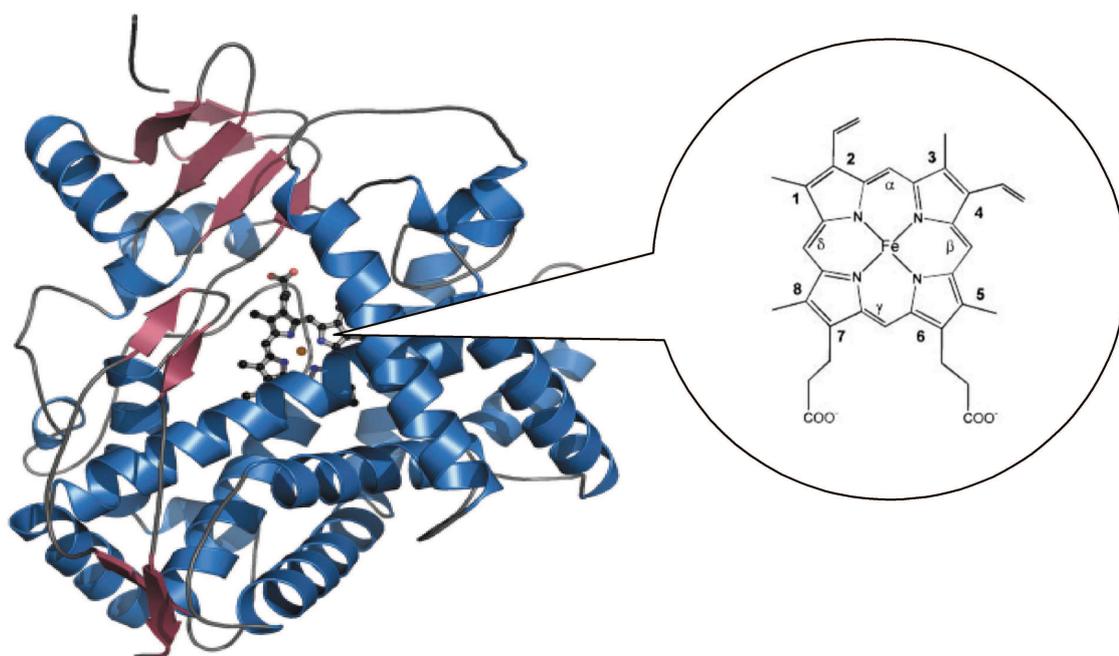


**Figure 1.4: Schematic organization of different cytochrome P450 systems.** (A) Class I, bacterial system; (B) class I, mitochondrial system; (C) class II, microsomal system; (D) class III, bacterial system; (E) class IV, bacterial thermophilic system; (F) class V, bacterial [Fdx]–[P450] fusion system; (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 P450<sub>nor</sub>; (K) class X, independent eukaryotic system (Hannemann *et al.* 2007)

#### 1.2.4. Structural features

Despite of sequence identity across the gene super family being less than 20%, cytochromes P450 share a common overall fold and topology (Hasemann 1995). In general, there are approximately thirteen  $\alpha$ -helices and four  $\beta$ -strands in P450 proteins (Peterson and Graham 1998). The conserved P450 structural core is formed by a four-helix bundle around the heme (three parallel helices D, L, I and one antiparallel helix E) (Presnell *et al.* 1989) (Figure 1.5). The prosthetic heme group (a tetrapyrrole heme b) is located between the distal I helix and proximal L helix and bound to the adjacent cystein heme- ligand loop containing the P450 signature amino acid sequence FxxGx(H/R)xCxG. The absolutely conserved cysteine is the proximal or “fifth” ligand

to the heme iron. Typically, the proximal cysteine forms two hydrogen bonds with neighboring backbone amides (Denisov *et al.* 2005). The long I helix forms a wall of the heme pocket and contains the signature amino acid sequence (A/G)-Gx(E/D)T which is centered at a kink in the middle of the helix. The I-helix also contains a highly conserved threonine that is positioned in the active site and is believed to be involved in catalysis (Imai *et al.* 1989; Martinis *et al.* 1989; Kimata *et al.* 1995). The 6 substrate recognition sites (SRS) proposed by Gotoh *et al.* (Gotoh *et al.* 1992) and are based on sequence alignments between members of the CYP2 family and P450cam. These protein regions are considered to be flexible and move upon substrate binding in an induced-fit mechanism to favor substrate binding and subsequently the catalytic reaction (Pylypenko and Schlichting 2004; Denisov *et al.* 2005). An example of cytochrome P450 structure is shown in Figure 1.5.



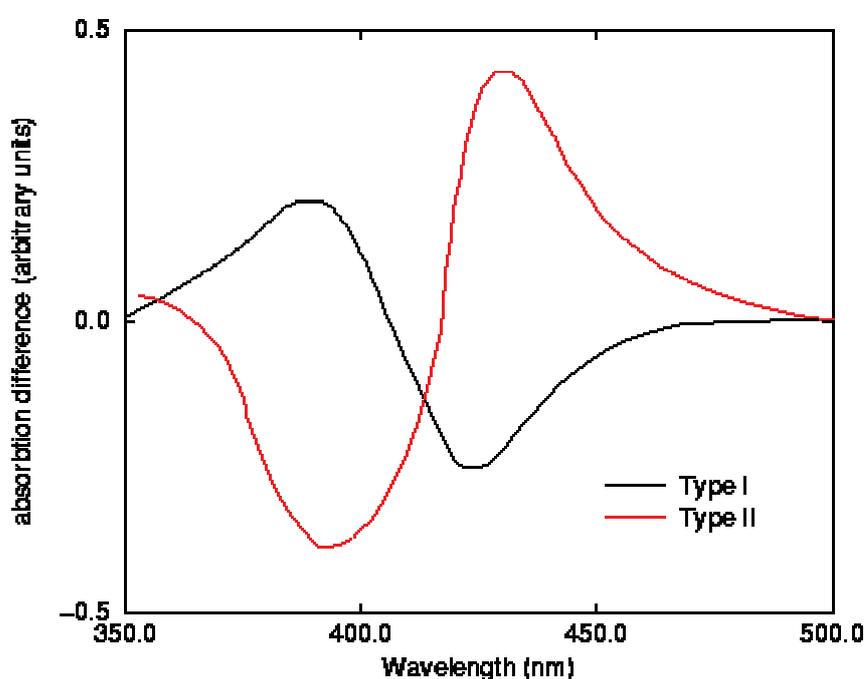
**Figure 1.5:** A ribbon diagram of the human cytochrome P450 2D6 structure.  $\alpha$ -helices are shown in blue,  $\beta$ -strands in mauve. The heme group is shown in ball and stick, and its structure is described in the callout

(<http://www.esrf.eu/Industry/UsersAndScience/Publications/Highlights/2006/SB/SB09>)

### 1.2.5. Interactions of cytochromes P450 with their ligands

Since the heme prosthetic group serves as the chromophore of a cytochrome P450, interactions between ligands and P450s occurring at the heme lead to a change in the

UV/vis spectra. In general, three types of binding spectra have been identified: type I, type II and type III (or reverse type I) (Schenkman *et al.* 1981). Type I shows a Soret absorption peak at 390 nm and a trough at 420 nm (Gibson and Skett 1994). This type is often observed in substrate bound forms. Type II is characterized by a decrease in absorption at around 390-405 nm and an increase at about 425-435 nm. Type III binding spectrum displays an increase of the 420 nm signal and a reduction of absorbance at 390 nm (Schenkman *et al.* 1981). The type II and III spectra are mostly observed upon the binding of an inhibitor instead of a substrate. However, some type II ligands can also act as substrates (Gigon *et al.* 1969; Kunze *et al.* 2006; Pearson *et al.* 2006). Type I and II are shown in Figure 1.6.



**Figure 1.6: The UV-difference spectra of cytochrome P450 in different form.** The type I is shown in black line, the type II is shown in red line (<http://www.tcm.phy.cam.ac.uk/~mds21/thesis/node50.html>)

P450s in the resting state exist in an equilibrium between low-spin (one unpaired electron,  $S = 1/2$ ) or high-spin iron (five unpaired electrons,  $S = 5/2$ ) (Denisov *et al.* 2007). The low-spin enzyme is hexacoordinated with water as the sixth axial ligand, and has an absorption maximum at around 418 nm (416-420 nm). The high-spin species is pentacoordinate with dissociated water and the iron is located slightly under the plane of the heme with a shortened iron–sulfur bond length (Conner *et al.* 2011). In this high spin state, the absorption maximum shifts to 390 nm (385-394 nm). The interaction of

ligands with the active site of CYPs leads to a spin-shift in the heme iron structure concomitant with a decrease in the reduction potential from -300 mV to a positive value (100 mV) (Guengerich *et al.* 1975; Williams-Smith and Cammack 1977; Guengerich 1983; Daff *et al.* 1997; Fantuzzi *et al.* 2004, Denisov *et al.* 2005). These ligand-induced changes can be explained as follows: In the resting state, the heme iron of P450s exists in the ferric  $\text{Fe}^{3+}$  form and the reduction potential of  $\text{Fe}^{3+}/\text{Fe}^{2+}$  is in the range of -400 to -170 mV (Guengerich 1975). In the absence of a ligand, a water molecule coordinates with the sixth distal axial coordination position of the heme and stabilizes the low spin state of the ferric iron. Upon ligand binding the water molecule is displaced by the ligand leading to a shift of the hexa-coordinated low spin state to the penta-coordinated high spin state. Consequently, the heme iron is transformed into a thermodynamically unstable state ( $\text{Fe}^{2+}$  form), and the redox potential is directed to a more positive value. The stronger the ability of the substrate to disturb the water interaction with the ferric heme, the more pronounced is the resulting positive shift of the redox potential. The equilibrium of the low spin and the high spin state is also influenced by ambient parameters such as temperature, pH as well as the presence of co-solvents (Denisov *et al.* 2005).

### 1.3. Terpene cyclases and terpenoid biosynthesis

#### 1.3.1. Terpenoids

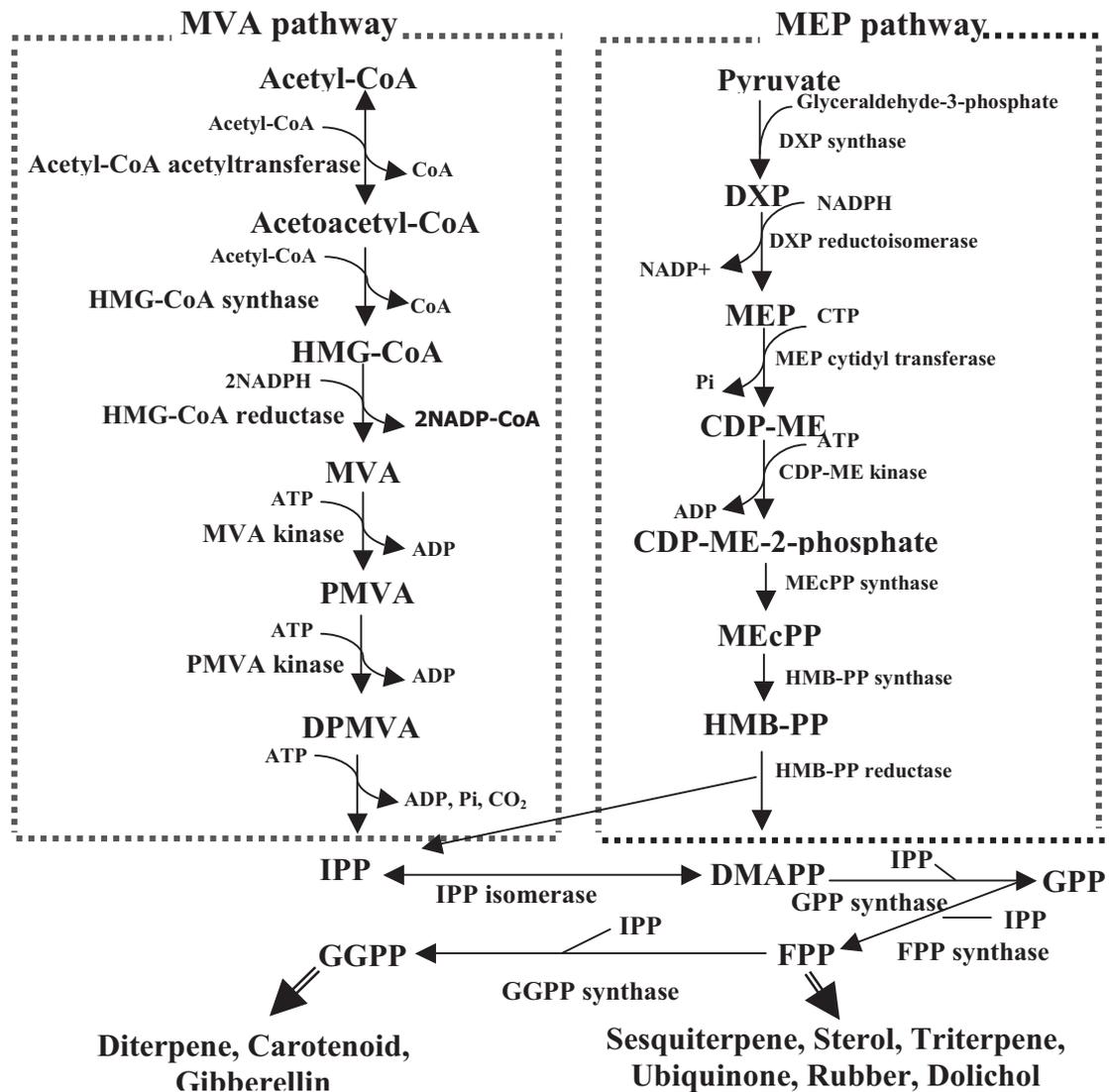
The terpenoids belong to the largest and most diverse class of natural products. More than 30000 individual terpenoids, the majority of which is isolated from plants, have been identified (Buckingham 1998). The building blocks of terpenoids are isoprenes (5C). Based on the number of the isoprene units, terpenes can be divided into hemiterpenes (C5), monoterpenes (C10), sesquiterpenes (C15), diterpenes (C20), triterpenes (C30), tetraterpenes (C40) and polyterpenes  $(\text{C}5)_n$  where n varies from 9 to 30000. The best-known monoterpenes are camphor, limonene, menthol. Other important terpenoid examples are sesquiterpenes, e.g. patchoulol, abscisic acid, pentalenolactone or diterpenes, e.g. taxol, casbene and gibberellin (McGarvey and Croteau 1995, Davis and Croteau 2000). Terpenoids often have multicyclic structures. Each differs from one another not only in functional groups but also in their basic carbon skeletons (Yermakov *et al.* 2010). In plants, they are involved in primary metabolism, as e.g. growth hormones (gibberellins, abscisic acid), photosynthetic

pigments (phytol, carotenoids, the side chain of chlorophyll), as electron carriers (ubiquinone, plastoquinone), mediators of polysaccharide assembly (polyprenyl phosphates) and structural components of membranes (phytosterols) (McGarvey and Croteau 1995). Besides physiological, metabolic and structural functions, terpenoid compounds are also involved in communication and defense, for example as attractants for pollinators and seed dispersers, competitive phytotoxins, antibiotics and herbivore repellents and toxins (Harborne 1991; McGarvey and Croteau 1995). Due to their biological activities and physico-chemical properties, many terpenoids are important renewable resources providing a range of commercially useful products, including solvents, flavorings and fragrances, adhesives, coatings and synthetic intermediates (Harborne 1991; McGarvey and Croteau 1995; Dawson 1994). They also provide industrially useful polymers (rubber, chicle) and a number of pharmaceuticals (limonene, artemisinin, taxol) (Crowell and Gould 1994; Van Geldre *et al.* 1997) and agrochemicals (pyrethrins, azadirachtin) (Harborne 1991). Some of these terpenoids are available in relatively large amounts as essential oils, resins, and waxes, but others which are naturally produced only in trace amounts are in short supply.

### 1.3.2. Terpenoid biosynthesis pathways

The biosynthesis of terpenoids can be divided into four major processes. The first stage is the biosynthesis of the active isoprene unit, isopentenyl pyrophosphate (IPP) and its following isomerization to dimethylallyl diphosphate (DMAPP) by IPP isomerase. Sequentially, this precursor can be converted by various prenyltransferases to higher order terpenoid building blocks, geranyl pyrophosphate (GPP; C<sub>10</sub>), farnesyl pyrophosphate (FPP; C<sub>15</sub>), and geranylgeranyl pyrophosphate (GGPP; C<sub>20</sub>). The next stage includes cyclization to create the basic parent skeletons of the various terpenoid families or self-condensation reactions to form the C<sub>30</sub> and C<sub>40</sub> precursors of sterols and carotenoids, respectively). Finally, oxidation, reduction, isomerization, conjugation, or other secondary transformations elaborate the unique and manifold character of the terpenoids (McGarvey and Croteau 1995).

The building blocks for the biosynthesis of terpenes, IPP and DMAPP, can be synthesized via two pathways: the mevalonate-dependent pathway (referred to as the MVA pathway) and the mevalonate-independent pathway (referred to as the MEP pathway or DXP pathway) (Rohmer 1993; Schwender 1996; Lichtenthaler 1997; Adam 1998; Rohmer 1999; Tholl 2006). These two pathways are illustrated in Figure 1.7.



**Figure 1.7: Isoprenoid biosynthetic pathway.** Isoprenoids are synthesized via two biosynthetic pathways: MVA pathway and MEP pathway. CoA is coenzyme A; HMG is 3-hydroxy-3-methylglutaryl; MVA is mevalonic acid; PMVA is mevalonate-5-phosphate; DPMVA is mevalonate-5-pyrophosphate; DXP is 1-Deoxy-D-xylulose 5-phosphate; MEP is 2-C-methylerythritol 4-phosphate; CDP-ME is 4-diphosphocytidyl-2-C-methylerythritol, MecPP is 2-C-methyl-D-erythritol 2,4-cyclopyrophosphate; HMBPP is (E)-4-hydroxy-3-methylbut-2-enyldiphosphate (Harada *et al.* 2009)

In general, eukaryotes use the MVA pathway exclusively to convert acetyl-CoA to IPP and DMAPP precursors, except plants using both the MVA and MEP pathways for isoprenoid synthesis (Boucher and Doolittle 2000; Rohdich *et al.* 2002). By contrast, in prokaryotes, isoprenoids are synthesized primarily via the MEP pathway (Boucher and Doolittle, 2000). For example, most *Streptomyces* strains have only the MEP pathway, but some of them use MVA pathway for the formation of isopentenyl diphosphate (Dairi 2005). For myxobacteria, the mevalonate pathway has been shown to be active in

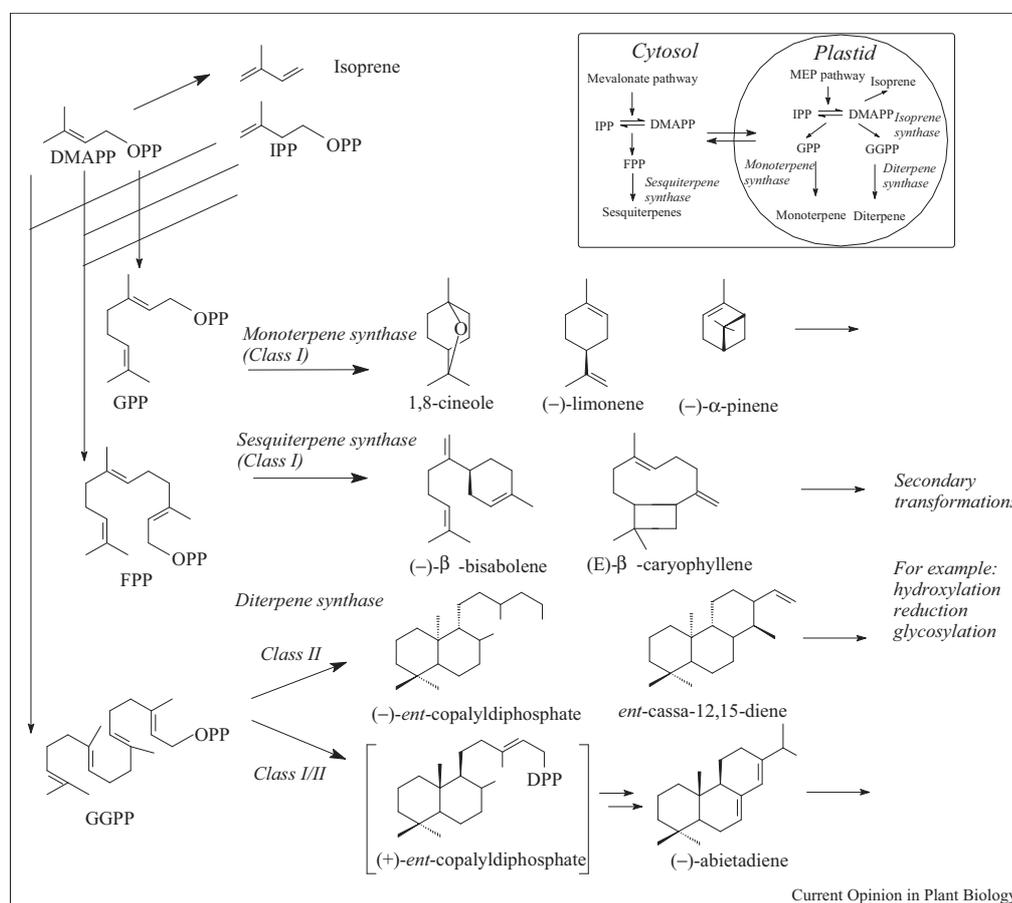
*Myxococcus xanthus* (Horbach *et al.* 1993) and *Stigmatella aurantiaca* (Bode *et al.* 2003). However, it cannot be excluded that myxobacteria have both MVA and MEP routes (Bode *et al.* 2003). In *E. coli*, IPP and DMAPP are essential for the prenylation of tRNAs and the synthesis of FPP, which is used for quinone and cell wall biosynthesis (Connolly and Winkler 1989; Apfel *et al.* 1999). Nevertheless, *E. coli* does not naturally produce appreciable quantities of GPP or GGPP, the precursors for mono- and diterpenes, respectively (Reiling *et al.* 2004).

Due to their important role in the isoprenoid biosynthesis, both MVA and MEP pathways are targets for metabolic engineering of terpenoid biosynthesis. In fact, many useful isoprenoids are present in a small quantity from their natural sources, and the chemical synthesis of these structurally complex compounds is often difficult and expensive. However, by introducing heterologous biosynthetic pathways into microbial host organisms, commonly into *E. coli* and *S. cerevisiae*, which are easy to manipulate, numbers of important terpenoid compounds have been produced at higher level (Takahashi *et al.* 2007; Muntendam *et al.* 2009; Misawa 2011). Successful metabolic engineering examples are the attempt in production of amorphadiene - precursor to antimalarial artemisinin in *E. coli* (Martin 2003) and the production of paclitaxel - a clinically important natural product in *S. cerevisiae* (Engels *et al.* 2008). Undoubtedly, understanding the isoprenoid pathway at molecular level is a key for a successful application of metabolic engineering – a new approach in exploiting natural terpenoid products from either plants or microorganism.

### 1.3.3. Terpene cyclases

As mentioned above, most terpenoids are cyclic and derived from linear prenyl pyrophosphate precursors (GPP, FPP, GGPP) (McGarvey and Croteau 1995). The primary enzymes responsible for the cyclizations of these isoprenoid precursors are *terpene cyclases* (or terpene synthases) (Tholl 2006). Monoterpene cyclase catalyzes the cyclization of GPP to monoterpenes, sesquiterpene cyclase creates sesquiterpenes from FPP and diterpene cyclases generate diterpenes from GGPP (Davis and Croteau 2000). Terpene cyclases are divided into two major classes based on their types of cyclization (Bohlmann *et al.* 1998). The cyclization reaction of the class I is initiated by the ionization of polyprenyl diphosphates to an allylic carbocation, followed by cyclization and deprotonation to the olefin. In another way, the cyclization reaction of the class II is initiated by protonation at the terminal-double bond of polyprenyl diphosphate

substrates (Dairi 2005). The enormous structural variety of terpenoid compounds is mostly due to the evolution of a large terpene synthase superfamily (Tholl 2006). Until now, dozens of terpene cyclase genes from eukaryotes, but only a limited amount from prokaryotes, have been reported (Dairi 2005; Ikeda 2007). Thus, this is a promising new field for future research.

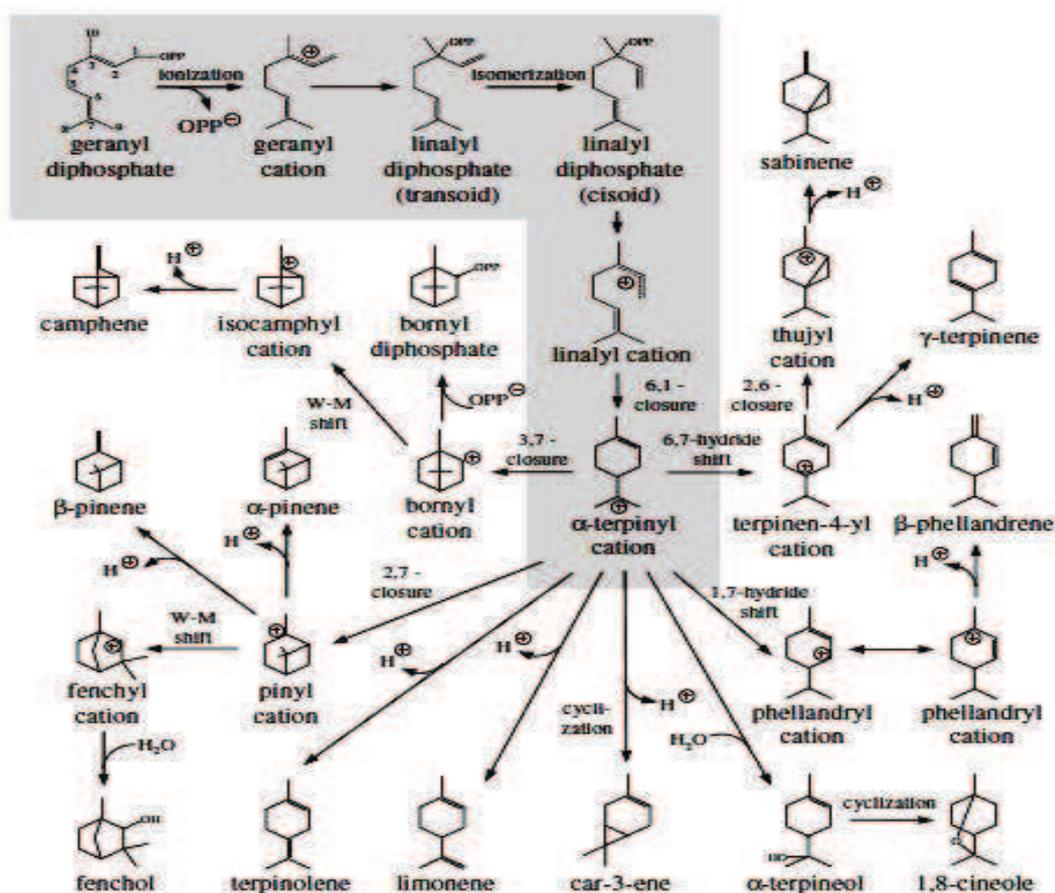


**Figure 1.8: Outline of the formation of plant terpenes (Tholl 2006)**

### 1.3.3.1. Catalytic mechanism

Terpene cyclases catalyze the cyclization of isoprenyl pyrophosphate precursors by a complex rearrangement reaction (Christianson 2006). The initiation step is a carbocationic reaction, where enzymes typically ionize the allylic pyrophosphate ester linkage of their substrate. Catalysis of this reaction is dependent on divalent metal ion co-factors, usually  $Mg^{2+}$ , which were hypothesized to bind to and assist ionization of the pyrophosphate (Christianson 2006; Zhou and Peters 2009). Typical mechanisms are the reactions of the monoterpene cyclases (Figure 1.9). These reactions start with the divalent metal ion-dependent ionization of the substrate GPP in the active site. The cationic intermediates then can undergo a series of cyclizations, hydride shifts or

rearrangements until the reaction is terminated by deprotonation or water capture (Christianson 2006; Degenhardt *et al.* 2009). This reaction mechanism was elucidated by Croteau and coworkers (Croteau 1987; Wise and Croteau 1998). However, monoterpene cyclases have an unusual feature, many of these enzymes produce multiple products. For example, besides the principal cyclic product limonene, limonene synthase generates smaller amounts of myrcene and  $\alpha$ - and  $\beta$ -pinene (Colby *et al.* 1993; Günnewich 2007). The cyclization catalyzed by sesquiterpene and diterpene cyclases is performed in a similar manner to the one used by monoterpene cyclases, although in most cases, a preliminary isomerization step is not required because of the greater flexibility of the longer C15 or C20 chain (McGarvey and Croteau 1995; Degenhardt *et al.* 2009). After the initial cyclization, a series of reactions is carried out including secondary cyclization, deprotonation to a neutral intermediate, hydride shifts, methyl shifts and Wagner–Meerwein rearrangements (Degenhardt *et al.* 2009).



**Figure 1.9:** The mechanisms of cyclic reaction of all monoterpene synthases start with the ionization of the geranyl diphosphate substrate. The resulting carbocation can undergo a range of cyclizations, hydride shifts and rearrangements before reaction is terminated by deprotonation or water capture. The formation of cyclic monoterpenes requires the preliminary isomerization of the geranyl cation to a linalyl intermediate capable of cyclization. The production of the initial cyclic species, the  $\alpha$ -terpinyl cation,

opens the door to secondary cyclizations. The route to the central  $\alpha$ -terpinyl intermediate is highlighted in grey and the reaction pathways to the major products of all cloned monoterpene synthases are shown. The formation of the acyclic monoterpenes linalool, myrcene and (E)- $\beta$ -ocimene might proceed either via the geranyl cation or via the linalyl cation. The numbering of carbon atoms of intermediates and products refers to that for GPP (Degenhardt *et al.* 2009)

### 1.3.3.2. Structure features

Plant terpene synthases have strongly conserved amino acid sequences, as well as similar intron-exon organization and exon sizes, suggesting a common evolutionary origin (Trapp and Croteau 2001; Komatsu *et al.* 2008). In contrast, microbial terpene synthases in general show no overall sequence similarity either to one another, except for ones synthesizing the same product (Hohn 1999; Cane and Watt 2003). Of the known microbial terpene cyclases, sesquiterpene cyclases have been the most extensively studied family because x-ray crystal structures of several sesquiterpene cyclase enzymes have been solved (Lesburg *et al.* 1997; Starks 1997; Caruthers 2000; Rynkiewicz 2001).

An important domain found in both plant and microbial terpene synthase family members is an aspartate-rich region, DDxxD (Figure 1.10) (Zhou *et al.* 2009; Degenhardt *et al.* 2009). It has been proven that this region is responsible for the binding of divalent metal ions that later react with the diphosphate moiety of the substrate (Lesburg *et al.* 1997; 1998; Starks *et al.* 1997). A highly conserved RxR motif located at about 35 amino acids upstream of the DDxxD has also been observed in terpene cyclases. This motif is involved in the complexation of the diphosphate function after ionization of the substrate to protect the carbo-cationic intermediates from a nucleophilic attack (Starks *et al.* 1997; Degenhardt *et al.* 2009). Another motif, NSE/DTE or (N/D)DXX(S/T)XXXE (metal binding residues in boldface), has been discovered by X-ray analysis of terpene cyclases (Rynkiewicz *et al.* 2001). It is believed to play a role as an additional metal cofactor-binding motif located on the opposite site of the active site entry. While the DDxxD motif is highly conserved throughout almost all plant terpene synthases, the NSE/DTE motif is likely less well conserved (Figure 1.10) (Christianson 2006; Zhou *et al.* 2009; Degenhardt *et al.* 2009). It has been suggested that a water molecule could substitute for the hydroxyl side chain of the central serine/threonine in terpene synthases when a glycine is located at the position of the NSE/DTE motif. Additionally, an arginine-rich motif has been found at about 60

residues from the N-terminus of many monoterpene synthases. It has a pattern like a tandem arginine (RR) which might participate in the isomerization of GPP to a cyclizable intermediate, such as the linalyl cation. This motif can be absent in monoterpene synthases producing only acyclic compounds (Williams *et al.* 1998; Degenhardt *et al.* 2009).

AgAS	(621)	DDL <sup>Y</sup> D	...	(765)	NDTKTYQAE
AtKS	(531)	DDFFD	...	(675)	NDIQGFKRE
CmKS	(536)	DDFYD	...	(680)	NDIRSYDRE
OsKSL10	(550)	DDFFD	...	(696)	NDSQTYRKE
OsKSL11	(553)	DDFFD	...	(698)	NDVMTYEKE
OsKSL8	(551)	DDL <sup>L</sup> FD	...	(696)	NDVMTYEKE
OsKSL5	(556)	DDFFD	...	(701)	NDMQTYEKE
OsKSL6	(556)	DDFFD	...	(701)	NDMQTYEKE
OsKS1	(552)	DDFFD	...	(695)	NDSQGFERE
OsKSL7	(577)	DDL <sup>L</sup> FD	...	(721)	NDIRGIERE
OsKSL4	(589)	DDFFD	...	(733)	NDIQSFERE

**Figure 1.10: Alignment of the terpene synthases divalent metal binding motifs (DDxxD and NSE/DTE) for selected diterpene synthases (Zhou *et al.* 2009)**

## 2. MATERIALS AND METHODS

### 2.1. Materials

All chemicals and reagents used in this work were analytic-grade. Valencene,  $\alpha$ -/ $\beta$ -ionone, (+)-nootkatone,  $\alpha$ -humulene, (-)- $\alpha$ -copaen, clovene, (-)-isolongifolene-9-one, (+)- $\alpha$ -longipinene, GPP, FPP, GGPP were purchased from Sigma-Aldrich, Germany. Restriction enzymes were purchased from New England Biolabs, Germany.

#### 2.1.1. Bacterial strains

**Table 2.1: *E. coli* strains used in this work**

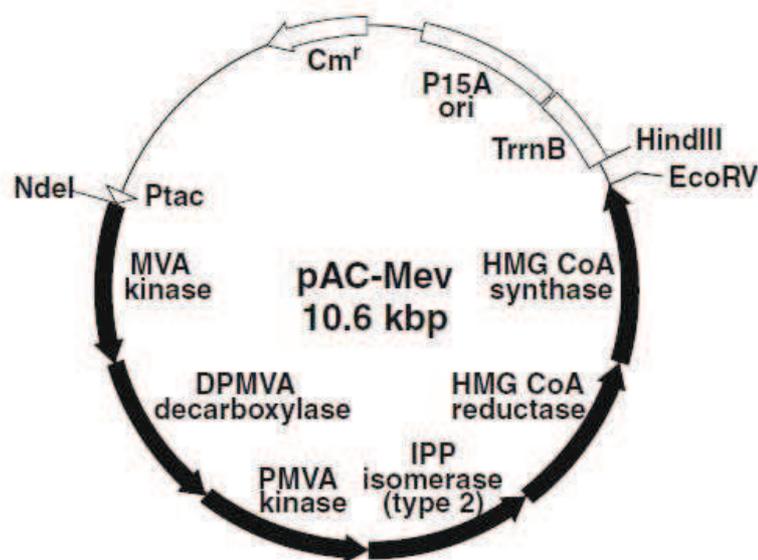
Strain	Genotype	Company
BL21(DE3)	<i>E. coli</i> B F <sup>-</sup> <i>dcm ompT hsd S (r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>)</i>	Novagen
C43(DE3)	<i>F<sup>-</sup> ompT gal dcm hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>)(DE3)</i>	(Miroux & Walker, 1996)
TOP10F'	<i>F' {lacIq, Tn10(TetR)} mcrA<sub>-</sub> (mrrhsdRMS- mcrBC) <math>\phi</math>80lacZ<sub>-</sub>M15<sub>-</sub>lacX74recA1araD1 39<sub>-</sub>(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG</i>	Invitrogen

#### 2.1.2. Genomic DNA

The genomic DNA of So ce56 was kindly provided by Dr Olena Perlova from Prof. Dr. Rolf Müller's group, Institute for Pharmaceutical Biotechnology, Saarland University. The genomic DNA was isolated using the Puregene<sup>®</sup> Genomic DNA Purification Kit (Gentra, Minneapolis, USA) according to the manufacturer's instructions as described (Perlova *et al.* 2006).

### 2.1.3. Vectors

Vector pET17b (the structure is illustrated in Appendix) was purchased from Invitrogen. Plasmid pGro12 containing the molecular chaperone *groES-groEL* was constructed by Nishihara (Nishihara *et al.* 1998). Vector pETC4 (Appendix) containing the CYP264B1 encoding gene was constructed previously by Dr. Yogan Khatri (Khatri PhD thesis 2009). Tricistronic vector pETMR5 (Appendix) containing CYP264A1, AdR and Adx encoding genes was kindly provided by Michael Ringle. The mevalonate vectors pAC-Mev (Figure 2.1) and pAC-Mv were kindly provided by Prof Norihiko Misawa, Research Institute for Bioresources and Biotechnology, Ishikawa Prefectural University, Japan. The plasmid pAC-Mev contains a cluster composed of five genes of the mevalonate pathway (encoding 3-hydroxy-3-methylglutaryl CoA (HMGCoA) synthase, HMG-CoA reductase, MVA kinase, phospho-mevalonate (PMVA) kinase, and diphosphomevalonate (DPMVA) decarboxylase) as well as the IPP isomerase gene from *Streptomyces sp.* strain CL190 (Harada *et al.* 2009). The plasmid pAC-Mv is derived from pAC-Mev by removing the genes encoding HMG-CoA reductase and HMG-CoA synthase and adding of a DNA fragment containing a multiple cloning site between *HindIII* and *EcoRV* (Yu *et al.* 2011)



**Figure 2.1:** Mevalonate vector pAC-Mev (Harada *et al.* 2009)

### 2.1.4. Primers

Oligonucleotides were synthesized by Eurofins MWG Synthesis Company (Ebersberg, Germany). The primers used in this work are listed in Table 2.1 (the rbs is underlined, restriction sites are *italic*, histidine-tag coding sequence is shown in **bold**)

**Table 2.1: List of primers**

Name	Oligonucleotide sequences
<i>NdeI</i> _GeoA_For	ccaat <i>catatg</i> tcatctgatcgaactagcg
<i>HindIII</i> _His_GeoA_Rev	ccaataagc <i>ttcagtgatggatggatg</i> atgcccagcgcgcacgatca
<i>EcoRI</i> _rbs_GeoA_For	ccaatga <i>attcaaggag</i> atataccatgcatctgatcgaactagcgttg
<i>NotI</i> _GeoA_Rev	ccaatgcggccgctcagtgatggatggatg
<i>AscI</i> -rbs- C4 -For	ccaatggcgcgcca <i>aggag</i> atatacatatgactcg
<i>SpeI</i> _C4_Rev	ccaatactagttcagtgatggatggatg

### 2.1.5. Proteins

Bovine AdR and Adx4-108 were kindly provided by Wolfgang Reinle. AdR was purified according to a procedure published by Sagara et al. (Sagara *et al.* 1993). Adx4-108 was purified following a protocol developed in our group (Uhlmann *et al.* 1994). Ferredoxin and ferredoxin reductase from *So ce56* were kindly provided by Dr. Kerstin Maria Ewen and purified as previously described (Ewen *et al.* 2009).

### 2.1.6. Culture media

Luria-Bertani (LB, Beckton Dickinson), Terrific broth (TB, Beckton Dickinson), Nutrient agar (NA, Sifin) media were used (Sambrook and Russell 2001). Ampicillin (Amp, Roth) of 100 µg/ml, kanamycin (Kan, Sigma) of 50 µg/ml, chloramphenicol (Cm, Sigma) of 30 µg/ml, isopropyl-beta-D-thiogalactopyranoside (IPTG, Sigma) of 1 mM, and arabinose (Ara, Sigma) of 4 mg/ml, delta-aminolevulinic acid ( $\delta$ -Ala, Fluka) of 0.5 mM final concentration were used in this study.

## 2.2. Methods

Common procedures, including preparation of *E. coli* competent cells, plasmid preparation, restriction digestions, plasmid transformations and other standard molecular biological techniques were carried out as previously described in Sambrook and Russell (Sambrook and Russell 2001). Polymerase chain reaction (PCR) was performed using phusion DNA polymerase (Finnzymes). Sodium dodecylsulphate polyacrylamid gelelectrophoresis (SDS-PAGE) was carried out according to Laemmli (Laemmli 1970). Protein was quantified using Bicinchoninic acid (BC) assay protein quantitation kit (Uptima Interchim, Montluçon, France). Material preparation for this part is illustrated in Appendix.

### 2.2.1. Plasmid construction

#### 2.2.1.1. Construction of pETG for terpene cyclase GeoA expression

The gene encoding the ORF of terpene cyclase (GeoA) was amplified from the genomic DNA of the myxobacterium *So ce56* with forward primer *NdeI\_GeoA\_For* and reverse primer *HindIII\_His\_GeoA\_Rev* using PCR with the temperature program as follows: denaturation at 98°C for 2 min, followed by 29 cycles of 98°C for 30 seconds, 50°C for 10 seconds, 72°C for 45 seconds and the final elongation at for 72°C 10 min. The target PCR fragment was then separated by agarose gel electrophoresis and afterwards purified using NucleoSpin (Macherey-Nagel) according to the manufacturer's instructions. After digestion with *NdeI* and *HindIII* (NEB), the PCR fragment was ligated into pET17b using Fast-Link™ DNA Ligation Kit (Invitrogen) according to the manufacturer's instructions, in a PCR machine with the following thermal program: 16°C for 15 min, 20°C for 15 min, 70°C for 15 min. The sequence of the terpene cyclase gene was confirmed by automatic DNA sequencing at MWG Biotech. The obtained vector was named pETG.

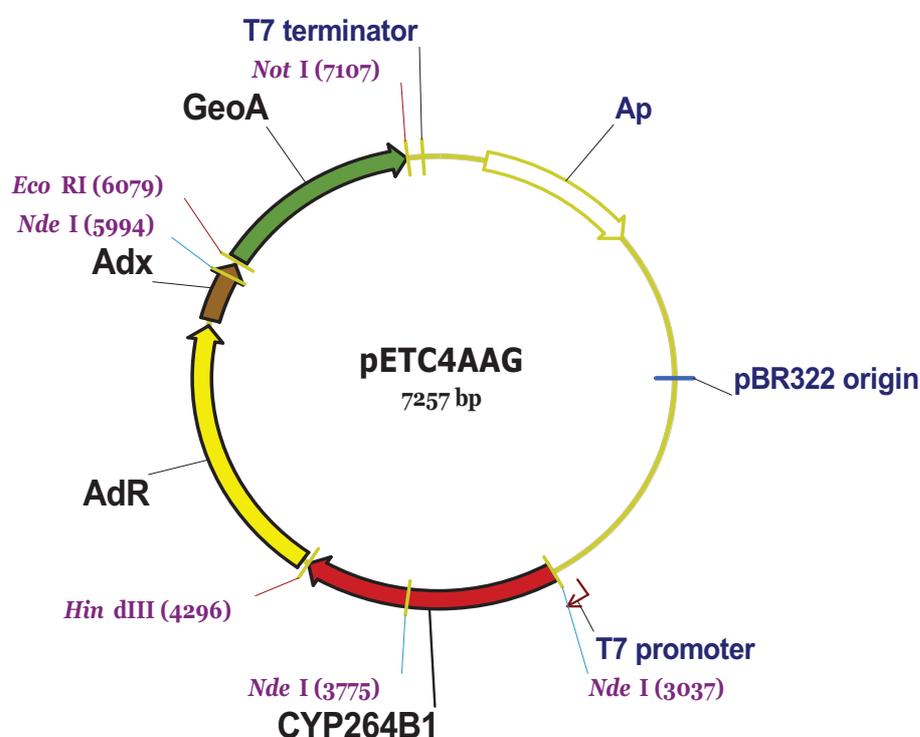
#### 2.2.1.2. Construction of pETC4AA for *in vivo* conversion

A DNA fragment containing adrenodoxin reductase (AdR) and truncated adrenodoxin (Adx) encoding genes with a Shine-Dalgarno site upstream of each gene was cut from the vector pETMR5 (Appendix) (a kind gift from Michael Ringle) by *HindIII* and *EcoRI* (NEB). This fragment was then inserted into vector pETC4 using the same

restriction sites, resulting in the tricistronic vector pETC4AA containing CYP264B1, AdR and Adx encoding genes.

### 2.2.1.3. Construction of pETC4AAG for terpene production

The gene encoding GeoA was amplified from vector pETG by PCR with thermal cycles as described in the section 2.2.1.1, using forward primer *EcoRI\_rbs\_GeoA\_For* and reverse primer *NotI\_GeoA\_Rev*. The *EcoRI\_rbs\_GeoA\_For* primer introduced an *EcoRI* site and a Shine-Dalgarno site upstream of the ORF of the GeoA gene. The target PCR fragment was then digested by *EcoRI* and *NotI* and inserted into vector pETC4AA using the same restriction sites, resulting in the tetracistronic vector pETC4AAG (Figure 2.2).



**Figure 2.2: Vector pETC4AAG.** Genes coding CYP264B1, AdR, Adx, GeoA are represented in red, yellow, brown and green color, respectively

### 2.2.1.4. Construction of pETGAA for terpene production

A DNA fragment containing adrenodoxin reductase (AdR) and adrenodoxin (Adx) encoding genes was cut from vector pETMRT5 by *HindIII* and *EcoRI*. This fragment was then inserted into vector pETG using the same restriction sites, resulting in the

tricistronic vector pETGAA containing terpene cyclase GeoA, AdR and Adx encoding genes.

#### **2.2.1.5. Construction of pAC-MvC4 for terpene production**

The gene encoding CYP264B1 was amplified from vector pETC4 by PCR reaction with thermal cycles as described in section 2.2.1.1, using forward primer *AscI\_rbs\_C4\_For* and reverse primer *SpeI\_C4\_Rev*. The target PCR fragment was digested and inserted into vector pAC-Mv between *AscI* and *SpeI*. The vector obtained was designated pAC-MvC4.

#### **2.2.2. Heterologous gene expression of CYP264B1 and GeoA**

*E. coli* strain BL21(DE3) was used for the expression of CYP264B1 and the terpene cyclase GeoA. Cells were co-transformed with the plasmid pET17b harboring *cyp264B1* (or *GeoA*) and the plasmid pGro12 carrying genes encoding the chaperones GroES and GroEL. Transformants were plated on nutrient agar (NA) plates containing Amp (100 µg/ml) and Kan (50 µg/ml) and incubated overnight at 37°C. A single colony was picked and inoculated in 10 ml of Luria-Bertani (LB) containing Amp and Kan at 37°C overnight under constant shaking. Overnight cultures were used to inoculate (1% v/v) 300 ml of terrific broth (TB) containing Amp and Kan in a 2-L baffled flask. The cultures were grown at 37°C under constant shaking until the optical density measured at 600 nm (OD<sub>600</sub>) reached 0.6-1.0. Cultures were then induced by addition of 1 mM of β-D-1 thiogalactopyranoside (IPTG) and 4 mg/ml arabinose and grown at 27°C for another 48 h under constant shaking. For CYP264B1 expression, 0.5 mM of delta-aminolevulinic acid (δ-Ala) was added at induction time to support the heme synthesis.

#### **2.2.3. Purification of His<sub>6</sub>-CYP264B1 and His<sub>6</sub>-GeoA**

Bacteria were harvested by centrifugation at 4000 g for 10min at 4°C and then resuspended in 10 mM potassium phosphate buffer, pH 7.4 containing 0.1 mM PMSF, 300 mM NaCl and 10% glycerol. Cells were disrupted by sonication (USD 30, WV) for 15 min on ice (Amplitude 25 µm, Puls 30 s, Ratio 1) and then centrifuged at 30.000 rpm (Rotor P45, centrifugator himac CP75β, Hitachi) for 30 minutes to remove cell debris. The supernatant was loaded onto an IMAC (immobilized metal ion affinity chromatography) column (TALON<sup>TM</sup>, Clontech). The purification was performed

according to the manufacturer's instructions. Fractions containing His<sub>6</sub>-tagged protein were then collected and concentrated using centrifugal filter (Utracel 30k, Amicon®Ultra, Leatherland). For CYP264B1, these fractions were recognized by its red colour. For GeoA, these fractions were recognized by using Bicinchoninic acid (BC) assay protein quantitation kit (Uptima Interchim, Montluçon, France) according to the manufacturer's instructions. The concentrated sample was applied to a Superdex 75 (GE Healthcare) column and the size exclusion chromatography was carried out at a constant flow rate of 0.1 ml min<sup>-1</sup> with 10 mM potassium phosphate buffer, pH 7.4. Protein fractions of sufficient purity were pooled, concentrated and stored at -80°C. The purity and molecular mass of the proteins were estimated by SDS-PAGE (Laemmli 1970). The concentration of GeoA was measured with the BC assay as mentioned. The concentration of CYP264B1 was measured according to the CO-reduced difference spectra method (Omura and Sato 1964) as described below (section 2.2.4).

#### 2.2.4. Determination of P450s concentration

The concentration of P450 was calculated using the CO-reduced difference spectra method (Omura and Sato 1964) and a double-beam spectrophotometer (UV2100PC, Shimadzu, Japan). The enzyme (in 10 mM potassium phosphate buffer, pH 7.4 containing 20% glycerol) was reduced with a few grains of sodium dithionite and divided into two cuvettes to record a base line between 400 and 500 nm. Then the sample cuvette was bubbled gently with carbon monoxide for 1 minute and a spectrum was recorded. The concentration of functional CYP264B1 was calculated according to the Beer–Lambert equation:

$$C_{P450} = \frac{\Delta A_{450-490} \times f_d}{\epsilon \times d}$$

$C_{P450}$ :	concentration of cytochrome P450 [mM]
$\Delta A_{450-490}$	absorptional difference at 450 nm and 490 nm
$f_d$	dilution factor
$\epsilon$	millimolar extinction coefficient [91 mM <sup>-1</sup> ×cm <sup>-1</sup> ]
$d$	diameter or length of the cuvette [cm]

### **2.2.5. Determination of redox partners for CYP264B1**

Redox partner determination was carried out based on the difference spectra described by Omura and Sato (Omura and Sato 1964). In these experiments, the heme iron of CYP264B1 (2  $\mu$ M) was reduced by NADPH (1 mM) instead of dithionite in the presence of various redox proteins as electron transfer mediators. During this experiment, the redox proteins from *S. cerevisiae* (5 ferredoxins and 2 ferredoxin reductases), from bovine adrenal (truncated adrenodoxin (Ad<sub>X4-108</sub>) and adrenodoxin reductase (AdR)) were tested. Different combinations of ferredoxin and ferredoxin reductase were analysed. The combinations giving a peak maximum at 450 nm were identified as suitable redox partners of CYP264B1.

### **2.2.6. Substrate determination**

#### **2.2.6.1. Substrate determination for CYP264B1 and analysis of substrate binding**

A substrate binding assay (based on the occurrence of a type I shift) was used to screen potential substrates for CYP264B1. The experiment was carried out at 25°C under aerobic conditions on a double-beam spectrophotometer (UV2100PC, Shimadzu, Japan) equipped with two tandem quartz cuvettes (Hellma, Müllheim, Germany). In each cuvette, one chamber contained 2  $\mu$ M CYP264B1 in 10 mM potassium phosphate buffer, pH 7.4, containing 20% glycerol, and the other chamber contained buffer. The substrate dissolved in DMSO was added equally (<2  $\mu$ l) in a final concentration of 5  $\mu$ M into the chamber containing the enzyme in case of the sample cuvette and into the chamber containing only buffer in case of the reference cuvette, respectively. After 5 min, spectra were recorded between 360 nm and 500 nm. Compounds causing a type I shift of the CYP264B1 spectrum were considered as potential substrates and used for enzyme activity assays.

For the determination of the binding constant ( $K_d$ ) to CYP264B1, substrate was titrated into enzyme solution containing 4  $\mu$ M CYP264B1 until the enzyme became saturated with substrate. The peak-to-trough difference was plotted against substrate concentration and fit into a nonlinear tight binding quadratic equation (Williams and Morrison 1979) using Origin 8.1 software. All titrations were carried out three times and the final values are the mean of the three experimental data sets.

### **2.2.6.2. Substrate determination for GeoA**

An *in vitro* assay was used to determine the substrate for GeoA. Universal substrates of terpene cyclases (GPP, FPP, GGPP) were tested. The *in vitro* assay was carried out in a volume of 500  $\mu$ l of 10 mM potassium phosphate buffer (pH 7.4) containing GeoA (3  $\mu$ M), substrate (50  $\mu$ M) and MgCl<sub>2</sub> (10 mM). Reactions were incubated for 1 hour at 30°C in a thermomixer (Eppendorf) under constant shaking. Reactions were stopped by adding 500  $\mu$ l of n-hexan and thoroughly mixing of the samples. Extraction was done twice with 2 volumes of n-hexan. The n-hexan phases were pooled and concentrated to 100  $\mu$ l in a vacuum-drier (SpeedVac Concentrator 5301, Eppendorf) before they were analysed by GCMS.

### **2.2.7. *In vitro* conversion**

#### **2.2.7.1. *In vitro* conversion with CYP264B1**

The *in vitro* reconstitution standard assays for substrates were carried out in a final volume of 500  $\mu$ l in 10 mM potassium phosphate buffer, pH 7.4, 20% glycerol. The reaction mixture contained 2  $\mu$ M CYP264B1, 200  $\mu$ M substrate, 40  $\mu$ M Adx<sub>4-108</sub>, 2  $\mu$ M AdR and a NADPH regenerating system consisting of 5 mM glucose-6-phosphate, 1 U glucose-6-phosphate dehydrogenase, and 1 mM MgCl<sub>2</sub>. Reactions were initiated by the addition of 0.1 mM NADPH. After incubation at 30°C for 30 min in a thermomixer (Eppendorf), reactions were stopped by adding 500  $\mu$ l chloroform. The samples were extracted twice and the extracts dried in a vacuum-drier. The residuals were dissolved in 200  $\mu$ l acetonitrile or 50  $\mu$ l ethylacetate for HPLC or TLC monitoring, respectively.

#### **2.2.7.2. *In vitro* conversion with GeoA**

The *in vitro* assay was described previously, in the section 2.2.6.2

#### **2.2.7.3. *In vitro* conversion with GeoA and CYP264B1**

The *in vitro* assays for substrates (FPP, GGPP) were carried out in a final volume of 500  $\mu$ l in 10 mM potassium phosphate buffer, pH 7.4, 20% glycerol. In general, this reaction mixture contained the combined components from the two reaction systems of CYP264B1 and of terpene cyclase as described above: CYP264B1 (2  $\mu$ M), GeoA (10  $\mu$ M), substrate (125  $\mu$ M), Adx<sub>4-108</sub> (40  $\mu$ M), AdR (2  $\mu$ M) and a NADPH regenerating

system consisting of 5 mM glucose-6-phosphate, 1 U Glucose-6-phosphate dehydrogenase, 10 mM MgCl<sub>2</sub>. The reaction of CYP264B1 was initiated by addition of 0.1 mM NADPH. After incubating at 30°C for 1 hour in a thermomixer, reactions were stopped by adding 500 µl chloroform, extracted twice and dried in a vacuum-drier. The residuals were dissolved in 100 µl ethyl acetate or 10 µl n-hexan for GCMS or TLC monitoring, respectively.

### 2.2.8. Kinetic analysis of CYP264B1

For enzyme kinetic experiments, substrate concentrations were varied in the range of 0-400 µM (nootkatone) and 0-200 µM ( $\alpha$ - and  $\beta$ -ionone), respectively. All other components were fixed as in the standard assay and reactions were carried out as described in 2.2.7.1. After 5 minutes, the reactions were stopped, extracted twice using chloroform and dried in a vacuum-drier. The residuals were dissolved in 100 µl acetonitrile and injected onto HPLC. Product quantification was performed by correlating the peak area of the respective product(s) with the combined peak area of the product(s) and the substrate. The product formation was plotted against the respective substrate concentration and fitted to the hyperbolic curve  $f(x) = ax/(b+x)$  (Michaelis – Menten kinetics) using the program SigmaPlot 9.0 to determine the values for the  $V_{max}$  and  $K_m$ ; standard deviations were derived from 3 independent experiments.

### 2.2.9. *In vivo* substrate conversion with CYP264B1

*In vivo* conversion was carried out for  $\beta$ -ionone. A single colony of *E. coli* C43(DE3) carrying plasmid pETC4AA was cultivated in LB medium containing Amp (100 µg/ml) at 37°C overnight under constant shaking. The overnight culture was used to inoculate (1% v/v) 50 ml of TB medium in a 300 ml baffled flask and grown at 37°C under constant shaking until the OD<sub>600</sub> reached 0.6-1.0. After the culture was induced by addition of 1 mM of IPTG and supplemented with 0.5 mM  $\delta$ -Ala it was further incubated at 30°C under constant shaking. After 24 hours, a stock solution of substrate dissolved in ethanol (1 M) was added to the culture to a final concentration of 300 µM and incubated for another 24-48 hours.

To analyze the *in vivo* conversion, aliquots of 0.5 ml culture were centrifuged to remove the cells. The supernatant was extracted twice with chloroform. The organic phase was

then evaporated in a vacuum-drier. The residuals were dissolved in 200  $\mu$ l chloroform or 50  $\mu$ l ethyl acetate for HPLC or TLC analysis, respectively.

### **2.2.10. Terpenoid production in mevalonate pathway-engineered *E. coli***

#### **2.2.10.1. Terpene production with GeoA**

*E. coli* C43(DE3) carrying plasmids pETG and pAC-Mev (or pAC-Mv) was precultured in TB medium containing Amp (100  $\mu$ g/ml) and chloramphenicol (Cm, 30  $\mu$ g/ml) at 37°C overnight under constant shaking. The overnight culture was used to inoculate (1% v/v) 30 ml TB medium containing Amp and Cm in a 50-ml flask. The culture was grown at 37°C under constant shaking. When the OD<sub>600</sub> reached 0.6-1.0, 1mM IPTG and 0.5 mg/ml of D-mevalonolactone (Sigma, Germany) were added and the cells were further cultivated at 27°C under constant shaking. After 3 hours, the culture was overlaid with 10 ml decane (Sigma, Germany) and incubated for another 48 - 72 hours under constant shaking. Aliquots of 100  $\mu$ l of the overlaid - decane phase were then taken for GCMS analysis.

#### **2.2.10.2. Terpene production with GeoA and CYP264B1**

*E. coli* C43(DE3) carrying plasmids pETC4AAG and pAC-Mev (or pAC-Mv) or plasmids pAC-MvC4 and pAAG was precultured in TB medium containing Amp (100  $\mu$ g/ml) and Cm (30  $\mu$ g/ml) at 37°C for overnight under constant shaking. The overnight culture was used to inoculate (1% v/v) 50 ml TB medium containing Amp and Cm in a 300-ml-baffle flask. The culture was grown at 37°C under constant shaking. When the OD<sub>600</sub> reached 0.6-1.0, 1mM IPTG and 0.5 mg/ml of D-mevalonolactone, 0.5 mM of  $\delta$ -Ala, 1.13% (w/v) of hydroxypropyl- $\beta$ -cyclodextrin (Sigma) were added and the cells were further cultivated at 30°C for another 48 hours with shaking. Aliquots of 0.5 ml of the culture were centrifuged to remove the cells. The supernatant were extracted twice using 1 ml chloroform. The chloroform phases were then evaporated in a vacuum-drier. The residuals were dissolved in 50  $\mu$ l ethyl acetate or 100  $\mu$ l n-hexan for TLC or GC-MS analysis, respectively.

### **2.2.11. Product purification and identification**

#### **2.2.11.1. Silicagel column**

In this work, a silicagel column was used to purify the products of  $\beta$ -ionone and nootkatone conversion. Products were prepared in an up-scaled *in vitro* reaction of 250 ml. Reactions were carried out in an incubator at 30°C under constant shaking. After 2 hours, reaction mixtures were extracted twice using 2 volumes of chloroform. The chloroform phases were then collected and evaporated in a rotary vacuum system (Büchi Rotavapor R-114). The residuals were dissolved in ethyl acetate and passed through a silica gel column (35-70  $\mu$ M particle size, 1,5 cm x 60 cm) using n-hexane/ethyl acetate (3:2) as eluent. Fractions were checked by TLC as described below (section 2.2.11.2). The fractions containing the product were combined and dried in a vacuum-drier.

#### **2.2.11.2. Thin layer chromatography**

TLC was used as a preliminary method to monitor the presence of compounds in the substrate conversion as well as to determine appropriate solvents for the purification using a silica gel column. Samples were spotted onto TLC aluminum plates (Silica Gel 60 F254, Fluka). The plates were developed in a solvent tank containing n-hexane/ethyl acetate. Different ratios of n-hexane/ethyl acetate were investigated to find out the appropriate developing solvent for each compound. After developing, the TLC plates were dried and visualized first at 254 nm (UV light). Then, plates were stained with a vanillin solution 1 % (w/v) in ethanol and sulfuric acid (95:5) and subsequently heated for 2 min at 110°C to visualize again.

#### **2.2.11.3. High performance liquid chromatography**

High performance liquid chromatography (HPLC) was used to analyze the substrate conversion. Aliquots of 0.5 ml of the substrate conversion were extracted with chloroform. The organic phases were evaporated in a vacuum-drier. The obtained residuals were dissolved in 200  $\mu$ M acetonitrile and applied onto a HPLC system (LC900, Jasco, Germany) using a reversed phase C18 column (Macherey-Nagel CC125/4 Nucleodur 100-5) and an isocratic solvent system containing acetonitrile/water

(55:45). A flow rate  $1 \text{ ml min}^{-1}$  was used. The products and substrates were monitored at 254 nm.

#### **2.2.11.4. Gas chromatography-mass spectrometry**

The products of the conversion by terpene cyclase were analyzed via gas chromatography-mass spectrometry (GC-MS). GC-MS analysis was performed by Dominik Pistorius, Department of Pharmaceutical Biotechnology, Saarland University on an Agilent 6890N Network GC-System combined with a 5973 Network Mass Selective Detector and a DB-WAX column (0.25 mm x 0.25 mm x 30 m, J&W Scientific) under the following conditions: EI 70 eV, source temperature 200 °C; carrier gas He; injection temperature 250 °C, interface temperature 300 °C; flow rate 1.0 ml/min, constant flow mode; splitless injection, column temperature program: Start at 50°C, hold for 1 min, then ramp to 150°C, 4°C/min, ramp to 300°C, 15°C /min, hold for 5 min, then ramp to 50°C, 15°C/min. All products were identified by comparison of their EI-MS spectra with those of the National Institute of Standards and Technology (NIST MS Search 2.0).

The products of the terpene cluster encoding enzymes were also analyzed via GC-MS. Measurements were performed by Nils Günnewich at CNRS-Strasbourg; on an Agilent 6890N (GC) and Agilent 5973N (MS) instrument under the following conditions: EI 70 eV with a scan width of  $m/z$  50-600, source temperature 250°C, column DB-5ms or HP-5ms (both Agilent Technologies); injection temperature 250°C, interface temperature 250°C (200°C); carrier gas He, flow rate 1.2 ml/ min., constant flow mode; splitless injection, column temperature program: 50°C hold for 0.5 min, then raised to 320°C at a rate of 10°C/min and then hold 5 min (which is in total 32.5 min). All products were identified by comparison of their EI-MS spectra with those of the NIST MS search 2.0.

#### **2.2.11.5. Nuclear magnetic resonance spectroscopy**

Nuclear magnetic resonance spectroscopy (NMR) was used to characterize the products of nootkatone and  $\beta$ -ionone conversions. NMR analysis was performed by Dr. Josef Zapp, Department of Pharmaceutical Biology, Saarland University. Ten milligrams of the products were prepared and used for  $^1\text{H}$  and  $^{13}\text{C}$  NMR with a Bruker DRX 500 NMR spectrophotometer. The spectra were recorded in  $\text{CDCl}_3$ . All chemical shifts are reported using the standard  $\delta$  notation in parts per million (ppm).

### 2.2.12. Homology modeling and substrate docking into CYP264B1 active site

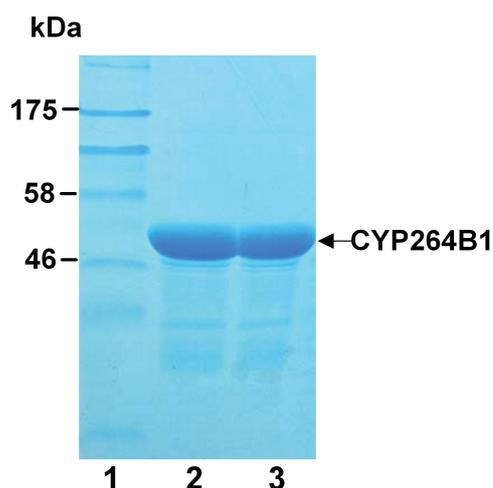
The homology modeling and substrate docking were performed by Dr. Michael C. Hutter, Center for Bioinformatics, Saarland University. P450eryF (CYP107A1) from *Saccharopolyspora erythraea* (accession code Q00441) shows a sequence identity of 30% to CYP264B1 of *S. cerevisiae*. The corresponding crystallographic structure (pdb entry 1Z8P) was chosen as a structural template using the automated mode of SWISS-MODEL (Arnold *et al.* 2006; Guex *et al.* 1997; Schwede *et al.* 2003). The coordinates of the heme-porphyrin atoms from the template structure were added subsequently to the obtained homology model. No further structural refinement of the model was performed, since the carboxylate groups of the heme formed ionic hydrogen-bonds with the side chains of Arg91, Arg287, and His343. Comparison of the obtained model with the template structure showed deviations only in the outer regions, whereas the region around the active center was strongly conserved. Compounds for docking were generated manually and energetically optimized using the MM+ force field as implemented in HYPERCHEM (version 6.02, Hypercube Inc., Gainsville FL, 1999). The program AUTODOCK (version 4.00) (Huey *et al.* 2007; Morris 1998) was applied for docking of  $\alpha$ -ionone (R- and S-isomer),  $\beta$ -ionone, and (+)-nootkatone into the homology model. The Windows version 1.5.2 of Autodock Tools was used to compute Kollman charges for the enzyme and Gasteiger-Marsili charges for the ligands (Sanner 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 ligating nitrogen atoms to -0.348 e. Flexible bonds of the ligands were assigned automatically and verified by manual inspection. A cubic grid box (54 x 54 x 54 points with a grid spacing of 0.375 Å) was centered 5 Å above the heme-iron. For each of the ligands 150 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.

### 3. RESULTS

#### 3.1. Expression and characterization of CYP264B1

##### 3.1.1. Protein expression and purification

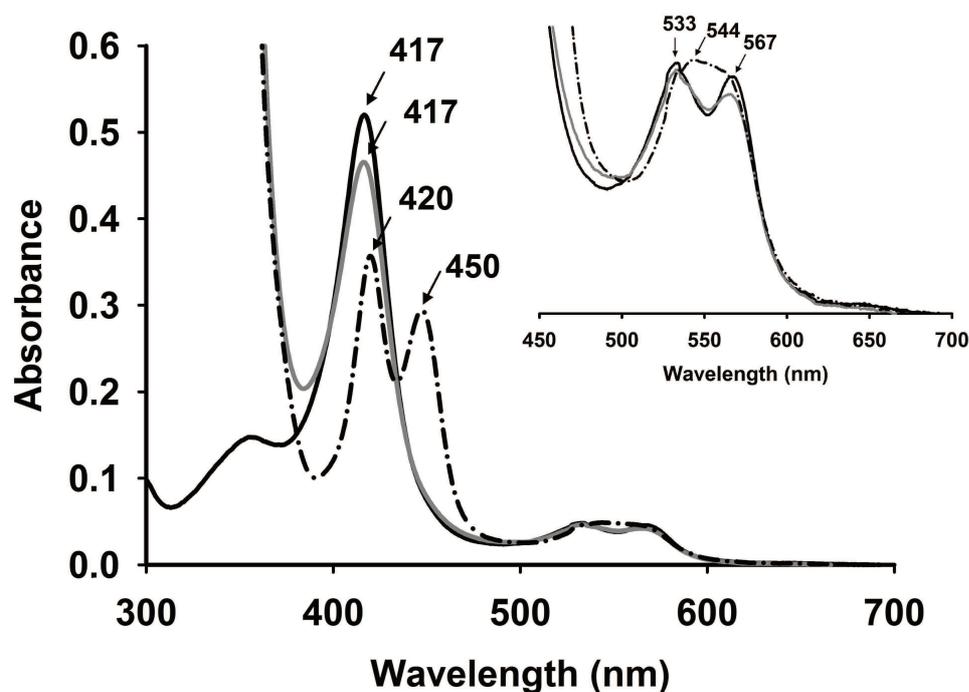
The gene coding for CYP264B1 in pET17b was expressed in *E. coli* BL21(DE3). To monitor the expression level, different temperatures (25°C, 28°C, 30°C, 32°C, 35°C, 37°C) were investigated. After 48 h of expression, cells were harvested, disrupted by sonification and centrifuged as described in section 2.2.3. The expression levels were calculated based on the reduced CO difference spectrum method (Omura and Sato 1964) for the crude extract enzyme from the supernatant. The result showed that the expression level was even less than 5 nmol L<sup>-1</sup> *E. coli* culture for all tested temperatures as mentioned above (data not shown). However, when temperature was decreased to 20°C, the expression level could be increased up to 257 nmol L<sup>-1</sup> *E. coli* culture. Moreover, co-expression of CYP264B1 with the chaperones GroES and GroEL at 28°C further improved the level of protein expression, leading to a yield of 820 nmol L<sup>-1</sup> *E. coli* culture. After affinity (IMAC) and size exclusion chromatography, the yield of the purified protein was 511 nmol L<sup>-1</sup>. Purity of protein was estimated by SDS-PAGE and the protein molecular weight was found to correspond well to the predicted value of 47 kDa (Figure 3.1).



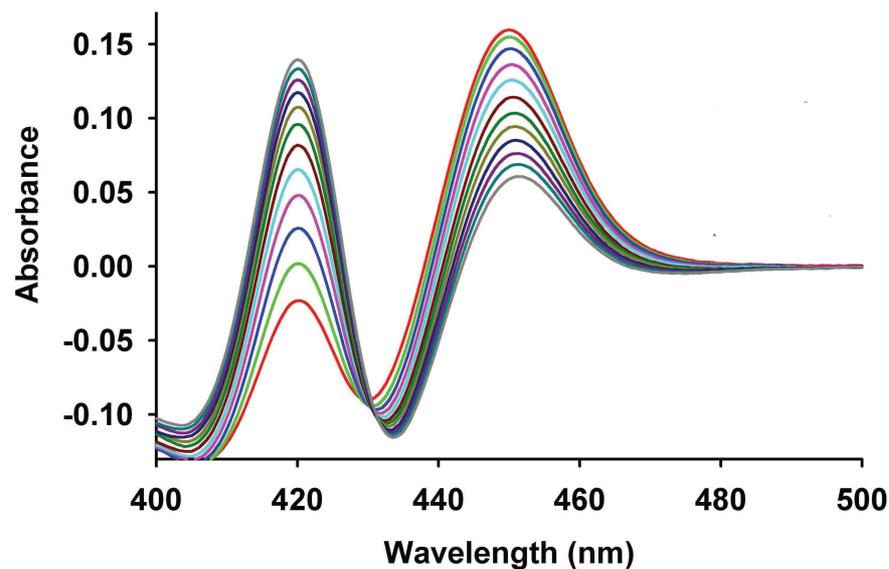
**Figure 3.1: SDS-PAGE of CYP264B1.** Lane 1 - prestained protein marker (BioLabs); Lane 2 and 3 - purified protein

### 3.1.2. Spectrophotometric characterization

To spectrophotometrically characterize CYP264B1, purified enzyme in potassium phosphate buffer (pH 7.4) containing 20% glycerol was measured in the UV-visible region from 300 to 700 nm. The result showed that the oxidized form of CYP264B1 exhibited a major ( $\gamma$ ) Soret peak at 417 nm and smaller  $\alpha$  and  $\beta$  bands at 567 nm and 533 nm, respectively (Figure 3.2), indicating that the protein was purified in its low-spin state. The dithionite reduced spectrum of CYP264B1 also showed the absorption maximum of the Soret band at 417 nm (Figure 3.2). The carbon monoxide bound form gave a typical peak maximum at 450 nm and a dominating shoulder at 420 nm (Figure 3.2). During a time interval of 0.5 min, the peak at 420 nm increased corresponding with the simultaneous decrease of the peak at 450 nm (Figure 3.3), which implied certain instability of CYP264B1 in its reduced form by dithionite.



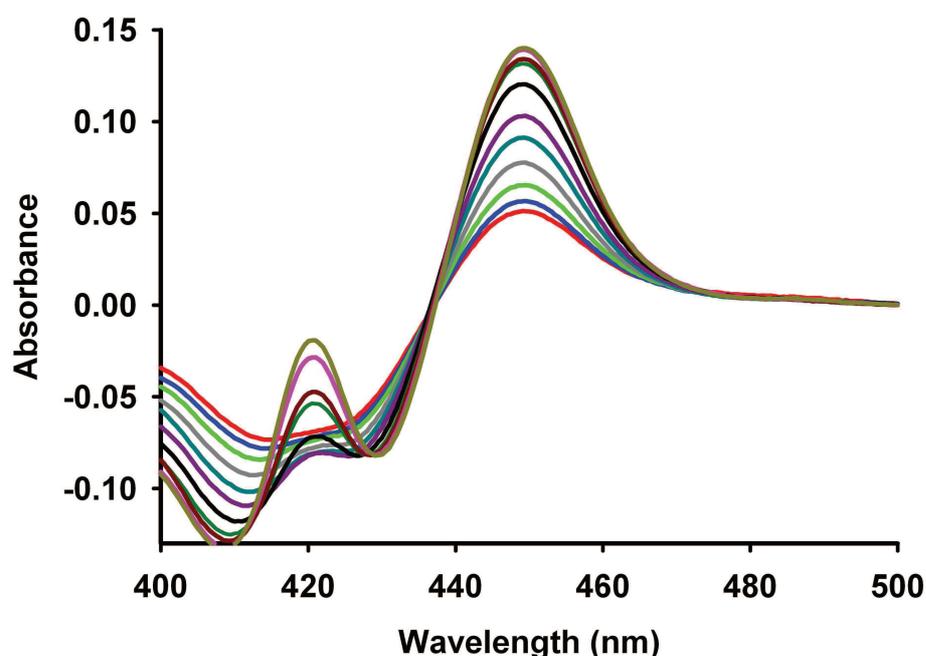
**Figure 3.2: UV-visible spectra of CYP264B1.** UV-visible absorption spectrum of oxidized (black, solid line), dithionite reduced (gray line), and reduced CO complexed (dashed with dotted line) enzyme. The inset shows a magnification of the spectrum in the  $\alpha/\beta$  band region. 2  $\mu$ M CYP264B1 was used



**Figure 3.3: CO-difference spectra of CYP264B1.** The time dependent development of the peak at 420 nm at 450 nm is shown. Red, green, blue, pink, cyan, dark red, dark green, dark yellow, dark blue, dark pink, dark cyan, dark gray line are CO-difference spectra of CYP264B1 at 0.5; 1; 1.5; 2; 2.5; 3; 3.5; 4; 4.5; 5; 5.5; 6 min, respectively

### 3.1.3. Redox partners

Most P450s perform their chemical reactions after receiving reduction equivalents from NADPH or NADH through electron transfer partners (Hannemann *et al.* 2007). Thus, the identification of the redox partners is the primary step necessary to characterize CYP264B1. In this work, 5 ferredoxins (Fdx1, Fdx2, Fdx3, Fdx5, Fdx8) and 2 ferredoxin reductases (FdRA and FdRB) from *So ce56*, as well as adrenodoxin (Adx<sub>4-108</sub>) and adrenodoxin reductase (AdR) from bovine adrenals were tested. Based on the CO-difference spectral method, various ratios and combinations of these autologous electron transfer proteins were tested as potential redox partners for CYP264B1. However, only with the combinations of Fdx2-FdRB and Fdx8-FdRB, a very low peak of the CO difference spectrum at 450 nm was observed (data not shown). In contrast, in the presence of the heterologous redox partners Adx and AdR, the typical CO-difference peak was observed with NADPH as electron donor. Reduction in this case occurs slower but the protein is more stable than upon the reduction by dithionite (Figure 3.4).



**Figure 3.4: CO-difference spectra of CYP264B1 using NADPH in combination with Adx and AdR as reducing agent.** The time dependent development of the peak at 420 nm at 450 nm is shown. Red, blue, green, dark gray, dark cyan, dark pink, black, dark green, dark red, pink, dark yellow are CO-difference spectra of CYP264B1 at 0.5; 1; 1.5; 2; 2.5; 3; 3.5; 4; 4.5; 5; 5.5 min, respectively. Optimal NADPH reduction was observed with the concentration of CYP264B1:Adx:AdR of 2  $\mu$ M:40  $\mu$ M:2  $\mu$ M

#### 3.1.4. Substrate screening

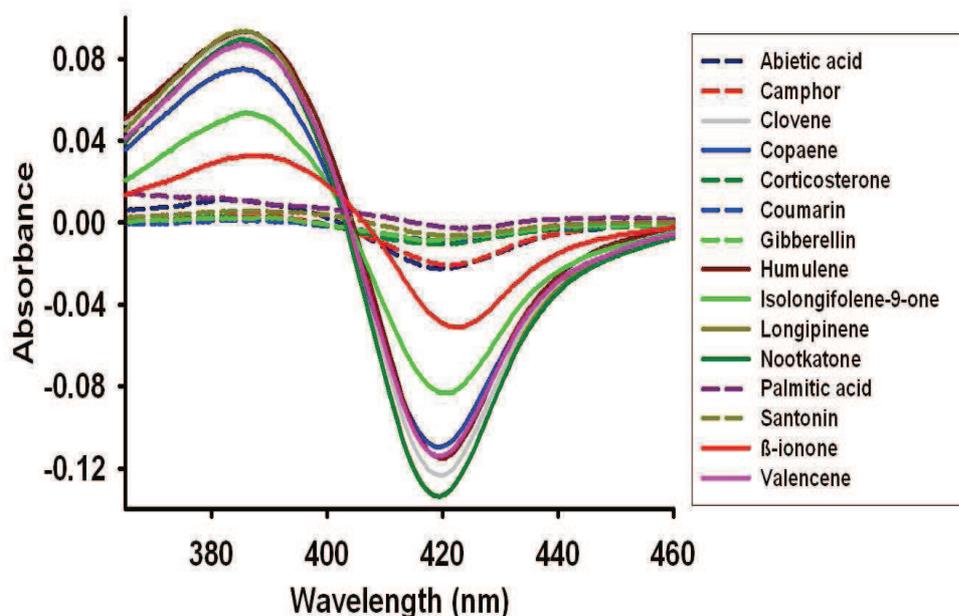
It has been shown that substrates can bind to the active site of the cytochrome P450 and displace the heme water ligand thereby causing a shift in the ferric heme iron from low-spin towards high-spin, which induces a so-called type I spectrum with the Soret absorption minimum around 420 nm and a maximum at about 390 nm (Denisov 2005; Lewis 2001). Based on this principle, 22 compounds (their structures are illustrated in Appendix) belonging to 9 groups (monoterpenoids, sesquiterpenoids, norisoprenoids, diterpene acids, steroids, fatty acids, and ketons and aroma related compounds) were investigated with regard to their ability to induce a spin-shift of CYP264B1 (Table 3.1).

**Table 3.1: Screening of potential substrates for CYP264B1**

(++) A clear spin shift was observed, (+) a slight spin shift was observed, (NS) no shift was observed.

Compound group	Name	Type I shift
Monoterpenoid	Camphor	+
Sesquiterpenoid	Nootkatone	++
	Valencene	++
	Isolongifolene-9-one	++
	Clovene	++
	Humulene	++
	$\alpha$ -longipinene	++
	Copaen	++
Norisoprenoid	$\alpha$ -ionone	++
	$\beta$ -ionone	++
Diterpenoid	Abietic acid	NS
	Gibberellin A3	+
Keton	Santonin	NS
Aroma	Coumarin	+
	7-ethoxy-coumarin	+
Steroid	Corticosterone	+
	Progesterone	NS
	Testosterone	NS
	17- $\alpha$ -methyl-testosterone	NS
Fatty acid	Palmitic acid	NS
	Myristic acid	NS
	<b>Stearic acid</b>	NS

Among 22 investigated compounds, the sesquiterpenes (valencene, nootkatone, humulene, copaen, clovene, longipinene, isolongifolene-9-one) and the norisoprenoids ( $\alpha$ -ionone,  $\beta$ -ionone) resulted in clear type I shifts of CYP264B1. Also with some of the other compounds (camphor, coumarin, gibberelline, santonin, corticosterone, progesterone), a slight type I shift was observed (Table 3.1 and Figure 3.5).



**Figure 3.5: Substrate screening based on type I shift of CYP264B1.** The interaction between 15 compounds and CYP264B1 are shown. The substrate dissolved in DMSO was added (final concentration of 5  $\mu\text{M}$ ) to a solution of CYP264B1 (2  $\mu\text{M}$ ) in 10 mM potassium phosphate buffer pH 7.4 containing 20% glycerol using tandem cuvettes as described in the section 2.2.6.1. Sesquiterpenes (clovene, copaene, humulene, isolongifolene-9-one, longipinene, nootkatone, valencene) and the norisoprenoid  $\beta$ -ionone (illustrated in solid lines), show a clear type I shift

### 3.1.5. Substrate binding constant

Because of the high selectivity of hydroxylation caused CYP264B1 (see section 3.1.6 below), only the products of  $\alpha$ -/ $\beta$ -ionone and nootkatone hydroxylation were further characterized. In the type I binding, the spectral change depends on substrate concentration in relation to the enzyme concentration. Therefore, the dissociation constant ( $K_d$ ) value of substrate and P450 can be determined based on this characteristic. In this work, the interactions between three compounds ( $\alpha$ -ionone,  $\beta$ -ionone and nootkatone) and CYP264B1 were studied. The substrates were stepwise added to 4  $\mu\text{M}$  enzyme until the enzyme became saturated. The peak-to-trough difference was plotted against substrate concentration and fit into a nonlinear tight binding quadratic equation (Williams and Morrison 1979) using Origin 8.1 software. The result showed that both the  $\alpha$ - and  $\beta$ -ionone-bound form of CYP264B1 gave a peak at 393 nm and a trough at 426 nm. The ionones showed a tight binding with  $K_d$  values of  $1.79 \pm 0.13 \mu\text{M}$  for  $\alpha$ -ionone and  $2.11 \pm 0.14 \mu\text{M}$  for  $\beta$ -ionone, respectively (Figures

3.6 a and b). The nootkatone-bound form of CYP264B1 gave a peak at 390 nm and a trough at 424 nm. This substrate also showed a tight binding with a  $K_d$  value of  $1.21 \pm 0.13 \mu\text{M}$  (Figure 3.6 c).

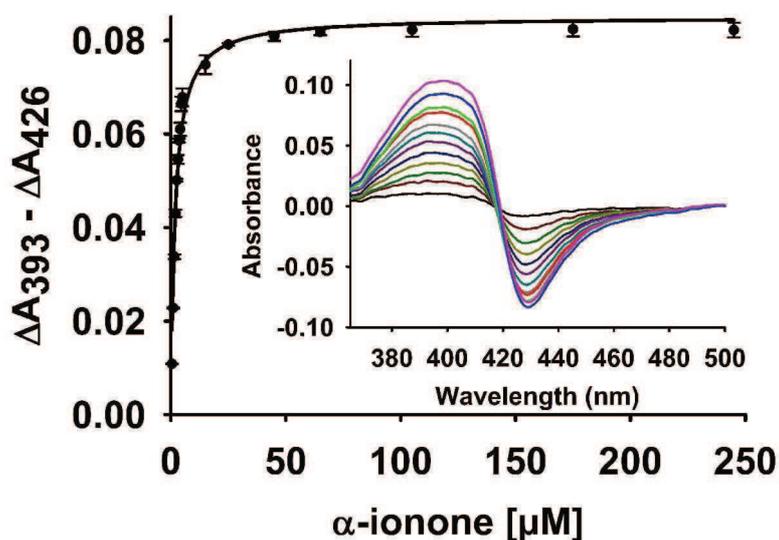


Figure 3.6 a: Spectral changes of CYP264B1 induced by various concentrations of  $\alpha$ -ionone. The inset illustrates spectral shifts induced by  $\alpha$ -ionone binding to CYP264B1

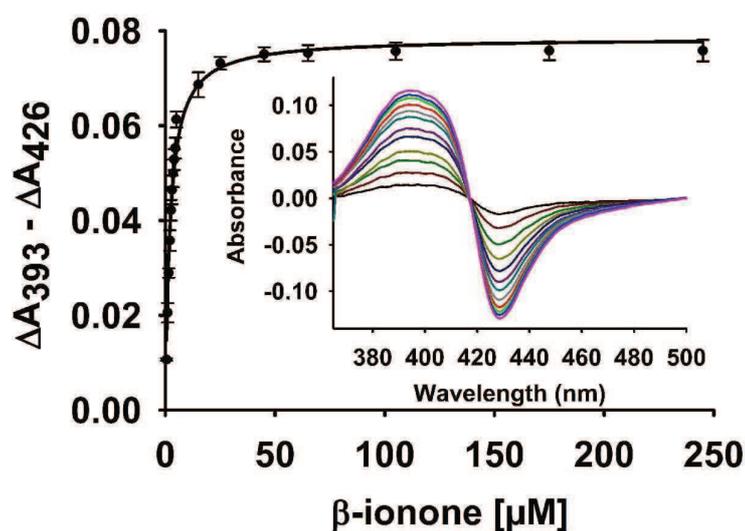
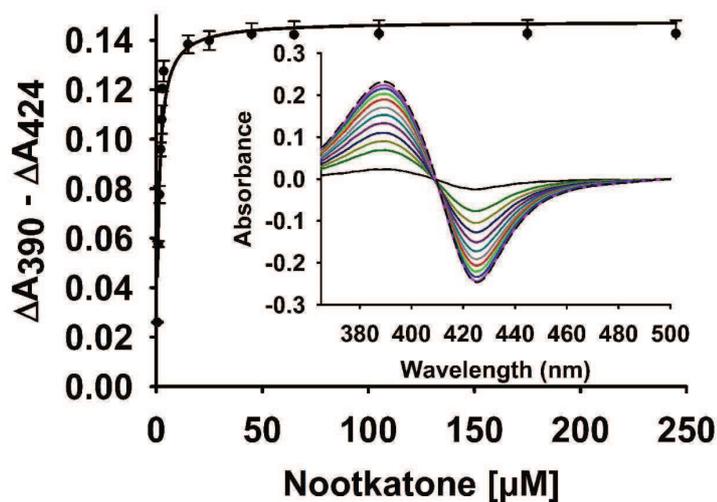


Figure 3.6 b: Spectral changes of CYP264B1 induced by various concentrations of  $\beta$ -ionone. The inset illustrates spectral shifts induced by  $\beta$ -ionone binding to CYP264B1



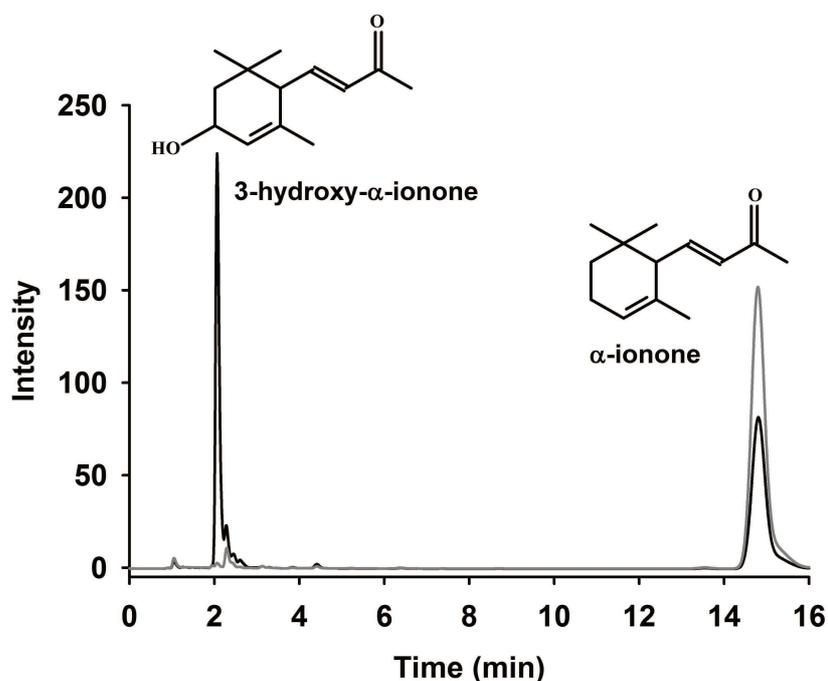
**Figure 3.6 c: Spectral changes of CYP264B1 induced by various concentrations of nootkatone.** The inset illustrates spectral shifts induced by nootkatone binding to CYP264B1

### 3.1.6. Substrate conversion

*In vitro* conversions were carried out for all compounds inducing type I shifts of CYP264B1 (see section 3.1.4). Reactions were performed at 30°C as described in the section 2.2.7.1. After an incubation time of 1 hour, reactions were stopped by adding 500  $\mu\text{l}$  chloroform. Products were extracted twice with chloroform and dried in a vacuum-drier. Product formations were monitored by TLC (see section 2.2.11.2) or HPLC (see section 2.2.11.3).

#### 3.1.6.1. Conversion of $\alpha$ -ionone

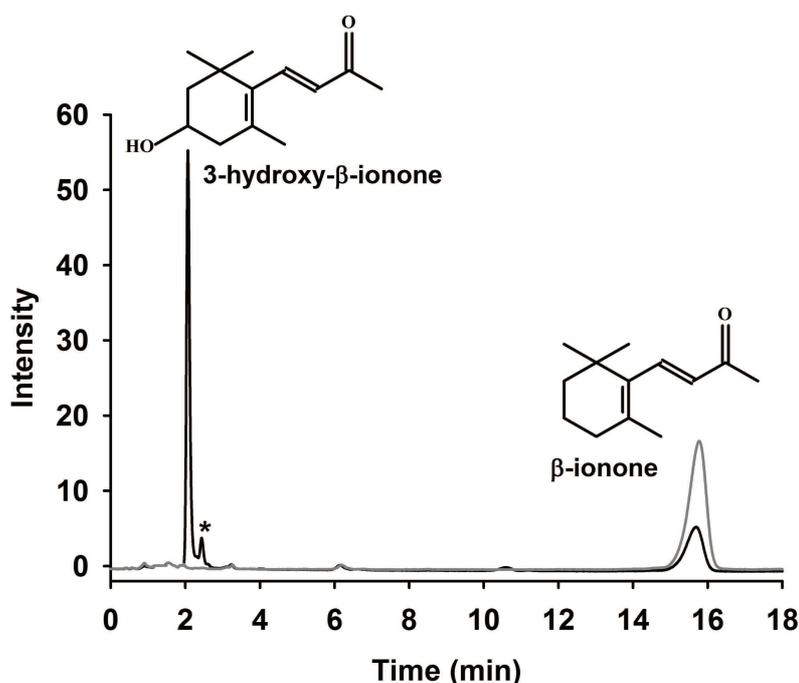
The *in vitro* conversion of  $\alpha$ -ionone by CYP264B1 was analyzed via HPLC. The chromatogram showed that CYP264B1 converted  $\alpha$ -ionone into a single product found at a retention time of 2 min. This product was compared with the authentic 3-hydroxy- $\alpha$ -ionone (Khatri *et al.* 2010) and finally confirmed as 3-hydroxy- $\alpha$ -ionone (Figure 3.7).



**Figure 3.7: Conversion of  $\alpha$ -ionone by CYP264B1.** HPLC chromatogram of  $\alpha$ -ionone conversion into 3-hydroxy- $\alpha$ -ionone; the conversion of substrate is shown in black line; the negative control without NADPH is shown in gray line

### 3.1.6.2. Conversion of $\beta$ -ionone

The conversion of  $\beta$ -ionone also resulted in a single product with a retention time of 2 min (Figure 3.8). Since no authentic hydroxylated product of  $\beta$ -ionone was available for comparison, the identification of the product was performed by NMR. To obtain a sufficient quantity of product for NMR analysis, the *in vitro* conversion was performed in a volume of 250 ml using crude extract enzyme. Under these conditions, another non-specific product was formed, which was also present in the negative control (crude extract of *E. coli* containing only Adx, only AdR or only CYP264B1) and detected at a retention time of 3 min (data not shown). After purification via a silicagel column, the CYP264B1-specific product was characterized by NMR. The  $^1\text{H}$  and  $^{13}\text{C}$  spectra revealed the structure of a monohydroxylated  $\beta$ -ionone and confirmed it as 3-hydroxy- $\beta$ -ionone.



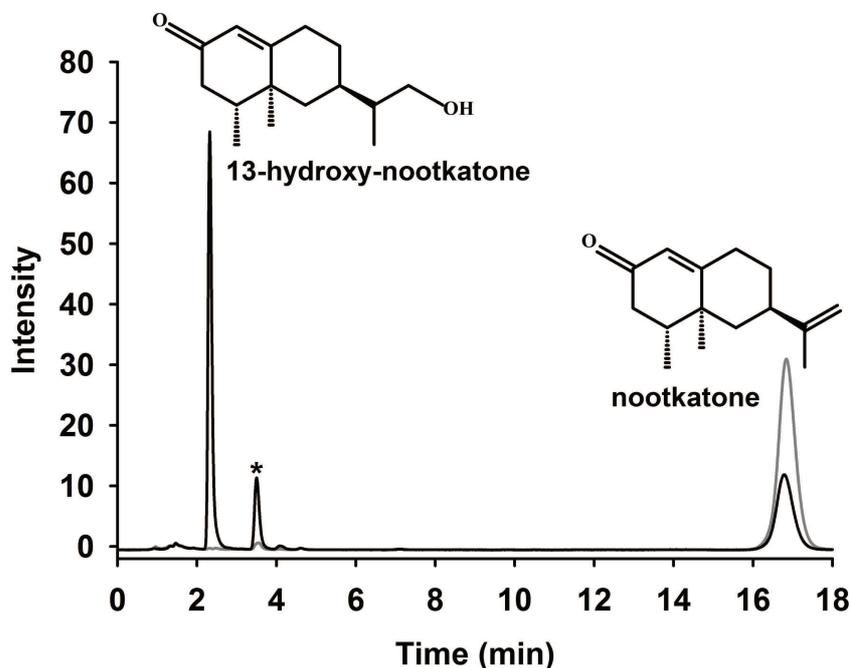
**Figure 3.8: Conversion of  $\beta$ -ionone by CYP264B1.** HPLC chromatogram of  $\beta$ -ionone conversion into 3-hydroxy- $\beta$ -ionone; the conversion of substrate is shown in black line; the negative control without NADPH is shown in gray line. Asterisk represents artifact

The detailed NMR data for 3-hydroxy- $\beta$ -ionone were obtained as:  $^1\text{H}$  NMR: ( $\text{CDCl}_3$ , 500 MHz),  $\delta$  1.09, *s*, 3H (H-12); 1.10, *s*, 3H (H-11); 1.47, *dd*,  $J = 12$  and 12 Hz (H-2a); 1.75, *s*, 3H (H-13), 1.77, *ddd*,  $J = 12$ , 3.5 and 2 Hz (H-2b); 2.06, *dd*,  $J = 17.5$  and 9.5 Hz (H-4a); 2.28, *s*, 3H (H-10); 2.4, 1 *brdd*,  $J = 17.5$  and 6 Hz (H-4b); 3.99, *dddd*,  $J = 12$ , 9.5, 6 and 3.5 Hz (H-3); 6.09, *brd*,  $J = 16.5$  Hz (H-8); 7.19, *brd*,  $J = 16.5$  Hz (H-7).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 125 MHz),  $\delta$  21.58 (C-13), 27.31 (C-10); 28.56 (C-12); 30.06 (C-11), 36.91 (C-1); 42.74 (C-4); 48.39 (C-2); 64.54 (C-3); 132.22 (C-5); 132.39 (C-8); 135.63 (C-6); 142.3 (C-7); 198.53 (C-9).

### 3.1.6.3. Conversion of (+)-nootkatone

The conversion of nootkatone by CYP264B1 revealed a major product (77%) at 2.3 min and a minor product (13%) at 3.5 min (Figure 3.9). Similar to  $\beta$ -ionone, no authentic hydroxylated product of nootkatone was available for comparison; thus the identification of the product was performed by NMR. Again a large-scale conversion in a volume of 250 ml using crude extract enzyme was carried out. Both, the major and minor product previously found in the conversion using purified enzyme, were detected in this case. However, the minor product was also present in the control experiment where crude extract of cells containing only Adx, AdR or CYP264B1 was used.

Therefore, only the major product was further characterized. After purification via a silicagel column, the major product was identified as 13-hydroxy-nootkatone by NMR.

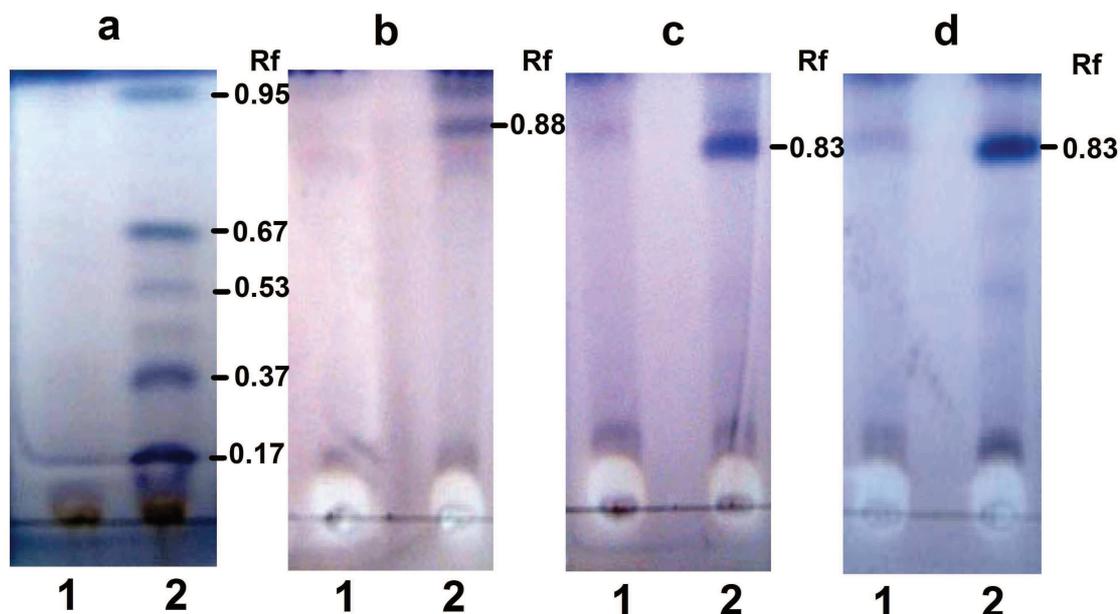


**Figure 3.9: Conversion of nootkatone by CYP264B1.** HPLC chromatogram of nootkatone conversion into 13-hydroxy-nootkatone; the conversion of substrate is shown in black line; the negative control without NADPH is shown in gray line. Asterisk represents the minor product

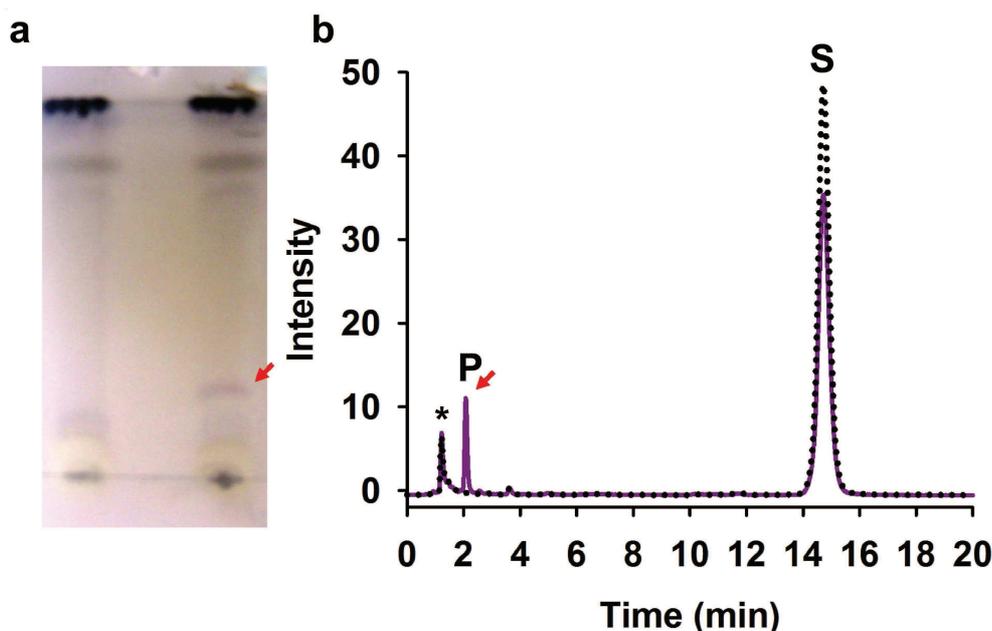
The detailed NMR data for 13-hydroxy-nootkatone are:  $^1\text{H}$  NMR: ( $\text{CDCl}_3$ ; 500 MHz)  $\delta$  0.95, *d*,  $J = 7$  Hz, 3H (H-14); 1.11, *s*, 3H (H-15); 1.15, *dd*,  $J = 13.5$  and 13.5 Hz (H-9b); 1.36, *m* (H-8b); 1.96, *m* (H-8a); 2.01, *m* (H-9a); 2.02, *m* (H-4); 2.22, *m* (H-3b); 2.29, *m* (H-3a); 2.37, *m* (H-6b); 2.43, *m* (H-7); 2.52, *m* (H-6a); 4.14, *s*, 2H (H-14); 4.90, *s* (H-12b); 5.07, *s* (H-12a); 5.76, *brs* (H-1).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ; 125 MHz)  $\delta$  14.90 (C-14); 16.80 (C-15), 32.11 (C-8); 33.07 (C-6); 36.05 (C-7); 39.44 (C-5); 40.42 (C-4); 42.06 (C-3); 44.42 (C-9); 65.32 (C-13); 109.19 (C-12); 124.82 (C-1); 152.56 (C-11); 170.07 (C-10); 199.54 (C-2). HMBC: (effects only for the methylene protons H-3, H-6, H-8 and H-9): H-3a to C-2, C-4, C-5 and C-14; H-3b to C-1, C-2, C-4, C-5 and C-14; H-6a to C-5 and C-7; H-6a to C-5, C-7 and C-15; H-6b to C-4, C-5, C-7, C-8, C-11 and C-15; H-8a/H-8b to C-9 and C-10; H-9a to C-1, C-8 and C-10; H-9b to C-1, C-7, C-8 and C-10.

### 3.1.6.4. Conversion of other sesquiterpenes

Among 7 tested sesquiterpenes (valencene, nootkatone, humulene, copaen, clovene, longipinene, isolongifolene-9-one), only copaene was not converted by CYP264B1. The enzyme was able to convert valencene into 5 products with retention factors ( $R_f$ ) of 0.17, 0.37, 0.53, 0.67, 0.95 on TLC (n-hexan/ethylacetate: 4/1),  $\alpha$ -longipinene into 1 product with  $R_f$  of 0.88 on TLC (n-hexan/ethylacetate: 3/2), clovene into 1 main product with  $R_f$  of 0.83 on TLC (n-hexan/ethylacetate: 3/2),  $\alpha$ -humulene into 1 main product with  $R_f$  of 0.83 on TLC (n-hexan/ethylacetate: 3/2) (Figure 3.10) and isolongifolene-9-one into 1 product with  $R_f$  of 0.22 on TLC (n-hexan/ethylacetate: 3/2) and retention time of 2 min on HPLC (acetonitrile/water: 55/45) (Figure 3.11). Products of these substrate conversions have not yet characterized due to lack of time.



**Figure 3.10: Conversion of of sesquiterpene compounds by CYP264B1** visualized on TLC using vallinin staining solution as described in the section 2.2.11.2. Conversion of valencene (a);  $\alpha$ -longipinene (b); clovene (c); humulene (d). Lane 1- negative control without NADPH; Lane 2- conversion with NADPH



**Figure 3.11: Conversion of isolongifolene-9-one by CYP264B1.** TLC (a): Lane 1 - negative control without NADPH; Lane 2 - conversion with NADPH. HPLC (b) dotted black line is the negative control without NADPH, dark pink line is the conversion with NADPH. Arrows represent formed product (P). S - substrate, asterisk - artifact

### 3.1.7. Enzyme kinetics

Catalytic activity of CYP264B1 was further characterized by enzyme dependent substrate conversion as described in the section 2.2.8. Activity of CYP264B1 was reconstituted with the heterologous redox partners, Adx and AdR. The ratio of CYP264B1/Adx/AdR was 2:20:1. Substrate concentrations were varied in the range of 0-400 mM (nootkatone) and 0-200 mM ( $\alpha$ - and  $\beta$ -ionone). Reaction was carried out in a thermal shaker at 30°C for 5 min. Assuming the absorbance properties of the products and substrate are the same, product quantification was performed by correlating the peak area of the respective product(s) with the combined peak area of the product(s) and the substrate. Product amount was plotted against the respective substrate concentration and fitted to the hyperbolic equation ( $f(x) = ax/(b+x)$ ); standard deviations were derived from 3 independent experiments. Results are shown in Figure 3.12.

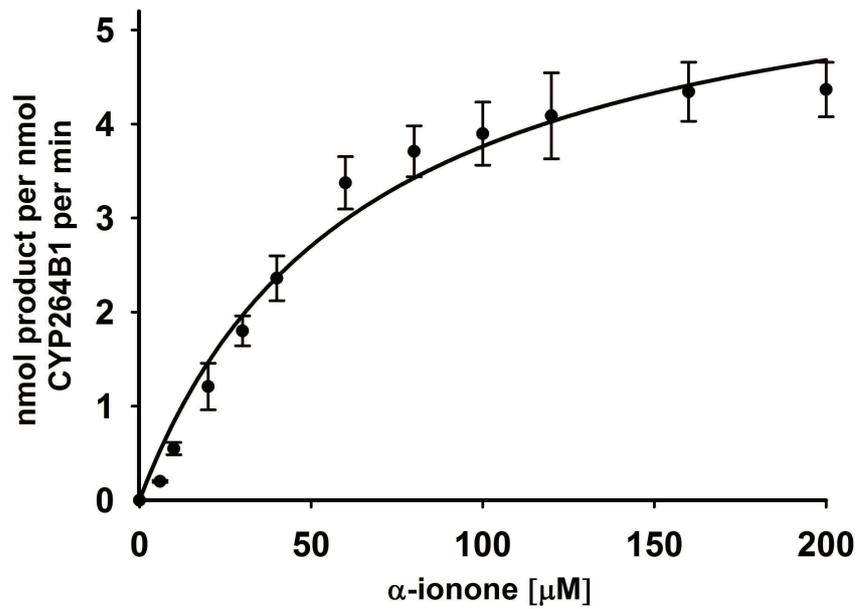


Figure 3.12 a: Determination of kinetic parameters for  $\alpha$ -ionone conversion catalyzed by CYP264B1

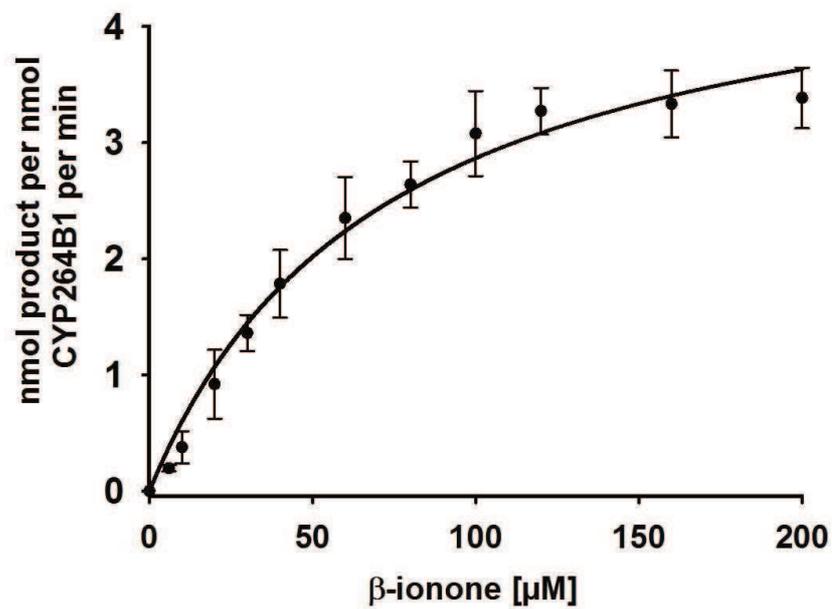
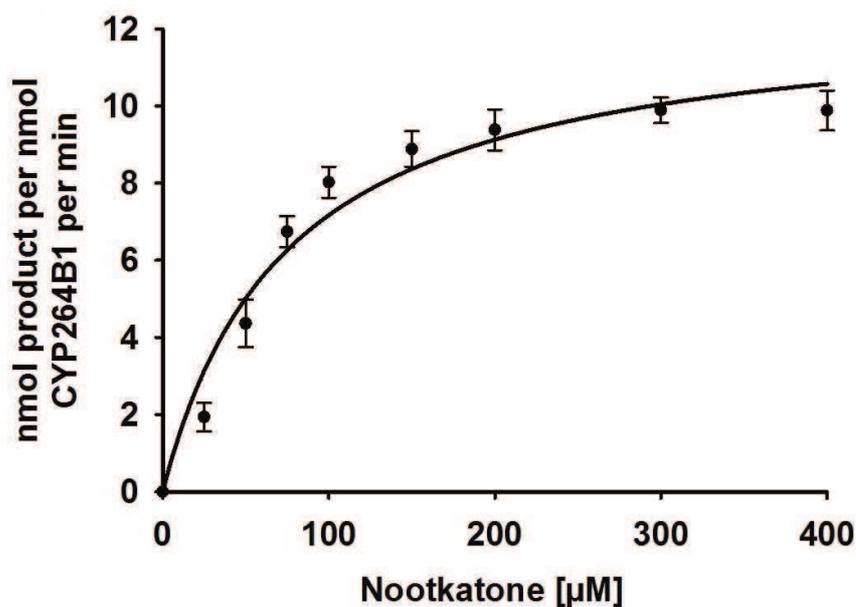


Figure 3.12 b: Determination of kinetic parameters for  $\beta$ -ionone conversion catalyzed by CYP264B1



**Figure 3.12 c: Determination of kinetic parameters for nootkatone conversion catalyzed by CYP264B1**

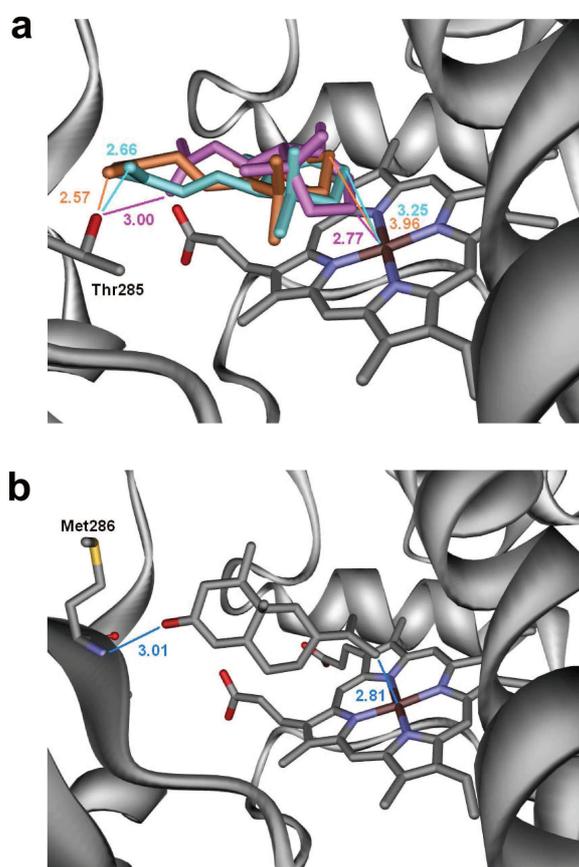
For conversion of  $\alpha$ -ionone into 3-hydroxy- $\alpha$ -ionone by CYP264B1, a  $K_m$  value of  $64.5 \pm 8.4 \mu\text{M}$  and a  $V_{\text{max}}$  value of  $6.19 \pm 0.33$  nmol product per minute per nmol CYP264B1 were determined ( $R^2 = 0.96$ ) (Figure 3.12 a). Similar values were also obtained for  $\beta$ -ionone ( $K_m$  and  $V_{\text{max}}$  values of  $72.72 \pm 11.43 \mu\text{M}$  and  $4.95 \pm 0.33$  nmol product per minute per nmol CYP264B1, respectively ( $R^2 = 0.96$ ) (Figure 3.12 b). Since the conversion of nootkatone gives one major and one minor product, the catalytic activity was calculated for the total product formation taking into account both products, with  $K_m$  and  $V_{\text{max}}$  values of  $75.34 \pm 12.57 \mu\text{M}$  and  $12.56 \pm 0.71$  nmol product per minute per nmol CYP264B1, respectively ( $R^2 = 0.95$ ) (Figure 3.12 c).

### 3.1.8. Homology modeling and substrate docking

To get a deeper insight into the structural basis for the selectivity of hydroxylation, a homology model of CYP264B1 was produced and the substrates were docked into this model using AUTODOCK. Both the R- and S-isomer of  $\alpha$ -ionone bind in an orientation that has the 3-position in a suitable distance to the heme-iron for hydroxylation, thus being in excellent agreement with the experimental results. Although the distance is shorter for the R-isomer ( $2.77 \text{ \AA}$ ) than for the S-isomer ( $3.25 \text{ \AA}$ ), the corresponding pose of the former is adopted less often than that of the S-isomer. The energetically most favorable docking position of ionone is likewise that showing the 3-position closest for hydroxylation ( $3.96 \text{ \AA}$ ). In all conformations found for  $\alpha$ -ionone and  $\beta$ -ionone that

allow hydroxylation in 3-position, the carbonyl oxygen is within hydrogen-bond distance to the hydroxyl group of Thr285. Moreover, these poses found suggest that hydroxylation is stereo-selective (Figure 3.13 a).

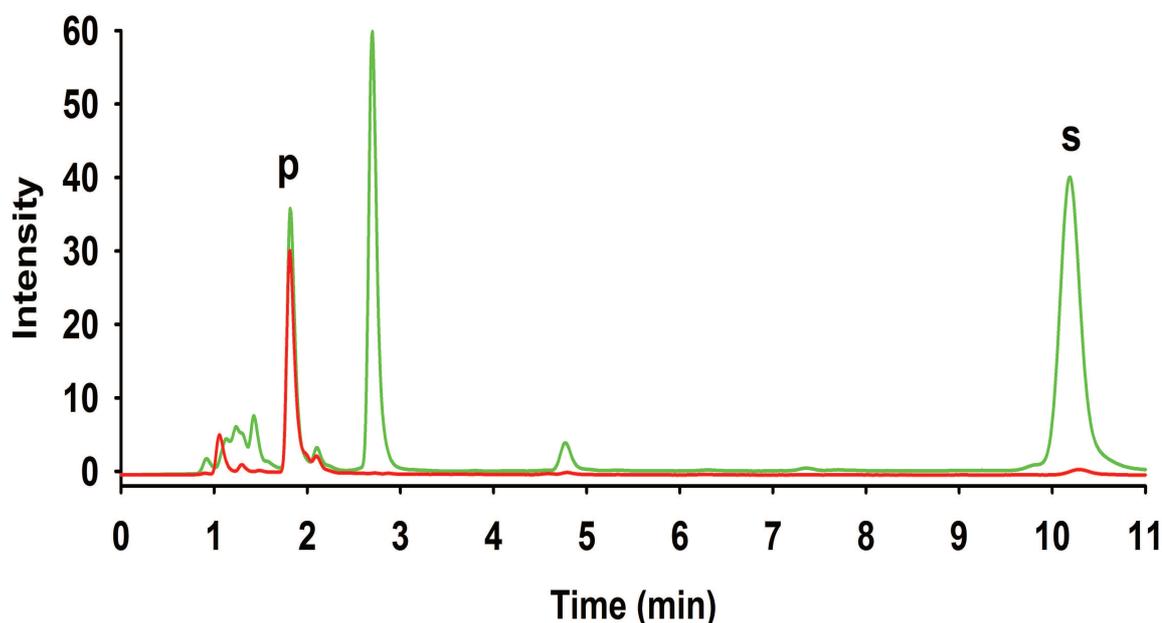
The most often adopted pose of (+)-nootkatone has the 13-position in a distance of 2.81 Å to the iron, whereby its carbonyl oxygen forms a hydrogen-bond with the backbone nitrogen of Met286. This corresponds to the major product confirmed experimentally. The docking position shows nootkatone oriented with the carbonyl oxygen towards the iron. Therefore, it is unlikely that the observed minor product is formed from this pose. Instead, it could be formed after reorientation of 13-hydroxy-nootkatone within the binding pocket, or after dissociation and rebinding in a different orientation (Figure 3.13 b).



**Figure 3.13: Obtained docking positions in the homology model of CYP264B1.** The R-isomer of  $\alpha$ -ionone is shown in pink, the S-isomer in cyan, and the  $\beta$ -ionone in orange. Likewise color coded are the distances (given in Ångstrom) between the hydroxyl oxygen of Thr285 and the carbonyl oxygens of the ionones, as well as between the 3-position and the heme iron, respectively (a). Distances (given in Ångstrom) are shown between the backbone nitrogen of Met286 and the carbonyl oxygen of nootkatone, as well as between position C-13 of nootkatone and the heme iron (b)

### 3.1.9. *E. coli* whole cell conversion for CYP264B1 substrate

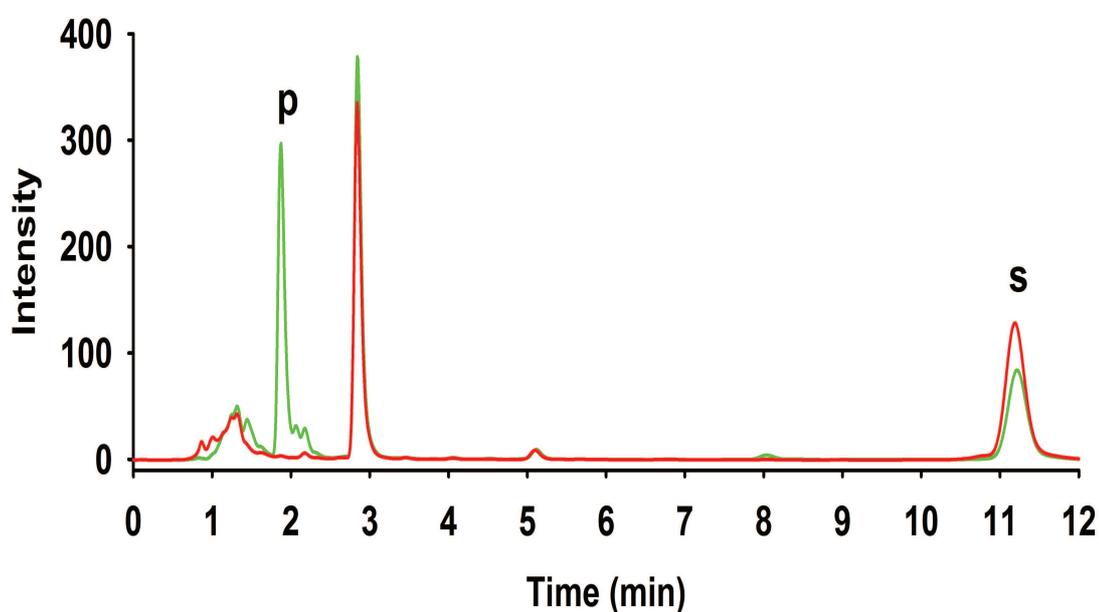
Since CYP264B1 was able to convert diverse sesquiterpene compounds, it is an attractive candidate to produce new sesquiterpene derivatives. For this purpose, using the isolated enzyme is not the solution of choice due to the high costs of pyridine nucleotide cofactors as well as the laborious protein purification steps. We therefore constructed a whole cell system for the conversion of CYP264B1 substrates. Genes coding for CYP264B1, Adx and AdR were introduced into vector pET17b, constructing the tricistronic vector pETC4AA. This vector was transformed into *E. coli* C43(DE3), which was subsequently tested for *in vivo* conversion of  $\beta$ -ionone 24 hours after induction,  $\beta$ -ionone was added to the culture at a final concentration of 300  $\mu$ M and incubation was continued. After another 22 and 48 hours, aliquots of 0.5 ml culture were centrifuged to remove the cells. The supernatant was extracted with chloroform and the bioconversion was analyzed by HPLC as described in the section 2.2.9. HPLC analyses of the *in vivo*  $\beta$ -ionone conversion are shown in Figure 3.14.



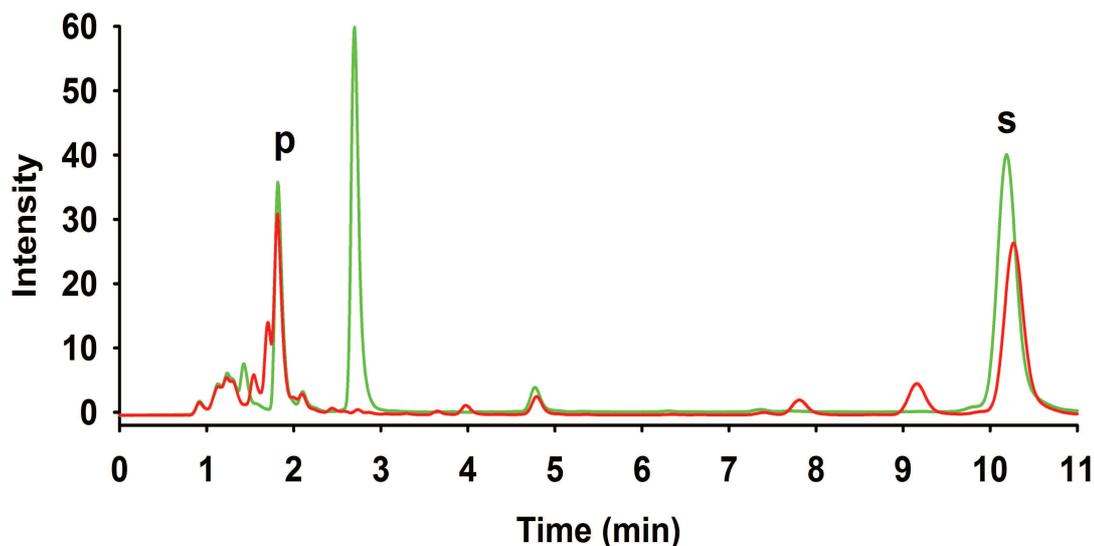
**Figure 3.14: HPLC of  $\beta$ -ionone product.** The red line and the green line indicate the product formation from the *in vitro* and *in vivo* conversion, respectively

As shown in Figure 3.14, in the *in vivo* conversion of  $\beta$ -ionone by CYP264B1 (green line), two products were formed. Comparison with the *in vitro* conversion (red line) showed that the specific product of CYP264B1 (3-hydroxy- $\beta$ -ionone) was formed and

could be detected at a retention time of 2 min. However, in the *in vivo* conversion, besides 3-hydroxy- $\beta$ -ionone product, another product with a retention time of 3 min was also found. In the control of *E. coli* carrying pETC4, only this side product was observed (Figure 3.15). This result indicated that 3-hydroxy- $\beta$ -ionone is the sole product of CYP264B1 and that the side product is produced by other enzymes of this strain. Yield of 3-hydroxy- $\beta$ -ionone was estimated to be 48 mg/L culture per day. In addition, Figure 3.16 shows that the product level after 48 hours was nearly similar to the level after 24 hours and the side product peak has dissappeared. This result indicated that the maximum yield of 3-hydroxy- $\beta$ -ionone product could be achieved at around 24 hours of conversion and the side product is unstable in time course.



**Figure 3.15: HPLC chromatogram of  $\beta$ -ionone conversions.** Comparison of the conversion of  $\beta$ -ionone by *E. coli* strain carrying pETC4AA (green line) and pETC4 (red green line)



**Figure 3.16: HPLC chromatogram of  $\beta$ -ionone conversions.** Comparison of the conversion of  $\beta$ -ionone by *E. coli* strain carrying pETC4AA after 22 hours (green line) and after 48 hours (red line)

## 3.2. Cloning, expression and characterization of GeoA

### 3.2.1. Computer-based identification of GeoA

The putative terpene cyclase GeoA with locus\_tag sce8552 (GenBank accession number YP\_001619202) contains 329 amino acids. To study the phylogenetic relationship between GeoA with other proteins in the NCBI database, the amino acid sequence of GeoA was used as a query for the local BLAST search of the GenBank using a recent non-redundant protein database. The amino acid sequence was identified to belong to the isoprenoid biosynthesis enzymes, class 1. The 10 closest homologues of GeoA were identified and are listed in Table 3.2. GeoA has the highest sequence identity (34%) with a terpene synthase metal-binding domain-containing protein, which was identified as a monoterpene cyclase for biosynthesis of the off-flavor terpenoid alcohol-2-methylisoborneol (2-MIB) of *Roseiflexus castenholzii* DSM 13941 (Komatsu *et al.* 2008). It also shares 32% and 30% identity with putative pentalene synthase of *So ce56* and pentalene synthase of *Streptomyces pristinaespiralis* ATCC 5486, respectively.

**Table 3.2: The closest homologues of the terpene cyclase**

Protein names	Organism	NCBI Reference Sequence	% Identity	Gaps %
Terpene synthase metal-binding domain-containing protein	<i>Roseiflexus castenholzii</i> DSM 13941	YP_001430766.1	34	3
Terpene synthase metal-binding domain-containing protein	<i>Roseiflexus sp.</i> RS-1	YP_001277817.1	33	2
Terpene synthase, metal-binding protein	<i>Rubrobacter xylanophilus</i> DSM 9941	YP_643279.1	30	1
Terpene synthase metal-binding domain protein	<i>Chitinophaga pinensis</i> DSM 2588	YP_003124367.1	29	2
Terpene synthase metal-binding domain-containing protein	<i>Nostoc punctiforme</i> PCC 73102	YP_001867159.1	30	1
Pentalenene synthase	<i>Sorangium cellulosum</i> So ce56	YP_001617016.1	32	1
Pentalenene synthase	<i>Streptomyces pristinaespiralis</i> ATCC 5486	ZP_06913376.1	30	4
Terpene synthase metal-binding domain protein	<i>Ktedonobacter racemifer</i> DSM 44963	ZP_06965495.1	27	2
Terpene synthase, metal-binding	<i>Anabaena variabilis</i> ATCC 29413	YP_322499.1	29	1
Terpene synthase metal-binding domain protein	<i>Chitinophaga pinensis</i> DSM 2588	YP_003123612.1	27	5

The genomic organization of genes coding for CYP264B1 and GeoA is similar to the sesquiterpene biosynthesis gene cluster of *Streptomyces avermitilis* (consisting of pentalenene synthase and CYP183A1) (Quaderer *et al.* 2006) and the sesquiterpene biosynthesis gene cluster of *Streptomyces coelicolor* A3 (consisting of epiisozizaene synthase and CYP170A1 (Zhao *et al.* 2008). Alignment of the terpene cyclase GeoA with the pentalene synthase and epiisozizaene synthase revealed that all contain two important domains characteristic for members of the terpene cyclase family: a

universally conserved aspartate-rich motif - DDXXD implicated in the binding of the pyrophosphate moiety of the substrate through chelation of the requisite divalent cation and a conserved triad of amino acid side chains - (D/N)DXX(S/T)XXXE - which plays an important role in Mg<sup>2+</sup> binding (Rynkiewicz *et al.* 2001; Degenhardt *et al.* 2009) (Figure 3.17). Percentage identity between GeoA and these two sesquiterpene cyclases are 25% (with penetalene synthase) and 28% (with epiisozizaene synthase), respectively.

```

sco5222  MHAFPHGTTATPTAIAVPPSLRLPVIEAAFPRLHPYWPKLQETTRTWLLEKRLMPADKV 60
sce8552  -----MSSDRTSVVVS-KRDAGGFYEPFAASCHPGREVTEQRTLAWVRRRLRVPDGRS 52
sav2998  -----MPQDVDFHIFPFSRRSPDFERARADHLSWPRALGLIGTDAA 41
                :. .* . * . :* **: .

sco5222  EEYADGLCYTDLMAGYYLGAPDEVLQAIADYSAWFFVWDDRHDRDIVHGRAGAWRRLRGL 120
sce8552  LSRLKATNFSHLAAWLLPSASTQTLQLASDFTAVLFLDDAYDEGQLSTDPESVEWLNEK 112
sav2998  AERHSRGGYADLAARFYPSATGADLDLGVDLMSWFFLFDDLFD-GPRGEDPQETRKLTA 100
                . . :.* * .*. *: * : **: ** .* . . . *

sco5222  LHTALDSPGDHLHHEDTLVAGFADSVRRLYAFLPATWNAFARHFHTVIEAYDREFHNRT 180
sce8552  YLGELFG-YTEADMSDPLTRGMLDVRERIRRSHPFFLNRLWLSHFQYYYEANLWEANNRK 171
sav2998  VAAALDG--PLPTSAPPIAHGFADVWRRTCQGMSPAWRARSARHWRNYFSGYVDEAVSRH 158
                * . . . *: * . * . : * **: . . * .*

sco5222  R-GIVPGVEEYLELRRLTFAHIWTDLLEPSSGCELPDAVRKHPAYRRAALLSQEFAAWY 239
sce8552  Q-MRVPHLEEYLMRRYSGAVYTYCDLLELLERPLPLEVQHPLIQTVRDICNDILCWT 230
sav2998  LNTPYDSAGHYLAMRRQTIGVQPTVDAERSCHCEVPQRVFDSAVLFAMLQIATDTNLIL 218
                .** **: . . ** * :* * . . :. :

sco5222  NDLCSLPKEIAGDEVHNLGISLITHHSLTLEEAIQEVRRRVEECITEFLAVERDALRFAD 299
sce8552  NDYFSLGKELTNGETHNLIVVLRNECVSTLEEAIQDRKMDHRRVAEYQGVKEKVL---A 287
sav2998  NDIASLEKEEARGELNMMVFILMREHGWRGRSIAHQDGVTRLEQFLLEACLPL---K 275
                ** ** * : . * **: . * . * . :* **: : : :

sco5222  ELADGTVRGKELSGAVRANVGNMNRNWFSSVYWFHESGRYM-VDSWDDRSTPPYVNEEA 358
sce8552  LWADDEIR-LYLDAVEAMIAGNQR-----WALEAGRYSGLESILVRAG----- 329
sav2998  VYDTFELTAQERESAKEYRMDGVRVIRGSYDWHRSSGRYAADYAIASQYQ-YLEELGS 334
                : . . . . * : :*** :

sco5222  GEK 361
sce8552  ---
sav2998  TL- 336

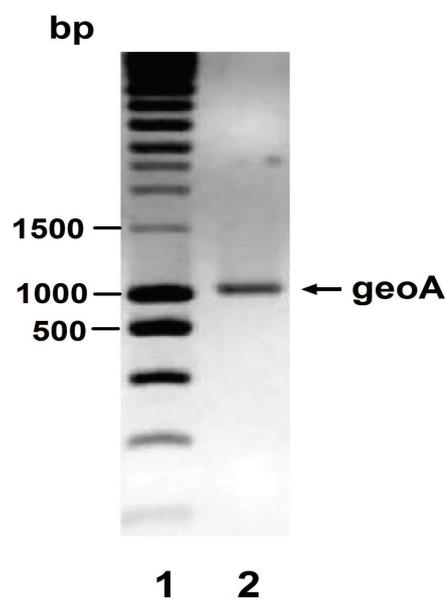
```

**Figure 3.17: Comparison of GeoA (sce8552) with pentalenene synthase (sav2998) from *Streptomyces avermitilis* and epiisozizaene synthase (Sco5222) from *Streptomyces coelicolor* A3.** The amino acid sequences of these enzymes were aligned by ClustalW2. Conserved residues are shaded. DDXXD: Universally conserved aspartate-rich motif, implicated in the binding of the pyrophosphate moiety of the substrate through chelation of the requisite divalent cation; (D/N)DXX(S/T)XXXE: Conserved triad of amino acid side chains, plays an important role in Mg<sup>2+</sup> binding

### 3.2.2. Cloning of the GeoA gene

The gene coding for GeoA was amplified from the genomic DNA of *So ce56* using primer pairs as listed in Table 2.1. The reverse primer was designed to add a region

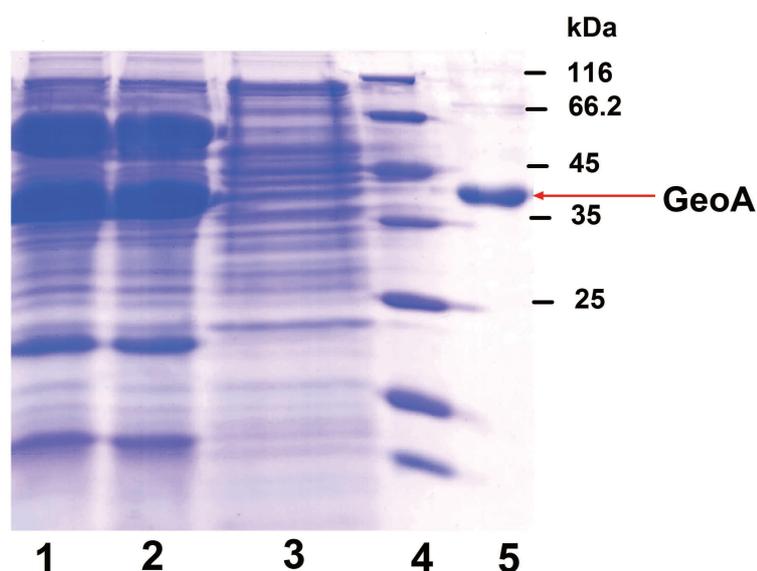
encoding a His<sub>6</sub>-tag to the gene. PCR amplification using phusion polymerase as described in the section 2.2.1.1, yielded a fragment of the expected size of 1008 bp (Figure 3.18). This fragment was then inserted into expression vector pET17b between *Nde*I and *Hind*III sites for heterologous protein expression.



**Figure 3.18: Agarose gel electrophoresis of terpene cyclase PCR product** (using 1% agarose gel). Lane 1 - Marker Smart Ladder (Eurogentec); Lane 2 - PCR product

### 3.2.3. Protein expression and purification

GeoA was overproduced as His-tagged fusion proteins in *E. coli* BL21(DE3). To monitor the expression level, different temperatures (from 25°C to 37°C) were investigated. The expression levels were estimated by SDS-PAGE for the crude extract. However, at first no expression was observed (data not shown). To optimize the protein expression, conditions similar to those used for the CYP264B1 expression were applied for GeoA. Thus, GeoA was co-expressed with the chaperones GroES/GroEL at 28°C and, indeed, a new band at the size expected for GeoA was observed very clearly (Figure 3.19). After affinity (IMAC) and size exclusion chromatography, the yield of the purified protein attained 16.3 mg/l culture. The purity of GeoA was estimated by SDS-PAGE and the protein's molecular weight was in good agreement with the predicted value of 38.92 kDa (Figure 3.19).



**Figure 3.19: SDS-PAGE represents the expression and purification of terpene cyclase GeoA.** Lane 1 and 2 - cell lysate after 48 hours of co-expression with chaperone; Lane 3 - cell lysate before induction; Lane 4 - protein marker (pEQLab); Lane 5 - purified terpene cyclase

#### 3.2.4. Substrate identification

Terpene cyclases are the primary enzymes responsible for catalyzing the formation of monoterpenes (C10), sesquiterpenes (C15) or diterpenes (C20) from the universal prenyl pyrophosphate precursors GPP, FPP or GGPP, respectively (Tholl 2006). Therefore, these three substrates were tested in *in vitro* conversion. After incubation with GeoA in a reaction mixture containing  $Mg^{2+}$ , compounds were extracted with n-hexan and analysed by GCMS as described in the section 2.2.11.4. A single product peak was detected for the reaction mixtures containing FPP (Figure 3.20) and GGPP (Figure 3.21) at retention times of 22 min and 32 min, respectively. In case of GPP, there was no product peak observed (data not shown). It is noteworthy, that the substrates were not observed in the chromatogram because they are hydrophilic compounds and thus do not enter the n-hexan phase. In the reaction without adding  $Mg^{2+}$ , no product was found (data not shown). This indicates that  $Mg^{2+}$  plays an important role for the activity of the terpene cyclase GeoA.

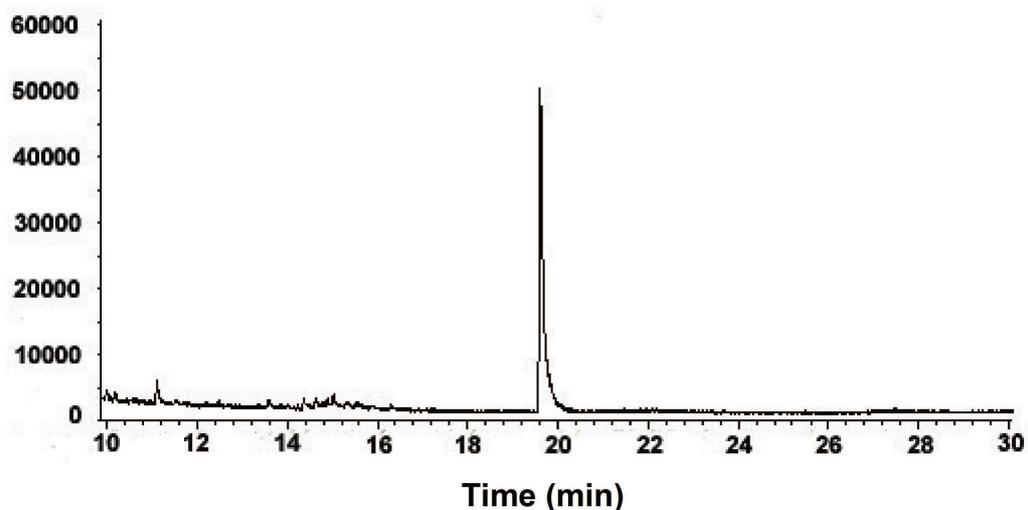
**Abundance**

Figure 3.20: Gas chromatogram of FPP conversion by terpene cyclase GeoA

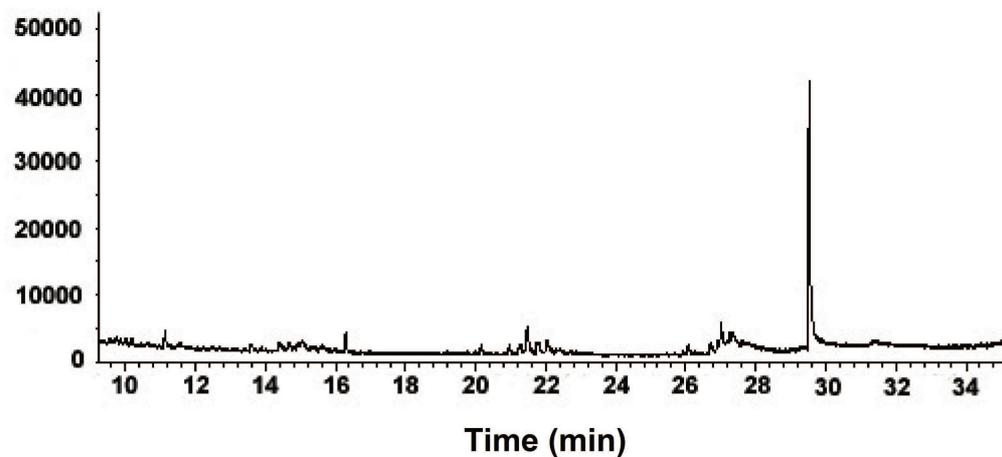
**Abundance**

Figure 3.21: Gas chromatogram of GGPP conversion by terpene cyclase GeoA

### 3.2.5. Product identification

Identification of the products of FPP and GGPP conversions was attempted based on comparison of the mass spectra to the mass spectra library of NIST 2.0 (National Institute of Standards and Technology). The product of FPP is a sesquiterpene with 89% similarity to valencene - a sesquiterpene that is an aroma component of citrus fruit and citrus-derived odorants (Figure 3.22). The product of GGPP is a diterpene with 80%

similarity to neocembrene A - a major component of the trail-following pheromone in the genus *Prorhinotermes* (Sillam-Dusses *et al.* 2005) (Figure 3.23).

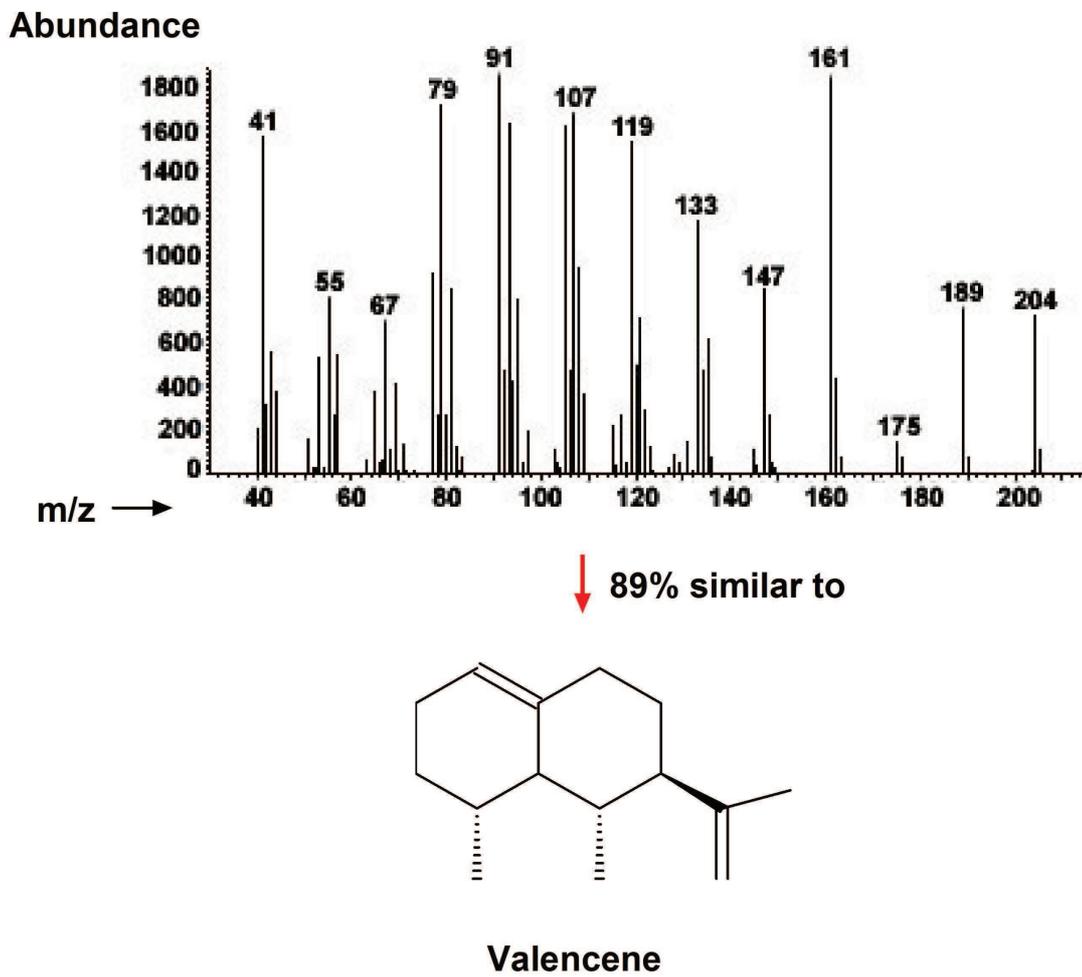


Figure 3.22: Mass spectra of terpene cyclase product from FPP

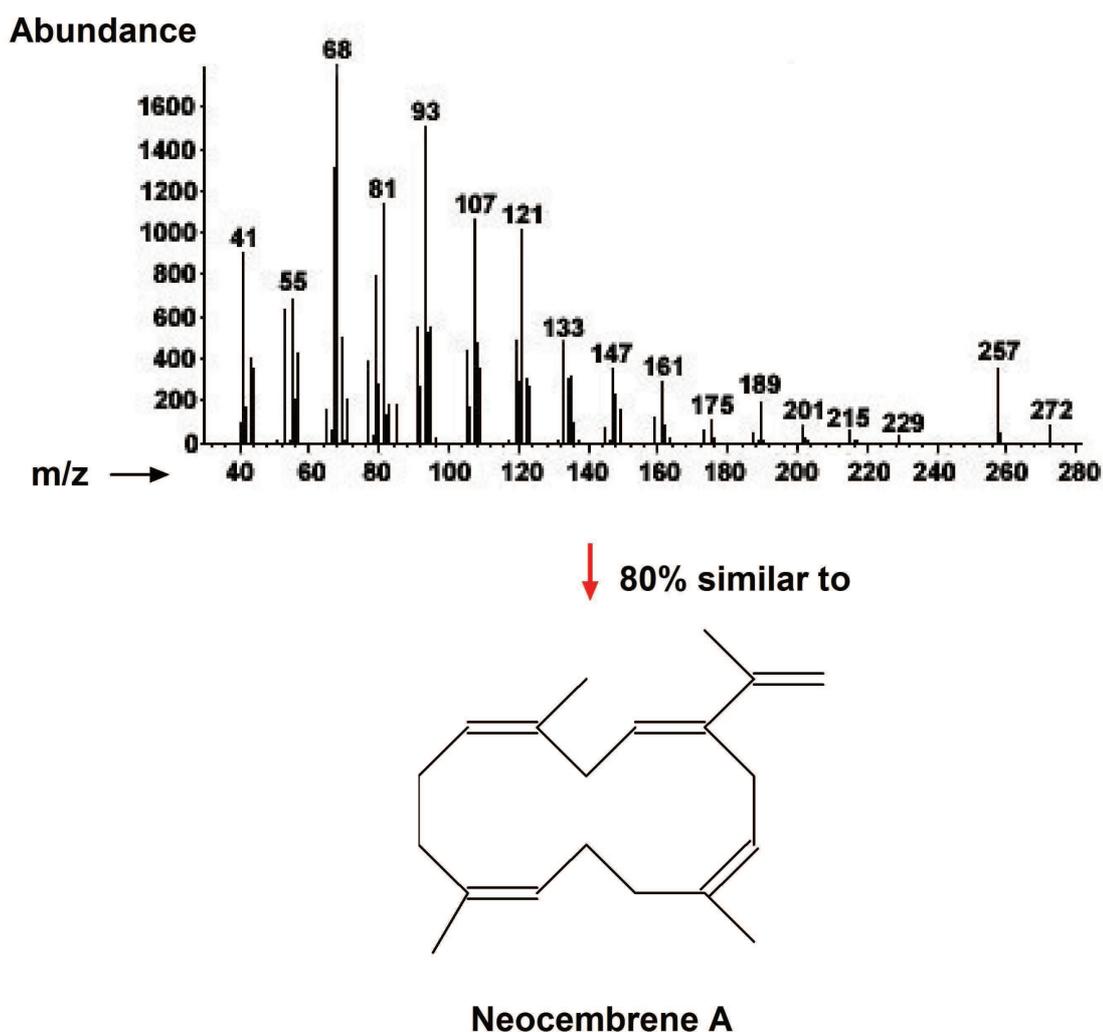
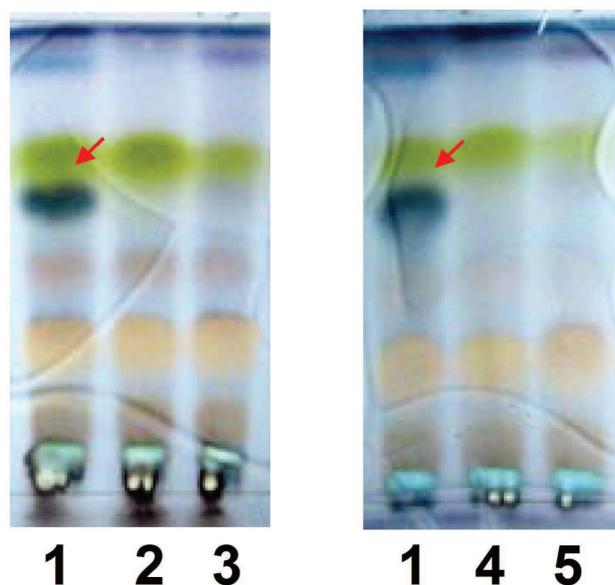


Figure 3.23: Mass spectra of terpene cyclase product from GGPP

### 3.3. Identification of products of the *TPSgc* encoding enzymes

#### 3.3.1. Product formation in the *in vitro* conversions

To study the genuine product formed by the enzymes of the *TPSgc*, CYP264B1 and GeoA were incubated in a reaction mixture containing substrate (FPP or GGPP) and all necessary components as described in the section 2.2.7.3. After incubating at 30°C for 1 hour in a thermomixer, reactions were stopped by adding chloroform, extracted twice and dried in a vacuum-drier. The samples were dissolved in 100  $\mu$ l of ethyl acetate and spotted on TLC using n-hexan/ethyl acetate (3/2) as developing solvent. After staining with vanilline, the results showed that CYP264B1 was able to convert the sesquiterpene product of GeoA into one product, which has a  $R_f$  of 0.63 (Figure 3.24).



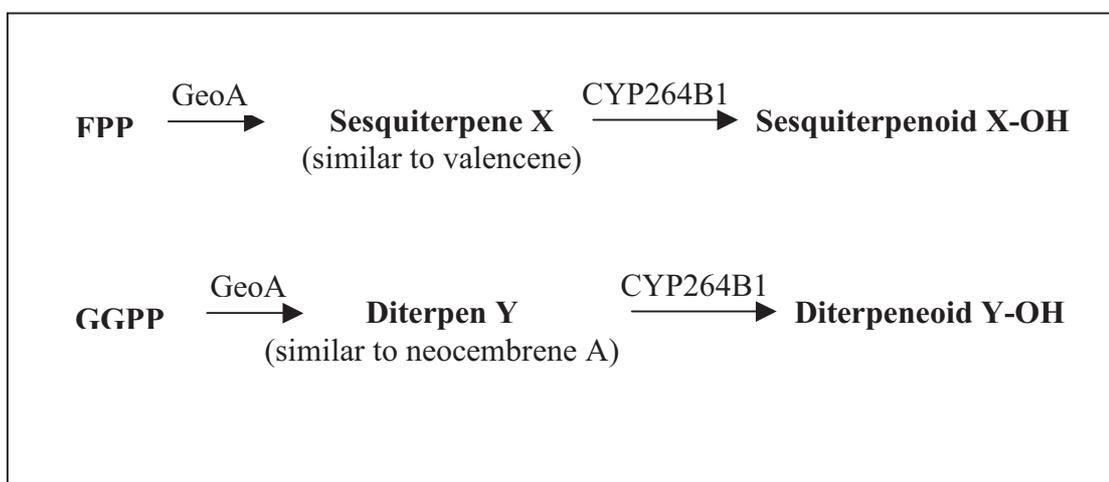
**Figure 3.24: TLC chromatogram of FPP conversion by the *TPSgc* encoding enzymes.** Lane 1 - FPP conversion; Lane 2 - negative control with boiled CYP264B1; Lane 3 - negative control with boiled GeoA; Lane 4 - negative control without substrate; Lane 5 - negative control without NADPH. Arrow represents the product of the *TPSgc* encoding enzymes

Similarly, CYP264B1 was able to convert the diterpene product of GeoA into one product, which has a  $R_f$  of 0.95 (n-hexan/ethyl acetate: 3/7) (Figure 3.25). In the negative controls (with boiled CYP254B1 or with boiled GeoA or without substrate or without NADPH) no products were detected. These results suggest that GeoA produced a sesquiterpene and a diterpene from FPP and GGPP, respectively, which can both serve as substrates for CYP264B1. The postulated route of terpene biosynthesis from prenyl pyrophosphate precursors by the *TPSgc* encoding enzymes CYP264B1-GeoA is illustrated in Figure 3.26.



**Figure 3.25: TLC chromatogram of GGPP conversion by the *TPSgc* encoding enzymes.** Lane 1 - negative control without NADPH; Lane 2 - GGPP conversion; Lane 3 - negative control with boiled CYP264B1; Lane 4 - negative control with boiled GeoA. Arrow represents the product of the *TPSgc* encoding enzymes

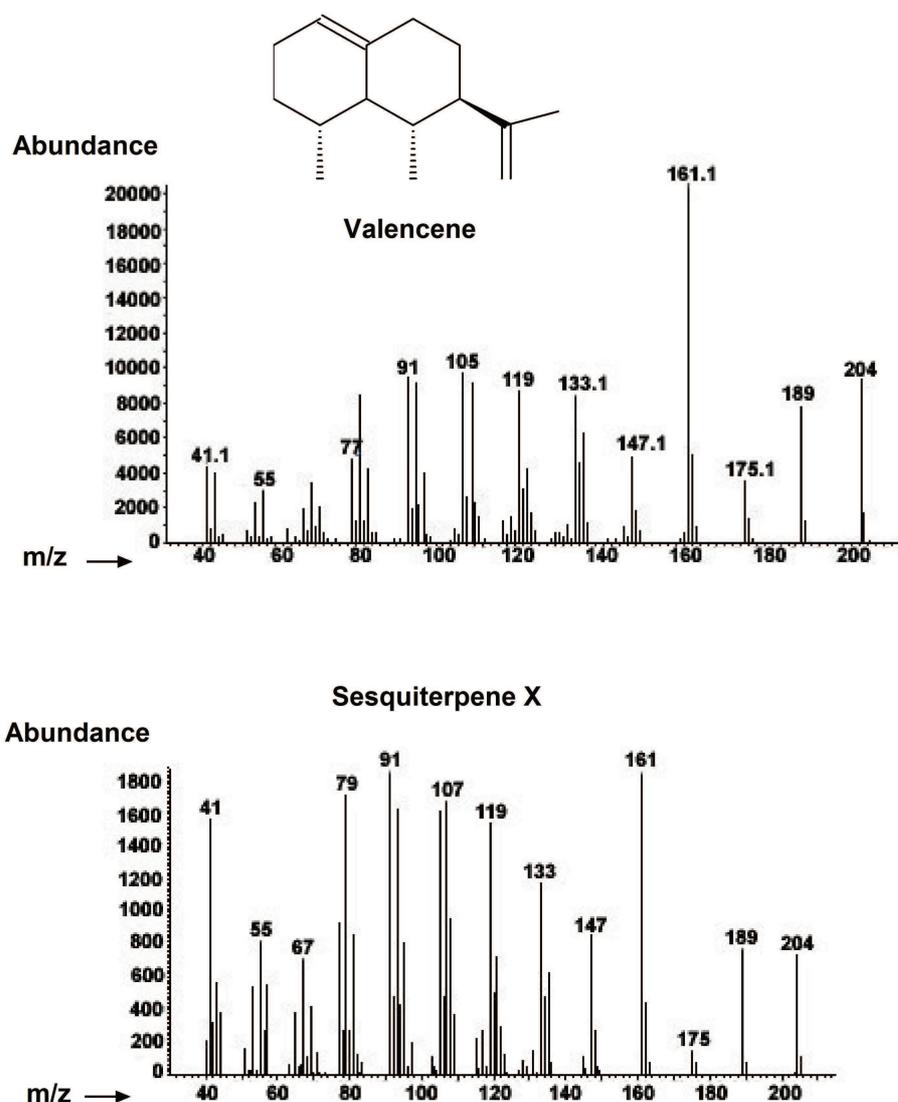
From these results, the routes of the biosynthesis of sesquiterpene X from FPP and diterpene Y from GGPP by the *TPSgc* encoding enzymes were postulated as shown in Figure 3.26.



**Figure 3.26: Postulated routes of terpene biosynthesis by the *TPSgc* encoding enzymes from FPP and GGPP**

### 3.3.2. Comparison of sesquiterpene product X with valencene

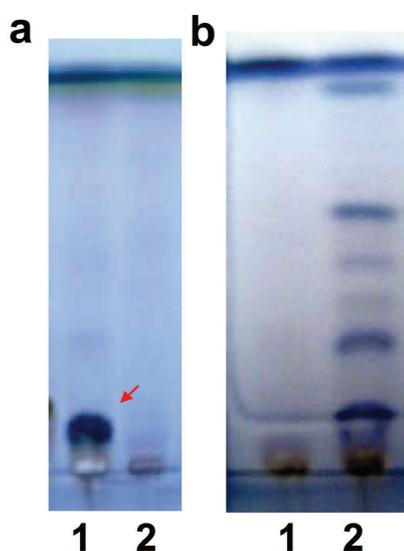
As mentioned above, GeoA converted FPP into a sesquiterpene X, which has 89% similarity to valencene. GC-MS showed that valencene and this sesquiterpene have a similar retention time of 20 min at the same conditions (data not shown); also the mass spectra of these two compounds are similar but are not identical (Figure 3.27).



**Figure 3.27:** Mass spectra of valencene and the sesquiterpene product of terpene cyclase

In order to confirm whether the sesquiterpene X is valencene or not, we compared the conversion of sesquiterpene X with the conversion of valencene by CYP264B1. Reactions were performed as described in the section 2.2.7.3. Analysis with TLC showed that CYP264B1 converted valencene into at least four different products whereas it converted the sesquiterpene X into one specific product (Figure 3.28).

Therefore, it must be concluded that the sesquiterpene product of GeoA is not valencene, but a new compound not present in the database of National Institute of Standards and Technology (NIST).

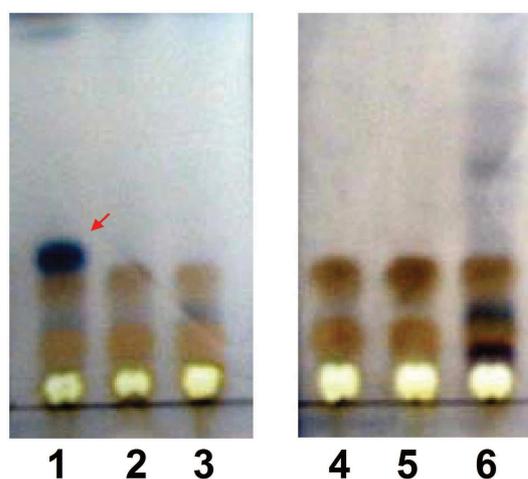


**Figure 3.28: TLC chromatogram.** Sesquiterpene product conversion by CYP264B1 (a) Lane 1 - FPP conversion by terpene cyclase and CYP264B1; Lane 2 - negative control without NADPH. Valencene conversion by CYP264B1 (b) Lane 1 - negative control without NADPH; Lane 2 - valencene conversion. The developing solvent containing n-hexan/ethylacetat (4/1) was used. Arrow represents the product of the *TPSgc* encoding enzymes

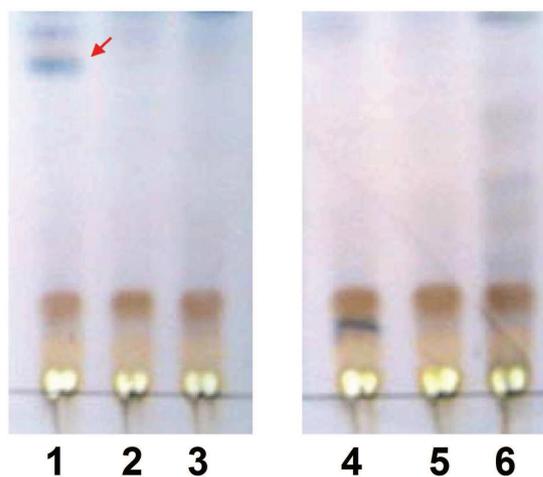
### 3.3.3. Conversion of the products of GeoA by different CYPs of *So ce56*

The result above showed that CYP264B1 was able to convert both the sesquiterpene X and the diterpene product Y of terpene cyclase. In order to determine whether the products of the terpene cyclase *sce8552* are also converted by other CYPs of the *So ce56* or if they are exclusively used by CYP264B1 as substrates, five additional *So ce56* cytochrome P450s (CYP264A1, CYP109C2, CYP109C1, CYP109D1, CYP260A1) were selected for *in vitro* conversion assays based on their close phylogenetic relationship to CYP264B1 (these CYPs were kindly provided by Dr. Yogan Khatri). These enzymes were known to be supported by Adx and AdR (Khatri PhD Thesis 2009). The reconstituted systems of these CYPs were incubated with FPP or GGPP together with GeoA as described in the section 2.2.7.3. Interestingly, 4 out of 5 CYPs (CYP264A1, CYP109C2, CYP109C1, CYP109D1) were not able to convert the products of GeoA (Figure 3.29 and Figure 3.30). Only CYP260A1 was able to convert

the products of GeoA but less specific than CYP264B1, since at least 2 products were observed, neither of which being identical to the product of CYP264B1 (Figure 3.29 and Figure 3.30). This result demonstrates that the sesquiterpenoid X-OH and diterpenoid Y-OH are the product of terpene cyclase and only one cytochrome P450 of So ce56 – CYP264B1.



**Figure 3.29: TLC chromatogram of the conversion of FPP by GeoA and different cytochromes P450 of So ce56.** Lane 1 - CYP264B1; Lane 2 - CYP264A1; Lane 3 - CYP109C2; Lane 4 - CYP109C1; Lane 5 - CYP109D1, Lane 6 - CYP260A1. The developing solvent containing n-hexan/ethylacetat (5/2) was used. Arrow represents the product of the *TPSgc* encoding enzymes



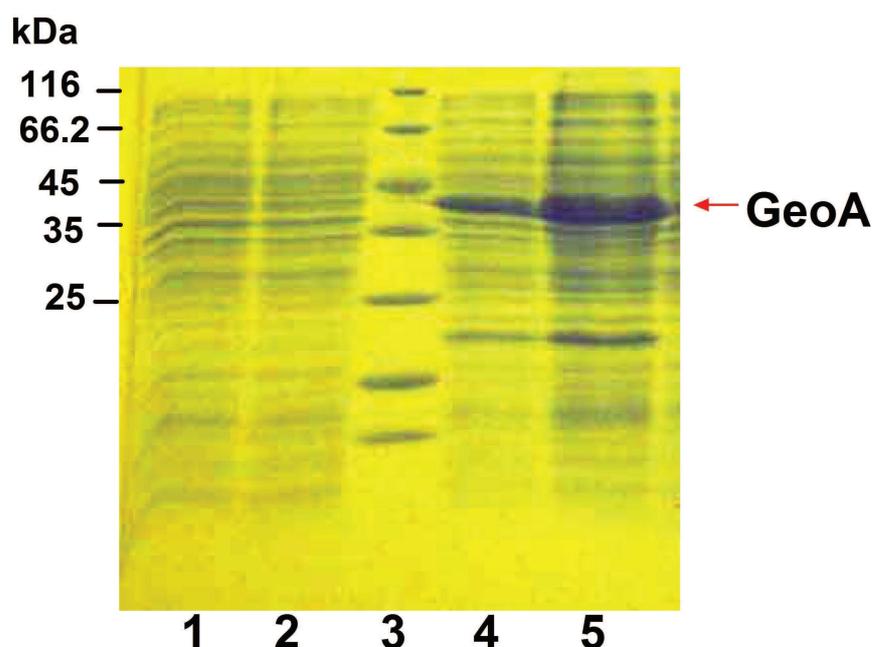
**Figure 3.30: TLC chromatogram of the conversion of GGPP by GeoA and different cytochrome P450s of So ce56.** Lane 1 - CYP264B1; Lane 2 - CYP264A1; Lane 3 - CYP109C2; Lane 4 - CYP109C1; Lane 5 - CYP109D1; Lane 6 - CYP260A1. The developing solvent containing n-hexan/ethylacetat (3/2) was used. Arrow represents the product of the *TPSgc* encoding enzymes

### 3.4. Investigation of terpenoid production in engineered *E. coli*

As described above, the terpene cyclase GeoA converted FPP and GGPP into a sesquiterpene X and a diterpene Y, respectively. The mass spectra of these products did not match with any compound included yet in the National Institute of Standards and Technology (NIST) database. Therefore, they are likely to be novel metabolites and, it is necessary to identify them by NMR method. To produce a sufficient amount of these compounds (sesquiterpene X and X-OH and diterpene Y and Y-OH) for analysis, the *in vitro* conversion requires 4 purified proteins (terpene cyclase, CYP264B1, Adx and AdR) and the expensive cofactor NADPH as well as the very expensive precursors FPP and GGPP ( $\approx 2600 \text{ € /10 mg}$ ). Thus, an *in vitro* conversion would not be economical, even if the terpenoids produced should prove to be of e.g. pharmaceutical interest. Due to such considerations, metabolically engineered microorganisms are nowadays widely used as host for the heterologous production of isoprenoid compounds (Martin 2003). In this work, we used *E. coli* for this purpose, because these bacteria offer advantages in terms of higher growth rate and easy genetic manipulation. Since *E. coli* produce a very limited amount of FPP through their native MEP pathway, a heterologous isoprenoid (MEV pathway) was introduced into this host to enhance isoprenoid flux for FPP production (Harada *et al.* 2009). We introduced the genes encoding the mevalonate-dependent isoprenoid pathway originating from *Streptomyces* sp. strain CL190 (vector pAC-Mev or pAC-Mv) and expressed them in *E. coli* (these vector were kindly provided by Prof Norihiko Misawa, Research Institute for Bioresources and Biotechnology, Ishikawa Prefectural University, Japan).

#### 3.4.1. *E. coli* strain and culture medium

Since protein expression of both CYP264B1 and GeoA require the co-expression with the chaperones GroEL and GroES, production of terpene compounds in this strains become difficult when introducing the mevalonate vector as a third vector. Therefore, C43(DE3), a mutant of the previously used of *E. coli* BL21(DE3), was tested for the expression of CYP264B1 and GeoA. This strain is often used to overcome the toxicity when recombinant proteins are over-expressed using the bacteriophage T7 RNA polymerase expression system (Dumon-Seignovert L *et al.* 2004). In addition, LB and TB culture media were also tested because the carbon source is known as a factor, which could have an influence on the protein expression level (Lee *et al.* 2004). The protein expression from crude extract was analyzed by SDS-PAGE (Figure 3.31).



**Figure 3.31: SDS-PAGE of the expression of GeoA in *E. coli* C43(DE3).** Lane 1 - cell lysate at 0 h of expression; Lane 2 - cell lysate after 20 h of expression in LB medium; Lane 3 - protein marker (pEQLab); Lane 4 - cell lysate after 44 h of expression in TB medium; Lane 5 - cell lysate after 20 h of expression in TB medium

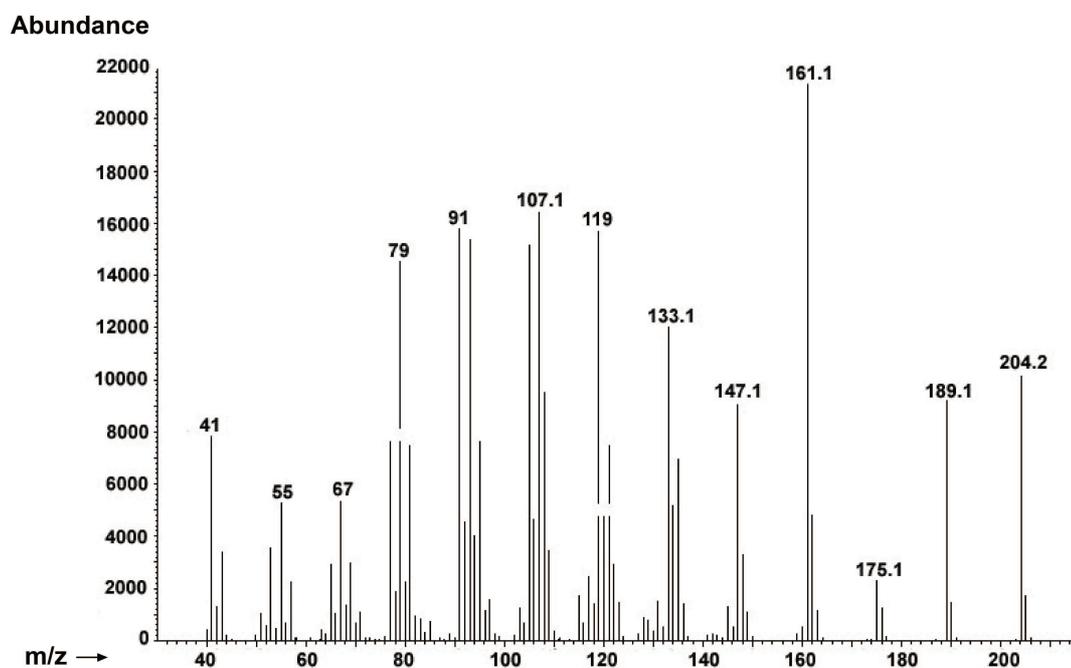
As shown in Figure 3.31, bands corresponding to GeoA were observed 20 hours after induction in TB medium. In contrast, no expression was observed when LB medium was used. The concentration of CYP264B1 determined by the CO difference spectra method was 513 nmol per litre culture. For this reason, the C43(DE3) strain and TB medium were used for the production of terpene compounds.

#### 3.4.2. Terpene production with GeoA

To produce the products of GeoA, *E. coli* C43(DE3) carrying plasmids pETG and pAC-Mev (or pAC-Mv) were cultured in 30 ml TB medium as described in section 3.2.10.1. Since the products of terpene cyclase are often volatile, the culture was overlaid with 10 ml decane to capture the released compounds. After 48 hours of incubation at 27°C, 100 µl of the overlay-decane phase was sampled for GCMS analysis. The result showed that there was a sesquiterpene present in the decane phase (data not shown).

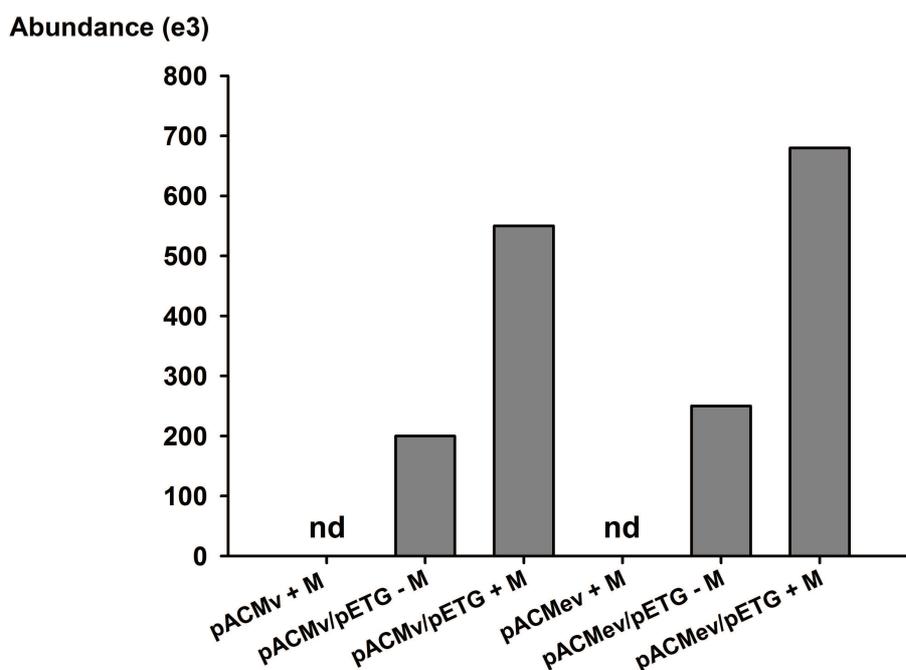
The mass spectrum of the detected sesquiterpene was compared to that of sesquiterpene X found in the *in vitro* conversion of FPP. The relative ion abundances for the sesquiterpene produced *in vivo* correlated with the values of sesquiterpene X for all ion fragments: 41, 55, 77, 91, 107, 120, 133, 147, 161, 175, 189, 204 m/z ions (Figure 3.32). In contrast, a control extract prepared from *E. coli* C43(DE3) transformed with

only pAC-Mev (or pAC-Mv) did not give any products, confirming the fact that *E. coli* does not produce any monoterpene, sesquiterpene or diterpene compounds (Connolly and Winkler 1989; Apfel *et al.* 1999). The results obtained here indicated that the sesquiterpene X product of GeoA was produced in the engineered *E. coli* strain.



**Figure 3.32: GC-MS of the detected sesquiterpene in decane phase overlaid the culture**

When mevalonate (0.5 mg/ml) was supplied to the medium, sesquiterpene X production was improved 2.5 fold (Figure 3.33). Since *E. coli* is not able to utilize mevalonate without the introduction of the foreign mevalonate pathway gene (Yoon *et al.* 2007), this result indicates that the introduced mevalonate pathway was active in this strain. Sesquiterpene X yield was estimated using valencene as external standard in GCMS. Recombinant *E. coli* harboring the complete mevalonate pathway (pAC-Mev) and pETG produced 0.5 mg of sesquiterpene X per 50 mL *E. coli* culture after 48 h, a yield 1.2-fold higher than that produced by the strain harboring the “bottom” pathway (pAC-Mv) and pETG (Figure 3.33).

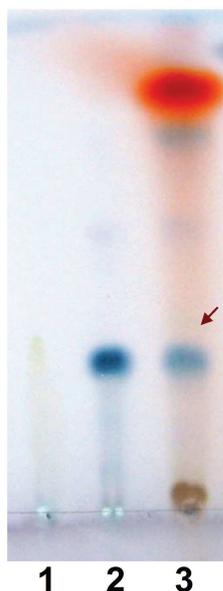


**Figure 3.33: Sesquiterpene X production by *E. coli* C43(DE3) harboring different plasmids.** “+M” or “-M” indicates “with addition of” or “without addition of” D-mevalonolactone, respectively; nd - not detected

### 3.4.3. Terpene production with CYP264B1 and GeoA

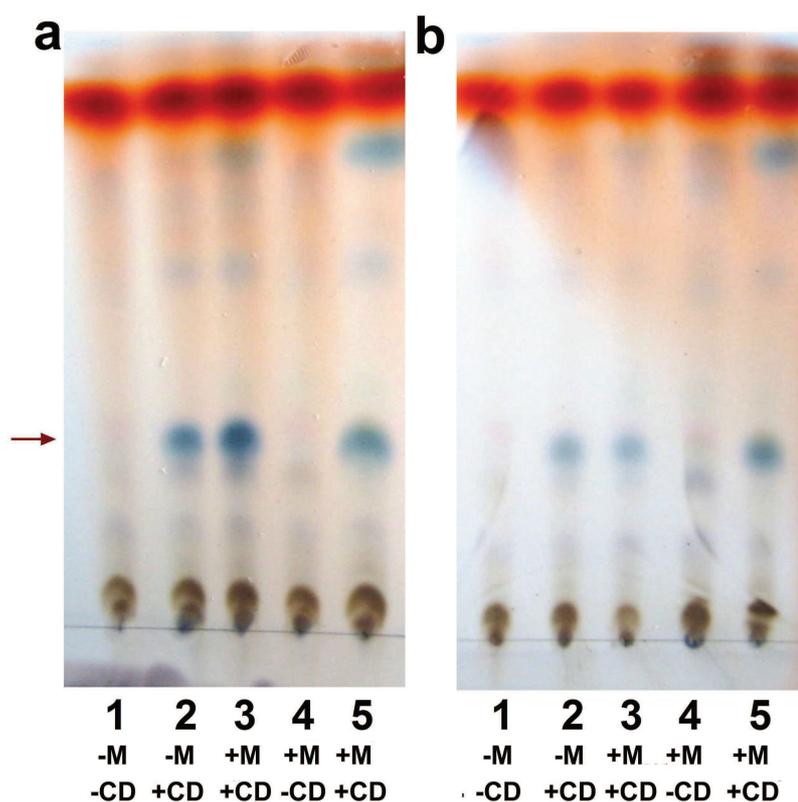
In order to produce product of the *TPSgc* encoding enzymes, *E. coli* C43(DE3) carrying plasmids pETC4AAG and pAC-Mev or plasmids pAC-MvC4 and pETGAA were cultured in TB medium; at the time of induction with IPTG, 0.5 mg/ml mevalonate (M) was added to the culture as described above. In this setting, besides the terpene cyclase GeoA and the enzymes of the mevalonate pathway, also CYP264B1 and its redox partners were expressed. To analyze the *in vivo* conversion, aliquots of 0.5 ml culture were centrifuged to remove the cells. Supernatant was extracted with chloroform and analyzed by TLC. The result showed that no product was detected in this case (data not shown). It was discussed by Yu and coworkers (Yu *et al.* 2011) in their work that the low yield of hydroxylated sesquiterpene in *E. coli* might be due to the inefficient transport of hydrophobic sesquiterpene from the cytosol to the hydroxylation sites. Therefore, we used hydroxypropyl- $\beta$ -cyclodextrin (CD) with the hope that it will be able to capture the released volatile sesquiterpene. This compound is known as an enhancer of the permeation of poorly soluble compounds through biological membranes (Másson *et al.* 1999). CD was added into the culture to get a final concentration of 1.13% (w/v) at the time of induction. After 48 hours, the terpenoid production was

analyzed by TLC. Interestingly, when CD was added to the culture, a product was detected which has an retention factor ( $R_f$  of 0.33) identical to the one of the product of FPP *in vitro* conversion by the *TPSgc* encoding enzymes CYP264B1-GeoA (Figure 3.34).



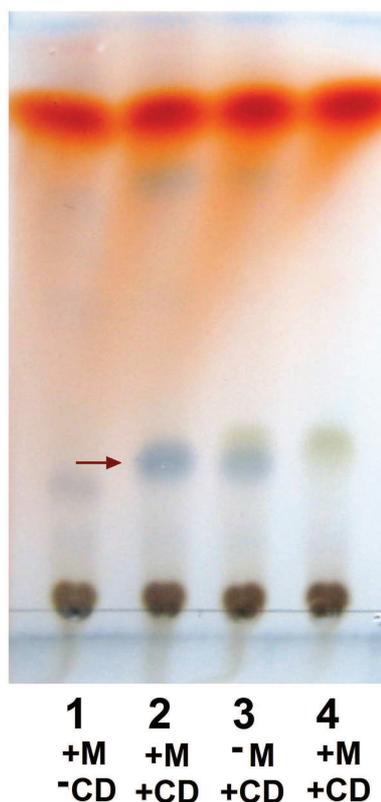
**Figure 3.34: TLC chromatogram of FPP conversion.** The *in vitro* conversion of FPP by the *TPSgc* encoding enzymes (Lane 2) and terpene production of *E. coli* C43(DE3) harboring pAC-MvC4/pETGAA (Lane 3); Lane 1 - negative control of *in vitro* FPP conversion without NADPH. The developing solvent containing n-hexan/ethylacetat (5/2) was used. Arrow represents the putative product of CYP264B1 and GeoA produced by *E. coli* C43(DE3)

In order to confirm whether this product was formed by the *TPSgc* encoding enzymes or not, different vector combinations were transformed and expressed in *E. coli* C43(DE3). The effects of presence or absence of hydroxypropyl- $\beta$ -cyclodextrin and D-mevalonolactone were also tested. As shown in Figure 3.35, the product with a  $R_f$  0.33 was not present in the cultures harboring pETC4AAG/pAC-Mev and pAC-MvC4/pETGAA when neither hydroxypropyl- $\beta$ -cyclodextrin nor D-mevalonolactone were added. In contrast, this product was observed in the cultures supplemented with hydroxypropyl- $\beta$ -cyclodextrin, irrespective of the addition of D-mevalonolactone.



**Figure 3.35: TLC chromatogram analysis of terpene production of *E. coli* C43(DE3) harboring different vector.** (a) Lane 1, 2, 3 and 4 - pETC4AAG/ pAC-Mev; Lane 5 - pAC-MvC4/pETGAA. (b) Lane 1, 2, 3 and 4 - pAC-MvC4/pETGAA; Lane 5 - pETC4AAG/pAC-Mev. (+) - addition of; (-) - without addition of; M - D-mevalonolactone; CD - hydroxypropyl- $\beta$ -cyclodextrin. The developing solvent containing n-hexan/ethylacetat (5/2) was used. Arrow indicated the retention factor of the putative product of CYP264B1 and GeoA produced by *E. coli* C43(DE3)

In addition, when pETC4AAG was transformed alone and the heterologous mevalonate pathway was not introduced in the host, the product was also detected in the culture supplemented both M and CD (data not shown). However, the product was not observed in the culture harboring pAC-MvC4 even when both M and DC were supplemented (Figure 3.36).



**Figure 3.36: TLC chromatogram analysis of terpene production of *E. coli* C43(DE3) harboring different vector.** Lane 1, 2, 3 - pAC-MvC4/pETGAA; Lane 4 - pAC-MvC4. The developing solvent containing n-hexan/ethylacetat (7/2) was used. Arrow represent the putative product of CYP264B1 and GeoA produced by *E. coli* C43(DE3)

Taken together, it can be stated that the mevalonate engineered *E. coli* C43(DE3) was able to produce sesquiterpene X when the gene coding for the terpene cyclase GeoA was introduced into this host. For production of the hydroxylated product X-OH of the *TPSgc* encoding enzymes, hydroxypropyl- $\beta$ -cyclodextrin might be used to overcome the loss of the volatile sesquiterpene X from the culture, so that the yield of the hydroxylated product might be improved. However, to confirm the hydroxylated product, mass spectra analysis for this experiment is necessary.

## 4. DISCUSSION

### 4.1. Characterization of CYP264B1

P450s are versatile biocatalysts and their regio- and stereo-selective activities are particularly attractive for biotechnological processes exploiting new substances of industrial value (Bernhardt 2006). In recent years, genome sequencing projects disclosed information on many new P450 genes which provides a valuable basis for potential industrial applications. In 2007, the whole genome sequence of the myxobacterium *Sorangium cellulosum* So ce56 was published, revealing 21 novel putative P450s. The strain So ce56 possesses the largest genome discovered in bacteria to date and produces several bioactive compounds including chivosazol, etnangien, and myxochelin (Schneiker *et al.* 2007; Wenzel and Muller 2009). Characterization of P450s of this strain therefore promises an interesting contribution for P450-centered biotechnology.

This study describes the expression and characterization of CYP264B1, one of 21 P450s of So ce56. In a previous work in our laboratory, the gene encoding CYP264B1 was cloned into the *E. coli* expression vector pCWori+ and expressed in *E. coli* BL21, but the expression level was very low (only 4 nmol/L culture). In order to improve the expression, the gene was cloned into pET17b and expressed in *E. coli* BL21(DE3), but the yield was still not improved (Khatri PhD thesis 2009). Since the high GC content of target genes is biased against *E. coli* codon usage (Redwan 2006), it could influence the protein expression of So ce56's P450s in heterologous hosts such as *E. coli*. Several reports have shown the improvement of protein expression from high GC content genes in *E. coli* by codon optimization (Panahi *et al.* 2004; Srivastava 2005). However, it was shown that among the 21 So ce56 P450s, several were expressed at high levels (more than 1000 nmol/L culture) in *E. coli* without changing the nucleotide sequence, including CYP109C1, CYP109C2 and CYP266A1. In addition, it was also reported that the elimination of rare codons does not increase the expression of the Rv3297 gene in *E. coli* (Guo *et al.* 2009). The reason for the difference in expression levels of So ce56 P450 genes remains elusive since a comparison of the occurrence of *E. coli* rare codons with these genes disclosed no correlation of the frequency of rare codons and the expression level observed in *E. coli* (Khatri PhD thesis 2009). It seems that other reasons, rather than codon bias caused by high GC content, are responsible for the low

protein expression in heterologous host *E. coli*. In their study, Han *et al.* (Han *et al.* 2010) pointed out that the high GC content is likely responsible for many stable stem structures on the mRNA of human peptide deformylase, which influence the transcription or translation of the N-terminal-tagged gene and inhibits the protein expression. The high GC content present on the translation initiation region is especially showing this effect (Guo *et al.* 2009; Gu *et al.* 2010). According to Francis and Page (Francis and Page 2010), protein expression in *E. coli* requires four elements: (1) the protein of interest, (2) a bacterial expression vector, (3) an expression cell line, and (4) the equipment/materials for bacterial cell culture. Thus, if genes coding of all 21 *So ce56*'s P450s were cloned into the same pCWori<sup>+</sup> vector and expressed under the same conditions, but the expression levels of different protein were not the same, it is because of the intrinsic properties of each protein itself (e.g. hydrophobic residues, complexity regions or disulfide bonds). When these P450s are expressed in heterologous hosts, where the internal microenvironment differs from that of *So ce56* (e.g. pH, osmolarity), the interaction of these proteins with new environments and mechanisms for folding are different, leading to differences in the expression level. Therefore, in this context, simple elements, such as expression temperature, should be first investigated. Lowering the expression temperature is often suggested, since at lower temperatures cell processes slow down leading to reduced rates of transcription, translation, and cell division (Chou 2007), while also leading to reduced aggregation and improved solubility of recombinantly expressed proteins (Shirano and Shibata 1990; Sahdev *et al.* 2008; Volonte *et al.* 2008). In addition, this also results in a reduction in the degradation of proteolytically sensitive proteins since most proteases are less active at lower temperatures (Francis and Page 2010; Hunke and Betton, 2003; Pinsach *et al.* 2008). Moreover, if high GC content is the reason as discussed above, lowering the temperature also can help because prokaryotes living in colder environments tend to have comparatively less stable mRNA secondary structure at the translation-initiation region (Gu *et al.* 2010). In our work, when lowering the expression temperature to 20°C, the yield of CYP264B1 was improved from 4 nmol to 257 nmol/L culture. However, this yield is still rather low and scarcely sufficient to obtain enough protein for further studies. The molecular chaperone GroEL of *E. coli* is known to be essential for the folding of many proteins *in vivo* and *in vitro* (Ellis 1991; Landry and Gierasch 1994); also the cochaperonin GroES, a member of the heptameric cpn10 family, is an essential additional component for the GroEL-mediated folding of some proteins *in vitro*

(Chandrasekhar *et al.* 1986). Therefore, the pET17b vector carrying the CYP264B1 gene was co-transformed with the plasmid pGro12 carrying the molecular chaperones GroES and GroEL into *E. coli* BL21(DE3). As expected, the protein expression level improved to 820 nmol/L culture which was sufficient for the characterization of CYP264B1. Spectrophotometric characterization of the carbon monoxide bound form of CYP264B1 gave a typical peak maximum at 450 nm and a dominating shoulder at 420 nm.

Since most P450s perform their chemical reactions after receiving reduction equivalents from NADPH or NADH through autologous or heterologous electron transfer partners (Hannemann *et al.* 2007), we sought to identify redox partners for CYP264B1. The genome of *S. cellulosum* So ce56 contains 8 ferredoxins (Fdx) and 2 ferredoxin reductases (FdR) genes, but none of them is located in the vicinity of a P450 gene (Ewen *et al.* 2009). Therefore, the prediction of natural redox partners for CYP264B1 was not possible using solely genomic information. Thus, the available myxobacterial redox proteins as well as the heterologous adrenodoxin (Adx<sub>4-108</sub>) and adrenodoxin reductase (AdR) from bovine adrenals were tested as potential redox partners for CYP264B1 based on the CO difference spectra method (Omura and Sato 1964). The result showed that the electron transfer capability of the heterologous redox proteins Adx-AdR to CYP264B1 was much better than the one of Fdx2-FdRB and Fdx8-FdRB (see section 3.1.3). This observation is in accordance with results obtained with CYP109D1, CYP264A1, and CYP266A1 (Khatri *et al.* 2010) and CYP260A1 (Ewen *et al.* 2009) from this bacterium, which also showed higher activity with Adx and AdR compared with the redox pairs Fdx2/FdRB or Fdx8/FdRB. Therefore, Adx and AdR were selected as redox partners of CYP264B1.

In the genome of *So ce56*, the CYP264B1 gene is clustered with a putative terpene cyclase gene with only 63 base pairs separating the two genes. Thus it seems likely that CYP264B1 is involved in terpene biosynthesis. While we tried to identify and produce the natural substrate of CYP264B1 (see below), we also decided to perform a screening with diverse terpenoid compounds and substances with similar structures as potential substrates for CYP264B1. In this work, seven sesquiterpenoids (nootkatone, valencene,  $\alpha$ -longipinene, clovene, humulene, isolongifolene-9-one and copaene), two norisoprenoids ( $\alpha$ -ionone and  $\beta$ -ionone) as well as ten compounds belonging to other groups (monoterpene, diterpene acid, steroid, keton and aroma related compounds) were tested. Among 22 investigated compounds, 14 compounds induced a type I spectrum of

CYP264B1 (see section 3.1.4) and were therefore considered as potential substrates. An *in vitro* conversion by CYP264B1 using Adx and AdR as redox partners was carried out for these 14 potential substrates. Interestingly, among them 8 compounds belonging to the sesquiterpenes and norisoprenoids were found to be converted by CYP264B1 (see section 3.1.6). Among the identified substrates, we focused on nootkatone and  $\alpha$ -/ $\beta$ -ionone because of the high selectivity of the CYP264B1-catalyzed hydroxylation of these substances. Thus the products of  $\alpha$ -/ $\beta$ -ionone and nootkatone conversions were further characterized.

Ionones and their derivatives are important constituents of essential oils, which are found in flowers, fruits, and leaves of many plants (Mikami 1981; Ohloff 1978). Since ionones are known as interesting substances for fragrance and flavor industry (Brenna 2002), over the last decades the regio-selective and/or enantio-selective conversion of ionones has been investigated with growing interest. However, it has been shown that the bio-transformations of  $\alpha$ -/ $\beta$ -ionone by using microorganisms (Mikami *et al.* 1981; Larroche *et al.* 1995; Lutz-Wahl *et al.* 1998; Urlacher *et al.* 2006) or even selected P450s (Celik *et al.* 2005; Girhard 2010) often resulted in mixtures of hydroxyl- and/or oxo-derivatives, which is not desirable for industrial application. In contrast with CYP264B1, we observed a highly selective hydroxylation of  $\alpha$ - and  $\beta$ -ionone resulting in the respective 3-OH products; these products might be important and might find an industrial application as intermediates for the synthesis of carotenoids, e.g. of astaxanthin and zeaxanthin (Loeber 1971), or deoxyabscisic acid, a synthetic analogue of the phytohormone abscisic acid (Larroche *et al.* 1995). Because of the potential industrial application of 3-OH ionone derivatives, several biotransformation studies were described in the literature, including the screening of 215 *Streptomyces* strains for the regio- and stereo-selective hydroxylation of  $\alpha$ -/ $\beta$ -ionone. However, none of them converted  $\beta$ -ionone into 3-hydroxy- $\beta$ -ionone (Lutz-Wahl *et al.* 1998). Even though we have demonstrated recently that CYP109D1 was able to convert  $\beta$ -ionone regioselectively, the 3-OH product of  $\beta$ -ionone has not been obtained so far (Khatri *et al.* 2010). Instead, 4-hydroxy- $\beta$ -ionone was found. Thus, it seems that C-3 of  $\alpha$ -ionone and C-4 of  $\beta$ -ionone were the most frequent positions of biooxidation systems for the hydroxylations. In terms of stereochemistry, C-3 of  $\alpha$ -ionone is often a favored position for hydroxylation because of the presence of two methyl groups at C-1 position that direct any oxidative attack toward C-3. Likewise, the C-4 position of  $\beta$ -ionone is also

favored because of the electronic activation of the allylic hydrogens by the double bond of the cyclohexane ring, which governs the regio-selective hydroxylation at this position. Nevertheless, in a recent publication, the conversion of  $\beta$ -ionone by CYP101B1 and CYP101C1 from *Novosphingobium aromaticivorans* was reported to result in a 3-hydroxylated product albeit this product was found in a mixture with 4-hydroxy- $\beta$ -ionone (Bell *et al.* 2010). In contrast to the *Novosphingobium* CYPs, CYP264B1 from *S. cellulosum* So ce56 is able to hydroxylate both  $\alpha$ - and  $\beta$ -ionone in a highly regio-selective manner at C-3 position, giving 3-hydroxy- $\alpha$ -ionone and 3-hydroxy- $\beta$ -ionone, respectively. This finding of a regio-selective hydroxylation of  $\alpha$ - and  $\beta$ -ionone by CYP264B1 at the C-3 position implies a very rigid orientation of the trimethylcyclohexane ring within the enzyme's active site. To gain more insight into the structural basis for the selectivity of hydroxylation, both,  $\alpha$ - and  $\beta$ -ionone, were docked into the active site of CYP264B1. As shown in Fig. 3.13 (section 3.1.8), the optimal orientation allowing a suitable arrangement for hydroxylation was only observed for the C-3 position. The computational data thus confirmed our experimental observation that CYP264B1 hydroxylates both  $\alpha$ - and  $\beta$ -ionone at the same position, C-3.

Nootkatone, the other substrate of choice, is also an interesting compound for detailed investigations because of its flavoring properties (Fraatz *et al.* 2009), pharmacological importance (Tassaneeyakul *et al.* 2000; Murase *et al.* 2010) and insecticidal activity (Miyazawa *et al.* 2000; Zhu *et al.* 2001; Panella *et al.* 2005). In this study, the hydroxylation position of nootkatone was identified to be at C-13. This makes the nootkatone product of CYP264B1 pharmaceutically highly interesting, since it was shown previously that 13-hydroxy-nootkatone displays an increased antiproliferative activity towards a cancer cell line compared to other hydroxylated products of nootkatone (Gliszczynska *et al.* 2011). Moreover, the presence of 13-hydroxy-nootkatone in the secrete of diseased wood of *Chamaecyparis lawsoniana* was suggested to play a role of this compound in controlling fungal infection (Takashi 2007). It is of special interest that although many biotransformations of nootkatone were investigated, a hydroxylation at C-13 was never observed so far (Furusawa *et al.* 2005; Borges 2009; Hegazy *et al.* 2006). This indicates that the selective hydroxylation of nootkatone at C-13 position seems to be difficult to obtain via biooxidation processes. Thus, CYP264B1 provides a novel and highly interesting system for the production of this attractive metabolite of nootkatone.

In summary, CYP264B1, a novel P450 from the myxobacterium *Sorangium cellulosum*, was successfully expressed, purified and characterized. In addition to two endogenous redox chains (FdRB-Fdx2 and FdRB-Fdx8) also the more efficient heterologous redox proteins constituting the electron transfer chain Adx-AdR were identified as interaction partners. Among the screened potential substrates, norisoprenoids ( $\alpha$ - and  $\beta$ -ionone) and sesquiterpenes were converted by CYP264B1. Since norisoprenoids can be considered as degraded sesquiterpenoids (Gomes-Carneiro *et al.* 2006), CYP264B1 is identified as a novel sesquiterpene hydroxylase.

#### 4.2. *E. coli* whole-cell biocatalyst for sesquiterpene conversion

Sesquiterpenes are the most diverse group of terpenoid compounds and common constituents of essential oils in plants (Fraga 2006; Edris 2007). Herbs and higher plants containing sesquiterpenoids have been exploited in traditional medicine since ancient times (Sauter 2011; Chu 2011; Galani 2010). The derivatives of the sesquiterpenes often possess desirable fragrances and flavoring or pharmaceutical properties (Fraga 2006). In addition, they can be used as starting materials to synthesize high-value compounds like amorpho-4,11-diene, the precursor for the potent antimalarial drug artemisinin, and germacrone, the starting compound in the synthesis of the anticancer drug  $\beta$ -elemene (Chang 2000; Barrero 2011). Since the great majority of terpenoids are biologically active, many of them have been modified to create novel substances with desired properties (Azerad 2000). The hydroxylation of sesquiterpenes, an essential step in the production of flavour molecules in flavour and fragrance industry, was often used to enhance their sensory properties (Faber 2000, Cankar *et al.* 2011). For this purpose, bio-oxidation systems are more advantageous in comparison to chemical synthesis by means of cost and time efficiency, as well as environmentally friendly habit. Furthermore, using oxidase enzymes such as cytochromes P450, the ability of introducing an atomic oxygen into allylic positions, double bonds or even into non-activated C–H bonds, which is difficult for chemical synthesis, can be achieved (Bernhardt 2006, Urlacher *et al.* 2004). As described above, CYP264B1 was identified as a novel sesquiterpene hydroxylase, which is able to convert diverse sesquiterpene substrates. According to the best of our knowledge, until now, besides CYP264B1, only one cytochrome P450, named premnaspirodiene oxygenase, from the plant *Hyoscyamus muticus* has been reported to be capable of hydroxylating diverse sesquiterpene substrates (Takahashi *et al.* 2007). However, the utilization of P450s from bacteria is more desirable than that

from plants and animals because they are not membrane associated and normally exhibit a higher stability and activity (Urlacher *et al.* 2004). Moreover, CYP264B1 displays interesting characteristics towards the hydroxylation of its substrates. It is able to hydroxylate these compounds at positions that are rarely targeted by other biooxidation systems known so far (see section 4.1). Thus, CYP264B1 is a promising candidate for novel biotechnological processes and applications, especially exploiting new sesquiterpene derivatives.

Due to the high costs of pyridine nucleotide cofactors as well as the laborious protein purification steps, it is not economic to use isolated enzymes for the production of these compounds in a large scale. In contrast, using a recombinant whole-cell biocatalyst is economically advantageous and very convenient, because not only the cofactor is provided by the host cell (it is regenerated simultaneously in the cellular metabolism), but also the P450 systems are stabilized (Maurer *et al.* 2005). The biggest challenge in the field of biocatalyst are such cofactor-dependent redox reactions (Zhao and van der Donk 2003; Urlacher *et al.* 2004; Hannemann 2006). To overcome this problem, several methods have been developed for the regeneration of cofactors using enzymes and whole cell based systems (Wandrey 2004; Wichmann and Vasic-Racki 2005). A recombinant *E. coli* whole-cell catalyst containing CYP106A2 and its redox partners was successfully applied for the conversion of 11-deoxycorticosterone (DOC) to 15 $\beta$ -hydroxy-DOC (Hannemann *et al.* 2006). Additionally, an efficient *E. coli* whole-cell catalyst by integrating heterologous cofactor regeneration, combining glucose facilitator and dehydrogenase, for the conversion of P450(BM-3)'s substrate was also reported (Schewe *et al.* 2008). Likewise, a new *Bacillus megaterium* whole-cell catalyst for the hydroxylation of the pentacyclic triterpene 11-keto- $\beta$ -boswellic acid has been successfully used (Bleif *et al.* 2011).

Since *E. coli* is a well studied bacterium and does not contain a P450 gene in its genome, the organism is a suitable host for the characterization and application of P450s (Agematu *et al.* 2006). We therefore created an *E. coli* whole cell system in which CYP264B1 was co-expressed with its heterologous redox partners AdR and Adx for the production of sesquiterpene derivatives. This system was tested for  $\beta$ -ionone conversion using *E. coli* C43(DE3) as host cell. Substrate was added into the culture medium 24 hours after induction. Then, after 24 or 48 hours, the compounds in the culture medium were extracted and monitored using a HPLC separation system. The result demonstrated that the specific product of CYP264B1 (3-hydroxy- $\beta$ -ionone) was

formed. However, besides this product, another product peak at later retention time was also found. In the control strain of *E. coli* carrying pET17b, pETMR5 or pETC4, only this side product was observed (data not shown). This result indicated that 3-hydroxy- $\beta$ -ionone is the sole product of CYP264B1 and that the side product is produced by other enzymes of this strain. As discussed above (see section 4.1),  $\beta$ -ionone could be transformed by various bio-systems into different derivatives, but not into 3-hydroxy- $\beta$ -ionone. Therefore, this is the first system for the conversion of  $\beta$ -ionone into 3-hydroxy- $\beta$ -ionone. The product yield of 3-hydroxy- $\beta$ -ionone was estimated to be 48 mg/L culture per day under non optimized laboratory conditions. In addition, the optimal time after induction for taking samples is around 24 h.

Taken together, an *E. coli* whole-cell system for the transformation of sesquiterpenes was constructed. CYP264B1 and its heterologous redox partners Adx and AdR were expressed functionally in C43(DE3). This system was successfully applied for the conversion of  $\beta$ -ionone and therefore provides an opportunity for the conversion of other substrates of CYP264B1, including valencene, clovene and humulene.

### 4.3. Characterization of GeoA and identification of the *TPSgc*

Terpenoids comprise a highly diverse class of natural products (Martin 2003) They play multi-functional roles in plants, human health and commerce (see introduction). The primary enzymes responsible for the synthesis of these compounds are terpene cyclases. Genome mining of the myxobacterium *So ce56* - a very promising source for the discovery of novel classes of secondary metabolites revealed three putative terpene cyclase genes (locus\_tag sce1440, sce6369 and sce8552). Among them, the terpene cyclase GeoA (locus\_tag sce8552) is located downstream of CYP264B1. The distance of 63 base pairs between them implies that they are likely arranged in an operon and together involved in the biosynthesis of terpenoids in *So ce56*. Due to the economic and biological significance of the terpenoids, we were interested in the characterization of the terpene cyclase GeoA, especially in the role of the *TPSgc* encoding enzymes CYP264B1-GeoA in the terpenoid biosynthesis of this myxobacterium.

The gene encoding for GeoA was cloned into vector pET17b and expressed in BL21(DE3). Similar to CYP264B1, the expression of GeoA in this strain required the support of the molecular chaperones GroEL and GroES. The explanation for this has been discussed for CYP264B1 (see section 4.1). To identify the substrate(s) of this

terpene cyclase, three common substrates of terpene cyclases (GPP, FPP, GGPP) were tested in the *in vitro* conversion. GC-MS analysis of the conversion of FPP and GGPP revealed a product peak at 20 and 30 min, respectively. This result indicated that FPP and GGPP were substrates of GeoA. In a similar manner to many terpene cyclases analyzed so far, GeoA required  $Mg^{2+}$  as a catalyst. No activity was detected when this divalent metal ion was absent. The important aspartate-rich region of all terpene synthase enzymes, the DDxxD, has been proved to be responsible for the binding of divalent metal ions that later react with the diphosphate moiety of the substrate (Degenhardt *et al.* 2009). This region was found to start at Asp-91 of GeoA. The conserved triad NDXXSXXXE motif was also found in this enzyme, beginning at Asn-231. This motif is believed to play a role as an additional metal cofactor binding motif located on the opposite site of the active site entry (Rynkiewicz *et al.* 2001). The deduced amino acid sequence of this terpene cyclase exhibits the highest identity to a monoterpene cyclase for the biosynthesis of the off-flavor terpenoid alcohol-2-methylisoborneol (2-MIB) of *Roseiflexus castenholzii* DSM 13941. Nevertheless, it is able to convert both FPP and GGPP into a sesquiterpene and diterpene, respectively, but not GPP into monoterpene. Thus, it can be stated that this enzyme belongs to both, the sesquiterpene and diterpene cyclase group. Similar characteristics were observed for pentalenene synthase (Cyc2) of *Streptomyces* strains *Kitasatospora griseola* - a terpenecin producer (Dairi 2005). GCMS analysis showed that GeoA converted FPP into a product which has 89% similarity with the sesquiterpene valencene. For the GGPP conversion, a product which has 80% similarity with the diterpene neocembrene A was found. According to the NIST database 2.0, the values above 80% are good matches. Therefore, valencene and neocembrene A are potential products of GeoA.

In the genome of *So ce56*, CYP264B1-coding gene is located upstream from *geoA*. For that reason, this P450 was believed to be responsible for modifying further the product of the terpene cyclase GeoA by creating specific functional groups on these compounds. Because CYP264B1 has been identified as a novel hydroxylase, it is suspected to hydroxylate the product of terpene cyclase GeoA. A terpenoid synthesis route through the *TPSgc* encoding enzymes CYP264B1-GeoA was thus postulated (see section 3.3.1; Figure 3.26). The first step is cyclization of FPP or GGPP by GeoA to sesquiterpene X or diterpene Y to establish the basic skeleton. The further step is oxygenation of these terpenes by CYP264B1 to the sesquiterpenoid X-OH or the diterpenoid Y-OH. To confirm this postulation, the CYP264B1 reconstitution system was added to the

incubation of GeoA with either FPP or GGPP. Under this reaction condition, the formation of the products X-OH and Y-OH was observed. On the other hand, these products were not formed by a reaction in which CYP264B1 was inactivated. This result confirmed that the sesquiterpene X and the diterpene Y were converted into X-OH and Y-OH by CYP264B1 and that they are the real products converted from FPP and GGPP by GeoA. As mentioned above, valencene is a possible sesquiterpene product of GeoA. However, CYP264B1 converts valencene into 5 products whereas it converts the sesquiterpene intermediate of GeoA into only one product (see section 3.3.2). Thus, the sesquiterpene X can not be valencene. This result indicates that the metabolites of the *TPSgc* encoding enzymes are new compounds which are not included in the NIST database yet.

In order to see whether other CYPs of *So ce56* are able to convert the products of GeoA or not, 5 enzymes (CYP264A1, CYP109C2, CYP109C1, CYP109D1, CYP260A1) were tested. These enzymes (kindly provided by Dr. Khatri) can be supported by the heterologous redox partners AdR and Adx. Among them, CYP109D1 was able to hydroxylate  $\alpha$ - and  $\beta$ -ionone (Khatri *et al.* 2010), CYP206A1 was able to hydroxylate nootkatone (Ewen 2009). Nootkatone and  $\alpha$ -/ $\beta$ -ionone are also substrates of CYP264B1. Interestingly, under similar condition applied for CYP264B1, none of the selected cytochromes P450 was able to convert the intermediates sesquiterpene X and diterpene Y of the terpene cyclase into further products. This result demonstrates that the sesquiterpene X-OH and the diterpene Y-OH are only formed by the cytochrome P450 CYP264B1 of the *So ce56*. Furthermore, it indicates that the *TPSgc* encoding enzymes could be involved in the biosynthesis of both sesquiterpenoids and diterpenoids in the *So ce56*. However, according to our observation, the role of this cluster was suspected to be responsible for the synthesis of sesquiterpenes rather than diterpenes in the myxobacterium *So ce56* because of three following reasons. Firstly, CYP264B1 in this cluster is capable of hydroxylate diverse sesquiterpene substrates (see section 3.1.6). Secondly, under the same conditions using the system with CYP264B1 and GeoA, the sesquiterpene product from FPP was observed much more clearly than the diterpene product from GGPP. Thirdly, the genomic organization of the gene cluster encoding for CYP264B1 and GeoA is similar to the sesquiterpene biosynthesis gene cluster of *Streptomyces avermitilis* (Quaderer *et al.* 2006) and the sesquiterpene biosynthesis gene cluster of *Streptomyces coelicolor* A3 (Zhao *et al.* 2008). For these

reasons, more attention should be address to the sesquiterpene synthesis through this gene cluster.

Taken together, the terpene cyclase GeoA was successfully expressed, purified and characterized. The substrates of this enzyme were found and the products were detected by GC-MS. The sesquiterpene product and diterpene product are similar to valencene and neocembrene A, respectively. In addition, the functional relation between CYP264B1 and GeoA has been shown and their products were detected. The *TPSgc* was thought to be involed in the sesquiterpene biosynthesis in the myxobacterium *Soc56*.

#### 4.4. Investigation of the terpene production in engineered *E. coli*

In present work, FPP and GGPP were identified as substrates of the terpene cyclase GeoA. Sesquiterpene and diterpene products were observed but they were proven to be novel metabolites which are not included in the NIST database yet (see section 3.3.2). Therefore, the identification of products (sesquiterpenes X and X-OH, diterpenes Y and Y-OH) of *TPSgc* encoding enzymes by NMR analyses is needed. To get sufficient amounts of products for NMR, *in vitro* conversion is not a desirable solution due to the very expensive precursor FPP, the expensive cofactor NADPH as well as purification of four proteins (GeoA, CYP264B1, Adx and AdR). An alternative approach is using metabolically engineered microorganisms as a host for heterologous production of isoprenoid compounds. For such purpose, *E. coli* is an attractive candidate because of extensive molecular genetic resources and the ease of genetic manipulation. This microorganism has been used as a host for the heterologous production of terpenoids (Martin *et al.* 2003, Harada 2009; Chang *et al.* 2006) and carotenoids (Albrecht *et al.* 1999; Alper *et al.* 2005; Farmer and Liao 2000; Vadali *et al.* 2005). It is known that *E. coli* does not synthesize isoprenoids such as monoterpenes, diterpenes, sesquiterpenes, but they synthesize IPP and DMAPP precursors through a non-mevalonate (MEP) pathway for the prenylation of tRNAs and the synthesis of FPP, which is used for quinone and cell wall biosynthesis (Connolly and Winkler 1989; Apfel *et al.* 1999; Martin 2003; Harada *et al.* 2009). However, the level of FPP is rather low and by no means sufficient for biotechnological applications. Therefore, in order to turn *E. coli* into a suitable host for the production of terpenoids, a series of biosynthesis genes encoding enzymes catalyzing several steps from FPP to the desired isoprenoids have been introduced into *E. coli* which enable it to synthesize these compounds from

intracellular FPP (Yokoyama *et al.* 1998; Harada *et al.* 2009). Many studies on pathway-engineering have been performed for increasing intracellular pools of IPP and DMAPP (Wang *et al.* 1999; Matthews and Wurtzel 2000; Kim and Keasling 2001; Reiling *et al.* 2004). For example, FPP amounts increased significantly when the intrinsic 1-deoxy-D-xylulose-5-phosphate synthase or 1-deoxy-D-xylulose 5-phosphate reductoisomerase gene in the MEP pathway was over-expressed in *E. coli*, resulting in an increase of the desired terpenoid yield (Albrecht *et al.* 1999). Similarly, the FPP content was enhanced several times when the IPP isomerase gene derived from green alga *Haematococcus pluvialis* or yeast *S. cerevisiae* was expressed in *E. coli* (Yuan *et al.* 2006). Although engineering the native MEP pathway has received a lot of attention, introducing the heterologous mevalonate pathway into *E. coli* seems to be more attractive, since it has been demonstrated to be a superior biosynthetic route for delivering high-level precursors to terpene synthases for large-scale production of isoprenoids (Kakinuma *et al.* 2001; Martin *et al.* 2003; Vadali *et al.* 2005; Yoon *et al.* 2006; Harada *et al.* 2009). But also introducing just a foreign partial (bottom) mevalonate pathway into *E. coli* and supplementing the culture medium with mevalonate significantly increased lycopene production in *E. coli* (Yoon *et al.* 2007). Based on these observations, we introduced enzymes of the mevalonate pathway in *E. coli* and coexpressed them with GeoA and CYP264B1 for sesquiterpene production. The two mevalonate vectors (pAC-Mev and pAC-Mv) used in this work have been reported in context with efficient synthesis of  $\alpha$ -humulene (Harada *et al.* 2009) and  $\beta$ -eudesmol (Yu *et al.* 2008) in *E. coli*.

Because poor expression of heterologous GeoA and CYP264B1 in *E. coli* would limit the terpene production, optimizing the expression conditions was a primary task. At the beginning of this work, BL21(DE3) was used for the expression of CYP264B1 and terpene cyclase GeoA. This strain is commonly used as a host for the synthesis of a foreign proteins using pET vectors. However, we found that both, CYP264B1 and GeoA, can only be expressed in this strain when they are co-expressed with the molecular chaperones GroEL and GroES (see section 3.1.1 and 3.2.3). Therefore, C43(DE3) - a mutant strain of *E. coli* BL21(DE3), was applied. This strain is often used to overcome the toxicity when recombinant proteins are over-expressed using the bacteriophage T7 RNA polymerase expression system (Dumon-Seignovert *et al.* 2004). We found that CYP264B1 and GeoA could be expressed in C43(DE3) without co-expression of chaperones. Furthermore, using this strain for the whole cell conversion

of  $\beta$ -ionone by CYP264B1, the product 3-hydroxy- $\beta$ -ionone was obtained with high yield (see section 3.1.9), confirming that CYP264B1 and its redox partners (Adx and AdR) were functionally expressed. For this reason, we used C43(DE3) for the production of terpene compounds.

We first investigated the efficiency of the production of the terpene cyclase products in *E. coli* C43(DE3). Cells carrying pAC-Mev or pAC-Mv together with pETG were cultured in TB medium supplemented with mevalonate. Since sesquiterpenes are volatile, we overlaid the culture with decane, an organic solvent with high hydrophobicity and low volatility to capture volatile compounds. This solvent has to be proved not to affect cell growth (Wang *et al.* 2010). GC-MS analysis of the decane phase 48h after induction showed that the sesquiterpene X was formed. The mass spectrum of the sesquiterpene product obtained in the *in vivo* conversion matched perfectly with the one of the product observed in the *in vitro* conversion of FPP by the terpene cyclase GeoA. Interestingly, when mevalonate (0.5 mg/ml) was supplied to the medium, sesquiterpene production was improved about 2.5 fold (see section 3.4.2). This result demonstrates that GeoA and enzymes of the mevalonate pathway are active in this strain. Using valencene as external standard for GCMS, the highest sesquiterpene yield was estimated to be 0.5 mg/50ml culture medium. Furthermore, the product yield in the strain carrying pAC-Mev was higher than in the strain carrying pAC-Mv (about 1.2 times). In the control reaction (i.e. *E. coli* carrying just pAC-Mev or pAC-Mv), sesquiterpene X was not detected, confirming that *E. coli* itself did not produce this compound from the isoprenoid precursor. However, the yield of sesquiterpene X ( $\approx 10$  mg/l) was much lower than the yield of  $\alpha$ -humulene ( $\approx 1$  g/l culture) achieved by Harada *et al.* (Harada *et al.* 2009) using the same vector pAC-Mev to boost isoprenoid synthesis. According to our observation, cell growth was considerably inhibited by decane, which might have caused the comparatively low yield. Although the conversion time was prolonged to 48 hours after induction, the optical density of the culture was very low ( $OD_{600} < 1.5$ ). Therefore, we believe that the sesquiterpene X production can be further improved by using other systems instead of decane to capture the volatile products.

In order to investigate the production of products of the *TPSgc* encoding enzymes in *E. coli*, we constructed the tetracistronic vector pETC4AAG containing genes coding for CYP264B1, Adx, AdR and GeoA. Co-expression of pETC4AAG and pAC-Mev in *E. coli* C43(DE3) was carried out in TB medium supplemented with mevalonate as

described in section 3.4.3. However, TLC analysis of the ethyl acetate extract from the culture revealed no product formation. Since CYP264B1, Adx and AdR was also shown to be expressed functionally in this strain (through the whole cell conversion of  $\beta$ -ionone, section 3.1.9), possible reason for this result is the loss (due to volatility) of the hydrophobic sesquiterpene X from the cells. A strategy to overcome this shortcoming comes from the question: how to retain the volatile compound in the culture but do not disturb cell growth? For this purpose, cyclodextrins (CDs) are attractive candidates since they are well known as an enhancer of the permeation of poorly soluble compounds through biological membranes (Másson *et al.* 1999). They are used not only in drug delivery but also in bioconversion due to their biocompatibility with microbes (Singh *et al.* 2002). For example, in the presence of cyclodextrin, microbial transformation of cholesterol to androst-4-ene-3,17-dione increased up to 90% (Bar 1989). The explanation for the highly interesting application of cyclodextrins comes from their structural feature. CDs are circular oligosaccharides composed of (1 $\rightarrow$ 4)-linked  $\alpha$ -D-glucosyl units with a shape as truncated cones (Szejtli 1998; Larsen 2002). The external surface of CDs is sufficiently hydrophilic to render them (or their complexes) soluble in water. In contrast, their interior is considerably less hydrophilic than the aqueous environment, thus it is able to hold appropriately sized non-polar moieties of other hydrophobic molecules within the cavity, without the formation or disturbance of covalent bonds (Munoz-Botella *et al.* 1995). The binding between the host CD and the guest molecule is rather in a dynamic equilibrium than being fixed or permanent (Jiang *et al.* 2010). Obviously, substrate must be taken up by the microbe before being biotransformed (Goswami *et al.* 1983). One hypothesis is that CDs extract and form inclusion complexes with lipids and other hydrophobic moieties from the cell membrane, making it more permeable, which facilitates uptake of substrate into the *E. coli* cells (Aachmann and Aune 2009). In short, we supposed that the sesquiterpene X would be more accessible to the biocatalyst of CYP264B1 when using CD as a strategy for overcoming the volatile loss. Among different type of CDs, hydroxypropyl- $\beta$ -cyclodextrin seems to be the most efficient for enhancing biotransformation (Singer *et al.* 1991; Mahato and Garai 1997). Therefore we used this compound for our experiment. Interestingly, after using this strategy, analysis of culture extracts by TLC, revealed a product which has a retention time identical to the one of the product of FPP *in vitro* conversion by CYP264B1 together with GeoA. As expected, this product was not present in the control cultures (i.e. the same *E. coli* strain carrying pAC-Mvc4) and

therefore thought to be the desired product. This result suggested that the production of sesquiterpene X-OH could be achieved by this system.

After having discussed about the product formation from FPP by *in vivo* conversion in detail, we want to draw attention to the diterpene product from GGPP conversion. Interestingly, in the *in vivo* conversion system, neither diterpene Y (80 % similarity to neocembrene A) nor the derivative Y-OH were detected. However, it is known that *E. coli* does not naturally produce appreciable quantities of GGPP, the precursor for diterpene biosynthesis (Reiling *et al.* 2004). Assuming this is the reason for the failed attempt to produce the diterpene Y and its hydroxylated product in *E. coli*, the obvious solution would be to introduce and over-express a GGPP synthase gene in *E. coli* to increase the intracellular GGPP pool. A suitable GGPP synthase gene could be cloned from the genomic DNA of the myxobacterium So ce56. Among three candidates from So ce56 (locus\_tag sce1722, sce3009 and sce4204), the most attention should be paid for the one of locus\_tag sce1722 since its product is probable geranylgeranyl pyrophosphate synthase (NCBI). Another interesting suggestion comes from Ohnuma (Ohnuma *et al.* 1996): they suggested that GGPP synthase could be produced by mutating native FPP synthase IspA from *E. coli* (Tyr80 of IspA was mutated to an Asp residue) to avoid the vagaries of foreign protein production in *E. coli* (codon usage and quaternary or tertiary protein structure). This might certainly be important regarding So ce56's high GC content.

Accordingly, our current results suggests that amounts of sesquiterpene X and diterpene Y sufficient for NMR analysis can be obtained by using *E. coli* C43(DE3) as host and introducing the genes encoding the enzymes needed for the production of the respective isoprenoid precursors. In general, the product yield could be improved by optimizing culture conditions. For example, the carbon source can be an important factor effecting terpenoid yield. Indeed, both terpenoid production and cell growth can be significantly enhanced by proportionally increasing the available amounts of mevalonate and glycerol (Alper *et al.* 2005; Yoon *et al.* 2007). Glycerol was reported to increase carotenoid production, whereas glucose was found to decrease it (Lee *et al.* 2004; Tao *et al.* 2005). Furthermore, acetoacetate - a cheaper compound could substitute D-mevalonate, using *E. coli* cells expressing acetoacetate-CoA ligase genes as well as mevalonate pathway genes (Harada *et al.* 2009). In addition, protein expression levels could likely be improved by substitution of the expression vector or host strain, or the use of codon-optimized genes (Harada *et al.* 2009; Martin 2003) using the *E. coli* codon

preferences. Higher yields can also be achieved by optimization of the culture conditions in a fermentor, where culture parameters like pH, dissolved oxygen, and substrate concentration can be controlled.

In conclusion, terpene production in *E. coli* C43(DE3) using the mevalonate gene cluster of pAC-Mev and pAC-Mv were investigated. Although the products have not been identified, the result suggests the possibility to produce sufficient product amounts for NMR analysis, to clarify the role of the *TPSgc* in the myxobacterium *Sorangium cellulosum* So ce56.

#### 4.5. Outlook

CYP264B1 is a potent sesquiterpene hydroxylase with the ability of performing the regioselective hydroxylation at unfavorable sites like C-13 of nootkatone and C-3 of  $\beta$ -ionone. Thus, CYP264B1 is an attractive candidate for exploitation of norisoprenoids and sesquiterpene biotransformation. Among the identified substrates, besides nootkatone and ionones, valencene and humulene also have interesting potential for further studies. Indeed, valencene is a precursor of nootkatone, the most important and expensive aromatic compound of grapefruit (Fraatz *et al.* 2009). Also, humulene is a precursor for the synthesis of zerumbone, a compound possessing anti-inflammatory and anti-HIV effects as well as anti-proliferative activities in a wide variety of cancer cell lines (Yu *et al.* 2011). To explore this intriguing P450, studies on protein crystallization and homology modeling for detailed knowledge about the structure of CYP264B1 should be performed. In addition, with the aid of molecular biology techniques, such as directed evolution or site-directed mutagenesis to improve enzyme activity, substrate range or product formation, would provide a great potential for the exploitation of this enzyme, especially towards sesquiterpene compounds.

The *E. coli* whole cell system which was successful applied for  $\beta$ -ionone conversion in this study could be applied for other substrates of CYP264B1. Especially, to overcome the loss of volatile substrates during biotransformation, using cyclodextrin might be a significant suggestion.

The identification of the substrates of the terpene cyclase GeoA and the demonstration that the metabolites of the *TPSgc* encoding enzymes are novel compounds, which are not included in the NIST database yet, are important results for the continuing work. In comparison to CYP264B1, GeoA is less stable after purification because it is easier precipitated during storage. Therefore, in order to characterize this enzyme kinetically,

optimal pH, temperature conditions as well as additives (e.g. 2-mercaptoethanol, and Tween 80) which might help to enhance the activity of terpene cyclase, should be investigated.

Although the question about the function of the *TPSgc* in the myxobacterium *sorangium cellulosum* So ce56 has not yet been fully answered, the investigation of the terpene production in mevalonate-engineered *E. coli* as well as a solution for overcoming the bottleneck during terpene biotransformation by using cyclodextrins, has pointed out that it is possible to get sufficient amounts of product for NMR analysis if laboratory conditions are optimized, and therefore the role of the *TPSgc* in the secondary metabolic pathways of So ce56 is going to be clarified.

## 5. SUMMARY

In the work presented here, CYP264B1 and the terpene cyclase GeoA of *Sorangium cellulosum* So ce56 have been characterized.

CYP264B1 is able to convert norisoprenoids ( $\alpha$ -ionone and  $\beta$ -ionone) and diverse sesquiterpene compounds, including nootkatone. Three products, 3-hydroxy- $\alpha$ -ionone, 3-hydroxy- $\beta$ -ionone and 13-hydroxy-nootkatone were characterized using HPLC and  $^1\text{H}$  and  $^{13}\text{C}$  NMR. CYP264B1 is the first enzyme reported to be capable to hydroxylate regioselectively both norisoprenoids at the position C-3 as well as nootkatone at the position C-13. The kinetics ( $K_m$  and  $V_{max}$ ) of the product formation were analyzed by HPLC. The results of docking  $\alpha$ -/ $\beta$ -ionone and nootkatone into a homology model of CYP264B1 revealed the structural basis of these selective hydroxylations. In addition, an *E. coli* whole cell system containing CYP264B1 and its redox partners was created for the biotransformation of CYP264B1 substrates. This system was applied successfully for  $\beta$ -ionone conversion.

FPP and GGPP were found to be substrates for GeoA. The sesquiterpene and diterpene products of GeoA are similar to valencene (89%) and neocembrene A (80%), respectively. However, these products are most likely new compounds. In order to characterize them by NMR, a whole cell system based on mevalonate pathway-engineered *E. coli* was created to facilitate the production of sufficient amounts. The terpene production using this system was investigated, showing that it is possible to obtain the amounts required for NMR analysis if laboratory conditions are optimized.

## ZUSAMMENFASSUNG

In der vorliegenden Arbeit wurden CYP264B1 und die Terpencyclase GeoA aus *S. cerevisiae* charakterisiert.

CYP264B1 ist in der Lage, Norisoprenoide ( $\alpha$ -Ionon und  $\beta$ -Ionon) und diverse Sesquiterpene, inklusive Nootkaton, umzusetzen. Drei Produkte (3-Hydroxy- $\alpha$ -Ionon, 3-Hydroxy- $\beta$ -Ionone und 13-Hydroxy-Nootkaton) wurden mittels HPLC und  $^1\text{H}$  und  $^{13}\text{C}$  NMR charakterisiert. Damit ist CYP264B1 das erste Enzym, das die Fähigkeit besitzt, regioselektiv Norisoprenoide an Position C-3, sowie Nootkaton an Position C-13 zu hydroxylieren. Die Kinetik der Produktbildung ( $V_{\max}$  und  $K_m$ ) wurde mittels HPLC analysiert. Durch das Docking von  $\alpha/\beta$ -Ionon und Nootkaton in das Homologiemodell von CYP264B1 konnte die strukturelle Grundlage dieser selektiven Hydroxylierungen aufgeklärt werden. Des Weiteren wurde ein *E. coli* Ganzellsynthesystem für die Umsetzung von CYP264B1 Substraten etabliert, das neben CYP264B1 auch seine Redox Partner enthält. Dieses System wurde erfolgreich für die Umsetzung von  $\beta$ -Ionon eingesetzt.

FPP und GGPP wurden als Substrate von GeoA identifiziert. Die jeweiligen Sesquiterpen- und Diterpen- Produkte wiesen Ähnlichkeit mit Valencen (89%), bzw. Neocembren (80%) auf. Jedoch handelt es sich bei den gebildeten Produkten höchstwahrscheinlich um neue Verbindungen. Um ausreichende Mengen dieser Verbindungen für eine NMR Analyse zur Verfügung zu stellen, wurde ein Ganzellsystem aufgebaut, das auf *E. coli* Zellen die heterolog zusätzliche Proteine des Mevalonat Stoffwechselweges exprimieren, basiert. Durch weitere Optimierung dieses Systems sollte es in Zukunft möglich sein, die für eine NMR Analyse erforderlichen Produktmengen produzieren.

## 6. REFERENCES

- Aachmann FL and Aune TEV (2009) Use of cyclodextrin and its derivatives for increased transformation efficiency of competent bacterial cells. *Appl Microbiol Biotechnol* 83:589–596
- Adam KP, Thiel R, Zapp J, Becker H (1998) Involvement of the Mevalonic Acid Pathway and the Glyceraldehyde–Pyruvate Pathway in Terpenoid Biosynthesis of the Liverworts *Ricciocarpos*. *Arch Biochem Biophys* 354:181–187
- Agematu H, Matsumoto N, Fujii Y, Kabumoto H, Doi S, Machida K, Ishikawa J, Arisawa A (2006) Hydroxylation of testosterone by bacterial cytochrome P450 using the *Escherichia coli* expression system. *Biosci Biotechnol Biochem* 70:307–311
- Albrecht M, Misawa N, Sandmann G (1999) Metabolic engineering of the terpenoid biosynthetic pathway of *Escherichia coli* for production of the carotenoids  $\beta$ -carotene and zeaxanthin. *Biotechnol Lett* 21:791–795
- Alper H, Jin YS, Moxley JF, Stephanopoulos G (2005) Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metab Eng* 7:155–164
- Alper H, Miyaoku K, Stephanopoulos G (2005) Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. *Nat Biotechnol* 23:612–616
- Apfel CM, Takacs B, Fountoulakis M, Stieger M, Keck W (1999) Use of genomics to identify bacterial undecaprenyl pyrophosphate synthetase: Cloning, expression, and characterization of the essential *uppS* gene. *J Bacteriol* 181:483–492
- Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* 22:195–201
- Azerad R (2000) Regio- and stereoselective microbial hydroxylation of terpenoid compounds. in: R.N. Patel (Ed.), *Stereoselective Biocatalysis*, Marcel Dekker Publ, New York, pp. 153–180
- Bar R (1989) Cyclodextrin aided bioconversions and fermentations. *Trends Biotechnol.* 7:2–4
- Barrero AF, Herrador MM, Quílez del Moral JF, Arteaga P, Meine N, Pérez-Morales MC, Catalán JV (2011) Efficient synthesis of the anticancer beta-elemene and other bioactive elemenes from sustainable germacrone. *Org Biomol Chem* 9:1118–1125

- Bell SG, Dale A, Rees NH, Wong LL (2010) A cytochrome P450 class I electron transfer system from *Novosphingobium aromaticivorans*. *Appl Microbiol Biotechnol* 86:163–175
- Bernhardt R (2006) Cytochromes P450 as versatile biocatalysts. *J Biotechnol* 124:128–145
- Bicas JL, Dionisio AP, Pastore GM (2009) Bio-oxidation of terpenes: an approach for the flavor industry. *Chem Rev* 109:4518–31
- Bleif S, Hannemann F, Zapp J, Hartmann D, Jauch J and Bernhardt R (2011) A new *Bacillus megaterium* whole-cell catalyst for the hydroxylation of the pentacyclic triterpene 11-keto- $\beta$ -boswellic acid (KBA) based on a recombinant cytochrome P450 system. *Appl Microbiol Biotechnol*. DOI 10.1007/s00253-011-3467-0
- Bode HB, Zeggel B, Silakowski B, Wenzel SC, Reichenbach H, Müller R (2003) Steroid biosynthesis in prokaryotes: identification of myxobacterial steroids and cloning of the first bacterial 2,3(S)-oxidosqualene cyclase from the myxobacterium *Stigmatella aurantiaca*. *Molecular Microbiology* 47: 471–481
- Bode HB and Müller R (2006) Analysis of myxobacterial secondary metabolism goes molecular. *J Ind Microbiol Biotechnol* 33:577–588
- Bohlmann J, Meyer-Gauen, Croteau R (1998) Plant terpenoid synthases: Molecular biology and phylogenetic analysis. *Proc Natl Acad Sci USA* 95: 4126–4133
- Borges KB, Borges W de Souza, Rosa Durán-Patrón, Pupo MT, Bonato PS, Collado IG (2009) Stereoselective biotransformations using fungi as biocatalysts. *Tetrahedron: Asymmetry* 20:385–397
- Boucher Y and Doolittle WF (2000) The role of lateral gene transfer in the evolution of isoprenoid biosynthesis pathways. *Mol Microbiol* 37:703–716
- Brenna E, Fuganti C, Serra S, and Kraft P (2002) Optically Active Ionones and Derivatives: Preparation and Olfactory Properties. *Eur J Org Chem*: 9672978
- Buckingham J (1998) Dictionary of natural products on CD-ROM vol 6.1. Chapman & Hall, London
- Cane DE and Watt RM (2003) Expression and mechanistic analysis of a germacradienol synthase from *Streptomyces coelicolor* implicated in geosmin biosynthesis. *PNAS* 100:1547–1551
- Cankar K, an Houwelingen A, Bosch D, Sonke T, Bouwmeester H, Beekwilder J (2011) A chicory cytochrome P450 mono-oxygenase CYP71AV8 for the oxidation of (+)-valencene. *FEBS Letters* 585:178–182

- Caruthers JM, Kang I, Rynkiewicz MJ, Cane DE, Christianson DW (2000) Crystal structure determination of aristolochene synthase from the blue cheese mold, *Penicillium roqueforti*. *J Biol Chem* 275:25533–25539
- Celik A, Flitsch SL, Turner NJ (2005) Efficient terpene hydroxylation catalysts based upon P450 enzymes derived from actinomycetes. *Org Biomol Chem* 3: 2930–2934
- Chandrasekhar GN, Tilly K, Woolford C, Hendrix R and Georgopoulos C (1986) Purification and properties of the groES morphogenetic protein of *Escherichia coli*. *J Biol Chem* 261:12414–12419
- Chang MCY, Eachus RA, Trieu W, Ro D-K and Keasling JD (2007) Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. *NATURE CHEMICAL BIOLOGY* 3:274–277
- Chang YJ, Song SH, Park SH, Kim SU (2000) Amorpha-4,11-diene synthase of *Artemisia annua*: cDNA isolation and bacterial expression of a terpene synthase involved in artemisinin biosynthesis. *Arch Biochem Biophys* 383:178–184
- Chou CP (2007) Engineering cell physiology to enhance recombinant protein production in *Escherichia coli*. *Appl Microbiol Biotechnol* 76:521–532
- Christianson DW (2006) Structural biology and chemistry of the terpenoid cyclases. *Chem Rev* 106:3412–42
- Chu SS, Jiang GH, Liu ZL (2011) Insecticidal compounds from the essential oil of Chinese medicinal herb *Atractylodes chinensis*. *Pest Manag Sci* 67:1253–1257
- Colby SM, Alonso WR, Katahira EJ, McGarvey DJ, Croteau R (1993) 4S-limonene synthase from the oil glands of spearmint (*Mentha spicata*). cDNA isolation, characterization, and bacterial expression of the catalytically active monoterpene cyclase. *J Biol Chem* 268:23016–23024
- Conner KP, Woods CM, Atkins WM (2011) Interactions of cytochrome P450s with their ligands *Archives of Biochemistry and Biophysics* 507:56–65
- Connolly DM and Winkler ME (1989) Genetic and physiological relationships among the miaA gene, 2-methylthio-N<sup>6</sup>-(delta<sup>2</sup>-isopentenyl) a denosine transfer RNA modification, and spontaneous mutagenesis in *Escherichia coli* K-12. *J Bacteriol* 171:3233–3246
- Croteau R (1987) Biosynthesis and catabolism of monoterpenoids. *Chem Rev* 87:929–954
- Crowell PL and Gould MN (1994) Chemoprevention and therapy of cancer by d-limonene. *CRC Crit Rev Oncogenesis* 5:1–22

- Curci R, D'Accolti L, Fusco C (2006) A novel approach to the efficient oxygenation of hydrocarbons under mild conditions. Superior oxo transfer selectivity using dioxiranes. *Acc Chem Res* 39:1–9
- Daff SN, Chapman SK, Turner KL, Holt RA, Govindaraj S, Poulos TL, Munro AW (1997) Redox control of the catalytic cycle of flavocytochrome P-450 BM3. *Biochemistry* 36:13816–13823
- Dairi T (2005) Studies on Biosynthetic Genes and Enzymes of Isoprenoids Produced by Actinomycetes. *J Antibiot* 58:227–243
- Dawson FA (1994) The amazing terpenes. *Naval Stores Rev* 6–12
- Degenhardt J, Köllner TG, Gershenzon J (2009) Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* 70:1621–1637
- Denisov IG, Makris TM, Sligar SG, and Schlichting I (2005) Structure and chemistry of cytochrome P450. *Chem Rev* 105:2253–2278
- Denisov IG, Baas BJ, Grinkova YV, Sligar SG (2007) Cooperativity in cytochrome P450 3A4: linkages in substrate binding, spin state, uncoupling, and product formation. *J Biol Chem* 282:7066–7076
- Dickschat JS, Bode HB, Mahmud T, Müller R and Schulz S (2005) A novel type of geosmin biosynthesis in myxobacteria. *J Org Chem* 70:5174–5182
- Dumon-Seignovert L, Cariot G, Vuillard L (2004) The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expression and Purification* 37:203–206
- Davis EM and Croteau R (2000) Cyclization Enzymes in the Biosynthesis of Monoterpenes, Sesquiterpenes, and Diterpenes: In *Topics in Current Chemistry* 209: 54–95
- Edris AE (2007) Pharmaceutical and therapeutic potentials of essential oils and their individual volatile constituents: a review. *Phytother Res* 21:308–23
- Ellis RJ (1991) Molecular chaperones. *Annu Rev Biochem* 60:321–347
- Engels B, Dahm P, Jennewein S (2008) Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards taxol (paclitaxel) production. *Metab Eng* 10:201–206
- Ewen KM, Hannemann F, Khatri Y, Perlova O, Kappl R, Krug D, Hüttermann J, Müller R, Bernhardt R (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* 284:28590–28598

- Faber K (2000) Biotransformations in Organic Chemistry: A Textbook, Springer, Berlin. pp. 220–272
- Fantuzzi A, Fairhead M, Gilardi G (2004) Direct electrochemistry of immobilized human cytochrome P450 2E1. *J Am Chem Soc* 126:5040–5041
- Farmer WR, Liao JC (2000) Improving lycopene production in *Escherichia coli* by engineering metabolic control. *Nat Biotechnol* 18:533–537
- Feyereisen R (1999) Insect P450 enzymes. *Annu Rev Entomol* 44:507–533
- Fraatz MA, Berger RG, Zorn H (2009) Nootkatone-a biotechnological challenge. *Appl Microbiol Biotechnol* 83:35–41
- Fraga BM (2006) Natural sesquiterpenoids. *Nat Prod Rep* 23:943–72
- Francis DM and Page R (2010) Strategies to Optimize Protein Expression in *E. coli*. *Current Protocols in Protein Science* DOI: 10.1002/0471140864.ps0524s61
- Furusawa M, Hashimoto T, Noma Y, Asakawa Y (2005) Biotransformation of citrus aromatics nootkatone and valencene by microorganisms. *Chem Pharm Bull (Tokyo)* 53:1423–1429
- Galani VJ, Patel BG, Rana DG (2010) *Sphaeranthus indicus* Linn.: A phytopharmacological review. *Int J Ayurveda Res* 1:247–253
- 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
- Gibson GG and Skett P (1994) Introduction to drug metabolism. London: Blackie Academic and Professional
- Gigon PL, Gram TE, and Gillette JR (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
- Girhard M, Klaus T, Khatri Y, Bernhardt R, Urlacher VB (2010) Characterization of the versatile monooxygenase CYP109B1 from *Bacillus subtilis*. *Appl Microbiol Biotechnol* 87: 595–607
- Gliszczynska A, Łysek A, Janeczko T, Świtalska M, Wietrzyk J, Wawrzeńczyk C (2011) Microbial transformation of (+)-nootkatone and the antiproliferative activity of its metabolites. *Bioorg Med Chem* 19:2464–2469
- Gomes-Carneiro MR, Dias DM, Paumgartten FJ (2006) Study on the mutagenicity and antimutagenicity of beta-ionone in the *Salmonella*/microsome assay. *Food Chem Toxicol* 44:522–527

- Goswami PC, Singh HD, Bhagat SD, Baruah JN (1983) Mode of uptake of insoluble solid substrates by microorganisms. I: Sterol uptake by an *Arthrobacter* species. *Biotechnol Bioeng* 25:2929–2943
- Griffin BW, Peterson JA, Estabrook RW (1979) Cytochrome P450: Biophysical properties and catalytic function. The 2nd ed, D. Dolphin. *The Porphyrins* 7, Academic Press New York:p333–75
- Gu W, Zhou T, Wilke CO (2010) A Universal Trend of Reduced mRNA Stability near the Translation-Initiation Site in Prokaryotes and Eukaryotes. *PLoS Computational Biology* 6:e1000664
- Guengerich FP, Ballou DP, Coon MJ (1975) Purified liver microsomal cytochrome P450. Electron-accepting properties and oxidation-reduction potential. *J Biol Chem* 250:7405–14
- Guengerich FP (1983) Oxidation-reduction properties of rat liver cytochromes P450 and NADPH-cytochrome P450 reductase related to catalysis in reconstituted systems. *Biochemistry* 22:2811–20
- Guengerich FP (2001) Uncommon P450-catalyzed Reactions. *Curr Drug Metab* 2:93–115
- Guengerich FP, Wu ZL and Bartleson CJ (2005) Function of human cytochrome P450s: characterization of the orphans. *Biochem Biophys Res Commun* 338:465–469
- Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling. *Electrophoresis* 18:2714–2723
- Gunsalus IC, Pederson TC, Sligar SG (1975) Oxygenase-catalyzed biological hydroxylations. *Annu Rev Biochem* 44:377–407
- Gunsalus IC, Sligar SG (1978) Oxygen reduction by the P450 monooxygenase systems. *Adv Enzymol Relat Areas Mol Biol* 47:1–44
- Günnewich N, Page JE, Köllner TG, Degenhardt J, Kutchan TM (2007) Functional expression and characterization of trichome-specific (-)-limonene synthase and (+)- $\alpha$ -pinene synthase from *Cannabis sativa*. *Natural Product Communications* 2:223–232
- Guo Y, Wallace SS, Bandaru V (2009) A novel bicistronic vector for overexpressing *Mycobacterium tuberculosis* proteins in *Escherichia coli*. *Protein Expr Purif* 65:230–237
- Han J-H, Choi Y-S, Kim W-J, Jeon YH, Lee SK, Lee B-J, Ryu K-S (2010) Codon optimization enhances protein expression of human peptide deformylase in *E. coli*. *Protein Expression and Purification* 70:224–230

- Hannemann F, Virus C, 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
- Hannemann F, Bichet A, Ewen KM, Bernhardt R (2007) Cytochrome P450 systems—biological variations of electron transport chains. *Biochim Biophys Acta* 1770:330–344
- Harada H, Yu F, Okamoto S, Kuzuyama T, Utsumi R, Misawa N (2009) Efficient synthesis of functional isoprenoids from acetoacetate through metabolic pathway-engineered *Escherichia coli*. *Appl Microbiol Biotechnol* 81:915–925
- Harborne JB (1991) Recent advances in the ecological chemistry of plant terpenoids. In *Ecological Chemistry and Biochemistry of Plant Terpenoids*. Harborne JB and Tomas-Barberan FA, eds (Oxford: Clarendon Press) p 399–426
- Hasemann CA, Kurumbail RG, Boddupalli SS, Peterson JA and Deisenhofer J (1995) Structure and function of cytochromes P450: a comparative analysis of three. *Structure* 3:41–62
- Hasler JA, Estabrook RW, Murray M, Pikuleva I, Waterman MR, Capdevila J, Holla V, Helvig C, Falck J, Farrell G (1999) Human cytochromes P450. *Mol Aspects Med* 20:1–137
- Hegazy M-EF, Kuwata C, Matsushima A, Ahmed AA, Hirata T (2006) Biotransformation of sesquiterpenoids having  $\alpha$ -,  $\beta$ -unsaturated carbonyl groups with cultured plant cells of *Marchantia polymorpha*. *Journal of Molecular Catalysis B: Enzymatic* 39:13–17
- Helvig C, Koener JF, Unnithan GC, 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 USA* 101:4024–4029
- Hohn TM (1999) Cloning and expression of terpene synthase genes. *Comprehensive Natural Products Chemistry: Isoprenoids Including Carotenoids and Steroids*, ed Cane DE (Elsevier, Oxford) pp 201–215
- Holton TA, Brugliera F, Lester DR, Tanaka Y, Hyland CD, Menting JG, Lu CY, Farcy E, Stevenson TW, Cornish EC (1993) Cloning and expression of cytochrome P450 genes controlling flower colour. *Nature* 366:276–279
- Horbach S, Sahm H, and Welle R (1993) Isoprenoid biosynthesis in bacteria: two different pathways?. *FEMS Microbiol Lett* 111: 135–140
- Huey R, Morris GM, Olson AJ, Goodsell DS (2007) A semiempirical free energy force field with charge-based desolvation. *J Comput Chem* 28:1145–1152

- Hunke S and Betton JM (2003) Temperature effect on inclusion body formation and stress response in the periplasm of *Escherichia coli*. *Mol Microbiol* 50:1579–1589
- Ichikawa Y, Yamano T (1967) Reconversion of detergent- and sulfhydryl reagent-produced P-420 to P-450 by polyols and glutathione. *Biochim Biophys Acta* 131:490–497
- Ikeda C, Hayashi Y, Itoh N, Seto H, Dairi T (2007) Functional analysis of eubacterial ent-copalyl diphosphate synthase and pimara-9(11),15-diene synthase with unique primary sequences. *J Biochem* 141:37–45
- 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 USA* 86:7823–7827
- Jiang J, He X, and Cane DE (2006) Geosmin biosynthesis: *Streptomyces coelicolor* germacradienol/germacrene D synthase converts farnesyl diphosphate to geosmin. *J Am Chem Soc* 128:8128–8129
- Jiang S, Li J-N, Jiang Z-T (2010) Inclusion reactions of  $\beta$ -cyclodextrin and its derivatives with cinnamaldehyde in *Cinnamomum loureirii* essential oil. *Eur Food Res Technol* 230:543–550
- Kahn R, Durst F (2000) Function and evolution of plant cytochrome P450. *Recent Adv Phytochem* 34:151–189
- Kakinuma K, Dekishima Y, Matsushima Y, Eguchi T, Misawa N, Takagi M, Kuzuyama T, Seto H (2001) New approach to multiply deuterated isoprenoids using triply engineered *Escherichia coli* and its potential as a tool for mechanistic enzymology. *J Am Chem Soc* 123:1238–1239
- Kalb VF, Woods CW, Turi TG, Dey CR, Sutter TR, and Loper JC (1987) Primary structure of the P450 lanosterol demethylase gene from *Saccharomyces cerevisiae*. *DNA* 6:529–537
- Khatri Y (2009) The cytochrome P450 complement of the myxobacterium *Sorangium cellulosum* So ce56 and characterization of two members, CYP109D1 and CYP260A1 PhD Thesis Saarland University, Saarbrücken.
- Khatri Y, Girhard M, Romankiewicz A, Ringle M, Hannemann F, Urlacher VB, Hutter MC, Bernhardt R (2010) Regioselective hydroxylation of norisoprenoids by CYP109D1 from *Sorangium cellulosum* So ce56. *Appl Microbiol Biotechnol* 88:485–95
- Khatri Y, Hannemann F, Ewen KM, Pistorius D, Perlova O, Kagawa N, Brachmann AO, Müller R, Bernhardt R (2010) The CYPome of *Sorangium cellulosum* So

- ce56 and Identification of CYP109D1 as a New Fatty Acid Hydroxylase. *Chemistry & Biology* 17: 1295–1305
- Khatri Y, Hannemann F, Perlova O, Müller R, Bernhardt R (2011) Investigation of cytochromes P450 in myxobacteria: excavation of cytochromes P450 from the genome of *Sorangium cellulosum* So ce56 *FEBS Lett* 585:1506–1513
- Kim SW, Keasling JD (2001) Metabolic engineering of the nonmevalonate isopentenyl diphosphate synthesis pathway in *Escherichia coli* enhances lycopene production. *Biotechnol Bioeng* 72:408–415
- Kimata Y, Shimada H, Hirose T, Ishimura Y (1995) Role of Thr-252 in cytochrome P450cam: a study with unnatural amino acid mutagenesis. *Biochem Biophys Res Commun* 208:96–102
- Komatsu M, Tsuda M, Omura S, Oikawa H, Ikeda H (2008) Identification and functional analysis of genes controlling biosynthesis of 2-methylisoborneol. *Proc Natl Acad Sci USA* 105:7422–7427
- Kunze KL, Nelson WL, Kharasch ED, Thummel KE, and Isoherranen N (2006) Stereochemical aspects of itraconazole metabolism in vitro and in vivo. *Drug Metab Dispos* 34:583–590
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685
- Lamb DC, Lei L, Warrilow AG, Lepesheva GI, Mullins JG, Waterman MR and Kelly SL (2009) The first virally encoded cytochrome P450. *J Virol* 83:8266–8269
- Landry SJ and Gierasch LM (1994) Polypeptide interactions with molecular chaperones and their relationship to *in vivo* protein folding. *Annu Rev Biophys Biomol Struct* 23:645–649
- Larroche C, Creuly C, Gros JB (1995) Fed-batch biotransformation of  $\beta$ -ionone by *Aspergillus niger*. *Appl Microbiol Biotechnol* 43:222–227
- Larsen KL (2002) Large cyclodextrins. *J Inclusion Phenom Macro* 43:1–13
- Lee PC, Mijts BN, Schmidt-Dannert C (2004) Investigation of factors influencing production of the monocyclic carotenoid torulene in metabolically engineered *Escherichia coli*. *Appl. Microbiol. Biotechnol* 65:538–546
- Lesburg CA, Zhai G, Cane DE, Christianson DW (1997) Crystal structure of pentalenene synthase: mechanistic insights on terpenoid cyclization reactions in biology *Science*. 277: 1820–1824

- Lesburg CA, Caruthers JM, Paschall CM, Christianson DW (1998) Managing and manipulating carbocations in biology: terpenoid cyclase structure and mechanism. *Curr Opin Struct Biol* 8:695–703
- Lewis DFV (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
- Lichtenthaler HK, Schwender J, Disch A, Rohmer M (1997) Biosynthesis of isoprenoids in higher plant chloroplasts proceeds via a mevalonate-independent pathway *FEBS Lett* 400:271–274
- Loeber DE, Russell SW, Toubé TP, Weedon BCL, and Diment J (1971) Carotenoids and related compounds. Part XXVIII. Syntheses of zeaxanthin,  $\beta$ -cryptoxanthin, and zeinoxanthin ( $\alpha$ -cryptoxanthin). *J Chem Soc* 404–408
- Lutz-Wahl S, Fischer P, Schmidt-Dannert C, Wohlleben W, Hauer B, Schmid RD (1998) Stereo- and regioselective hydroxylation of alpha-ionone by *Streptomyces* strains. *Appl Environ Microbiol* 64:3878–81
- Mahato SB, Garai S (1997) Advances in microbial steroid biotransformation. *Steroids* 62:332–45
- Mansuy D (1998) The great diversity of reactions catalyzed by cytochromes P450. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol* 121:5–14
- Martin VJ, Pitera DJ, Withers ST, Newman JD, Keasling JD (2003) Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat Biotechnol* 21:796–802
- Martinis SA, Atkins WM, Stayton PS, Sligar SG (1989) A conserved residue of cytochrome P450 is involved in heme-oxygen stability and activation. *J Am Chem Soc* 111:9252–9253
- Másson M, Loftssona T, Mássonb G, Stefánsson E (1999) Cyclodextrins as permeation enhancers: some theoretical evaluations and *in vitro* testing. *Journal of Controlled Release* 59:107–118
- Matthews PD, Wurtzel ET (2000) Metabolic engineering of carotenoid accumulation in *Escherichia coli* by modulation of the isoprenoid precursor pool with expression of deoxyxylulose phosphate synthase. *Appl Microbiol Biotechnol* 53:396–400
- Maurer SC, Kühnel K, Kaysser LA, Eiben S, Schmid RD, Urlacher VB (2005) Catalytic hydroxylation in biphasic systems using CYP102A1 mutants. *Adv Synth Catal* 347:1090–1098

- McGarvey DJ and Croteau R (1995) Terpenoid Metabolism. *The Plant Cell* 7:1015–1026
- Mikami Y, Fukunaga Y, Arita M, and Kisaki T (1981) Microbial transformation of  $\beta$ -ionone and  $\beta$ -methylionone. *Appl Environ Microbiol* 41:610–617
- Miroux B and Walker JE (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* 260: 289–98
- Misawa N (2011) Pathway engineering for functional isoprenoids. *Current Opinion in Biotechnology* 22:1–7
- Miyazawa M, Nakamura Y, Ishikawa Y (2000) Insecticidal sesquiterpene from *Alpinia oxyphylla* against *Drosophila melanogaster*. *J Agric Food Chem* 48:3639–3641
- Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK and Olson AJ (1998) Automated Docking Using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function. *J Comput Chem* 19:1639–1662
- Munoz-Botella S, del Castillo B, Martin MA (1995) Cyclodextrin properties and applications of inclusion complex formation. *Ars Pharm* 36:187–98
- Muntendam R, Melillo E, Ryden A, Kayser O (2009) Perspectives and limits of engineering the isoprenoid metabolism in heterologous hosts. *Appl Microbiol Biotechnol* 84:1003–1019
- Murakami K, Mason HS (1967) An electron spin resonance study of microsomal fex. *J Biol Chem* 242:1102–1110
- Murase T, Misawa K, Haramizu S, Minegishi Y, Hase T (2010) Nootkatone, a characteristic constituent of grapefruit, stimulates energy metabolism and prevents diet-induced obesity by activating AMPK. *Am J Physiol Endocrinol Metab* 299:266–275
- Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW, Gunsalus IC, Nebert DW (1996) P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* 6:1–42
- Nishihara K, Kanemori M, Kitagawa M, Yanagi H, 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
- Ogata J, Kanno Y, Itoh Y, Tsugawa H, Suzuki M (2005) Plant biochemistry: anthocyanin biosynthesis in roses. *Nature* 435:757–758

- Ohloff G (1978) Recent development in the field of naturally occurring aroma components. *Fortschr Chem Org Nat* 35:431–527
- Ohnuma S-I (1996) Conversion from Farnesyl Diphosphate Synthase to Geranylgeranyl Diphosphate Synthase by Random Chemical Mutagenesis. *The journal of biological chemistry* 271:10087–10095
- Omura T and Sato R (1964) The Carbon Monoxide-Binding Pigment of Liver Microsomes. II. Solubilization, Purification, and Properties. *J Biol Chem* 239:2379–85
- Panahi M, Alli Z, Cheng X, Belbaraka L, Belgoudi J, Sardana R, Phipps J, Altosaar I (2004) Recombinant protein expression plasmids optimized for industrial *E. coli* fermentation and plant systems produce biologically active human insulin-like growth factor-1 in transgenic rice and tobacco plants. *Transgenic Res* 13:245–259
- Panella NA, Dolan MC, Karchesy JJ, Xiong Y, Peralta-Cruz J, Khasawneh M, Montenieri JA, Maupin GO (2005) Use of novel compounds for pest control: insecticidal and acaricidal activity of essential oil components from heartwood of *Alaska yellow cedar*. *J Med Entomol* 42:352–358
- 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
- Perlova O, Gerth K, Kaiser O, Hans A, Müller R (2006) Identification and analysis of the chivosazol biosynthetic gene cluster from the myxobacterial model strain *Sorangium cellulosum* So ce56. *J Biotechnol* 121:174–191
- Peterson JA and Graham SE (1998) A close family resemblance: the importance of structure in understanding cytochromes P450. *Structure* 6:1079–85
- Petzoldt K. Process for the preparation of 11-beta-hydroxy steroids. Schering Aktiengesellschaft (Berlin, Germany). Patent 4353985
- Pinsach J, de Mas C, Lopez-Santin J, Striedner G, and Bayer K (2008) Influence of process temperature on recombinant enzyme activity in *Escherichia coli* fed-batch Cultures. *Enzyme Microb Technol* 43:507–512
- Presnell SR and Cohen FE (1989) Topological distribution of four-alpha-helix bundles. *Proc Natl Acad Sci USA* 86:6592–6596
- Pylypenko O, Schlichting I (2004) Structural aspects of ligand binding to and electron transfer in bacterial and fungal P450s. *Annu Rev Biochem* 73:991–1018
- Quaderer R, Omura S, Ikeda H, and Cane DE (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
- Redwan EL (2006) The optimal gene sequence for optimal protein expression in *Escherichia coli*: principle requirements. Arab J Biotechnol 9:493–510
- Reichenbach H and Hofle G (1993) Biologically active secondary metabolites from myxobacteria. Biotech Adv 11:219–277
- Reichenbach H (2001) Myxobacteria, producers of novel bioactive substances. Journal of Industrial Microbiology & Biotechnology 27:149–156
- Reiling K, Yoshikuni Y, Martin VJJ, Newman J, Bohlmann J, Keasling JD (2004) Mono and Diterpene Production in *Escherichia coli*. Biotechnology and bioengineering 87:200–212
- Rohdich F, Hecht S, Gärtner K, Adam P, Krieger C, Amslinger S, Arigoni D, Bacher A, Eisenreich W (2002) Studies on the nonmevalonate terpene biosynthetic pathway: metabolic role of IspH (LytB) protein. Proc Natl Acad Sci USA 99:1158–1163
- Rohmer M, Knani M, Simonin P, Sutter B, Sahn H. (1993) Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate. Biochemical Journal 295:517–524
- Rohmer M (1999) The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. Nat Prod Rep 16:565–574. doi:10.1039/a709175c. PMID 10584331
- Ruckpaul K, Rein H, Blanck J (1989) Regulation mechanism of the activity of the hepatic endoplasmic cytochrome P-450. In: Ruckpaul K, Rein H. (Eds.), Basis and Mechanism of Regulation of Cytochrome P450, 1. Akademie-Verlag, Berlin, pp1–55
- Rynkiewicz MJ, Cane DE, Christianson DW (2001) Structure of trichodiene synthase from *Fusarium sporotrichioides* provides mechanistic inferences on the terpene cyclization cascade. Proc Natl Acad Sci USA 98:13543–13548
- Sagara Y, Wada A, Takata Y, Waterman MR, Sekimizu K and Horiuchi T (1993) Direct expression of adrenodoxin reductase in *Escherichia coli* and the functional characterization. Biol Pharm Bull 16:627–630
- Sahdev S, Khattar SK, and Saini KS (2008) Production of active eukaryotic proteins through bacterial expression systems: A review of the existing biotechnology strategies. Mol Cell Biochem 307:249–264
- Sakaki T, Sawada N, Nonaka Y, Ohyama Y, Inouye K, (1999) Metabolic studies using recombinant *Escherichia coli* cells producing rat mitochondrial CYP24 CYP24

- can convert 1 $\alpha$ ,25- dihydroxyvitamin D<sub>3</sub> to calcitroic acid. *Eur J Biochem* 262:43–48
- Sambrook J, Russell DW (2001) *Molecular Cloning*, 3rd ed, Cold Spring Harbor Laboratory Press: New York
- Sanner MF (1999) Python: A Programming Language for Software Integration and Development. *J Mol Graph Model* 17:57–61
- Sauter IP, Dos Santos JC, Apel MA, Cibulski SP, Roehe PM, von Poser GL, Rott MB (2011) Amoebicidal activity and chemical composition of *Pterocaulon polystachyum* (Asteraceae) essential oil. *Parasitol Res*
- Schenkman JB, Sligar SG, and Cinti DL (1981) Substrate interaction with cytochrome P-450. *Pharmacol Ther* 12:43–71
- Schewe H, Kaup BA, Schrader J (2008) Improvement of P450(BM-3) whole-cell biocatalysis by integrating heterologous cofactor regeneration combining glucose facilitator and dehydrogenase in *E. coli*. *Appl Microbiol Biotechnol* 78:55–65
- Schiffler B, Bernhardt R (2003) Bacterial (CYP101) and mitochondrial P450 systems- how comparable are they?. *Biochem Biophys Res Commun* 312:223–228
- Schneiker S, Perlova O, Kaiser O, Gerth K, Alici A, Altmeyer MO, Bartels D, Bekel T, Beyer S, Bode E, Bode HB, Bolten CJ, Choudhuri JV, Doss S, Elnakady YA, 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 AC, Merai M, Meyer F, Mormann S, Munoz-Dorado J, Perez J, Pradella S, Rachid S, Raddatz G, Rosenau F, Rückert C, Sasse F, Scharfe M, Schuster SC, Suen G, Treuner-Lange A, Velicer GJ, Vorhölter FJ, Weissman KJ, Welch RD, Wenzel SC, Whitworth DE, Wilhelm S, Wittmann C, Blöcker H, Pühler A and Müller R (2007) Complete genome sequence of the myxobacterium *Sorangium cellulosum*. *Nat Biotech* 25:1281–1289
- Schuler MA, and Werck-Reichhart D (2003) Functional genomics of P450s. *Annu Rev Plant Biol* 54:629–667
- Schwede T, Kopp J, Guex N, and Peitsch MC (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res* 31:3381–3385
- Schwender J, Seemann M, Lichtenthaler HK, Rohmer M (1996) Biosynthesis of isoprenoids (carotenoids, sterols, prenyl side-chains of chlorophylls and plastoquinone) via a novel pyruvate/glyceraldehyde 3-phosphate non-

- mevalonate pathway in the green alga *Scenedesmus obliquus*. *Biochem J* 316:373
- Shirano Y and Shibata D (1990) Low temperature cultivation of *Escherichia coli* carrying a rice lipoxygenase L-2 cDNA produces a soluble and active enzyme at a high level. *FEBS Lett* 271:128–130
- Sillam-Dusses D, Semon E, Moreau C, Valterov I, Sobotnik J, Robert A, Bordereau C (2005) Neocembrene A, a major component of the trail-following pheromone in the genus *Prorehinotermes* (Insecta, Isoptera, Rhinotermitidae). *Chemoecology* 15:1–6
- Singer Y, Shity H, and Bar R (1991) Microbial transformations in a cyclodextrin medium. Part 2. Reduction of androstenedione to testosterone by *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 35:731–737
- Singh M, Sharma R, Banerjee UC (2002) Biotechnological applications of cyclodextrins. *Biotechnology Advances* 20:341–359
- Sono M, Roach MP, Coulter ED, Dawson JH (1996) Hemecontaining oxygenases. *Chem Rev* 96: 2841–2888
- Srivastava P, Bhattacharaya P, Pandey G, Mukherjee KJ (2005) Overexpression and purification of recombinant human interferon alpha2b in *Escherichia coli*. *Protein Expr Purif* 41:313–322
- Starks CM, Back K, Chappell J, Noel JP (1997) Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. *Science* 277: 1815–1820
- Szejtli J (1998) Introduction and General Overview of Cyclodextrin Chemistry. *Chem Rev* 98:1743–1753
- Takahashi S, Yeo Y, Greenhagen BT, McMullin T, Song L, Maurina-Brunker J, Rosson R, Noel JP, and Chappell J (2007) Metabolic engineering of sesquiterpene metabolism in yeast. *Biotechnol Bioeng* 97:170–181
- Takahashi S, Yeo YS, Zhao Y, O'Maille PE, Greenhagen BT, Noel JP, Coates RM, Chappell J (2007) Functional characterization of premnaspirodiene oxygenase, a cytochrome P450 catalyzing regio- and stereo-specific hydroxylations of diverse sesquiterpene substrates. *J Biol Chem* 282:31744–31754
- Takashi A, Tomoya K, Hiroshi Y (2007) Terpenes Isolated from Bark and Wood on the Seiridium Canker Infected *Chamaecyparis lawsoniana*. *Bull Col Edu Ibaraki Univ (Nat Sci)* 56:15–22
- Tao L, Jackson RE, Rouviere PE, Cheng Q (2005) Isolation of chromosomal mutations that affect carotenoid production in *Escherichia coli*: Mutations alter copy number of ColE1-type plasmids. *FEMS Microbiol Lett* 243:227–233

- Tassaneeyakul W, Guo LQ, Fukuda K, Ohta T, Yamazoe Y (2000) Inhibition selectivity of grapefruit juice components on human cytochromes P450. *Arch Biochem Biophys* 378:356–63
- Taylor M, Lamb DC, Cannell R, Dawson M, Kelly SL (1999) Cytochrome P450105D1 (CYP105D1) from *Streptomyces griseus*: Heterologous expression, activity, and activation effects of multiple xenobiotics. *Biochem Biophys Res Commun* 263:838–842
- Tholl D (2006) Terpene synthases and the regulation, diversity and biological roles of terpene metabolism *Current Opinion in Plant Biology* 9:1–8
- Trapp SC and Croteau RB (2001) Genomic organization of plant terpene synthases and molecular evolutionary implications. *Genetics* 158:811–832
- 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
- Urlacher VB, Lutz-Wahl S, Schmid RD (2004) Microbial P450 enzymes in biotechnology. *Appl Microbiol Biotechnol* 64:317–325
- Urlacher VB, Makhsumkhanov A, Schmid RD (2006) Biotransformation of beta-ionone by engineered cytochrome P450 BM-3. *Appl Microbiol Biotechnol* 70:53–9
- Urlacher VB and Eiben S (2006) Cytochrome P450 monooxygenases: perspectives for synthetic application. *TRENDS in Biotechnology* 24:324–330
- Urlacher VB and Girhard M (2011) Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. *Trends Biotechnol*
- Vadali RV, Fu Y, Bennett GN, San KY (2005) Enhanced lycopene productivity by manipulation of carbon flow to isopentenyl diphosphate in *Escherichia coli*. *Biotechnol Prog* 21:1558–1561
- van Beilen JB, Duetz WA, Schmid A, Witholt B (2003) Practical issues in the application of oxygenases. *Trends Biotechnol* 21:170–177
- Van Geldre E, Vergauwe A and Van den Eeckhout E (1997) State of the art of the production of the antimalarial compound artemisinin in plants. *Plant Mol Biol* 33:199–209
- Vijaya Sree N, Udayasri P, V.V. Aswani kumar Y, Ravi Babu B, Phani kumar Y, Vijay Varma M (2010) Advancements in the Production of Secondary Metabolites. *Journal of Natural Products* 3:112–123

- Volonte F, Marinelli F, Gastaldo L, Sacchi S, Pilone MS, Pollegioni L, and Molla G (2008) Optimization of glutaryl-7- aminocephalosporanic acid acylase expression in *E-coli*. *Protein Expr Purif* 61:131–137
- Wandrey C (2004) Biochemical reaction engineering for redox reactions. *Chem Rec4*:254–265
- Wang C, Yoon SH, Shah AA, Chung YR, Kim JY, Choi ES, Keasling JD, Kim SW (2010) Farnesol production from *Escherichia coli* by harnessing the exogenous mevalonate pathway 107:421-429
- Wang C-W, Oh M-K, Liao JC (1999) Engineered isoprenoid pathway enhances astaxanthin production in *Escherichia coli*. *Biotechnol Bioeng* 62:235– 241
- Wenzel SC and Müller R (2007) Myxobacterial natural product assembly lines: fascinating examples of curious biochemistry. *Nat Prod Rep* 24:1211–1224
- Wenzel SC and Müller R (2009) Myxobacteria—‘microbial factories’ for the production of bioactive secondary metabolites. *Mol BioSyst* 5:567–574
- Werck-Reichhart D, Feyereisen R (2000) Cytochromes P450: a success story. *Genome Biol* 1: REVIEWS3003
- Wichmann R, Vasic-Racki D (2005) Cofactor regeneration at the lab scale. *Adv Biochem Eng Biotechnol* 92:225–260
- Williams DC, McGarvey DJ, Katahira EJ, Croteau R (1998) Truncation of limonene synthase preprotein provides a fully active 'pseudomature' form of this monoterpene cyclase and reveals the function of the amino-terminal arginine pair. *Biochemistry* 37:12213–12220
- Williams JW and Morrison JF (1979) The kinetics of reversible tight-binding inhibition. *Methods Enzymol* 63: 437–67
- Williams-Smith DL, Cammack R (1977) Oxidation-reduction potentials of cytochromes P450 and ferredoxin in the bovine adrenal. Their modification by substrates and inhibitors. *Biochim Biophys Acta*. 499:432–442
- Wise ML and Croteau R (1998) in *Comprehensive Natural Products Chemistry: Isoprenoid Biosynthesis*, ed. Cane DE (Pergamon, Oxford), in press
- Yermakov AI, Khlaifat AL, Qutob H, Abramovich RA, Khomyakov YY (2010) Characteristics Of The GC-MS Mass Spectra Of Terpenoids (C<sub>10</sub>H<sub>16</sub>). *Chemical Sciences Journal CSJ*-7
- Yokoyama A, Shizuri Y, Misawa N (1998) Production of new carotenoids, astaxanthin glucosides, by *Escherichia coli* transformants carrying carotenoid biosynthetic genes. *Tetrahedron Lett* 39:3709–3712

- Yoon SH, Lee YM, Kim JE, Lee SH, Lee JH, Kim JY, Jung KH, Shin YC, Keasling JD, Kim SW (2006) Enhanced lycopene production in *Escherichia coli* engineered to synthesize isopentenyl diphosphate and dimethylallyl diphosphate from mevalonate. *Biotechnol Bioeng* 94:1025–1032
- Yoon SH, Park HM, Kim JE, Lee SH, Choi MS, Kim JY, Oh DK, Keasling JD, Kim SW (2007) Increased  $\beta$ -Carotene Production in Recombinant *Escherichia coli* Harboring an Engineered Isoprenoid Precursor Pathway with Mevalonate Addition *Biotechnol. Prog* 23:599–605
- Yu F, Harada H, Yamasaki K, Okamoto S, Hirase S, Tanaka Y, Misawa N, Utsumi R (2008) Isolation and functional characterization of a beta-eudesmol synthase, a new sesquiterpene synthase from *Zingiber zerumbet* Smith. *FEBS Lett* 582:565–572
- Yu F, Okamoto S, Harada H, Yamasaki K, Misawa N, Utsumi R (2011) *Zingiber zerumbet* CYP71BA1 catalyzes the conversion of  $\alpha$ -humulene to 8-hydroxy- $\alpha$ -humulene in zerumbone biosynthesis. *Cell Mol Life Sci* 68:1033–1040
- Yuan LZ, Rouviere PE, Suh W (2006) Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in *E. coli*. *Metab Eng* 8:79–90
- Zhao B, Lin X, Lei L, Lamb DC, Kelly SL, Waterman MR, and Cane DE (2008) Biosynthesis of the sesquiterpene antibiotic albaflavenone in *Streptomyces coelicolor* A3(2). *J Biol Chem* 283:8183–8189
- Zhao H, van der Donk WA (2003) Regeneration of cofactors for use in biocatalysis. *Curr Opin Biotechnol* 14:583–589
- Zhou K, Peters RJ (2009) Investigating the conservation pattern of a putative second terpene synthase divalent metal binding motif in plants. *Phytochemistry* 70:366–369
- Zhu BC, Henderson G, Chen F, Maistrello L, Laine RA (2001) Nootkatone is a repellent for Formosan subterranean termite (*Coptotermes formosanus*). *J Chem Ecol* 27:523–531

## 7. APPENDIX

### 7.1. Media

#### 7.1.1. Luria-Bertani (LB) Broth

Bacto-Tryptone	10 g
Yeast extract	5 g
NaCl	10 g
Add H <sub>2</sub> O	1000 ml

#### 7.1.2. Terrific Broth (TB)

Typtone	12 g
Yeast extract	24 g
Glycerin	4 ml
Add H <sub>2</sub> O	900 ml

After autoclaving, add 100 ml of sterilized potassium phosphate buffer:

K <sub>2</sub> HPO <sub>4</sub>	0.72 M
KH <sub>2</sub> PO <sub>4</sub>	0.17 M

#### 7.1.3. SOC medium

Tryptone	20 g
Yeast extract	5 g
NaCl	0.5 g

Add H<sub>2</sub>O 980 ml

After autoclaving, add 20 ml of sterilized 1M glucose

## 7.2. Preparation of chemically competent cells

### 7.2.1. Transfer buffer I (TFPI)

1M CH <sub>3</sub> COOK	3 ml
1M KCl	10 ml
1M CaCl <sub>2</sub>	1 ml
Glycerol	12 ml

Filled up to 80 ml with distilled water, adjusted pH to 6.1

Add 1M MnCl<sub>2</sub> 5 ml

Filled up to 100 ml with distilled water, pH should be 5.8.

### 7.2.2. Transfer buffer II (TFPII)

0.2M MOPS	1 ml
1M KCl	0.2 ml
1M CaCl <sub>2</sub>	1.5 ml
Glycerol	2.4 ml

Filled up to 20 ml with distilled water, pH should be 6.5-7.0.

Both of TFPI and TFPII was sterilized by filtration and stored in the refrigerator at 4°C before used

## 7.3. Buffers and solutions for protein purification

**7.3.1. Cell lysis buffer:** 50 mM Sodium phosphate buffer, pH 8.0; 300 mM NaCl; 20% Glycerin; 0.1 mM PMSF (stock solution: 100 mM in isopropanol)

### 7.3.2. Buffer for affinity chromatography:

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

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

**7.3.3. Buffer for size exclusion chromatography:** Potassium phosphate buffer (10 mM, pH 7.4) containing 20% glycerol.

## 7.4. Sodium dodecylsulphate (SDS) polyacrylamid gelelectrophoresis

### 7.4.1. Gel components

Stock solutions	Separating gel (15%)	Stacking gel (5%)
4 x LT	3.75 ml	-
4 x UT	-	2.5 ml
10% APS	75 $\mu$ l	50 $\mu$ l
dest.H <sub>2</sub> O	ad 15 ml	Ad 10 ml
30% AA/Bis	7.5 ml	1.6 ml
TEMED	10 $\mu$ l	5 $\mu$ l

### 7.4.2. Buffer and solution

- **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; 10%  $\beta$ -mercaptoethanol; 0.004% Bromphenol blue
- **Staining solution:** 0.1% Coomassie Brilliant Blue G-250; 40% Methanol; 10% Acetic acid
- **Destaining solution:** 25% Methanol; 10% Acetic acid

### 7.5. The DNA sequence of gene coding for CYP264B1

ATGACTCGACTTAATCTGTTCCGCCCCGAGGTCCGAGAGAACCCGTACCCCTTCTACGCGGCGCTGCGCC  
 GCGAGTCGCCCGTGTGTGTCAGGTGGACCCCAACGGGATGTGGGTGGTCACCCGGTATGATGACATCGTGGC  
 GGCGTTCAAGAACACGCAGGTGTTCTCGTCCGCGGGGCTGCGCATGGCCACCGAGCCTCCTTACCTGCGG  
 CGTCAAAATCCATTGTCCGGATCCATGATCCTCGCGGACCCGCCGCGGCACGGACAGCTGCGCAGCATCA  
 GCAGCCGTGCGTTACCCGCCAACATGGTGAGCAGCTGGAGCACCACATGCGATCCATGGCCGTGCGGCT  
 CACGGACGACCTGGTGCACCGTCGCGTCTCGAGTTTATCTCCGAGTTCGCGAGCCGCGCGAGGTCAGC  
 GTCCCTCGCCAAGTTGATCCGGATTTGATCCAGGGCTAGAGGGCACTTCAAGCGATGGGGCGACTGACCTGG  
 TCATCGTTGGCGTCATACCTCCGGAGGATCATGCGCGGATCGCCGAGGTGCGCAGGACCATCGATGAGAT  
 GGAGCAGTACATGCTGGGCCTGCTGGCGAGCCGCCGCGTCAATTTGGAGAACGATCTGGTAAGTGAGCTG  
 CTCCGATCTCGGAGGGACGACGATGGCATCACGGACCAGGACCTCGTGTCCCTCCTCTCGCTGCTCCTCG  
 TGGCAGGGCTCGAAACGTCGACATCGCTGATGACCCATATGGTCTTGATCCTCGCCAGCGGCCGATGTG  
 GATGGATCGCCTACGTGCGGAGCCCGCGCTCATCCCGATTTTCATCGAGGAGGTGATGCGTTTTGAGGGC  
 CCTGTGCATGCCACGATGCGACTGACCGTCACCGAGACGGAGCTCGGCGGTACGCGTCTGCCCGCTCACG  
 CCGTGGTTCGCGTCTCATCAGCTCGGGTCTGCGCGATGAGGCGAGGTTCCAGGAGCCGGACCGCTTCAA  
 CCCAGAGCGCGGCGATCAGGCCAACCTGGCATTTCGGCCATGGTGCCCACTTCTGCCTCGGGGTCTTCTG  
 GCCCGGTGTCAGGCCCGCATCGTGTGGAGGAGCTCCTCCGTGCGGTGCCACCGCATCGTGTCTCGTACCG  
 ACCGTCTCGAGTGGCAAGCGGCTCTCAACACGCGCTCCCCGGTGGCGCTGCCCATCGAGGTGATCCCCGT  
 CTCGACTACCGCCGCACGTGAGTCGCCGGTTCGTCCAGGGCATCTGGTAG

### 7.6. The protein sequence of CYP264B1

MTRLNLFAPVRENYPYFYAALRRESPVCQVDPNGMWVTRYDDIVAAFKNTQVFSSAGLRMATEPPYLR  
 RQNPLSGSMILADPPRHGQLRSISSRAFTANMVSTLEHHMRSMVRLTDDLVRHRRVVEFISEFASRAQVS  
 VLAKLIGFDPGLEGHFKRWATDLVIVGVIPEPDHARIAEVRRTIDEMEQYMLGLLASRRRHLENDLVSEL  
 LRSRRDDGITDQDLVSLLSLLLAVAGLETSTSLMTHMVLILAQRPMWMDRLRAEPALIPHFIEEVMRFEA  
 PVHATMRLTVTETELGGTRLPAHAVVALLISSGLRDEARFQEPDRFNPERGDQANLAFGHGAHFCLGVFL  
 ARVQARIVLEELLRRCHRIVLRTDRLEWQAALNTRSPVALPIEVI PVSTTAARESPPVQGIW

### 7.7. The DNA sequence of gene coding for GeoA

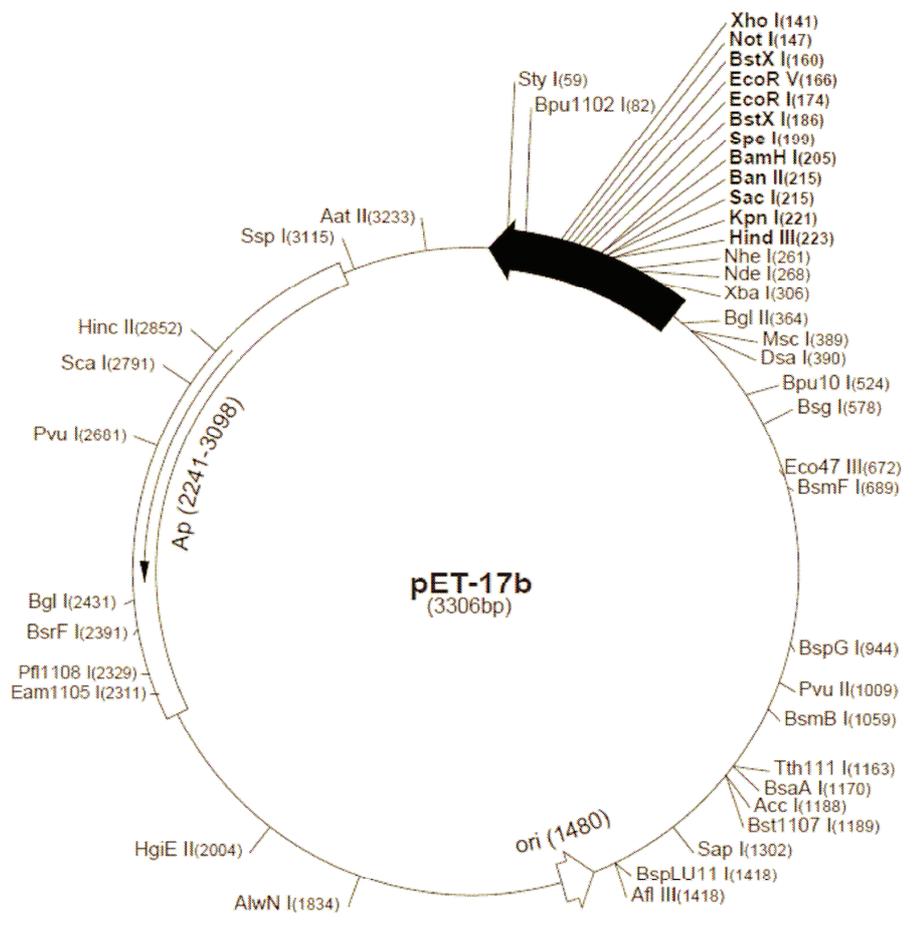
ATGTCATCTGATCGAACTAGCGTTGTCGTGAGCAAGCGCGACGCTGGTGGATTGGAATATCCGTTTGCTG  
 CGTCGTGTCACCCCGGGAGAGAGGTCACGGAGCAAAGGACCCTGGCGTGGGTTTCGCCGTTGCGTCTGGT  
 GCCAGACGGTTCGGTTCGCTTTACGTCTCAAGGCGACGAACTTCTCCACCTGGCGGCATGGCTCTTACCG  
 TCCGCGAGCACGCAGACGCTCCAGCTGGCATCGGACTTACGGCCGTGCTGTTCTCCTGGATGATGCGT  
 ACGACGAGGGGCGACTCAGCACGGACCCGGAATCCGTGAGTGGCTCAACGAGAAGTACCTGGGTGAGCT  
 CTTCCGGTACACCGAAGCGGACATGAGCGACCCCTGACGCGCGGGATGCTCGATGTACGTGAGCGCATC  
 CGACGCAGCCATCCGCACTTCTTCTGAATCGCTGGCTCTCGCACTTTCAGTACTACTACGAGGCCAACC  
 TGTGGGAGGCGAATAACCGAAAGCAGATGAGGGTGCCGCATCTTGAAGAATATTTGATGATGCGGCGCTA  
 CAGCGGCGCTGTCTATACCTACTGTGACCTGCTCGAGCTGCTGCTCGAACCGCCGCTGCCGCTCGAGGTC  
 GTCCAGCACCCGTTGATTGAGACGGTTCGAGACATCTGTAATGACATCCTCTGCTGGACGAACGACTACT  
 TCTCGCTCGGGAAGGAGCTGACCAACGGCGAGACGCACAACCTGATCGTTCGTTCTCCGGAACGAATGCGT  
 GAGCACGCTGGAAGAGGCGATCGACCGGCTCAAGGACATGCACGACCCCGGGTTCGCGGAGTACCAAGGC  
 GTGAAGGAGAAGGTGCTCGCGCTCTGGGCGGACGACGAGATCCGCCTTACCTGGACGCGGTTCGAGGCCA  
 TGATCGCCGGAATCAGCGTGGGCGCTCGAGGCCGGTTCGCTATTCTGGACTAGAAAGTTTGTATCGTGC  
 CGCTGGATGA

### 7.8. The protein sequence of GeoA

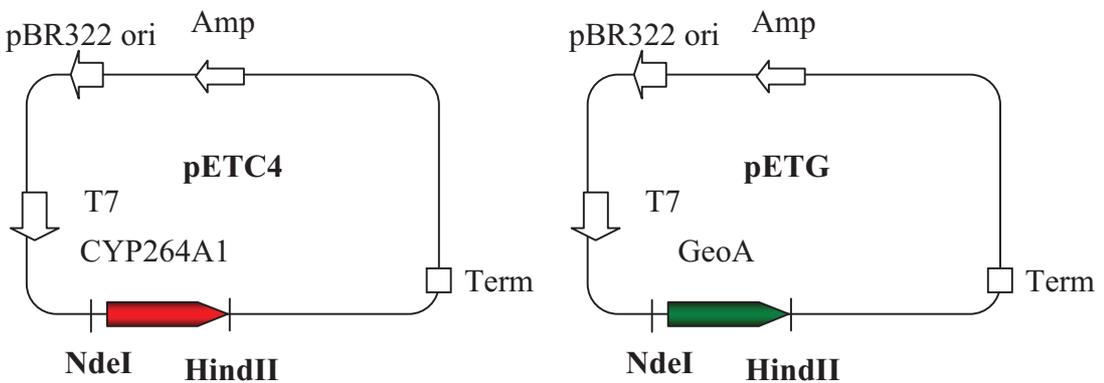
MSSDRTSVVVSKRDAGGFYEPFAASCHPGREVTEQRTLAWVRRLRLVDPGRSLRSLKATNFSHLAAWLLP  
 SASTQTLQLASDFTAVLFLDDAYDEGQLSTDPESVEWLNEKYLGELEFGYTEADMSDPLTRGMLDVRERI  
 RRSHPHFLLNRWLSHFQYYEANLWEANNRQMRVPHLEEYLMRRYSGAVYTYCDLLELLERPLPLEV  
 VQHPLIQTVRDI CNDI LCWTDYFSLGKELTNGETHNLI VVLRNECVSTLEEAIDRLKDMHRRVAEYQG  
 VKEKVLALWADDEIRLYLDAVEAMIAGNQRWALEAGRYSGLLESIVRAG

## 7.9. Vector maps

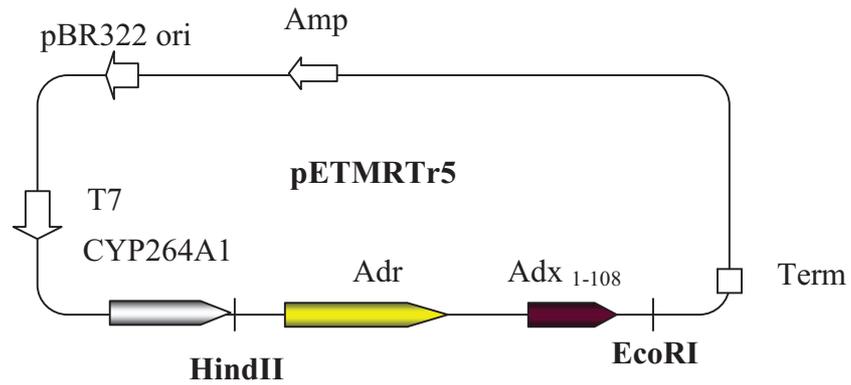
### 7.9.1. pET17b



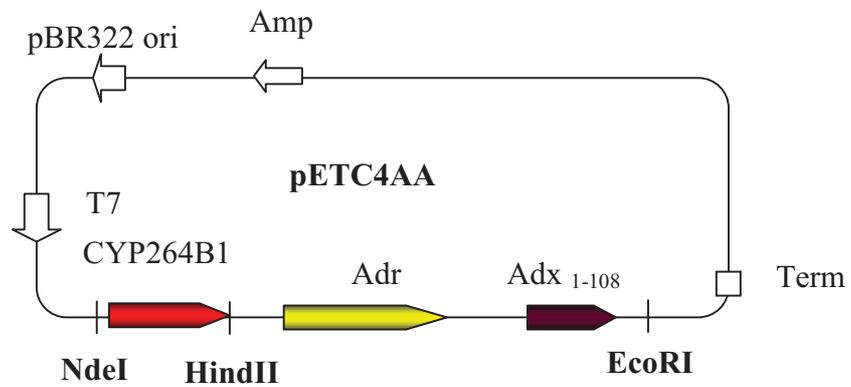
### 7.9.2. pETC4 and pETG



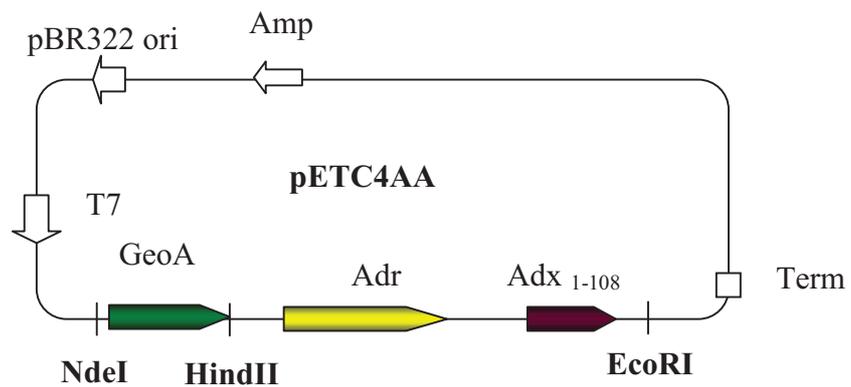
## 7.9.3. pETMR5



## 7.9.4. pETC4AA

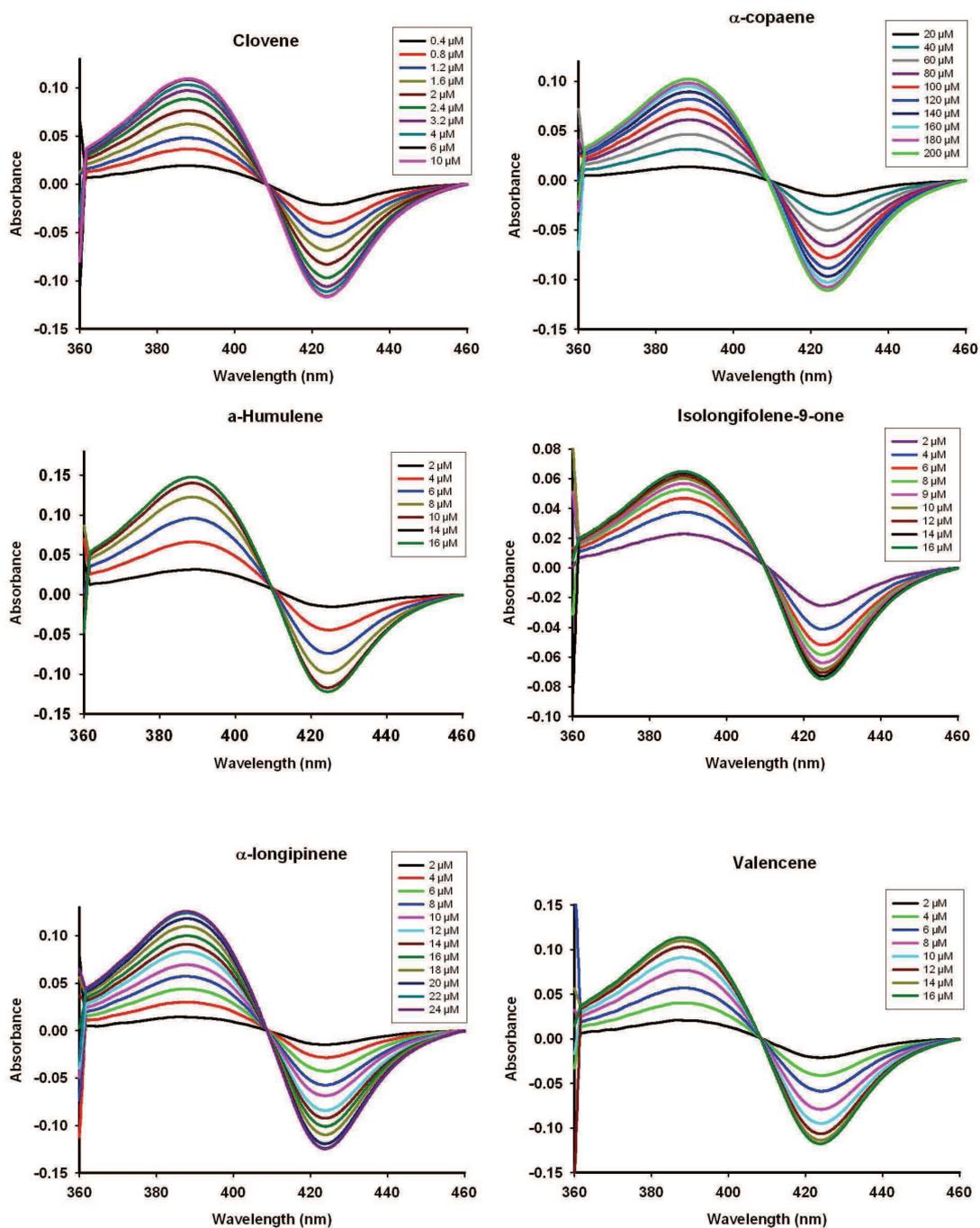


## 7.9.5. pETGAA





## 7.11. Type I spectra changes of CYP264B1 upon binding of some sesquiterpenes



## **STATEMENT / ERKLÄRUNG**

I hereby declare that I have independently done this dissertation. I did not use any unauthorized assistance and unmentioned materials.

Hiermit erkläre ich an Eides statt, die vorliegende Dissertation selbstständig angefertigt zu haben. Ich habe keine unerlaubten sowie unerwähnten Hilfen benutzt.

Saarbücken, 13.10.2011

Thuy Thi Bich Ly

## ACKNOWLEDGMENT

First of all, I would like to express my deepest gratitude to my supervisor, Prof. Rita Bernhardt for introducing me to this project and for providing me the opportunity to work in this very interesting research area as well as excellent scientific facilities. She kindly supported me financially during my prolongation.

I am grateful to Prof. Dr. Rolf Müller for his cooperation and collaboration with Prof. Bernhardt's group, letting me work in myxobacterium *Sorangium cellulosum* project.

I would like to express my deepest thanks to Department of Vietnamese International Education Development, Ministry of Education and Training for granting me a Ph.D. scholarship. My thanks also go to the German Academic Exchange Service (Deutscher Akademischer Austausch Dienst) for supporting me a language course as well as useful and interesting occasions to meet other international students in Germany.

I am grateful to Dr. Frank Hannemann for his scientific suggestions. His questions used to be special valuable for me in directing this work.

I am so thankful to Dr. Yogan Khatri for his kind help. He not only guided me at the beginning of my work here, but also is always willing to share his excellent experiences.

My special thanks go to Dr. Kerstin Maria Ewen for her suggestions in our group meetings. I so appreciate for her critical reading, language correction as well as valuable comments for my draft. Many thanks go to Dipl. Chem. Michael Ringle for his valuable suggestions for my work.

I am so grateful to Dipl. Biol. Britta Wilzewski for her kind help and sharing with me when I had healthy trouble.

I am indebt to Dr. Norio Kagawa and so thankful for sharing his profound experiences of P450 expression during his stay in our group. My special thanks go to Mr. Wolfgang Reinle for excellent purifications of AdR and Adx.

I would like to thank Mr. Dominik Pistorious, Department of Pharmaceutical Biotechnology, Saarland University, for his kind help in GCMS analysis. My thanks go to Dr. Josef Zapp, Department of Pharmaceutical Biology, Saarland University, for helping me in NMR analysis, and many thanks also go to Dr. Michael C. Hutter, Center for Bioinformatics, Saarland University, for his kind help in Enzyme homology modeling and substrate docking.

Many thanks go to Dr. Nils Günnewich for his cooperation in working with terpene cyclase and for his kind help to solve personal things.

I am so thankful to Dipl. Chem. Alexander Shiffin for his language checking for my draft as well as for his language translation. I also thank Dr. Sandra Giovanna Salamanca-Pinzón for her language checking for my draft, and thank Dipl. Biol. Jens Neunzig for his language translation.

I would like to thank our secretary Ms. Gabi Schon for helping me in every administrative aspect.

My thanks also go to my colleagues in our laboratory: Dipl. Biol. Adrian Gerber, Dipl. Biol. Michael Kleser, M. Sc. in Biotech. Anna Hobler, Dipl. Biol. Simon Janocha, Dipl. Biol. Simone Anderko, Dipl. Biol. Elisa Brill, Dipl. Biol. Danella Rauf, Dr. Huy Hoang Nguyen, Dr. Kim Thoa Nguyen, Dr. Andreas Bichet, Dr. Anja Berwanger, Dr. Berna Mersinli, Dr. Tarek Hakki, Dipl. Chem. Sabrina Bleif, Dipl. Biol. Tin Ming Kwai, Dipl. Biol. Bernd Lauer, Dr. Martin Wörner and our technician Katharina Bompais and Dipl. Biol. Antje Edien Plach for their help during my staying here and for sharing nice time and memories.

I am so grateful to Prof. Dr. Nguyen Thi Ngoc Dao for introducing me to our work group.

My deep thanks and love go to my family: my husband and son, my mother, my sister, my brother-in-law, especially my parent-in-law for their moral support, understanding, encouragement and taking care of our son during my PhD time.

Finally, I would like to express my admiration to Germany, a beautiful country with nice people, excellent-organized society and many nice things. I believe that the time I have been here is the most memorable time of my life. Thanks for everything.