Enabling the P450 complement of *Beauveria bassiana* and *Rhodococcus jostii* for biocatalysis

Claudia Spandolf

PhD

University of York

Chemistry

October 2015

Abstract

Enzyme discovery today is proceeding at an enormous pace due to ever-growing technology development. As a result of this, more than 21000 cytochrome P450s have been identified in all kingdoms of life to date, making a wide range of enzyme resources with outstanding potential for biocatalytical implementation available. P450s from filamentous fungi, such as *Beauveria bassiana*, represent particularly compelling targets for the discovery of novel enzymes as these organisms have a long history of application in industrial hydroxylation reactions, many of which are believed to be P450-dependent. In addition, the genome sequence of *B. bassiana* has recently been completed revealing 83 putative P450s.

In order to uncover new cytochrome P450-based biocatalysts from the fungus *Beauveria bassiana* extensive bioinformatics analysis of the *Beauveria* CYPome were performed. As a result 7 genes encoding for heme domains with possible alkane hydroxylase function and one encoding a naturally fused P450 with homology to $P450_{foxy}$ from *Fusarium oxysporum* could be identified for subsequent cloning, heterologous expression and characterization. Different expression hosts as well as various expression conditions have been investigated. Despite our efforts, delivery of active biocatalysts could not be realized. However, empirical data acquired in this project will be of value for future studies of fungal P450s.

In addition, 23 cytochrome P450 heme domains from *Rhodococcus jostii* fused to the P450 reductase domain (RhfRED) of cytochrome $P450_{Rhf}$ from *Rhodococcus* sp. NCIMB 9784 have been investigated in a further strand of this work and provided a screening platform that could be applied for industrial purposes.

Table of Contents

Abstract ..... **III**
Table of Contents ..... **V**
List of Figures ..... **IX**
List of Tables ..... **XIII**
Acknowledgements ..... **XV**
Author's Declaration ..... **XVI**
1. Introduction **1**
1.1 Cytochrome P450 1
1.1.1 Discovery and Nomenclature 1
1.1.2 Structure and Classes 5
1.1.3 Catalytic Reaction Mechanism 12
1.1.4 Evolution and Catalysed Reactions 16
1.1.5 Applications 18
1.1.5.1 Potential and Limitations of P450s in biocatalysis 18
1.1.5.2 Commercial applications 21
1.1.5.3 Medicine 28
1.1.5.4 Bioremediation 29
1.1.6 The fungal Kingdom 31
1.1.6.1 Fungi – the double-edged sword 31
1.1.6.2 Fungal P450 33
1.2 *Beauveria bassiana* 35
1.3 *Rhodococcus jostii* 37
1.4 Project Aim 38
2. General Material and Methods **39**
2.1 Chemicals 39
2.2 Strains and Plasmids 39
2.2.1 *Escherichia coli* strains 39
2.2.2 Yeast strains 40
2.2.3 Plasmids 40
2.3 Media 42
2.3.1 *Escherichia coli* 42
2.3.2 Yeast 43
2.3.3 Supplements 44

2.4 Glycerol stocks 44
2.5 Working with nucleic acids 45
2.5.1 Enzymatic restriction of DNA 45
2.5.2 Plasmid DNA preparation 45
2.5.3 Purification of DNA fragments 45
2.5.4 Agarose gel electrophoresis 46
2.5.5 *In vitro*-Amplification of DNA by PCR 46
2.5.6 Ligation independent cloning method 47
2.5.7 In Fusion cloning 49
2.6 Preparation of recombinant microorganisms 51
2.6.1 Preparation of $CaCl_2$-competent *E. coli*-cells 51
2.6.2 Plasmid transformation into BL21 (DE3) and Rosetta2 (DE3) competent *E. coli* cells 52
2.6.3 Plasmid transformation into NovaBlue competent *E. coli* cells 52
2.6.4 Plasmid transformation into *Saccharomyces cerevisiae* WAT11 52
2.7 Working with Proteins 53
2.7.1 Cell growth and protein expression 53
2.7.1.1 Small scale expression test 54
2.7.1.2 Scaled up protein expression 54
2.7.2 Protein purification 54
2.7.3 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) 55
2.7.4 Western blot 57
2.7.4.1 Western blot set up and protein transfer 57
2.7.4.2 Immunoprecipitation 58
3. Selection of *Beauveria bassiana* P450 targets 60
3.1 In situ analysis of CYPome 60
3.2 Phylogenetic tree 64
3.3 Discussion 65
4. Construction of a library of *B. bassiana* P450 genes and their expression in *E.coli* 69
4.1 Materials and Methods 69
4.1.1 Modification of P450s in preparation of synthesis 69
4.1.2 Cloning of P450s 70
4.1.3 Protein characterization 71
4.1.3.1 Spectrophotometric characterisation 71
4.1.3.2 Whole cell activity assay 71

4.1.3.3 Gas chromatography 71
4.2 Results 72
4.2.1 Cloning into LICRED and LIC-3C 72
4.2.2 Expression test 74
4.2.2.1 Media optimization 76
4.2.2.2 Co-expression with chaperones 80
4.2.3 Whole cell activity assay 84
4.2.4 Spectrophotometric characterisation 85
4.2.5 Purification of CYP539B1_LICRED 86
4.3 Discussion 87
5. Construction of library of a *B. bassiana* P450s in pYeDP60 and expression in *Saccharomyces cerevisiae* WAT11 91
5.1 Materials and Methods 91
5.1.1 Reintroduction of the N-terminal region 91
5.1.2 Cloning of P450 91
5.1.3 Expression 93
5.1.4 Microsome preparation 94
5.2 Results 95
5.2.1 Cloning into pYeDP60 95
5.2.2 Transformation into *S. cerevisiae* 97
5.2.3 Expression test in *Saccharomyces cerevisiae* WAT11 98
5.3 Discussion 100
6. A P450 fusion library of heme domains from *Rhodococcus jostii* RHA1 and its evaluation for the biotransformation of drug molecules 104
6.1 Introduction 104
6.2 Material and Methods 105
6.2.1 Screening for drug metabolites using resting whole cells 105
6.2.2 Screening for drug metabolites using purified Ro07-RhfRED 106
6.2.3 *In vitro* activity assays towards imipramine using purified Ro07-RhfRED 107
6.2.4 Analysis using UPLC-MS^E and data processing 107
6.3 Results 108
6.3.1 Expression test in *E. coli* 108
6.3.2 Screening for drug metabolites using resting whole cells 111
6.3.3 Purification and characterisation of Ro07-RhfRED 112
6.3.4 Biotransformation of imipramine 115
6.4 Discussion 115

7. Final discussion **120**
Appendix A **125**
Appendix B **130**
Appendix C: LC-MS/MS analysis of *Beauveria bassiana* proteome **135**
C.1 Introduction 135
C.2 Materials and methods 135
C.3 Results and discussion 138
Appendix D **140**
Abbreviations **151**
References **158**

List of Figures

Figure 1.1: Schematic distribution of cytochromes P450 among the different kingdoms (August 2013) 2

Figure 1.2: Selected milestones in cytochrome P450 research (1962-2000) 3

Figure 1.3: Scientific articles published on National Center for Biotechnology Information. 4

Figure 1.4: Structure of cytochrome P450 using the *Pseudomonas putida* camphor hydroxylase $P450_{cam}$ (CYP101A1) as an example. 5

Figure 1.5: Topographic map showing the secondary elements commonly found among P450 enzymes[44] 6

Figure 1.6: Map of the signature motifs in the P450 proteins of an ER-bound P450 protein (class II enzyme) 7

Figure 1.7: Prosthetic of cysteinato-heme enzymes 13

Figure 1.8: The three electronic configurations of molecular oxygen, O_2. 14

Figure 1.9: The cytochrome P450 catalytic cycle. 15

Figure 1.10: Schematic summary of the diverse P450-catalysed reactions[2] 17

Figure 1.11: P450-catalysed C–C bond coupling in isoquinoline alkaloid biosynthesis of plants 18

Figure 1.12: Industrial potential of CYP450-catalysed bioprocesses ([139], altered) 22

Figure 1.13: Schematic representation of the engineered artemisinic acid biosynthetic pathway[146] 24

Figure 1.14: Schematic representation of the final step in progesterone biosynthesis by engineered yeast. 25

Figure 1.15: Anthocyanidins and flavonoid biosynthetic pathway relevant to flower colour. 27

Figure 1.16: Schematic representation of gene directed enzyme pro-drug therapy (GDEPT) 29

Figure 1.17: Oxidation of PCDDs by rat CYP1A1 and CYP1A2 expressed in *S. cerevisiae*. 30

Figure 1.18: Phylogeny of the fungal kingdom[174] 32

Figure 1.19: Growth of annotated genomes in MycoCosm (fungal genomics portal; update 2013).[190] 33

Figure 1.20: Relative P450ome sizes of fungal species belonging to different fungal (sub)phyla, grouped by their morphology.[1] 34

Figure 1.21: A spider colonized by *B. bassiana* 36

Figure 1.22: Variety of *B. bassiana*'s reaction attributable to cytochrome P450 37

Figure 2.1: Vector map of LICRED 41

Figure 2.2: LIC-RED cloning site 41

Figure 2.3: pYeDP60 shuttle vector map 42
Figure 2.4: Scheme of T4 polymerase treatment 48
Figure 2.5: Assembly scheme of the In-Fusion® HD Cloning System 50
Figure 2.6: Western blot set up 58
Figure 2.7: Immunoprecipitation after Western blot 59
Figure 3.1: Sequence logos of the conserved CYP motifs from 47 tested fungi[177] 60
Figure 3.2: Alignment of 33 P450s with relevant CYP-numbers from *B. bassiana* in accordance with their conserved motifs 63
Figure 3.3: Alignment of self-sufficient P450s from *B. megaterium*, *F. oxysporum* and *B. bassiana* with regard to conserved motifs 64
Figure 3.4: Phylogenetic tree of 33 P450 heme domains and 1 putative natural fusion gene from *B. bassiana* in relation to 2 well-known self-sufficient P450s 65
Figure 4.1: linearized LICRED vector and PCR products after gel extraction 72
Figure 4.2: Initial screen of positive transformants using restriction enzymes *Xba*I and *Nhe*I 73
Figure 4.3: Initial screen of positive transformants using restriction enzymes XbaI and *Nhe*I 73
Figure 4.4: Initial screen of positive transformants using restriction enzymes *Xba*I and *Nhe*I 73
Figure 4.5: Western blot analysis of selected P450s 74
Figure 4.6: Western blot analysis of selected P450s 74
Figure 4.7: Western blot analysis of selected P450s 75
Figure 4.8: SDS gel of in *E. coli* expressed CYP52G6 and CYP52G8 grown on variable carbon sources 77
Figure 4.9: SDS gel of in *E. coli* expressed CYP584E2 and CYP52T1 grown on variable carbon sources 78
Figure 4.10: SDS gel of in *E. coli* expressed CYP52T1_2 and CYP52T1_3 grown on variable carbon sources 78
Figure 4.11: SDS gel of in *E. coli* expressed CYP539B1 grown on variable carbon sources 79
Figure 4.12: SDS gel of in *E. coli* expressed CYP505A1 without and with transmembrane domain grown on variable carbon sources 79
Figure 4.13: SDS gel of CYP52G6, CYP52G8 and CYP52T1_3 co-expressed with chaperones in *E. coli* grown on variable carbon sources 81
Figure 4.14: SDS gel of CYP52T1_2, CYP52T1 and CYP584E2 co-expressed with chaperones in *E. coli* grown on variable carbon sources 82
Figure 4.15: SDS gel of CYP539B1 and CYP505A1 co-expressed with chaperones in *E. coli* grown on variable carbon sources 82
Figure 4.16: Western blot of soluble expressed P450s using the GroEL/ES chaperon system and glucose as carbon source 83

Figure 4.17: Western blot of soluble expressed P450s using the GroEL/ES chaperon system and glycerol as carbon source 83

Figure 4.18: UV-visible absorbance spectra for CYP539B1_LICRED....85

Figure 4.19: Nickel affinity chromatography chromatogram for the purification of 539B1_LICRED. 86

Figure 4.20: SDS-PAGE analysis of CYP539B1_LICRED fractions after nickel affinity chromatography....87

Figure 5.1: *S. cerevisiae* colonies on SGI plate. 93

Figure 5.2: PCR products incorporating the C-terminal his-tag after gel extraction....95

Figure 5.3: Initial screen of positive transformants containing a C-terminal his-tag using restriction enzymes *Bam*HI and *Kpn*I 96

Figure 5.4: Colony PCR of transformed *S. cerevisiae* (pYeDP60 constructs containing genes of interest with C-terminal his-tag) 97

Figure 5.5: Colony PCR of transformed *S. cerevisiae* (pYeDP60 constructs containing genes of interest with C-terminal his-tag) 98

Figure 5.6: UV-visible absorbance spectra for CYP52T1_DP60 containing either no his tag, N-terminal his-tag or C-terminal his-tag in comparison to empty pYeDP60 99

Figure 6.1: Phylogenetic tree constructed of *Rhodococcus jostii* RHA1 gene targets (blue) Included with the RHA1 P450 genes are various cytochromes P450 heme domains of known structure designated by their PDB codes....104

Figure 6.2: Oxidative activities attributed to P450 from Rhodococcus jostii RHA1....105

Figure 6.3: experimental setup for *E. coli* whole cell activity assays and UPLC-MS^E analysis 106

Figure 6.4: Western blot analysis of Ro08-RhfRED expression carried out at varied temperatures 109

Figure 6.5: Western blot analysis of soluble expressed X-RhfRED fusions at 16°C 110

Figure 6.6: Western blot analysis of soluble expressed X-RhfRED fusions at 16°C 110

Figure 6.7: Drug molecules from the screen that showed significant levels of transformation against negative controls when incubated with P450RHA1-RhfRED fusions. 111

Figure 6.8: Nickel affinity chromatography chromatogram for the purification of Ro07-RhfRED 113

Figure 6.9: SDS-PAGE analysis of Ro07 RhfRED fractions after nickel affinity chromatography....113

Figure 6.10: UV-visible absorbance spectra of purified Ro07-RhfRED 114

Figure 6.11: Biotransformation of imipramine with purified Ro07-RhfRED 115

Figure 6.12: Biotransformation with purified Ro07-RhfRED of diltiazem at different substrate concentrations[279]....118

Figure 6.13: Conversions of imipramine 1 and analogs 2-11 and their relationship to physico-chemical properties of the substrates catalysed by Ro07-RhfRED [279] 118

Figure 7.1: Biotransformation of lauric acid using microsomes from yeast expressing *CYP52X1* 122

Figure C.1: substrates added for biotransformation in *B. bassiana* 136

List of Tables

Table 1.1: Classes of P450 redox partners [50, adapted] 9
Table 1.2: Challenges and limitations for biotechnological application of CYPs 20
Table 2.1: List of *E. coli* strains used for cloning and recombinant expression 39
Table 2.2: *Saccharomyces cerevisiae* strain used in this work 40
Table 2.3: Plasmids for gene cloning and enzyme expression....40
Table 2.4: Supplements for bacteria and yeast media....44
Table 2.5: List of restriction endonucleases....45
Table 2.6: Components of the PCR reaction....47
Table 2.7: PCR reaction conditions 47
Table 2.8: Protocol for T4 polymerase treatment 48
Table 2.9: In-Fusion cloning reaction set up....49
Table 2.10: Yeast transformation mix suitable for 1 transformation 53
Table 2.11: SDS-PAGE gel solutions 56
Table 2.12: BIO-RAD Protein Molecular Weights in dalton 56
Table 3.1: CYP family distribution of 33 analysed *B. bassiana* P450s in accordance to similarity with P450s of known or assumed function 67
Table 4.1: Properties of selected heme-domains from *B. bassiana* for cloning and expression....69
Table 4.2: Primer used for PCR amplification....70
Table 4.3: Relative soluble and insoluble expression of Cytochrome P450s in the presence of different carbon sources at 16 ˚C. 80
Table 5.1: Primer pairs used for PCR amplification incorporating a N-terminal his tag.....92
Table 5.2: Primer pairs used for PCR amplification incorporating a C-terminal his tag.....92
Table 5.3: Primer pairs used for PCR amplification incorporating no additional his tag....93
Table 5.4: Primer pair used for colony PCR 94
Table 5.5: Relative nucleotide length of amplification products 96
Table 5.6: relative nucleotide length of colony PCR amplification products 98
Table 5.7: comparison of codon usage of 4 major aa in *E. coli* and *S. cerevisiae* 101
Table 5.8: Proportion of low usage codons in gene-sequences of *B. bassiana* heme-domains optimised for *E. coli* expression 102
Table 6.1: experimental set up for imipramine kinetic assay with purified Ro07-RhfRED.... 107
Table 6.2: Biotransformation spectra of heterologously expressed cytochrome P450 fusion proteins from *R. jostii* RHA1 in whole cells. 112
Table A. 1: Cytochromes P450 in *B. bassiana* ARSEF 2860....125

Table A.2: Cytochromes P450 in *Rhodococcus jostii* RHA1 126

Table A.3: Primer pairs used for PCR amplification of N-terminal anchor and truncated P450 heme domain from *B. bassiana* 127

Table A.4: Primer pairs used for PCR amplification of P450 targets from *R. jostii* 128

Table C.1: P450 hits on exposure to hydroxylation substrates in *Beauveria bassiana* using LC-MS analysis 139

Table D.1: List of the 65 substrates used in activity assays of *Rhodococcus jostii*..... 140

Acknowledgements

Generating a PhD thesis is without a doubt not an easy task to accomplish, both professionally and personally, but I never imagined it to turn into one of the greatest challenges of my life. A challenge I would not have succeeded if it had not been for some very special people, I wish to express my deepest gratitude for.

First and foremost, I would like to thank my genius supervisor **Gideon Grogan**, not only for giving me the opportunity to be part of the P4FIFTY network, which allowed me to meet and work with great scientists and travel to amazing places throughout the world, but also for being – what I consider – a friend. At this point I can honestly say that these last 3 years would have turned out to be more than just a struggle without your guidance and support. For this I am truly grateful.

I also would like to thank all the people in the Grogan group, in particular **Chantel Jensen, Henry Man** (Wingman), **Muhiadin Omar** (Muhi), and **Elizabeth Wells** (Lilly) [listed in decreasing order of priority ☺], who have been with me from the very beginning and always made me feel welcome. I will miss you!

Special thanks go to all the people from the P4FIFTY network who never failed to make progress meetings so much more interesting and enjoyable. **Justyna Kulig**, **Sandra Notonier** and **Tina Ilc** deserve particular attention as they did an incredible good job in supervising (and entertaining) me during my internships. I also would like to express my gratitude towards **Neil Bruce** and **Margaret Cafferky** who spared no effort in creating this network and make it such a success. Last but not least I would like to mention **Maria Razalan** who has been by my side throughout this whole project and beyond.

Lastly, the contribution of love and support from **friends** and **family** are not to be understated in all life's endeavours and this was certainly no exception.

Author's Declaration

I declare that I am the sole author of the work in this thesis and that it is original except where indicated by special reference in the text. No part of this degree has been submitted for any other degree to any other institution.

1. Introduction

1.1 Cytochrome P450

Cytochrome P450 enzymes (P450s) can generally be described as heme-thiolate proteins widely distributed in all forms of life from prokaryotes (archaea, bacteria) and lower eukaryotes (fungi and insects) to higher eukaryotes (plants and animals including humans).[1] They are involved in numerous processes, which include the metabolism of aliphatic, alicyclic and aromatic molecules in reactions resulting in hydroxylation, epoxidation, dealkylation, sulfoxidation, deamination, desulfuration, dehalogenation, and *N*-oxide reduction.[2] The fact that these systems catalyse such a vast and interesting set of reactions combined with the availability of new genetic engineering techniques makes them especially promising candidates for preparative biocatalysis.

1.1.1 Discovery and Nomenclature

The foundation that preceded the discovery of cytochrome P450 (P450) was set by an important finding in cell biology in the early 1940s. In 1943 Albert Claude and his collaborators discovered a novel subcellular fraction, called microsomes, after establishing a method for cell fractionation of animal tissue.[3] As a result, many scientists, especially in the field of biochemistry, showed particular interest in these newly found subcellular fractions and started to investigate morphological and biochemical properties of these microsomes. Thereby, and with the advancement in available analysis tools like mass spectrometry and spectrophotometry, the first microsomal NAD(P)H-linked electron-transfer components: NADPH-cytochrome *c* reductase[4], cytochrome b_5[5] and NADH-cytochrome b_5 reductase[6] were discovered. Although their physiological functions were speculated upon, no evidence for their natural role was elucidated in the 1950s. However, the knowledge and techniques acquired in cytochrome b_5 studies[7] paved the way for the identification of P450 proteins.

Other observations in the mid-1950s also contributed to the classification of cytochrome P450. Axelrod demonstrated in his studies of drug metabolism in liver microsomes that NADPH and oxygen were required[8-9] and Ryan and Engels suggested in 1957 that steroid hydroxylases of adrenal cortex microsomes hold a heavy metal, possibly iron, due to the inhibition by carbon monoxide which was reversible by white light irradiation.[10] Yet, one of the most important observations was made in 1958 by Klingenberg, while studying the

spectral properties of rat liver microsomes, and would later be accountable for the naming of P450s. The microsomes showed, in the presence of a reducing agent and carbon monoxide, a prominent optical absorption peak at 450 nm which was unusual among the known coloured proteins at that time.[11] However, the nature of these enzymes remained unknown until the discovery of P450s by Omura in 1962. Omura repeated Klingenbergs experiments and noticed a new spectral peak at 420 nm upon treatment of the enzyme with detergents while the 450 nm peak disappeared. This new spectrum had a lot in common with that of hemoglobin, but contamination of the original microsomes with hemoglobin was precluded. To distinguish the original microsome-bound pigment from its detergent-solubilized form, the former was named P450.[12]

After these first simple spectral studies, the identification and characterization of P450s leapt forward and started to become a major topic in drug metabolism. Selected milestones of the key scientific discoveries in cytochrome P450 research during the past 50 years are highlighted in Figure 1.2.

Today, cytochrome P450 hemoprotein enzymes constitute one of the largest protein superfamilies found in nature with more than 21000 officially by the P450 Nomenclature Committee accepted P450s (http://drnelson.uthsc.edu/cytochromeP450.html) (Figure 1.1).

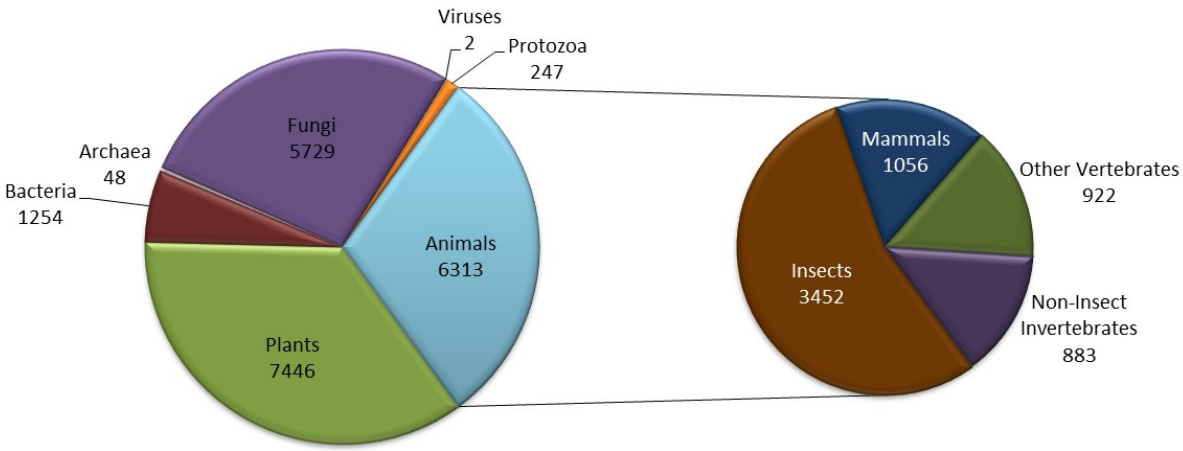


Figure 1.1: Schematic distribution of cytochromes P450 among the different kingdoms (August 2013)

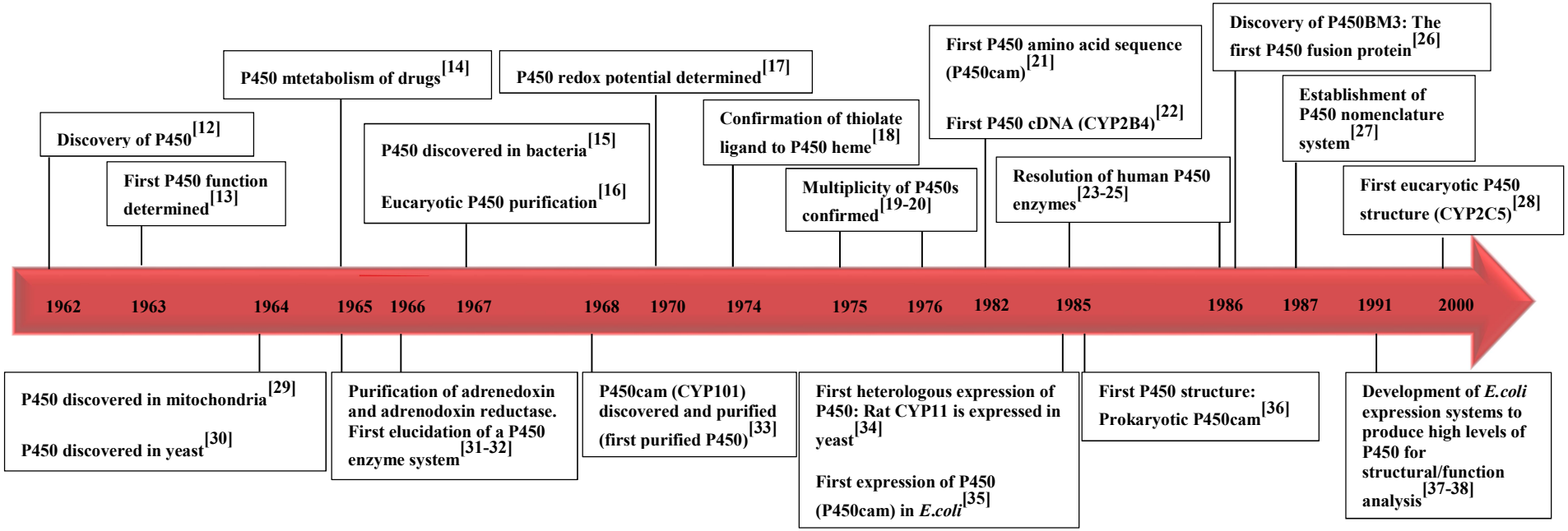


Figure 1.2: Selected milestones in cytochrome P450 research (1962-2000)

Ever since the discovery of cytochrome P450 in 1962[12], more than 37000 papers (NCBI, http://www.ncbi.nlm.nih.gov/) related to P450 research have been published and the number of academic research paper increases steadily with about 2000 per year (Figure 1.3) still being added (http://webtools.mf.uni-lj.si/public/medsum.html).

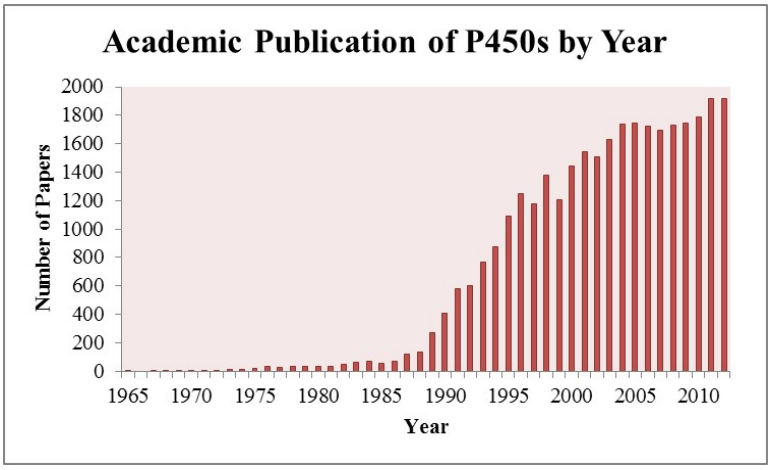


Figure 1.3: Scientific articles published on National Center for Biotechnology Information.

With the emergence of multiple forms and isoforms of P450s[19-20, 39] in this rapidly growing research field, a classification system for the naming of P450 soon became a requirement. But in order to classify P450s based on the similarity of their primary sequences the accessibility of complete amino acid sequences was necessary. *Pseudomonas putida* $P450_{cam}$ was the first primary sequence determined[21] followed by a phenobarbital-induced P450 of rat liver derived from cloned cDNA[22] and many more were to come due to the rapid expansion of sequence data in the 1980s. It so happened that in 1987 Nebert and collaborators proposed a commonly accepted nomenclature system for cytochrome P450 (CYP)[27] with follow-ups in 1989[40] and 1991[41]. P450s that share >40% amino acid identity are grouped into the same family which will then be divided into subfamilies when members have >55% identity. Designated Arabic numbers indicate the gene family (CYP1, CYP2, *etc.*) and a subsequent capital letter the subfamily (CYP1A, CYP1B, CYP2A, CYP2B, *etc.*). Individual genes, on the other hand, are specified by another number (CYP1A1, CYP1A2, CYP2A1, CYP2A2, *etc.*). Moreover, additional nomenclature describing allelic variants are applied when P450 sequences share more than 97% amino acid identity (v1, v2, *etc.*). Eventually, David Nelson established, out of a need for unlimited space to present nomenclature and annotation information, a publicly accessible homepage (http://drnelson.utmem.edu/CytochromeP450.html) in 1995.[42] Despite facilitating the naming of enzymes, it should be noted that the naming system does

not reflect on the function of the enzymes. Therefore, members within the same family may have completely different functions and enzymes with similar roles may belong to different families.

1.1.2 Structure and Classes

The first crystal structure of a P450, namely CYP101A1 ($P450_{cam}$) from *Pseudomonas putida*, was determined by Poulus and coworkers 23 years after the first discovery of P450s.[36] Since then, the number of crystal structures identified for P450s increased rapidly with more than 600 entries in the Protein Data Bank for P450s to date (http://www.rcsb.org). With the availability of these structures, it has become apparent that the general structural fold across the P450 superfamily, including eukaryotic and mitochondrial P450s, is strongly conserved (as shown in Figure 1.4) despite often low sequence identity between the different families (10-30%)[43], and with only three absolutely conserved amino acids across this gene superfamily.[44-45] Furthermore, no other proteins apart from cytochromes P450 showed this structural arrangement. Not even prokaryotic and eukaryotic enzymes like nitric oxide synthases, which execute similar reaction chemistry to P450s and have cysteinate-coordinated heme iron, displayed any similarity in their structural topology.[46-47]

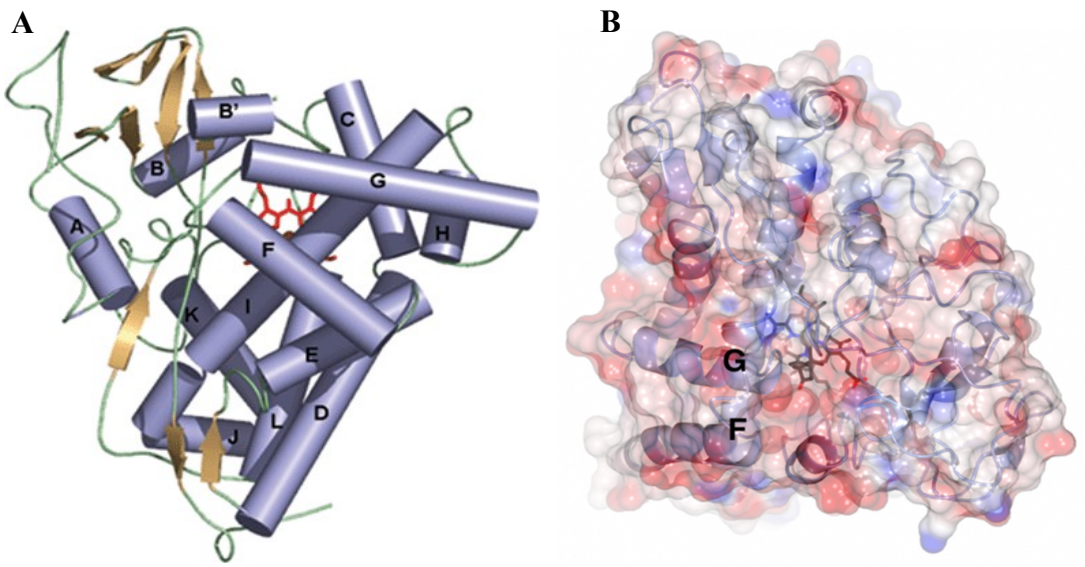


Figure 1.4: Structure of cytochrome P450 using the *Pseudomonas putida* camphor hydroxylase $P450_{cam}$ (CYP101A1) as an example.
A: overall structure[48]; selected helices labelled according to standard nomenclature[49]; α helices = blue cylinders; β sheet components = brown arrows; Interconnecting loop regions = cyan string; heme cofactor = red; **B**: space-filling model showing the access channel with camphor above the heme (PDB code: 2CPP)

The general shape of P450s resembles a trigonal prism with the heme cofactor concealed at the center. The overall structure of P450 enzymes contain four β-sheets (β5 is variable) and approximately 13 α-helices that can be distributed into the α (helix-rich) domain and the β (sheet-rich) domain (the top left of the $P450_{cam}$ structure in Figure 1.4 **A**) as shown in Figure 1.5.[48, 44]

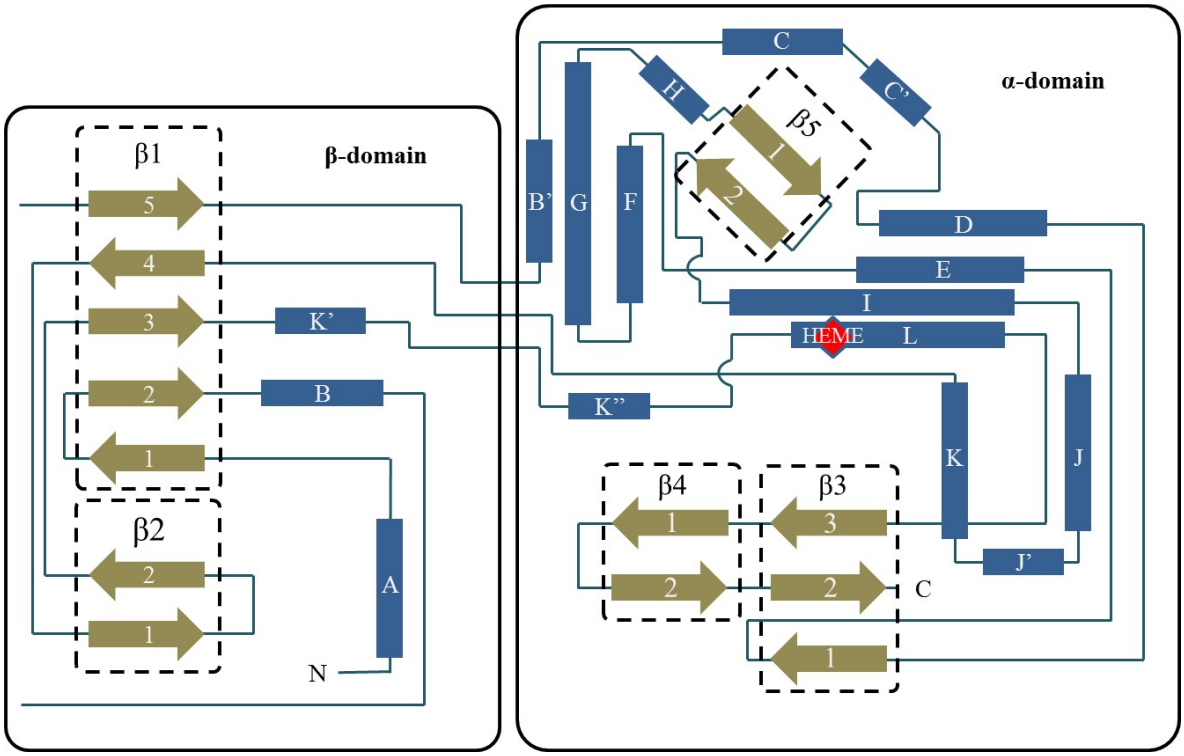


Figure 1.5: Topographic map showing the secondary elements commonly found among P450 enzymes[44]
α helices = blue cylinders; β sheet components = brown arrows; Interconnecting loop regions = cyan string; heme cofactor = red

β1 and β2 of the smaller β domain are involved in the formation of the hydrophobic substrate access channel.[44] The highest structural conservation is found in the structural core which is composed of helices D, E, I, and L and helices J and K, two sets of β sheets, and a coil called the "meander".[50] The "meander" is a structurally conserved region that is spanned by 10-15 amino acid residues between K and the Cys-pocket [loop region (the β-bulge) preceding the L helix which contains the thiolate heme ligand] and is supposed to take part in heme binding and stabilization of the tertiary structure. Major structural elements include the long I helix, which is located diagonally across the heme cofactor and the L helix that runs behind the cofactor.[48] Those positions of structural elements close to the heme are generally well conserved and have regions of highly conserved amino acid motifs that center on the heme thiolate ligand and oxygen activation

chemistry. Figure 1.6 shows the most common amino acids motifs among the P450 superfamily using the primary structure of a class II enzyme as example (additional transmembrane domain).

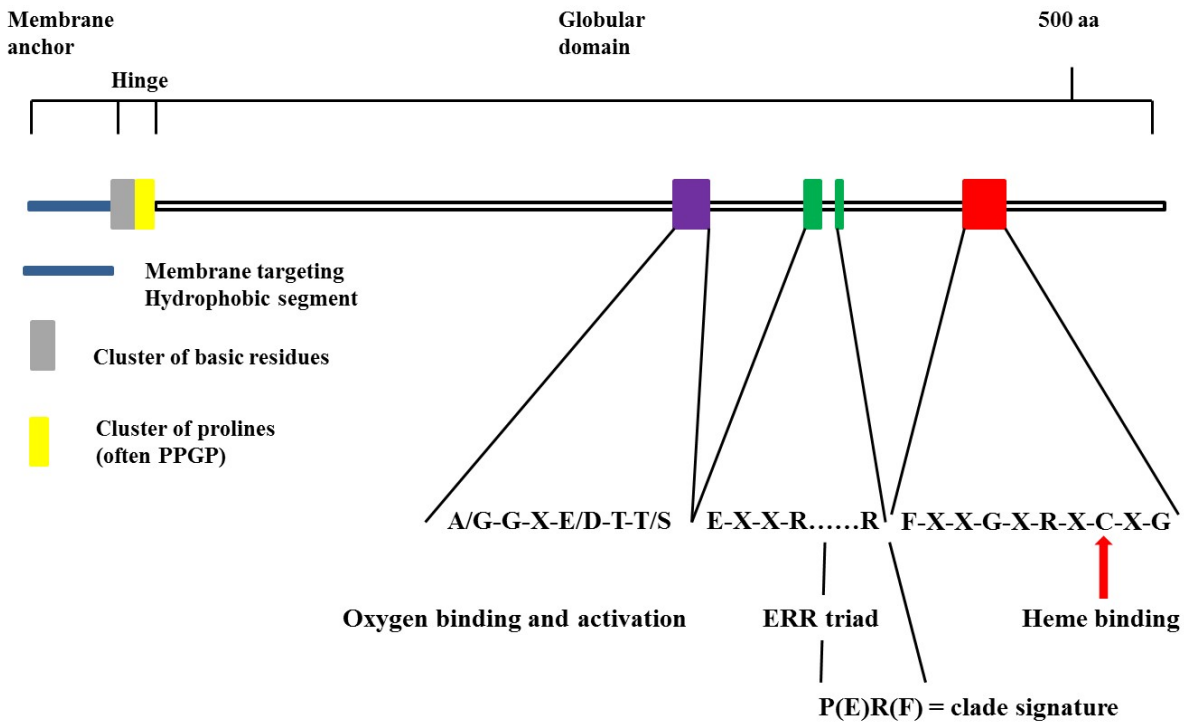


Figure 1.6: Map of the signature motifs in the P450 proteins of an ER-bound P450 protein (class II enzyme)

The cysteine ligand to the heme iron is absolutely conserved and critical to P450 oxygenase function. It is located just prior the L helix in the heme-binding loop, which contains the most characteristic P450 consensus sequence (Phe-X-X-Gly-X-Arg-X-Cys-X-Gly). The structural organization of this segment protects the Cys ligand and enables it to accept hydrogen bonds from peptide NH-groups. The only other motif considered absolutely conserved in the P450 superfamily is an EXXR motif in the K helix.[50] These residues appear to be important for hydrogen bonding and maintaining the "meander" region. The EXXR motif forms the so-called ERR triad using the R of the meander region (consensus sequence: PERF) as second "R" in the motif. The ERR triad is believed to play an important role in heme binding due to single point mutation experiments of the K-helix glutamate[51] and arginine[52-54] or the meander arginine[55] in four separate enzymes that resulted in inactive protein. Helix I contains another consensus sequence considered as P450 signature (Ala/Gly-Gly-X-Asp/Glu-Thr-Thr/Ser) and includes the highly conserved Thr with important roles in coupling of electron transfer to substrate oxygenation and in substrate specificity.[48, 50]

Although the topology of almost all crystallised P450s are highly similar (Figure 1.4) they differ in substrate specificity and their electron transfer partners. These differences are facilitated by the diverse regions, which vary in both sequence and structure.[56] For instance, Gotoh described in 1992 six regions based on the comparison between the CYP2 family and $P450_{cam}$, the so called SRS (substrate recognition sites), that are participating in substrate recognition and binding. These variable regions include helices A, B, B', F, and G and their flanking loops.[57] The loop region B–B' and B'–C line the active site and harbour SRS1 while helices F and G and corresponding loop, that form part of the access channel and ceiling of the active site, hold SRS2 in the C-terminal end of F and SRS3. The N-terminus of β-strand 1–4 houses SRS-5 and the β-turn at the end of β sheet 4 (β 4-1) SRS6. SRS4 lies in the center portion of the I helix but is in comparison to the other SRSs conserved throughout the P450 family.[44]

SRSs are generally associated to be involved in substrate orientation in the pocket since most of them are actually in the active-site/heme pocket of the P450 enzymes. Substrate recognition on the other hand is achieved through helix A, the F–G loop and β-strands 1–1 and 1–2 given that mutation in this region influence the substrate binding.[58] Furthermore, these regions, except for helix A, are believed to be associated with the membrane in the membrane-associated P450s because of their increased hydrophobicity.[59]

In addition, P450 enzymes require electrons for catalysis which are usually supplied by cofactors NADPH or NADH. In order to shuttle required electrons from cofactor to the heme domain, most P450s are associate with a NAD(P)H-linked reductase or reducing system in the cell. Interactions between P450 and its redox partner have to be well regulated to overcome the issue of heterogeneity among the P450 population in general and also their outnumbering of reductase molecules.[44, 60] It is believed that a complementary charge interaction is involved in redox-partner docking which allow P450s to interact with their electron donor at the proximal face of the molecule.[43] While the P450 itself has positively charged residues at the surface as indicated by mutation studies[54, 61-64], the reductase has conserved negative charges[65-67] that positively affect the alignment between the redox-partner and the P450. However, alignment studies of structurally known P450s shown different, but nonetheless unique charge distributions on the proximal face indicating that the binding between the P450 and the reductase will differ between P450 families.[56] Regarding P450 – redox partner arrangement, there are 10 different classes of P450 systems described to date (summarized in Table 1.1).

Table 1.1: Classes of P450 redox partners [50, adapted]
red: P450; light brown: ferredoxin (Fdx), flavodoxin (FMN); dark brown: ferredoxin reductase (FdR), [2Fe–2S] ferredoxin domain (Fe/S), flavin adenine dinucleotide (FAD); yellow: P450 reductase, CPR; blue: 2-oxoacid:ferredoxin oxidoreductase (OFOR) domain;

Class/ Source	Domain Organization	Localization/ Remarks	Example	Reference
Class I				
Bacterial	P450 Fdx FdR; NAD(P)H + H^+ → $NAD(P)^+$	Cytosolic Soluble	CYP101 CYP105D1	[33], [68] [69]
Mitochondrial		**P450**: inner mitochondrial membrane **FdR**: membrane associated **Fdx**: mitochondrial matrix, soluble	CYP11A1	[32]
Class II				
Bacterial	P450 CPR; NAD(P)H + H^+ → $NAD(P)^+$	Cytosolic Soluble	CYP105A3	[70]
Microsomal A		ER Membrane anchored	CYP3A4	[71]
Class III				
Bacterial	P450 FMN FAD; NAD(P)H + H^+ → $NAD(P)^+$	Cytosolic Soluble	CYP176A	[72], [73]
Class IV				
Bacterial	CoA → Acetyl-CoA; P450 Fdx OFOR; Pyruvate → CO_2	Cytosolic Soluble	CYP119	[74], [75]
Class V				
Bacterial	P450 Fdx FdR; NADPH + H^+ → $NADP^+$	Cytosolic Soluble	CYP51FX	[76]
Class VI				
Bacterial	FdR FMN P450; NAD(P)H + H^+ → $NAD(P)^+$	Cytosolic Soluble	CYP177A1	[77], [78]
Class VII				
Bacterial	P450 FMN Fe/S; NAD(P)H + H^+ → $NAD(P)^+$	Cytosolic Soluble	CYP116B2 CYP116B3	[79] [80]

Class/ Source	Domain Organization	Localization/ Remarks	Example	Reference
Class VIII				
Bacteria Fungi	P450 FMN FAD NADPH + H^+ $NADP^+$	**Prokaryotic**: Cytosolic, Soluble **Eukaryotic**: membrane anchored	CYP102A1 CYP505A1	[81] [82]
Class IX				
Fungi	P450 NADH + H^+ NAD^+	Cytosolic Soluble Only NADH dependent	CYP55	[83]
Class X				
Plants Mammals	P450 PGH_2 TXA_2	ER Membrane bound	CYP74A	[84], [85], [86]

Class I P450s (**I**) are three-component systems consisting of a flavin adenine dinucleotide (FAD)-containing reductase (FDR) that transfers electrons supplied by a cofactor to an iron-sulfur protein (ferredoxin, Fdx) which then reduces the P450. This system occurs solubly in prokaryotes and usually utilizes NADH as electron source but can also be found in mitochondria of eukaryotes in which the P450s and reductase are membrane bound and prefer NADPH.[87-88] It was first elucidated in *Pseudomonas putida* (CYP101) in 1968 by Katagiri and co-workers who were able to separately purify all elements of the electron transport chain and restore its activity.[33]

Eukaryotic organisms usually possess the class II P450 system (**II**), a microsomal system typically located in the endoplasmatic reticulum (ER), consisting of 2 integral membrane proteins. One is a single NADPH specific P450 reductase (CPR) that contains both FAD and FMN domains for mediating electron transfer from NADPH[89-90] and the other is the P450 itself. Microsomal P450s in mammalian liver also utilize cytochrome b_5 as a second source of electrons, where the reductant is NADH and electron transfer is mediated by cytochrome b_5 reductase which is an FAD-containing oxidoreductase.[91] One prokaryotic soluble class II monooxygenase system has been described in *Streptomyces carbophilus*. It consists of a NADH-dependent P450 reductase containing both FAD and FMN and the CYP105A3 ($P450_{sca}$).[70]

In 2002, Hawkes and collaborators observed a unique arrangement of P450 and putative redox partner (**III**) that deviate from the originally proposed class I and class II systems.

CYP176A1 from the bacterium *Citrobacter braakii* contains in contrast to bacterial class I systems not an iron-sulfur domain as second mediator protein but a flavodoxin, so that electrons are transferred *via* the redox centres FAD and FMN.[72]

Class IV systems (**IV**) which receive electrons from a non-NAD(P)H-dependent reductase were discovered in extremophile bacteria like *Sulfolobus solfataricus*[92], and *Sulfolobus tokodaii*[93]. The P450 CYP119 obtains its reducing equivalents from 2-oxoacid:ferredoxin oxidoreductase (OFOR)[94-95] that uses pyruvic acid instead of NAD(P)H as the source of electrons.[96]

A fifth class of P450-redox partner association (**V**) was found in the bacterium *Methylococcus capsulatus*, one of the few bacteria that can synthesise sterols *de novo*.[76] The CYP51 P450 heme domain is fused at the C-terminus to a ferredoxin domain utilizing an oxogenous ferredoxin reductase as the third electron-transfer protein component.[97]

Another novel P450 fusion arrangement was identified in *Rhodococcus rhodochrous* (strain 11Y) (**VI**) consisting of a fusion between a soluble C-terminal P450 domain (XplA) and an FMN-containing N-terminal flavodoxin domain.[77] XplA is encoded in a gene cluster that includes the putative redox-partner upstream of *xplA,* the FAD-containing reductase protein (XplB). XplA catalyses the degradation of nitramine explosive and pollutant hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) by reductive denitration and allows *R. rhodochrous* 11Y to use it as its only nitrogen source.[98]

Another natural P450 fusion enzyme (**VII**) was found in the bacterium *Rhodococcus* sp. NCIMB 9784.[99] The protein was named CYP11B2 ($P450_{Rhf}$) and comprises a soluble P450 heme domain fused to a FMN-and a 2Fe2S-containing reductase at the C-terminus. Homologues of this fusion system, which endogenous function is not known, were also discovered in *Ralstonia metallidurans* (CYP116B1) and *Rhodococcus ruber* (CYP116B3).[80]

In 1986, a naturally occurring P450 fusion enzyme (CYP102A1; $P450_{BM3}$) was discovered in *Bacillus megaterium* that for the first time deviated from the standard class I and class II systems.[26] Enzymes belonging in the class VIII system (**VIII**) are catalytically self-sufficient monooxygenases in which the N-terminal heme domain is linked to an eukaryotic-like diflavin reductase domain (CPR) at the C-terminus. $P450_{BM3}$ has amongst all P450 monooxygenases the highest reported turnover rate which is probably due the rapid intratransfer of electrons between the fused protein domains.[100] Homologues of

CYP102A1 could be found in *Bacillus subtilis* (CYP102A2 and CYP102A3) as well as in *Ralstonia metallidurans*, *Bradyrhizobium japonicum*, and various bacilli and streptomycetes.[101] Apart from class VIII systems in prokaryotes, eukaryotic counterparts exist like the ER membrane bound CYP505A1 ($P450_{foxy}$) from *Fusarium oxysporum*.[102]

Unlike most P450s that need a redox partner for catalysis, P450s of class IX (**IX**) function without participation of any NAD(P)H-linked reductase or reducing system. CYP55A1 ($P450_{nor}$) of the fungus *Fusarium oxysporum* is the first example ever been described for this class.[103] In case of $P450_{nor}$ the heme prostetic group receives electrons directly from NADH to catalyse the reduction of nitric oxide to nitrous oxide. Homologues could be found in *Trichosporon cutaneum*[104], *Cylindrocarpon tonkinense*, *Histoplasma capsulatum* and *Aspergillus oryzae*[105].

The final class of P450s (**X**) can catalyse the rearrangement of the oxygen atoms in the substrate molecule itself and therefore don't require NAD(P)H or any redox partners. CYP8A1 for example, can synthesize prostaglandin H2 from prostacycline[106] whereas the microsomal P450 CYP5A1 utilizes prostaglandin H2 to build tromboxane A2[107] without the supply of reducing equivalents. Other examples can be found in plants in the CYP74 family which encode allene oxide synthases[108].

1.1.3 Catalytic Reaction Mechanism

The classical reaction catalysed by cytochromes P450 is a monooxygenase reaction (alternative reaction see chapter 1.1.4) in which one oxygen atom, originating from molecular O_2, is incorporated into the substrate while the second oxygen atom is reduced to water. This reaction utilizes two electrons usually derived from reduced pyridine nucleotides (NADH or NADPH)[109].

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

Oxygen is a highly abundant element in aerobic organisms and although it is generally regarded as being a chemically reactive element, the dioxygen molecule is relatively inert because of the high stability of the dioxygen bond (497 kJ mol/l) and its unusual electronic structure[110-111]. In order to utilize its inherent oxidative properties for metabolism, it is crucial to activate oxygen. This activation can be achieved by metal-dependent oxygenases, like cytochromes P450 which contain a prosthetic group constituted of an iron

(III) protoporphyrin-IX with thiolate as fifth (axial) ligand, leaving the sixth coordination site open to bind and activate molecular oxygen (Figure 1.7)[112].

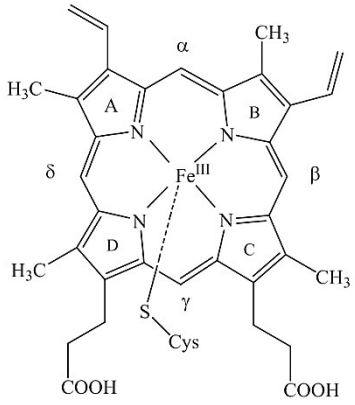


Figure 1.7: Prosthetic of cysteinato-heme enzymes
An iron-(III) protoporphyrin-IX linked with proximal cysteine ligand of the cytochrome P450

In order to understand the reason for, and mechanism of, oxygen activation by cytochromes P450, it is important to consider the different spin- and redox states of the dioxygen molecule as well as the iron atom present in the heme group. In dioxygen, each oxygen atom contributes eight electrons to the molecule. Half of these electrons occupy sigma bonding and antibonding orbitals, formed by overlap of the 1s and the 2s orbitals. Of the remaining eight electrons, six occupy, in accordance to the aufbau principle, the 2σ and the 2π bonding orbitals, leaving two electrons which must be placed in the two $2\pi^*$ antibonding orbitals. These electrons will each - according to Hund's Rule - occupy one of the degenerate (same energy) orbitals, and they are therefore unpaired and have parallel spins.[113] Figure 1.8 A illustrates this ground-state of dioxygen which is commonly referred to as the triplet ground state. The triplet electronic state can be traced back to the formula for the spin quantum number [(2S + 1); S = ½ for each unpaired electron] which yields a value of 3 for the overall spin state as the total S = 1 (½ + ½). Thus, the triplet state is regarded as 1 plus the number of unpaired electrons. This feature, makes oxygen paramagnetic and according to Pauli's exclusion principle very unlikely to participate in reactions with organic molecules which are usually in a singlet state because they generally don't contain unpaired electrons (S = 0; 2S + 1 = 1). In order to overcome the spin restriction, the triplet oxygen has to form the singlet state by absorption of sufficient energy to reverse the spin on one of the unpaired electrons (Figure 1.8 B and C). Another mechanism for oxygen activation involves monovalent reduction. [110, 113-114] However, the iron in P450s is able to interact with O_2 when both are in their high-spin states.

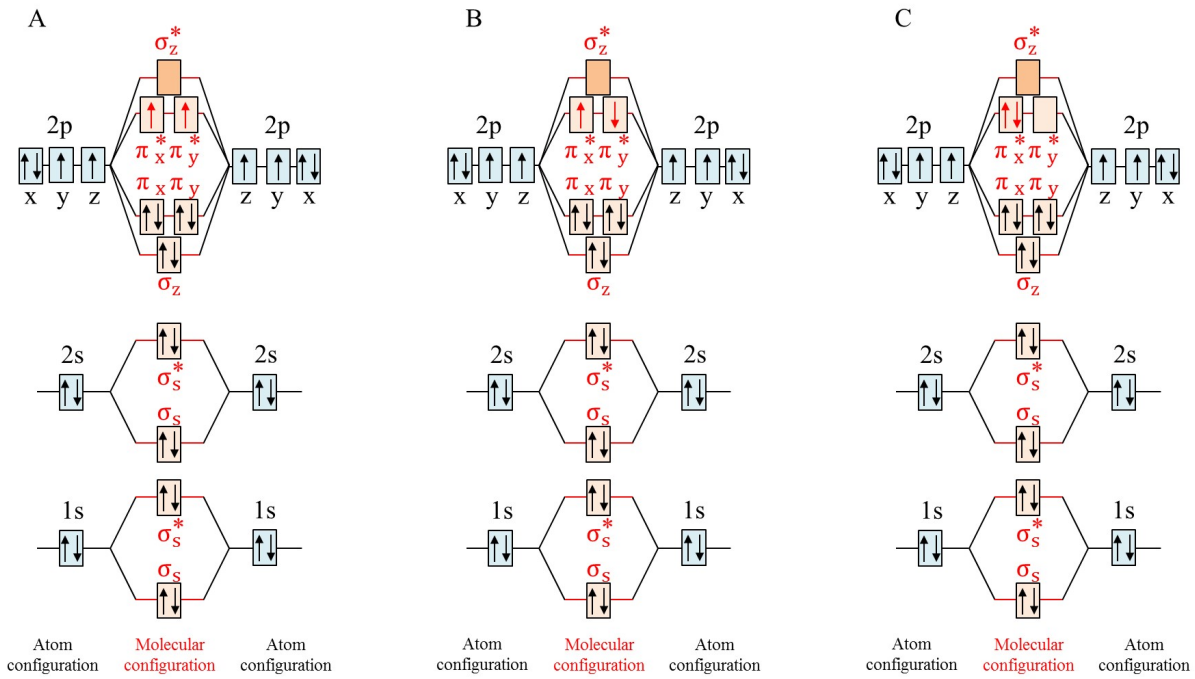


Figure 1.8: The three electronic configurations of molecular oxygen, O_2.
A: the triplet ground state $1\Sigma_g^-$; B: the singlet oxygen $^1\Sigma_g^+$ excited state; C: the singlet oxygen $^1\Delta_g$ excited state.

When the P450 enzyme is in its resting state (prior substrate binding), the ferric iron [Fe^{3+} or Fe(III); valence shell electrons: $3d^5$], that lies within the plane of the heme group, is predominantly low spin (two electron pairs and only one unpaired electron in the d orbital). Substrate binding at the enzymes' active site, however, leads to a change in the ferric iron spin-state from low spin to high spin due to the displacement of the distal water molecule (Figure1.9 **A-B**) which consequently causes the iron to move out of plane of the porphyrin ring as the central cavity is too small to accommodate the increased cationic size. The transition to a high spin state (degenerate orbitals occupied with five unpaired electrons; spin-free) and the changes in heme geometry, trigger the reduction of the P450 enzyme by its redox partner as substrate binding lowers the Fe(III) redox potential of the P450 such that the reduction from ferric to ferrous state [Fe^{2+} or Fe(II)] is favoured (Figure 1.9 **C**). It is believed that Fe(II) in its high spin form (four unpaired electrons in the d orbital) ensures a spin-favoured interaction with the triplet ground-state dioxygen as both are in their high spin states which leads to the activation of O_2 to superoxide (Figure 1.9 **D**, ferric superoxide complex) by substrate induced spin and redox changes.[110-111, 115-116]

Once the ferric superoxide complex is formed, the P450 undergoes a second reduction in order to complete its cycle. This second electron transfer, that is usually the rate limiting step of the catalytic cycle, generates the ferric peroxy anion which is protonated to form

the ferric hydroperoxo species known as compound 0 (Figure 1.9 **E**). The ferric hydroperoxo intermediate, however, readily undergoes a second protonation as it is unstable and that gives rise to a porphyrin radical cation Fe(IV) species (alternative species: protein radical cation Fe(IV) **F'** or Fe(V) species **F''**) also known as compound I (Figure 1.9 **F**). This ferryl intermediate is able to attack the substrate which produces the hydroxylated metabolite (Figure 1.9 **G**). The enzyme will then return to its resting ferric state due to product release and reequilibration with water to complete the cycle.[117]

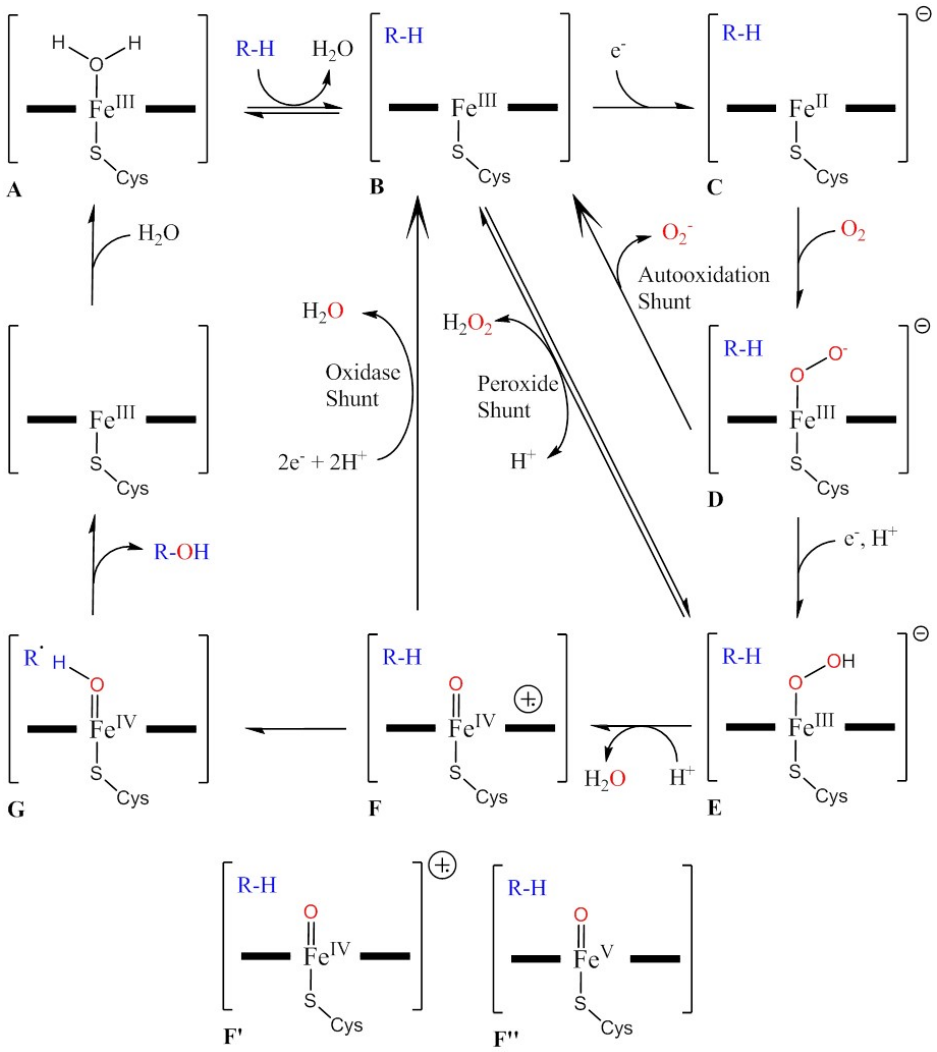


Figure 1.9: The cytochrome P450 catalytic cycle.
heme group: solid bars; iron: Fe; cysteine thiolate: Cys-S; substrate hydrocarbon: RH; hydroxylated product: ROH; The +. over one of the heme bars indicates the radical cation is located on the porphyrins, whereas its placement besides the brackets indicates that the radical is located somewhere on the protein. See main text for detailed description.

During the P450 reaction cycle, forms of active oxygen species ($O_2^{-\bullet}$, O_2^{2-}, $OH^{\bullet}$) are produced ,which are usually only formed *in situ*. However, in the event of a so called ″uncoupling″ there are three reaction mechanisms that lead to the break down of the oxo

intermediates. The first is the autoxidation of the ferric superoxide complex **D** which reforms the ferric P450 and produces superoxide radicals. In the second uncoupling scenario, known as the peroxide shunt, compound 0 **E** collapses and hydrogen peroxide is released. This reaction is reversible and can be exploited by some P450s to bypass the need of a redox partner by introducing the oxygen from hydrogen peroxide into the substrate. In the third reaction (oxidase shunt) the compound I **F** is protonated and reduced again by the dissociation of water instead of an insertion of the oxygen into the substrate.[48, 118]

1.1.4 Evolution and Catalysed Reactions

P450s are well known for their broad substrate range and the variety of reactions they catalyse. The functional diversity among this gene family can be caused through a variety of genetic mechanisms (*e.g.*, exon shuffling, alternative splicing and RNA editing).[119] The most commonly acknowledged mechanism driving the expansion and diversification, however, is gene duplication.[120]

It is proposed that the first cytochrome P450 originated from an ancestral gene that existed three and a half billion years ago[121], thus before the accumulation of an oxygen-rich atmosphere and the advent of eukaryotes. It is therefore believed that cytochrome P450s originally performed nitroreductase or endoperoxide isomerase reactions.[122] With the buildup of significant levels of molecular oxygen in the earth's atmosphere (approximately 2.8 billion years ago) the function of cytochrome P450s may have then shifted to protect early life forms from oxygen toxicity.[123] Additionally, repeated rounds of expansion by gene and genome duplication lead most likely to the modern form of P450s. One of the first expansions occurred probably one and a half billion years ago and evoked the development of cytochrome P450 families that may played a role in maintaining the membrane integrity of early eukaryotic cells and were therefore involved in the metabolism of endogenous fatty acids, cholesterol and its derivatives (*e.g.,* CYP11 and CYP4 families).[124] The most recent and possibly most dramatic expansion started about four hundred million years ago and was driven by two major events: (1) emergence of aquatic organisms onto land and (2) exposure of terrestrial organisms to hydrocarbon-based combustion products in the atmosphere).[125] It appears that the process of co-evolution between plants, animals, fungi and prokaryotes sustained the expansion and diversification of the cytochrome P450 superfamily and led to a multiplicity of reactions catalysed by P450s.

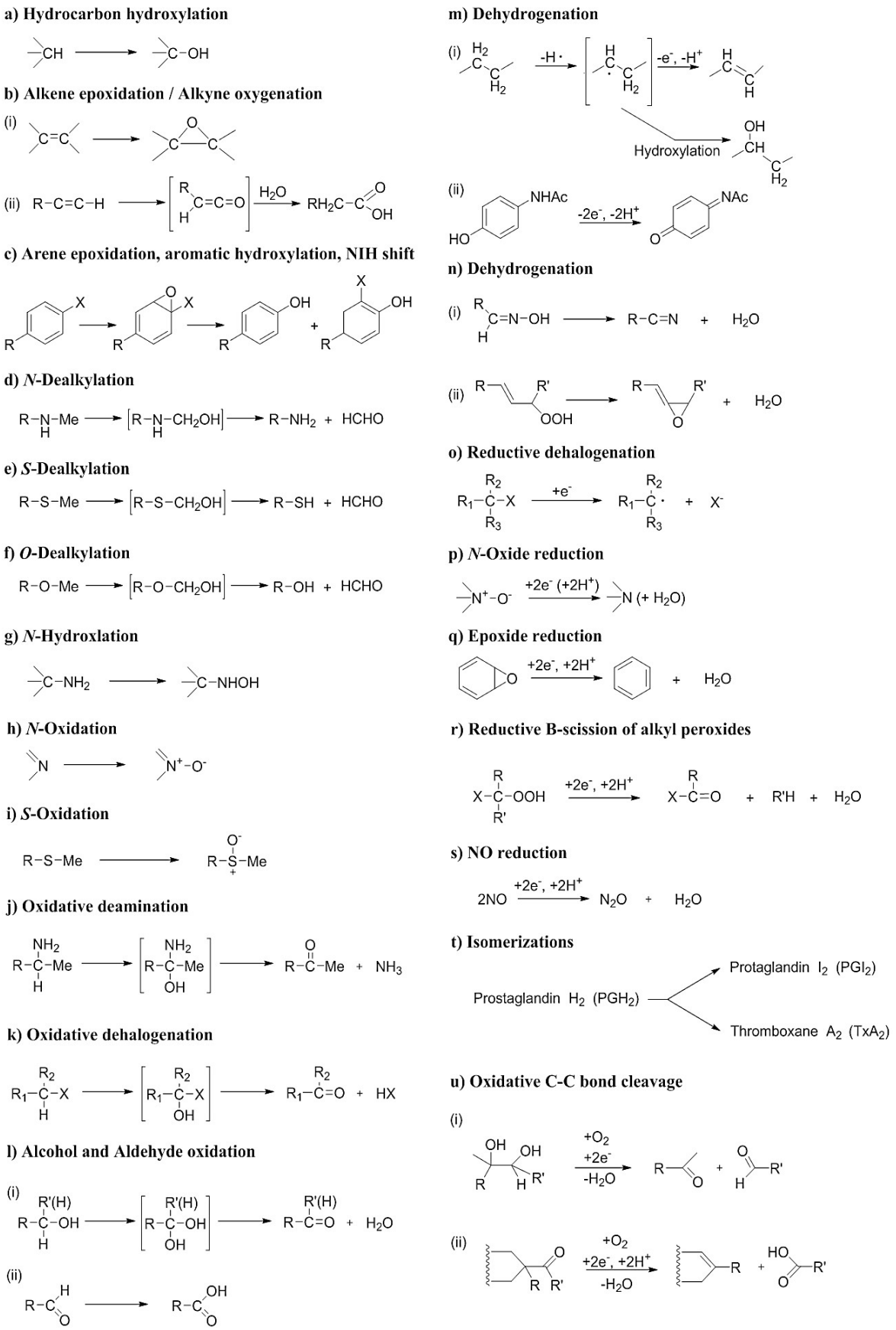


Figure 1.10: Schematic summary of the diverse P450-catalysed reactions[2]

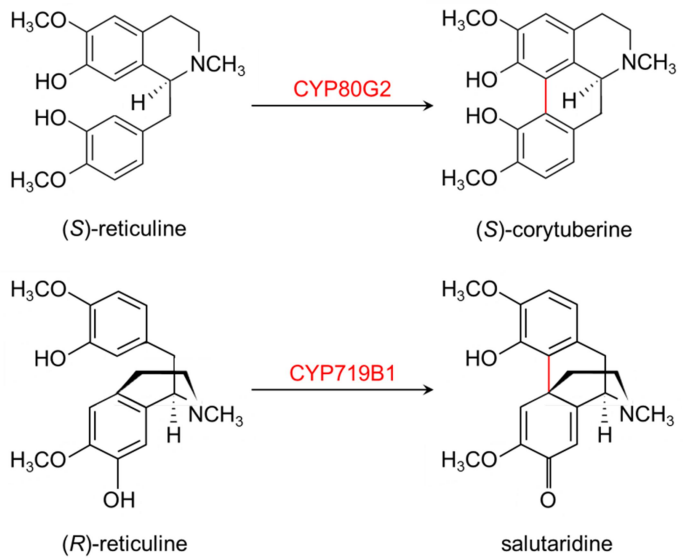


Figure 1.11: P450-catalysed C–C bond coupling in isoquinoline alkaloid biosynthesis of plants

Modern P450s play important roles in catalysing key steps of many different pathways and their reaction repertoire extends from hydroxylation and oxidation, which are the most common P450 reactions, to alkylation, dealkylation, epoxidation, demethylation, aryl migration and many more. A summary of the most common P450 reactions is given in the review by Sono *et al.* (Figure 1.10) describing more than 20 different reaction types.[2] In addition, P450-catalysed C–C bond couplings are common in the metabolism of many plants. These reactions are often critical in the biosynthesis of plant secondary metabolites (*e.g.,* alkaloidbiosynthesis, Figure 1.11).[126] Guengerich and Munro provided a recent description of unusual P450 reactions including nitration of tryptophan, cyclopropanation *via* carbene transfer, and intramolecular C-H amination.[126]

1.1.5 Applications

1.1.5.1 Potential and Limitations of P450s in biocatalysis

For a long time, enzyme catalysis was not acknowledged as a first choice option in organic synthesis because of observed or often assumed disadvantages (*e.g.*, narrow substrate range, limited stability of enzymes, low efficiency) and was only applied in the production of chemicals that were too difficult to synthesize in conventional ways. However, over the last 20-30 years enzymes have been recognized as viable tools for biotechnological application due to the tremendous progress in enzyme discovery, enzyme engineering, process development and, nonetheless, the rising pressure to design greener processes.[127] In order to meet green chemistry criteria, enzymes offer a major advantage because they

are usually able to operate at ambient temperature and pressure conditions which support lower energy processes. Furthermore, enzymes present the probability to reduce or even end the need for dangerous, toxic reagents as well as organic solvents in synthetic chemistry. However, several criteria have to be considered in order to apply biotransformations in an industrial context to maximise cost efficiencies. High conversion rates enabling for consumption of the maximal substrate amount and thus contributing to a decrease in substrate costs are of particular interest. Moreover, a high enantio- and regioselectivity is very desirable for the resulting product. Furthermore, substrate and product concentrations have to be considered for biotransformations as well as volumetric productivities.[128]

One of the main reason P450s attracted so much attention in the first place is their ability to oxyfunctionalize non-activated C-H bonds which still remains a major challenge in chemistry.[129] In addition, P450s are able to catalyse numerous different reaction types beside oxygenation (see 1.1.4)[2, 126] under ambient conditions and offer the advantage of a broad substrate range (*e.g.,* xenobiotics, antibiotics, steroids, terpenes, alkanes, fatty acids, alkaloids, *etc*.) often exhibiting high regio-, chemo-, and/ or stereoselectivity,[130] which makes them excellent starting materials to engineer new powerful biocatalysts. The most powerful source in P450 engineering is, however, the wide-ranging knowledge that has been acquired about P450s in last 50 years.

Although CYPs are generally regarded as biocatalyst with extraordinary potential, commercial applications of P450 catalysed reactions are rather scarce due to intrinsic drawbacks that continue to create major challenges when working with this class of enzymes (Table 1.2). One of the main issues is their natural occurring low turnover rate when compared to other enzyme classes. Also, the dependence on expensive cofactors like NAD(P)H as a source of electrons that have to be transferred to the heme domain using complex electron-transfer systems causes problems for the implementation of P450s to perform in an cost efficient industrial scale. Furthermore, the P450 origin has to be considered. While most bacterial P450s are easily expressed as soluble, stable enzymes in recombinant hosts and are therefore easier to handle than eukaryotic P450s, which are generally membrane bound and often unstable or inactive in the purified form, the substrate spectra and reactions catalysed by eukaryotic P450s are often more amenable to industrial applications.[130-133]

Table 1.2: Challenges and limitations for biotechnological application of CYPs

Limitation	Explanation/Cause	Possible solution(s)
General for CYPs		
Low CYP activity	Natural role of P450s	Protein engineering
Need for electron transfer partners	Channeling of electrons to active site	Use of heterologous redox partners
		Peroxide shunt
		Electrochemical and photo-chemical reduction of heme iron
Limitations of electron transfer rate	Complexity of catalysis	Fusions between CYP and electron transfer proteins
	Mismatch between redox partners	Redox chain optimization
Uncoupling	Poor fit of substrate to active site	Protein engineering
Overoxidation	Product is also substrate	*In situ* product removal
NAD(P)H requirement	Source of electrons	Enzymatic cofactor regeneration
		Coupling with plant photosynthesis
		Electrochemical heme reduction
Low substrate solubility	Substrates are often hydrophobic and/or poorly soluble in water	Two-liquid phase systems (dissolving substrate in an inert organic solvent)
		Addition of co-solvents or cyclodextrins
Specific for whole-cell biocatalysts		
Limited substrate uptake	Hydrophobic compounds disrupt cell membranes	Alternative hosts with altered uptake profiles
Limited product efflux	Lack of specific efflux pumps in host	Cell membrane permeabilization Modulation of efflux system
Limited oxygen transfer rate	Competition with endogenous respiration	Addition of oxygen
Substrate or product toxicity	General toxicity of polar compounds	Alternative hosts with enhanced resistance
Product degradation		Two-liquid phase systems

Efforts to overcome limitation concerning enzyme activity and substrate specificity include predominantly enzyme engineering *via* site directed mutagenesis[134-135] or direct evolution.[136-137] With regard to cofactor-dependency for biotransformation, whole cell systems are often employed to take advantage of cell internal cofactor recycling and thus bypass the constant addition of expensive material. Although this seem to be an elegant solution, significant bottlenecks related to whole cell systems, like limited substrate uptake, toxicity of the substrate or product, and product degradation need to be taken into account as well.[131] Nonetheless, it was possible to employ several well engineered P450s as

catalysts for various applications due to the extended research in the field of cytochromes P450.

1.1.5.2 Commercial applications

As already mentioned above, the implementation of cytochrome P450s in fine chemical synthesis has to meet certain requirements for industrial (minimum space–time yield of 0.1 g l^{-1} h^{-1}; minimum final product concentration of 1 g l^{-1}) as well as in pharmaceutical (minimum space-time yield of 0.001 g l^{-1} h^{-1}; minimum final product concentration of 0.1 g l^{-1}) production.[138] An analysis of 12 CYP-based processes by Julsing and coworkers (shown in Figure 1.12) commenced in 2008 demonstrates the difficulties related to P450 linked industrial application.[139] At that time, only one of the analysed cytochrome P450s fulfilled the minimal requirements defined for fine chemical synthesis (**A** Figure 1.12), while 4 CYP450s couldn't even meet the minimum requirements set for the pharmaceutical industry (**I**, **J**, **K** and **L** Figure 1.12). Furthermore, the operational time window of CYP450-based processes appears very high when compared to nonheme oxygenases. Nonetheless, it becomes apparent how subsequent optimization unlocks completely new perspectives for CYP application when looking at the example of the engineered *S. cerevisiae* strain (**H** Figure 1.12) that was part of Julsing's analysis.

In that particular case, a *S. cerevisiae* strain has been facilitated to produce artemisinic acid (precursor of the antimalarial drug artemisinin) from simple sugars.[146] Figure 1.13 shows the schematic representation of the engineered artemisinic acid biosynthetic pathway in the *S. cerevisiae* strain EPY224 expressing amorphadiene synthase (ADS), CYP71AV1 and its cognate CPR from the plant *Artemisia annua*. First consideration in the engineering process was to increase the production of farnesyl pyrophosphate (FPP) and decrease its use for sterols. To increase FPP production in *S. cerevisiae*, Ro and coworker upregulated the expression level of several genes responsible for FPP synthesis and downregulated one gene responsible for FPP conversion to sterols (not shown in Figure 1.13) using chromosomal integration to implement these modifications. The second step, after the generation of a high FPP producing yeast strain, was the introduction of the amorphadiene synthase gene (*ADS*) from *A. annua*, which has been characterized and used for *de novo* production of amorphadiene in *E. coli*.[147-148] The third and final obstacle to overcome was the production of artemisinic acid from amorphadiene.

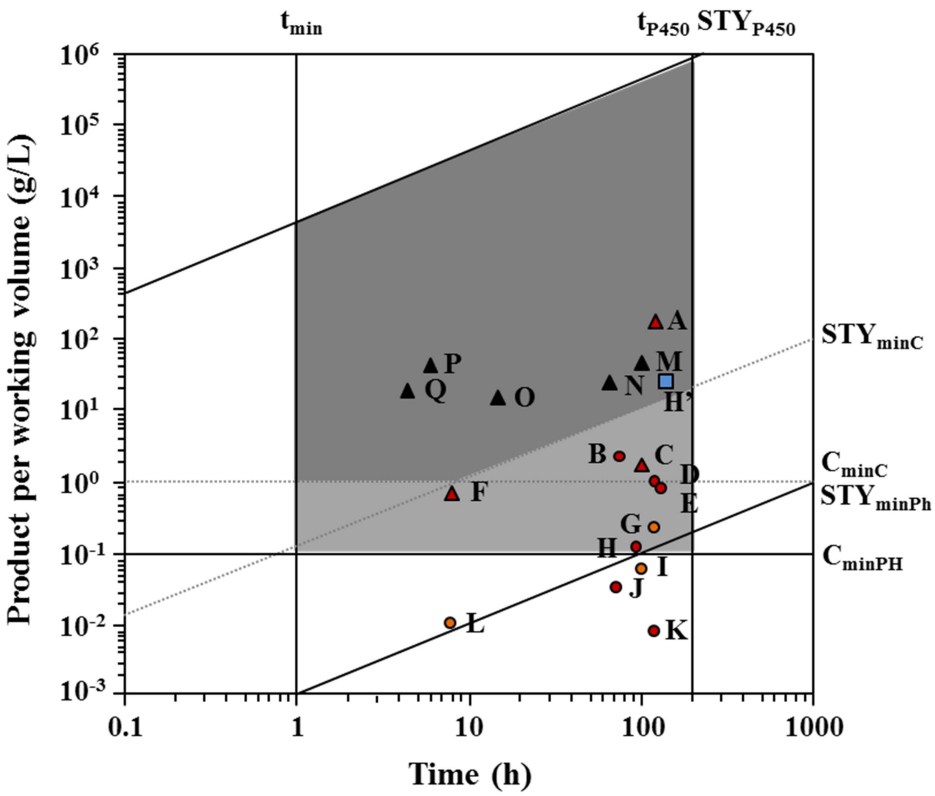


Figure 1.12: Industrial potential of CYP450-catalysed bioprocesses ([139], altered)
dark grey area: requirements for fine-chemical industry; light and dark grey area together: requirements for pharmaceutical industry; o: production of pharmaceutical products; △: production of fine chemicals; black filling: oxygenases containing nonheme-iron or flavin centers; red filling: bacterial P450s; orange filling: mammalian P450s; ■: engineered *Saccharomyces cerevisiae* strain exploiting a plant P450 for production of pharmaceutical products (**[140]**, see text below); t_{min}: minimum process running time; t_{P450}; longest reported CYP450 enzyme activity time; STY_{P450}: theoretical maximum space–time yield for CYP450-based processes; STY_{minC}: minimum required space–time yield for bioprocesses in the chemical industry; STY_{minPh}: minimum required space–time yield for bioprocesses in the pharmaceutical industry; C_{minC}: minimum required product concentration for bioprocesses in the fine-chemical industry; C_{minPH}: minimum required product concentration for bioprocesses in the pharmaceutical industry; **A**: *Candida tropicalis* strain engineered for the P450 catalysed production of long-chain dicarboxylic acids; **B**: (CYP153A14) from *Mycobacterium* sp. expressed in *Pseudomonas putida* converts limonene to the anticancer drug perillyl alcohol; **C**: CYP102A1 mutant used in a biphasic system to hydroxylate cyclohexane, octane and myristic acid; **D**: P450 catalysed bioconversion of compactine into pravastatin by *Streptomyces* sp.; **E**: optimization of fed-batch condition resulting in 25-fold increase of CYP102A1 expression level in *E. coli*; **F**: P450BM3 expression in *E. coli* whole cells lead to controlled regioselective oxidation of fatty acids; **G**: Human CYP2D6 overexpressed in *Schizosaccharomyces pombe* catalyses the oxidation 4'-hydroxymethyl-α-pyrrolidinobutyrophenone; **H**: engineered *Saccharomyces cerevisiae* strain produces antimalarial drug precursor artemisinic acid; **H'**: further engineered *Saccharomyces cerevisiae* strain produces antimalarial drug precursor artemisinic acid (**[140]**; see text below) **I**: engineered yeast catalyses self-sufficient biosynthesis of pregnenolone and progesterone; **J**: addition of cyclodextrin enhances P450 catalysed hydroxylation of vitamin D3 to 25-hydroxyvitamin D3 and 1α,25-dihydroxyvitamin D3 by *Amycolata autotrophica*; **K**: Transformation of vitamin D3 to 1α,25-dihydroxyvitamin D3 *via* 25-hydroxyvitamin D3 using *Amycolata* sp. strains; **L**: Bioconversion using immobilized recombinant flocculent yeast cells carrying a fused enzyme gene in an 'intelligent' bioreactor; **M-Q**: comparison of biosynthesis processes catalysed by oxygenases other than P450s, see references **[141] [142] [143] [144] [145]**

For that, the *CYP71AV1* gene, also from *A. annua*, was employed to oxyfunctionalize amorpha-4,11-diene and produce artemisinic acid in a three-step oxidation. As a result, this system yielded up to 0.1 g l^{-1} artemisinic acid, and thus barley fulfilled the minimal requirements for pharmaceutical compounds as defined by Julsing.[139] Advances within the next 4 years after Julsings publication to further improve this strain, made it, however, possible to yield a final concentration of 1.6 g l^{-1} of artemisinic acid.[149] Moreover, this artificial multi-enzyme cascade system originally developed by Ro and coworkers also paved the way for further improved yeast systems. In 2009 Teoh and coworkers reported a yeast system that employed, in addition to the synthase and the CYP71AV1, an alcohol and aldehyde dehydrogenases from *A. annua*.[150] All these findings and subsequent improvements finally led to an engineered yeast system that could produce up to 25 g l^{-1} in fermentation experiments as very recently reported by Paddon *et al.*.[140] The bioprocess catalysed by this strain is shown in Figure 1.12 (**H'**) in comparison to the original strain (**H**) and strinkingly demonstrates the potential of bioprocess engeneering. Based on this engineered yeast strain, Sanofi (a Paris-based pharmaceutical company) officially launched a new production facility in 2013 to produce artemisinin in large scale (http://www.rsc.org/chemistryworld/2013/04/sanofi-launches-malaria-drug-production).

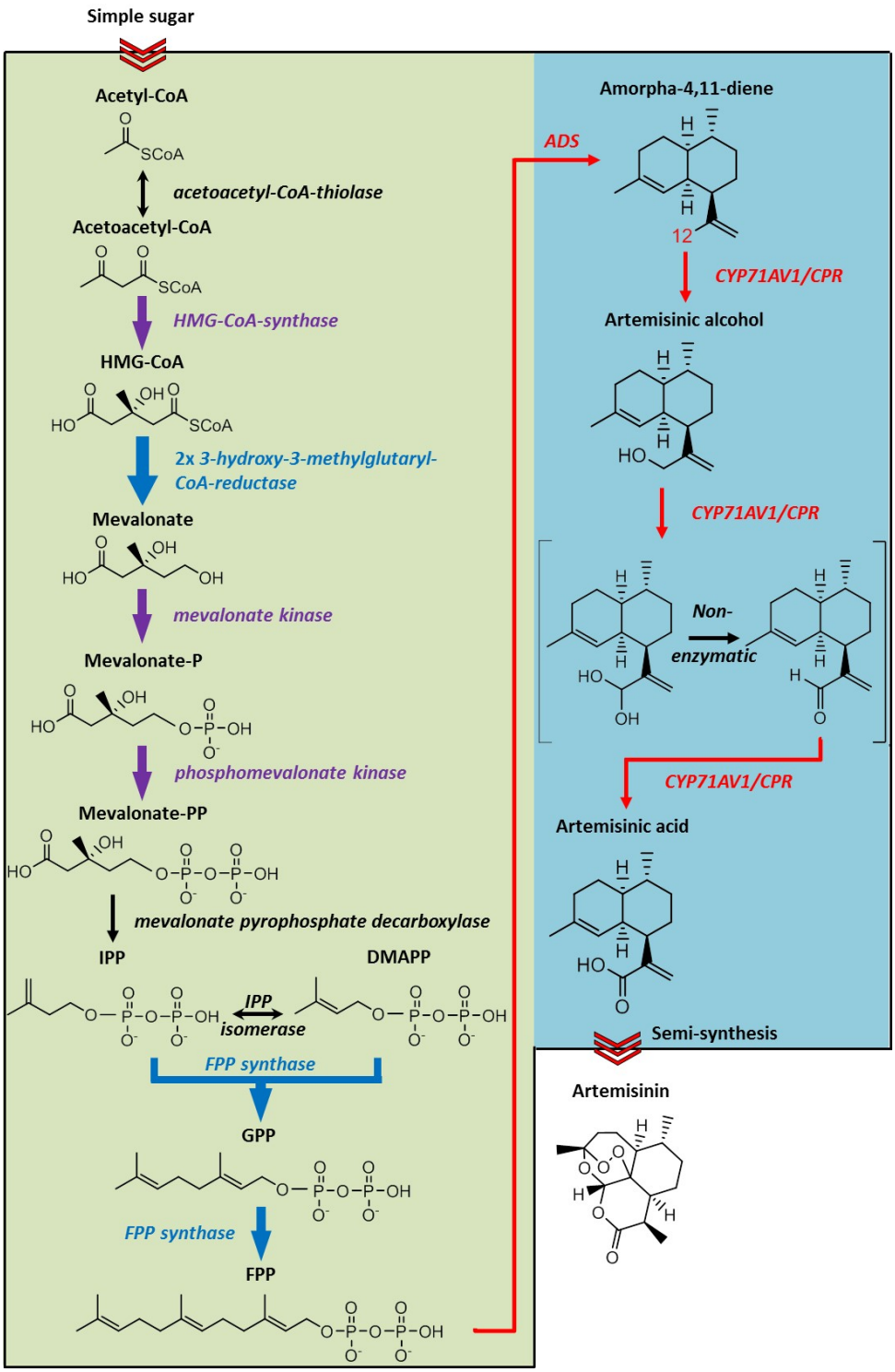


Figure 1.13: Schematic representation of the engineered artemisinic acid biosynthetic pathway[146]
the green background indicates the engineered mevalonate pathway in *S. cerevisiae* strain EPY224; the blue background shows the introduced artemisinic acid pathway; blue arrows: directly upregulated genes; purple arrows: *upc2-1* expression (global transcription factor regulating the biosynthesis of sterols) indirectly upregulated genes; red arrows: introduced biochemical pathway leading from farnesyl pyrophosphate (FPP) to artemisinic acid; IPP: isopentenyl pyrophosphate; DMAPP; dimethyl allyl pyrophosphate; GPP: geranyl pyrophosphate; ADS: amorphadiene synthase; see main text for description.

Another example of CYP implementation in artificial multi-enzyme cascades is the synthesis of the steroids pregnenolone and progesterone, mediated by bovine $P450_{scc}$ (CYP11A1) in recombinant *S. cerevisiae*, co-expressing adrenodoxin and adrenodoxin reductase to support activity of CYP11A1 (Figure 1.14).[151] The strain was able to produce a total pregnenolone concentration of 60 mg l^{-1}. Further improvement of this strain by Szczerbara and co-worker led to a full biosynthetic pathway for the production of hydrocortisone *via* progesterone, 17-hydroxy-progesterone and 11-deoxycortisol. The improved pathway involved 13 genes of which four were P450s.[152] Sanofi-Aventis, finally, adapted this system for industrial production. Further well-established commercial applications in the P450 based biotransformation of steroids is the 11-β-hydroxylation of Reichstein S to hydrocortisone and other antiinflammatory corticosteroids by *Curvularia* sp. (Schering, now Bayer)[153] as well as the conversion of progesterone to cortisone by *Rhizopus* sp. (Upjohn, now Pfizer). [154]

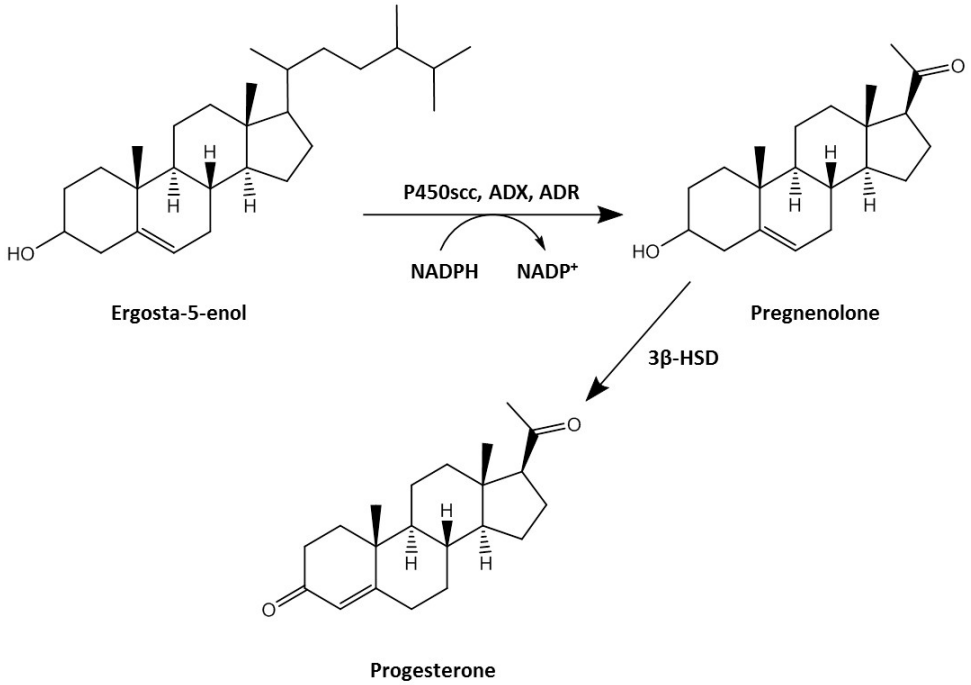


Figure 1.14: Schematic representation of the final step in progesterone biosynthesis by engineered yeast.
ADR: Adrenodoxin reductase; ADX: adrenodoxin; P450scc: bovine CYP11A1; 3β-HSD: human 3β-hydroxysteroid dehydrogenase/isomerase

In addition to drug and drug metabolite production, CYPs are utilized for several other commercial products. Amazing examples for this are transgenic plants that are able to produce flowers with unusual colours due to transferred CYP colouration genes from other plants.[130] Three classes of pigments (flavonoids, carotenoids, and betalains) primarily attribute to the colour of flowers, whereas the anthocyanins (class of flavonoids) cover colour ranges from orange over red to blue and violet as shown in Figure 1.15.[155] In order to obtain flowers with a blue or violet colour the biosynthesis of the pigment delphinidin by flavonoid 3',5'-hydroxylase (CYP75A/F3'5'H) is essential. The gene encoding F3'5'H was first isolated from petunia[156] and many other plants subsequently, and although it diverged before the specification of higher plants, numerous plants aren't able to synthesize delphinidin due to deficiency of F3'5'H.[157] Among them are *Rosa hybrida* (rose), *Chrysanthemum moriflolium* (chrysanthemum), *Dianthus caryophyllus* (carnation) and *Lilium* spp. (lily), which occupy more than 50 % of the cut-flower market. In order to gain access to an untapped market like this researchers develop genetically engineered transgenic "blue roses" by introducing CYPs from blue pansy and dihydroflavonol reductase from petunia into the rose with simultaneous suppression of dihydroflavonol reductase (Suntory Ltd, Japan; Calgene Pacific, now Florigene Pty Ltd, Australia).[158] Others plants like mauve carnations ("Moon dust") and violet carnation ("Moon shadow") have been developed and are available on the market in Japan and Australia.

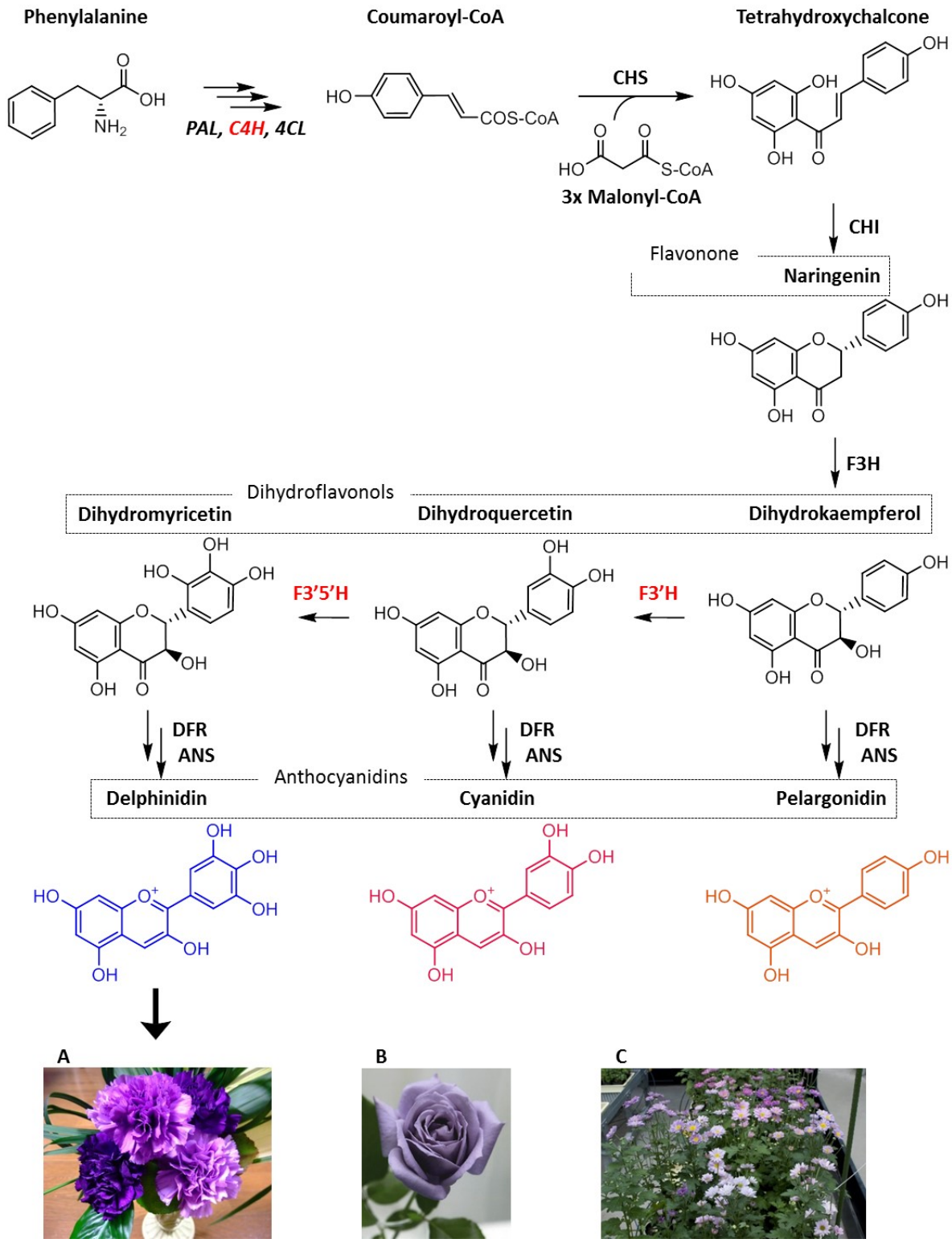


Figure 1.15: Anthocyanidins and flavonoid biosynthetic pathway relevant to flower colour. red letters: P450 enzymes; *PAL*: phenylalanine ammonia lyase; *C4H*: cinnamic acid 4-hydroxylase, *4CL*: 4-coumarate CoA ligase; CHS: chalcone synthase; CHI: chalcone isomerase; F3H: flavanone 3-hydroxylase; F3'H: flavonoid 3'-hydroxylase; F3'5'H: flavonoid 3'5'-hydroxylase; DFR: dihydroflavonol 4-reductase; ANS: anthocyanidin synthase; colour-modified transgenic flowers by expression of F3'H and/or F3'5'H in carnation (A), rose (B) and chrysanthemum (C).

1.1.5.3 Medicine

In recent years increased attention regarding the use of P450s in the medical sector has been concentrated on gene-directed enzyme prodrug therapy (GDEPT). It's a strategy that comprises a three-component system (Figure 1.16) and aims to convert prodrugs into active therapeutic metabolites in the cancer cell itself and thus improve the efficacy and reduce toxicity of cancer chemotherapeutic agents.[159] Although, no GDEPT product is currently available on the market it has shown great promise in cancer therapy.

One of the most intensively studied prodrug/enzyme system involving P450 gene transfer is the cytochrome P450/oxazaphosphorines system. Oxazaphosphorines include cyclophosphamide (CPA) which holds substantial antitumour activity but requires activation by hepatic cytochrome P450 (CYP2B6) catalysed 4-hydroxylation to yield 4-hydoxyl derivatives which are further metabolized to cytotoxic phosphoramide mustard.[160] Phosphoramide mustard, however, is in contrast to the 4-hydroxy metabolites unable to cross cell membranes. Consequently, activation of the prodrug in the liver can't produce effective local tumour concentrations as well as bystander cytotoxic effect.[161] The P450 GDEPT approach conquers this issue by introducing P450-expressing genes into tumour cells and therefore generating high levels of 4-OH-CPA directly in the tumour cells which in turn leads to strong bystander effects. Trials applying this system using human CYP2B6, CYP2C18 and CYP3A4 in 9L-gliosarcoma cells showed a strong cytotoxicity with CPA.[162] Even more encouraging are the results of phase II clinical trials using CYP2B6 conducted by Hunt and coworkers which led to the selective death of the tumour cells inflicting only minimal harm to normal cells.[163] Other CYPs apart from human P450s have been tested and shown great promise as well. The canine CYP2B11 for example showed in comparison to CYP2B6 an increased intratumoural concentration of 4-OH-CPA when tested *in vivo*.[164]

As P450s are responsible for drug metabolism in the body they also attracted attention for medical applications as biosensors to monitor drug levels in blood plasma.[165] In order to bypass the need for an electron donor (NADPH) and electron transfer partner (CPR), CYP enzymes that are employed as biosensors are usually immobilized on an electrode which directly transfers required electrons to the enzyme that either binds to or converts the drug. By unravelling the P450 substrate profile of a drug, biosensors allowing for example for the prediction of damaging drug-drug or food-drug interactions caused by the CYP metabolism.[166]

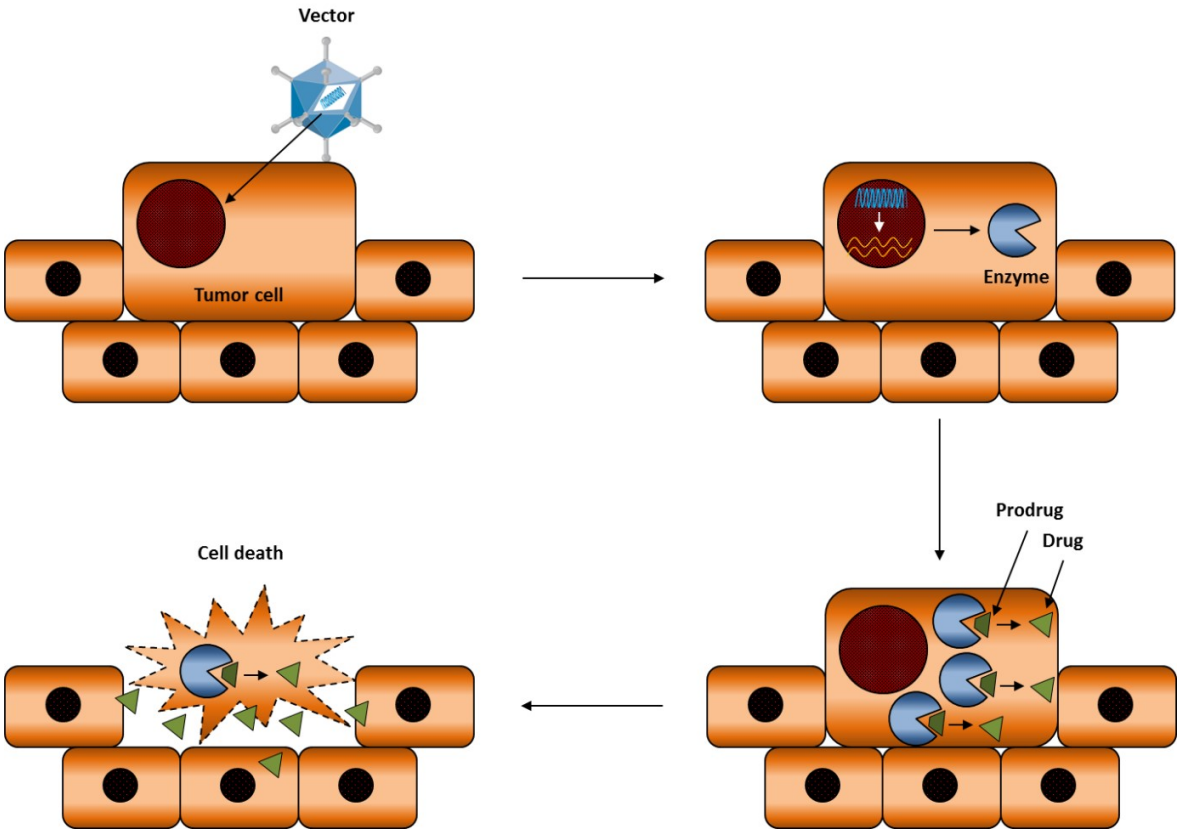


Figure 1.16: Schematic representation of gene directed enzyme pro-drug therapy (GDEPT)
The 3 component system usually comprises: an inactive drug (prodrug), a gene coding for an enzyme that converts inactive prodrug to an active drug, and a carrier (vector) that delivers the gene to a tumour cell with or without carriers.

1.1.5.4 Bioremediation

Soil and water pollution due to industrial chemicals like polycyclic aromatic hydrocarbons (PAHs); polychlorinated dibenzo-p-dioxins (PCDDs); and polychlorinated biphenyls (PCBs) as well as herbicides and explosives are a leading cause for environmental contamination. Detoxification of affected areas still remains a major challenge but is of crucial importance in order to exploit unused contaminated land, increase crop production and improve underground water quality, which is often used as drinking water. An alternative treatment technique to already existing physicochemical methods, applies enzymes that are able to transform or degrade these chemicals.[167-168] Recombinant bacteria expressing P450s that are involved in deactivation or degradation of toxic xenobiotics can for example be used for waste water treatment while transgenic plants can be used for detoxifacation of contaminated soil.

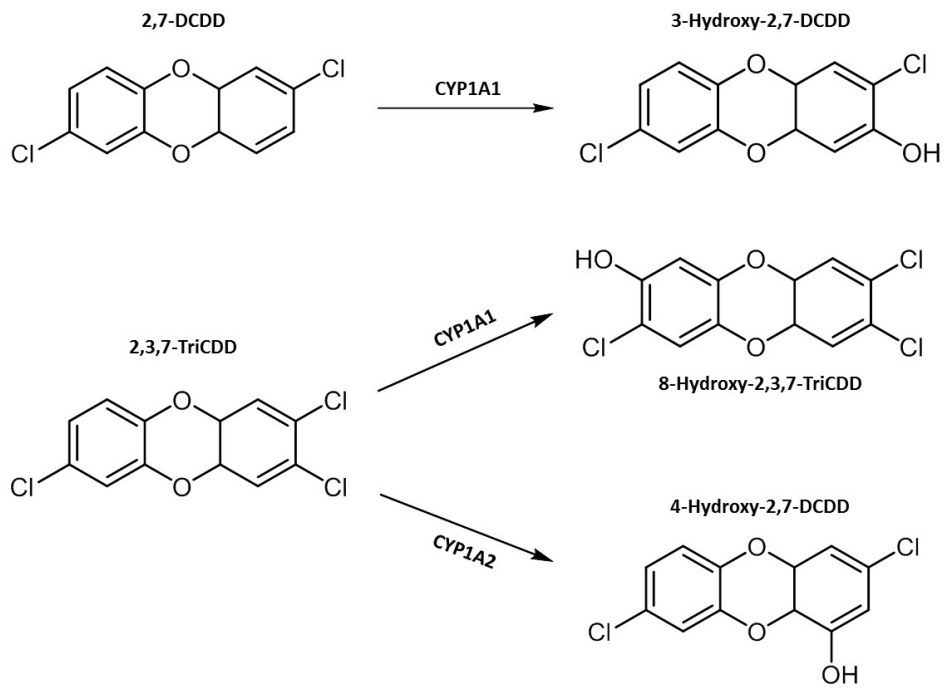


Figure 1.17: Oxidation of PCDDs by rat CYP1A1 and CYP1A2 expressed in *S. cerevisiae*

Numerous wild-type P450s of plant, bacterial and mammalian origin have been investigated with regard to potential detoxification capacities. Human CYP1A1 and CYP2B6, for example, could metabolize several herbicides as well as PAHs (CYP1A1) and PCBs (CYP2B6).[168] CYP76B1 from *Helianthus tuberos*[169] and CYP71A10 from soybeans[170] showed activity against herbicides such as phenylureas. Furthermore, enzymes engineered by either rational or directed evolution can be employed for possible bioremediation as shown for engineered $P450_{BM3}$ and $P450_{cam}$. Mutants of $P450_{BM3}$ showed enhanced activity towards terpenes and gaseous alkanes, while some $P450_{cam}$ mutants were able to metabolize PAHs and PHBs.[168] Construction of recombinant cells and transgenic plants expressing CYPs already showed high potential. Sakaki and coworker were able to express rat CYP1A1 and CYP1A2 in yeast which as a result were able to biodegrade PCDDs (Figure 1.17).[171] Moreover, some transgenic plants showed herbicide resistance as well as phytoremediation of environmental contaminants. Transgenic potato and rice plants generated by Inui and Ohkawa, for example, metabolized several herbicides, insecticides and industrial chemicals.[172] Another exceptional example for utilization of transgenic plants is an *Arabidopsis thaliana* strain that has been engineered to express a P450 fusion (XplA fused to flavodoxin redox partner) from *Rhodococcus rhodochorus* which enabled the plant to degrade RDX (hexahydro-1,3,5-trinitro-1,3,5-triazine) from contaminated soil.[77]

1.1.6 The fungal Kingdom

1.1.6.1 Fungi – the double-edged sword

Fungi are members of a large group of lower eukaryotic microorganisms, morphologically clustered into yeast, filamentous or dimorphic fungi.[1] They are classified as a kingdom, Fungi[173], which is separate from plants and animals, even though they share key components in cellular physiology and genetics with other kingdoms. Like plant cells, fungal cells possess a cell wall, which, however, lacks cellulose and contains chitin (a component of arthropod exoskeletons) instead. They live, similarly to animals, in a predominantly heterotrophic mode and inhabit a broad range of environments. Many play fundamental roles in nutrient cycling by decomposing organic matter, while others derive their nutrients from animal or plant hosts as either obligate or opportunistic pathogens, as well as through symbiotic relationships.[1, 174] The estimated origin of fungi is very inconsistent (660 Ma to 2.5 Ga)[175], but recent studies support an estimated time span between 760 Ma and 1.06 Ga.[176-177]

Within these approximately 900 million years of evolutionary history, the fungal kingdom developed a enormous diversity of taxa (Figure 1.18) with an estimated number of 1.5[178] to 5.1 million[179] species which are organised in four major groups, *i.e.*, ascomycetes, basidiomycetes, zygomycetes, and chytrids.[174]

The members of this kingdom have adapted to diverse ecological niches and therefore affecting nearly all other forms of life in either beneficial or detrimental manner. Moreover, it is assumed that symbiotic relationships of fungi with photosynthesizing organism paved the way in the establishment of eukaryotic life on land and thus affected Earth's atmosphere, climate, and evolution of animals.[180-181] Close relationships between fungi and plants still play a decisive role today as shown by the fact that 95% of all plant families have associated mycorrhizal fungi.[182] Furthermore, fungi play an important role in the global carbon cycle and in degrading organic material. Although fungi take up many beneficial roles in nature, they can have equally devastating impacts in form of pathogens for animals, plants and humans, respectively. Fungal plant pathogens which infect all major crop plants, for example, not only cause tremendous loss of crop yields (*e.g.*, *M. grisea* destroys enough rice annually to feed 60 million people)[183] but also lead to food contamination through the production of mycotoxins.

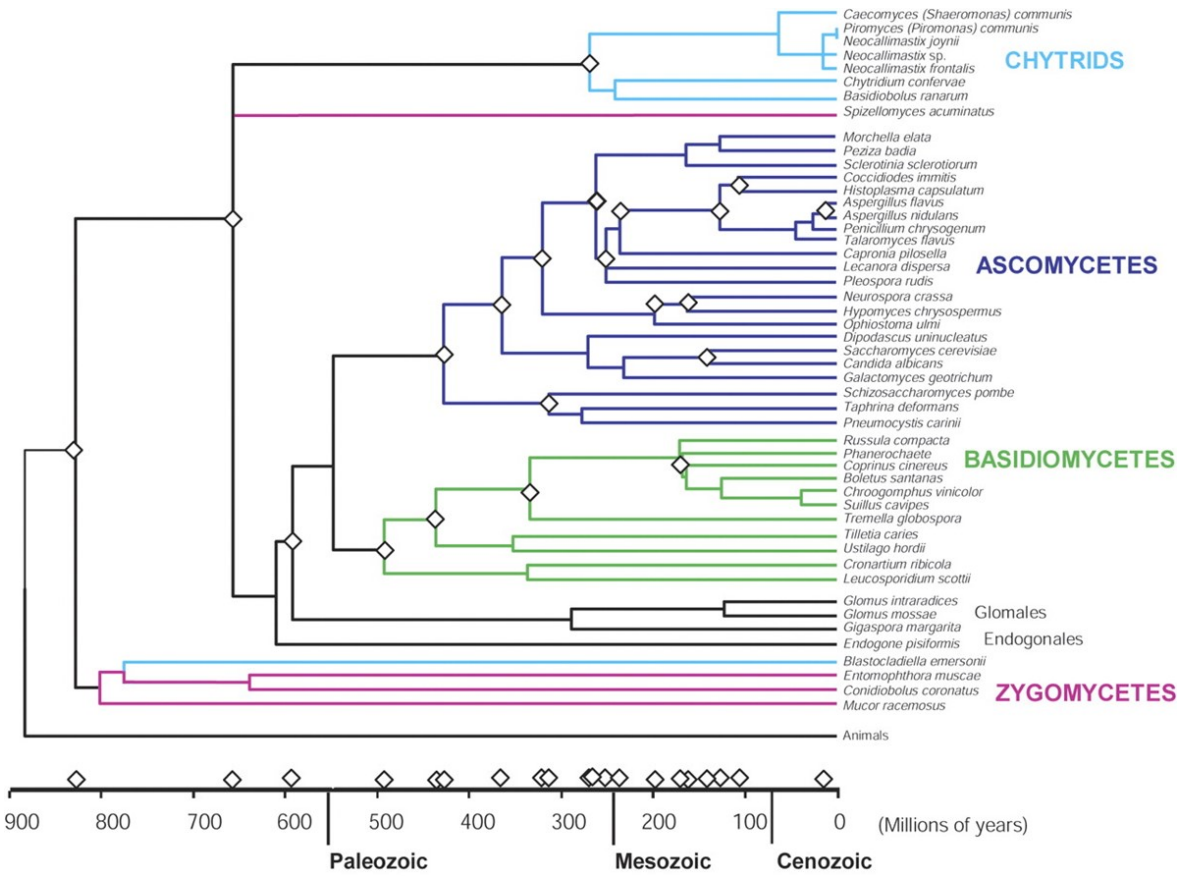


Figure 1.18: Phylogeny of the fungal kingdom[174]
Diamonds indicate evolutionary branch points, and their approximate dating (time line is displayed at the bottom of the image).

An astonishing example of how fungal plant pathogens even influence cultural development in human civilization is the fungus *Hemileia vastatrix*, which is responsible for the preference of tea over coffee in the British Empire due to infection of coffee plants in the 1870s and the consequential use of these fields for tea.[174] Furthermore, fungal pathogens like *Candida* and *Aspergillus* species pose a dramatic risk to immunocompromised or therapeutically immunosuppressed patients causing mortality rates between 20%[184] and 49%.[185] But also the healthy population get threatened by fungal infections as demonstrated by an outbreak of *C. neoformans* (2002, Vancouver Island, British Columbia)[186] and the annual increasing cases of *Coccidioides* infections in the USA.[187] Effective measures against fungal diseases are, however, hard to come by due to the eukaryotic biology they share with humans which usually leads to serious side effects of most existing antifungal drugs.[188] Nonetheless, fungi, particularly filamentous fungi, are of tremendous beneficial use for humans as they produce a vast array of secondary metabolites which are of significance for biomedicine, agriculture and industry.[189]

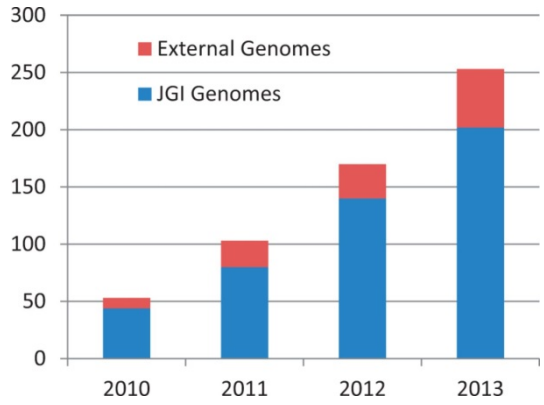


Figure 1.19: Growth of annotated genomes in MycoCosm (fungal genomics portal; update 2013).[190]
The genomes sequenced by JGI (Joint Genome Institute) are shown in blue and those sequenced by others are shown in red.

The growing interest in bioactive fungal metabolites, may they beneficial or detrimental, as well as the advancement of powerful available genomic tools in the last decade, have driven efforts in the development of genome-wide functional studies for an increasing number of fungal species. Several large-scale genomics initiatives like the 1000 Fungal Genomes Project contribute towards an ever growing number of complete fungal genomes (Figure 1.19) and thus providing an unique opportunity to study the biology and evolution of an entire eukaryotic kingdom.[190]

1.1.6.2 Fungal P450

Fungi, especially filamentous fungi, exhibit a long history of successfully occupying various ecological niches due to their ability to adapt to changing environments, which is owed to a large array of enzymatic mechanisms. Cytochrome P450 enzymes notably contribute to this fitness and fertility as they participate in a wide variety of physiological reactions. CYPome analysis of the white rot fungus *Phanerochaete chrysosporium,* which belongs to the only known group of microorganisms in nature that are capable of completely breaking down lignin to carbon dioxide and water, for example, suggest that several P450s are involved in the degradation of lignin and a broad range of environmental toxic chemicals.[191] CYPs also play important roles in the production of secondary metabolites like mycotoxins (*e.g.,* aflatoxin[192], trichothecenes[193], and fumonisins[194]) which usually serve to disable host defence responses or defend the fungus against other microorganisms[195] but pose potential problems to both public health and economics. Furthermore, cytochromes P450 are believed to be associated with fungal pathogenicity.[196] In addition to these highly specialized P450 functions, there are also

CYPs that are essential to the primary metabolism and participate, for example, in membrane ergosterol biosynthesis.[1] Consequently, these housekeeping CYPs became popular antifungal targets.[197]

20 years have passed by since the first complete fungal genome (*S. cerevisiae*)[198] was reported. With the sequencing of the first filamentous fungus (*Neurospora crassa*) [199], however, it became apparent that diversity within the fungal kingdom is much more substantial than previously assumed, as the newly acquired data revealed nearly twice as many genes in *N. crassa* and the lack of homologs to known proteins for over 40% of these genes. The same applies to fungal CYPomes which sizes can vary greatly from one P450 found in *Eremothecium cymbalariae* to over 300 in *Moniliophthora perniciosa* and *Postia placenta* (Figure 1.20), albeit relative species usually show some similarities in CYPome distribution.[1, 177] In general, filamentous fungi possess high numbers of CYP genes, while yeast-form fungi contain only very few CYP genes. It is therefore believed that the CYPome size reflects both evolutionary history and adaptation to the environment.[1]

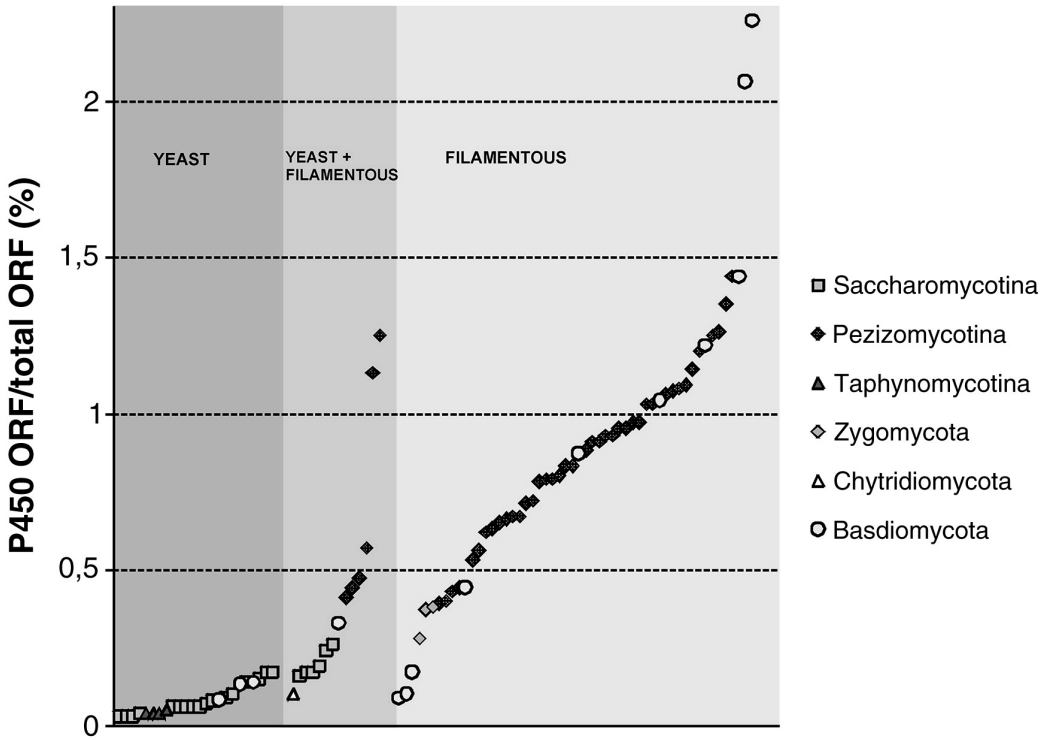


Figure 1.20: Relative P450ome sizes of fungal species belonging to different fungal (sub)phyla, grouped by their morphology.[1]

Although the CYP number in fungi is generally lower than that in plants, filamentous fungi have the highest density of CYPs considering the relative genome size leading to a great diversity. The family diversity of P450 genes, which mainly arose through events of gene

duplication and horizontal gene transfer,[200] differs considerably between species even among relatives. For instance, the Aspergilli *A. flavus*, *A fumigatus* and *A. nidulans* possess 93, 57 and 90 family types, respectively whereas they only share 45.[177] Although, fungal CYP families (399 families)[97] are currently allocated to the groups CYP51-CYP69, CYP501-CYP699 and CYP5001-CYP6999 the vast majority of their functions and biological roles still remain unclear. Furthermore, there are still many fungal CYPs that need to be newly assigned. However, most fungi share two global highly conserved families, CYP51 and CYP61, which play essentials roles in fungi housekeeping functions (sterol biosynthesis)[1, 201] and could provide important information on evolution on fungi and CYPome.

1.2 *Beauveria bassiana*

The impact of insects on the economic and health sector, either as a result of agricultural pests or through vector-induced diseases in humans and animals, is a serious problem for society. Attempts to control insect infestations usually employ methods such as the use of chemical insecticides. The increase of environmental awareness, public concern about food safety as well as development of insect resistance to chemicals, however, led to an advanced interest in alternative strategies. Biopesticides based on entomopathogenic fungi like *Beauveria bassiana* (Figure 1.21) offer therefore a perfect solution as they already play major roles in the natural regulation of insect populations.[202] Unlike other insect-pathogenic microorganisms, entomopathogenic fungi infect target hosts by penetrating through their cuticle and thus act by contact and do not require ingestion, which makes them very promising in the development of biological control agents (BCAs).[203] The epicuticle itself is the outermost insect surface, composed of very long-chain hydrocarbons with variable amounts of fatty alcohols, fatty acids and wax esters and therefore represents the first barrier against chemical and biological attacks, as well as desiccation.[204-205] Infection of target insects usually begins with adhesion of fungal spores or conidia to the surface of the insect cuticle. The conidium will then germinate and form a germ tube that couples mechanical pressure to the secretion of degrading enzymes in order to penetrate the host cuticle. Once the fungus has passed through the cuticle, invasion of the insect body and circulatory system (hemolymph) follows, leading to the death of the host by physiological starvation in 3-7 days.[202, 206] Of the approximately 750 known species of entomopathogenic fungi only 12 are used for currently available commercial products,

whereas *B. bassiana* accounts for nearly 40% of these mycoinsecticides (*e.g.*, Ostrinil, Boverin, Boveriol).[202, 207]

Figure 1.21: A spider colonized by *B. bassiana*

Beauveria bassiana belongs to the Division of Ascomycota and is one of the best-known species of entomopathogenic fungi. It was first discovered by Agostino Bassi de Lodi in 1835 during his investigation of the heavy decline in larval silkworms caused by muscardine disease.[208] Since then more 700 host species have been reported for *B. bassiana* including major economic insect pests (*e.g.*, European corn borer, *Ostrinia nubilalis*; codling moth, *Laspeyresia pomonella*; Japanese beetle, *Popillia japonica*; Colorado potato beetle, *Leptinotarsa decemlineata*; chinch bug, *Blissus leucopterus*; and the European cabbageworm, *Pieries brassicae*)[209] as well as mosquito[210] and flea vectors.[211] Furthermore, *B. bassiana* can also adopt a saprophytic lifestyle and recent studies indicate interaction of *B. bassiana* with plants in form of endophytes (endosymbiont that lives within a plant) as an adaptive protection against herbivorous insects.[202] For instance, *B. bassiana* has been found in plant tissues of corn (*Zea mays*)[212], cacao (*Theobroma cacao*)[213] and poppy (*Papaver somniferum*)[214]. The underlying mechanisms that allow for such physiological plasticity of *B. bassiana* are still poorly understood. Enzymes taking part in the degradation of the insect cuticle like chitinases, lipases, proteases and, as of recently, putative cytochromes P450 are believed to contribute to *B. bassiana* pathogenesis.[215-217] Also secreted protein toxins (bassiacridin) and primary metabolites (oxalic acid) as well as several secondary metabolites (beauvericin, bassianolide) produced by *B. bassiana* appear to play significant roles as virulence factors, but may also serve beyond that, in form of antibiotics, antifungals,

insecticides and nematicides, in order to defend against competing parasites and saprophytes in nature.[218]

In addition to its use as biopesticide, *Beauveria bassiana* ATCC 7159 is also employed as whole cell biocatalysts in chemical and industrial applications and is second only to *Aspergillus niger* as the most frequently used fungal biocatalyst. In 2000 Grogan and Holland summarized the range and applications of *Beauveria*'s biocatalytic reactions, many of which are believed to involve cytochrome P450s (Figure 1.22), corroborating its potential as a powerful tool in synthetic chemistry, especially for hydroxylation processes.[219] Furthermore, the genome of *B. bassiana* strain ARSEF 2860 has recently been sequenced[208], and revealed 83 open reading frames (ORF) that putatively encode P450s, each of which have yet to be analysed in more detail.

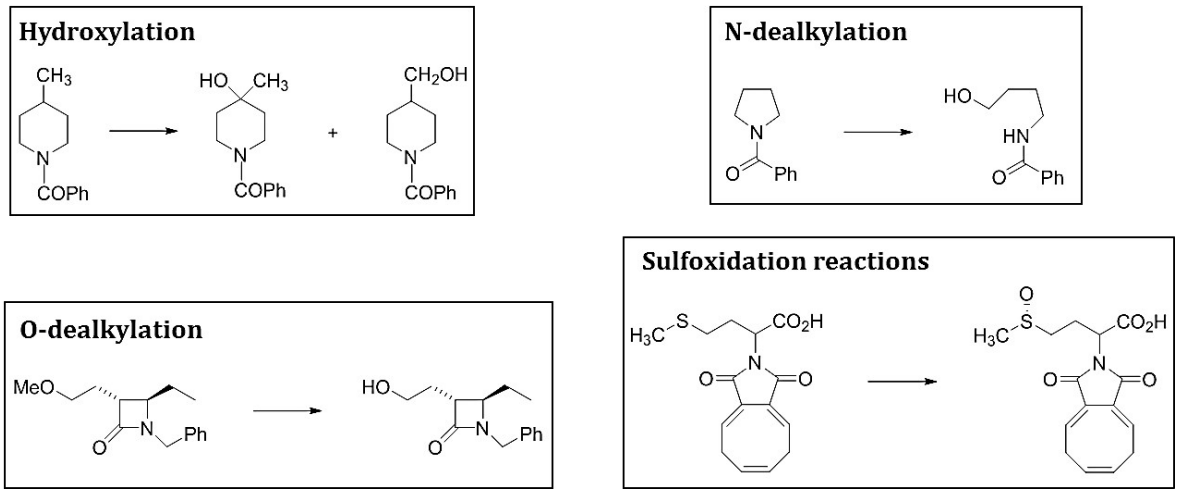


Figure 1.22: Variety of *B. bassiana*'s reaction attributable to cytochrome P450

1.3 *Rhodococcus jostii*

Rhodococci belong to the genus of aerobic, non-motile and non-sporulating bacteria with DNA containing high GC content (63-73 %). They are commonly found in soil, where they degrade a wide range of organic compounds but inhabit also environmental niches like salt water, foams in activated sludge, guts of blood-sucking arthropods and herbivore dung[220] as well as animals and plants.[221] Because of their robust nature and their outstanding ability to degrade aromatic compounds, steroids and a vast variety of other chemicals, they are used as biocatalysts in industry.[220, 222-223] Applications of Rhodococci include bioactive steroid production, fossil fuel biodesulfurization, and the production of acrylamide and acrylic acid.[223] This outstanding metabolic versatility is most likely attributed to the large number of genes encoding oxygenase-like proteins such as

cytochrome P450s. Genome analysis of *Rhodococcus jostii* by McLeod *et al.* in 2006[223] revealed at least 25 ORFs for putative P450s. However, to date not much is known about the *in vivo* role of CYPs identified in *Rhodococcus jostii* RHA1, except for CYP125 and CYP257A1, which are thought to be involved in bacterial sterol/steroid degradation.[224-225] Further investigation of *Rhodococcus jostii* RHA1 with regard to the diversity of P450s may thus result in the identification of interesting biocatalysts for applications in bioconversion and biotransformation processes, but also increase our fundamental knowledge about physiological and metabolic properties of members of the genus Rhodococcus.

1.4 Project Aim

Cytochrome P450 enzymes from filamentous fungi, such as *Beauveria bassiana*, present attractive targets for industrial applications due to their involvement in the catalysis of various hydroxylation reactions. Genome sequencing programs revealed very recently 83 genes encoding putative P450s of mostly unknown function in *B. bassiana* that needed to be investigated in more detailed.

But also bacteria such as *Rhodococcus* sp. have a long standing history in industrial biocatalysis and offer, just like filamentous fungi, an interesting set of cytochrome P450 enzymes which have yet to be studied to a greater extend.

The aim of this project is to identify P450s of *B. bassiana* for cloning, recombinant expression and as subjects for functional studies for further characterisation and industrial application. Extensive bioinformatics studies of the *Beauveria* CYPome will be conducted and diverse expression systems as well as expression conditions will be tested to obtain functional enzymes.

A further aim of this project is focused on a library of 23 chimeric P450 fusion enzymes from *Rhodococcus jostii*. Optimal expression conditions for these fusion enzymes will be determined in order to provide a screening platform that can be applied in an industrial context.

2. General Material and Methods

2.1 Chemicals

Chemicals used in this study were purchased from AGTC Bioproducts Ltd (Hessle, UK), Alfa Aesar (Heysham, UK), Fisher Scientific UK Ltd. (Loughborough, UK), Merck Chemicals Ltd. (Nottingham, UK), Scientific Laboratory Supplies Ltd. (Nottingham, UK), Sigma-Aldrich Company Ltd. (Dorset, UK), Takara Bio Europe Clontech (St Germain-en-Laye, France), and VWR International Ldt. (Lutterworth, UK).

Limonene was donated by the Giulia Paggiola of the Green Chemistry Department (University of York, UK). 65 commercial available drugs used for screening assays of *Rhodococcus jostii* chimeric fusion were kindly provided by AstraZeneca.

Restriction enzymes were bought from New England Biolabs (Ipswich, UK), Promega UK Ltd. (Southampton, UK), and Thermo Fisher Scientific Biosciences GmbH (St. Leon-Rot, Germany).

PCR Primers were synthesised by Eurofins Scientific (Wolverhampton, UK) and genes by GeneArt (now Life Technologies Ltd, Paisley, UK).

2.2 Strains and Plasmids

2.2.1 *Escherichia coli* strains

Bacterial strains are summarized in Table 2.1.

Table 2.1: List of *E. coli* strains used for cloning and recombinant expression

Strain name	Genotype	Manufacturer
BL21(DE3) Singles™	F^- *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm* (DE3)	Novagen-Merck
NovaBlue Singles™	*endA1 hsdR17* (r_{K12}^- m_{K12}^+) *supE44 thi-1 recA1 gyrA96 relA1 lac* F′[*proA*$^+$*B*$^+$ *lacI*^{q}*Z*Δ*M15*::Tn*10*] (Tet^R)	Novagen-Merck
Rosetta™ 2(DE3)	F^- *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3) pRARE2 (Cam^R)	Novagen-Merck

2.2.2 Yeast strains

The properties of the strain used for microsome preparation are summarized in Table 2.2.

Table 2.2: *Saccharomyces cerevisiae* strain used in this work

Strain name	Genotype	optained from
Saccharomyces WAT11	*MATa*; *ade*2–1; *his*3–11,-15; *leu*2-3,-112; *ura*3–1; can^R; cyr^+ (a derivative of the W303-B strain)	Prof. D. Werck Reichhart (CNRS-Institute de Biologie moléculaire des plantes, Strasbourg)

2.2.3 Plasmids

Plasmids for gene cloning and enzyme expression are shown in Table 2.3.

Table 2.3: Plasmids for gene cloning and enzyme expression

vector	antibiotic resistance	features	source
LIC	Kanamycin	pETYSBLIC3C vector, cleavable his tagged N-terminus	Dr. Gideon Grogan[226]
LICRED	Kanamycin	pETYSBLIC3C vector, cleavable his tagged N-terminus, contains Rhf reductases from *Rhodococcus* sp.	Dr. Federico Sabbadin[227]
pGro7	Chloramphenicol	Chaperone groES-groEL, *araB* Promotor	Takara
pYeDP60	Ampicillin	GAL10-CYC1 promoter	Prof. Daniele Werck-Reichart[228]

The LICRED[227] (Figure 2.1) and the LIC-vector are based on pETYSBLIC3C vector[228], which contains an additional cleavable his tagged N-terminus.

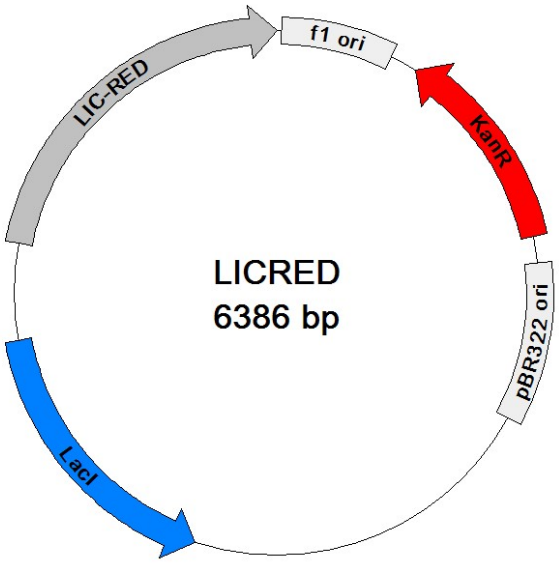


Figure 2.1: Vector map of LICRED
LICRED = Ligation independent cloning site with additional P450 reductase domain (RhfRED) of cytochrome $P450_{Rhf}$ from *Rhodococcus* sp. NCIMB 9784, f1 ori = f1 phage origin of replication, pBR322 ori = origin of replication of the plasmid pBR322, KanR = kanamycin resistence gene, LacI = repressor gene for IPTG induction

Figure 2.2 shows the LIC-RED region of the LICRED-vector[227] in more detail.

Figure 2.2: LIC-RED cloning site
pT7 = T7 promotor, oLAC = Lac operator, RBS = ribosome binding site, His = 6x histidines residues, which can be cleaved, 3C = HRV 2C protease site, tT7 = T7 terminator, Gene = exchangeable cytochrome P450, RhfRED = permanent P450 reductase domain of cytochrome P450Rhf from *Rhodococcus* sp. NCIMB 9784

The vector pYeDP60 (Figure 2.3) is used as a shuttle vector for cloning in *E. coli* but expression in yeast.[228]

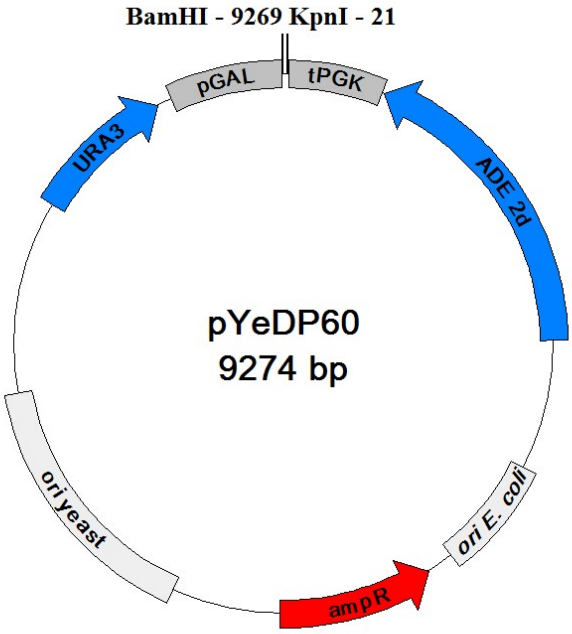


Figure 2.3: pYeDP60 shuttle vector map
tPGK = phosphoglycerate kinase terminator, pGAL = galactose promoter, ADE 2d and URA3 = selection marker for adenine and uracil auxotrophy, *amp*R = ampicillin resistence gene, *Bam*HI and *Kpn*I = restriction site, ori *E. coli* = origin of replication for *E. coli*, ori yeast = origin of replication for yeast

2.3 Media

2.3.1 *Escherichia coli*

Escherichia coli were grown in Lysogeny broth (LB) and M9 minimal media. Super Optimal broth with Catabolite repression (SOC) was used for recovering *E. coli* cells after plasmid DNA transformation.

LB-Medium[229]

Tryptone		10 g
Yeast extract		5 g
NaCl		10 g
H_2O	fill up to	1000 ml

The pH of this medium ranges between 7 and 7.5. 1.5 % [w/v] agar was added to LB prior autoclaving for the preparation of the solid agar plates. Depending on requirements supplements listed in Table 2.4 were added.

M9-Medium[229]

Preparation of M9 Salt stock:

Na_2HPO_4		33.9 g
KH_2PO_4		15 g
NaCl		2.5 g
NH_4Cl		5.0 g
H_2O	fill up to	1000 ml

➔ Autoclave

Preparation of M9 minimal medium:

M9 salts		200 ml
20 % glycerol/glucose*		20 ml
1 M $MgSO_4$*		2 ml
1 M $CaCl_2$*		100 µl
H_2O*	fill up to	1000 ml

∗ sterile filtrated prior usage

Depending on requirements supplements listed in Table 2.4 were added.

SOC-Medium[229]

Tryptone		20 g
Yeast Extract		5 g
$MgSO_4$		4.8 g
Glucose*		3.603 g
NaCl		0.5g
KCl		0.186 g
H_2O	fill up to	1000 ml

∗Sterile filtrated glucose was added after autoclaving

2.3.2 Yeast

Saccharomyces WAT11 was cultivated in YPAG medium while SGI medium was employed to select positive transformants. YPGE medium was used for expression of microsomes subsequently.

YPGA-Medium[230]

Bactopeptone		10 g
Yeast extract		10 g
Glucose		20 g
Adenine		200 mg
H_2O	fill up to	1000 ml

2 % [w/v] pastagar was added to YPGA prior autoclaving for the preparation of the solid agar plates.

SGI-Medium[230]

Bactocasamino acids		1 g
Yeast nitrogen base		7 g
Glucose		20 g
Tryptophan		20 mg
H_2O	fill up to	1000 ml

2 % [w/v] pastagar was added to YPGA prior autoclaving for the preparation of the solid agar plates.

YPGE-Medium[230]

Bactopeptone		10 g
Yeast extract		10 g
Glucose		5 g
Ethanol*		30 ml
H_2O	fill up to	1000 ml

∗Ethanol was added after autoclaving

2.3.3 Supplements

Table 2.4: Supplements for bacteria and yeast media

Supplement	Stock solution	Concentration in medium
Chloramphenicol	20 mg/ml in ethanol	20 µg/ml
Kanamycin	34 mg/ml in H_2O	34 µg/ml
δ-aminolevulinic acid	0.5 M in H_2O	0.5 mM
$FeCl_3$	0.5 M in H_2O	0.5 mM
IPTG	1 M in H_2O	1 mM
(L)-arabinose	20 g/l in H_2O	3 mg/ml
Galactose	200 g/l in H_2O	10 g/l

2.4 Glycerol stocks

E. coli laboratory stocks were prepared for long-term storage by adding 500 µl of sterile 80% glycerol to 500 µl of logarithmic growing cells in LB. The addition of glycerol stabilizes the frozen bacteria, prevents damage to the cell membranes and keeps the cells alive. The glycerol stocks were stored stably at -80°C.

Yeast glycerol stocks were prepared by adding 40 % sterile glycerol in a ratio of 1:1 to overnight cultures of yeast grown in the appropriate medium.

2.5 Working with nucleic acids

2.5.1 Enzymatic restriction of DNA

Sequence specific hydrolysis of nucleic acids by restriction endonucleases is used to create fragments with defined ends. For preparative digests, DNA was purified prior to the reaction by gel purification (2.5.3). Restriction of DNA was performed using the buffer system supplied by the manufacturer (Table 2.5). Plasmid DNA and PCR-fragments were incubated for 2-4 h at 37 °C. The digest of 1 µg DNA required a volume of at least 10 µl with no more than 10 % enzyme in the total volume. The reaction stop was initiated by either adding 6x Loading Dye or by heat inactivation. Restriction endonucleases used in this project are listed in Table 2.5.

Table 2.5: List of restriction endonucleases

Restriction endonuclease	Manufacturer	Cut site	Buffer
*Nhe*I	Thermo Scientific	5'-G^CTAGC-3'	Tango™
*Xba*I	Promega	5'-T^CTAGA-3'	Buffer D
*Bse*RI	NEB	5'-GAGGAG$(N)^{10}$^-3'	NEBuffer 2
*Bam*HI	Thermo Scientific	5'-G^GATCC-3'	Buffer *Bam*HI
*Kpn*I	Thermo Scientific	5'-GGTAC^C-3'	Buffer *Bam*HI

2.5.2 Plasmid DNA preparation

5 ml LB medium (2.3.1) containing the appropriate antibiotics (2.3.3) were inoculated with *E. coli* cells comprising the anticipated plasmid and incubated overnight at 37 ºC and 180 r.p.m.. The cultures were harvested by centrifugation at 13000 r.p.m. for 1 min (Sigma 2k15 centrifuge). A GenElute (TM) HP Plasmid Miniprep Kit (Sigma) was used to purify the plasmids according to the manufacturer's instructions.

2.5.3 Purification of DNA fragments

DNA fragments resulting from PCR amplification and restriction digest were separated using agarose gel electrophoresis (2.5.4) and excised from the gel before purification. A GenElute Gel Extraction Kit (SigmaAldrich) was used according to the manufacturer's protocol to purify DNA fragments cut out of agarose gels.

2.5.4 Agarose gel electrophoresis

Horizontal agarose gel electrophoresis was used for preparative and analytical separation of DNA. Separation took place in gel electrophoresis cells (Hybaid or Whatman Biometra). Depending on the size of the expected DNA fragments, agarose concentration varied between 0.8-1.5 % (w/v). Agarose was dissolved in 1x Tris-acetate-EDTA (TAE) buffer, prepared from a 50x TAE stock. To visualise DNA under UV light SYBR® Safe DNA Gel Stain (Fisher Scientific Ltd) was added directly to the gel. Samples have been mixed with 0.2 Vol. 6x Loading Dye (Promega) for visual tracking of DNA migration during electrophoresis. Furthermore, the presence of glycerol ensures that the DNA in the ladder and sample forms a layer at the bottom of the well. To determine the product size 1 kb DNA ladder (NEB) was used according to the manufacturer's instruction. Electrophoresis usually took place at a constant voltage of 100 V using 1x TAE running buffer. Running times were set between 0.75 to 1.5 h.

50x TAE-Puffer[229]

Tris		242 g	10 mM
Glacial acetic acid		57 ml	1 mM
EDTA (0.5 M, pH 8.0)		100 ml	50 mM
H_2O	fill up to	1000 ml	

Adjusting of pH 7.5 takes place using HCl.

2.5.5 *In vitro*-Amplification of DNA by PCR

Fragments for cloning reactions have been amplified by polymerase chain reaction (PCR) using KOD Hot Start DNA Polymerase (Merck Chemicals Ltd). KOD Hot Start DNA Polymerase generates blunt-ended PCR products suitable for cloning LIC Vector Kits. Primers for ligation independent cloning have been designed manually or with HiTel software (TF Protein Production Laboratory, University of York, http://bioltfws1.york.ac.uk/cgi-bin/primers.cgi?). Primers for In-Fusion cloning were designed using the TaKaRa Clontech online tool (http://www.clontech.com/US/Support/xxclt_onlineToolsLoad.jsp?citemId=https://www.takara-bio.co.jp/infusion_primer/infusion_primer_form.php&section=16260&xxheight=1800). Eurofins mwg/operon synthesised the in this work used Primers. PCR components and reaction conditions are shown in table Table 2.6 and Table 2.7.

Table 2.6: Components of the PCR reaction

Component	Volume (µl)
Template DNA*	1
Forward primer	1
Reverse primer	1
$MgSO_4$ (25 mM)	1
dNTP's (2 mM each)	5
KOD hot start polymerase buffer (10x)	5
KOD hot start polymerase	1
water to a final volume of 50 µl	

∗ for colony PCR: cell material from selected colonies using sterile pipette tip

Table 2.7: PCR reaction conditions

	Temp (°C)	Time	N° Cycles
Initial Denature	95	2 min	1
Denature	95	30 s	35
Anneal	55-58	30 s	
Extension	72	30 s per 500 bp	
Final Extension	72	10 min	1
Hold	4	n	1

2.5.6 Ligation independent cloning method

The cloning method used was based on T4 polymerase treatment of insert and vector resulting in complementary long overhangs and therefore required no ligation step in between.

Genes, amplified by PCR (2.5.5), were cloned into LIC-3-C or LICRED (Table 2.3) vector. Both vectors needed to be linearized prior to T4 polymerase treatment by digestion with restriction enzyme *Bse*RI (Table 2.5). The linearized vectors as well as the PCR products were then separated on a 1% agarose gel (2.5.4) and purified by gel elution (2.5.3) with a GenElute Gel Extraction Kit (SigmaAldrich).

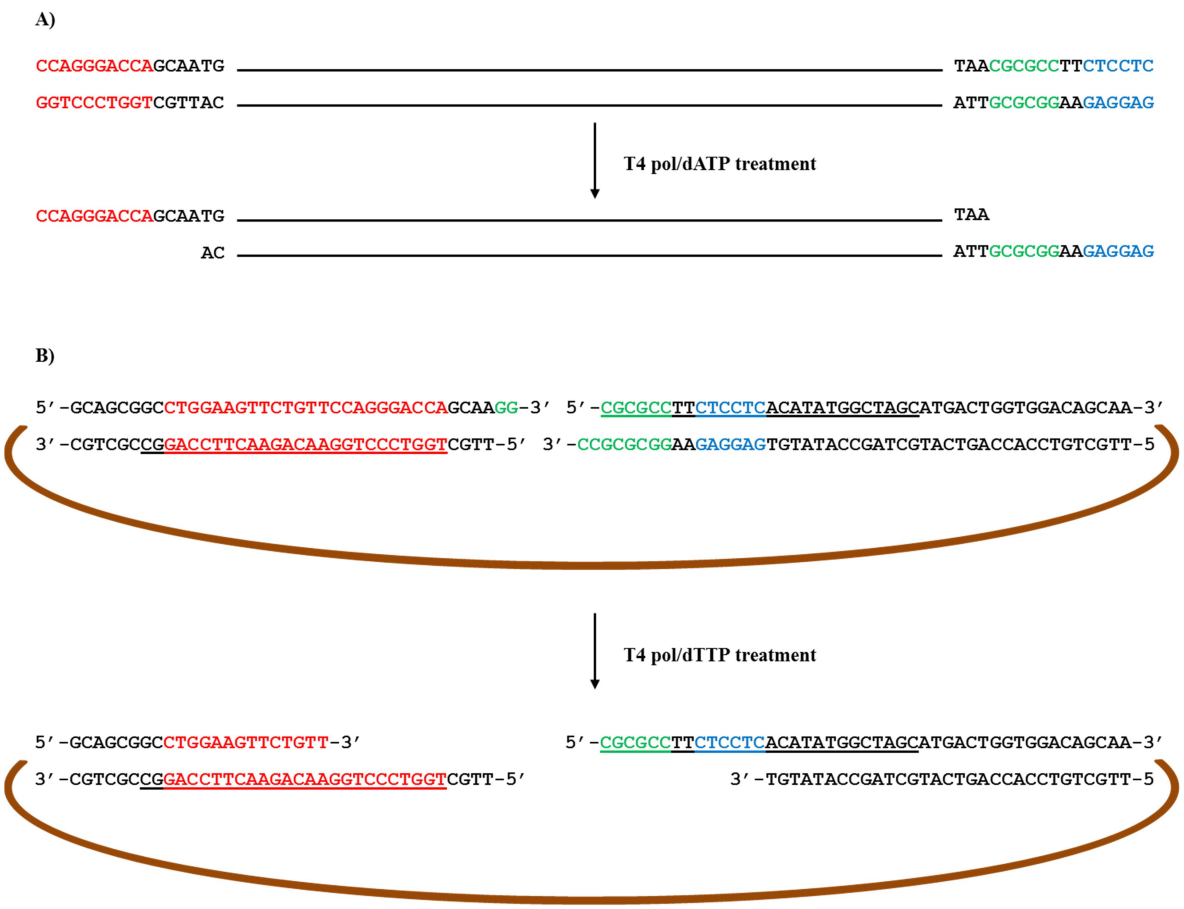


Figure 2.4: Scheme of T4 polymerase treatment
A) Insert, B) LIC-vector. *Bse*RI restriction site (blue letters), *Asc*I restriction site (green letters), 3C protease site (red letters); see text below for further description.

The purified insert and vector are treated with T4 polymerase (LIC qualified, Novagen-Merck) which has a 3′→ 5′ exonuclease activity and therefore creates complementary long sticky ends when supplied with bases T (in vector; Figure 2.4 A) and A (in insert; Figure 2.4 B). The protocol for the T4 polymerase treatment is shown in Table 2.8.

Table 2.8: Protocol for T4 polymerase treatment

vector		**insert**	
40 µl	10x T4 pol buffer	2 µl	10x T4 pol buffer
4 pmol	linearised vector	0.2 pmol	insert
20 µl	100 mM DTT	1 µl	100 mM DTT
10 µl	100 mM dTTP	0.5 µl	100 mM dATP
8 µl	T4 DNA polymerase*	0.4 µl	T4 DNA polymerase*
water to a final volume of 400 µl		water to a final volume of 20 µl	

* 2.5U/ l LIC qualified T4 DNA polymerase, Novagen/Merck

Adding 1 µl LIC prepared vector (~50 ng/µl) to 2 µl insert, incubation at room temp (~20-22 ºC) for 10 min, followed by addition of 1 µl EDTA (25 mM) to give a final volume of 4 µl and further 10 min incubation at room temperature, leads to an annealing reaction. 2 µl of LIC annealing reaction is then transformed (2.6.3) into NovaBlue Single competent cells (Table 2.1).

2.5.7 In Fusion cloning

The In-Fusion® HD Cloning System (Clontech) employs the In-Fusion HD enzyme which is able to fuse PCR-generated sequences to linearized vectors by recognizing a 15 bp overlap at their ends. Therefore, this system provides the advantage to clone any PCR fragment or multiple fragments into any linearized vector in a single reaction without the need for additional vector dephosphorylation, blunt-end polishing or PCR fragment digestion.

Figure 2.5 shows the principle of the In-Fusion® HD Cloning System which is based on a ligase-independent mechanism. The ability of the In-Fusion enzyme to create single-stranded regions at the ends of the PCR insert and linearized vector exposes the 15 bp complementary regions. The insert and vector DNA molecules will then spontaneously anneal through base pairing. Any single-stranded gaps will be repaired by transformation into *E. coli*, subsequently. Furthermore, this method allows for >95% cloning efficiency according to the manufacturer.

Table 2.9 gives an overview of the set up for the In-Fusion reaction comprising reaction components and quantity. The mix was then incubated for 15 minutes at 50 ˚C and placed on ice, subsequently. Transformation into NovaBlue competent *E. coli* cells followed as described in 2.6.3.

Table 2.9: In-Fusion cloning reaction set up

Reaction Component	Quantity
Purified PCR fragment	10–200 ng*
Linearized vector	50–200 ng**
5x In-Fusion HD Enzyme Premix	2 µl
Deionized Water	to 10 µl

*<0.5 kb: 10–50 ng, 0.5 to 10 kb: 50–100 ng, >10 kb: 50–200 ng
**<10 kb: 50–100 ng, >10 kb: 50–200 ng

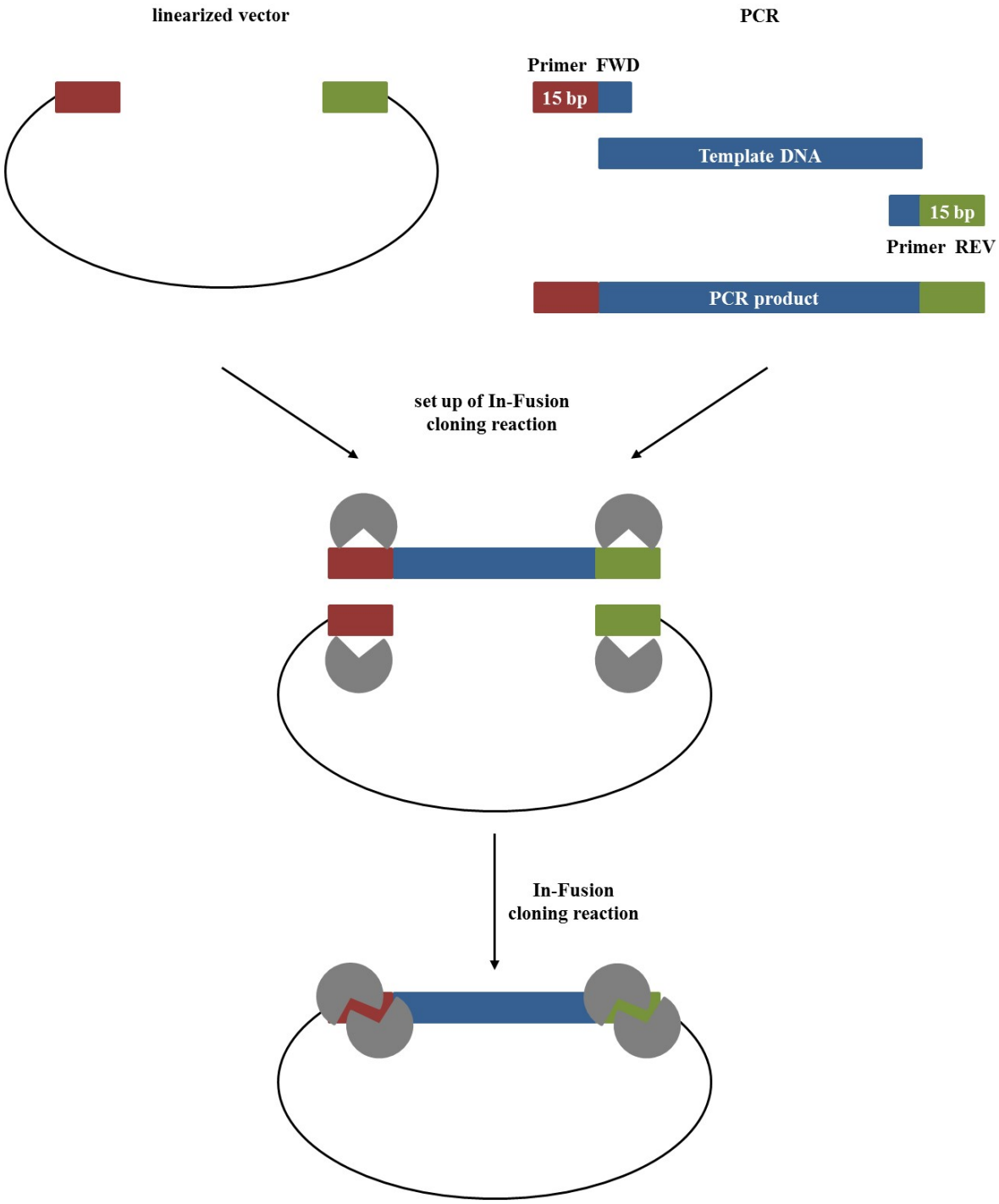


Figure 2.5: Assembly scheme of the In-Fusion® HD Cloning System
Primers for the PCR (top right) are colour-coded to emphasize their homology to the DNA template (blue) and cloning sites of the linearized vector (top left). The gene specific areas of forward (Primer FWD) and reverse primer (Primer REV) are indicated in blue and the 15 bp extension homologous to the corresponding sites in the vector are shown in either red or green. The In-Fusion enzyme (grey) fuses the PCR-generated sequence and linearized vector by recognizing a 15 bp overlap at their ends.

2.6 Preparation of recombinant microorganisms

2.6.1 Preparation of $CaCl_2$-competent *E. coli*-cells

Since *E. coli* is not naturally transformable, the ability to take up DNA or competency must be induced by chemical methods using divalent and multivalent cations (calcium, magnesium, manganese, rubidium, or hexamine cobalt)[231-232]. Alteration in the permeability of the membranes allows DNA to cross the cell envelope of *E. coli* which is composed of an outer membrane, an inner membrane, and a cell wall. The negative charges of the incoming DNA, however, are repelled by the negatively charged portions of the macromolecules on the bacterium's outer surface. The addition of $CaCl_2$ serves to neutralize the unfavourable interactions between the DNA and the polyanions of the outer layer. The reaction mixture is then exposed to a brief period of heat-shock at 42 ˚C. The change in temperature alters the fluidity of the semi-crystalline membrane state achieved at 0 ˚C thus allowing the DNA molecule to enter the cell through the zone of adhesion.

Procedure[231]: 5 ml overnight culture was prepared using LB medium (2.3.1) containing required antibiotics (2.3.3). Fresh LB medium was inoculated the following morning with overnight culture in a 1:100 dilution and incubated at 37 °C (shaking: 180 r.p.m.) until OD_{600} reached 0.4 - 0.6. Cells were then harvested by centrifugation (Sorvall RC 5B plus centrifuge, SS34 rotor) at 4000 r.p.m. for 5 min. The cell pellet was resuspended in about half culture volume ice-cold sterile 50 mM $CaCl_2$ and incubated on ice for 30 min subsequently. After incubation, cells were spun down (Sorvall RC 5B plus centrifuge, SS34 rotor) at 1500 r.p.m. for 10 min (4 ˚C) and the cell pellet resuspended in 1/50th of the original culture volume ice-cold sterile 50 mM $CaCl_2$, followed by additional incubation on ice for at least 30 min (ideally 2 h). 200 µl of the cell suspension were then dispensed into pre-cooled Eppendorf tubes and transformed (2.6.2) with appropriate vector.

Cells will remain competent for up to 24 h at 4˚C. Transformation efficiency increases four- to six-fold during this time. Alternatively, it is possible to use 50 mM cold $CaCl_2$ solution (50 mM $CaCl_2$, 15% glycerol, 10 mM piperazine-N,N'-bis-2-hydroxypropane-sulfonic acid, pH 7) instead of only 50 mM $CaCl_2$ for long-term storage at -70 ˚C.

2.6.2 Plasmid transformation into BL21 (DE3) and Rosetta2 (DE3) competent *E. coli* cells

Rosetta2 (DE3) cells (2.2.1) were used in this project because they are designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*.

25 µl competent cells were defrosted on ice for 10 – 15 min before 1 µl of the plasmid was added. Incubation of the mixture for another 30 min on ice was followed by a 45 s heat shock at 42 °C. 1 ml LB medium (2.3.1) was added afterwards and cells were incubated shaking for 60 min at 37 °C. Cells were then centrifuged (Sigma 2k15 centrifuge) at 4000 r.p.m. for 5 min. The cell pellet was resuspended in 150 µl LB medium and spread on LB agar (2.3.1) containing appropriate antibiotics (2.3.3). Incubation of the plates took place overnight at 37 ˚C.

2.6.3 Plasmid transformation into NovaBlue competent *E. coli* cells

Due to their high-efficiency transformation, NovaBlue competent cells (2.2.1) were only used for cloning.

25 µl competent cells were defrosted on ice for 10 – 15 min before 1 µl of the plasmid was added. The mixture was incubated on ice for 5 min followed by a 30 s heat shock at 42 °C and an additional 2 min incubation on ice. 150 µl SOC medium (2.3.1) was added to the reaction mixture and cells were incubated shaking for 60 min at 37 °C. Cells were then spread on LB agar (2.3.1) containing appropriate antibiotics (2.3.3). Incubation of the plates took place overnight at 37 ˚C.

Several colonies were picked for an initial screen of positive transformants by control digest with *Xba*I and *Nhe*I or *Bam*HI and *Kpn*I (2.5.1) prior to the final confirmation by sequencing (GATC).

2.6.4 Plasmid transformation into *Saccharomyces cerevisiae* WAT11

The transformation was carried out using a modified protocol developed in 1989 by Schiestl & Gietz.[230]

20 ml overnight cultures were prepared in YPGA medium (2.3.2) using a single yeast colony as inoculum and incubated shaking at 200 r.p.m. and 30 °C. The OD_{600} of the overnight culture was measured the following morning in a 1:100 dilution to calculate the appropriate volume of overnight culture necessary to inoculate 50 ml YPGA main culture

to an OD_{600} of 0.15. The main culture was then incubated for 5 h at 200 r.p.m. and 30°C. Carrier DNA and transformation mix were prepared as described in Table 2.10 30 min prior to the end of the incubation period. Cells were harvested by centrifugation (Sorvall RC 5B plus centrifuge, SS34 rotor) at 4500 r.p.m. for 5 min. The cell pellet was washed in 25 ml of cold sterile water and spun down subsequently at 4500 r.p.m. (Sorvall RC 5B plus centrifuge, SS34 rotor) for 5 min at 4 ˚C. Cells were resuspended in 1ml cold sterile water and 100 µl aliquots of the cell suspension were prepared in pre-cooled, sterile 1.5 ml Eppendorf tubes. Eppendorf tubes containing the cell suspension were centrifuged (Sigma 2k15 centrifuge) at high speed for 30 s and the supernatant discarded. Prepared transformation mix (Table 2.10) was added to the cell pellet and cells were transformed by heat shock at 42°C for 30 min. The reaction mix was placed on ice for at least 2 min subsequently to the heat shock followed by centrifugation (Sigma 2k15 centrifuge) at high speed for 30 s to remove the transformation mix. The cell pellet was washed once in 1 ml sterile cold water and was subsequently resuspended in 1ml sterile cold water. 1/20th of the cell suspension was then spread on appropriate selective plate (SGI, 2.3.2) and incubated at 30 °C for 3 to 4 days.

Table 2.10: Yeast transformation mix suitable for 1 transformation

Reaction Component	Quantity
PEG 3500 50% w/v	240 µl
Li Acetate 1M	36 µl
Single stranded carrier DNA (2 mg/ml)*	50 µl
Plasmid DNA	250 ng
Deionized water	to 360 µl

* Boil prior use for 20 minutes and chill on ice

2.7 Working with Proteins

2.7.1 Cell growth and protein expression

E. coli BL21 (DE3) (2.2.1), which has been made competent using 50 mM $CaCl_2$ (2.6.1) and transformed with pGro7 plasmid (Table 2.3) or *E. coli* Rosetta 2 (DE3) (2.2.1), were used for transformation with the plasmid containing the gene of interest.

2.7.1.1 Small scale expression test

Starter cultures of positive transformants were grown overnight at 37 °C and 180 r.p.m. in 5 ml LB (2.3.1) containing the required supplements (2.3.3). M9 minimal medium (2.3.1) was used as growth medium (unless stated otherwise) for 10 ml main cultures which have been inoculated with 0.5 ml starter culture and incubated shaking at 37 °C until an OD_{600} = 0.6 - 0.9 was reached. Gene expression was then induced by adding 1 mM IPTG, 0.5 mM $FeCl_3$ and 0.5 mM ALA (δ-aminolevulinic acid, a heme ring precursor) (2.3.3). The proteins were expressed with shaking (180 r.p.m.) at 16 °C overnight (unless stated otherwise). Cells were harvested by centrifugation (Sorvall RC 5B plus centrifuge, SS34 rotor) at 5000 r.p.m. for 15 min. Supernatant was discarded and the cell pellet resuspended in 600 µl of buffer A (50 mM Tris/HCl pH 7.5, 300 mM NaCl). The resuspension was then sonicated and centrifuged at 13000 r.p.m. (Sigma 2k15 centrifuge) for 30 min to separate insoluble (pellet after centrifugation; resuspended in buffer A) and soluble fraction (supernatant).

2.7.1.2 Scaled up protein expression

In order to perform protein expression in larger scale, one colony has been picked from plates with positive transformants to inoculate 5 ml of LB (2.3.1) containing the required antibiotics (2.3.3). This starter culture was grown for 6 h at 37 ˚C and 180 r.p.m.. 0.5 ml starter LB culture was used to inoculate 10 ml of M9 medium (2.3.1) preculture which was incubated overnight at 37 ˚C and 180 r.p.m.. 2 ml preculture was used as inoculum for a 200 ml main M9 medium culture. Induction and the harvest of cells (Sorvall RC 5B plus centrifuge, F10-6x500 rotor) was performed as described for small scale expression test (2.7.1.1). Depending on further use of cells, the pellet was resuspended in $1/10^{th}$ to $1/50^{th}$ original culture volume buffer A (50 mM Tris/HCl pH 7.5, 300 mM NaCl) and used for whole cell assays (4.1.3.2 and 6.3.2) or purification (2.7.2 and 6.3.3).

2.7.2 Protein purification

Cells were grown to a total volume of 2 l (10 x 200 ml) and harvested as described in 2.7.1.2. The resuspension was sonicated at 4 ˚C in 6 x 30 s intervals with 30 s delay between each interval. Soluble and insoluble fractions have then been separated by high speed centrifugation (~ 16000 r.p.m., Sorvall RC 5B plus centrifuge, F10-6x 500 rotor) for 30 min. The obtained cell lysate was purified by nickel affinity chromatography. A 5 ml

HisTrap FF Crude nickel column (GE Healthcare Life Sciences) was loaded with the clear supernatant. However, the column needed to be prepared prior to protein loading as follows: first, the column was rinsed with 5 column volume distilled, filtered water, subsequently with 5 column volume 0.5 M NaCl solution containing 0.1 M EDTA and 5 column volume 0.5 M NaCl solution and once more with 5 column volume distilled, filtered water to strip off the storage solution; the second step was to charge the column by rinsing it with 5 column volume 0.1 M Ni_2SO_4 solution; the equilibration of the column with 5 column volume Buffer A (50 mM Tris/HCl pH 7.5, 300 mM NaCl) marked the last preparation step.

After column loading, the protein was eluted with a gradient of imidazole (0–500 mM) over 20 column volumes using a programmed ÄKTA purifier system operated by UNICORN 5 control software. Column fractions containing anticipated protein were pooled and concentrated by centrifugation (SIGMA 3-16PK centrifuge, Swing-out rotor for 4 buckets) using Amicon Ultra-15 Centrifugal Filter Units (Millipore).

2.7.3 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE[233] is a method to separate proteins according to their molecular weight at denaturating conditions. Electrophoresis took place in Mini Vertical Electrophoresis Units (Hoefer, Inc.) using gel solutions listed in Table 2.11. Stacking gels constantly contained 4 % acrylamide [w/v] while the amount of acrylamide in the resolving gels varied between 12 % and 10 %. Samples were mixed with 2x SDS loading buffer in a 1:1 ratio and denaturated at around 100 ˚C for 5 min before applied to the gel. Electrophoresis usually took place at a constant voltage of 200 V and room temperature using 1x running buffer, until running front reached the end of the gel. Size estimation of samples was carried out by applying a broad or low range marker (BIO-RAD) which contains proteins of defined molecular weight (Table 2.12).

After separation, the gel was carefully transferred to a plastic container filled with Coomassie stain solution and heated up in a microwave for 70 s. Subsequently, the Coomassie stain solution has been removed and filtered for reuse. The stained gel was briefly rinsed with water to remove residual staining solution and incubated with destain solution to visualize protein bands.

Table 2.11: SDS-PAGE gel solutions

	Stacking gel 4%	Resolving gel 10%	Resolving gel 12%
H_2O	3.2 ml	4 ml	3.2 ml
4x resolving buffer	-	2.5 ml	2.5 ml
4x stacking buffer	1.3 ml	-	-
Acrylamide	0.5 ml	3.3 ml	4.2 ml
APS (10 % [w/v])	25 µl	50 µl	50 µl
TEMED	8 µl	8 µl	8 µl

Table 2.12: BIO-RAD Protein Molecular Weights in dalton

Protein	Molecular weight	Broad range	Low range
Myosin	200,000	X	
ß-galactosidase	116,250	X	
Phosphorylase b	97,400	X	X
Serum albumin	66,200	X	X
Ovalbumin	45,000	X	X
Carbonic anhydrase	31,000	X	X
Trypsin inhibitor	21,500	X	X
Lysozyme	14,400	X	X
Aprotinin	6,500	X	

4x Resolving buffer

Tris-HCl (pH 8,8) 1,5 M

SDS (w/v) 0,4 %

4x stacking buffer

Tris-HCl (pH 6,8) 0,5 M

SDS (w/v) 0,4 %

1X Running Buffer

4X SDS-PAGE Running Buffer 250 ml

- Tris Base 12 g
- Glycine 58 g
- H_2O fill up to 1000 ml

H_2O 750 ml

10% SDS 10 ml

2x SDS loading buffer

Glycerol		1 ml
β-mercaptoethanol		0.5 ml
10 % SDS		2 ml
0.1 % bromphenol blue		0.5 ml
0.5 M Tris-HCl (pH 6.8)		1.2 ml
H_2O	fill up to	10 ml

Coomassie stain solution

Propan-2-ol		250 ml
Glacial acetic acid		100 ml
Coomassie brilliant blue R		2 g
H_2O	fill up to	1000 ml

Destain solution

Propan-2-ol		50 ml
Glacial acetic acid		70 ml
H_2O	fill up to	1000 ml

2.7.4 Western blot

Western blots[234] can be used to transfer proteins, prior separated by SDS-PAGE (2.7.1), onto a nitrocellulose membrane by electroelution (Figure 2.6). The probing of proteins takes place in 1 instead of 2 steps using antibodies, which recognize the protein of interest (specific to the His-tag) and contain a detectable label. A horseradish peroxidase linked to the antibody serves as label. It has the ability to cleave a chemiluminescent agent producing a luminescent reaction product that can be detected in proportion to the amount of protein.

2.7.4.1 Western blot set up and protein transfer

A Immobilon™ PVDF membrane (0.45 µm, Millipore) and Whatman® 3MM paper (Sigma) was used for the set up. The transfer took place in a Trans-Blot® SD semi-dry transfer cell (BIO-RAD) and constant voltage, power, and current was provided by a PowerPac 1000 Power Supply (BIO-RAD).

Nitrocellulose membrane and Whatman paper needed to be equilibrated prior to the final Western blot set up. The membrane was incubated in 100 % methanol for 15 s, then transferred into Milli-Q water for 2 min and equilibrated in transfer buffer for additional 10 min while the Whatman paper has only been soaked in transfer buffer. The Western blot was set up in the following order: 4 layers of Whatman paper followed by nitrocellulose

membrane, polyacrylamide gel and another 4 layers of Whatman paper subsequently (Figure 2.6). After residual fluid was removed to provide semidry transfer condition, the transfer cell was close and connected to PowerPac 1000.

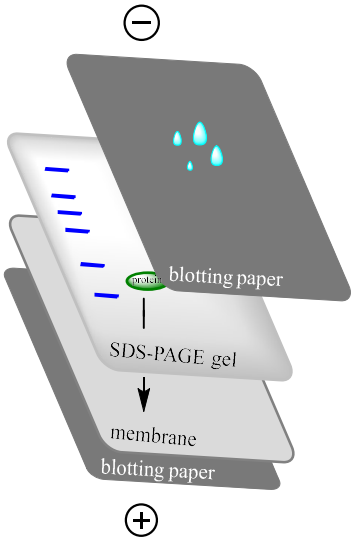


Figure 2.6: Western blot set up
SDS gel containing separated protein samples is placed against a membrane, and current is passed across the gel to the membrane, transferring the proteins onto the membrane.

Protein transfer took place by applying a constant voltage of 25 V and maximum current of 500 mA for 55 min. The membrane was then removed from the transfer apparatus and rinsed in 1x TBST to remove loose acrylamide. Ponceau Red was added to stain the membrane for 5 min and visualize proteins. Marker bands have then been marked with a pencil.

2.7.4.2 Immunoprecipitation

Non-specific binding sites were blocked by immersing the membrane in 1x TBST buffer containing 5% non-fat dried milk for 1 h at room temperature on an orbital shaker. The membrane was then briefly rinsed with three changes of 1x TBST, followed by 60 min incubation with the antibody (Monoclonal Anti-polyHistidine Peroxidase Conjugate, 1:2000 dilution; Sigma) diluted in TBST plus 5% milk at room temperature. Subsequently to the antibody treatment, 3 washing steps at room temperature for at least 10 min each were required using fresh changes of 1x TBST buffer. Chemifluorescent detection took place by using ECL Plus Western blotting Detection Reagents (GE Healthcare UK Ltd.) according to the manufactures instruction. Figure 2.7 gives an overview of the steps necessary to develop proteins after transfer to a membrane A G:Box Syngene system and GeneSnap software was used for the documentation of the Western blot results.

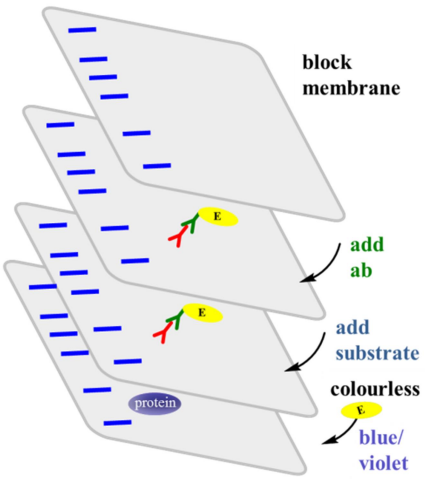


Figure 2.7: Immunoprecipitation after Western blot
The transferred protein is detected using specific primary and secondary enzyme labeled antibody. Antibodies bind to specific sequences of amino acids. Substrate that reacts with the enzyme linked to the secondary anitbody is used to visualise the antibody/protein complex.

Transfer Buffer (pH 8.3)

Tris		25 mM
Glycine		192 mM
Methanol		20%
H_2O	fill up to.	1000 ml

10X TBST pH 7.5

Tris		100 mM
NaCl		1 M
Tween 20		1 %
H_2O	fill up to.	1000 ml

3. Selection of *Beauveria bassiana* P450 targets

3.1 In situ analysis of CYPome

The overall nucleotide and amino acid sequences for *in situ* analysis were downloaded from NCBI (http://www.ncbi.nlm.nih.gov/gene). Sequences were determined *via* shotgun sequencing from *Beauveria bassiana* ARSEF 2860.[208] 83 mRNA sequences annotated as cytochrome P450 could be identified for detailed review.

The 83 annotated P450 sequences (Appendix A, Table A.1) have been analysed using bioinformatic tools in order to pick suitable targets for cloning and expression. Several sequences couldn't be considered to encode for functional CYPs and were dismissed before looking at actual conserved domains. 5 genes which code for less than 400 amino acids (aa) as well as 9 genes encoding unusually long runs of amino acids in a row were therefore dismissed before the analysis started. Of the remaining 69 sequences, 5 encode for proteins that consist of more than 1000 aa and were consequently regarded separately.

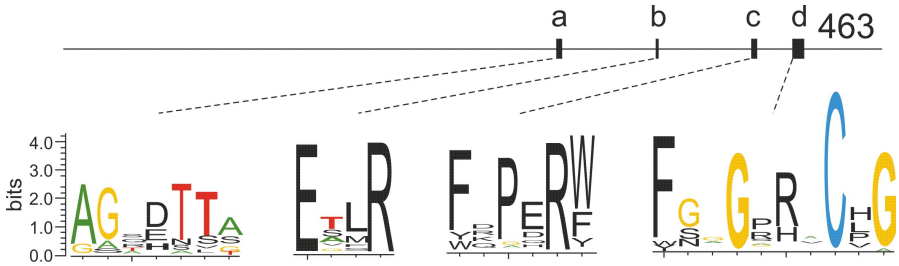


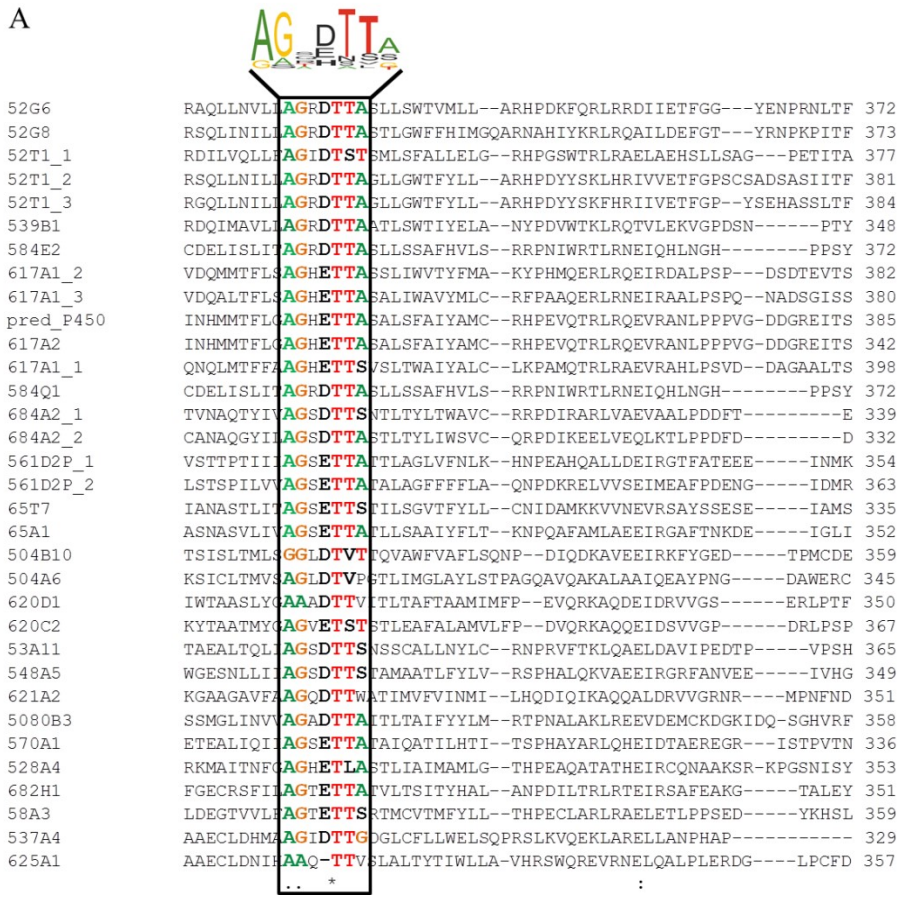
Figure 3.1: Sequence logos of the conserved CYP motifs from 47 tested fungi[177]
The four regions a, b, c, and d correspond to the positions 273–279, 330–333, 383–388, and 405-414, respectively. Bits shown on the left side specify the abundance (high bits = high abundance) of the individual aa among the tested fungi. See text below for description of conserved motifs.

Although sequence similarity shared amongst fungal CYPs is rather low, some conserved domains exist relating to key characteristics of CYPs. Figure 3.1 summarizes the characteristic motifs of 47 investigated fungal CYPs by Wanping Chen *et al.* in 2014.[177] They are fairly similar to those of animal, plant, and even archaea and bacteria but display predominance for specific aa.[235] According to P450 crystal structure comparison of various organisms, the highest structural conversation is found in the core of the protein around the heme. The heme-binding domain FXXGXXXCXG, which is the most conserved region, contains the axial cysteine ligand to the heme (Figure 3.1 d). The second absolutely conserved motif is E-X-X-R (Figure 3.1 b) and PER (Figure 3.1 c) which form the E–R–R triad. The triad stabilizes the core structure and is important for locking the

heme pocket into position. Another consensus sequence considered as P450 signature (AGXDTT, Figure 3.1 a) can be found in the central part of the I helix and contributes to oxygen binding and activation.[50, 200-201]

In order to identify CYPs which have at least the potential of functionality, only P450s of *Beauveria bassiana* that possessed all 4 above described structural properties could be considered for further investigation. After alignment and sequence comparison of the 65 putative heme domains only 33 P450s containing these consensus sequences remained for construction of a phylogenetic tree. Figure 3.2 shows the alignment results of these 33 P450s with regard to the above mentioned consensus sequence using Clustal W version 2.0.[236] The 5 sequences with more than 1000 aa have not been included into this alignment as they interfere with the alignment quality. After investigation of the 5 sequences it, however, became evident that only one could be considered as potential P450 (CYP505A1) as the other four lack the most conserved FXXGXXXCXG region. Furthermore, it needs to be mentioned that one of the 69 analysed P450s (*i.e.,* CYP505A2) didn't contain P450 consensus sequences, but flavodoxin and a FAD binding domain. Thus, this gene probably corresponds to a cytochrome P450 reductase.

A



B

```
52G6          ANLKACTYLQRVMTEVLRLFPPLPMNARYATCDTSLPRGGGPDAESPVYVKKG----QAV 428
52G8          EGLKNLTYLQWCINETLRLYPIVPMNGRAAVKDTVLPLGGGPDGRSPILVKKG----QDI 429
52T1_1        GQLKDCVFLQNVVKETLRLYPPVPINSREAIRDTVLPTGGGADGSKPVFVPKG----TSL 433
52T1_2        ESLKACHPVQHLLSEALRLHPVVPENGRRAVRDTTLPRGGGPDGQSPVFIRKG----QDV 437
52T1_3        ESLKACSHLQNLLSEVLRLHPVVPENSRRATRNTMLPRGGGVDGNAPVYIRKG----EEV 440
539B1         EDIKGITYLTHAISETLRLYPAVPYNIRSCLQDSTLTGAPGQ---PDIACFKG----NHV 401
584E2         EQLRNLKFVKYIINETLRLYPPVFRNARKAVRDTILPTGGGPNGTSPVFVPKD----TGV 428
617A1_2       AVIDSMPYLNAVCSEVLRVHSPVAQSIRVATQDTTIQ---------GQFIPKE----TLL 429
617A1_3       ADIDSLPYLNAVCDEVLRLYGPVPATIRVNNCNTTVQ---------GQFIPRD----TMV 427
pred_P450     MEIDRLPYLAAVCSEVLRMYSPVPQTIRETVRDTTIQ---------GQPLPRG----TRL 432
617A2         MEIDRLPYLAAVCSEVLRMYSPVPQTIRETVRDTTIQ---------GQPLPRG----TRL 389
617A1_1       VDFDRLPYLNAVVNETLRRYPSVPVTGRVAMQDTSVL---------GVPVARG----TSV 445
584Q1         EQLRNLKFVKYIINETLRLYPPVFRNARKAVRDTILPTGGGPNGTSPVFVPKD----TGV 428
684A2_1       ADTRDLRYLEQVIAETMRLHAAAPCSLARVVPPEG-VCLSG------YYIEGG----TTV 388
684A2_2       DSLRNMPILGNVIDETLRLFPAAPSGLPRETPPEG-ATIAG------YRVDGG----TVL 381
561D2P_1      SAAK-LQYLQACLHETLRVYPPVPETTPRHSPG---ALING------DYIPKG----TRL 400
561D2P_2      STAR-LEYLHAALDETLRLYPPVAVTPPRVSPG---AEIGG------YYIPKDASCPTLI 413
65T7          TTAQRLPYMSACFLEAFRLYPPVPSGLQRMVPSEGRTRVSS------YDIPPN----TKV 385
65A1          TTQG-LPYMQAVLDEALRLYPPVAGGGSPRKVAKGGVQIAG------YFVPEN----TLV 401
504B10        LDDQKCPYVVALVKESLRYYTVLRLALPRASIKDVAYGT--------ATIPKG----SIL 407
504A6         LVEEKVPYVTALVKETLRFFTVIPICLPRTNIKDIPYEN--------TVIPAG----TTF 393
620D1         ADRESLPYIDALVKEASRWWPISPMGFPHTATEDVEYRG--------MHIPKG----SLL 398
620C2         QDRARLPYTSNVVTELLRWWPVLPMGVVHTAQEDIMYRD--------YLIPKG----AIL 415
53A11         DMVRDLPYLNSVLNETLRHHSTSGIGLPRQIPPGSPGVTIRG-----HYFPAG----TVL 416
548A5         PALGGCTYLRACIDEAMRLSPSVGGILPREILAGGMTIDG-------HALPVG----TVV 398
621A2         RQHPGLRYIEYLLQETLRWCPVSPIGVPHRSIKDDIYNG--------YFIPAG----SFI 399
5080B3        QDAKNMPYLQAVITEGMRLHSVAGLPLWRDVAEGGVELGG-------QYFPEG----TTV 407
570A1         AEAQSLAYLQAVILEGLRVFPPAAALYPKVCDTDEILCG--------VSVPAG----TNV 384
528A4         ADALELVYTRAVIREAMRLYPVISMSLPRVVPLHGSGLQLSG-----LRIPPG----ATV 404
682H1         QEVKALPYLTGVINEALRTSNSISARLVRYSDVADLQYQQ-------FFLPRG----TYI 400
58A3          PALENLPYLTAVIHEGLRLTFGPISRSSRVATEQELVYG-------EYTIPVG----TPV 408
537A4         --LDQLPYLDAVVCEGLR---PIPMSLPRYVPEGGRMVDG-------HWLPEK----TVA 373
625A1         AIRKKAPVLEACLCESLRTRPLSSGRAERIVPETNLYDD--------VMIPAG----TII 405
                                *  *
```

C

```
52G6          LYNAHILH-RRTDIWGP------DAGEFNPDRWEGRKGG---------------------- 460
52G8          GYSVHVMH-HRTDLWGA------DADDFRPERWEKRKPG---------------------- 461
52T1_1        KYSPYVMH-RRSDLYGP------DAMKWKPDRWLGRSHG---------------------- 465
52T1_2        LYSVNVMH-RRKDLWGD------DAHEFRPERWADRKHG---------------------- 469
52T1_3        IYNVNVMH-RRKDIWGD------DADEFRPQRWIGSKHG---------------------- 472
539B1         IYSTYAMQ-RRRDLYPPVSETFADPDIYSPDRWDHWTPR---------------------- 439
584E2         VYSAWAMH-RRTDLYGP------DATEFNPERWATQRHG---------------------- 460
617A1_2       GLVPWATN-TDPKQWGD------DAHEFNPERWLTPENGGKNATN-----------AASG 471
617A1_3       IIPIWAIN-TDPLLWGP------DAHDFKPERWLKSHDDGTGHTY-----------TASA 469
pred_P450     LLVPWATN-LDTRFWGS------DGGEFKPERWLVARGGGDDEDDGLGGRAASVKAASVG 485
617A2         LLVPWATN-LDTRFWGP------DGGEFKPERWLVARGGGDDEDDGLGGTAASVKAASVG 442
617A1_1       TIPQWAIN-VDKTLWGD------DAEDFNPDRWLDARTSDDGGVT----------LNNTG 488
584Q1         VYSAWAMH-RRTDLYGP------DATEFNPERWATQRHG---------------------- 460
684A2_1       GCQGYTMH-RDRDIFP-------EPDEFKPERWENPTR----------------------- 418
684A2_2       CAQPYSMH-RDPAVFP-------NPEEFDPHRWQEPTK----------------------- 411
561D2P_1      SVYQWATY-RNPNNFT-------DPDIYRPQRWLPATHPLYDERY---------------- 437
561D2P_2      SVHQWATF-RNPANFH-------LPEQFLPERWLPASHPRYEARF---------------- 450
65T7          SVHPLAAY-NDARNWH-------KPEIFLPERWLPEAKSDPSSLF---------------- 422
65A1          ENDMWALH-HNPKYFA-------QPESFVPERWLGHKD------F---------------- 432
504B10        FLNAWACN-MDSNVWS-------DPDYFRPERW--YEQP---------------------- 436
504A6         FMNAYAAD-YDDRRFK-------MPEQFIVERF--LNDH---------------------- 422
620D1         LPAVWWFL-HDPDVYA-------DPEEFDPDRFLTPRS----------------------- 428
620C2         MPMVWWFC-HDPGSYA-------DPAEFDPDRFSKAGSN---------------------- 446
53A11         SVPTYTMH-HSKDIWGP------DAEEFRPERWETLTQR---------------------- 448
548A5         GVPHYTVH-HNEDYFP-------DPYSYVPERWLPGAKSSALGRETS----------QDD 440
621A2         YANARAMT-HDPDIYL-------DPDEFDPDRYIPEKEG---------------------- 430
5080B3        GVNGWCAH-YNESVFG------EDAHEFRPERWLVKDEDK--------------------- 440
570A1         AWSPWTIM-RNADIFGA------DADDFRPERWLGAGKT---------------------- 416
528A4         GCNPVSLH-RNEDICGP------DPDEFDPRRWLAGSGTPDN------------------ 439
682H1         SVSMDDLH-MNSVIFPE-------PEEFNPDRWLNQADYKR------------------- 433
58A3          SQSTYFMH-SNESIFP-------DYKEFDPERFIKAAKEG--------------------- 440
537A4         SCQAFSVHRINDDMFP-------EPDAFQPERWLQLDGDAD------------------- 407
625A1         STSTIAIQ-HNPHVFQN-------PHEYDPGRWLRADKDE--------------------- 437
                                        :   *:
```

D

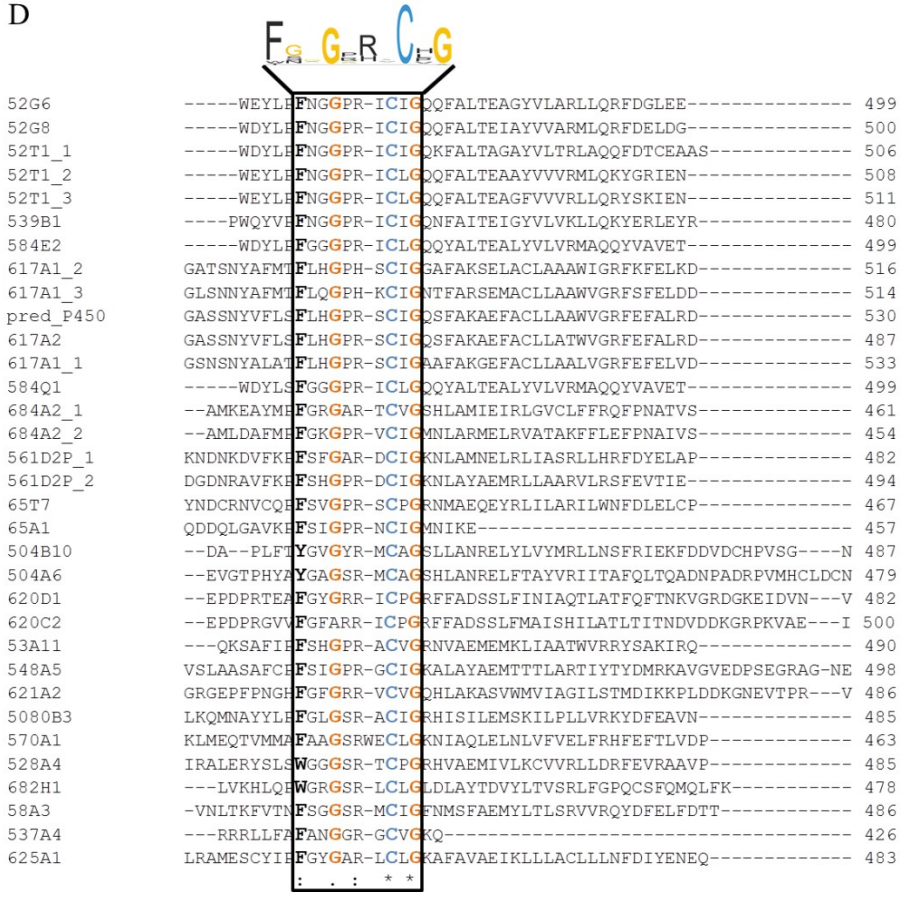


Figure 3.2: Alignment of 33 P450s with relevant CYP-numbers from *B. bassiana* in accordance with their conserved motifs

The AGXDTT motif (**A**) contributes to oxygen binding and activation; EXXR motif (**B**) and PER motif (**C**) form the E–R–R triad which stabilizes the core structure; the heme binding motif FxxGxxxCxG (**D**) contains the conserved cysteine residue that ligates to the Fe of the heme cofactor; numbers on the right side indicate amino acid position; CYP-numbers are located left of the alignment

CYP505A1 showed high identity to a catalytically self-sufficient fatty acid hydroxylase from *Fusarium oxysporum* ($P450_{foxy}$) which is known as the eukaryotic counterpart of $P450_{BM3}$ from *Bacillus megaterium*. In addition to all conserved motifs (shown in Figure 3.3) contains CYP505A1 flavodoxin as well as a FAD binding domain which are typical for CYP reductases.

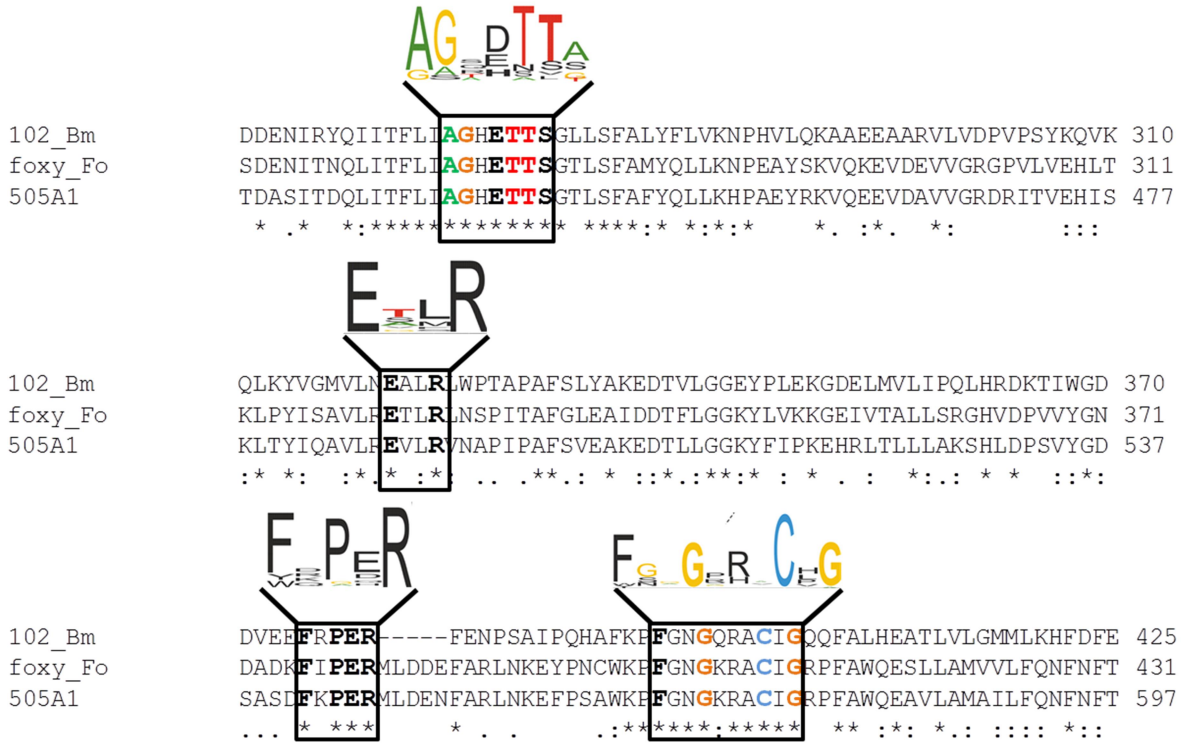


Figure 3.3: Alignment of self-sufficient P450s from *B. megaterium*, *F. oxysporum* and *B. bassiana* with regard to conserved motifs
102_Bm: $P450_{BM3}$ from *B. megaterium*; foxy_Fo: $P450_{foxy}$ from *F.* oxysporum; 505A1: CYP505A1 from *B. bassiana*; conserved motifs are in accordance with motifs shown in Figure 3.2; numbers on the right side indicate amino acid position.

3.2 Phylogenetic tree

In order to generate the phylogenetic tree, sequences were used as input into the PhylOgenetic WEb Repeater (POWER) which performs MSA in ClustalW[236] and builds a dendrogram based on the PHYLIP package.[237] The tree was constructed by using the neighbour joining method with an associated bootstrap in which 1000 datasets were generated by SEQBOOT. The final tree was drawn in Drawtree 3.66[238] and is displayed in Figure 3.4.

The 33 P450s are clustered in 5 groups. It was possible to classify these groups by potential catalysis function using *Basic Local Alignment Search Tool* (BLAST) and comparing them to P450s of known or assumed function. For example: the majority of heme domains shown with a yellow background have similarity to P450s with benzoate-4-monooxygenase activity while clustered heme domains with a green background have more in common with phenylacetate hydroxylases. P450s with a purple background, however, couldn't be classified as no known or described P450 showed any similarity to this group when using BLAST analysis. As expected, CYP505A1 is clustered with the

self-sufficient P450 from *B. megaterium* (102 BM) and *F. oxysporum* (foxy Fo), although it seems to be phylogenetically closer related to $P450_{foxy}$.

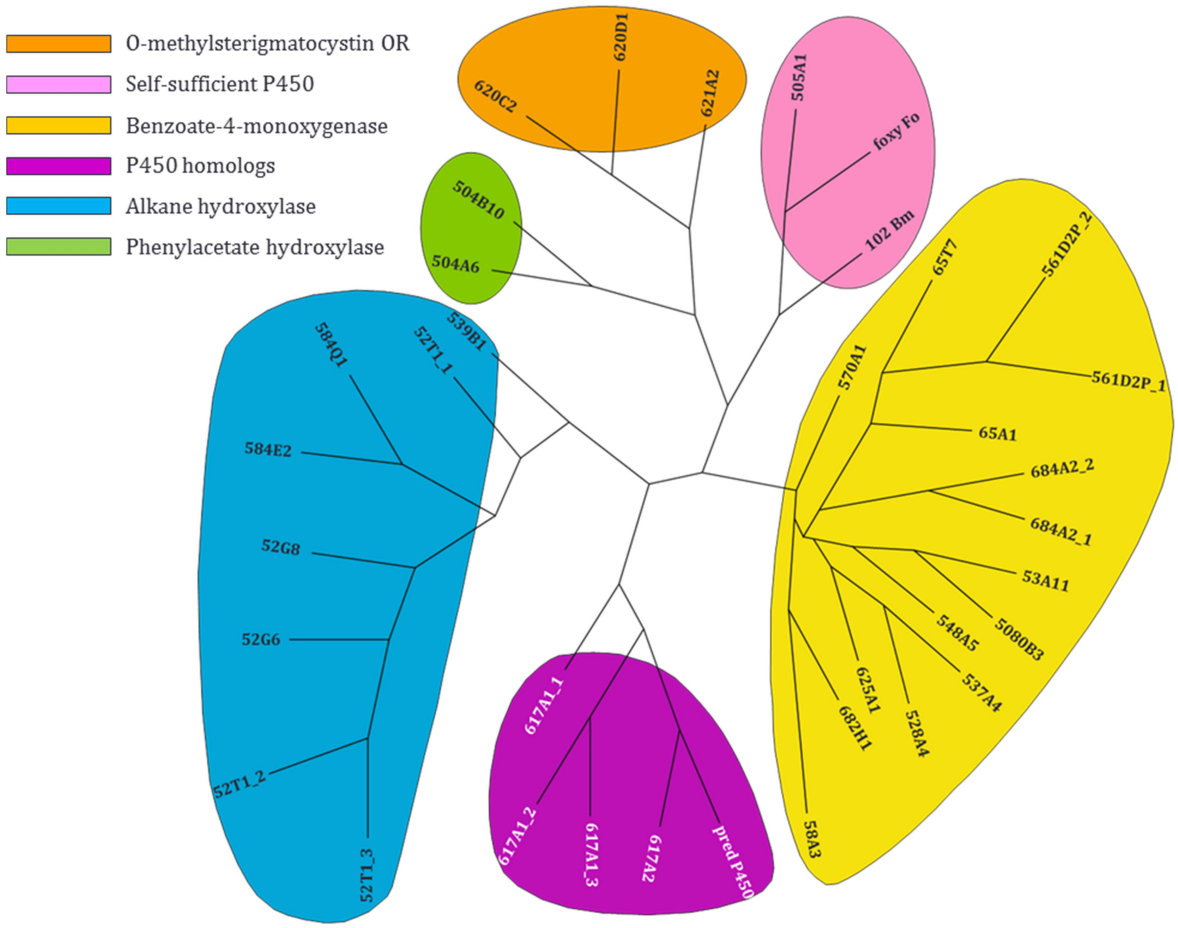


Figure 3.4: Phylogenetic tree of 33 P450 heme domains and 1 putative natural fusion gene from *B. bassiana* in relation to 2 well-known self-sufficient P450s.
The gene encoding the presumed fusion protein CYP505A1 (displayed as '505A1') from *B. bassiana* is clustered with genes coding for the self-sufficient $P450_{BM3}$ from *B. megaterium* (displayed as '102 BM') and the self-sufficient $P450_{foxy}$ from *F. oxysporum* (displayed as 'foxy Fo'). *B. bassiana* P450 heme domains (presented by CYP numbers, compare Appendix, Table A.1) are clustered in 5 groups. Comparison of individual *B. bassiana* P450s with P450s of known or assumed function indicates a correlation between gene clusters and subsequent function of their respective proteins with regard to biocatalysis. The background colour of clustered P450s correlates to the potential catalysis function displayed in the image legend (upper left corner).

3.3 Discussion

Despite considerable variation in sequence, P450 enzymes maintain a conserved P450 fold perpetuated by very few invariant residues known as the P450 motifs (Figure 3.1). As the availability of sequences increases, exceptions of these motifs become, however, more apparent. The conserved Thr in the AGXDTT motif, for example, was long considered to be invariant owing to mutagenesis studies which resulted in loss of activity when it was

substituted.[239] P450s such as CYP107A1 ($P450_{eryF}$)[240] and CYP176A ($P450_{cin}$)[241] proved this belief to be wrong as they lack this conserved Thr and are still fully functional. The same applies to the EXXR motif, which is missing in the CYP157C1 of *Streptomyces* sp but still yields a correctly folded and spectrally normal P450.[242] Moreover, there are examples of orphan P450s in which the haem axial ligand Cys - allegedly the only common feature of P450s – is replaced.[200] These exceptions are often a clue to a peculiarity of P450 catalysis. Especially modifications in the heme-binding domain indicate that oxygen is not required for catalytic activity and thus indicate novel catalytic activities.[177] This project is, however, focused on the identification and exploitation of cytochrome P450-based biocatalysts, thus P450s from *Beauveria bassiana* with biocatalytic activity. Since there is only little knowledge about the functionality of individual P450s in *B. bassiana,* as well as the fact that no more than eight of these 83 P450s were to be selected for investigation, it appeared to be more reasonable to consider P450s with typical CYP motifs, as they most likely correspond to catalytically active proteins. On this basis, 33 out of the original 83 P450s remained for further investigation. Dismissed genes with insufficient conformation in the CYP motif, however, may have still catalytic activity as affirmed by the above mentioned examples. A noteworthy exception is the protein annotated as CYP505A2, which completely lacks the CYP motifs. It contains, like the potential *B. bassiana* natural fusion P450, CYP505A1, a flavodoxin and a FAD binding domain instead and acts presumably as natural reductase partner for *B. bassiana* P450s.

The nomenclature system for CYPs is based upon amino acid identity (40% identity and above place a CYP in the same family, more than 55% identity places them in the same subfamily[27]) grouping fungal CYP families currently into more than 337 CYP gene families.[201] The 33 remaining *B. bassiana* heme domains (plus CYP505A1) are distributed into 20 families. Table 3.1 gives an overview of the distribution of these 33 P450s among the 20 different families with regard to the phylogenetic clustering (Figure 3.4) and corresponding similarity towards P450s of known or assumed function. In accordance to a distribution study of 47 fungal CPYomes from 4 different phyla, Ascomycota seem to have some frequently present CYP families (*e.g.*, CYP52, CYP56, CYP65, CYP68, CYP505, CYP532, CYP537, CYP539, CYP540, CYP548, CYP578, CYP584, CYP617, CYP682, CYP53, and CYP504).[177] Nearly 60 % of the selected *B. bassiana* P450s belong to these families, leading to the assumption that these CYPs are

possibly biocatalytically active proteins as they are widely distributed among this specific phylum. The CYP52 family in particular appears to be very interesting as they are known to participate in the degradation of alkanes and insect epicuticle.[243]

Table 3.1: CYP family distribution of 33 analysed *B. bassiana* P450s in accordance to similarity with P450s of known or assumed function

CYP family	Number of members	Similarity to
CYP504	2	Phenylacetate hydroxylase
CYP505	1	Self-sufficient P450
CYP52	5	Alkane hydroxylase
CYP539	1	
CYP584	2	
CYP617	4	P450 homolog
CYP620	2	O-methylsterigmatocystin OR
CYP621	1	
CYP5080	1	Benzoate-4-monoxygenase
CYP5282	1	
CYP53	1	
CYP537	1	
CYP548	1	
CYP561	2	
CYP570	1	
CYP58	1	
CYP625	1	
CYP65	2	
CYP682	1	
CYP684	2	

Additionally, the 20 families seem to cluster into 5 clans (groups of CYP families that consistently cluster together on phylogenetic trees) in accordance to proposed biocatalytic functions. Similar cluster behaviour could be observed in Chen's study of the 47 fungal CPYomes in which the families CYP52, CYP539 and CYP584 as well as CYP53, CYP58, CYP65, CYP528, CYP537, CYP548, CYP561, CYP570, CYP684 and CYP5080 form regular clans and are thus most likely linked to a common ancestor gene.[177] CYP505A1 on the other hand forms a single clan with the self-sufficient fatty acid hydroxylase from *Fusarium oxysporum* ($P450_{foxy}$) and its eukaryotic counterpart of $P450_{BM3}$ from *Bacillus megaterium,* undermining the close relationship between these natural fusions.

After careful consideration, sequences coding for CYP505A1 and 7 P450s with possible alkane hydroxylase function were picked for cloning and expression in *E. coli* since they are most likely encode active proteins. This assumption is substantiated by the fact that 5 of the picked heme domains have high sequence similarity or identity to cytochrome P450

enzymes that are, according to a study from Pedrini *et al.,* implicated in insect hydrocarbon degradation in *Beauveria bassiana*.[203] Moreover, the selected P450s are expected to catalyse oxidation of hydrophobic substrates due to their similarity to alkane hydroxylases. In addition, they are, with regard to substrate screening, probably easier to analyse. The selected heme domain CYP52G8, which is identical to one of the heme domains investigated by Pedrini and co-workers (CYP52X1), has already been reported to catalyse lauric acid hydroxylation and thus facilitates activity assays.[217]

4. Construction of a library of *B. bassiana* P450 genes and their expression in *E.coli*

4.1 Materials and Methods

4.1.1 Modification of P450s in preparation of synthesis

Eukaryotic P450s are usually membrane associated through an N-terminal hydrophobic membrane anchor of around 25 to 70 amino acids. The association to the membrane is the main cause of the insolubility of these proteins when expressed in *E. coli*. The removal of the N-terminus potentially leads to an increase in solubility without reducing the enzyme activity due to the fact that the catalytic centre is situated in the cytoplasm.[244-245]

In order to increase the solubility of expressed proteins, transmembrane areas predicted by software like TMpred[246] and SOSUI [247-249] were removed in the selected heme-domains. Table 4.1 shows the calculated transmembrane regions and protein properties of the selected P450s. It should be mentioned that, in accordance with the predicted transmembrane domain for CYP505A1, the first 73 aa should have been cut off. Alignment analysis with $P450_{foxy}$, however, showed a lack of the first 165 aa in *F. oxysporum*. Therefore, the first 165 aa were cut off CYP505A1 instead of only 73aa. The prepared raw sequences of the P450s were codon optimized (listed in Appendix B) as an attempt to improve the expression in *E. coli* and were send to GeneArt® to be synthesized.

Table 4.1: Properties of selected heme-domains from *B. bassiana* for cloning and expression
amino acid (aa) length, molecular weight (MW) in Dalton (Da) as well as software predicted transmembrane areas are shown for corresponding cytochromes P450. The N-terminal beginning and C-terminal end of the calculated hydrophobic membrane anchor is specified by numbers corresponding to the aa position within the protein. The complete transmembrane region for each P450 is represented as aa sequence.

CYP name	Length (aa)	MW (Da)	Transmembrane region		
			N terminal	transmembrane region aa	C terminal
CYP52G6	528	59113.6	1	MALTAILIGLVVVTFVLR	18
			N terminal	transmembrane region aa	C terminal
CYP539B1	515	58720.4	7	NTTTVALAIPVCLLLFVIVNWLT	29
			N terminal	transmembrane region aa	C terminal
CYP52T1_1	534	58271.6	1	MALHAAYLFIAATLVAVYLTRSI	23

CYP name	Length (aa)	MW (Da)	Transmembrane region		
CYP52T1_2	536	60519.7	N terminal	transmembrane region aa	C terminal
			4	LSSSFAAVLLSAVIAAYILKVVW	26
CYP52T1_3	541	61462.9	N terminal	transmembrane region aa	C terminal
			5	PISTFLAGAAILYLARWVCIEIN	27
CYP52G8	528	59480.6	N terminal	transmembrane region aa	C terminal
			6	VISLPALLVSLTVAFILLQVIEY	28
CYP584E2	528	59929.2	Soluble Protein		
CYP505A1	1232	136465.7	N terminal	transmembrane region aa	C terminal
			51	PYAMVAMVAMVAMVAMEAMV SEA	73

4.1.2 Cloning of P450s

Truncated versions (truncation of hydrophobic N-terminal region) of the 7 codon optimised *B. bassiana* P450s were amplified using KOD hot start polymerase (2.5.5) and corresponding primers (Table 4.2) before cloning into the LICRED-vector (2.5.6). Cloning of the self-sufficient P450 (CYP505A1) in form of a truncated and a natural state version was also performed using the ligation independent method but employing the LIC-3C vector, as there is no additional reductase required for electron shuttling. The primers used for CYP505A1 amplification resulting in a PCR product incorporating the transmembrane area (CYP505A1_w_FWD) or a PCR product without N-terminal anchor (CYP505A1_wo_FWD) are listed in Table 4.2.

Table 4.2: Primer used for PCR amplification

Primer	Sequence
CYP52G6_F	CCAGGGACCAGCAATGCGTGTTCTGGAAAGCCTGCGTCATGC
CYP52G6_R	GAGGAGAAGGCGCGCACTGCTTCATGGACACGAACTTTAAC
CYP52G8_F	CCAGGGACCAGCAATGGTTCGTTTTCGTAGCAAAGCAGCACGT
CYP52_G8_R	GAGGAGAAGGCGCGTTCATCAAAGTGAACTTTCAGGGTAAC
CYP52T1_F	CCAGGGACCAGCAATGCATTGGCTGTTTGCACGTAAACTGGGT
CYP_52T1_R	GAGGAGAAGGCGCGCAGCTGTGCTGCTTCAGAAAAACGAAC
CYP52T1_2_F	CCAGGGACCAGCAATGCGTCAGGTTAGCTATCAGAGCCTGGC
CYP52T1_2_R	GAGGAGAAGGCGCGATTTGCCTGCTGCAGACGAACGGCAAC
CYP52T1_3_F	CCAGGGACCAGCAATGCGTAGCATTCAGCGTCGTAATGCACG
CYP52T1_3_R	GAGGAGAAGGCGCGGCTGCTATTACCCAGGCTCACCGGAAC
CYP539B1_F	CCAGGGACCAGCAATGGTTGCCTATAAAGTTGCAAAAAGCAGC
CYP539B1_R	GAGGAGAAGGCGCGGGCCGTTTTTTCATCGGTTTTTTCCGG
CYP584E2_F	CCAGGGACCAGCAATGGCACTGGGTCAGCTGGCACCGACCGTT

Primer	Sequence
CYP584E2_R	GAGGAGAAGGCGCGGTTTTTGGCACGGGTCAGTTTACAATT
CYP505A1_wo_F	CCAGGGACCAGCAATGGCAGAAAGCATTCCGATTCCGGAACCGC
CYP505A1_w_F	CCAGGGACCAGCAATGGATGTTGATCAGATTGCCTTTCGTTAT
CYP505A1_R	GAGGAGAAGGCGCGTCAGTCAAACACATCGGCCATAAAACG

4.1.3 Protein characterization

4.1.3.1 Spectrophotometric characterisation

Purified enzyme or concentrated cell lysate was employed for spectrophotometric investigation. Analysis has been performed at room temperature using a Varian Cary 50 BioUV/Vis Spectrophotometer (Agilent Technologies) and UV-transparent disposable cuvettes (ultra-micro, 15 mm window height, BrandTech Scientific). Absorption measurements took place between the wavelength 200 and 600 nm with one point for each nanometre.

4.1.3.2 Whole cell activity assay

Cells were grown as described in 2.7.1.2 and resuspended in a relevant volume of buffer A (50 mM Tris/HCl pH 7.5, 300 mM NaCl) to yield a suspension of $100g_{cww}$ /l. Assays have been performed in a total volume of 6 ml containing no more than 0.2 mg/ml of substrate. Incubation took place in a 50 ml Erlenmeyer flask at 37 °C (shaking at 180 r.p.m.). 300 µl samples were taken at time point 0 and after 1, 2, 4 and 22 h, and were prepared for GC analysis (4.1.3.3). Cells containing empty LIC-3C or LICRED vector have been used as negative control.

4.1.3.3 Gas chromatography

300 µl samples taken from whole cell activity assays (4.1.3.2) were prepared as follows: First, transfer of sample to an Eppendorf tube filled with 500 µl volume of ethyl acetate and vortexing at maximum speed; secondly, 2 min centrifugation (Sigma 2k15 centrifuge) at room temperature and 13,000 r.p.m. to facilitate phase separation; and finally, transfer of the upper organic phase into a screw top vial (Agilent Technologies) to be analysed using GC.

When using lauric acid as substrate a second extraction round was performed followed by evaporation and derivatisation with 60 µl of N,O-bis(trimethylsilyl)trifluoroacetamide (BSTFA) and 60 µl of methyl tert-butyl ether (MTBE) for 30 min at 70 ˚C, subsequently.

Analysis of substrates was performed using a Agilent 6890N network gas chromatograph with a flame ionization detector (FID). Helium served as carrier gas and a Agilent J&W HP-5 column (0.32 mm, 0.25 µm) was used for separation. Data analysis took place using Agilent ChemStation software.

4.2 Results

4.2.1 Cloning into LICRED and LIC-3C

The *B. bassiana* $P450_{foxy}$ homolog (CYP505A1) was cloned into LIC-3-C vector which lacks the additional reductase but follows the same principles as the LICRED cloning used for the establishment of the P450 – fusion library. Figure 4.1 shows the success of PCR (P450 heme-domains: lane 5-11; CYP505A1 without transmembrane-domain: lane 3; CYP505A1 with transmembrane domain: lane 4).

After annealing of prepared vector and insert (2.5.6) and subsequent transformation into NovaBlue Single competent cells (2.2.1), several colonies were picked for an initial screen of positive transformants by control digest with X*ba*I and N*he*I (2.5.1) prior to the final confirmation using sequencing. Positive clones for all LICRED constructs are shown in Figure 4.2 and Figure 4.3. The images display 2 selected clones for each P450 fusion construct containing the empty LICRED-vector and the insert of the expected length after digest. Sequencing by GATC confirmed the success of the cloning process.

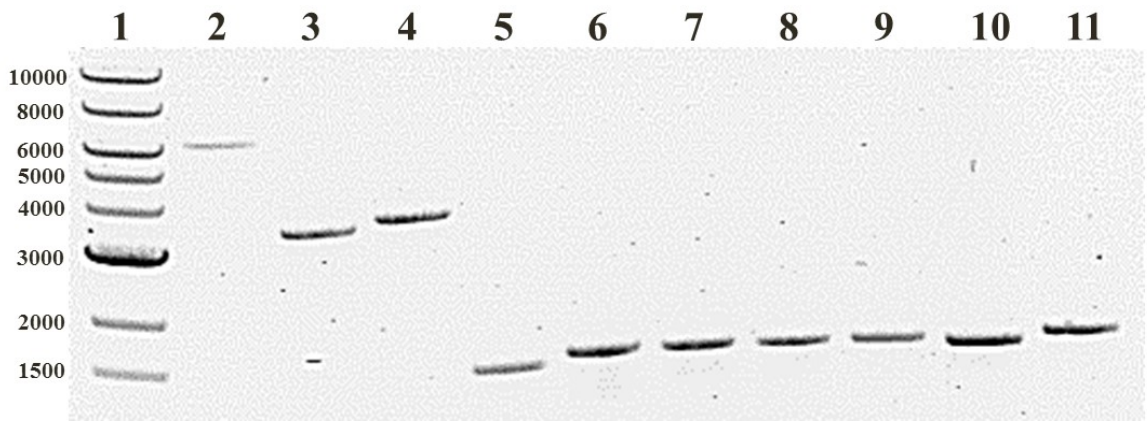


Figure 4.1: linearized LICRED vector and PCR products after gel extraction
lane 1: 1 kb DNA ladder; lane 2: linearized LICRED-vector consisting of 6386 bp; lane 3: CYP505A1_without transmembrane domain (3231 bp); lane 4: CYP505A1_with transmembrane domain (3726 bp); lane 5: CYP539B1 (1491 bp); lane 6: CYP52T1 (1575 bp); lane 7: CYPT1_3 (1575 bp); lane 8: CYPT1_2 (1535 bp); lane 9: CYP52G6 (1563 bp); lane 10: CYP52G8 (1533 bp); lane 11: CYP584E2 (1614 bp)

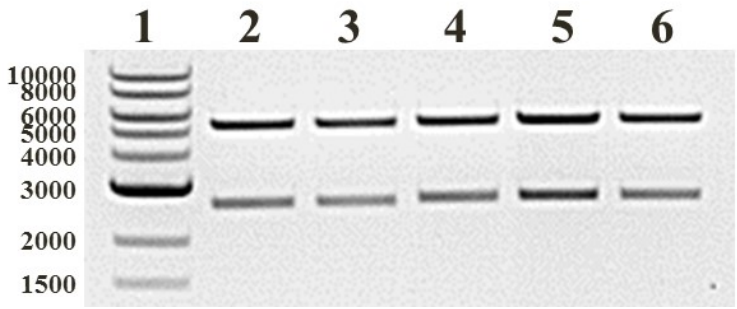


Figure 4.2: Initial screen of positive transformants using restriction enzymes *Xba*I and *Nhe*I
lane 1: 1 kb DNA ladder; lane 2: CYP52G8_p1_LICRED (2636 bp); lane 3: CYP52G8_p2_LICRED (2636 bp); lane 4: CYP52T1_3_p1_LICRED (2678 bp); lane 5: CYP52G6_p1_LICRED (2666 bp); lane 6: CYP52G6_p2_LICRED (2666 bp)

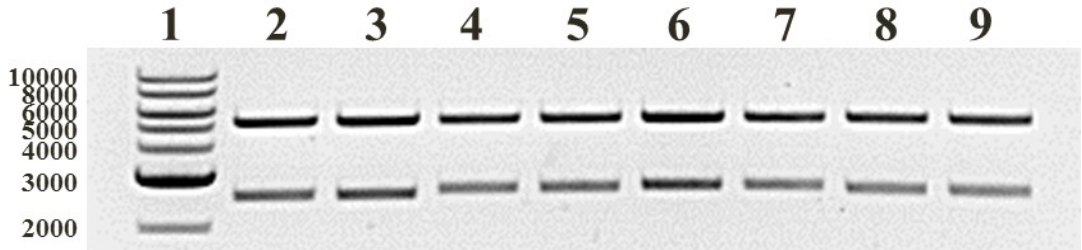


Figure 4.3: Initial screen of positive transformants using restriction enzymes XbaI and *Nhe*I
lane 1: 1 kb DNA ladder; lane 2: CYP539B1_p1_LICRED (2594 bp); lane 3: CYP539B1_p2_LICRED (2594 bp); lane 4: CYP52T1_p1_LICRED (2678 bp); lane 5: CYP52T1_p2_LICRED (2678 bp); lane 6: CYP584E2_p1_LICRED (2717 bp); lane 7: CYP584E2_p2_LICRED (2717 bp); lane 8: CYP52T1_2_p1_LICRED (2636 bp); lane 9: CYP52T1_2_p2_LICRED (2636 bp)

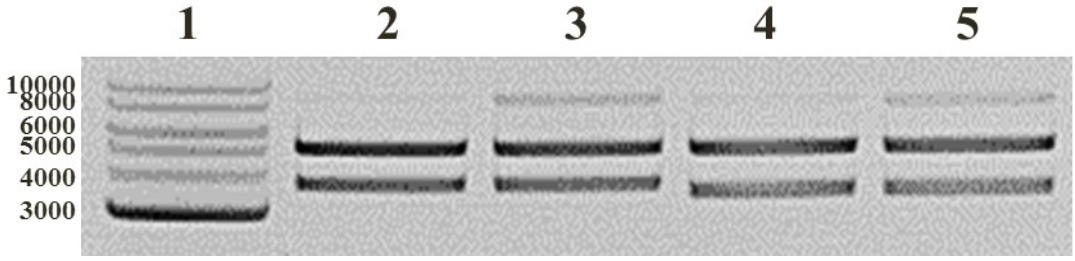


Figure 4.4: Initial screen of positive transformants using restriction enzymes *Xba*I and *Nhe*I
lane 1: 1 kb DNA ladder; lane 2: CYP505A1_w_LIC-3C_p1 (3726 bp); lane 3: CYP505A1_w_LIC-3C_p2 (3726 bp); lane 4: CYP505A1_wo_LIC-3C_p1 (3231 bp); lane 5: CYP505A1_wo_LIC-3C_p2 (3231 bp)

CYP505A1 has been cloned into the LIC-3C vector once without the transmembrane-domain (CYP505A1_wo_LIC-3C) and once comprising the whole gene sequence (CYP505A1_w_LIC-3C). Potential positive clones were picked and screened in the same manner as described for the LICRED system. Figure 4.4 illustrates the successful insertion of gene sequences with the calculated bp length into the vector which could be confirmed by sequencing data received from GATC.

4.2.2 Expression test

In order to confirm overexpression of selected P450s, small scale expression tests were performed as described in 2.7.1.1. Unfortunately, no signal was noticeable in SDS-PAGE (data not shown), neither in soluble nor insoluble fraction. Therefore, the high sensitive method of Western blotting (2.7.4) was used to verify overexpression.

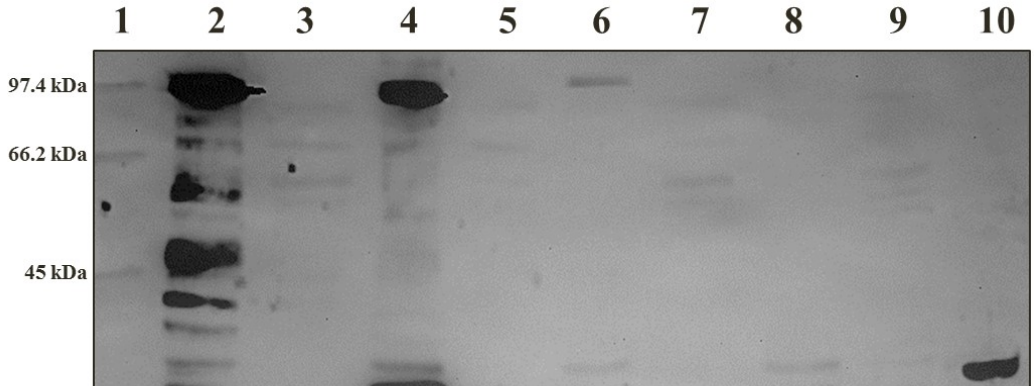


Figure 4.5: Western blot analysis of selected P450s
Insoluble fractions are marked with ins while soluble fractions are labelled sol: lane 1: Low-Range-Marker; lane 2: CYP52G8_LICRED_ins (94.5 kDa); lane 3: CYP52G8_LICRED_sol (94.5 kDa); lane 4: CYP52G6_LICRED_ins (95.4 kDa); lane 5: CYP52G6_LICRED_sol (95.4 kDa); lane 6: CYP52T1_3_LICRED_ins (94.1 kDa); lane 7: CYP52T1_3_LICRED_sol (94.1 kDa); lane 8: CYP52G8_LICRED _ins_uninduced; lane 9: CYP52G8_LICRED _sol_uninduced 10: positive control

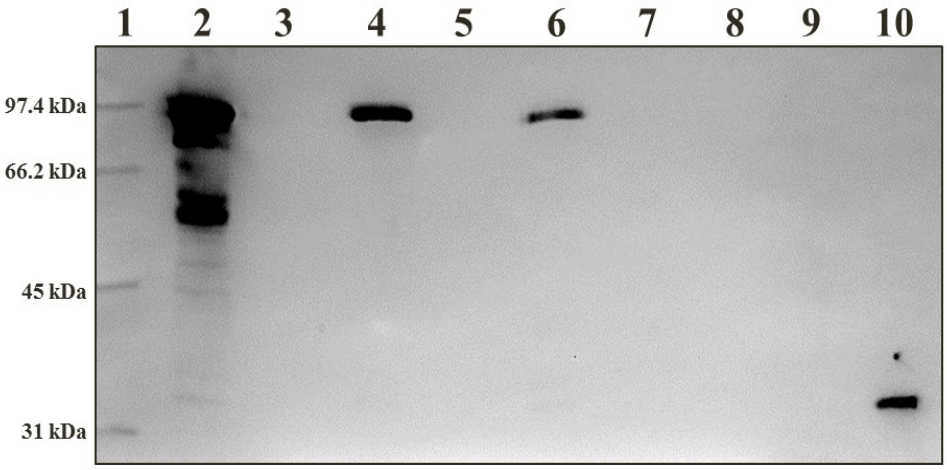


Figure 4.6: Western blot analysis of selected P450s
Insoluble fractions are marked with ins while soluble fractions are labelled sol: lane 1: Low-Range-Marker; lane 2: CYP539B1_LICRED_ins (93.5 kDa); lane 3: CYP539B1_LICRED_sol (93.5 kDa); lane 4: CYP52T1_LICRED_ins (94.2 kDa); lane 5: CYP52T1_LICRED_sol (94.2 kDa); lane 6: CYP52T1_2_LICRED_ins (95.7 kDa); lane 7: CYP52T1_2_LICRED_sol (95.7 kDa); lane 8: CYP584E2_LICRED _ins (97.8 kDa); lane 9: CYP584E2_LICRED _sol (97.8 kDa) lane 10: positive control

Clear signals of protein expression of predicted molecular weight could be detected (Figure 4.5) for CYP52G8 (lane 2) and CYP52G6 (lane 4) as well as a weak signal in case of CYP52T1_3 (lane 6), while there was no signal in uninduced cells. Unfortunately, none of these proteins seemed to be expressed solubly since bands of the expected length only occur in lanes containing insoluble samples. Furthermore, CYP52G8 possesses several bands below 90 kDa indicating that the protein gets easily degraded or is unstable. The same applies to CYP52G6 which appears to have 2 distinctive cleavage sites, resulting in to 2 additional signals that appear at about 35 kDa.

Figure 4.6 displays the outcomes of Western blotting for CYP539B1, CYP52T1, CYP52T1_2 and CYP584E2. There is no expression detectable in the case of CYP584E2 either in the soluble or in the insoluble fraction. A rather strong band corresponding to anticipated molecular weight can be seen in lane 2 containing the insoluble phase of CYP539B1 and 2 weaker signals are visible in lane 4 (insoluble fraction of CYP52T1) and 6 (insoluble fraction of CYP52T1_2). Like in CYP52G8 and CYP52G6 shown above, other bands lower than 90 kDa also appear in the insoluble phase of CYP539B1 leading to the conclusion that this protein gets easily degraded as well. Also cleavage at the linker area between P450 domain and reductase domain (16 aa linker) might contribute to the degradation process and unstability of the chimeric fusion proteins.

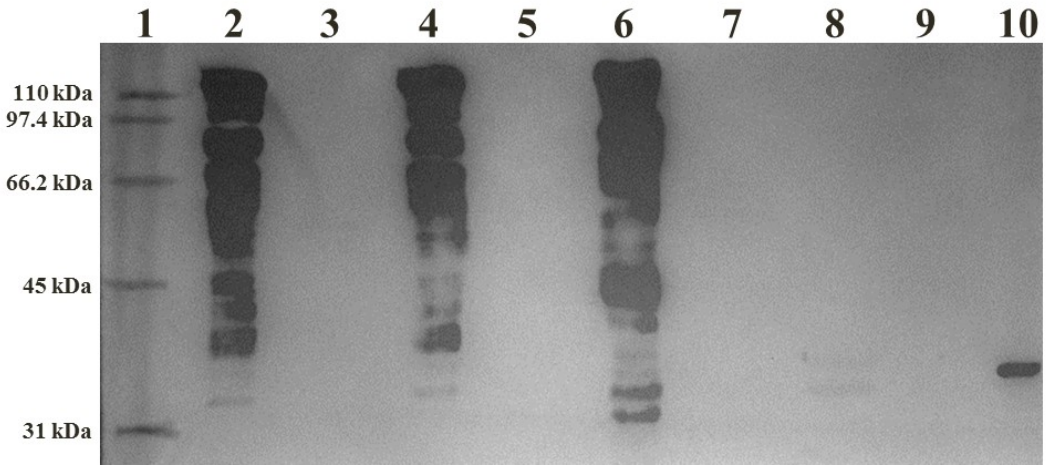


Figure 4.7: Western blot analysis of selected P450s
Insoluble fractions are marked with ins while soluble fractions are labelled sol: lane 1: Broad-Range-Marker; lane 2: CYP505A1_wo_LIC-3C_ins (120.6 kDa); lane 3: CYP505A1_wo_LIC-3C_sol (120.6 kDa); lane 4: CYP505A1_wo_LIC-3C_ins (120.6 kDa); lane 5: CYP505A1_wo_LIC-3C_sol (120.6 kDa); lane 6: CYP505A1_w_LIC-3C_ins (139 kDa); lane 7: CYP505A1_w_LIC-3C_sol (139 kDa); lane 8: CYP505A1_wo_LIC-3C_ins_uninduced; lane 9: CYP505A1_wo_LIC-3C_sol_uninduced; lane 10: positive control

Western blot analysis of *E. coli* strains expressing CYP505A1 (with and without transmembrane domain) shown in Figure 4.7 confirm overexpression. Bands above 110 kDa can clearly be seen in lanes 2 (without TM; insoluble), 4 (without TM; insoluble) and 6 (with TM; soluble) while there are no bands in the uninduced negative control displayed in lane 8 and 9. As already pointed out earlier for P450 Rhf fusion proteins, there is also no soluble expression for any of the LIC-3C constructs in *E. coli*. Furthermore, all insoluble expressed proteins appear to be particularly degraded, signified by rather strong signals below the 110 kDa mark.

Since signals observed in insoluble phases are fairly strong, exposure time while detecting the Western blot was limited and as a result, very weak signals couldn't be visualized. In order to ultimately confirm that no soluble expression has been conducted by any of the strains Western blotting comprising only the soluble phases was carried out (data not shown). Unfortunately, most signals detected are unspecific (background signals) and don't agree with the predicted molecular weight.

In summary, it can be said that all P450 fusions were expressed insolubly, except for CYP584E2, but no signals could be spotted when analysing the soluble phases. The same applied for *B. bassiana* $P450_{foxy}$ homolog cloned into LIC-3C vector.

4.2.2.1 Media optimization

The expression system used in this project is based on a pET system which contains components of the lac operon. The lac operon can be regulated in a positive and a negative way. Negative regulation is mediated by a lac repressor, which binds to the lac operator - a specific DNA sequence between the promoter and the coding regions - in the operon. This event can be prevented by adding specific inducers of the lac operon (*e.g.*, IPTG) to the expression system. IPTG binds to the lac repressor and substantially decrease its binding affinity to the lac operator. Positive regulation on the other hand requires the presence of cyclic AMP (cAMP) and cyclic AMP receptor protein, called CRP or CAP. The CAP/cAMP complex is able to stimulate transcription by binding to an element upstream of the lac promoter. Hence, the level of cAMP is a decisive factor of effective expression and is strongly influenced by the carbon source present in the medium. In the present of glucose an effect shared by a number of *E. coli* operons called catabolite repression occurs which leads to low cAMP levels and therefore poor transcription from the lac promoter.[250]

Cultures for small scale expression trails (2.7.1.1) have so far only been grown in M9 medium (2.3.1) containing 4 mg/ml glucose as sole carbon source. In order to increase expression of anticipated proteins various alternative carbon sources have been tested to raise the cAMP level and therefore transcription.

Arabinose or glycerol was tested as sole carbon source in the medium. Both have been supplemented from a 20 % (w/v) stock solution to ensure a final concentration of 4 mg/ml in the M9 medium. In addition, a combination of arabinose and glycerol or glucose and arabinose was tested. Carbon sources tested in combination had a final concentration of 2 mg/ml each in the medium.

Table 4.3 summarizes the relative expression levels seen after Western blot analysis for each protein and Figure 4.8 to Figure 4.12 show P450s after SDS gel separation. Data for arabinose expression are not shown, because cells didn't grow when used as only carbon source.

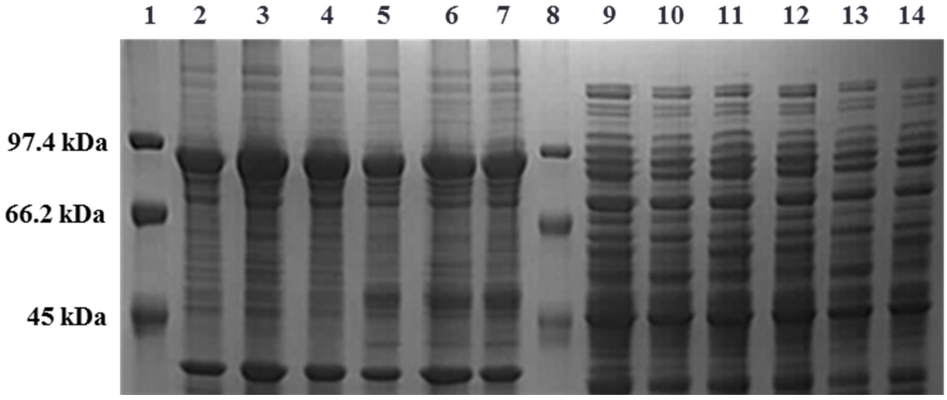


Figure 4.8: SDS gel of in *E. coli* expressed CYP52G6 and CYP52G8 grown on variable carbon sources
Glc = Glucose; Ara = Arabinose; Gly = Glycerol; **lane 1**: Low-Range-Marker; **lane 2-4**: insoluble expressed CYP52G6_LICRED (94.5 kDa) grown in medium containing Glc & Ara (lane 2), Gly (lane 3), and Gly & Ara (lane 4); **lane 5-7**: insoluble expressed CYP52G8_LICRED (94.5 kDa) grown in medium containing Glc & Ara (lane 5), Gly (lane 6), and Gly & Ara (lane 7); **lane 8**: Low-Range-Marker; **lane 9-11**: soluble expressed CYP52G6_LICRED (94.5 kDa) grown in medium containing Glc & Ara (lane 9), Gly (lane 10), and Gly & Ara (lane 11); **lane 12-14**: soluble expressed CYP52G8_LICRED (94.5 kDa) grown in medium containing Glc & Ara (lane 12), Gly (lane 13), and Gly & Ara (lane 14).

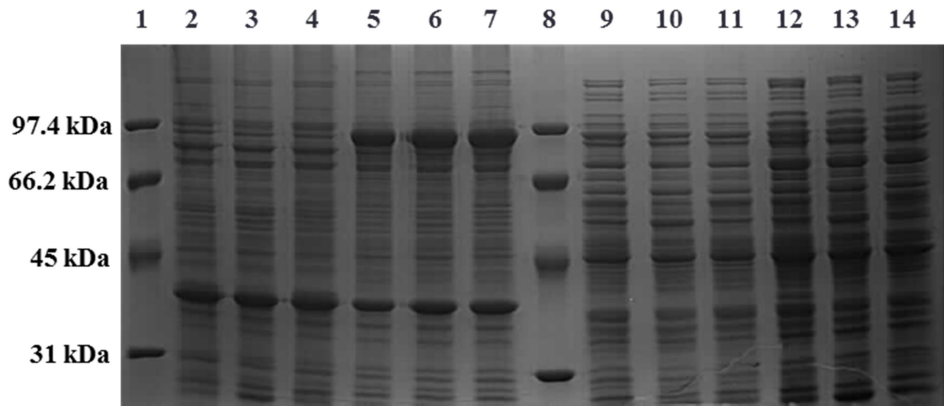


Figure 4.9: SDS gel of in *E. coli* expressed CYP584E2 and CYP52T1 grown on variable carbon sources
Glc = Glucose; Ara = Arabinose; Gly = Glycerol; **lane 1**: Low-Range-Marker; **lane 2-4**: insoluble expressed CYP584E2_LICRED (97.8 kDa) grown in medium containing Glc & Ara (lane 2), Gly (lane 3), and Gly & Ara (lane 4); **lane 5-7**: insoluble expressed CYP52T1_LICRED (94.2 kDa) grown in medium containing Glc & Ara (lane 5), Gly (lane 6), and Gly & Ara (lane 7); **lane 8**: Low-Range-Marker; **lane 9-11**: soluble expressed CYP584E2_LICRED (97.8 kDa) grown in medium containing Glc & Ara (lane 9), Gly (lane 10), and Gly & Ara (lane 11); **lane 12-14**: soluble expressed CYP52T1_LICRED (94.2 kDa) grown in medium containing Glc & Ara (lane 12), Gly (lane 13), and Gly & Ara (lane 14).

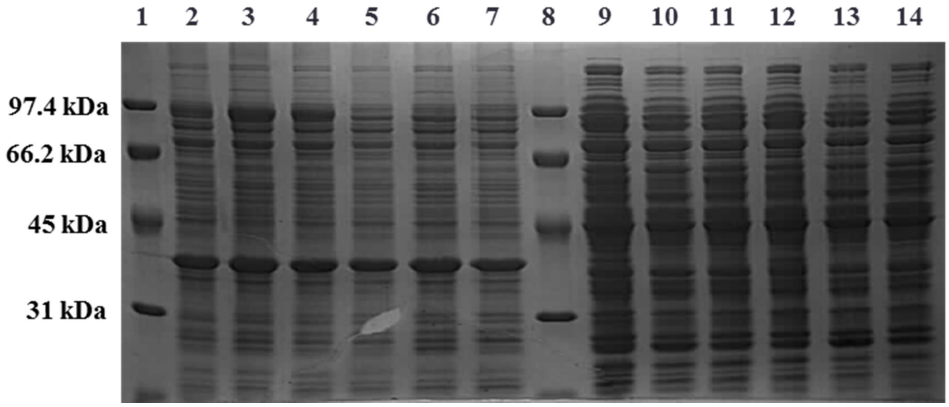


Figure 4.10: SDS gel of in *E. coli* expressed CYP52T1_2 and CYP52T1_3 grown on variable carbon sources
Glc = Glucose; Ara = Arabinose; Gly = Glycerol; **lane 1**: Low-Range-Marker; **lane 2-4**: insoluble expressed CYP52T1_3_LICRED (94.1 kDa) grown in medium containing Glc & Ara (lane 2), Gly (lane 3), and Gly & Ara (lane 4); **lane 5-7**: insoluble expressed CYP52T1_2_LICRED (95.7 kDa) grown in medium containing Glc & Ara (lane 5), Gly (lane 6), and Gly & Ara (lane 7); **lane 8**: Low-Range-Marker; **lane 9-11**: soluble expressed CYP52T1_3_LICRED_sol (94.1 kDa) grown in medium containing Glc & Ara (lane 9), Gly (lane 10), and Gly & Ara (lane 11); **lane 12-14**: soluble expressed CYP52T1_2_LICRED (95.7 kDa) grown in medium containing Glc & Ara (lane 12), Gly (lane 13), and Gly & Ara (lane 14).

Clear visible bands of proteins with predicted molecular weight could already be detected in SDS gels for insoluble fractions whether cells are grown only in glycerol or in combination with either glucose or arabinose. The only exceptions were CYP584E2 and CYP52T1_2 which only showed a moderate to no expression improvement at all.

However, none of the P450 constructs revealed visible bands in the soluble fraction when run on a SDS gel.

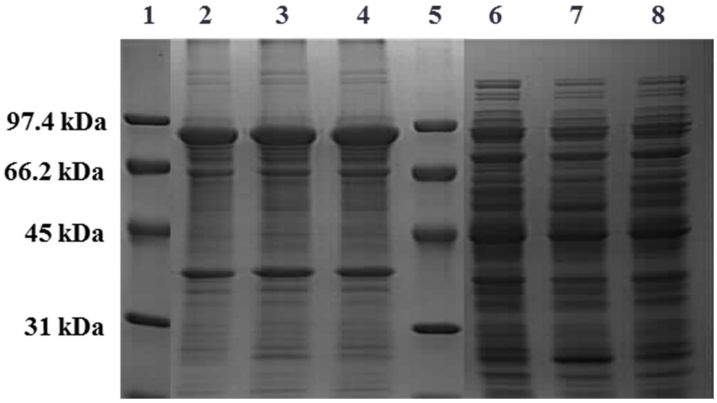


Figure 4.11: SDS gel of in *E. coli* expressed CYP539B1 grown on variable carbon sources
Glc = Glucose; Ara = Arabinose; Gly = Glycerol; **lane 1**: Low-Range-Marker; **lane 2-4**: insoluble expressed CYP539B1_LICRED (93.5 kDa) grown in medium containing Glc & Ara (lane 2), Gly (lane 3), and Gly & Ara (lane 4); **lane 5**: Low-Range-Marker; **lane 6-8**: soluble expressed CYP539B1_LICRED (93.5 kDa) grown in medium containing Glc & Ara (lane 6), Gly (lane 7), and Gly & Ara (lane 8)

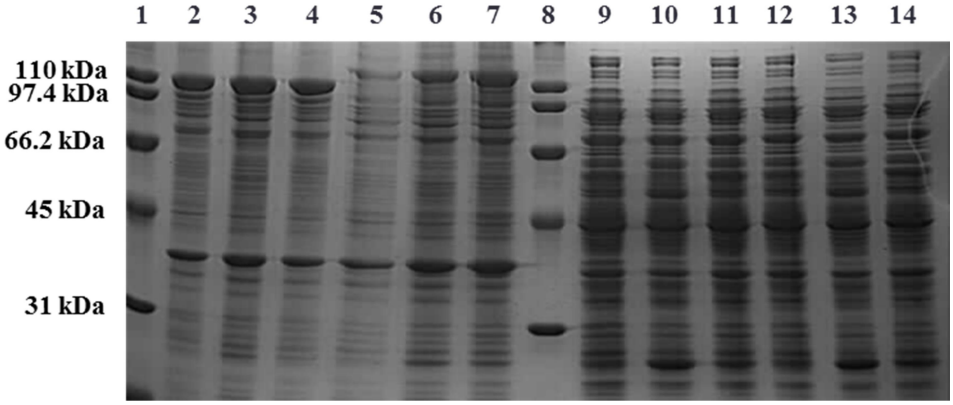


Figure 4.12: SDS gel of in *E. coli* expressed CYP505A1 without and with transmembrane domain grown on variable carbon sources
Glc = Glucose; Ara = Arabinose; Gly = Glycerol; wo = without transmembrane; w = with transmembrane **lane 1**: Broad-Range-Marker; **lane 2-4**: insoluble expressed CYP505A1_LIC-3C without transmembrane domain (120.6 kDa) grown in medium containing Glc & Ara (lane 2), Gly (lane 3), and Gly & Ara (lane 4); **lane 5-7**: insoluble expressed CYP505A1_LIC-3C with transmembrane domain (139 kDa) grown in medium containing Glc & Ara (lane 5), Gly (lane 6), and Gly & Ara (lane 7); **lane 8**: Broad-Range-Marker; **lane 9-11**: soluble expressed CYP505A1_LIC-3C without transmembrane domain (120.6 kDa) grown in medium containing Glc & Ara (lane 9), Gly (lane 10), and Gly & Ara (lane 11); **lane 12-14**: soluble expressed CYP505A1_LIC-3C with transmembrane domain (139 kDa) grown in medium containing Glc & Ara (lane 12), Gly (lane 13), and Gly & Ara (lane 14)

Table 4.3: Relative soluble and insoluble expression of Cytochrome P450s in the presence of different carbon sources at 16 ˚C.
increased expression is indicated by the symbol + whereas the intensity of the increase is linked to the rising amount of the + symbols; unvaried expression is symbolized by 0.

Cytochrome P450s	Insoluble Expression			Soluble Expression		
	Glucose & Arabinose	Glycerol	Glycerol & Arabinose	Glucose & Arabinose	Glycerol	Glycerol & Arabinose
CYP52G6_LICRED (95.4 kDa)	++	+++	++	0	0	0
CYP52G8_LICRED (94.5 kDa)	++	+++	+++	0	0	0
CYP52T1_LICRED (94.1 kDa)	++	+++	+++	0	0	0
CYP52T1_2_LICRED (95.7 kDa)	0	0	0	0	0	0
CYP52T1_3_LICRED (96.5 kDa)	+	++	++	0	0	0
CYP539B1_LICRED (93.5 kDa)	++	+++	+++	0	0	0
CYP584E2_LICRED (97.8 kDa)	0	0	0	0	0	0
CYP52A1_w_LIC-3C (139 kDa)	0	+	+	0	0	0
CYP52A1_wo_LIC-3C (120.6 kDa)	++	++	++	0	0	0

Western blot analysis confirmed the observations made by SDS PAGE analysis for insoluble and also soluble expression. No distinct signal could be detected to prove an increase in soluble expression. Nevertheless, an increase of expression in general could be evidently verified, especially if glycerol is the only carbon source applied during cell growth.

4.2.2.2 Co-expression with chaperones

Co-expression of chaperones is an approach known to increase not only insoluble but also soluble expression. In 1996 Mayhew *et al*. were able to show that co-expressing with the GroEL/ES chaperon system enhances correct protein folding and therefore protein yield.[251] For most prokaryotic CYPs as well as for CYP505A1 ($P450_{foxy}$), co-expression of GroEL/ES greatly improved the P450 concentration in the soluble protein fraction as well.[252]

A commercially available pGro7 plasmid set (2.2.3) coding groES-groEL was used to transform in BL21 (DE3) competent cells (2.2.1) following the protocol described in 2.6.2. In order for the transformed BL21 (DE3) cells to take up expression plasmid for the target protein, competency was induced by using $CaCl_2$ (2.6.1). After transformation with both plasmids, small scale expression tests (2.7.1.1) have been conducted in M9 medium with glycerol or glucose as carbon source and necessary antibiotics (2.3.1; 2.3.3). Furthermore, supplement of L-arabinose (3 mg/ml) to the medium is required to induce chaperone expression.

Figure 4.13 to Figure 4.15 display SDS gel analysis of all P450 fusions as well as CYP505A1 (Figure 4.15).

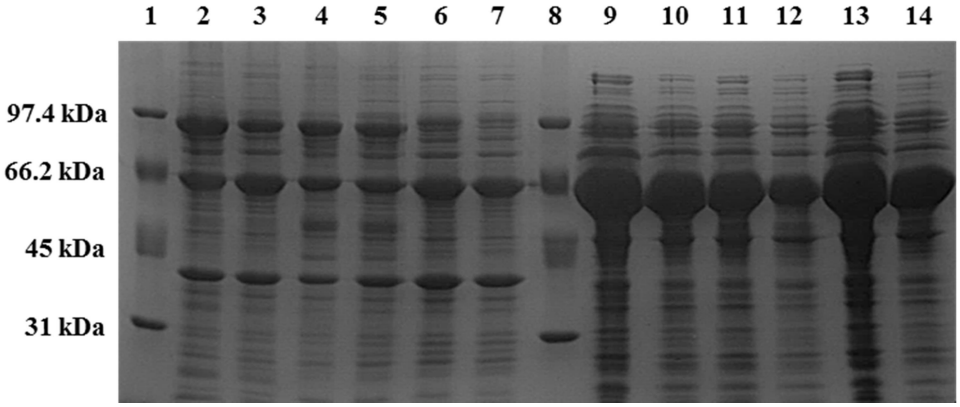


Figure 4.13: SDS gel of CYP52G6, CYP52G8 and CYP52T1_3 co-expressed with chaperones in *E. coli* grown on variable carbon sources
Glc = Glucose; Gly = Glycerol; **lane 1**: Low-Range-Marker; **lane 2-3**: insoluble expressed CYP52G6_LICRED (94.5 kDa) grown in medium containing Glc (lane 2) or Gly (lane 3); **lane 4-5**: insoluble expressed CYP52G8_LICRED (94.5 kDa) grown in medium containing Glc (lane 4) or Gly (lane 5); **lane 6-7**: insoluble expressed CYP52T1_3_LICRED (94.1 kDa) grown in medium containing Glc (lane 6) or Gly (lane 7); **lane 8**: Low-Range-Marker; **lane 9-10**: soluble expressed CYP52G6_LICRED (94.5 kDa) grown in medium containing Glc (lane 9) or Gly (lane 10); **lane 11-12**: soluble expressed CYP52G8_LICRED (94.5 kDa) grown in medium containing Glc (lane 11) or Gly (lane 12); **lane 13-14**: soluble expressed CYP52T1_3_LICRED (94.1 kDa) grown in medium containing Glc (lane 13) or Gly (lane 14).

The overexpression of chaperones is clearly visible in the soluble fractions indicated by strong bands around 57 kDa corresponding to GroEL. Bands correlating to P450s of expected molecular weight could only be detected in the insoluble fractions except for CYP584E2. Furthermore, these bands appear to be stronger for P450s that have been expressed with glucose as carbon source. However, no distinct signals could be detected in SDS gels for soluble expressed P450s. Western blot analysis where therefore conducted for soluble fractions.

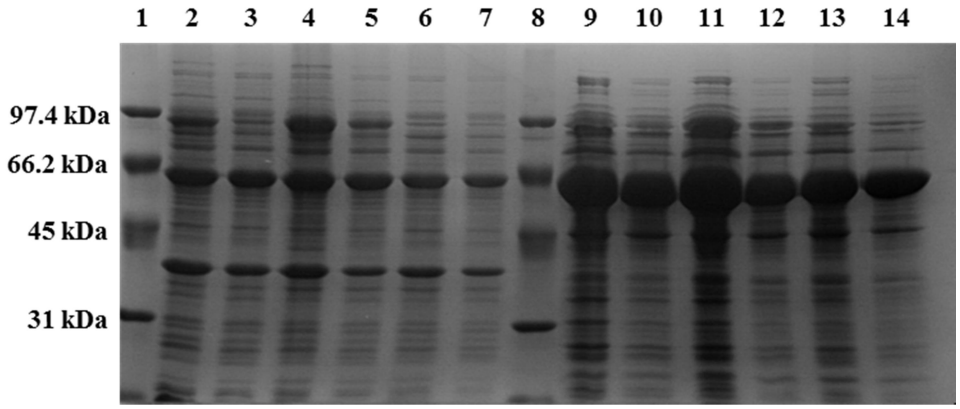


Figure 4.14: SDS gel of CYP52T1_2, CYP52T1 and CYP584E2 co-expressed with chaperones in *E. coli* grown on variable carbon sources
Glc = Glucose; Gly = Glycerol; **lane 1**: Low-Range-Marker; **lane 2-3**: insoluble expressed CYP52T1_2_LICRED (95.7 kDa) grown in medium containing Glc (lane 2) or Gly (lane 3); **lane 4-5**: insoluble expressed CYP52T1_LICRED (94.2 kDa) grown in medium containing Glc (lane 4) or Gly (lane 5); **lane 6-7**: insoluble expressed CYP584E2_LICRED (97.8 kDa) grown in medium containing Glc (lane 6) or Gly (lane 7); **lane 8**: Low-Range-Marker; **lane 9-10**: soluble expressed CYP52T1_2_LICRED (95.7 kDa) grown in medium containing Glc (lane 9) or Gly (lane 10); **lane 11-12**: soluble expressed CYP52T1_LICRED (94.2 kDa) grown in medium containing Glc (lane 11) or Gly (lane 12); **lane 13-14**: soluble expressed CYP584E2_LICRED (97.8 kDa) grown in medium containing Glc (lane 13) or Gly (lane 14).

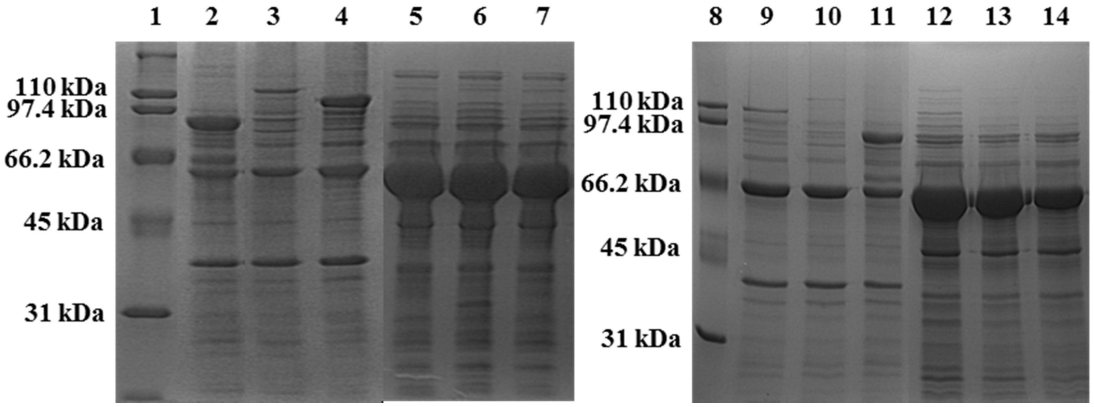


Figure 4.15: SDS gel of CYP539B1 and CYP505A1 co-expressed with chaperones in *E. coli* grown on variable carbon sources
Glc = Glucose; Gly = Glycerol; **lane 1**: Broad-Range-Marker; **lane 2-5**: insoluble expressed CYP539B1_LICRED (93.5 kDa, lane 2), CYP505A1_LIC-3C with transmembrane domain (139 kDa, lane 3), and CYP505A1_LIC-3C without transmembrane domain (120.6 kDa, lane 4) grown in medium containing Glc; **lane 5-7**: soluble expressed CYP539B1_LICRED (93.5 kDa, lane 5), CYP505A1_LIC-3C with transmembrane domain (139 kDa, lane 6), and CYP505A1_LIC-3C without transmembrane domain (120.6 kDa, lane 7) grown in medium containing Glc; **lane 8**: Broad-Range-Marker; **lane 9-11**: : insoluble expressed CYP539B1_LICRED (93.5 kDa, lane 9), CYP505A1_LIC-3C with transmembrane domain (139 kDa, lane 10), and CYP505A1_LIC-3C without transmembrane domain (120.6 kDa, lane 11) grown in medium containing Gly; **lane 12-14**: soluble expressed CYP539B1_LICRED (93.5 kDa, lane 12), CYP505A1_LIC-3C with transmembrane domain (139 kDa, lane 13), and CYP505A1_LIC-3C without transmembrane domain (120.6 kDa, lane 14) grown in medium containing Gly.

Figure 4.16 displays Western blot analysis of all P450 constructs co-expressed with GroEL/ES in M9 medium containing glucose as carbon source whereas Figure 4.17 shows the equivalent expressed in glycerol as carbon source.

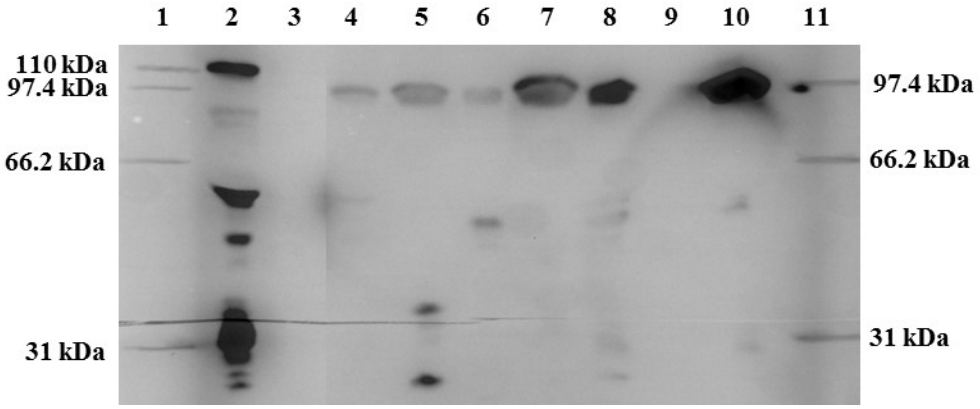


Figure 4.16: Western blot of soluble expressed P450s using the GroEL/ES chaperon system and glucose as carbon source
lane 1:Broad-Range-Marker; lane 2: CYP505A1_wo_LIC-3C (120.6 kDa); lane 3: CYP505A1_w_LIC-3C (139 kDa); lane 4:CYP539B1_LICRED (93.5 kDa); lane 5: CYP52G6_LICRED (95.4 kDa); lane 6: CYP52G8_LICRED (94.5 kDa); lane 7: CYP52T1_3_LICRED (96.5 kDa); lane 8: CYP52T1_2_LICRED (94.7 kDa); lane 9:CYP584E2_LICRED (97.8 kDa); lane 10: CYP52T1_LICRED (94.1 kDa); lane 11:Low-Range-Marker

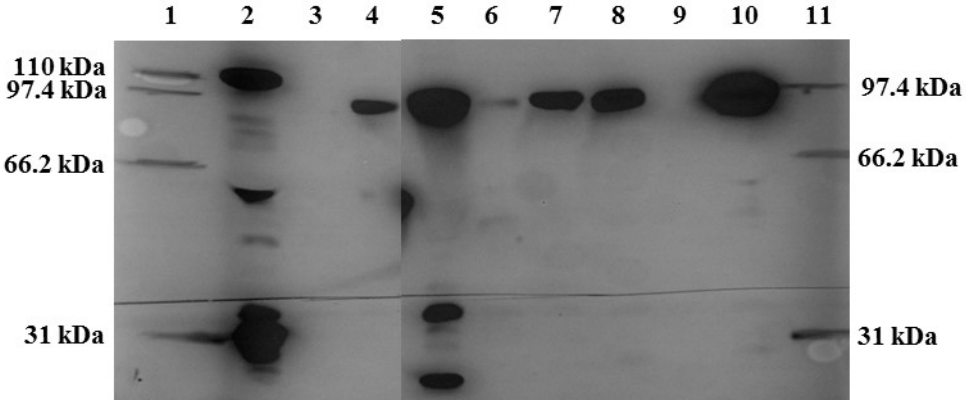


Figure 4.17: Western blot of soluble expressed P450s using the GroEL/ES chaperon system and glycerol as carbon source
lane 1: Broad-Range-Marker; lane 2: CYP505A1_wo_LIC-3C (120.6 kDa); lane 3: CYP505A1_w_LIC-3C (139 kDa); lane 4:CYP539B1_LICRED (93.5 kDa); lane 5: CYP52G6_LICRED (95.4 kDa); lane 6: CYP52G8_LICRED (94.5 kDa); lane 7: CYP52T1_3_LICRED (96.5 kDa); lane 8: CYP52T1_2_LICRED (94.7 kDa); lane 9: CYP584E2_LICRED (97.8 kDa); lane 10: CYP52T1_LICRED (94.1 kDa); lane 11: Low-Range-Marker

In both cases soluble expression was detectable for all P450s after only 30 s of exposure, except for CYP584E2 which didn't show any traceable signal of the expected size.

CYP52T1_LICRED, CYP52T1_2_LICRED, CYP52T1_3_LICRED and CYP505A1_wo_LIC-3C obtained the highest level of insoluble expression when grown with glucose in comparison to CYP539B1_LICRED, CYP52G6_LICRED and CYP52G8_LICRED which bands seem to be much weaker in Western blot analysis.

Expression levels of P450s grown in M9 medium containing glycerol, surpassed expression levels of P450s grown with glucose many times over, when detected with the same exposure time in Western blot analysis. Highest expression level have been reached by CYP52G6_LICRED (lane 5) and CYP52T1_LICRED (lane 10). Nevertheless, respectable overexpression could also be detected for CYP505A1_wo_LIC-3C, CYP539B1_LICRED (lane 4), CYP52T1_3_LICRED (lane 7) and CYP52T1_2_LICRED (lane 8). CYP52G8_LICRED (lane 6) revealed only a weak band in Western blot whereas CYP584E2_LICRED didn't show any signal.

In summary, it can be said that the soluble expression level of all P450 fusions could be improved using the GroEL/ES chaperon system. Only exceptions are CYP505A1 expressed with the transmembrane anchor and CYP584E2_LICRED which generally seemed to express rather poorly as indicated by earlier expression tests. Highest expression level could be reached by GroEL/ES co-expression and cell growth in M9 media using glycerol as carbon source. Hence, cell cultivation for whole cell activity assays, spectrophotometric characterization and purification were performed applying designated condition.

4.2.3 Whole cell activity assay

Even though the native substrates of *Beauveria bassiana*'s P450s are unknown so far, substrates were picked for initial activity testing. Whole cells were tested against *N*-benzylpyrrolidine as classical hydroxylation substrate for *Beauveria,* as well as *N*-benzoylpiperidine (*Beauveria* is known for catalysing the hydroxylation of *N*-protected heterocycles[219]) and limonene for the reason that alkane hydroxylases have been reported to show activity towards monterpene molecules.[253] Furthermore CYP52G8_LICRED was tested in whole cell activity assays against different concentrations (lowest concentration of 50 µM) of lauric acid as it has been shown to hydroxylate lauric acid with strict regioselectivity on the terminal methyl by Zhang *et al.*[217]

Assays were performed as described in 4.1.3.2 and analysed by GC (4.1.3.3). No activity could be detected for any of the P450s when tested against limonene (data not shown) and

N-benzoylpiperidine. The initial test against *N*-benzylpyrolidine on the other hand seemed promising. Absorption peaks that didn't appear in the negative control could be identified for most P450s and a general increase of products was observed (data not shown). However, the mentioned result couldn't be successfully reproduced. All peaks that appeared in P450 GC samples were also noticeable in the negative control. The first round of positive hits was therefore most likely the result of contaminated samples. Moreover, no conversion of lauric acid to 12-hydroxylauric acid by CYP52G8_LICRED could be shown, leading to the conclusion that the chimeric fusion is possibly inactive. Spectrophotometric analysis substantiates (4.2.4) this possibility.

4.2.4 Spectrophotometric characterisation

Cytochromes P450 have, in comparison to other cytochromes, rather uncommon spectral properties that allow for simple identification of active proteins using spectrophotometric analysis. Unreduced P450s have typically an absorption maximum at 420 nm which shifts of about 30 nm to 450 upon reduction and subsequent CO binding.[254]

Concentrated *E. coli* cell lysate with expressed P450s cultivated as described in 2.7.1.2 was reduced with sodium dithionite and then bubbled with carbon monoxide. Spectrophotometric analysis was performed for every step subsequently, as described in 4.1.3.1.

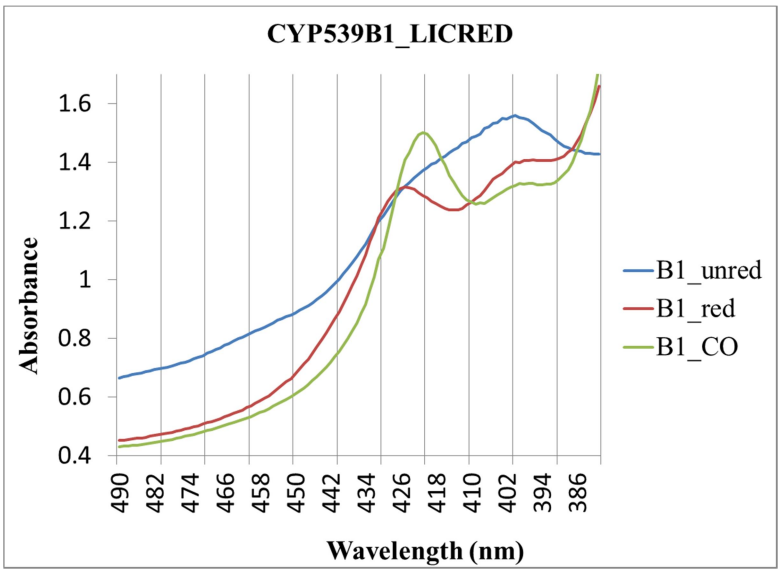


Figure 4.18: UV-visible absorbance spectra for CYP539B1_LICRED
oxidised form: blue line; sodium dithionite reduced form: red line; reduced, CO-bound form: green line.

Figure 4.18 displays the recorded CO-difference spectra for CYP539B1_LICRED as representative for the other *Beauveria* P450 fusions. Instead of an absorption maximum at 420 nm when unreduced, a rather broad peak at around 400 - 410 nm was observed for all strains. This absorption maximum shifted after reduction with sodium dithionite to a more distinct peak at 425 - 430 nm and seemed to move back to 420 nm upon CO binding. Host *E. coli* proteins in the cell lysate or not properly folded protein and therefore inactive protein might have been the reason of this untypical CO-difference spectra.

4.2.5 Purification of CYP539B1_LICRED

In order to preclude interference of external influences (*e.g.* host *E. coli* proteins and chaperones) CYP539B1_LICRED was chosen to be purified and subsequently be tested for the characteristic absorption peak upon reduction and CO binding at 450 nm. Purification was performed as described in 2.7.2.

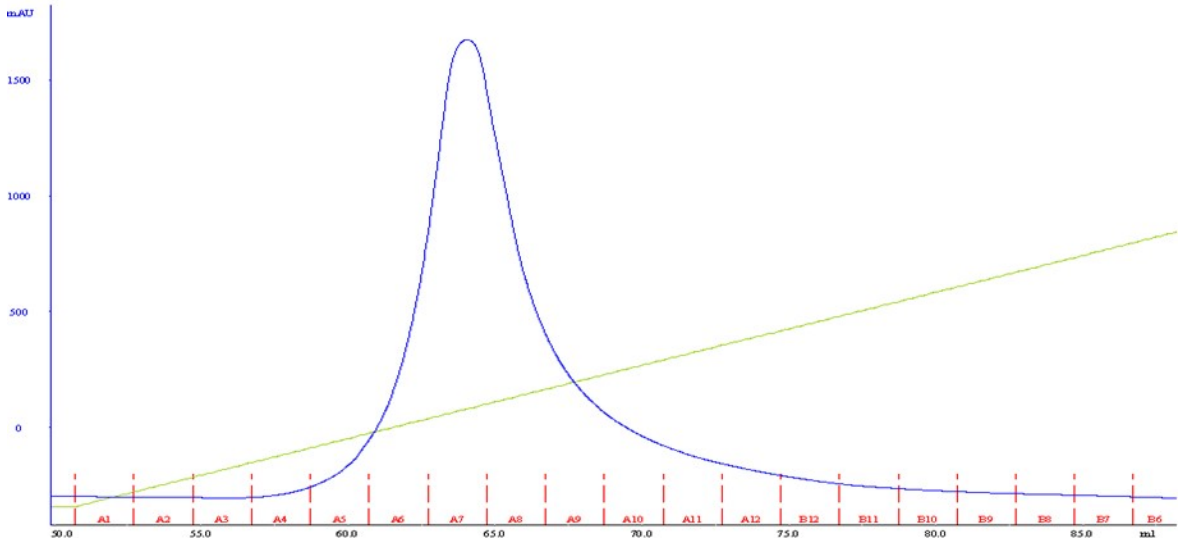


Figure 4.19: Nickel affinity chromatography chromatogram for the purification of 539B1_LICRED.
blue line: UV absorbance at 280 nm; green line: imidazole concentration gradient; Fraction A1-B6 are displayed along the x axis.

A distinct peak after nickel affinity chromatography could be verified for 539B1_LICRED as shown in Figure 4.19. Analysis of the collected fractions using SDS gel electrophoresis, however, revealed that the peak corresponded only to overexpressed chaperones (Figure 4.20) while no bands of the expected length could be detected for 539B1_LICRED.

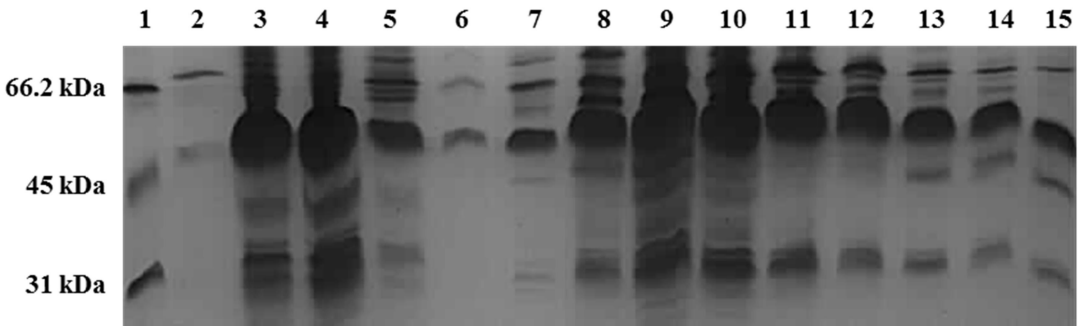


Figure 4.20: SDS-PAGE analysis of CYP539B1_LICRED fractions after nickel affinity chromatography
lane 1: 1 kb DNA ladder; lane 2: insoluble fraction; lane 3-4: collected fractions after washing of column with buffer A; lane 5: flow-through after column loading; lane 6-15: samples corresponding to the fractions A4, A5, A6, A7, A8, A9, A10, A11, A12, and B12 collected in nickel affinity chromatography (Figure 4.19).

Therefore, purification wasn't possible using nickel affinity chromatography, either because of steric hindrance of the polyhistidine-tag by the fusion-protein itself or as a result of strong interaction with the co-expressed chaperones.

4.3 Discussion

E. coli was initially selected for heterologous expression of *B. bassiana* P450s as it provides advantages such as low cost of maintenance, ease of use, usually high yield of protein in comparison to yeast and mammalian cells and the lack of endogenous P450s.[255] Despite these advantages, the use of bacterial cells for the expression of CYPs has until recently been limited primarily to the soluble prokaryotic CYPs as the expression of membrane-bound cytochrome P450s can cause expression problems due to different codon bias and protein misfolding. However, advances of more efficient techniques and methods led to well established systems capable of producing large yields of catalytically active protein.

N-terminal modification, for example, is an important tool for heterologous expression of cytochromes P450. With regard to the expression objective there are two categories of N-terminal modification: (1) modification to direct expression to the plasma membrane of *E. coli* and (2) modification to localize the protein in the cytoplasm. The first modification form is usually applied in order to retain the membrane binding characteristics. Membrane bound proteins possess signal sequences at the N-terminus (several hydrophobic aa) which are recognized by a signal recognition particle (SRP) and determine their membrane localization within the cell.[256] Although prokaryotes have analogous systems for directing protein to the membrane (plasma membrane), foreign signal sequences from eukaryotic

P450s are usually not recognized by bacterial systems and require further changes to increase expression levels as well as to obtain catalytically active protein.[257] Ways to achieve this include: altering (mutating) the native N-terminus in order to increase recognition of host SRP[258], insertion of previously established sequences to the N-terminus[259] and combination of truncation and substitution with a modified sequence.[260] The number of variations using this system is vast and has to be regarded separately for every P450. It was therefore more reasonable to select the second modification system for this project and simply truncate the N-terminal region in order to obtain soluble expressed proteins. Furthermore, in the event of soluble expressed protein, structural information could have been obtained using crystallization. However, CYP expression applying N-terminal truncation of the anchor sequence has been reported with varying results and is no guarantee for soluble expression as the F-G helices of cytochrome P450s are also involved in membrane association. Li and coworker, for example, were able to express 85% truncated rat hepatic cholesterol 7a-hydroxylase in the cytosol[261] while 50% of expressed truncated CYP46A1 by Mast and coworkers remained membrane bound.[262]

As mentioned above, *E.coli* lacks P450s and although it has endogenous electron transport systems it can't support the full catalytic activity of a heterologously expressed P450 without an additional CYP reductase. In order to bypass the necessity to express and process multiple proteins separately, P450 heme domains were fused to a redox partner, except for CYP505A1 which is a natural fusion already. Another advantage of the fusion system is that it simplifies the control of relative expression levels as P450 and reductase are produced in equivalent amounts and may also improve catalytic performance. In this project a non-natural redox partner from P450-Rhf was employed as it has been reported to be a versatile fusion partner. Although artificial fusion constructs utilising the P450-Rhf reductase (RhfRED) have been described by several different research groups[263-265] the LICRED system[227] developed by Sabbadin and coworkers appeared to be the most useful system, since it was generated as a high throughput tool for novel P450s and it was readily available in the university of York. The utility of this system was demonstrated for various bacterial targets and showed also promising results beyond class I type fusions. Schückel applied the LICRED system creating the first active bacterial–plant fusion P450 enzyme by fusing an N-terminally modified IFS1 (membrane anchor removed) to RhfRED.[266] Furthermore, the LICRED vector is based on the pET plasmid which is one of the most frequently used vectors for protein expression in *E. coli* and provided therefore a reliable

and controllable design for the LICRED system. A T7 RNA polymerase is needed for transcription of the gene of interest as it recognizes the T7 viral promoter resulting in selective high-level expression. This polymerase is, however, not naturally produced by bacterial cells and must be introduced into the host chromosome.[267] Thus, *E. coli* BL21 (DE3) and *E. coli* Rosetta (DE3) were used in this project since they are suitable for T7-expression systems.

Another factor to consider when expressing eukaryotic proteins in *E. coli* is the codon usage. Codons used in eukaryotes can differ from codons used in *E. coli* and as elongation proceeds it may happen that the ribosomes fail to recognize particular codon sequences due to the lack of corresponding available tRNA which can in turn influence the transcription rate and extent of heterologously expressed genes.[268] Furthermore, formation of secondary mRNA structures may hinder ribosomal binding but could be prevented by simple replacement of key nucleotides. Silent mutations offer therefore a simple method to increase the expression yield without altering the protein's integrity. Thus, for this project genes have been synthesized for expression with regard to codon optimisation for *E. coli*.

Expression problems can also occur through protein misfolding due to the absence of post-translational machinery in *E. coli* and misfolded proteins tend to be rapidly degraded.[269] The GroEL-ES system has therefore been employed in this study in order to assist the correct folding of unfolded proteins by providing the appropriate hydrophilic environment. This system was especially interesting as it has been reported to greatly improve fungal cytochrome P450 concentration in the soluble protein fraction[252] and to increase yields of eukaryotic cytochrome P450 proteins as well. Yamasaki and coworkers, for example, were able to increase CYP2J3 protein yields twelve-fold by simply adding the GroES-GroEL chaperone expression plasmid.[270]

Although all the above mentioned factors were considered, soluble expression of active proteins wasn't possible in *E. coli*. Nonetheless, insoluble expression levels could be greatly improved by modifying the expression medium using different carbon sources. It is arguable wether the insolubility of the chimeric fusions was caused by the eukaryotic nature of the heme domains or may be triggered by the addition of a bacterial reductase domain. Experiments expressing solely the truncated heme domains CYP539B1 and CYP52G8 employing the LIC-3C vector (data not shown), however, had similar results when using the same expression conditions. No soluble expression could be verified for heme domains leading to the assumption that the general insolubility of the chimeric fusion

is owed to the eukaryotic nature of the heme domain rather than to the fused bacterial Rhf reductase domain. Thus, further changes in the expression strategy were required in order to improve solubility.

Although soluble expression of chimeric fusions as well as truncated CYP505A1 could be increased by introduction of chaperonin GroEL-GroES, the increase in solubility seemed to be mainly caused through strong interaction between the P450 protein and the chaperones. As a result, the GroEL-GroES complex appeared to be permanently attached to the P450 and thus, forced solubility as the fusions were impeded to either interact with the plasma membrane or form aggregates. This is substantiated by the purification results of 539B1_LICRED which was co-expressed with chaperones. Furthermore, protein integrity checks using CO-different spectra of crude lysate containing P450s co-expressed with chaperones indicated the loss of activity of P450 enzymes as it was only possible to measure wavelength maxima at 420 nm instead of 450 nm. Moreover, whole cell activity assays failed to confirm the presence of catalytically active protein. It is arguable that the absence of product formation could be not only caused by inactivity of the proteins but also the lack of binding affinity of the P450s to deployed substrates. However, in the case of the CYP52G8 chimera, active protein should have been able to convert lauric acid to 12-hydroxylauric acid. The reason for this assumption is given by the fact that native CYP52G8 is identical to a by Zhang *et al.* published *B. bassiana* P450 (CYP52X1) that has been shown to perform this reaction. [217]

In conclusion, the *E.coli* expression system wasn't suitable for soluble expression of neither natural and truncated chimeric *B. bassiana* fusions nor sole *B. bassiana* heme domains. Further changes would have had to be applied to this system in order to improve protein expression and cellular distribution regarding each P450 individual. In consideration of remaining project time and number of variables to be accounted for in order to improve the *E. coli* expression system it was decided to simply move on to an eukaryotic expression host that was already reported to express *B. bassiana* P450s[217] instead of improving the bacterial host system.

5. Construction of library of a *B. bassiana* P450s in pYeDP60 and expression in *Saccharomyces cerevisiae* WAT11

Heterologous expression of eukaryotic P450s using bacterial hosts such as *E. coli* can be very challenging. A major issue that leads to often poor or no soluble expression is due to the interference of the N-terminal anchor. In order to provide conditions that enhance the probability of soluble protein expression in *E. coli,* genes of the *B. bassiana* P450 domains have been synthezised without the N-terminal anchor as described in 4.1.1. However, cytochrome P450 expression in *S. cerevisiae* requires in comparison to expression in *E. coli* N-terminal structures that drive localisation of the protein to the endoplasmic reticulum (ER) membranes. It is important to integrate cytochromes P450 in the ER membranes of the organism in order to enable interaction with the membrane bound cytochrome P450 reductase which acts as redox partner and supplies the P450 with electrons for biocatalysis.[60] Furthermore, the redox-environment created in the *S. cerevisiea* ER is more similar to the natural surroundings of membrane bound P450 enzymes and thus may have a positive impact on the protein integrity and activity. It was therefore necessary to reintroduce the N-terminal anchor for yeast expression.

5.1 Materials and Methods

5.1.1 Reintroduction of the N-terminal region

The sequences of the removed N-terminal regions (4.1.1) were send to GeneArt® to be synthesized. Amplification of the anchor regions (see sequences in Appendix B) and the truncated P450 heme domains took place using KOD hot start polymerase (2.5.5) and corresponding primers (Appendix A, Table A.3). PCR product of the N-terminal anchor, PCR product of appropriate P450 heme domain and the linearized pYeDP60 vector (2.2.3) were then combined in one reaction using the In-Fusion® HD Cloning System (2.5.7). Sequencing by GATC confirmed the success of the reintroduction process and the vector containing the full length sequence of the *B. bassiana* P450 heme domain could be used as template for the actual cloning described below.

5.1.2 Cloning of P450

The 7 selected *B. bassiana* P450 heme domains were amplified in full length using KOD hot start polymerase (2.5.5) and corresponding primers (Table 5.1, Table 5.2, and

Table 5.3). Amplification products deployed for cloning incorporated either a N-terminal his-tag, C-terminal his-tag or no his-tag. The full length self-sufficient P450 (CYP505A1) was amplified with addition of a C-terminal his-tag only. PCR products were cloned into the linearized pYeDP60 vector (2.2.3) using the In-Fusion® HD Cloning System (2.5.7). The pYeDP60 vector was linearized by double digest with *Bam*HI and *Kpn*I and serves as a shuttle vector for cloning in *E. coli* and expression in yeast.[228]

Table 5.1: Primer pairs used for PCR amplification incorporating a N-terminal his tag

Primer	Sequence
52G6_DP60_n-tag_F	CTAAATTACCGGATCCATGCATCATCACCATCACCAC
52G6_DP60_n-tag_R	GCGAATTCGAGCTCGGTACCTTACACTGCTTCATGCAGACG
52G8_DP60_n-tag_F	CTAAATTACCGGATCCATGCATCATCACCATCACCAC
52G8_DP60_n-tag_R	GCGAATTCGAGCTCGGTACCTTATTCATCAAAGTGAACTTTC AGG
52T1_DP60_n-tag_F	CTAAATTACCGGATCC ATGCATCATCACCATCACCAC
52T1_DP60_n-tag_R	GCGAATTCGAGCTCGGTACCTTAGCTGCTATTACCCAGG
52T1_2_DP60_n-tag_F	CTAAATTACCGGATCCATGCATCATCACCATCACCAC
52T1_2_DP60_n-tag_R	GCGAATTCGAGCTCGGTACCTTAATTTGCCTGCTGCAGACG
52T1_3_DP60_n-tag_F	CTAAATTACCGGATCCATGCATCATCACCATCACCAC
52T1_3_DP60_n-tag_R	GCGAATTCGAGCTCGGTACCTTACAGCTGTGCTGCTTCACA
539B1_DP60_n-tag_F	CTAAATTACCGGATCCATGCATCATCACCATCACCA
539B1_DP60_n-tag_R	GCGAATTCGAGCTCGGTACCTTAGGCCGTTTTTTCATCG
584E2_DP60_n-tag_F	AAATTAATAATGACCGGATCCATGGGCAGCAGCCATCATC
584E2_DP60_n-tag_R	CGCGAATTCGAGCTCGGTACCTTAGTTTTTGGCACGGGTCAG

Table 5.2: Primer pairs used for PCR amplification incorporating a C-terminal his tag

Primer	Sequence
52G6_DP60_c-tag_F	CTAAATTACCGGATCCATGGCGCTCACTGCTATCC
52G6_DP60_c-tag_R	CGAGCTCGGTACCTTAGTGGTGATGGTGATGATGCACTGCTTC ATGCAG
52G8_DP60_c-tag_F	CTAAATTACCGGATCCATGGCGGTACTATCCGTTA
52G8_DP60_c-tag_R	CGAGCTCGGTACCTTAGTGGTGATGGTGATGATGTTCATCAA AGTGAAC
52T1_DP60_c-tag_F	CTAAATTACCGGATCCATGGCTCTCCACGCTGCC
52T1_DP60_c-tag_R	CGAGCTCGGTACCTTAGTGGTGATGGTGATGATGGCTGCTATT ACCCAG
52T1_2_DP60_c-tag_F	CTAAATTACCGGATCCATGAACTTGTTGTCATCGTCTT
52T1_2_DP60_c-tag_R	CGAGCTCGGTACCTTAGTGGTGATGGTGATGATGATTTGCCTG CTGCAG
52T1_3_DP60_c-tag_F	CTAAATTACCGGATCCATGACTCTCTCGCCTATTTCTAC
52T1_3_DP60_c-tag_R	CGAGCTCGGTACCTTAGTGGTGATGGTGATGATGCAGCTGTG CTGCTTC
539B1_DP60_c-tag_F	CTAAATTACCGGATCCATGCTTATCGAGGCTGTCA
539B1_DP60_c-tag_R	CGAGCTCGGTACCTTAGTGGTGATGGTGATGATGGGCCGTTTT TTCATC
584E2_DP60_c-tag_F	AAATTAATAATGACCGGATCCATGGCACTGGGTCAGCTG
584E2_DP60_c-tag_R	CGAGCTCGGTACCTTAGTGGTGATGGTGATGATGGTTTTTGGC ACGGGT

Primer	Sequence
505A1_DP60_c-tag_F	CTAAATTACCGGATCCATGGATGTTGATCAGATTGC
505A1_DP60_c-tag_R	CGAGCTCGGTACCTTAGTGGTGATGGTGATGATGGTCAAACA CATCGGC

Table 5.3: Primer pairs used for PCR amplification incorporating no additional his tag

Primer	Sequence
52G6_DP60_F	CTAAATTACCGGATCCATGGCGCTCACTGCTATCC
52G6_DP60_R	GCGAATTCGAGCTCGGTACCTTACACTGCTTCATGCAGACG
52G8_DP60_F	CTAAATTACCGGATCCATGGCGGTACTATCCGTTA
52G8_DP60_R	GCGAATTCGAGCTCGGTACCTTATTCATCAAAGTGAACTTTCAGG
52T1_DP60_F	CTAAATTACCGGATCCATGGCTCTCCACGCTGCC
52T1_DP60_R	GCGAATTCGAGCTCGGTACCTTAGCTGCTATTACCCAGG
52T1_2_DP60_F	CTAAATTACCGGATCCATGAACTTGTTGTCATCGTCTT
52T1_2_DP60_R	GCGAATTCGAGCTCGGTACCTTAATTTGCCTGCTGCAGACG
52T1_3_DP60_F	CTAAATTACCGGATCCATGACTCTCTCGCCTATTTCTAC
52T1_3_DP60_R	GCGAATTCGAGCTCGGTACCTTACAGCTGTGCTGCTTCACA
539B1_DP60_F	CTAAATTACCGGATCCATGCTTATCGAGGCTGTCA
539B1_DP60_R	GCGAATTCGAGCTCGGTACCTTAGGCCGTTTTTTCATCG
584E2_DP60_F	AAATTAATAATGACCGGATCCATGGCACTGGGTCAGCTG
584E2_DP60_R	CGCGAATTCGAGCTCGGTACCTTAGTTTTTGGCACGGGTCAG

5.1.3 Expression

Expression was performed under conditions requiring appropriate sterilization of media and consumables by autoclaving.

S. cerevisiae colonies containing relevant plasmid were restreaked on a selective plate (SGI, 2.3.2) as shown in Figure 5.1 in order to harvest a reasonable amount of cell mass that could serve as inoculum for the SGI liquid preculture. Colonies used for expression tests have been analysed by colony PCR (2.5.5) prior to inoculation to ensure the presence of the CYP_pYeDP60 construct. The designed primers (Table 5.4) correspond to internal vector sequences resulting in PCR products incorporating 203 nucleotides in front of the cloned gene sequence and 84 nucleotides behind.

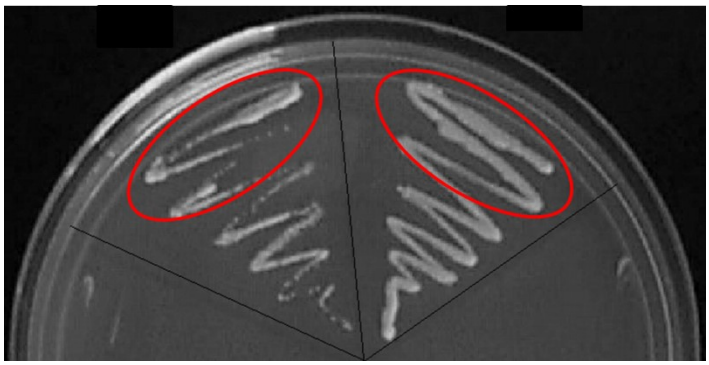


Figure 5.1: *S. cerevisiae* colonies on SGI plate.
Cell amount serving as inoculum for preparation of preculture for expression is circled in red.

Table 5.4: Primer pair used for colony PCR

Primer	Sequence
FWD_pYeDP60_int	TAGTGCTGACACATACAGGC
REV_pYeDP60_int	CCTGGCAATTCCTTACCTTC

For expression, 20 ml of SGI liquid medium (2.3.2) was inoculated with positive transformants, as described above, and incubated for 18 h at 28 °C (160 r.p.m.). The OD_{600} of the overnight culture was measured the following morning in a 1:10 dilution to calculate the appropriate volume of overnight culture necessary to inoculate 200 ml YPGE (2.3.2) main culture to an OD_{600} of 0.15. Main cultures were incubated shaking (160 r.p.m.) at 28 °C. After 30 h incubation, the main source of carbon in the medium (glucose) which is responsible for the repression of transcription should be used up, and thus forcing the yeast to switch metabolism towards ethanol. In order to induce transcription 10 ml galactose (2.3.3) were added to each flask and incubated shaking (160 r.p.m.) for 18 h at 18 °C. In presence of the inducer (galactose) Gal4p, a regulatory protein, binds to sites in the UAS (upstream activation sequence) and activates transcription. The yeast was then harvested and the microsomes isolated (5.1.4).

5.1.4 Microsome preparation

Microsomes were prepared by differential centrifugation.[228] Yeast cells were harvested by centrifugation at 4250 r.p.m. (Sorvall RC 5B plus centrifuge, F10-6x500 rotor) for 10 min at 4 °C. The cell pellet was then resuspended in 30 ml ice-cold TEK buffer and recovered by additional centrifugation. Cells were resuspended in 2 ml TES buffer and glass beads (glass beads, 0.40-0.60 mm, Sartorius) were added up to the level of cell suspension. The cell walls were disrupted mechanically by hand shaking in 6 x 1 min intervals with 1 min delay between each interval on ice. Glass bead suspension was washed three times with 20 ml TES buffer and the washings were transferred into fresh tubes. The solution was centrifuged (Sorvall Evolution RC centrifuge, SLC-3000 rotor) for 20 min at 7500 r.p.m. (4 °C) to remove cell membranes, leaving membrane fractions and soluble proteins in the supernatant. The supernatant was filtered through miracloth (Calbiochem, Merck) to remove all glass beads and then ultracentrifuged at 35000 r.p.m. at 4 °C (Beckman Avanti J-HC ultracentrifuge, 45 Ti-rotor) for 1 – 2 h to pellet the microsomes. The pellet was transferred into a potter homogeniser, 500 µl TEG added and homogenised on ice avoiding the build up of air bubbles.

To test the presence of active P450, sodium dithionite and CO was added by bubbling the microsome solution in the cuvette for 30 s in order to create the reduced CO bound form of the P450. Absorption was measured between 200 and 500 nm.

TEK buffer

2M Tris-HCl (pH 7.5)		12.5 ml
0.5M EDTA (pH 8.0)		1 ml
KCl		3.7 g
H_2O	fill up to	500 ml

TES buffer

2M Tris-HCl (pH 7.5)		25 ml
0.5M EDTA (pH 8.0)		2 ml
Sorbitol		109 g
H_2O	fill up to	1000 ml

10 g BSA and 120 µl β-mercaptoethanol were added to 1 l buffer on the day of use.

TEG buffer

2M Tris-HCl (pH 7.5)		2.5 ml
0.5M EDTA (pH 8.0)		100 µl
Glycerol		30 ml
H_2O	fill up to	100 ml

5.2 Results

5.2.1 Cloning into pYeDP60

7 *B. bassiana* heme domains were cloned into the pYeDP60 plasmid as described in 5.1.2. Figure 5.2 shows the success of PCR for all amplification products that incorporate the C-terminal his-tag.

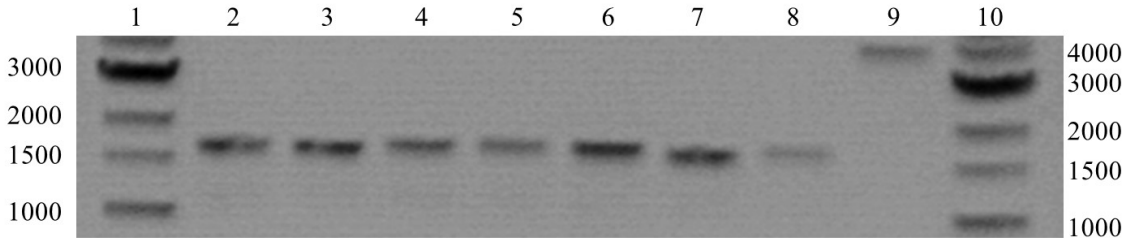


Figure 5.2: PCR products incorporating the C-terminal his-tag after gel extraction
lane 1: 1 kb DNA ladder; lane 2: CYP52G6; lane 3: CYP52G8; lane 4: CYP52T1_2; lane 5: CYP52T1_3; lane 6: CYP52T1; lane 7: CYP539B1; lane 8: CYP584E2; lane 9: CYP505A1; lane 10: 1 kb DNA ladder

PCR results comprising DNA sequences without his-tag and with N-terminal his-tag are not shown as the individual sequence length of the corresponding gene is very similar to the C-terminal his-tagged amplification product and would not show differences on an

agarose gel. Table 5.5 lists the relative nucleotide length for all amplification products in correspondence to molecular changes.

Table 5.5: Relative nucleotide length of amplification products

Template sequence	base pair length of nucleotide sequence		
	Without his-tag	N-terminal his-tag	C-terminal his tag
CYP52G6	1623	1641	1634
CYP52G8	1623	1641	1634
CYP52T1	1641	1659	1652
CYP52T1_2	1647	1665	1658
CYP52T1_3	1662	1680	1673
CYP539B1	1584	1602	1595
CYP584E2	1630	1648	1639
CYP505A1	-	-	3746

After annealing of prepared vector and insert and subsequent transformation into NovaBlue Single competent cells (2.6.3), several colonies were picked for an initial screen of positive transformants by control digest with *Bam*HI and *Kpn*I (2.5.1) prior to the final confirmation using sequencing. Positive clones for all constructs containing the C-terminal his-tag are shown in Figure 5.3.

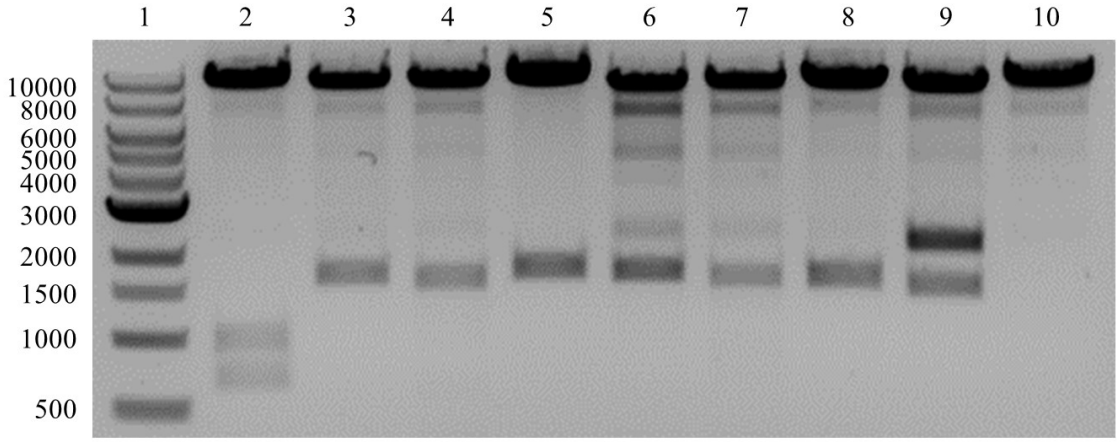


Figure 5.3: Initial screen of positive transformants containing a C-terminal his-tag using restriction enzymes *Bam*HI and *Kpn*I
lane 1: 1 kb DNA ladder; lane 2: CYP52G6_DP60; lane 3: CYP52G8_DP60; lane 4: CYPT1_2_DP60; lane 5: CYPT1_3_DP60; lane 6: CYP52T1_DP60; lane 7: CYP539B1_DP60; lane 8: CYP584E2_DP60; lane 9: CYP505A1_DP60; lane 10: empty pYeDP60

The image displays 1 selected clone for each P450 construct with C-terminal his-tag containing the empty pYeDP60 vector and the insert of the expected length after digest. The nucleotide length of the inserts after digest is similar to the nucleotide length of the individual amplification products listed in Table 5.5 since the restriction sites flank the inserted DNA sequences directly. CYP52G6_DP60, CYPT1_2_DP60 and CYP505A1_DP60 each possess a binding site for the applied restriction enzymes and thus

display 2 fragments in agarose gel (CYPT1_2_DP60's second fragment is too small to see in agarose gel). Sequencing by GATC confirmed the success of the cloning process.

5.2.2 Transformation into *S. cerevisiae*

Transformation was performed as described in 2.6.4. In order to confirm the insertion of the correct pYeDP60 construct, individual *S. cerevisiae* colonies were picked using a sterile pipette tip and restreaked on a selective plate (SGI, 2.3.2) as shown in Figure 5.1. On the pipette tip remaining cell material was then used for colony PCR (2.5.5). Figure 5.4 and Figure 5.5 display the colony PCR of two to three selected transformants for each P450 construct incorporating a C-terminal his-tag. Colony PCR results for P450 constructs containing no his-tag or an N-terminal his-tag are not displayed for the same reason mentioned in 5.2.1.

Positive transformation could not be verified for all selected colonies. In case of CYP52G6_DP60, for example, only one out of three colonies was tested positive using colony PCR. *S. cerevisiae* that wasn't transformed should not be able to grow on selective agar plates and *S. cerevisiae* transformed with the original vector (without insert) would have shown a band of around 300 bp on the agarose gel. This, however, was not the case leading to the assumption that too much employed cell material and therefore high concentrations of DNA obstructed the PCR.

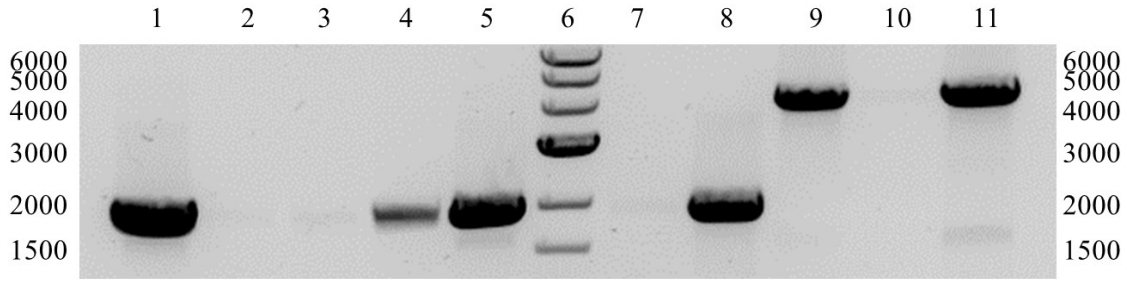


Figure 5.4: Colony PCR of transformed *S. cerevisiae* (pYeDP60 constructs containing genes of interest with C-terminal his-tag)
lane 1: CYP52G6_DP60_p1; lane 2: CYP52G6_DP60_p2; lane 3: CYP52G6_DP60_p3; lane 4: CYPT1_3_DP60_p1; lane 5 CYPT1_3_DP60_p2; lane 6: 1kb DNA ladder; lane 7: CYPT1_DP60_p1; lane 8: CYPT1_DP60_p2; lane 9: CYP505A1_DP60_p1; lane 10: CYP505A1_DP60_p2; lane 11: CYP505A1_DP60_p3

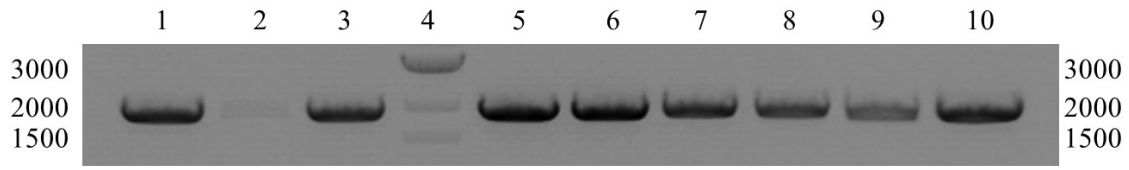


Figure 5.5: Colony PCR of transformed *S. cerevisiae* (pYeDP60 constructs containing genes of interest with C-terminal his-tag)
lane 1: CYP52G8_DP60_p1; lane 2: CYP52G8_DP60_p2; lane 3: CYP52G8_DP60_p3; lane 4: 1kb DNA ladder; lane 5 CYP584E2_DP60_p1; lane 6: CYP584E2_DP60_p2; lane 7: CYPT1_2_DP60_p1; lane 8: CYPT1_2_DP60_p2; lane 9: CYP539B1_DP60_p1; lane 10: CYP539B1_DP60_p2;

Table 5.6 lists the relative nucleotide length for all amplification products expected after colony PCR in correspondence to molecular changes.

Table 5.6: relative nucleotide length of colony PCR amplification products

Template sequence	base pair length of nucleotide sequence	
	Without his-tag	N-terminal his-tag or C-terminal his tag
CYP52G6	1874	1892
CYP52G8	1874	1892
CYP52T1	1892	1910
CYP52T1_2	1898	1916
CYP52T1_3	1913	1931
CYP539B1	1835	1853
CYP584E2	1874	1892
CYP505A1	-	4004

5.2.3 Expression test in *Saccharomyces cerevisiae* WAT11

Colonies used for expression tests have been analysed by colony PCR prior to inoculation as described in 5.1.3 in order to unsure the presence of the CYP_pYeDP60 constructs. After the harvest of cells (5.1.3) microsomes were prepared using differential centrifugation (5.1.4). Membrane-bound P450s expressed in microsomes have usually a rather low concentration and are therefore not easily verified *via* SDS page. No signals of the expected sizes could be detected in SDS page (data not shown). Thus, other identification methods had to be applied in order to confirm the presence of overexpressed P450s in the microsome fractions. Spectrophotometric analysis (CO-difference spectra) and the high sensitive method of Western blotting are suitable for this task and have been used in this study.

A 1:4 dilution of the microsome fraction in TEG buffer was reduced with sodium dithionite and then bubbled with carbon monoxide. Absorption for unreduced and reduced, CO treated microsome solution was measured between 200 and 600 nm. Figure 5.6

displays the recorded CO-difference spectra for CYP52T1 as representative for the other *Beauveria* P450s. Membrane fractions of induced *S. cerevisiae* containing the empty pYeDP60 have been analysed as a negative control in order to judge the influence of endogenously expressed P450s (Figure 5.6 **A**). Here, only a small peak around 425-430 nm could be detected after reduction and subsequent CO-treatment. No typical 450 nm peak was visible leading to the assumption that the expressed endogenous P450s are most likely not correctly folded and therefore inactive. The same applies for all microsome fractions that should contain overexpressed fungal P450s (Figure 5.6 **B**, Figure 5.6 **C**, and Figure 5.6 **D**). In fact, considering peak position, height and width, the CO-difference spectra appear suspiciously similar to the CO-difference spectrum of microsome fractions containing only endogenous P450s and therefore, signifying no expression at all. However, CO-difference spectra are not always dependable. Thus, the more sensitive and reliable method of Western blotting was performed for final validation. Unfortunately, Western blot analysis confirmed the results indicated by the spectrophotometric analysis. No signals of expressed proteins with the predicted molecular weight could be detected (data not shown).

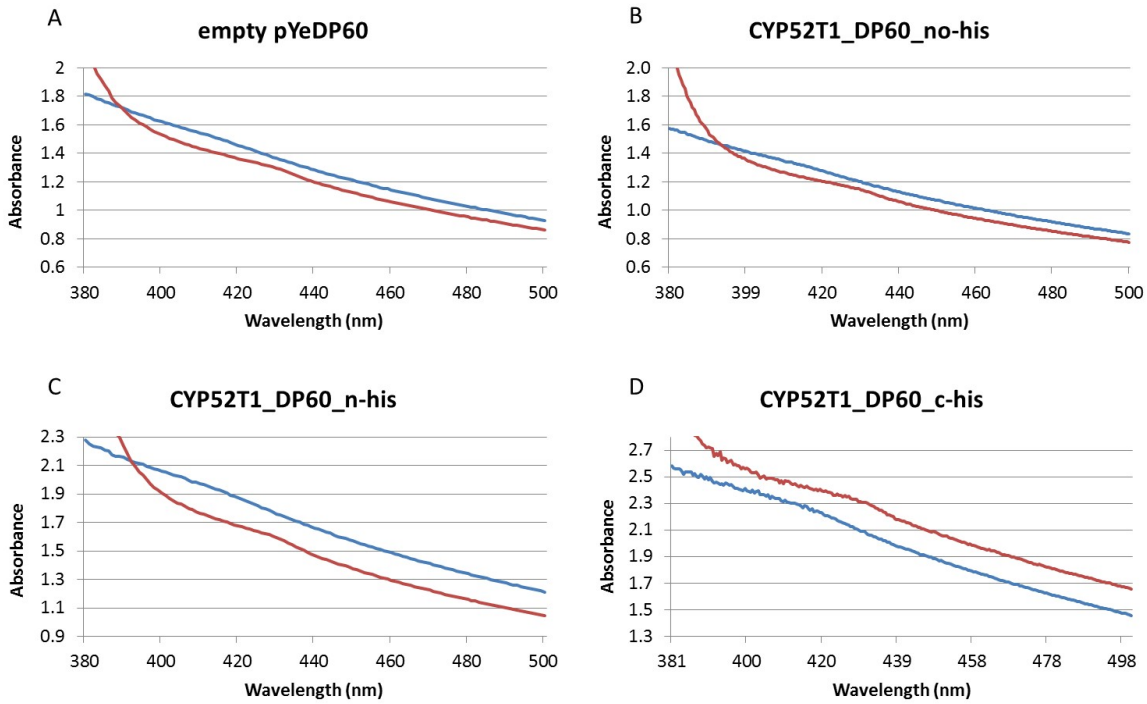


Figure 5.6: UV-visible absorbance spectra for CYP52T1_DP60 containing either no his tag, N-terminal his-tag or C-terminal his-tag in comparison to empty pYeDP60
oxidised form: blue line; reduced, CO-bound form: red line. The microsome fraction from induced *S. cerevisiae* containing the empty pYeDP60 vector (A) and microsome fractions expressing *CYP52T1* without his-tag (B), N-terminal his-tag (C) and C-terminal his-tag (D) are shown.

5.3 Discussion

The reason to select *S. cerevisiae* as aeukaryotic heterologous expression host for *B. bassiana* P450s is reflected by its obvious advantages. First of all, *Saccharomyces cerevisiae* is the first system reported to successfully express mammalian[34] and plant[271] P450 proteins and has advanced to the most frequently used system for functional expression of plant P450s.[272] Furthermore, it shows great promise for expression of functional fungal cytochrome P450 heme domains.[273-274] The most important advantage of the yeast system in comparison to other eukaryotic expression systems is, however, the combination of prokaryotic simplicity of growth (*e.g.*, low cost culture media, rapid growth) and manipulation (optimized yeast systems for P450 expression) with eukaryotic complexity of protein machinery (*e.g.*, posttranslational folding) and the presence of an ER membrane environment.

The pYeDP60 vector system[228], which has so far been the most successfully used for cytochrome P450, and the WAT11 yeast strain were selected for the expression of the *B. bassiana* P450s. *S. cerevisiae* WAT11 has been initially developed for expression of plant P450s as it contains a chromosomally integrated *Arabidopsis* P450 reductase gene (*ATR1*) to overcome coupling deficiencies with the endogenous *S. cerevisiae* P450 reductase and thus provides a redox environment that is optimal for plant P450 activities. Despite that, utilisation of the WAT11 strain in combination with the pYeDP60 vector system was still considered the most promising strategy for functional expression of *B. bassiana* P450s as it was already successfully applied for 5 *B. bassiana* P450s by Pedrini and co-worker.[274] Moreover, 2 (CYP52X1, CYP5337A1) of these 5 P450s possess identical amino acid sequences with CYP52G8 and CYP52T1 which were selected in this study and should have therefore been successfully expressed when using the same expression system and conditions.

P450s expressed by Pedrini *et al.* have either been cloned with an N-terminal his-tag or C-terminal his-tag. Although the insertion of a N-terminal his-tag is usually arguable since it is expected to be cleaved together with the signal sequence which conveys protein localization to the membrane, P450s are processed differently in comparison to other proteins as their N-terminal region is recognized by the signal recognition particle but not cleaved.[257] Thus, it was decided to use both systems for this study as well as untagged P450s in order to determine possible differences with regard to substrate assay performances of functional expressed P450s. Unfortunately, no expression could be

detected for any of the P450 constructs. Although the absence of a typical CO difference spectral shift is not necessarily proof for a lack of expression, as shown for 4 out of 5 expressed *B. bassiana* P450s by Pedrini *et al.*,[274] the absence of protein bands in Western blot analysis on the other hand is evident.

The nature of the utilized pYeDP60 vector system presents one possible explanation. In this system, the GAL10-CYC1 promoter is fully repressed by the glucose in the medium and derepression is only achieved when glucose is exhausted and cells are forced to rely on ethanol utilization for growth followed by full induction of protein expression through the addition of galactose.[228] It seemed possible that not all of the glucose was consumed by the time of induction and thus no expression was induced. However, the exchange of glucose containing medium at the end of the exponential growth phase with induction medium containing galactose instead led to similar results in the UV-visible absorbance spectra and Western blot analysis.

A more reasonable explanation can be found when looking at the nucleotide sequences of the cloned genes. As already mentioned for the *E. coli* expression system, codon bias of the expression host can affect the efficiency of translation, often resulting in proteins being poorly expressed or mistranslated in heterologous organisms. In the case of the in this study selected P450 enzymes, template sequences for PCR and additional cloning derived from *E. coli* optimised synthesised genes whereas Pedrini *et al.* were using the original fungal cDNA as template for PCR. Although the average GC content of the fungal P450s (ranging from 49.8% for CYP539B1 to 61.7% for CYP52G8) has been decreased to about 48-50% and as a consequence is significantly closer to the average GC content of yeast genes (40%)[275] the codon bias in yeast is very different.

Table 5.7: comparison of codon usage of 4 major aa in *E. coli* and *S. cerevisiae*

Amino acid	Codon	Codon usage in %	
		E. coli	*S. cerevisiae*
Leucine (Leu, L)	CTG	55	9
Serine (Ser, S)	AGC	26	9
Proline (Pro, P)	CCG	56	9
Arginine (Arg, R)	CGT	45	17
	CGC	37	4

Among the most unfavourable codons with reference to the bias in highly expressed genes in yeast, are the codons CTC for Leu, AGC for Ser, CCG for Pro and CGT or CGC for Arg (Table 5.7).[276] In *CYP505A1* they represent 17.44 % of the total codons while the

percentage in the non-fusion genes is even higher (25-30 %). More alarming is the fact that 100 % of the amino acids Leu and Arg are coded by rare codons. Low usage codons for Pro and Ser are utilized in a rate of 79-100 % (Table 5.8). It is therefore reasonable to assume that the presence of these low-usage codons have caused translational inhibition especially under consideration of the fact that protein sequence for 2 *B. bassiana* P450s as well as expression system and conditions have been identical to that of Pedrini *et al.*. Furthermore, it has been previously reported that codon bias in *S. cerevisiae* can have an enormous influence of protein expression.[277-278] Batard and coworkers, for example, were able to increase the expression of CYP73A17 to 300 pmol P450 mg^{-1} protein after recoding 111 base pairs while no expression was detectable when using the original gene sequence.[278] In addition, it has to be mentioned that the *B. bassiana* genes were expressed using the very strong GAL10-CYC1 promotor that might have contributed to the translational inhibition by simply inducing the synthesis of a critical number of transcripts which could have then led to complete depletion in rare tRNA and thus an excessive slowing down of the ribosomes.

Table 5.8: Proportion of low usage codons in gene-sequences of *B. bassiana* heme-domains optimised for *E. coli* expression

Gene	Codon (aa)	Rare codon used for aa (%)	Rare codon used in total aa sequence (%)	Total rare codon usage (%)
CYP505A1	CTG (Leu)	100	8.11	17.44
	AGC (Ser)	88	2.76	
	CCG (Pro)	93	2.59	
	CGT CGC (Arg)	100	3.98	
CYP52G6	CTG (Leu)	100	11.15	26.55
	AGC (Ser)	81	3.9	
	CCG (Pro)	79	4.3	
	CGT CGC (Arg)	100	7.2	
CYP52G8	CTG (Leu)	100	8.9	26.78
	AGC (Ser)	85.3	5.48	
	CCG (Pro)	100	4.9	
	CGT CGC (Arg)	100	7.5	
CYP52T1	CTG (Leu)	100	11.79	30.61
	AGC (Ser)	85.71	6.74	
	CCG (Pro)	93.54	5.34	
	CGT CGC (Arg)	100	6.74	

Gene	Codon (aa)	Rare codon used for aa (%)	Rare codon used in total aa sequence (%)	Total rare codon usage (%)
CYP52T1_2	CTG (Leu)	100	9.8	26.15
	AGC (Ser)	80	4.47	
	CCG (Pro)	92	4.28	
	CGT CGC (Arg)	100	7.6	
CYP52T1_3	CTG (Leu)	100	10.5	25.28
	AGC (Ser)	82.35	5.2	
	CCG (Pro)	91.3	3.87	
	CGT CGC (Arg)	100	5.71	
CYP539B1	CTG (Leu)	100	7.75	24.54
	AGC (Ser)	92.6	4.88	
	CCG (Pro)	93.93	6.05	
	CGT CGC (Arg)	100	5.86	
CYP584E2	CTG (Leu)	100	10.79	28.95
	AGC (Ser)	95.45	3.97	
	CCG (Pro)	79.31	4.35	
	CGT CGC (Arg)	100	9.84	

In conclusion, the *S. cerevisiae* expression system was expected to deliver functionally expressed P450 because it has been demonstrated to be suitable for *B. bassiana* P450 expression by Pedrini and coworkers.[274] Despite of the use of identical expression system and conditions, expression wasn't successful which was most likely due to the codon bias in yeast as the gene sequences used for expression have been optimised for *E. coli* and not for *S. cerevisiae*. Consequently, new gene sequences would have had to be applied to this system in order to improve protein expression and carry out substrate assays. In consideration of remaining project time, however, it wasn't possible to redo all the steps necessary to achieve functional fungal P450s in the microsomes of *S. cerevisiae*.

6. A P450 fusion library of heme domains from *Rhodococcus jostii* RHA1 and its evaluation for the biotransformation of drug molecules

23 *Rhodococcus jostii* P450 domains that had already been cloned into the LICRED vector by Ralph Hyde (University of York) have been investigated in the interests of identifying activities for the transformation of drug molecules. In addition, the applicability of the LICRED system as a high throughput tool could be reevaluated.

6.1 Introduction

In order to gain an overview of the 25 annotated P450 enzymes from *Rhodococcus jostii* RHA1, a phylogenetic analysis of the putative gene targets was performed by Lindsay Eltis and is shown in Figure 6.1.[279] A list of the gene targets can be found in Appendix A (Table A.2).

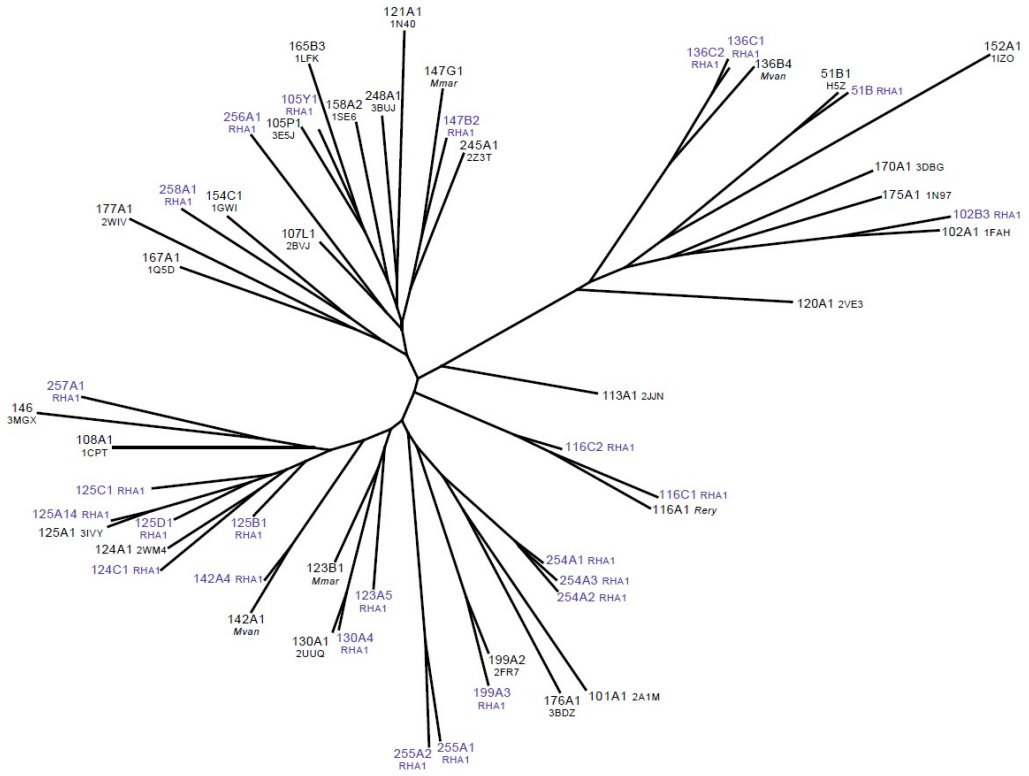


Figure 6.1: Phylogenetic tree constructed of *Rhodococcus jostii* RHA1 gene targets (blue) Included with the RHA1 P450 genes are various cytochromes P450 heme domains of known structure designated by their PDB codes

R. jostii features one homolog of CYP51B (*ro04671*), which has been assigned the role of a sterol demethylase in strains of *Aspergillus*[280]; four homologs of CYP125 (*ro04679*,

ro04667, *ro02651* and *ro02355*), including one that has been shown to initiate side-chain degradation in steroids through C26 hydroxylation (Figure 6.2 **1**)[224], and one of CYP257A1 (*ro11069*), which has been shown to catalyse the *N*-demethylation of the alkaloid dextromethorphan (Figure 6.2 **3**)[225].

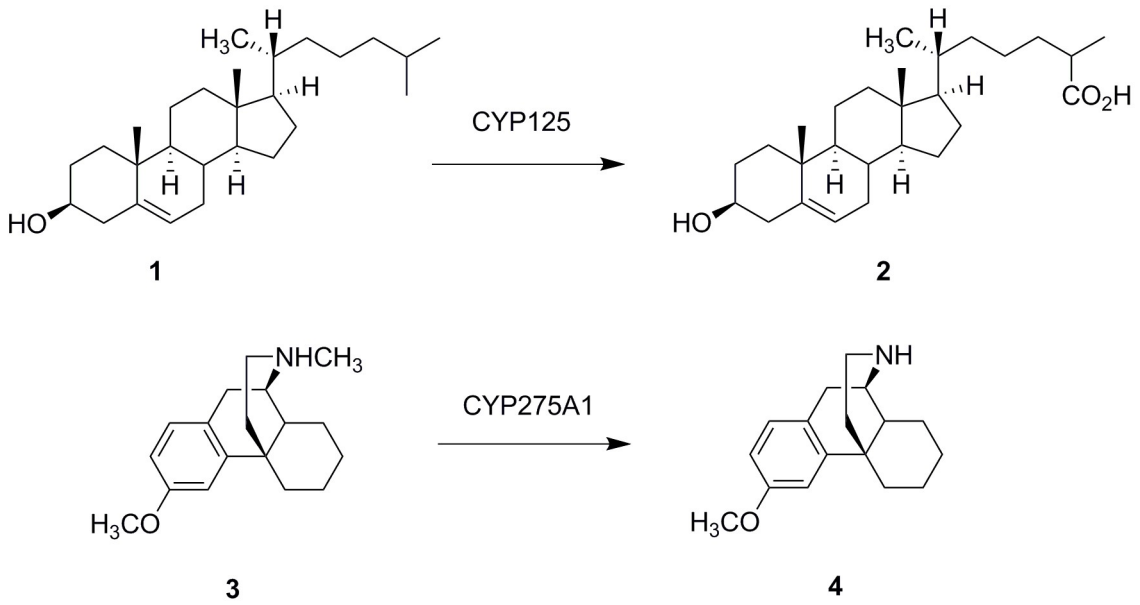


Figure 6.2: Oxidative activities attributed to P450 from Rhodococcus jostii RHA1
CYP125 catalyses the oxidation of sterols such as cholesterol **1** at the C26 position; CYP257A1 catalyses the N-demethylation of the alkaloid dextromethorphan **3**

Based on the phylogenetic analysis Ralph Hyde (University of York) created a P450RHA1-RhfRED fusion library for 23 of the *R. jostii* RHA1 targets by applying the LICRED technology (2.5.6) using genomic DNA as a template for the amplification. Primers for cloning can be found in Appendix A (Table A.4). Further investigations of the fusion library were commenced in this study.

6.2 Material and Methods

6.2.1 Screening for drug metabolites using resting whole cells

All whole-cell biotransformation tests were carried out with resting cells in buffer A (50 mM Tris/HCl, 300 mM NaCl, pH 7.5). 1 mM stock solutions of all substrates were prepared in methanol prior to biotransformation experiments. Biotransformations of freshly harvested cells with a cell wet weight of 100 mg/ml were carried out by addition of 12 µM (3.6 µl) of the relevant stock solution to give a total volume of 300 µl.

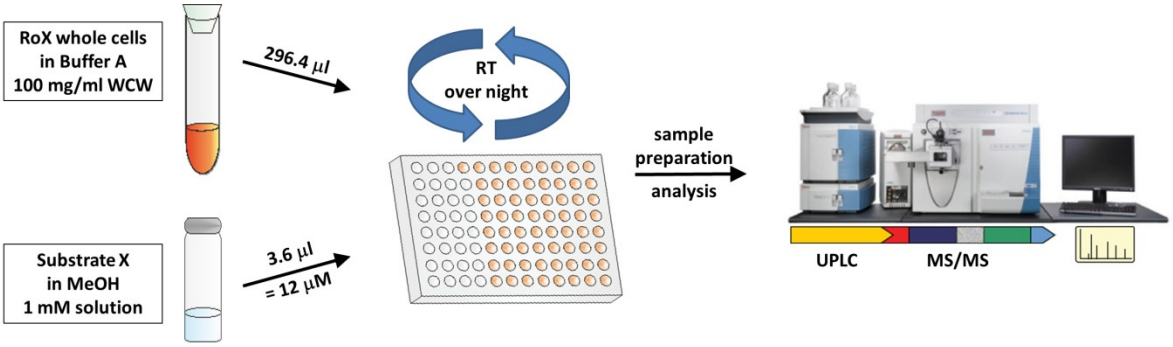


Figure 6.3: experimental setup for *E. coli* whole cell activity assays and UPLC-MS^E analysis
RoX: *E. coli* strain transformed with LICRED containing *R. jostii* RHA1 heme domain X; resting whole cells with a cell wet weight of 100 mg/ml are incubated with 12 µM of substrate X in 96-well plates; biotransformations of cells expressing the empty LICRED vector, overnight reactions of substrate X in buffer A and 40 % acetonitrile serve as negative controls

Reactions were incubated with constant vigorous shaking at room temperature in deep 96-well plates (Thermo Scientific Nunc, USA) of 1 ml well volume. Samples of 50 µl were taken at intervals: t = 0 h, t = 3 h and t = 18 h and transferred to a conical 96-well plate (Thermo Scientific, Denmark). The biological material was precipitated by the addition of 100 µl acetonitrile and the samples centrifuged at 4000 r.p.m. (Hettich, Germany) at 4° C for 20 min. 50 µl of the supernatant was then transferred into 150 µl 40% (v/v) acetonitrile in water and analysed by UPLC-MS^E (6.2.4). Figure 6.3 illustrates the experimental setup described above.

6.2.2 Screening for drug metabolites using purified Ro07-RhfRED

150 µM NADPH was added to a solution of pure Ro07-RhfRED (0.8 mg/ml, 9.4 µM) containing 12 µM of the drug compound, added from a 1 mM stock solution in methanol, to give a final reaction volume of 100 µl. Each reaction was incubated with shaking at room temperature (22–25°C) in deep 96 well plates (Thermo Scientific Nunc, USA) of 1 ml well volume. Samples were taken at time point t = 0 and t = 18 h. 25 µl of the sample was then removed and the biological material precipitated by the addition of 50 µl 100% acetonitrile, after which the sample was centrifuged at 4000 r.p.m. (Hettich, Germany) for 20 min at 4 °C. 25 µl of the supernatant was then transferred into 75 µl 40% (v/v) acetonitrile in water and analysed by UPLC-MS^E (6.2.4).

6.2.3 *In vitro* activity assays towards imipramine using purified Ro07-RhfRED

The activity of Ro07-RhfRED was tested at different concentrations of imipramine ranging from 1-50 µM. All reactions were carried out in the biotransformation buffer A (50 mM Tris/HCl, 300 mM NaCl, pH 7.5). Reactions were prepared as listed in Table 6.1. Briefly, 0.15 mM of NADPH was added to 100 µl volume of buffer A containing 1 mg/ml of the purified enzyme. 1–50 µM of imipramine was then added from a 0.06 mM stock in buffer A. The final reaction volume of 1 ml was obtained by adding buffer A. The reactions were incubated at 30 °C with constant shaking at 500 r.p.m.. Reactions were carried out in duplicates and reaction progress was measured at intervals t = 0 and 12 h. Reactions were stopped by addition of pure acetonitrile followed by the removal of precipitated proteins by centrifugation at 4 °C for 20 min at 4000 r.p.m... Subsequently, the reaction mixtures were quenched using solution of 40% acetonitrile (v/v) in water and analysed by UPLC-MS^E (6.2.4).

Table 6.1: experimental set up for imipramine kinetic assay with purified Ro07-RhfRED

Sample name	Imipramine (Conc. in µM)	Enzym (10 mg/ml) (in µl)	NADPH (3 mM) (in µl)	BufferA (in µl)	Substrate (0,06mM) (in µl)
control (no enzym)	50	0	50	100	850
control (no NADPH)	50	100	0	50	850
1	0	100	50	850	0
2	1	100	50	833	17
3	2	100	50	816	34
4	3	100	50	799	51
5	4	100	50	782	68
6	5	100	50	765	85
7	7	100	50	731	119
8	10	100	50	680	170
9	25	100	50	425	425
10	50	100	50	0	850

6.2.4 Analysis using UPLC-MS^E and data processing

Samples (10 µL) were injected and analysed using a Waters ACQUITY UPLC liquid chromatography system coupled to a Waters Synapt HDMS instrument (Waters, Milford, MA, USA) and equipped with an electrospray ionisation (ESI) source. Chromatographic separation was carried using an ACQUITY UPLC BEH C18 column (130 Å, 1.7 µm × 2.1 mm × 100 mm; Waters, Milford, MA, USA) at flow rate of 0.5 ml/min and column temperature of 45 °C. For separations using liquid chromatography mobile phases

consisting of ultra-pure water supplemented with formic acid (0.1% v/v; mobile phase A) and pure acetonitrile (mobile phase B) were employed. The gradient applied for separation was as follows: 0.0-6.0 min (10–70% mobile phase B); 6.0-6.7 min (70–90% mobile phase B), followed by a return to the initial mobile phase composition over 0.01 min.

The MS^E analysis was run on a Waters Synapt HDMS operating under positive electrospray ionization (ESI) conditions in V-mode. A generic method with two scan functions was used as follows: m/z 80–1000, cone voltage 20 V and 0.1 s scan time, the trap collision energy (CE) in function 1 was 20 V and in function 2 an energy ramp of 15–45 V was applied, the transfer cell CE was 12 V. Data were collected in a centroid mode. Leucine-enkephaline was used as a lock mass (m/z 556.2771) for internal calibration at a concentration of 250 pg/ml and a flow rate of 0.04 ml/min.

The MS^E data were processed in MetaboLynx 4.1 (Waters, Milford, MA, USA) using both the mass defect filter (MDF) and the dealkylation tool.

6.3 Results

6.3.1 Expression test in *E. coli*

In order to confirm overexpression of selected P450s, small scale expression tests were performed as described in 2.7.1.1. The strain *E. coli* Rosetta 2 (DE3) (2.2.1) was used to express *R. jostii* P450s cloned into the LICRED vector. Different conditions in respect of growth medium and expression temperature were tested. Initially, expression was tested in 10 ml LB medium and confirmed *via* SDS-PAGE and Western blot (data not shown). Although the production of insoluble and partly even soluble protein could be verified in all 23 cases, expression in general proved to be more efficient using M9 minimal medium for growth. All results shown below are taken from Western blot analysis (2.7.4) with cells grown in M9 medium. The optimal expression temperature was determined by growing cells at 16˚C, 30˚C and 37˚C. Only experiments that took place in 16˚C showed any soluble expression. Figure 6.4 displays Western blot analysis for Ro08-RhfRED grown at varied temperatures as representative for the other *R. jostii* P450 fusions.

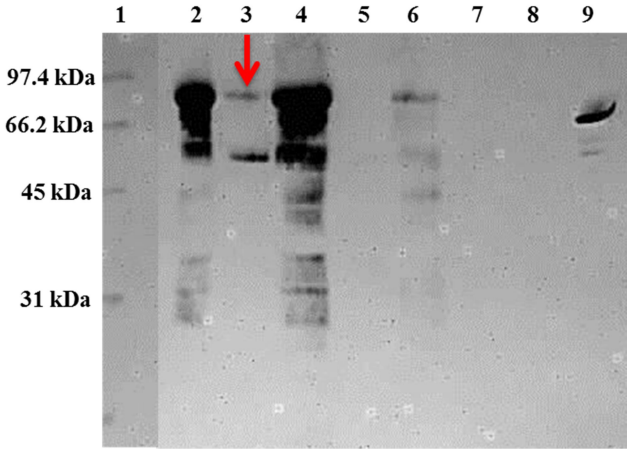


Figure 6.4: Western blot analysis of Ro08-RhfRED expression carried out at varied temperatures
lane 1: Low-Range-Marker; **lane 2, 4, and 6**: insoluble Ro08-RhfRED expressed at 16 ˚C (lane 2), 30 ˚C (lane 4), and 37 ˚C (lane 6); **lane 3, 5 and 7**: soluble Ro08-RhfRED expressed at 16 ˚C (lane 3), 30 ˚C (lane 5), and 37 ˚C (lane 7); lane 8: empty; lane 9: positive control.

All P450 fusions, except Ro03-RhfRED, were expressed insolubly and displayed bands of the expected height corresponding to a molecular weight of about 80 to 87 kDa in Western blot analysis. Soluble expression could be verified for Ro04-RhfRED, Ro05-RhfRED, Ro08-RhfRED, Ro11-RhfRED, Ro18-RhfRED and Ro22-RhfRED. However, exposure time whilst detecting Western blot was limited to the signal strength of the insoluble phase. These signals were usually fairly strong and masked the whole blot when exposed too long. Thus, weaker signals occurring in the soluble fraction could not be detected. Western blotting of all soluble phases was carried out in order to check for additional weaker signals (Figure 6.5 and Figure 6.6).

Significant signals of the predicted molecular weight (80 to 87 kDa) for soluble expressed fusions of RhfRED were detected by Western blot analysis for Ro01-RhfRED, Ro02-RhfRED, Ro04-RhfRED, Ro05-RhfRED, Ro06-RhfRED, Ro07-RhfRED, Ro08-RhfRED, Ro09-RhfRED, Ro11-RhfRED, Ro12-RhfRED, Ro16-RhfRED, Ro18-RhfRED, Ro20-RhfRED, Ro21-RhfRED and Ro22-RhfRED. Among these soluble expressed fusions, Ro01-RhfRED, Ro02-RhfRED and Ro09-RhfRED (Figure 6.5: lane 1, 2 and 11) as well as Ro16-RhfRED, Ro20-RhfRED, and Ro21-RhfRED (Figure 6.6: lane 6, 10 and 11) showed comparatively weak expression. The moderate (Ro04-RhfRED, Ro05-RhfRED, Ro06-RhfRED, Ro07-RhfRED and Ro11-RhfRED; Figure 6.5: lane 6, 7, 8, 9 and 13) or strongly (Ro08-RhfRED, Ro12-RhfRED, Ro18-RhfRED and Ro22-RhfRED; Figure 6.5: lane 10 and 14; Figure 6.6: lane 8 and 12) expressed genes on the other hand displayed evidence of

cleavage at the heme domain-reductase linkage illustrated by the blot response at approximately 45 and 55 kDa.

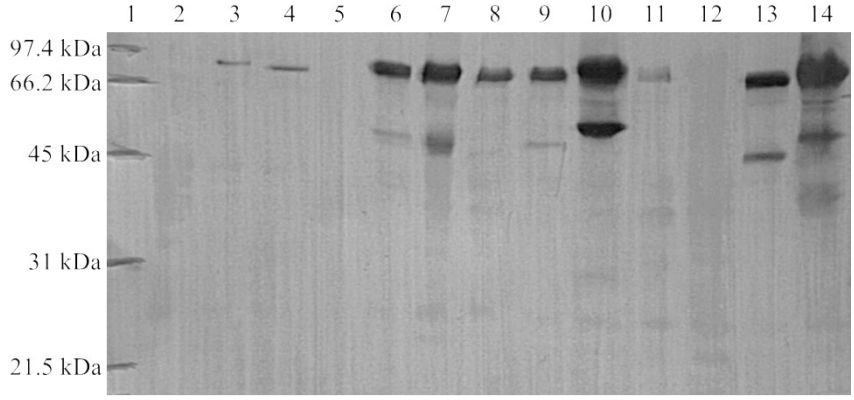


Figure 6.5: Western blot analysis of soluble expressed X-RhfRED fusions at 16°C
lane 1: Low-Range-Marker; lane 2: extract from non-induced cells; lane 3: Ro01-RhfRED; lane 4: Ro02-RhfRED; lane 5: Ro03-RhfRED; lane 6: Ro04-RhfRED; lane 7: Ro05-RhfRED; lane 8: Ro06-RhfRED; lane 9: Ro07-RhfRED; lane 10: Ro08-RhfRED; lane 11: Ro09-RhfRED; lane 12: Ro10-RhfRED; lane 13: R011-RhfRED; lane 14: Ro12-RhfRED

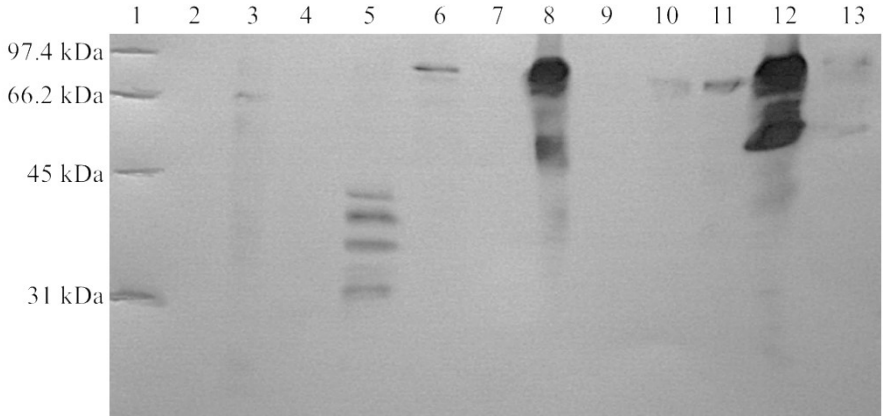


Figure 6.6: Western blot analysis of soluble expressed X-RhfRED fusions at 16°C
lane 1: Low-Range-Marker; lane 2: extract from non-induced cells; lane 3: Ro13-RhfRED; lane 4: Ro14-RhfRED; lane 5: Ro15-RhfRED; lane 6: Ro16-RhfRED; lane 7: Ro17-RhfRED; lane 8: Ro18-RhfRED; lane 9: Ro19-RhfRED; lane 10: Ro20-RhfRED; lane 11: Ro21-RhfRED; lane 12: Ro22-RhfRED; lane 13: Ro23-RhfRED

7 of the 15 recombinant cell strains expressing soluble fusion proteins were picked to be further analysed with regard to their oxidative activity as whole-cell biocatalysts in the biotransformation of drugs. Homologs of CYP51B, CYP125 and CYP257A1 (see introduction: 6.1) are represented by Ro22-RhfRED, Ro05-RhfRED and Ro07-RhfRED, respectively. In addition, Ro04-RhfRED, Ro08-RhfRED, Ro11-RhfRED and Ro18-RhfRED have been randomly selected for further investigation.

6.3.2 Screening for drug metabolites using resting whole cells

Whole cells were grown as described in 2.7.1.2. After cell harvest, the cell pellet was washed twice with buffer A (50 mM Tris/HCl pH 7.5, 300 mM NaCl) and was subsequently resuspended in a proportionated volume of buffer A to give a cell wet weight of 100 mg/ml. Whole cells were then challenged with 65 commercially available drug compounds (provided by AstraZeneca; see list: Appendix D), including alkaloids (*e.g.*, dextromethorphan), steroids (*e.g.*, ethinylestradiol), anti-inflammatory- (*e.g.*, amodiaquine, diclofenac, indomethacin), antidiabetic- (*e.g.*, pioglitazone, rosiglitazone) and cardiovascular agents (*e.g.*, clopidogrel, propranolol, verapamil) and natural antibiotics (*e.g.*, erythromycin). Cells containing only the empty LICRED vector as well as overnight drug reaction in cell free buffer A and acetonitrile were used as negative control in order to exclude background reactions by either *E. coli* internal biotransformation or as part of drug interactions with buffer or acetonitrile.

In order to obtain a first impression of the enzyme activity, cells were incubated with 12 µM compound as described in 6.2.1 and subsequently analysed using UPLC-MS^E (6.2.4). Each of the 7 recombinant cell strains performed biotransformations including hydroxylation and *N*-demethylation reactions. The most promising biotransformations are listed in Table 6.2. Corresponding structures are shown in Figure 6.7.

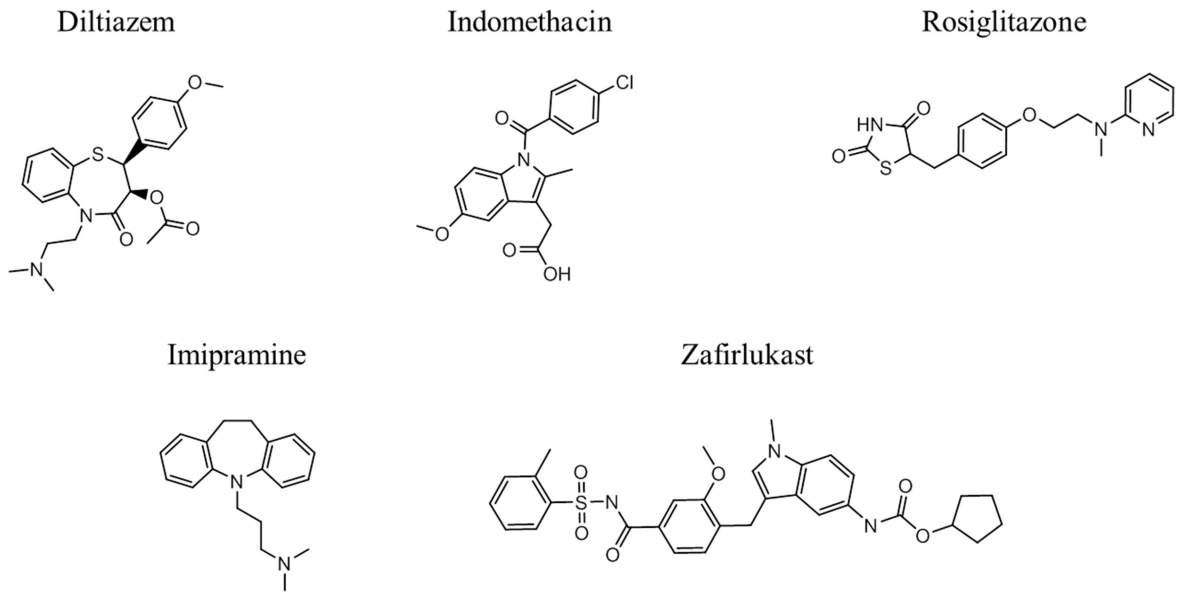


Figure 6.7: Drug molecules from the screen that showed significant levels of transformation against negative controls when incubated with P450RHA1-RhfRED fusions.

Table 6.2: Biotransformation spectra of heterologously expressed cytochrome P450 fusion proteins from *R. jostii* RHA1 in whole cells.
Biotransformations of cells expressing the empty LICRED vector, overnight reactions in buffer A and 40 % acetonitrile were included as negative control; grey areas signify biotransformation processes.

Biocatalyst X-RhfRED	Substrate				
	Diltiazem	*Indomethacin*	*Rosiglitazone*	*Imipramine*	*Zafirlukast*
	Conversion (%)				
	Demethylation	Hydroxylation	Demethylation	Demethylation	Hydroxylation
Ro04	**1.34**	–	**0.41**	**6.00**	–
Ro05	0.30	**5.67**	–	0.27	**3.02**
Ro07	**9.52**	–	**2.75**	**8.53**	-
Ro08	**1.07**	**11.41**	**0.44**	**0.56**	**6.15**
Ro11	**0.93**	**7.20**	**0.46**	**0.54**	**3.86**
Ro18	**1.45**	**8.32**	0.17	**0.73**	**6.24**
Ro22	**0.87**	**5.55**	**0.45**	**0.69**	**4.56**
LICRED	0.61	–	0.29	0.53	–
40% ACN	–	–	–	0.17	–
buffer A	0.18	–	–	0.23	–

With regard to enzyme activity it was noticeable that some of the fusions were more likely to perform hydroxylation reactions while others favoured *N*-demethylation reactions. Ro05-RhfRED, Ro08-RhfRED, Ro11-RhfRED, Ro18-RhfRED and Ro22-RhfRED hydroxylated indomethacin and zafirlukast with relatively high conversions but showed rather low *N*-demethylation activity towards diltiazem, rosiglitazone and imipramine. Ro07-RhfRED and Ro04-RhfRED on the other hand performed exclusively *N*-demethylation reactions. Ro07-RhfRED, however, was more successful in catalyzing demethylation reactions, giving a 9.5% and 8.5% conversion to demethylated metabolites of diltiazem and imipramine, respectively. It was, therefore, selected for a more detailed investigation, through studies of the pure enzyme.

6.3.3 Purification and characterisation of Ro07-RhfRED

Cells were grown to a total volume of 3 l and harvested as described in 2.7.1.2. Purification was mainly performed as described in 2.7.2. However, the imidazole gradient to elute Ro07-RhfRED from the column was optimised in accordance with previous purification trials. The loaded column was initially washed with about 12 column volumes buffer A. Impurities could be eluted by increasing the gradient of imidazole (within 20 min) from 0-20 mM (Figure 6.8: ascending green line). Finally, desired protein could be eluted from

the column by raising the imidazole concentration from 20 mM to 100 mM (Figure 6.8: descending green line) in a timeframe of 60 min.

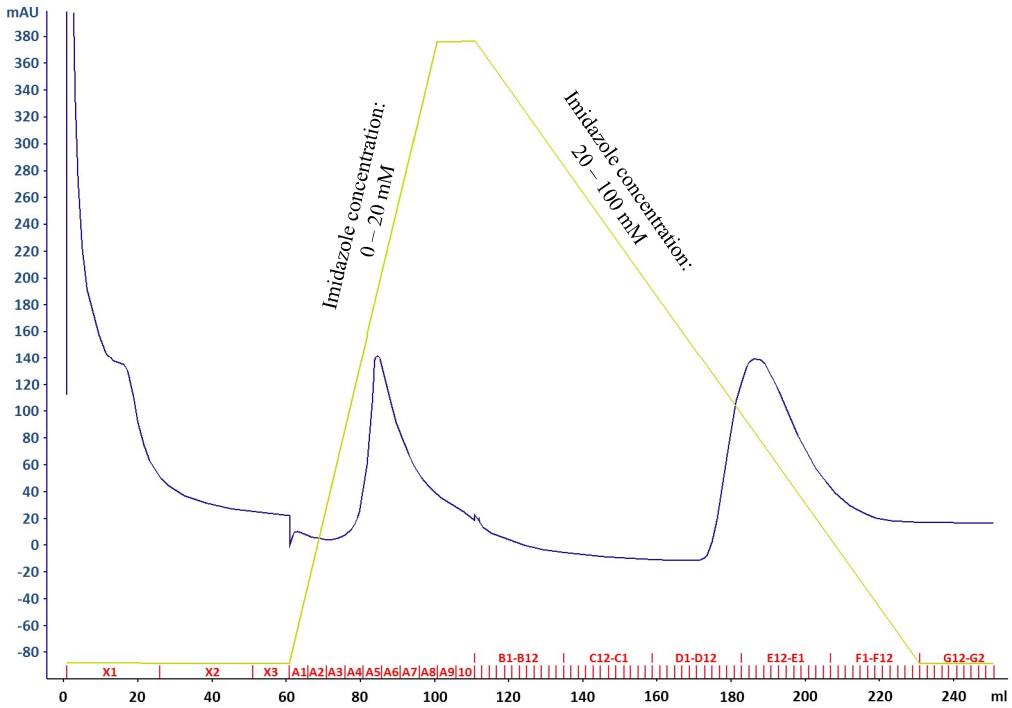


Figure 6.8: Nickel affinity chromatography chromatogram for the purification of Ro07-RhfRED
blue line: UV absorbance at 280 nm; green line: imidazole concentration gradient; Fractions are displayed along the x axis.

The chromatogram for the purification of Ro07-RhfRED is shown in Figure 6.8. In order to verify the purity of these fractions SDS gel electrophoresis has been carried out (Figure 6.9) and showed more than 90% purity.

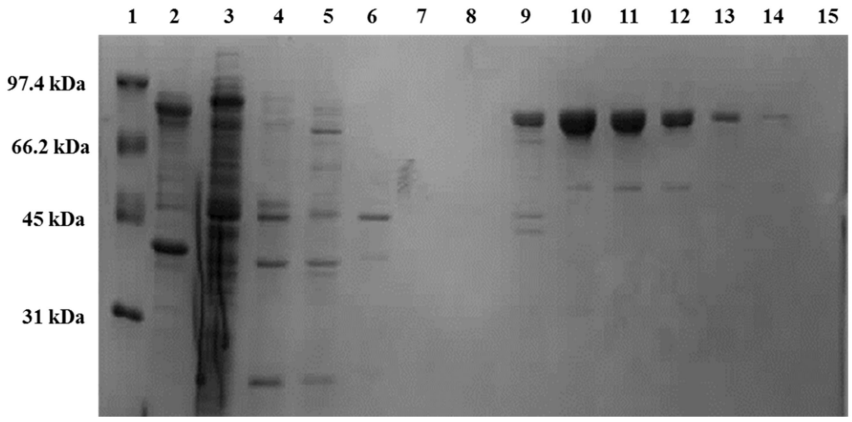


Figure 6.9: SDS-PAGE analysis of Ro07 RhfRED fractions after nickel affinity chromatography
lane 1: 1 kb DNA ladder; lane 2: insoluble fraction; lane 3: flow-through after column loading; lane 4-15: samples corresponding to the fractions X1, A5, A10, D7, D8, D11, E10, E7, E3, F2, F6, and F12 collected in nickel affinity chromatography (Figure 6.8).

In order to validate the integrity of the purified enzyme, CO-difference spectra were performed. Ro07-RhfRED was reduced with sodium dithionite and then bubbled with carbon monoxide. Spectrometric analysis was performed for every step subsequently, as described in 4.1.3.1.

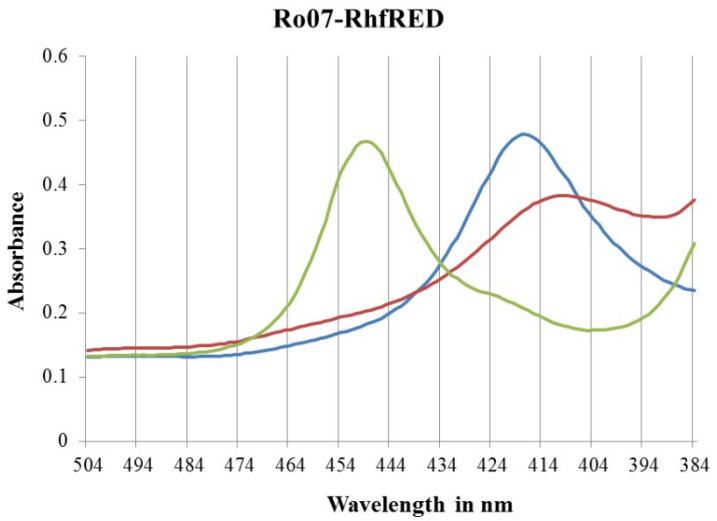


Figure 6.10: UV-visible absorbance spectra of purified Ro07-RhfRED
oxidised form: blue; sodium dithionite reduced form: red; reduced, CO-bound form: green

Figure 6.10 displays the recorded CO-difference spectra for Ro07-RhfRED. Unreduced P450s have typically an absorption maximum at ~ 420 nm which shifts of about 30 nm to 450 nm upon reduction and subsequent CO binding. This characteristic Soret shift could be observed in all tested fractions containing purified protein, indicating that the chimeric fusion was correctly folded and had the ability to bind substrates with high affinity. Furthermore, the Soret shift could still be observed after 2 weeks of storage of the purified protein in the fridge and thus suggest high stability of the fusion protein. Comparative pure fractions have then been pooled and concentrated for further use.

In order to validate the activity of the purified Ro07-RhfRED, it was challenged with the 65 drugs used for the whole cell screening (6.3.2). Assays were performed as described in 6.2.2. The results obtained with the purified Ro07-RhfRED confirmed its activity. Moreover, the conversion of imipramine to the *N*-demethylated product almost doubled when compared to the screening results with whole cells. Imipramine was therefore selected for more detailed investigation of the catalytic properties of the Ro07-RhfRED fusion.

6.3.4 Biotransformation of imipramine

In order to exclude possible negative effects of additives like methanol on the enzyme, stock solutions of substrate, NADPH as well as purified Ro07-RhfRED were prepared in buffer A. Samples for conversion of imipramine (Figure 6.11 **B 1**) to *N*-desmethyl imipramine (Figure 6.11 **B 2**) were prepared as described in 6.2.3.

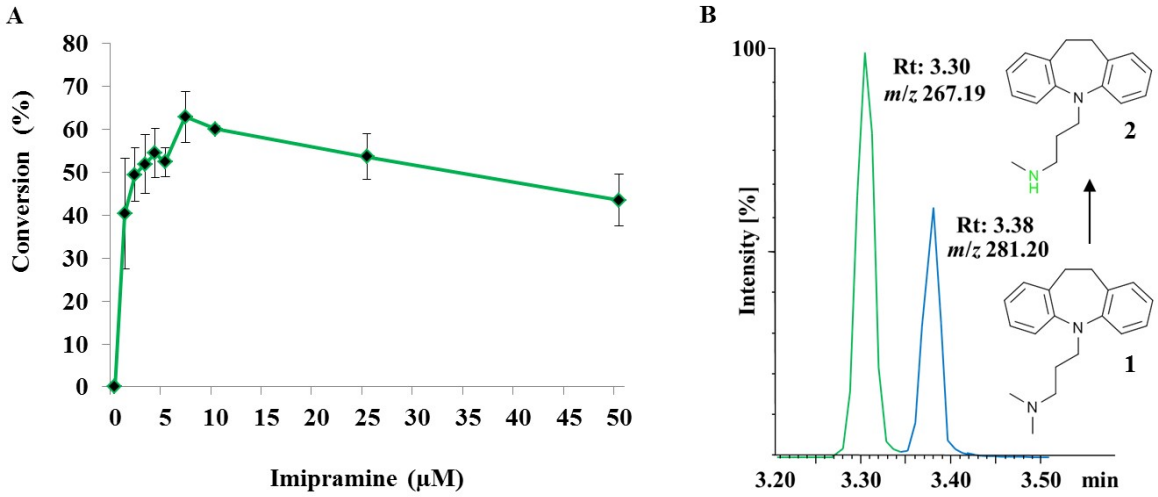


Figure 6.11: Biotransformation of imipramine with purified Ro07-RhfRED
A: Conversion (%) of imipramine after 12 h by purified Ro07-RhfRED at different substrate concentrations; B: UPLC/MS^E chromatogram of imipramine (25 µM) conversion after 12 h; blue: imipramine (**1**); green: demethylated metabolite (**2**)

Conversion rates of up to 60 % could be detected at low substrate concentration (Figure 6.11 **A**: 7 µM). Imipramine conversion decreased once its concentration exceeded 7 µM indicating substrate inhibition. However, conversion of high rates still took place with imipramine concentration of 50 µM (40 %) and led to the best yield of *N*-desmethyl imipramine overall. Furthermore, *N*-demethylation did not only depend on the substrate concentration but was also influenced by enzyme concentration. Conversion experiments performed by Justyna Kulig in AstraZeneca using only 0.5 mg/ml enzyme in a total reaction volume of 300 µl (0.15 mM NADPH) led to a dramatic decrease in conversion. The highest conversions of only 36.2 % and 35.4 % were measured for imipramine concentrations of 2.5 µM and 5 µM after 4 h incubation. In addition, no further conversion could be observed after the 4 h incubation.[279]

6.4 Discussion

Characterization of genomic sequences and thus identification of novel P450 enzymes has advanced rapidly in the last 20 years and created new challenges with regard to the

elucidation of these newly discovered proteins. New sequences are arising much faster than gene product function can be validated. *R. jostii* RHA1, for example, has a remarkable diversity of genes that encode oxidative enzymes of which the majority are, however, of unknown function. Inspired by natural P450-redox fusion proteins in which the electron supply components are fused to the P450 catalytic domain, researchers have focused on generating non-natural fusions of P450s as dealing with a single enzyme provides advantages in comparison to a multiprotein system and can simplify high throughput screening and therefore characterization procedures. $P450_{BM3}$ (CYP101A2) from *Bacillus megaterium*[26] and $P450_{Rhf}$ from *Rhodococcus* sp. NCIMB 9784 (CYP116B2) [79] are well known examples of such naturally occurring fusion systems. The $P450_{BM3}$ system, however, is not a very suitable tool for high throughput analysis as it relies on intermolecular electron transfer between two monomers and thus is only active as a dimer.[281] Chimeric fusions with BM3 and non-native heme domains show often a high degree of uncoupling[282] most likely owing to essential specific interactions between the domains of the two monomers not being formed. Chimeric fusions with RhfRED on the other hand were shown to transfer electrons from NADPH primarily intramolecularly to the P450 heme domain.[227]

The first report of a non-natural fusion with RhfRED was by Misawa and colleagues who produced functional chimeras that acted on the natural substrates.[263] Other reports employing non-natural fusions of RhfRED with varying degrees of success were to follow. However, before the development of the LICRED vector by Sabbadin and colleagues, no methods had been available to generate libraries of P450 biocatalysts in a high throughput manner.[227] The LICRED vector contains the RhfRED gene downstream of a ligation independent cloning site, allowing for rapid insertion of P450 genes and generation of large libraries of P450-RhfRed fusions. Although, the utility of this system was already demonstrated by fusions with P450cam, XplA and 22 novel P450s isolated from *Nocardia farcinica*, [227] it provided the perfect basis to examine heterologous expressed genes for useful activities with simultaneous re-evaluation of the systems potential.

By employing the LICRED system it was possible to create a fusion library of 23 cytochrome P450 heme domains from *R. jostii.* 15 of them could be expressed soluble in *E. coli* Rosetta (DE3). 7 of these have been analysed in more detail in whole cell assays and revealed many different activities including hydroxylation and demethylation reactions when challenged with 65 commercially available drugs and thus proved the generic

applicability of the LICRED platform. One recombinant strain, expressing Ro07-RhfRED (homolog of CYP257A1), catalysed the *N*-demethylation of diltiazem and imipramine and could be successfully purified and further characterized. The *N*-demethylation activity presented by Ro07-RhfRED complies with previous observations of this enzyme‘s activity as a demethylase of alkaloid substrates (dextromethorphan).[225] Dextromethorphan itself, however, was not transformed in significant quantity by the Ro07-RhfRED fusion protein in whole cells or purified form. The reason for this might be the fused reductase domain itself which could have affected substrate specificity of the enzyme, as has been observed for $P450_{MycG}$.[283] Nevertheless, Ro07-RhfRED was able to convert up to 63 % of imipramine when applied in cell-free biotransformations although its activity seemed to be strongly dependent on the substrate and enzyme concentration. Dependence on substrate concentration could also be demonstrated by Justyna Kulig through analysis of the biotransformation of diltiazem by cell-free Ro07-RhfRED (Figure 6.12).[279]

As with imipramine most efficient reactions were obtained at low substrate concentrations of between 1.0 (26.0 %) and 7.5 µM (25.6 %), but the highest turnover was measured for concentration of 50 µM. Furthermore, no conversion was observed after 4 h either due to substrate inhibition or activity loss of the protein.

Besides substrate screening it was also possible to investigate parameter of the *N*-demethylase activity of Ro07-RhfRED in more detail as also shown by Justyna Kulig (Figure 6.13).

The activity of Ro07-RhfRED towards analogues of imipramine demonstrated that the exchange of different chemical groups can have enormous influence on the *N*-demethylase activity. The introduction of hydroxy groups (Figure 6.13 **2** and **7**) or the presence of a second side chain in the substrate (Figure 6.13 **10**) and the corresponding rise of the polar surface area of the substrate, for example, reduced conversion rates dramatically. Substrate alteration that led to a lower polar surface area on the other hand didn’t seem to hinder *N*-demethylation activity as drastically. In fact, the introduction of a methyl group in the alkyl chain (Figure 6.13 **3**) or a chlorine atom on the aromatic ring (Figure 6.13 **4**) had only little influence on the activity of Ro07-RhfRED. Furthermore, it could be observed that the chimeric fusion possesses clear stereoselective preferences when tested on enantiomeric substrates (Figure 6.13 **8** and **9**).

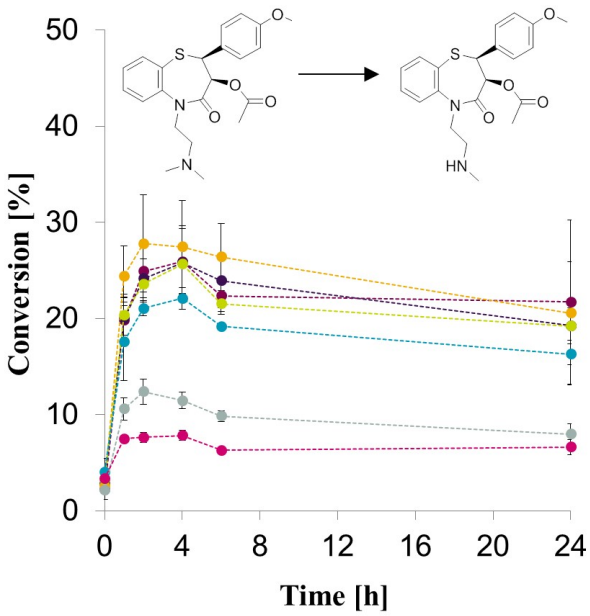


Figure 6.12: Biotransformation with purified Ro07-RhfRED of diltiazem at different substrate concentrations[279]
substrate concentrations: – • – 1.0 µM, – • – 2.5 µM, – • – 5.0 µM, – • – 7.5 µM, – • – 10.0 µM, – • – 25.0 µM, – • – 50.0 µM

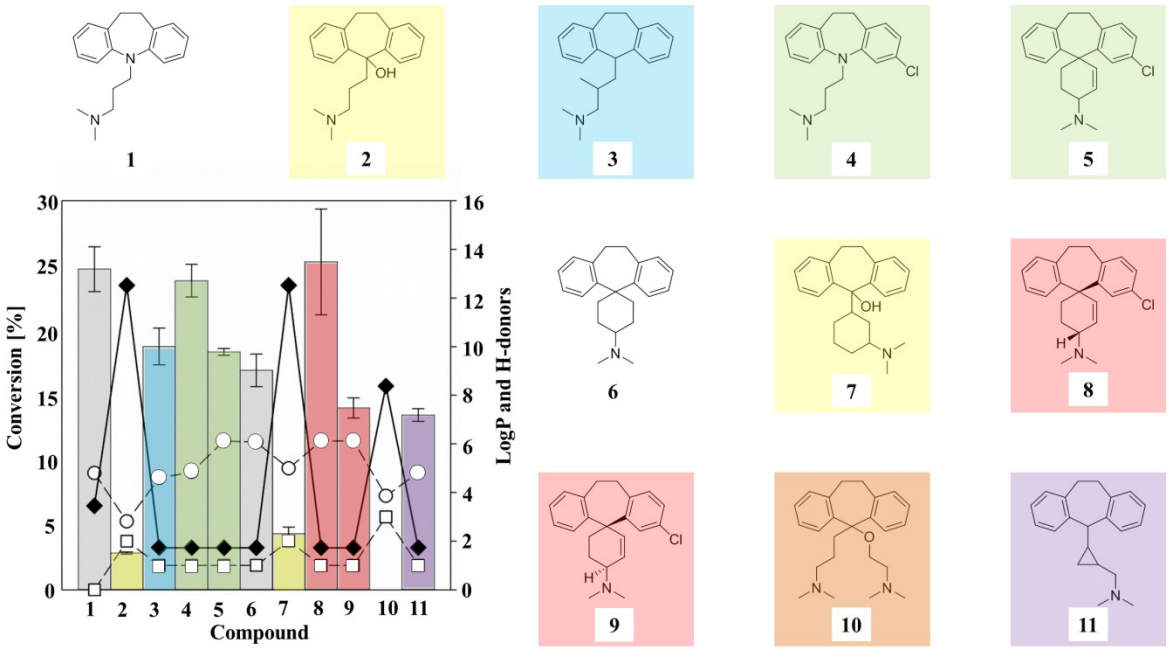


Figure 6.13: Conversions of imipramine 1 and analogs 2-11 and their relationship to physico-chemical properties of the substrates catalysed by Ro07-RhfRED [279]
♦ = polar surface area (PSA); ○ = octanol-water partition coefficient (logP); □ = H-donors; yellow: negative influence of hydroxyl groups; green: neutral influence of chlorine; blue: slightly negative influence of additional methyl group; purple: 2-fold decrease with additional cyclopropane group; red: stereopreference; orange: no reaction with additional side chain

This study could clearly substantiate the applicability of the by Sabbadin and colleagues conceived LICRED plasmid as strategy to create and screen libraries of diverse P450 fusions. Although the substrate specifity for natural substrates might have been affected by the fused reductase domain, as shown for Ro07-RhfRED, this system led to the identification of several enzymes from *R. jostii* competent to hydroxylate or demethylate various substrates of industrial interest and might be further applied in industry in order to screen for activities of use in other compounds of interest to medicinal chemistry. It could be further demonstrated that physico-chemical parameters like the polar surface area correlate with conversion rates and thus might be useful data for future selection of substrates prior to *in vitro* evaluation. However, the LICRED platform serves only as a screening tool for identifying useful P450 activities. Scale up processes of industrial relevance might need to be optimised for individual chimeras.

7. Final discussion

Enzymes - a precious resource of incomparable value for industry, accounting for more than 500 industrial products today.[284] Even though the capacity of natural catalysts (enzymes) has been harnessed by humans for thousands of years, dating back to 7000-6600 BCE with fermentation of sugars to alcohol by yeast[285], it wasn't until the late 19th century - spawned by the awareness of the existence of enzymes - that the curiosity and interest to elucidate biochemical pathways and enzyme mechanisms led to the development of enzyme-directed research. The implementation potential of natural catalysts transforming non-natural organic compounds in industry however, was mainly recognized during the 1980s, with the rise of enantioselective synthesis and the advance in molecular biology.[286] Since then, the number of on industrial scale performed biotransformations grew rapidly and is expected to continue growing as the demand for novel enzymes rises, impelled by a growing need for sustainable solutions.

Oxygenases in particular are of enormous interest and are often used to synthesize compounds that are not accessible by chemical routes.[132] Among them, cytochrome P450s attracted probably the most attention due to their involvement in the biosynthesis of a broad range of bioactive natural products in all kingdoms of life, often mediated with remarkable regio- and stereoselectivity. Most P450s catalyse monooxygenase reactions with the assistance of a redox partner by utilizing molecular oxygen and a reducing co-factor.[109] P450s that fulfil distinct roles in primary metabolism act with high catalytic efficiency and refined substrate selectivity abetting for the classical idea that enzymes have evolved to catalyse one particular chemical reaction. The majority of, and most likely the most striking P450s however, show more substrate ambiguity, usually as part of a protection mechanism to remove extrinsic natural products and thus allowing for an exceptional catalytic versatility.[287] P450s that mediate drug metabolism in the human liver, for example, show a remarkable substrate ambiguity and facilitate the clearance of almost 75 % of prescribed drugs.[288] Also many fungal P450s are suspected to act as multifunctional enzymes.[289] Although nature provides a vast amount of enzyme resources, the discovery and development of a suitable cytochrome P450 for biocatalysis is a demanding and time-consuming procedure. The classical biochemical approach represents a rather straightforward strategy starting with *in vivo* observations. These are then replicated in an *in vitro* assay that can be used to purify the enzyme of interest which in turn can be employed to obtain cDNA and determine the nucleotide sequence to identify

the corresponding gene.[290] However, enzyme discovery today has been revolutionised by various genome sequencing programs that led to an abundance of available nucleotide sequence data and thus resulting in the need for strategies to work backwards from gene to function. Especially microorganism that have already been applied in an industrial context or shown interesting reactions for industrial purposes are promising for more detailed investigation. In this context, this study concentrated on the exploration of the CYPome of the entomopathogenic fungus *B. bassiana* which has been sequenced in 2012[208] and was once one of the most frequently used whole cell biocatalysts.[219] A further area of focus was the analysis of 23 P450s from *Rhodococcus jostii* which had already been constructed as chimeric fusions by Ralph Hyde using the LICRED system.

83 sequences that were annotated as cytochrome P450 enzymes were subjected to thorough bioinformatic investigation resulting in the selection of 7 genes that encode for putative CYPs with possible alkane hydroxylation function and 1 putative natural fusion enzyme ($P450_{foxy}$ homolog) for subsequent cloning, over-expression and activity screening. The fact that 2 of the selected cytochrome P450s shared 100% sequence identity with CYP52X1 and CYP5337A1, which have shown to be involved in insect hydrocarbon degradation,[203] allowed for the rather confident assumption that the selected P450s play a physiological role in *B. bassiana* and thus should yield active proteins in the event of successful expression. Although, heterologous expression of eukaryotic P450s using bacterial hosts such as *E. coli* can be very challenging due to the interference of the N terminal anchor, different codon bias as well as the need for an external NADPH-cytochrome reductase, it was in the interest of this study to apply a system that can be easily handled and thus is more approachable from an industrial point of view. Hence, modifications of the selected genes have been realized on molecular level (codon optimization for *E. coli* and cleavage of N-terminal anchor) to ensure optimal expression in *E. coli*. Furthermore, encouraged by the remarkable success of the artificial fusion system designed by Sabbadin and co-workers, which proved not only to be efficacious when applied for prokaryotic P450 but also for eukaryotic P450 expression in *E. coli*[227, 266], and in order to bypass the need for a co-expressed reductase, a fusion library was constructed for the 7 selected *B. bassiana* P450s. Expression test which were performed using M9 minimal medium with glucose as only carbon source resulted in rather low expression levels and proteins were only detectable in the insoluble fractions. Improvement of expression could be realized by altering the carbon source in the medium most likely due

to the avoidance of a phenomenon called catabolite repression that can lead to poor transcription when using a lac promoter.[250] Of the tested carbon sources glycerol seemed to have the biggest impact on expression performance. However, signals indicating soluble expression were still missing. The truncated version of the natural fusion (CYP505A1) as well as all chimeric fusions, except CYP584E2_LICRED, could finally solubly expressed using a chaperone co-expression system which confirmed previous reports stating similar findings.[252, 270] However, functional analysis using whole cells for biotransformation of substrates that are known to be hydroxylated by *Beauveria bassiana* or/and alkane hydroxylases failed to determine P450 identity. Failed activity assays that focused on lauric acid as substrate have been of particularly significance to confirm the shortage of active biocatalysts as CYP52G8 (identical to CYP52X1) has been reported to convert lauric acid to 12-hydroxylauric acid (Figure 7.1).[217] In addition, spectrophotometric characterization of whole cells and concentrated *E. coli* cell lysate confirmed activity assay results as characteristic Soret shift peaks couldn't be identified either. Purification of soluble expressed chimeric fusions wasn't possible due to the strong interaction with the co-expressed chaperones.

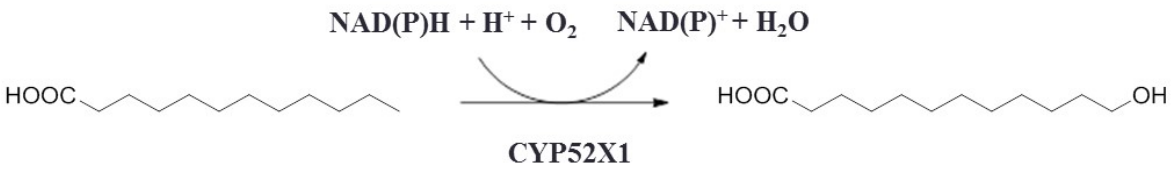


Figure 7.1: Biotransformation of lauric acid using microsomes from yeast expressing *CYP52X1*

In an attempt to preclude chaperone interference with protein activity and the impairment of the fused Rhf-reductase on solubility, 2 selected *B. bassiana* P450s have been cloned and expressed without the addition of chaperones and reductase. In the event of soluble expressed protein, strains co-expressing the natural *B. bassiana* reductase, which has been identified in this study (CYP505A2), could have been generated for detailed investigation of protein activities. However, soluble expression couldn't be detected leading to the assumption that *E. coli* is not a suitable host for the production of active fungal P450s. Further changes of the genes on the molecular level are most likely necessary in order to enhance soluble expression of active proteins using *E. coli* as a host.

As a result of the abortive heterologous protein expression in *E. coli*, *S. cerevisiae* has been employed for expression of *Beauveria* heme domains. There was good reason to believe

that this system would allow for expression of functional membrane associated *B. bassiana* P450s in microsomes since its functionality had been demonstrated in previous studies.[217, 274] In order to ensure functional expression it was considered important to recreate the parameters used in these studies. Hence, expression strain (*S. cerevisiae* WAT11), vector (pYeDP60) as well as conditions have not been varied. Each of the 7 selected *B. bassiana* heme domains were constructed with the addition of N-terminal or C-terminal his-tag (18 bp histidine tag). In addition, constructs that have been generated containing no his-tag were thought to serve as comparison group to evaluate possible his-tag induced activity impairment. However, active fungal P450s could not be detected in *S. cerevisiae* microsomes most likely due to the codon bias in yeast as the gene sequences used for expression in this study derived from the *E. coli* codon optimised gene sequences while Zhang and co-workers employed fungal cDNA for PCR amplifications.

One of the most challenging issues in this project was the management of a multitude of enzyme targets, particularly with regard to gene expression as investigation of individual requirements for functional protein production could not be realized. A more reasonable tactic to identify less P450s but P450s of particular interest might be a proteomic approach allowing for the identification of a protein in correspondence to a transformation reaction. First steps in this direction could be achieved at the end of this project. The original *B. bassiana* strain has been exposed to various hydroxylation substrates and the corresponding concentrated *B. bassiana* cell lysates were then employed for high-resolution mass spectrometry in order to identify possible P450 targets responsible for these hydroxylation reactions (see Appendix C). P450 hits after LC-MS/MS analysis including the negative control are shown in Table C.1 (Appendix C). Although, only a total of 4 P450 enzymes could be identified, 2 of which were also observed in the negative control, it is noteworthy that CYP52T1 as well as the NADPH-cytochrome P450 reductase seem to be more strongly induced in *Beauveria bassiana* when 1-benzylpyrolidine (Table C.1, sample 2), lauric acid (Table C.1, sample 3) or 2-phenoxypropionic acid (Table C.1, sample 4) has been added to the growing fungus. This could therefore be a clue with regard to the physiological role of CYP52T1 in *B. bassiana*. It also needs to be mentioned that the acquired data are not quantitative as no replication took place. However, it could be demonstrated that this form of analysis could be useful for target selection in future studies especially with more refined sampling strategies. The only reason that this analysis

technology has not been used in the beginning of this project was the lack of available equipment and techniques.

Eventhough the study of *Beauveria bassiana* P450s did not lead to the delivery of biocatalysts applicable to industry it was possible to investigate 23 chimeric fusions (P450 heme domains cloned in LICRED vector) from *Rhodococcus jostii* in more detail. An extensive screen of expression conditions in respect of growth medium and expression temperature were conducted using *E. coli* Rosetta (DE3) as host. Soluble expressed proteins could be detected at 16 ˚C (Western blot) in 15 out of 23 cases using M9 minimal medium for growth. This result is not too surprising as it has been previously reported that recombinant protein expression at reduced growth temperatures can increase solubility of proteins that are prone to aggregation.[291] 7 of the P450 fusions have been further analysed using whole cells for biotransformation of substrates of industrial interest. Each of the strains expressing a chimeric fusion protein showed activity towards at least 1 or 2 of the tested substrates and substantiated the potential of the applied LICRED system as screening method for enzyme discovery. The fusion protein expressed in the *E. coli* strain with the most promising catalytic activity was further characterized in purified form and revealed interesting catalytic properties, such as the activity dependency on substrate and enzyme concentration as well as the correlation of physico-chemical parameters with conversion rates. The here presented results have recently been published in *Bioorg. Med. Chem.* (doi:10.1016/j.bmc.2015.07.025) and might have an impact of future screening assays in industry.

In summary, it can be stated that although it wasn't possible to deliver active biocatalysts from *B. bassiana* for industrial implementation, data of empirical value could be obtained that might impact future studies in regard to heterologous expression strategies for fungal P450s. Furthermore, industrial partners have been provided with a valuable screening platform (*Rhodococcus* P450 fusion library) allowing for identification of biocatalysts capable of producing drug metabolites.

Appendix A

Table A. 1: Cytochromes P450 in *B. bassiana* ARSEF 2860

CYP family	Annotated CYP name	Length (aa)	NCBI Accession number
CYP503	CYP503B1	526	EJP67620.1
CYP504	CYP504E5	296	EJP67926.1
	CYP504A6	518	EJP66903.1
	CYP504B10	533	EJP63215.1
	CYP504E1	555	EJP67927.1
	CYP504B10	576	EJP66859.1
CYP505	CYP505A2	690	EJP66462.1
	CYP505D4	1044	EJP67978.1
	CYP505A1	1232	EJP64169.1
CYP5060	CYP5060A1	601	EJP70021.1
CYP5080	CYP5080B3	511	EJP68696.1
CYP5099	CYP5099A1	466	EJP64478.1
	CYP5099A1	522	EJP63211.1
CYP51	CYP51F2	516	EJP64120.1
	CYP51F1	525	EJP70591.1
CYP52	CYP52G6	528	EJP65762.1
	CYP52G8	528	EJP68426.1
	CYP52T1	534	EJP70096.1
	CYP52T1	536	EJP68380.1
	CYP52T1	536	EJP62931.1
	CYP52T1	541	EJP62098.1
CYP5202	CYP5202A1	475	EJP69240.1
CYP526	CYP526H1	602	EJP63805.1
CYP5262	CYP5262A3	491	EJP62710.1
CYP528	CYP528A4	517	EJP65064.1
CYP5280	CYP5280A1P	526	EJP66155.1
CYP5282	CYP5282A1	528	EJP67855.1
CYP5293	CYP5293A1	364	EJP64645.1
	CYP5293A2	470	EJP61325.1
	CYP5293A1	479	EJP66212.1
CYP53	CYP53A11	513	EJP61564.1
CYP534	CYP534C2	286	EJP68773.1
CYP537	CYP537A4	429	EJP62906.1
CYP539	CYP539B1	515	EJP65570.1
CYP540	CYP540B16	550	EJP67254.1
CYP541	CYP541A2	463	EJP70012.1
CYP542	CYP542B2	138	EJP61051.1
	CYP542B1	199	EJP61998.1
	CYP542B3	540	EJP61494.1
	CYP542B3	540	EJP68382.1
CYP548	CYP548A5	530	EJP69550.1
CYP551	CYP551C1	523	EJP62242.1
CYP56	CYP56C1	493	EJP61452.1

CYP family	Annotated CYP name	Length (aa)	NCBI Accession number
CYP561	CYP561D2P	512	EJP62441.1
	CYP561D2P	525	EJP68106.1
CYP570	CYP570E2	460	EJP69670.1
	CYP570A1	490	EJP64736.1
	CYP570H1	496	EJP64697.1
CYP578	CYP578A2	522	EJP63651.1
CYP58	CYP58A3	526	EJP62372.1
CYP584	CYP584D4	496	EJP65354.1
	CYP584G1	510	EJP66430.1
	CYP584E7	523	EJP66886.1
	CYP584E2	528	EJP68333.1
CYP586	CYP586B1	546	EJP61887.1
CYP6001	CYP6001C8	1117	EJP67251.1
CYP6003	CYP6003A1	1121	EJP69120.1
CYP6004	CYP6004A2	1070	EJP66691.1
CYP61	CYP61A1	534	EJP64151.1
CYP617	CYP617A2	520	EJP68630.1
	CYP617A1	547	EJP65976.1
	CYP617A1	549	EJP67360.1
	CYP617A1	566	EJP62088.1
CYP620	CYP620C2	531	EJP70368.1
	CYP620D1	537	EJP61298.1
CYP621	CYP621A2	519	EJP68308.1
CYP623	CYP623C1	490	EJP63692.1
CYP625	CYP625A1	516	EJP64344.1
	CYP625A1	522	EJP67065.1
CYP628	CYP628A2	537	EJP70675.1
CYP639	CYP639A3	520	EJP67629.1
CYP645	CYP645A1	550	EJP69853.1
CYP65	CYP65T7	497	EJP62380.1
	CYP65T7	497	EJP62380.1
CYP655	CYP655C1	529	EJP63691.1
CYP660	CYP660A2	535	EJP64784.1
CYP68	CYP68N1	508	EJP68182.1
CYP682	CYP682H1	509	EJP67697.1
	CYP682N1	530	EJP67851.1
	CYP682H1	544	EJP67232.1
CYP684	CYP684A2	483	EJP64479.1
	CYP684A2	490	EJP63437.1
	CYP684B2	509	EJP66339.1

Table A.2: Cytochromes P450 in *Rhodococcus jostii* RHA1

Number	*ro number*	Accession Number	Annotation
Ro01	02510	YP_702473.1	147B2
Ro02	02948	YP_702911.1	199A3
Ro03	04669	YP_704613.1	123A5

Number	*ro number*	Accession Number	Annotation
Ro04	02604	YP_702567.1	105Y1
Ro05	04667	YP_704611.1	125B1
Ro06	02651	YP_702614.1	125C1
Ro07	11069	YP_708874.1	257A1
Ro08	11320	YP_709125.1	116C2
Ro09	04588	YP_704532.1	142A4
Ro10	00423	YP_700417.1	254A1
Ro11	11277	YP_709082.1	258A1
Ro12	00377	YP_700371.1	254A3
Ro13	08608	YP_707273.1	116C1
Ro14	05719	YP_705651.1	130A4
Ro15	00393	YP_700387.1	254A2
Ro16	02355	YP_702318.1	125D1
Ro17	04679	YP_704623.1	125A14
Ro18	08984	YP_708186.1	256A1
Ro19	02382	YP_702345.1	255A2
Ro20	03876	YP_703834.1	255A1
Ro21	03826	YP_702614.1	124C1
Ro22	04671	YP_704615.1	51B
Ro23	04627	YP_704571.1	102B3
Absent	05210	YP_705149.1	136C1
Absent	03076	YP_703037.1	136C2

Table A.3: Primer pairs used for PCR amplification of N-terminal anchor and truncated P450 heme domain from *B. bassiana*

CYP number	Target area	Forward Primer (5'-3')	Reverse Primer (5'-3')
CYP52G6	N-terminal anchor	CTAAATTACCGGATCCATGGCGCTCACTGCTATCC	GCTTTCCAGAACACGTCGAAGCACAAAGGTGACGAC
	truncated heme domain	CGTGTTCTGGAAAGCCTGC	GCGAATTCGAGCTCGGTACCTTACACTGCTTCATGCAGACG
CYP52G8	N-terminal anchor	CTAAATTACCGGATCCATGGCGGTACTATCCGTTA	GCTACGAAAACGAACGTACTCAATCACCTGGAGG
	truncated heme domain	GTTCGTTTTCGTAGCAAAGCAG	GCGAATTCGAGCTCGGTACCTTATTCATCAAAGTGAACTTTCAGG
CYP52T1	N-terminal anchor	CTAAATTACCGGATCCATGGCTCTCCACGCTGCC	ACGCTGAATGCTACGAGTCAGATAGACGGCCACC
	truncated heme domain	CGTAGCATTCAGCGTCGT	GCGAATTCGAGCTCGGTACCTTAGCTGCTATTACCCAGG

CYP number	Target area	Forward Primer (5'-3')	Reverse Primer (5'-3')
CYP52T1_2	N-terminal anchor	CTAAATTACCGGATCCATGAACTTGTTGTCATCGTCTT	ATAGCTAACCTGACGCCATACCACCTTGAGAATATAC
	truncated heme domain	CGTCAGGTTAGCTATCAGAGC	GCGAATTCGAGCTCGGTACCTTAATTTGCCTGCTGCAGACG
CYP52T1_3	N-terminal anchor	CTAAATTACCGGATCCATGACTCTCTCGCCTATTTCTAC	TGCAAACAGCCAATGGTTTATCTCGATGCATACCCAC
	truncated heme domain	CATTGGCTGTTTGCACGTAAAC	GCGAATTCGAGCTCGGTACCTTACAGCTGTGCTGCTTCACA
CYP539B1	N-terminal anchor	CTAAATTACCGGATCCATGCTTATCGAGGCTGTCA	AACTTTATAGGCAACCGTGAGCCAGTTGACAATG
	truncated heme domain	GTTGCCTATAAAGTTGCAAAA	GCGAATTCGAGCTCGGTACCTTAGGCCGTTTTTTCATCG

Table A.4: Primer pairs used for PCR amplification of P450 targets from *R. jostii*

Number	Gene Target	Forward Primer (5'-3')	Reverse Primer (5'-3')
Ro01	*ro02510*	CCAGGGACCAGCAATGTCGCGTGCGCAGCTGTTCCAGC	GAGGAGAAGGCGCGCCCACAGTCCACCAGCAGGTGC
Ro02	*ro02948*	CCAGGGACCAGCAATGTTCGCCCCCGCTCCCGCGTTGC	GAGGAGAAGGCGCGGGGCAGCCGCACCTTGACCG
Ro03	*ro04669*	CCAGGGACCAGCAATGACCGCAGCCTCGATGCCCG	GAGGAGAAGGCGCGGGCATCGCGGACCTGCACTTTCATGG
Ro04	*ro02604*	CCAGGGACCAGCAATGAGTAAAACGGTCGCCGTCCCGTACG	GAGGAGAAGGCGCGCCGGGACGTTTCCGTCCAGGTGACC
Ro05	*ro04667*	CCAGGGACCAGCAATGACCGCATCTCGGATTGATCTGAAATGCC	GAGGAGAAGGCGCGCACCGGACTGCGCGGATGAATCCG
Ro06	*ro02651*	CCAGGGACCAGCAATGACGTTATCGGCCCTGCACACCGACC	GAGGAGAAGGCGCGTGACGTGGCCTTTCCGAGGTGAACG
Ro07	*ro01169*	CCAGGGACCAGCAATGACAGTCAGGACTGAACTGCAAGGGATTCCG	GAGGAGAAGGCGCGGCCGACCGCCACGGTCACTGG
Ro08	*ro11320*	CCAGGGACCAGCAATGACCGAAACCAGCGCGGAGCCTC	GAGGAGAAGGCGCGGGGCCGGTCCTCAGGCACTGG
Ro09	*ro04588*	CCAGGGACCAGCAATGCGGAGGAGTGCCTCGATGACGAACAGG	GAGGAGAAGGCGCGGAGCTTCACCGGCATCTCCTCGAGC

Number	Gene Target	Forward Primer (5'-3')	Reverse Primer (5'-3')
Ro10	*ro00423*	CCAGGGACCAGCAATGACGACAACCCAACTACCCGACAGG	GAGGAGAAGGCGCGGGACGCGGGTGTAAAGGTGATCGG
Ro11	*ro11277*	CCAGGGACCAGCAATGCTCGATCCGTTTGCGCCCGGCCTG	GAGGAGAAGGCGCGGATCAGGTTTACTGGCATCTCAGTGGGGCC
Ro12	*ro00377*	CCAGGGACCAGCAATGGCTTTCGAAGGGTGTTCTGCCTTAACAACTC	GAGGAGAAGGCGCGCAGTTCCGGAGTGAACGTGAGCGGTAGACG
Ro13	*ro08068*	CCAGGGACCAGCAATGACCGAAACCACTCCGGCACCG	GAGGAGAAGGCGCGCGGTCGGTCCTCGGGGAGC
Ro14	*ro05719*	CCAGGGACCAGCAATGGGCAATGTGTCTCACTATCGATACATGACC	GAGGAGAAGGCGCGTGCTTCACGCACCCGGAACGGGACG
Ro15	*ro00393*	CCAGGGACCAGCAATGACCTCGACACAGCTCGATCTTCC	GAGGAGAAGGCGCGGTGCGTCGGCGGCGTGAACG
Ro16	*ro02355*	CCAGGGACCAGCAATGCAGCTCGCCGACGTCGATCTCTACAACC	GAGGAGAAGGCGCGGAGGGACGAGGGCGTGAACGTCACC
Ro17	*ro04679*	CCAGGGACCAGCAATGGGCTCCTTTCCTTGTCCTCAGAAAATAGAACAGG	GAGGAGAAGGCGCGGTGTCTGACCGGGCAACCGCCTGACG
Ro18	*ro08984*	CCAGGGACCAGCAATGACCACTGTCGACGAACCGGTACTCG	GAGGAGAAGGCGCGCCAGGTCACGGGCAGCATAGTGATCGG
Ro19	*ro02382*	CCAGGGACCAGCAATGACCGCCACCCTGTCTTGGATCGACG	GAGGAGAAGGCGCGCACCTCCCACGTGGCCCGGAG
Ro20	*ro03876*	CCAGGGACCAGCAATGACCACCAGCACCACGTGGATCGAATCG	GAGGAGAAGGCGCGGACCTCCCAGGTGACGTGCAGG
Ro21	*ro03826*	CCAGGGACCAGCAATGACGGGTTCTGCGTCCGAAAAAATCATCTCGC	GAGGAGAAGGCGCGCCGTCCGTCACGTGTCCGGGCCGGCAGTGC
Ro22	*ro04671*	CCAGGGACCAGCAATGAACCTGGCTACCCCGCAACGTGTGTCC	GAGGAGAAGGCGCGCGTGCTTCGCTTCCGGTACCTCACG
Ro23	*ro04627*	CCAGGGACCAGCAATGCCAGTCACCGTACTGCTTGACACC	GAGGAGAAGGCGCGACGCCGGCGGAGCTGCAGC

Appendix B

***E. coli* codon optimized gene sequences of *B. bassiana* P450s**

Cytochrome P450 CYP505A1

ATGGATGTTGATCAGATTGCCTTTCGTTATCAGGATAGCGTTCGTGTTATTGATGCAATTCATAGC
GCACTGGCAGATCATCAGACCATGAACATTCAGCATCAGGATCCGCAGGATCGTACCGCAAGCCAG
AAACAGTGTGGTTTTAGCCCGTATGCAATGGTTGCCATGGTGGCTATGGTAGCCATGGTTGCAATG
GAAGCAATGGTGAGCGAAGCACATCTGGCAACCGCACCGCAGGTTCTGACCGATGCAGCACAGGGT
CGTCGTCAGGCACTGGAACCGCGTGAAATTGCAGTTCTGCGTGAAGCACGCGAAATTGCCGCACTG
CGTAAAGCATTTGCAGAACAGGTTCCGACCCGTGCAGTTAAAGCAGATGGTATTCGTCATCAGGGT
GATCTGCAGGCCGAAGCCGATGGCATTAAACGTGAAAATGAAAGCTGGAAAGGTGACCTGAAAACA
CCGGCAACCCCGAGCAGCCGTACCACCACCATTACCGCAGAAAGCATTCCGATTCCGGAACCGCCT
GGTCTGCCGTTTATTGGTAATCTGGGTGAAATGCGTACCAGCCCGATTAATGATTTTAAACGTCTG
GCAGATACCTACGGCGAAATTTATCGTATGCATCTGGGTGGTAGCGCATTTTGTGTTGTTAGCAGC
CGTGAACTGGTTAATGATGCATGTGATGAACGTCGCTTCAAAAAAACCGTTGGTGGCACCCTGGGT
AAAGTTCGTAATGCCATTCATGATGGTCTGTTTACCGCAGATAGCGAAACCGAACCGAATTGGGGT
AAAGCACATCGTATTCTGGTTCCGGCATTTGGTCCGCTGAGCATTCGTAATATGTTTGATGAAATG
CATGATATCGCCAGCCAGATGGCAATGAAATTTGCACGTCATAGCGGTGATCGTATTAATGCCAGT
GATGATTTCACCCGTCTGGCCCTGGATACCGTTGCACTGTGTGCAATGGATTATCGCTTTAACAGC
TATTATCGCGAAGAACTGCATCCGTTTGTTCGTGCAATGGGTGATTTTCTGACAGAAAGCGGTGCA
CGTAATCGTCGTCCGACCTTTGCACCGCAGTTTTTCTATCGTGCAGTGGATGAAAAATATGAGAAA
GATATCAAAACCATGCGTGATGTTGCCGATGAAGTTGTTGCAAATCGTCGTGCAAATCCGAGCGAA
CGTAAAGATCTGCTGAGCGCAATGCTGGATGGTAAAGATCCTCAGGATGGTCAGCGTCTGACAGAT
GCAAGCATTACCGATCAACTGATTACCTTTCTGATTGCCGGTCATGAAACCACCTCAGGCACCCTG
AGCTTTGCATTTTATCAGCTGCTGAAACATCCGGCAGAATATCGTAAAGTTCAAGAAGAAGTTGAC
GCAGTTGTTGGTCGTGATCGCATTACCGTTGAACATATTAGCAAACTGACCTATATTCAGGCCGTT
CTGCGCGAAGTGCTGCGTGTTAATGCACCGATTCCAGCATTTAGCGTTGAAGCAAAAGAAGATACC
CTGCTGGGAGGCAAATATTTCATTCCGAAAGAACATCGTCTGACCCTGCTGCTGGCAAAAAGCCAT
CTGGATCCGAGCGTTTATGGTGATAGCGCAAGCGATTTCAAACCGGAACGTATGCTGGACGAAAAT
TTTGCCCGTCTGAATAAAGAATTTCCGAGCGCATGGAAACCGTTTGGTAATGGCAAACGTGCATGT
ATTGGTCGTCCGTTTGCATGGCAAGAGGCAGTTCTGGCAATGGCAATTCTGTTTCAGAACTTTAAT
TTCACCCTGGACGATCCGAATTATACCCTGGAAATTCAAGAAACCCTGACCCTGAAACCGCACAAC
TTTTTTATGCGTGCAACCCTGCGTCATGGTATGAGCGCAACCGAACTGGAAGATCAGCTGAAAGGT
GGCACCGTTCATAGCAAAAGCAGTGATGGTGCAGATACACCGGTTGCAGCAAGTGCCGGTGATGGT
AAACCGCTGAGCGTTTTTTATGGTAGCAATAGCGGCACCTGTGAAGCACTGGCACAGCGTGCAGCA
GCAGATGCCAGCGCACATGGTTTTAAAGTTACCGAAATCGGTCCGCTGGATAATGTTAATCAGAAA
CTGCCGACCGATCGTCCGGTTGTTATTGTTACCGCCAGCTATGAAGGTGAACCGCCTAGCAATGCA
GCCCATTTTGTTGATTGGCTGAAAAGCCTGAAAGGCGACGAACTGAAAAATGTTAGCTATGCGGTT
TTTGGTTGCGGTCATCATGATTGGGCACAGACCTTTCATAAAATTCCGAAACTGGTTGATGCGACC
ATGGCCGAACGTGGTGCCGATCGTATCATTCCGATGACCGGTACAGATGCAGCAGATCGTGATATG
TTTAGCGATTTTGAAACCTGGGAAGATGAATGTCTGTGGCCTGCACTGAAAAAAAAATATGGTGCG
GATGAAACCAAAGATGGTCAGGGTGCAAGCGCACTGACCGTTGAAATTACCCATCCGCGTAAAACC
ACCCTGCGCCAGGATGTTGAAGAAGCAGCAGTTATTGATACCAAAGTGCTGACCAAAGGCACCCAG
AGCGTTAAAAAACATATTGAAATTCGTCTGCCGACGGGCATGACCTATAAAGCCGGTGATTATCTG
GCAGTGCTGCCTTTTAATCCGGCAGCCACCATTGCCCGTGTTTTTCGTCGTTTTAGCATTAGCTGG
GATGCAACCTTTACCATTACCTCAAATGGTCCGACAACCCTGCCGACAGGTGTTCCGATTAGCGCA
ACCAATGTTCTGGGTGCATATGTTGAACTGAGCCAGCCTGCAACCAAACGTAACATTCAGGCAATG
ATTGATAGCACCGAAGATGAAAAAACAGTGACCGCACTGAAAGGCCTGATTGGTGATAAATTCAGC
GAAGAAGTTACCGCAAAACGTCTGAGTGCACTGGACCTGCTGGAAAAATTTCCGGCAGTGGGTCTG
CTGTTTGAAAGTTTTCTGGCCATGCTGCCTCCGATGCGTGTTCGTCAGTATAGCATTAGTAGCAGT
CCGCTGGTTGATCCTACCCGTGTTACCCTGACCTATTCACTGCTGGATGTTCCGGCACATAGCGGT
CAGGGTCGCCATGTTGGTGTTGCCAGCCATTACCTGTTTAGCCTGCATGAAGGCGATAAAATTCAT
GTTGCAGTTCGTCCGAGTGCAGCCTTTCATCTGCCTGCCGATACCGAAAAAACCCCGATTATTTGT

GTTGCAGCAGGCACCGGTCTGGCTCCGTTTCGTGGTTTTGCCGAAGAACGTGCAGCCATGATTGCA
GCCGGTCGTAAACTGGCTCCGGCACTGCTGTTTTTTGGCTGTCGTGCACCGGGTGAAGATGATCTG
TATGCCGAACAGTTTGCAGAGTGGCAGAAAATGGGTGTGATTGATGTTCGTCATGCATATAGCCGT
GCAACCGATAAATCTCATGGTTGTAAATATGTGCAAGAACGCCTGAGCAATGATCGTGATGAAATC
TATAAACTGTGGGATCAGGGTGCGCGTCTGTATCTGTGTGGTAGTCGTGCCGTGGGTAAAGGTATT
GAAGATGCATGCGTTGAACTGGTGAAAGAAAATGCAGAACGCGAAAAAAGCAAAGAAGTGACCGAC
GAAGCAGCACGTGCATGGCTGGATAACCTGCGTAATGAACGTTTTATGGCCGATGTGTTTGACTAA

Cytochrome P450 CYP52G6
Transmembrane area was removed (nucleotides 4-54) for gene synthesis.

ATGCGTGTTCTGGAAAGCCTGCGTCATGCAAAAAATGCACGTGAACTGGGTTGTAAACCGCCTCCGCTGGCAC
CGATTAAAGATCCGCTGGGTATTCTGAGCCTGCTGGAAATGATTCAGGCAGATAAAGAAAAACGTGTTCCGGC
ACTGACCGAACAGCGTGTTAACAAAATGCGTGATGATAATGGTGGCAATTATGTTACCACCATGCGTCTGCGT
ACCGGTGCAGTTGAAAATATTCTGACCATTGATCCGAAAAACATTCAGGCAATTCTGGCCACCCAGTTTAAAG
AATTTTGTGTTGGTGCACAGCGTGAAAGCTGTATGGGTCCGCTGCTGGGTGCAGGTATTTTTACCACCGATGG
TCCGGCATGGTCACATAGCCGTGCAATGCTGCGTCCGCAGTTTACCCGTGATCAGATTAGCGATCTGAGTCTG
GAAGAGGTTCATGTTCAGAATGCCTTTAAAGTTATGCCTCCGGTTAATAATCAGGGTTGGACCGAAGTTGATA
TTCAGACCGTTTTTTTTCGTCTGACCCTGGATAGCGCAACCGAACTGCTGTTTGGTGAAAGTTGTAAAAGCCA
GCTGGTTGCACTGGATAATGCCAATAATGATAAAGAATTCAGCGCACGCGGTACAGATTTTGGTGCAAATTTT
GATCGTGGTCAGTGGTATCTGAGCCAGCGTGTTCGTACCCCGTTTCTGAAATTTCTGTATAACGGTGACGAGT
TTAAAAACTGCTGCAAAGAAGTGCACCGTTTTGTTGATCAGTGTGTTGAACGTGCCCTGCGTGAAACCAGCAA
AAAAAAACTGGATGCAGATGGTAAACCGCTGGAAGGTGGTGAACATTATGTTTTTCTGCATGCAATGGCAGCC
GAAACCCAGGATCCGATTGAACTGCGTGCACAGCTGCTGAATGTTCTGCTGGCAGGTCGTGATACCACCGCAA
GCCTGCTGAGCTGGACCGTTATGCTGCTGGCACGTCATCCGGATAAATTTCAGCGCCTGCGTCGTGATATCAT
TGAAACCTTTGGTGGTTATGAAAATCCGCGTAATCTGACCTTTGCAAATCTGAAAGCATGTACCTATCTGCAG
CGTGTGATGACCGAAGTGCTGCGTCTGTTTCCGCCTCTGCCGATGAATGCACGTTATGCAACCTGTGATACCA
GCCTGCCTCGTGGTGGTGGTCCGGATGCAGAAAGTCCGGTTTATGTTAAAAAAGGTCAGGCCGTGCTGTATAA
TGCACATATTCTGCATCGTCGTACCGATATTTGGGGTCCTGATGCCGGTGAATTTAACCCGGATCGTTGGGAA
GGTCGTAAAGGTGGTTGGGAATATCTGCCGTTTAATGGCGGTCCGCGTATTTGTATTGGTCAGCAGTTTGCCC
TGACCGAAGCAGGTTATGTTCTGGCACGCCTGCTGCAGCGTTTTGATGGCCTGGAAGAACTGAATCCGAGCAG
CAAAGTTAGCTGGGGTCTGACACTGGTTAGCCAGCCTGGTGAAAGCGTTAAAGTTCGTCTGCATGAAGCAGTG
TAA

Cytochrome P450 CYP52G8
Transmembrane area was removed (nucleotides 4-84) for gene synthesis.

ATGGTTCGTTTTCGTAGCAAAGCAGCACGTCTGGGTTGTAAAAGCGCACCGAGCGGTATTAGCAGCGATTGGA
GTGGTATTAGCCTGATGCGTAAAGGTCTGAAAGCACAGCGTGAAAAAAATGTTCCGAATTGGATGCGTCATGA
ATTTGCACGTCTGAGCGCACGTGAAGGTCGTCCGGTTGGCACCTTTGAAATGAGCGCACCGCTGTTTCGTCGT
GTTCTGTTTACCAGCGAACCGGAAAACATTAAAACCATTCTGGCAACCAGCTTCAAAGATTTTAGCCTGGGTG
ATAATCGTCGCGGTAACTTTAAACCGCTGCTGGGTGAAGGTATTTTTGCAAGTGATGGTAAAAAATGGGAACA
TAGCCGTGCAATGCTGCGTCCGCAGTTTGTTCGTAGCCAGGTTAGCGATATTAGTCTGGAAGAAACCCATGTT
CAGAATCTGATGACCGTTCTGGATAGCCATCTGGATAAAACCACCGGTTGGAGCGGTGCAGTTGATCTGCAGC
CGCTGTTTTTTCGTCTGACCCTGGATAGCGCAACCGAATTTCTGTTTGGTGAAAGCGTTAATAGCCAGCTGCG
TCGTGAAGGTGATGCAGATAATGATGCACATGCATTTGCGACCAGCTTTGATGCAAGCCAGAATCAGCTGGCA
GTTGCAGGTCGTTATGGTAGCAATTATTGGATTGGTCATACCAAAGCCTTTCGTGAAGATGTTCGTATCTGCC
ACGAATTCATCGATTACTTTGTTCAGAAAGCACTGAATGGTCAGCGTGATACCAGTGAAAAAGCAGATGCCGA
TAAAGAAGAACGCTACGTTTTTCTGGAAGCAATTGCACGTGAAACCAGCGATCCGGTTGAACTGCGTAGTCAG
CTGATTAACATTCTGCTGGCAGGTCGCGATACCACCGCAAGCACCCTGGGTTGGTTTTTTCATATTATGGGTC
AGGCACGTAACGCCCATATCTATAAACGTCTGCGTCAGGCAATTCTGGATGAATTTGGCACCTATCGTAATCC
GAAACCGATTACCTTTGAAGGCCTGAAAAATCTGACCTATCTGCAGTGGTGTATTAATGAAACCCTGCGTCTG
TATCCGATTGTTCCGATGAATGGTCGTGCAGCAGTTAAAGATACCGTTCTGCCGCTGGGTGGTGGTCCGGATG
GTCGTAGCCCGATTCTGGTTAAAAAAGGTCAGGATATTGGTTATAGCGTGCATGTTATGCATCATCGTACCGA
TCTGTGGGGTGCAGATGCAGATGATTTTCGTCCGGAACGTTGGGAAAAACGCAAACCGGGTTGGGATTATCTG
CCGTTTAATGGCGGTCCGCGTATTTGTATTGGTCAGCAGTTTGCACTGACCGAAATTGCATATGTTGTTGCAC
GTATGCTGCAGCGTTTTGATGAACTGGATGGTAGCACCCTGAGCGCAGAAAGCCATGGTCTGGGTCTGACCAA
TTGTCCGGGTGAAGGCGTTACCCTGAAAGTTCACTTTGATGAATAA

Cytochrome P450 CYP52T1
Transmembrane area was removed (nucleotides 4-60) for gene synthesis.

```
ATGCGTAGCATTCAGCGTCGTAATGCACGTGATCGTCTGGCACGTCAGCATGGTTGTGAACCGCTGACCCTGG
CATATAACAAACTGCCGTTTGGTCTGGATCGTAAATGGCAGATTGTTACCCATCGTGGTAATATTCTGGATGA
TCTGATTACCACCCGTTTTGCAGAACTGGGTAGCTATATCTATACCGATAATCAGTGGGGTAGCCCTCCGATT
ATTTGTGCAGAACCGGCAGCAATTAAAGCAGTTCTGAGCACCAAATTTCGTGATTGGGATATGGATAGCAATC
GTTATCCGGCACTGGGTCCGTGGCTGGGTCGTGGTGTTCTGGTTAGCAGCCATCAGGGTAAAGGTAGCCTGTG
GCTGACCGCACGTACCCTGCTGCGTCCGATGTTTGCAAGCGTTGCAACCTATAATCATGCCCTGATGGAAAAA
AGCGTTCAGGATTTTCTGAGTACCATGAGCCGTGTTAATGAAGATAGCGCAACCCGTAGCGATCTGCTGCCGC
TGATTCGTCGTCTGAATATTGATATTATCACCGCCATTTTTTGCGGTGGTAGCATTGGTGCACAGAAAAAAGG
TCTGGAAGCAGGTCCGCGTGCAGCAGCAGCTGCAGCAGCCGCAAGTCCGGGTAAAAAACCGACCCTGGAAGAG
GCATTTGATGCAATTGAACCGATTGCAGGTCTGCGCCTGCAGACCGGTAGTCTGTATTGGCTGTTTACCAGCA
AACCGTTTCGTGATGCATGTGATACCTTTAGCGAACTGGCAAATGGTTGGATTAATCAGGCACTGCGTAAAAC
CCATGAAAAAAGCAGTCCGCAGGGTGGTGGCCTGGATGGTGTTGCAGAAGCAGCACAGAGCTTTACCGAAGAA
CTGGTGAGCAGCACCGAAGATCGTGAACTGCTGCGTGATATTCTGGTTCAGCTGCTGTTTGCAGGTATTGATA
CCAGCACCAGTATGCTGAGCTTTGCACTGCTGGAACTGGGTCGTCATCCGGGTAGCTGGACCCGTCTGCGTGC
GGAACTGGCAGAACATAGCCTGCTGAGCGCAGGTCCGGAAACCATTACCGCAGGTCAGCTGAAAGATTGTGTT
TTCCTGCAGAATGTTGTGAAAGAAACCCTGCGTCTGTATCCGCCTGTTCCGATTAATAGCCGTGAAGCAATTC
GTGATACCGTTCTGCCGACCGGTGGTGGTGCAGATGGTAGCAAACCGGTGTTTGTTCCGAAAGGCACCAGCCT
GAAATATAGCCCGTATGTTATGCATCGTCGTAGTGATCTGTATGGTCCGGATGCAATGGAATGGAAACCGGAT
CGTTGGCTGGGACGTAGCCATGGTTGGGATTATCTGCCGTTTAATGGTGGTCCGCGTATTTGTATTGGTCAGA
AATTTGCACTGACCGCAGGCGCATATGTTCTGACACGTCTGGCCCAGCAGTTTGATACCTGTGAAGCAGCAAG
CAGCAATAAAGGTCCGCTGGAAAGCAAACTGGGTGCCGTTCTGATTCCGGCAGCCGGTGTTCCGGTGAGCCTG
GGTAATAGCAGCTAA
```

Cytochrome P450 CYP52T1_2
Transmembrane area was removed (nucleotides 4-78) for gene synthesis.

```
ATGCGTCAGGTTAGCTATCAGAGCCTGGCACGTCGTGCAGGTTGTAAACCGCCTCCGGCACGTCCGTATCGTC
TGCCGTTTGCAATTGATAATATTGTTCGTACCATGCGTGCCATTCTGGATCATACCCTGCAGAATGATGAAGT
TGCAGTGTATGAAGAAATGGGTTGTCCGGCAACCTGGCGTCAGAACATTCTGGGTGTTTGGTATCATGCAACC
GCAGATCCGGAAAACATTAAAGCACTGCTGGCAACCCAGTTTAACGATTTTGAACTGGGTAGCATTCGTCTGG
ATCACATGGGTCCGCTGATTGGTCATGGTATTTTTACCAGTGATGGTAAAGAATGGCAGCAGCAGCGTAGCAT
GCTGCGTCCGCAGTTTACCCGTGCACAGATTAGCAATCTGACCCTGCTGGAAGCACATGTTCAGAACCTGTTT
CAGCATTTTGATAGTCCGCATGCAGGTAGCTGGACCGCAGAAGTTGATCTGGCACCGCTGTTTTTTAACCTGA
CCCTGGATGCAGCAACCGAATTTCTGTTTGGTCAGAGCGTTGAAAGCCAGATTCATCATGGCAAAAAAAGCCA
TGGTGGTAGCGATAGCGGTAAAAAAAGTGGTCAGGATGGTCTGATTAGTGGTAAAGATTGGAGTAGCTTTGGT
CGTGCATTTGATCGTGCAAATGCAACCATTGCACTGCGTGGTATGCTGATGGATTTCTATTATCTGTATCGTC
CGAGCAGCCTGGCCCAGGATTGTCGTGAAGTTCATAAATTTGCCGATCATTTTGTTCAGCGTGCACTGAATAC
CGAAGTTCAGGATACCGAAGGTGATAGCGAAACCGAAGCATATGTTTTTCTGCGTGAACTGGTTAAAACCACC
CGTGATCCGTATGTTCTGCGTAGCCAGCTGCTGAATATTCTGCTGGCAGGTCGTGATACCACCGCAGGTCTGC
TGGGTTGGACCTTTTACCTGCTGGCACGCCATCCGGATTATTACAGCAAACTGCATCGTATTGTTGTGGAAAC
CTTTGGTCCGAGCTGTAGCGCAGATAGCGCAAGCATTATTACCTTTGAAAGCCTGAAAGCATGTCATCCGGTT
CAGCATCTGCTGAGCGAAGCACTGCGTCTGCATCCGGTTGTTCCGGAAAATGGTCGTCGTGCCGTTCGCGATA
CCACCCTGCCTCGTGGTGGCGGTCCGGATGGTCAGAGTCCGGTTTTTATTCGTAAAGGCCAGGATGTTCTGTA
TAGCGTTAATGTTATGCATCGTCGTAAAGATCTGTGGGGTGATGATGCACATGAATTTCGTCCGGAACGTTGG
GCAGATCGTAAACATGGTTGGGAATATCTGCCTTTTAATGGTGGTCCGCGTATTTGTCTGGGTCAGCAGTTTG
CACTGACCGAAGCAGCCTATGTTGTTGTTCGTATGCTGCAGAAATATGGTCGCATTGAAAATCTGGATCCGGA
TACCGTTACCCGTCATCGTTATACCCTGACCACCGCACCGGTTAAAGTTGCCGTTCGTCTGCAGCAGGCAAAT
TAA
```

Cytochrome P450 CYP52T1_3
Transmembrane area was removed (nucleotides 4-81) for gene synthesis.

```
ATGCATTGGCTGTTTGCACGTAAACTGGGTTGTAAACCGGCACATGTTCGTCGTAGCAAACTGCCGCTGGGTC
TGGATAATGTTCTGCGTATGGCAAAAGCAGCCAAAAATCAAGAACTGCAGAATGATGATCAGCTGGTGCATCA
AGAAATGGGTTGTCCGAGCACCTGGGTTCAGAATTTTCTGGGTTTTTGGTGTCATACCACCGTTGATCCGGAA
AACATTAAAGCAATTCTGGCCACGCAGTTCAAAGATTTTGAAATGGGTCCGTTTCGTACCGATACCCTGAGTC
CGCTGCTGGGTCATGGTATTTTTACCAGCAATGGTAAAGAATGGTACATCGTGAATTTCGATCCGCTGAAACG
TGATGTTCTGCTGACCTTTATGTTTACCCGTAATCAGATTAGCAGCCTGGAACTGGAAGAAGTTCATATTCAG
CACCTGTTTGGTCGTTTTCATCATGGTGCAGATGGTAGCTGGACCAGCCCGATTGATCTGGGTCCGCTGTTTT
TTAACCTGACCCTGGATAGCGCAACCGAATTTCTGTTTGGCCAGAGCGTTGATAGCCAGCTGCTGGATAGCCC
GAATGCAGCAAAAGCAACCAGCAGCGAACATGAAACCAAAGCAAATCGTGATAGCAAAGATTGGAGCAGCTTT
GGTCGTGCATTTGATCGTGCAAATACCACCATTAGCTTTAAAGGTATGCTGATGGATTTCCACTTTCTGTATA
GTCCGAAAAGCTTTACCGATGATTGCAATACCGTTCATCGTTTCGCCGATTATTTTGTTCAGCAGGCACTGAA
TGAAGAACAAGAAGTTAGCGGTAGTCCGGATAGTAAAGGTGAAACCGAAGCATTTGTTTTTCTGCGTGAACTG
GTTAAAAGCACCAAAGATCCGAAAGCACTGCGTGGTCAGCTGCTGAATATTCTGCTGGCAGGTCGTGATACCA
CCGCAGGTCTGCTGGGTTGGACCTTTTACCTGCTGGCACGTCATCCGGATTATTATAGCAAATTTCACCGCAT
TATCGTGGAAACCTTTGGTCCGTATAGTGAACATGCAAGCAGCCTGACCTTTGAAAGCCTGAAAGCATGTAGC
CATCTGCAGAATCTGCTGAGCGAAGTTCTGCGTCTGCATCCGGTTGTGCCGGAAAATTCACGTCGTGCAACCC
GTAATACCATGCTGCCTCGTGGTGGTGGTGTTGATGGTAATGCACCGGTTTATATTCGCAAAGGTGAAGAGGT
GATTTACAACGTGAATGTTATGCATCGTCGCAAAGATATTTGGGGTGATGATGCAGATGAATTTCGTCCTCAG
CGTTGGATTGGTAGCAAACATGGTTGGGAATATCTGCCGTTTAATGGCGGTCCGCGTATTTGTCTGGGTCAGC
AGTTTGCACTGACCGAAGCAGGTTTTGTTGTTGTTCGTCTGCTGCAGCGTTATTCCAAAATTGAAAATCTGGA
TACCGAAACCGTGACCAAACATCAGTATACCCTGACCACCGCACCGGTTAAAGTTCTGGTTCGTTTTTGTGAA
GCAGCACAGCTGTAA
```

Cytochrome P450 CYP539B1
Transmembrane area was removed (nucleotides 4-87) for gene synthesis.

```
ATGGTTGCCTATAAAGTTGCAAAAAGCAGCGGTGTTCGTGCACCGAGCATTGGTGATAATCCGATTAGCGCAA
TTCGTGTTAGCCTGACCGCAGTTAAATATCAGAATCAGAATCGCCTGTACGATTTCTTTAAAAGCATTTTCGA
TGCAGGTACACCGGAATGTCCGAATGCAGTTGAAGCACTGTTTTTTGGTCGTCGCATTATCTTTACCCAAGAA
CCGGAACATATCAAAACCGTTCTGACCGCAAAATTTGCCGATTATGGTAAAGGTCCGAAATTTCATGAAGTTT
GGGCACCGTTTCTGGGCGATAGCATTTTTACCACCGATGGTGCACAGTGGCATGATAGCCGTACCCTGATTCG
TCCGATGTTTGTTAAAGATCGTGTTCGCGATATGGGCATTTTTGAACGTTGGAGCGATAAACTGATTAGCAAA
CTGCCTGCAAGCGGTGAAACCGTTGATATGTGTGACCTGTTTTATCGTATGACCCTGGATCTGACCACCGATT
TTCTGCTGGGTAGCGGTGTTGGTAGCCTGGATAACCCGAATAGCGAATTTAGCAATGCATTTACCGTTGTTCA
GCGTCTGCAGATGATTCTGACCATTATGCTGCCGTTTCGTCGTTTTATTCCGCAGCAGCAGTATCGTGATGGT
ATTAAAACCCTGGAAAAATTCATGACCCCGTATATTCAGCAGACCCTGAGCCTGACACCGGAAGAACTGGAAA
AACTGAGCAAATCCGATAAACAGTTTACCTTTCTGCATAACATTGCCCTGTTTAGCCGTGATCCGAAAGTTAT
TCGTGATCAGATTATGGCAGTTCTGCTGGCAGGTCGTGATACCACCGCAGCCACCCTGAGCTGGACCATTTAT
GAACTGGCAAATTATCCGGATGTTTGGACCAAACTGCGTCAGACCGTTCTGGAAAAAGTTGGTCCGGATAGCA
ATCCGACCTATGAAGATATTAAAGGCATTACCTATCTGACCCATGCAATTAGCGAAACCCTGCGTCTGTATCC
GGCAGTTCCGTATAACATTCGTAGCTGTCTGCAGGATAGCACCCTGACAGGTGCACCGGGTCAGCCGGATATT
GCATGTTTTAAAGGTAACCATGTGATCTATAGCACCTATGCAATGCAGCGTCGTCGTGATCTGTATCCTCCGG
TGAGCGAAACCTTTGCAGATCCGGATATCTATTCACCGGATCGTTGGGATCATTGGACACCGCGTCCGTGGCA
GTATGTTCCGTTTAATGGTGGTCCGCGTATTTGTATTGGCCAGAATTTTGCCATTACCGAAATTGGTTATGTG
CTGGTTAAACTGCTGCAGAAATATGAACGTCTGGAATATCGTGACGATTGGAATGCACAGTTTCATAAAGCAG
AAATTGTTGGTTGTCCGGGTCGTGGTGTTCCGGTGGCCTTTTTTGAACCGGAAAAAACCGATGAAAAAACGGC
CTAA
```

Cytochrome P450 CYP584E2

ATGGCACTGGGTCAGCTGGCACCGACCGTTGCAGTTCTGAGCGATCATCCGGTTCGTGATCTGATTCTGCTGA
GCCTGATTGCCGTTCTGGGTTATCCGATTTTTCGCCATCTGATTGCATGGCGTGCCCTGGGTCGTGCATATGC
CCTGCATGATTGTGAACGTCCGGCAGCATATCCGCATAAAGATCCGCTGTTTGGTTTTGATGCAATTCTGGCA
AATATTGCAGCAGCACGTAATCATCGTTTTACCGCAAGCGAACTGCAGCGTTTTCAGGATATTGGTGCAGATA
CCTATTACACCTGGATTCAGGGTCGTCAGATTGTTGTTACCCGTGATCCGGATAATGTTCGTTGTATTCTGGG
CACCAATTGCAAAGATTATAGCATTGGTGGTCGTCGTGCACTGTTTGGTCGTTTTCTGGGTAATGGTATTTTT
GTGAGCGAAGGTGAAGAATGGGTTCGTAGCCGTGCACGTCTGCGTCGTAATTTTAGCCGTGAACAGGTTGCAG
ATCTGGCAATGCTGGAACGTCATGTTGCAAAACTGTTTCAGGTTCTGCCTGGTGATGATCAGCCGAATGCAGT
TGTTGATCTGCAGGATTGTTTTCTGCGTTTTACCACCGATAGCAGCAGCGAATTTCTGTTTGGCCATAGCACC
GATACCCTGGTTCGTCCGAGCGCACGTGATGTTGCATTTGGTGAAGCATTTACCCTGAGCCTGGATATTATCA
CACAGAAAATGCGTCGTGGTCCGCTGAATCGTTTTTATCCGAAAGATCCTCGTGAAGATAGCGCATTTCAGAT
TGTGCGTGATTATGTTCGTGCCTTTGTTGATGAAGCAGCAGCCCTGCGTGAGAAAAAAACCGCAGATGGTGAA
CTGCTGGAAGGCCATGAAGATGCAGGTCCGGAACATCGTTATCTGATCCTGCGTGAACTGGTTCGCGATTTTG
ATGATAAAGAACGCATTTGCGACGAACTGATTAGCCTGATTACCGCAGGTCGTGATACCACCGCAAGCCTGCT
GAGCAGCGCCTTTCATGTTCTGAGCCGTCGTCCGAATATTTGGCGTACCCTGCGTAATGAAATTCAGCATCTG
AATGGTCATCCGCCTAGCTATGAACAGCTGCGCAACCTGAAATTTGTGAAATACATTATCAATGAAACCCTGC
GTCTGTATCCGCCTGTTTTTCGTAATGCACGTAAAGCAGTTCGCGATACCATTCTGCCGACCGGTGGTGGTCC
GAATGGCACCAGTCCGGTTTTTGTTCCTAAAGATACCGGTGTTGTTTATAGCGCATGGGCAATGCATCGTCGT
ACCGATCTGTATGGTCCGGATGCAACCGAATTTAATCCGGAACGTTGGGCAACCCAGCGTCATGGTTGGGATT
ATCTGCCGTTTGGTGGCGGTCCGCGTATTTGTCTGGGCCAGCAGTATGCACTGACCGAAGCACTGTATGTTCT
GGTTCGTATGGCACAGCAATATGTTGCAGTTGAAACAGCAGATGATACCCCGTGGACCGAACATATTTGCCTG
ACCCTGGCAATTAAAGATGGTGTTAATTGTAAACTGACCCGTGCCAAAAACTAA

<u>Sequences of the truncated transmembrane areas of *B. bassiana* P450s</u>

Cytochrome P450 CYP52G6

ATGGCGCTCACTGCTATCCTCATCGGCCTCGTCGTCGTCACCTTTGTGCTTCGA

Cytochrome P450 CYP52G8

ATGGCGGTACTATCCGTTATTTCGCTTCCGGCGCTGCTGGTCTCTTTGACCGTTGCCTTTATCCTCCTCCAGG
TGATTGAGTAC

Cytochrome P450 CYP52T1

ATGGCTCTCCACGCTGCCTACCTATTTATTGCAGCGACGCTGGTGGCCGTCTATCTGACT

Cytochrome P450 CYP52T1_2

ATGAACTTGTTGTCATCGTCTTTTGCTGCAGTTCTTCTATCGGCTGTCATTGCCGCGTATATTCTCAAGGTGG
TATGG

Cytochrome P450 CYP52T1_3

ATGACTCTCTCGCCTATTTCTACTTTTCTCGCCGGTGCTGCTATCCTTTACCTTGCTAGGTGGGTATGCATCG
AGATAAAC

Cytochrome P450 CYP539B1

ATGCTTATCGAGGCTGTCAATACAACGACGGTGGCTCTGGCCATTCCTGTCTGTCTGCTTTTGTTTGTCATTG
TCAACTGGCTCACG

Appendix C: LC-MS/MS analysis of *Beauveria bassiana* proteome

C.1 Introduction

As a consequence of the unsuccessful heterologous expression of P450s selected from *B. bassiana* in *E. coli* and *S. cerevisiae* and thus the absence of active biocatalysts that could be tested against drugs of industrial significance or that could be investigated for interesting reactions, it was considered important to provide at least information that might be useful for future projects. It was therefore decided to investigate the applicability of the LC-MS/MS Protein Identification technology as a tool for possible identification of specific P450s from *B. bassiana*. This approach would allow the identification of an enzyme in correspondence to a transformation reaction and offers immediate data regarding the nature of the enzyme. Therefore, we designed an experimental setup that exposed the fungus to selected drugs (known to be hydroxylated by *B. bassiana*) in the hope to initiate expression of particular cytochrome P450s. A negative control (*B. bassiana* grown in liquid medium without addition of substrate) was thought to act as reference for the analysis. The only reason that this technology has not been used and refined in the beginning of this project was the lack of available equipment and techniques.

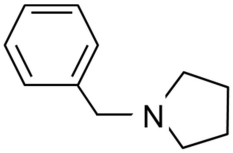
C.2 Materials and methods

The *Beauveria bassiana* strain used for proteomic analysis was ordered at CBS culture collection (CBS number: 209.27) and resuscitated on solid Oatmeal Agar (SigmarAldrich, 72.5 g per liter). The plates were incubated at 25 °C until an appreciable level of biomass was grown (7-10 days) and then stored at 4 °C for further use.

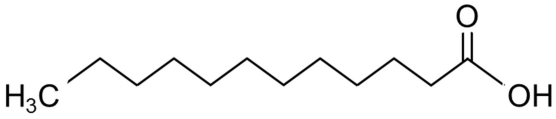
A loopful of the white fungal biomass from the oatmeal agar plate was transferred into 50 ml sterile liquid corn-steep medium (7.5 g/l corn steep solids, 10 g/l glucose, pH 4.8) using a nichrome loop. The fungus was then grown in an orbital shaker at 28 °C and 150 r.p.m. After 72 h of growth, substrate (see Figure C.1) dissolved in small amount of ethanol (0.5 ml) was added to the liquid medium (final concentration of 0.1 mg/ml) to initiate biotransformation.

GC samples were taken after 0 h, 4 h, 24 h and 48 h as described in 4.1.3.3 in order to confirm metabolization of the added substrate. On the basis of the GC analysis, cell harvest of freshly grown 50 ml liquid cultures that served as the actual sample for proteome

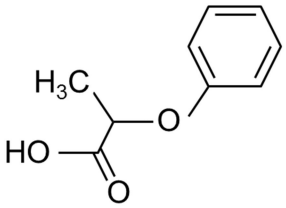
analysis took place 24 h after substrate addition. At that time, biotransformation was already in progress and still ongoing and should therefore account for a steady expression level of the protein(s) involved in the biotransformation of the substrate.



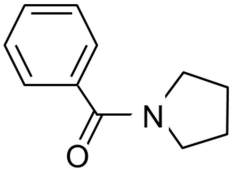
1-Benzylpyrolidine



Lauric acid



2-Phenoxypropionic acid



1-Benzoylpiperidine

Figure C.1: substrates added for biotransformation in *B. bassiana*

Cells were harvested by centrifugation at 4500 r.p.m. for 20 min. Supernatant was discarded and the cell pellet resuspended in 40 ml of buffer A (50 mM Tris/HCl pH 7.5, 300 mM NaCl). Cells were disrupted at high pressure (35000 psi) using a French Press and cell debris were then pelleted by centrifugation at 13000 r.p.m. (30 min, 4 °C). The cell pellet was discarded and the obtained cell lysate was 10 x concentrated. Concentrated proteins were prepared for SDS-PAGE as described in 2.7.3. The electrophoresis was stopped after approximately 8 min. The gel was stained in order to visualize the protein band which was then excised from the gel using a scalpel and send to the Technology Facility (University of York) for protein identification by LC-MS/MS. Analysis of prepared protein samples were performed as follows:

<u>In-gel Digestion</u>

In-gel tryptic digestion was performed after reduction with DTE and *S*-carbamidomethylation with iodoacetamide. Gel pieces were washed two times with 50 % (v:v) aqueous acetonitrile containing 25 mM ammonium bicarbonate, then once with acetonitrile and dried in a vacuum concentrator for 20 min. Sequencing-grade, modified porcine trypsin (Promega) was dissolved in 50 mM acetic acid, then diluted 5-fold with 25 mM ammonium bicarbonate to give a final trypsin concentration of 0.02 µg/µL. Gel pieces were rehydrated by adding 25 µL of trypsin solution, and after 10 min enough 25 mM

ammonium bicarbonate solution was added to cover the gel pieces. Digests were incubated overnight at 37 ^{o}C. Peptides were extracted by washing three times with 50 % (v:v) aqueous acetonitrile containing 0.1 % trifluoroacetic acid (v:v), before being dried down in a vacuum concentrator and reconstituting in aqueous 0.1 % trifluoroacetic acid (v:v).

LC-MS/MS

Samples were loaded onto a nanoAcquity UPLC system (Waters) equipped with a nanoAcquity Symmetry C_{18}, 5 µm trap (180 µm x 20 mm Waters) and a nanoAcquity HSS T3 1.8 µm C_{18} capillary column (75 µm x 250 mm, Waters). The trap wash solvent was 0.1 % (v/v) aqueous formic acid and the trapping flow rate was 10 µl/min. The trap was washed for 5 min before switching flow to the capillary column. Separation used a gradient elution of two solvents (solvent A: aqueous 0.1 % (v/v) formic acid; solvent B: acetonitrile containing 0.1 % (v/v) formic acid). The capillary column flow rate was 300 nl/min and the column temperature was 60 °C. The gradient profile was linear 2-30 % B over 125 min then linear 30-50 % B over 5 mins. All runs then proceeded to wash with 95 % solvent B for 2.5 min. The column was returned to initial conditions and re-equilibrated for 25 min before subsequent injections.

The nanoLC system was interfaced with a maXis HD LC-MS/MS system (Bruker Daltonics) with CaptiveSpray ionisation source (Bruker Daltonics). Positive ESI-MS and MS/MS spectra were acquired using AutoMSMS mode. Instrument control, data acquisition and processing were performed using Compass 1.7 software (microTOF control, Hystar and DataAnalysis, Bruker Daltonics). Instrument settings were: ion spray voltage: 1,450 V, dry gas: 3 l/min, dry gas temperature: 150 °C, ion acquisition range: *m/z* 150-2,000, MS spectra rate: 5 Hz, MS/MS spectra rate: 5 Hz at 2,500 cts to 20 Hz at 250,000 cts, cycle time: 1 s, quadrupole low mass: 300 *m/z*, collision RF: 1,400 Vpp, transfer time 120 ms. The collision energy and isolation width settings were automatically calculated using the AutoMSMS fragmentation table, absolute threshold 200 counts, preferred charge states: 2 – 4, singly charged ions excluded. A single MS/MS spectrum was acquired for each precursor and former target ions were excluded for 0.8 min unless the precursor intensity increased fourfold.

Database Searching

Tandem mass spectra were searched against the *Beauveria bassiana* subset of the UniProt database (22,786 sequences; 11,343,437 residues) using a locally-running copy of the

Mascot program (Matrix Science Ltd., version 2.5.1), through the Bruker ProteinScape interface (version 2.1). Results were passed through Mascot percolator to achieve a false discovery rate of <1% and further filtered to accept only peptides with an expect score of 0.05 or lower.

C.3 Results and discussion

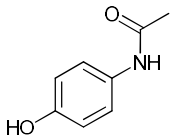
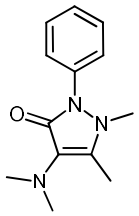
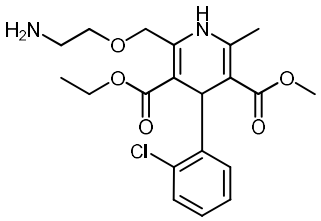
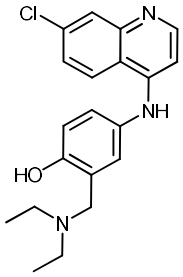
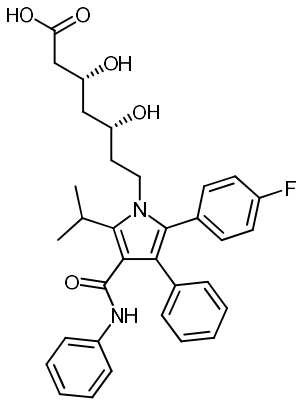
At first glace, the outcome of this experiment was rather disappointing as it was only possible to verify four P450 enzymes (*i.e.*, 2 x CYP52T1, CYP539B5 and CYP52A12) and the NADPH-cytochrome P450 reductase (see Table C.1). Moreover, two of the identified cytochromes P450 were also observed in the negative control, indicating a possible housekeeping function of these enzymes in the organism and thus were most likely not expressed as a direct response to the addition of the substrate. On the other hand, it was noticed that CYP52T1 as well as the NADPH-cytochrome P450 reductase seem to be more strongly induced in samples 2, 3 and 4 (Table C.1) which in turn leaves room to actually speculate about its physiological role regarding the metabolization of corresponding substrates in *B. bassiana*. It is important to emphasize the fact that this experiment was performed at the end of this project to explore the possibilities of proteome analysis for P450s in *B. bassiana*. Experimental setups as well as preparation of the protein samples need to be improved in order to aquire quantitative data that could lead to the identification of specific enzymes. Nonetheless, it could be demonstrated that it is possible to identify P450s in the cell lysate of *B. bassiana* and thus confirmed the potential use of high-resolution mass spectrometry for target selection in future studies.

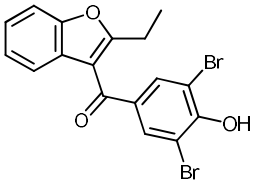
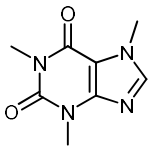
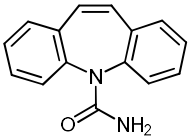
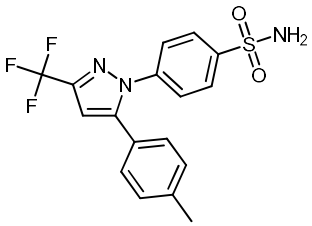
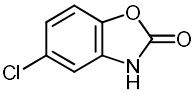
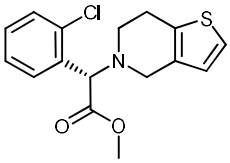
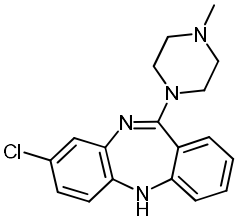
Table C.1: P450 hits on exposure to hydroxylation substrates in *Beauveria bassiana* using LC-MS analysis
sample 1: no substrate (negative control); sample 2: 1-benzylpyrolidine; sample 3: lauric acid; sample 4: 2-phenoxypropionic acid; sample 5: 1-benzoylpiperidine

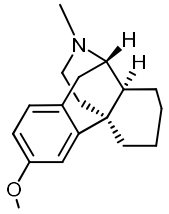
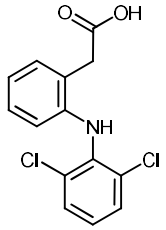
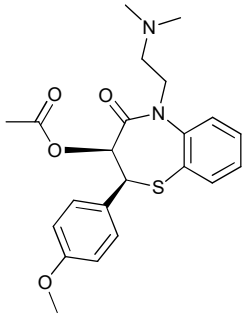
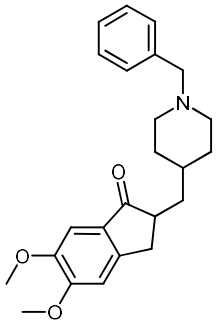
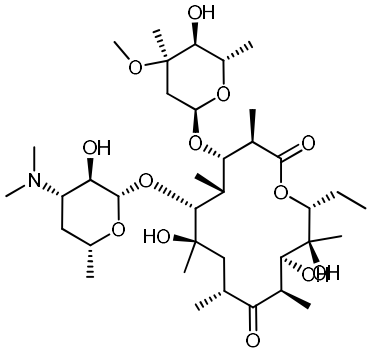
Sample	Family	Accession	Score	Mass	No. Spectra	No. Sequences	emPAI	Molar %	Description
1	443	J4UN70	60	76854	4	3	0.13	0.03	NADPH--cytochrome P450 reductase OS=Beauveria bassiana (strain ARSEF 2860)
1	501	J4WEG4	47	60824	3	2	0.1	0.02	Cytochrome P450 CYP52T1 OS=Beauveria bassiana (strain ARSEF 2860)
1	603	E2EAF9	35	59082	2	1	0.05	0.01	Cytochrome P450 monooxygenase CYP539B5 OS=Beauveria bassiana
2	240	A0A0A2V9Y1	102	76849	8	4	0.17	0.06	NADPH--cytochrome P450 reductase OS=Beauveria bassiana D1-5
2	249	J4WEG4	96	60824	4	3	0.16	0.06	Cytochrome P450 CYP52T1 OS=Beauveria bassiana (strain ARSEF 2860)
3	325	J4WEG4	83	60824	3	3	0.16	0.05	Cytochrome P450 CYP52T1 OS=Beauveria bassiana (strain ARSEF 2860)
3	533	J4UN70	43	76854	3	3	0.13	0.04	NADPH--cytochrome P450 reductase OS=Beauveria bassiana (strain ARSEF 2860)
4	223	A0A0A2V9Y1	102	76849	10	5	0.22	0.07	NADPH--cytochrome P450 reductase OS=Beauveria bassiana D1-5
4	396	J4WEG4	47	60824	4	2	0.1	0.03	Cytochrome P450 CYP52T1 OS=Beauveria bassiana (strain ARSEF 2860)
4	757	J5JHI3	16	60064	1	1	0.05	0.01	Cytochrome P450 CYP52T1 OS=Beauveria bassiana (strain ARSEF 2860)
5	370	A0A0A2V9Y1	60	76849	4	3	0.13	0.04	NADPH--cytochrome P450 reductase OS=Beauveria bassiana D1-5
5	615	E2EAF9	26	59082	1	1	0.05	0.01	Cytochrome P450 monooxygenase CYP539B5 OS=Beauveria bassiana
5	742	A0A0A2WAI9	18	60755	1	1	0.05	0.01	Cytochrome P450 52A12 OS=Beauveria bassiana D1-5
Crudely	Lower = more		Higher = more		Higher = more	Higher = more		Higher = more	

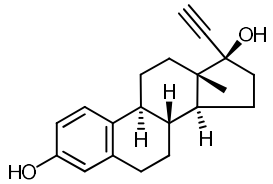
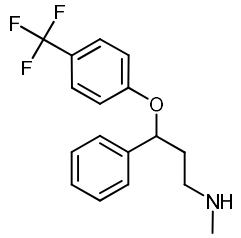
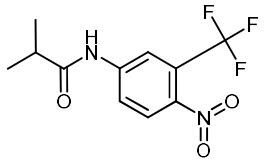
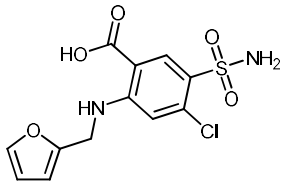
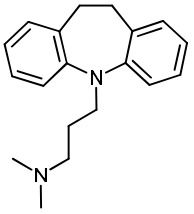
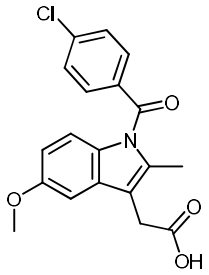
Appendix D

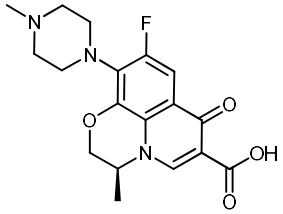
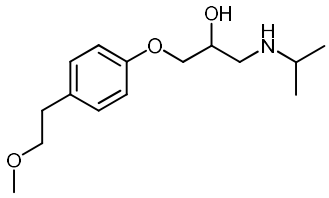
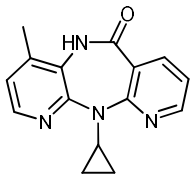
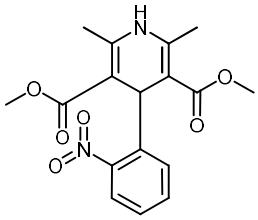
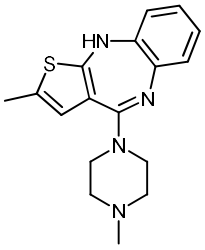
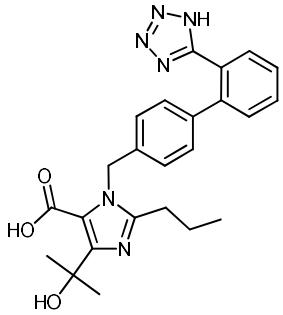
Table D.1: List of the 65 substrates used in activity assays of *Rhodococcus jostii*.

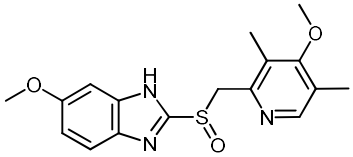
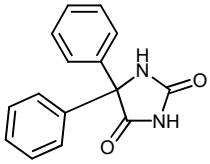
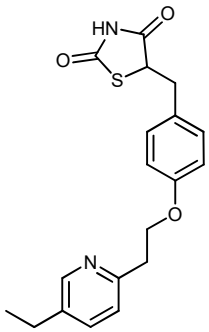
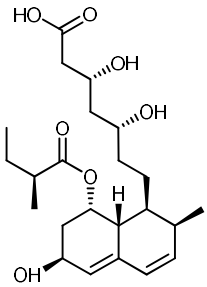
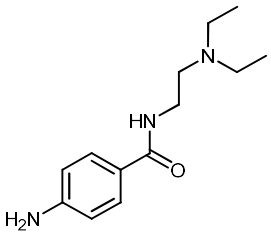
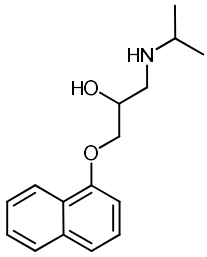
No.	Substrate	Structure	MW
1	Acetominophen		151.20
2	Aminopyrine		231.30
3	Amlodipine		408.90
4	Amodiaquine		355.90
5	Atorvastatin		558.60

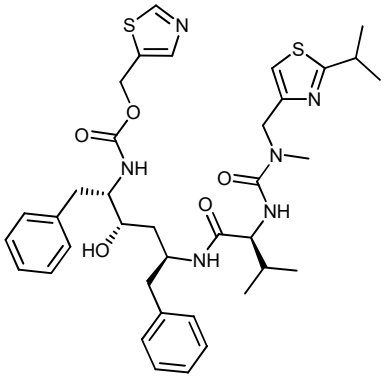
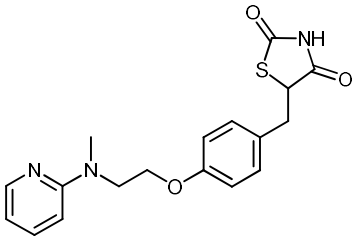
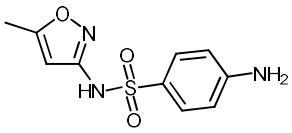
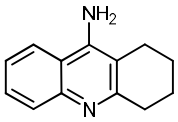
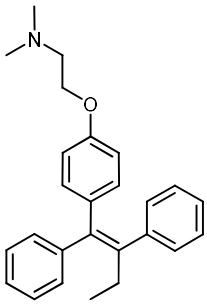
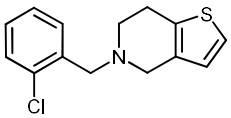
No.	Substrate	Structure	MW
6	Benzbromarone		424.10
7	Caffeine		194.20
8	Carbamazepine		236.30
9	Celecoxib		381.40
10	Chloroxazone		169.60
11	Clopridogrel		321.80
12	Clozapine		326.80

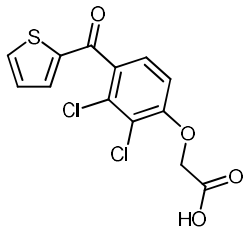
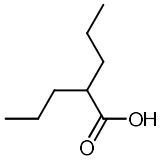
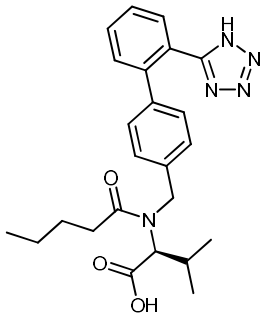
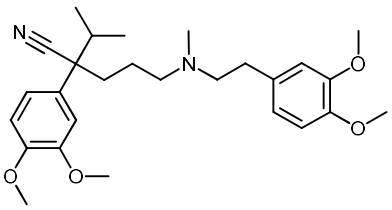
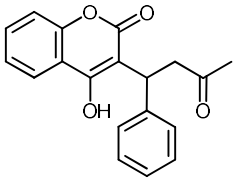
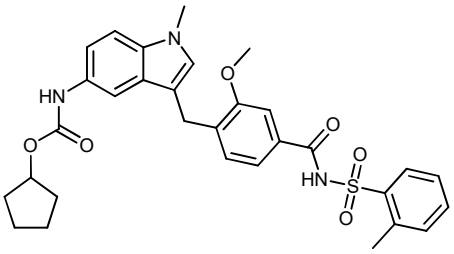
No.	Substrate	Structure	MW
13	Dextromethorphan		271.40
14	Dichlofenac		296.20
15	Diltiazem		414.50
16	Donepezil		379.50
17	Erythromycin		733.90

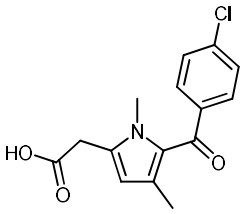
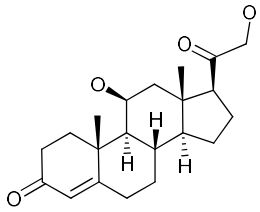
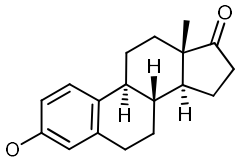
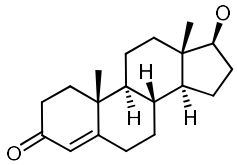
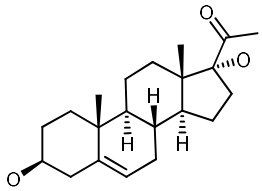
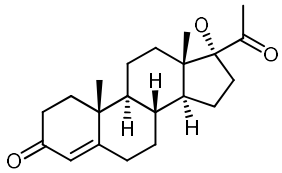
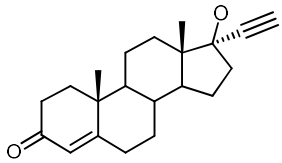
No.	Substrate	Structure	MW
18	Ethinylestradiol		296.40
19	Fluoxetine		309.30
20	Flutamide		276.20
21	Furosemide		330.70
22	Imipramine		280.40
23	Indomethacin		357.80

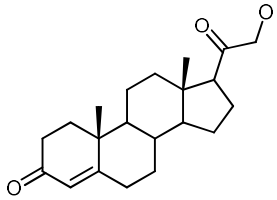
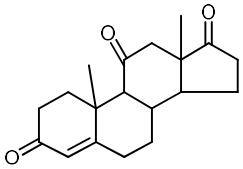
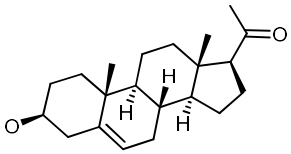
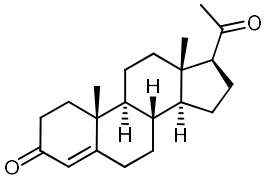
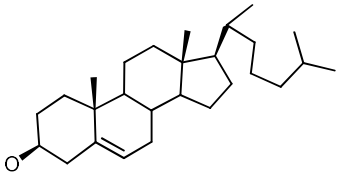
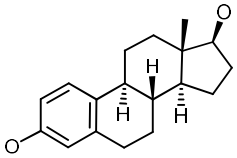
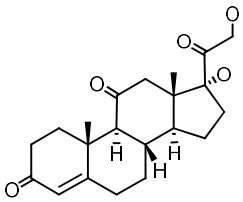
No.	Substrate	Structure	MW
24	Levofloxacin		361.40
25	Metoprolol		267.40
26	Nevirapine		266.30
27	Nifedipine		346.30
28	Olanzapine		312.40
29	Olmesartan		

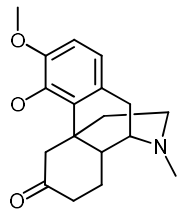
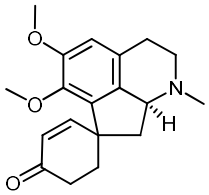
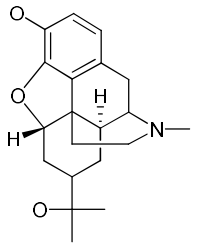
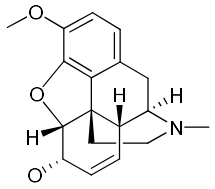
No.	Substrate	Structure	MW
30	Omeprazole		345.40
31	Phenytoin		252.30
32	Pioglitazone		356.40
33	Pravastatin		424.50
34	Procainamide		235.30
35	Propranolol		259.30

No.	Substrate	Structure	MW
36	Ritonavir		
37	Rosiglitazone		357.40
38	Sulfamethoxazole		253.30
39	Tacrine		198.30
40	Tamoxifen		371.50
41	Ticlodipine		263.80

No.	Substrate	Structure	MW
42	Tienilic acid		331.20
43	Valproic acid		144.20
44	Valsartan		435.50
45	Verapamil		454.60
46	Warfarin		308.30
47	Zafirlukast		575.70

No.	Substrate	Structure	MW
48	Zomepirac		291.70
49	Cortisone		346.46
50	Estrone		270.37
51	Testosterone		288.43
52	17-a-hydroxy-pregnenolone		332.48
53	17-a-progesterone		330.47
54	Ethisterone		312.45

No.	Substrate	Structure	MW
55	Deoxycortico-sterone		330.47
56	Andrenosterone		300.40
57	Pregnenolone		316.48
58	Progesterone		314.47
59	Cholesterol		386.66
60	b-estradiol		272.39
61	Corticosterone		360.45

No.	Substrate	Structure	MW
62	(+-)-Dihydro-thebainon		301.38
63	(8a'S)-5',6'-dimethoxy-1'-methyl-2',3',8',8a'-tetrahydro-1'H-spiro[cyclohex[2]ene-1,7'-cyclopenta[ij]isoquinolin]-4-one		313.40
64	(5a,14a)-7-(2-hydroxypropan-2-yl)-17-methyl-4,5-epoxymorphinan-3-ol		329.43
65	(5a,6a)-3-methoxy-17-methyl-7,8-didehydro-4,5-epoxymorphinan-6-ol		299.36

Abbreviations

4CL	4-coumarate CoA ligase
A	alanine or adenine
aa	amino acids
Abs	optical absorbance
ADE 2d	selection marker for adenine auxotrphy
ADR	adrenoxodin reductase
ADS	amorphadiene synthase
ADX	adrenodoxin
Ala	alanine
ALA	δ-aminolevulinic acid
AMP	adenosine monophosphate
*amp*R	ampicillin resistence gene
ANS	anthocyanidin synthase
araB	gene encoding for AraB
Arg	arginine
Asp	aspartic acid
ATR1	*Arabidopsis thaliana* cytochrome P450 reductase 1
BCAs	biological control agents
BLAST	Basic Local Alignment Search Tool
bp	base pairs
BSA	bovine serum albumine
BSTFA	N,O-bis(trimethylsilyl)trifluoroacetamide
C	cysteine or cytosine
C4H	cinnamic acid 4-hydroxylase
cAMP	cyclic adenosine monophosphate
*cam*R	chloramphenicol resistence gene

*can*R	resistance to canavanine sulphate
CAP	cyclic AMP receptor protein
CBS	Centraalbureau voor Schimmelcultures
cDNA	complementary DNA
CHI	chalcone isomerase
CHS	chalcone synthase
CO	carbon monoxide
CPA	cyclophosphamide
CPR	cytochrome P450 reductase
CYP	cytochrome P450
cyr+	encodes adenylate cyclase: required for cAMP production and cAMP-dependent protein kinase signaling
Cys	cysteine
D	aspartic acid
Da	dalton
dATP	deoxyadenosine triphosphate
DCDD	dichlorinated dibenzo-p-dioxins
dcm	cytosine methylation at second C of CCWGG sites
DE3	Lysogen that encodes T7 RNA polymerase
DFR	dihydroflavonol 4-reductase
DMAPP	dimethyl allyl pyrophosphate
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTP	dinucleotide triphosphate
dTTP	deoxythymidine triphosphate
E	glutamic acid
e.g.	*exempli gratia*

EDTA	ethylenediaminetetraacetic acid
endA1	gene encoding DNA-specific endonuclease I
ER	endoplasmic reticulum
et al.	*et alii*
etc.	*et cetera*
F’	F plasmid: a self-transmissible, low-copy plasmid used for the generation of single-stranded DNA when infected with M13 phage; may contain a resistance marker to allow maintenance and will often carry the lacI and lacZΔM15 genotypes
F^-	host strain does not carry the F plasmid
f1 ori	*f1 phage origin of replication*
F3'5'H	flavonoid 3’,5‘-hydroxylase
F3H	flavanone 3-hydroxylase
F3'H	flavonoid 3‘-hydroxylase
FAD	flavin adenine dinucleotide
FdR	ferredoxin reductase
Fdx	ferredoxin
Fe	iron, *ferrum*
FMN	flavin mononucleotide
FPP	farnesyl pyrophosphate
G	glycine or guanine
g	gram
Ga	gigaannum: 10^9 (1,000,000,000) years
Gal4p	novel yeast gene required for glucose repression
GAL10-CYC1	hybrid yeast promoter
GDEPT	gene directed enzyme pro-drug therapy
Glu	glutamic acid
Gly	glycine

GPP	geranyl pyrophosphate
gyrA96	DNA gyrase mutant: produces resistance to nalidixic acid
i.e.	*id est*
h	hour(s)
his	6x histidines residues
HiTEL	High Throughput Expression Laboratory
HMG-CoA	3-hydroxy-3-methylglutaryl-coenzyme A
HRV	human rhinovirus
HSD	human 3β-hydroxysteroid dehydrogenase/isomerase
HsdR17	endonuclease R, host restriction of foreign DNA
hsdSB	specificity determinant for HsdM and HsdR (both restriction and methylation of certain sequences is deleted from the strain)
IPP	isopentenyl pyrophosphate
IPTG	isopropyl β-D-1-thiogalactopyranoside
kb	kilo base pairs
kDa	kilodaltons
kJ	kilojoules
L	leucine
l	litre
LacI	repressor gene for IPTG induction
lacIqZΔM15	Allows α-complementation for blue/white selection of recombinant colonies in lacZ mutant hosts
LacZ	ß-galactosidase
LB	Lysogeny Broth
Leu	leucine
LIC	Ligation independent cloning
LICRED	Ligation Independent Cloning plasmid with P450Rhf reductase domain
M	molar

Ma	megaannus: 10^6 (1,000,000) years
MATa	mating type locus in yeats: MATa haploids express the genes *a1* and *a2* from the MAT locus
min	minutes
MTBE	methyl tert-butyl ether
MW	molecular weight
NAD^+	nicotinamide adenine dinucleotide, oxidised form
NADH	nicotinamide adenine dinucleotide, reduced form
$NADP^+$	nicotinamide adenine dinucleotide phosphate, oxidised form
NADPH	nicotinamide adenine dinucleotide phosphate, reduced form
NCBI	National Center for Biotechnology Information
nm	nanometers
OD_{600}	optical density at 600 nm
OFOR	2-oxoacid:ferredoxin oxidoreductase
oLac	Lac operator
ompT	mutation in the protease VII protein in outer membrane: reduces degradation of heterologous expressed proteins
ORF	open reading frame
ori	origin of replication
P	proline
PAGE	polyacrylamide gel electrophoresis
PAH	polycyclic aromatic hydrocarbons
PAL	phenylalanine ammonia lyase
pBR322 ori	origin of replication of the plasmid pBR322
PCBs	polychlorinated biphenyls
PCDD	polychlorinated dibenzo-p-dioxins
PCR	polymerase chain reaction
PEG	poly-ethylene glycol

pGAL	galactose promoter
PGH_2	prostaglandin H2
Phe	phenylalanine
pRARE2	plasmid which carries seven rare-codon tRNA genes
Pro	proline
$proA^+B^+$	mutations in proline biosynthesis; strain requires proline for growth in minimal medium
pT7	T7 promotor
R	arginine
RBS	ribosome binding site
RDX	hexahydro -1,3,5-trinitro -1,3,5-triazine
recA1	mutation in a gene responsible for general recombination of DNA
relA1	locus regulates the coupling between transcription and translation; relaxed phenotype: permits RNA synthesis in absence of protein synthesis
RNA	ribose nucleic acid
r.p.m.	revolutions per minute
S	serine or sulphur
SDS	sodium dodecyl sulfate
s	second(s)
Ser	serine
SGI	synthetic media for yeast
SOC	Super Optimal Broth
SRP	signal recognition particle
SRS	substrate recognition site
STY	space–time yield
subE44	tRNA glutamine suppressor of amber
T	threonine or thymine

TAE	Tris-acetate-EDTA
TBST	mixture of Tris-buffered saline and Tween 20
TEG	Tris buffer, EDTA and Glucose
TEK	Tris buffer, EDTA and KCl
TEMED	N, N, N', N'-Tetramethylethylendiamine
TES	Tris buffer, EDTA and sorbitol
*tet*R	tetracycline resistance gene
thi-1	requires thiamine for growth on minimal media
Thr	threonine
Tn10	transposon carrying tetracycline resistance
tPGK	phosphoglycerate kinase terminator
TriCDD	trichlorinated dibenzo-p-dioxins
Tris	2-amino-2-hydroxymethyl-1,3-propanediol
tT7	T7 terminator
TXA_2	tromboxane A2
UAS	upstream activation sequence
URA3	selection marker for uracil auxotrophy
UV	ultra violet (light)
v/v	volume to volume ratio
w/v	weight to volume ratio
YPGA	yeast medium containing yeast extract, bactopeptone, glucose and adenine
YPGE	yeast medium containing yeast extract, bactopeptone, glucose and ethanol

References

[1] Črešnar B., Petrič S. (2011) Cytochrome P450 enzymes in the fungal kingdom. *Biochim Biophys Acta*. **1814** (1): 29-35.

[2] Sono M., Roach M.P., Coulter E.D., Dawson J.H. (1996) Heme-containing oxygenases. *Chem Rev*. **96** (7): 2841-2888.

[3] Claude A. (1943) The Constitution of Protoplasm. *Science*. **97** (2525): 451-456

[4] Horecker B.L. (1950) Triphosphopyridine nucleotide-cytochrome c reductase in liver. *J Biol Chem*. **183**: 593-605.

[5] Strittmatter C.F., Ball E.G. (1952) A Hemochromogen Component of Liver Microsomes. *Proc Natl Acad Sci USA*. **38** (1): 19–25.

[6] Strittmatter P., Velick S.F. (1956) A microsomal cytochrome reductase specific for diphosphopyridine nucleotide. *J Biol Chem*. **221** (1): 277-286.

[7] Chance B., Williams G. R. (1954) Kinetics of cytochrome b5 in rat liver microsomes. *J Biol Chem*. **209** (2): 945-951.

[8] Axelrod J. (1955) The enzymatic demethylation of ephedrine. *J Pharmacol Exp Ther*. **114** (4): 430-438.

[9] Axelrod J. (1955) The enzymatic deamination of amphetamine (benzedrine). J Biol Chem. 214 (2): 753-763.

[10] Ryan K.J., Engel L.L. (1957) Hydroxylation of steroids at carbon 21. *J Biol Chem*. 225 (1): 103-114.

[11] Klingenberg M. (1958) Pigments of rat liver microsomes. *Arch Biochem Biophys*. **75** (2): 376-386.

[12] Omura T., Sato R. (1962) A new cytochrome in liver microsomes. *J Biol Chem*. **237**: 1375-1376.

[13] Estabrook R.W., Cooper D.Y., Rosenthal O. (1963) The Light Reversible Carbon Monoxide Inhibition of the Steroid C21-Hydroxylase System of the Adrenal Cortex. *Biochem Z*. **338**: 741-755.

[14] Cooper D.Y., Levin S., Narasimhulu S., Rosenthal O., Estabrook R.W. (1965) Photochemical Action Spectrum of the Terminal Oxidase of Mixed Function Oxidase Systems. *Science*. **147** (3656): 400-402.

[15] Appleby A.C. (1967) A soluble haemoprotein P 450 from nitrogen-fixing Rhizobium bacteroids. *Biochim Biophys Acta*. **147** (2): 399-402.

[16] 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** (3): 490-497.

[17] Waterman M.R., Mason H.S. (1970) The redox potential of liver cytochrome P-450. *Biochem Biophys Res Commun*. **39** (3): 450-454.

[18] Stern J.O., Peisach J. (1974) A model compound study of the CO-adduct of cytochrome P-450. *J Biol Chem*. **249** (23): 7495-7498.

[19] Haugen D.A., van der Hoeven T.A., Coon M.J. (1975) Purified liver microsomal cytochrome P-450. Separation and characterization of multiple forms. *J Biol Chem*. **250** (9): 3567-3570.

[20] Haugen D.A., Coon M.J. (1976) Properties of electrophoretically homogeneous phenobarbital-inducible and beta-naphthoflavone-inducible forms of liver microsomal cytochrome P-450. *J Biol Chem*. **251** (24): 7929-7939.

[21] Haniu M., Armes L.G., Tanaka M., Yasunobu K.T., Shastry B.S., Wagner G.C., Gunsalus I.C. (1982) The primary structure of the monoxygenase cytochrome P450CAM. *Biochem Biophys Res Commun*. **105** (3): 889-894.

[22] Fujii-Kuriyama Y., Mizukami Y., Kawajiri K., Sogawa K., Muramatsu M. (1982) Primary structure of a cytochrome P-450: coding nucleotide sequence of phenobarbital-inducible cytochrome P-450 cDNA from rat liver. *Proc Natl Acad Sci USA*. **79** (9): 2793-2797.

[23] Distlerath L.M., Reilly P.E., Martin M.V., Davis G.G., Wilkinson G.R., Guengerich F.P. (1985) Purification and characterization of the human liver cytochromes P-450 involved in debrisoquine 4-hydroxylation and phenacetin O-deethylation two prototypes for genetic polymorphism in oxidative drug metabolism. *J Biol Chem*. **260** (15): 9057-9067.

[24] Shimada T., Shea J.P., Guengerich F.P. (1985) A convenient assay for mephenytoin 4-hydroxylase activity of human liver microsomal cytochrome P-450. *Anal Biochem*. **147** (1): 174-179.

[25] Guengerich F.P., Martin M.V., Beaune P.H., Kremers P., Wolff T., Waxman D.J. (1986) Characterization of rat and human liver microsomal cytochrome P-450 forms involved in nifedipine oxidation a prototype for genetic polymorphism in oxidative drug metabolism. *J Biol Chem*. **261** (11): 5051-5060.

[26] Narhi L.O., Fulco A.J. (1986) Characterization of a catalytically self-sufficient 119000-dalton cytochrome P-450 monooxygenase induced by barbiturates in Bacillus megaterium. *J Biol Chem*. **261** (16): 7160-7169.

[27] Nebert D.W., Adesnik M., Coon M.J., Estabrook R.W., Gonzalez F.J., Guengerich F.P., Gunsalus I.C., Johnson E.F., Kemper B., Levin W., Phillips I.R., Sato R., Waterman M.R. (1987) The P450 gene superfamily: recommended nomenclature. *DNA*. **6** (1): 1-11.

[28] Williams P.A., Cosme J., Sridhar V., Johnson E.F., McRee D.E. (2000) Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity. *Mol Cell*. **5** (1): 121-131.

[29] Harding B.W., Wong S.H., Nelson D.H. (1964) Carbon monoxide-combining substances in rat adrenal. *Biochim Biophys Acta*. **92**: 415-417.

[30] Lindenmayer A., Smith L. (1964) Cytochromes and other pigments of baker's yeast grown aerobically and anaerobically. *Biochim Biophys Acta*. **93**: 445-461.

[31] Suzuki K., Kimura T. (1965) An Iron Protein As A Component of Steroid 11-Beta-Hydroxylase Complex. *Biochem Biophys Res Commun*. **19**: 340-345.

[32] Omura T., Sanders E., Estabrook R.W., Cooper D.Y., Rosenthal O. (1966) Isolation from adrenal cortex of a nonheme iron protein and a flavoprotein functional as a reduced triphosphopyridine nucleotide-cytochrome P-450 reductase. *Arch Biochem Biophys*. **117** (3): 660-673.

[33] Katagiri M., Ganguli B.N., Gunsalus I.C. (1968) A soluble cytochrome P-450 functional in methylene hydroxylation. J Biol Chem. 243 (12): 3543-3546.

[34] Oeda K., Sakaki T., Ohkawa H. (1985) Expression of rat liver cytochrome P-450MC cDNA in Saccharomyces cerevisiae. *DNA*. **4** (3): 203-210.

[35] Koga H., Rauchfuss B., Gunsalus I.C. (1985) P450cam gene cloning and expression in Pseudomonas putida and Escherichia coli. *Biochem Biophys Res Commun*. **130** (1): 412-417.

[36] Poulos T.L., Finzel B.C., Gunsalus I.C., Wagner G.C., Kraut J. (1985) The 2.6-A crystal structure of Pseudomonas putida cytochrome P-450. *J Biol Chem*. **260** (30): 16122-16130.

[37] Larson J.R., Coon M.J., Porter T.D. (1991) Alcohol-inducible cytochrome P-450IIE1 lacking the hydrophobic NH2-terminal segment retains catalytic activity and is membrane-bound when expressed in Escherichia coli. *J Biol Chem*. **266** (12): 7321 7324.

[38] Barnes H.J., Arlotto M.P., Waterman M.R. (1991) Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in Escherichia coli. *Proc Natl Acad Sci USA*. **88** (13): 5597-5601.

[39] Coon M.J. Cytochrome P450: nature's most versatile biological catalyst. *Annu Rev Pharmacol Toxicol* vol. 45 pp. 1-25 2005.

[40] Nebert D.W., Nelson D.R., Adesnik M., Coon M.J., Estabrook R.W., Gonzalez F.J., Guengerich F.P., Gunsalus I.C., Johnson E.F., Kemper B., *et al.* (1989) The P450 superfamily: updated listing of all genes and recommended nomenclature for the chromosomal loci. *DNA*. **8** (1): 1-13.

[41] Nebert D.W., Nelson D.R., Coon M.J., Estabrook R.W., Feyereisen R., Fujii-Kuriyama Y., Gonzalez F.J., Guengerich F.P., Gunsalus I.C., Johnson E.F., *et al.* (1991) The P450 superfamily: update on new sequences gene mapping and recommended nomenclature. *DNA Cell Biol* **10** (1): 1-14.

[42] Nelson D.R. (2009) The cytochrome p450 homepage. *Hum Genomics*. **4** (1): 59-65.

[43] Hasemann C.A., Kurumbail R.G., Boddupalli S.S., Peterson J.A., Deisenhofer J. (1995) Structure and function of cytochromes P450: a comparative analysis of three crystal structures. *Structure*. **3** (1): 41-62.

[44] Graham S.E., Peterson J.A. (1999) How similar are P450s and what can their differences teach us? *Arch Biochem Biophys*. **369** (1): 24-29.

[45] Poulos T.L., Johnson E. F. (2005) Cytochrome P450: Structure Mechanism and Biochemistry in *Structures of cytochrome P450 enzymes*. Kluwer Academic/Plenum Publishers, New York, pp 87–114.

[46] Raman C.S., Li H., Martásek P., Král V., Masters B.S., Poulos T.L. (1998) Crystal structure of constitutive endothelial nitric oxide synthase: a paradigm for pterin function involving a novel metal center. *Cell*. **95** (7): 939-950.

[47] Pant K., Bilwes A.M., Adak S., Stuehr D.J., Crane B.R. (2002) Structure of a nitric oxide synthase heme protein from Bacillus subtilis. *Biochemistry*. **41** (37): 11071 11079.

[48] Munro A.W., Girvan H.M., McLean K.J. (2007) Variations on a (t)heme--novel mechanisms redox partners and catalytic functions in the cytochrome P450 superfamily. *Nat Prod Rep*. **24** (3): 585-609.

[49] Poulos T.L., Finzel B.C., Howard A.J. (1987) High-resolution crystal structure of cytochrome P450cam. *J Mol Biol*. **195** (3): 687-700.

[50] Werck-Reichhart D., Feyereisen R. (2000) Cytochromes P450: a success story. *Genome Biol*. **1** (6): reviews3003.1–3003.9.

[51] Yoshikawa K., Noguti T., Tsujimura M., Koga H., Yasukochi T., Horiuchi T., Go M. (1992) Hydrogen bond network of cytochrome P-450cam: a network connecting the heme group with helix K. *Biochim Biophys Acta*. **1122** (1): 41-44.

[52] Kitamura M., Buczko E., Dufau M.L. (1991) Dissociation of hydroxylase and lyase activities by site-directed mutagenesis of the rat P45017 alpha. *Mol Endocrinol*. **5** (10): 1373-1380.

[53] Chen S., Zhou D. (1992) Functional domains of aromatase cytochrome P450 inferred from comparative analyses of amino acid sequences and substantiated by site-directed mutagenesis experiments. *J Biol Chem*. **267** (31): 22587-22594.

[54] Shimizu T., Tateishi T., Hatano M., Fujii-Kuriyama Y. (1991) Probing the role of lysines and arginines in the catalytic function of cytochrome P450d by site-directed mutagenesis. Interaction with NADPH-cytochrome P450 reductase. *J Biol Chem*. **266** (6): 3372-3375.

[55] Furuya H., Shimizu T., Hirano K., Hatano M., Fujii-Kuriyama Y., Raag R., Poulos T.L. (1989) Site-directed mutageneses of rat liver cytochrome P-450d: catalytic activities toward benzphetamine and 7-ethoxycoumarin. *Biochemistry*. **28** (17): 6848-6857.

[56] Peterson J.A., Graham S.E. (1998) A close family resemblance: the importance of structure in understanding cytochromes P450. *Structure*. **6** (9): 1079-1085.

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

[58] Oliver C.F., Modi S., Primrose W.U., Lian L.Y., Roberts G.C. (1997) Engineering the substrate specificity of Bacillus megaterium cytochrome P-450 BM3: hydroxylation of alkyl trimethylammonium compounds. *Biochem J*. **327** (2): 537 544.

[59] Graham-Lorence S., Amarneh B., White R.E., Peterson J.A., Simpson E.R. (1995) A three-dimensional model of aromatase cytochrome P450. *Protein Sci*. **4** (6): 1065 1080.

[60] Peterson J.A., Ebel R.E., O'Keeffe D.H., Matsubara T., Estabrook R.W. (1976) Temperature dependence of cytochrome P-450 reduction. A model for NADPH-cytochrome P-450 reductase:cytochrome P-450 interaction. *J Biol Chem*. **251** (13): 4010-4016.

[61] Stayton P.S., Sligar S.G. (1990) The cytochrome P-450cam binding surface as defined by site-directed mutagenesis and electrostatic modeling. *Biochemistry*. **29** (32): 7381-7386.

[62] Stayton P.S., Poulos T.L., Sligar S.G. (1989) Putidaredoxin competitively inhibits cytochrome b5-cytochrome P-450cam association: a proposed molecular model for a cytochrome P-450cam electron-transfer complex. *Biochemistry*. **28** (20): 8201-8205.

[63] Koga H., Sagara Y., Yaoi T., Tsujimura M., Nakamura K., Sekimizu K., Makino R., Shimada H., Ishimura Y., Yura K., Goc M., Ikeguchid M., Horiuchi T. (1993) Essential role of the Arg112 residue of cytochrome P450cam for electron transfer from reduced putidaredoxin. *FEBS Lett*. **331** (1-2): 109-113.

[64] Wada A., Waterman M.R. (1992) Identification by site-directed mutagenesis of two lysine residues in cholesterol side chain cleavage cytochrome P450 that are essential for adrenodoxin binding. *J Biol Chem*. **267** (32): 22877-22882.

[65] Bernhardt R., Pommerening K., Ruckpaul K. (1987) Modification of carboxyl groups on NADPH-cytochrome P-450 reductase involved in binding of cytochromes c and P-450 LM2. *Biochem Int*. **14** (5): 823-832.

[66] Coghlan V.M., Vickery L.E. (1991) Site-specific mutations in human ferredoxin that affect binding to ferredoxin reductase and cytochrome P450scc. *J Biol Chem*. **266** (28): 18606-18612.

[67] Geren L.M., O'Brien P., Stonehuerner J., Millett F. (1984) Identification of specific carboxylate groups on adrenodoxin that are involved in the interaction with adrenodoxin reductase. *J Biol Chem*. **259** (4): 2155-2160.

[68] Schiffler B., Bernhardt R. (2003) Bacterial (CYP101) and mitochondrial P450 systems-how comparable are they? *Biochem Biophys Res Commun*. **312** (1): 223 228.

[69] Taylor M., Lamb D.C., Cannell R., Dawson M., Kelly S.L. (1999) Cytochrome P450105D1 (CYP105D1) from Streptomyces griseus: heterologous expression activity and activation effects of multiple xenobiotics. *Biochem Biophys Res Commun*. **263** (3): 838-842.

[70] Serizawa N., Matsuoka T. (1991) A two component-type cytochrome P-450 monooxygenase system in a prokaryote that catalyses hydroxylation of ML-236B to pravastatin a tissue-selective inhibitor of 3-hydroxy-3-methylglutaryl coenzyme A reductase. *Biochim Biophys Acta*. **1084** (1): 35-40.

[71] Williams P.A., Cosme J., Vinkovic D.M., Ward A., Angove H.C., Day P.J., Vonrhein C., Tickle I.J., Jhoti H. (2004) Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. *Science*. **305** (5684): 683-686.

[72] Hawkes D.B., Adams G.W., Burlingame A.L., Ortiz de Montellano P.R., De Voss J.J. (2002) Cytochrome P450(cin) (CYP176A) isolation expression and characterization. *J Biol Chem*. **277** (31): 27725-27732.

[73] Meharenna Y.T., Li H., Hawkes D.B., Pearson A.G., De Voss J., Poulos T.L. (2004) Crystal structure of P450cin in a complex with its substrate 18-cineole a close structural homologue to D-camphor the substrate for P450cam. *Biochemistry*. **43** (29): 9487-9494.

[74] Puchkaev A.V., Wakagi T., Ortiz de Montellano P.R. (2002) CYP119 plus a Sulfolobus tokodaii strain 7 ferredoxin and 2-oxoacid:ferredoxin oxidoreductase constitute a high-temperature cytochrome P450 catalytic system. *J Am Chem Soc*. **124** (43): 12682-12683.

[75] Zhang Q., Iwasaki T., Wakagi T., Oshima T. (1996) 2-oxoacid:ferredoxin oxidoreductase from the thermoacidophilic archaeon Sulfolobus sp. strain 7. *J Biochem.* **120**: (3): 587-599.

[76] Jackson C.J., Lamb D.C., Marczylo T.H., Warrilow A.G., Manning N.J., Lowe D.J., Kelly D.E., Kelly S.L. (2002) A novel sterol 14alpha-demethylase/ferredoxin fusion protein (MCCYP51FX) from Methylococcus capsulatus represents a new class of the cytochrome P450 superfamily. *J Biol Chem*. **277** (49): 46959-46965.

[77] Rylott E.L., Jackson R.G., Edwards J., Womack G.L., Seth-Smith H.M., Rathbone D.A., Strand S.E., Bruce N.C. (2006) An explosive-degrading cytochrome P450 activity and its targeted application for the phytoremediation of RDX. *Nat Biotechnol.* **24** (2): 216-219.

[78] Seth-Smith H.M., Rosser S.J., Basran A., Travis E.R., Dabbs E.R., Nicklin S., Bruce N.C. (2002)Cloning sequencing and characterization of the hexahydro-135-Trinitro-135-triazine degradation gene cluster from Rhodococcus rhodochrous. *Appl Environ Microbiol.* **68** (10): 4764-4771.

[79] Roberts G.A., Grogan G., Greter A., Flitsch S.L., Turner N.J. (2002) Identification of a new class of cytochrome P450 from a Rhodococcus sp. *J Bacteriol*. **184** (14): 3898-3908.

[80] De Mot R., Parret A.H. (2002) A novel class of self-sufficient cytochrome P450 monooxygenases in prokaryotes. *Trends Microbiol*. **10** (11): 502-508.

[81] Warman A.J., Roitel O., Neeli R., Girvan H.M., Seward H.E., Murray S.A., McLean K.J., Joyce M.G., Toogood H., Holt R.A., Leys D., Scrutton N.S., Munro A.W. (2005) Flavocytochrome P450 BM3: an update on structure and mechanism of a biotechnologically important enzyme. *Biochem Soc Trans*. **33** (4): 747-753.

[82] Kitazume T., Takaya N., Nakayama N., Shoun H. (2000) Fusarium oxysporum fatty-acid subterminal hydroxylase (CYP505) is a membrane-bound eukaryotic counterpart of Bacillus megaterium cytochrome P450BM3. *J Biol Chem*. **275** (50): 39734-39740.

[83] Daiber A., Shoun H., Ullrich V. (2005) Nitric oxide reductase (P450nor) from Fusarium oxysporum. *J Inorg Biochem*. **99** (1): 185-193.

[84] Wang L.H., Kulmacz R.J. (2002) Thromboxane synthase: structure and function of protein and gene. *Prostaglandins Other Lipid Mediat*. **68-69**: 409-422.

[85] Grechkin A.N. (2002) Hydroperoxide lyase and divinyl ether synthase. Prostaglandins Other Lipid Mediat. 68-69: 457-470.

[86] Tijet N., Brash A.R. (2002) Allene oxide synthases and allene oxides. *Prostaglandins Other Lipid Mediat*. **68-69**: 423-431.

[87] Degtyarenko K.N., Archakov A.I. (1993) Molecular evolution of P450 superfamily and P450-containing monooxygenase systems. *FEBS Lett.* **332** (1-2): 1-8.

[88] Omura T. (2006) Mitochondrial P450s. *Chem Biol Interact*. **163** (1-2): 86-93.

[89] Black S.D., Coon M.J. (1982) Structural features of liver microsomal NADPH-cytochrome P-450 reductase. Hydrophobic domain hydrophilic domain and connecting region. *J Biol Chem*. **257** (10): 5929-5938.

[90] Schenkman J.B., Greim H. (1993) *Cytochrome P450*. (Handbook Expl. Pharmacol. vol. 105) Springer-Verlag, Berlin

[91] Lewis D.F., Hlavica P. (2000) Interactions between redox partners in various cytochrome P450 systems: functional and structural aspects. *Biochim Biophys Acta*. **1460** (2-3): 353-374.

[92] Wright R.L., Harris K., Solow B., White R.H., Kennelly P.J. (1996) Cloning of a potential cytochrome P450 from the archaeon Sulfolobus solfataricus. *FEBS Lett*. **384** (3): 235-239.

[93] Oku Y., Ohtaki A., Kamitori S., Nakamura N., Yohda M., Ohno H., Kawarabayasi Y. (2004) Structure and direct electrochemistry of cytochrome P450 from the thermoacidophilic crenarchaeon Sulfolobus tokodaii strain 7. *J Inorg Biochem*. **98** (7): 1194-1199.

[94] Fukuda E., Kino H., Matsuzawa H., Wakagi T. (2001) Role of a highly conserved YPITP motif in 2-oxoacid:ferredoxin oxidoreductase: heterologous expression of the gene from Sulfolobus sp.strain 7 and characterization of the recombinant and variant enzymes. *Eur J Biochem*. **268** (21): 5639-5646.

[95] Kojoh K., Matsuzawa H., Wakagi T. (1999) Zinc and an N-terminal extra stretch of the ferredoxin from a thermoacidophilic archaeon stabilize the molecule at high temperature. *Eur J Biochem*. **264** (1): 85-91.

[96] Puchkaev A.V., Ortiz de Montellano P.R. (2005) The Sulfolobus solfataricus electron donor partners of thermophilic CYP119: an unusual non-NAD(P)H-dependent cytochrome P450 system. *Arch Biochem Biophys*. **434** (1): 169-177.

[97] Yamazaki H. (2014) *Fifty Years of Cytochrome P450 Research*. Springer, Tokyo.

[98] Seth-Smith H.M., Rosser S.J., Basran A., Travis E.R., Dabbs E.R., Nicklin S., Bruce N.C. (2002) Cloning sequencing and characterization of the hexahydro-135-Trinitro-135-triazine degradation gene cluster from Rhodococcus rhodochrous. *Appl Environ Microbiol*. **68** (10): 4764-4771.

[99] Roberts G.A., Celik A., Hunter D.J., Ost T.W., White J.H., Chapman S.K., Turner N.J., Flitsch S.L. (2003) A self-sufficient cytochrome p450 with a primary structural organization that includes a flavin domain and a [2Fe-2S] redox center. *J Biol Chem*. **278** (49): 48914-48920.

[100] Miles J.S., Munro A.W., Rospendowski B.N., Smith W.E., McKnight J., Thomson A.J. (1992) Domains of the catalytically self-sufficient cytochrome P-450 BM-3. Genetic construction overexpression purification and spectroscopic characterization. *Biochem J.* **288** (2): 503-509.

[101] Gustafsson M.C., Roitel O., Marshall K.R., Noble M.A., Chapman S.K., Pessegueiro A., Fulco A.J., Cheesman M.R., von Wachenfeldt C., Munro A.W. (2004) Expression purification and characterization of Bacillus subtilis cytochromes P450 CYP102A2 and CYP102A3: flavocytochrome homologues of P450 BM3 from Bacillus megaterium. *Biochemistry*. **43** (18): 5474-5487.

[102] Nakayama N., Takemae A., Shoun H. (1996) Cytochrome P450foxy a catalytically self-sufficient fatty acid hydroxylase of the fungus Fusarium oxysporum. *J Biochem*. **119** (3): 435-440.

[103] Nakahara K., Tanimoto T., Hatano K., Usuda K., Shoun H. (1993) Cytochrome P 450 55A1 (P-450dNIR) acts as nitric oxide reductase employing NADH as the direct electron donor. *J Biol Chem*. **268** (11): 8350-8355.

[104] Zhang L., Takayam N., Kitazumem T., Kondom T., Shoun H. (2001) Purification and cDNA cloning of nitric oxide reductase cytochrome P450nor (CYP55A4) from Trichosporon cutaneum. *Eur J Biochem*. **268** (11): 3198-3204.

[105] Omura T. (2010) Structural diversity of cytochrome P450 enzyme system. *J Biochem*. **147** (3): 297-306.

[106] DeWitt D.L., Smith W.L. (1983) Purification of prostacyclin synthase from bovine aorta by immunoaffinity chromatography. Evidence that the enzyme is a hemoprotein. *J Biol Chem*. **258** (5): 3285-3293.

[107] Haurand M., Ullrich V. (1985) Isolation and characterization of thromboxane synthase from human platelets as a cytochrome P-450 enzyme. *J Biol Chem*. **260** (28): 15059-15067.

[108] Song W.C., Brash A.R. (1991) Purification of an allene oxide synthase and identification of the enzyme as a cytochrome P-450. *Science*. **253** (5021): 781-784.

[109] Bernhardt R. (2006) Cytochromes P450 as versatile biocatalysts. *J Biotechnol*. **124** (1): 128-145.

[110] Lewis D.F. (2002) Oxidative stress: the role of cytochromes P450 in oxygen activation. *J Chem Technol Biotechnol*. **77** (10): 1095–1100.

[111] da Silva J.J.R.F, Williams R.J.P. (1991) *The Biological Chemistry of the Elements*. Clarendon Press, Oxford

[112] Meunier B., de Visser S.P., Shaik S. (2004) Mechanism of oxidation reactions catalysed by cytochrome p450 enzymes. *Chem Rev*. **104** (9): 3947-3980.

[113] Petrucci R., Harwood W., Herring G., Madura J. (2007) *General Chemistry: Principles and Modern Applications 9th Edition*. Pearson Education, New Jersey.

[114] McKersie B.D., Lesheim Y. (1994) *Stress and Stress Coping in Cultivated Plants*. Kluwwer Academic Publishers, Dordrecht-Boston-London.

[115] Lewis D.F.V., Ito Y., Goldfarb P.S. (2005) Cytochrome P450 structures and their substrate interactions. *Drug Dev Res*. **66**: 19–24.

[116] Lewis D.F.V. (2001) *Cytochromes P450: Structure Function and Mechanism*. Taylor and Francis, London.

[117] Ortiz de Montellano P.R. (2010) Hydrocarbon hydroxylation by cytochrome P450 enzymes. *Chem Rev*. **110** (2): 932-948.

[118] Joo H., Lin Z., Arnold F.H. (1999) Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation. *Nature*. **399** (6737): 670-673.

[119] Danielson P.B. (2002) The Cytochrome P450 Superfamily: Biochemistry Evolution and Drug. *Curr Drug Metab*. **3** (6): 561-597.

[120] Chothia C. (1994) *Protein families in the metazoan genome*. Dev Suppl: 27-33.

[121] Loomis W.F. (1988) *Four Billion Years An Essay on the Evolution of Genes and Organisms.* Sinauer Associates Inc., Sunderland Massachusetts.

[122] Nebert D.W., Feyereisen R. (1994) *In Cytochrome P450 Biochemistry Biophysics and Molecular Biology (Lechner M.C. Ed.).* Libbey Eurotext, London.

[123] Schlezinger J.J., White R.D., Stegeman J.J. (1999) Oxidative inactivation of cytochrome P-450 1A (CYP1A) stimulated by 33'44'-tetrachlorobiphenyl: production of reactive oxygen by vertebrate CYP1As. *Mol Pharmacol*. **56** (3): 588 597.

[124] Nebert D.W., Gonzalez F.J. (1985) Cytochrome P450 gene expression and regulation. *Trends Pharmacol Sci*. **6**: 160–164.

[125] Nelson D.R., Strobel H.W. (1987) Evolution of cytochrome P-450 proteins. *Mol Biol Evol.* **4** (6): 572-593.

[126] Guengerich F.P., Munro A.W. (2013) Unusual cytochrome P450 enzymes and reactions. J Biol Chem. 288 (24): 17065-17073.

[127] Whittall J., Sutton P. (2010) *Practical methods for biocatalysis and biotransformations.* J. Wiley, Chichester.

[128] Groger H., Yasuhisa Asano Y. (2012) *Introduction - Principles and Historical Landmarks of Enzyme in Enzyme Catalysis in Organic Synthesis Third Edition*. Wiley-VCH, Weinheim, pp 1-42.

[129] Labinger J.A., Bercaw J.E. (2002) Understanding and exploiting C-H bond activation. *Nature*. **417** (6888): 507-514.

[130] Bernhardt R., Urlacher V.B. (2014) Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl Microbiol Biotechnol*. **98** (14): 6185-6203.

[131] Urlacher V.B., Eiben S. (2006) Cytochrome P450 monooxygenases: perspectives for synthetic application. *Trends Biotechnol*. **24** (7): 324-330.

[132] van Beilen J.B., Duetz W.A., Schmid A., Witholt B. (2003) Practical issues in the application of oxygenases. *Trends Biotechnol*. **21** (4): 170-177.

[133] O'Reilly E., Köhler V., Flitsch S.L., Turner N.J. (2011) Cytochromes P450 as useful biocatalysts: addressing the limitations. *Chem Commun (Camb)*. **47** (9): 2490-2501.

[134] Domanski T.L., Halpert J.R. (2001) Analysis of mammalian cytochrome P450 structure and function by site-directed mutagenesis. *Curr Drug Metab*. **2** (2): 117 137.

[135] Yasutake Y., Nishioka T., Imoto N., Tamura T. (2013) A single mutation at the ferredoxin binding site of P450 Vdh enables efficient biocatalytic production of 25-hydroxyvitamin D(3). *Chembiochem*. **14** (17): 2284-2291.

[136] Lee H., Kim J.H., Han S., Lim Y.R., Park H.G., Chun Y.J., Park S.W., Kim D. (2014) Directed-evolution analysis of human cytochrome P450 2A6 for enhanced enzymatic catalysis. *J Toxicol Environ Health A*. **77** (22-24): 1409-1418.

[137] Dougherty M.J., Arnold F.H. (2009) Directed evolution: new parts and optimized function. *Curr Opin Biotechnol*. **20** (4): 486-491.

[138] Straathof A.J., Panke S., Schmid A. (2002) The production of fine chemicals by biotransformations. *Curr Opin Biotechnol*. **13** (6): 548-556.

[139] Julsing M.K., Cornelissen S., Bühler B., Schmid A. (2008) Heme-iron oxygenases: powerful industrial biocatalysts?. *Curr Opin Chem Biol*. **12** (2): 177-186.

[140] Paddon C.J., Westfall P.J., Pitera D.J., Benjamin K., Fisher K., McPhee D., Leavell M.D., Tai A., Main A., Eng D., Polichuk D.R., Teoh K.H., Reed. D.W., Treynor T., Lenihan J., Fleck M., Bajad S., Dang G., Dengrove D., Diola D., *et al.* (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature*. **496** (7446): 528-532.

[141] Shibasaki T., Mori H., Ozaki A. (2000) Enzymatic production of trans-4-hydroxy-L-proline by regio- and stereospecific hydroxylation of L-proline. *Biosci Biotechnol Biochem*. **64** (4): 746-750.

[142] Kiener A. (1995) Biosynthesis of functionalized aromatic N-heterocycles. *Chemtech*. **25**: 31-35.

[143] Bühler B., Bollhalder I., Hauer B., Witholt B., Schmid A. (2003) Use of the two-liquid phase concept to exploit kinetically controlled multistep biocatalysis. *Biotechnol Bioeng.* **81** (6): 683-694.

[144] Lilly M.D., Woodley J.M. (1996) A structured approach to design and operation of biotransformation processes. *J Ind Microbiol*. **17** (1): 24-29.

[145] Park J.B., Bühler B., Habicher T., Hauer B., Panke S., Witholt B., Schmid A. (2006) The efficiency of recombinant Escherichia coli as biocatalyst for stereospecific epoxidation. *Biotechnol Bioeng*. **95** (3): 501-512.

[146] Ro D.K., Paradise E.M., Ouellet M., Fisher K.J., Newman K.L., Ndungu J.M., Ho K.A., Eachus R.A., Ham T.S., Kirby J., Chang M.C., Withers S.T., Shiba Y., Sarpong R., Keasling J.D. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*. **440** (7086): 940-943.

[147] Mercke P., Bengtsson M., Bouwmeester H.J., Posthumus M.A., Brodelius P.E. (2000) Molecular cloning expression and characterization of amorpha-411-diene synthase a key enzyme of artemisinin biosynthesis in Artemisia annua. *Arch Biochem Biophys*. **381** (2): 173-180.

[148] Martin V.J., Pitera D.J., Withers S.T., Newman J.D., Keasling J.D. (2003) Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nat Biotechnol*. **21** (7): 796-802.

[149] Westfall P.J., Pitera D.J., Lenihan J.R., Eng D., Woolard F.X., Regentin R., Horning T., Tsuruta H., Melis D.J., Owens A., Fickes S., Diola D., Benjamin K.R., Keasling J.D., Leavell M.D., McPhee D.J., *et al.* (2012) Production of amorphadiene in yeast and its conversion to dihydroartemisinic acid precursor to the antimalarial agent artemisinin. *Proc Natl Acad Sci USA*. **109** (3): 111-118.

[150] Teoh K.H., Polichuk D.R., Reed D.W., Covello P.S. (2009) Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis in Artemisia annua. *Botany-Botanique*. **87** (6): 635–642.

[151] Duport C., Spagnoli R., Degryse E., Pompon D. (1998) Self-sufficient biosynthesis of pregnenolone and progesterone in engineered yeast. *Nat Biotechnol*. **16** (2): 186 189.

[152] Szczebara F.M., Chandelier C., Villeret C., Masurel A., Bourot S., Duport C., Blanchard S., Groisillier A., Testet E., Costaglioli P., Cauet G., Degryse E., Balbuena D., Winter J., Achstetter T., Spagnoli R., Pompon D., Dumas B. (2003) Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nat Biotechnol*. **21** (2): 143-149.

[153] Sonomoto K., Hoq M.M., Tanaka A., Fukui S. (1983) 11beta-Hydroxylation of Cortexolone (Reichstein Compound S) to Hydrocortisone by Curvularia lunata Entrapped in Photo-Cross-Linked Resin Gels. *Appl Environ Microbiol*. **45** (2): 436 443.

[154] Hogg J.A. (1992) Steroids the steroid community and Upjohn in perspective: a profile of innovation. *Steroids*. **57** (12): 593-616.

[155] Tanaka Y., Sasaki N., Ohmiya A. (2008) Biosynthesis of plant pigments: anthocyanins betalains and carotenoids. *Plant J*. **54** (4): 733-749.

[156] Holton T.A., Brugliera F., Lester D.R., Tanaka Y., Hyland C.D., Menting J.G., Lu C.Y., Farcy E., Stevenson T.W., Cornish E.C. (1993) Cloning and expression of cytochrome P450 genes controlling flower colour. *Nature*. **366** (6452): 276-279.

[157] Tanaka Y., Brugliera F. (2013) Flower colour and cytochromes P450. *Philos Trans R Soc Lond B Biol Sci.* **368** (1612): 20120432.

[158] Ogata J., Kanno Y., Itoh Y., Tsugawa H., Suzuki M. (2005) Plant biochemistry: anthocyanin biosynthesis in roses. *Nature*. **435** (7043): 757-758.

[159] Chen L., Waxman D.J. (2002) Cytochrome P450 gene-directed enzyme prodrug therapy (GDEPT) for cancer. *Curr Pharm Des*. **8** (15): 1405-1416.

[160] Zhang J., Tian Q., Yung Chan S., Chuen Li S., Zhou S., Duan W., Zhu Y.Z. (2005) Metabolism and transport of oxazaphosphorines and the clinical implications. *Drug Metab Rev*. **37** (4): 611-703.

[161] Roy P., Waxman D.J. (2006) Activation of oxazaphosphorines by cytochrome P450: application to gene-directed enzyme prodrug therapy for cancer. *Toxicol In Vitro*. **20** (2): 176-186.

[162] Jounaidi Y., Hecht J.E., Waxman D.J. (1998) Retroviral transfer of human cytochrome P450 genes for oxazaphosphorine-based cancer gene therapy. *Cancer Res*. **58** (19): 4391-4401.

[163] Hunt S. (2001) Technology evaluation: MetXia-P450 Oxford Biomedica. *Curr Opin Mol Ther.* **3** (6): 595-598.

[164] Chen C.S., Jounaidi Y., Su T., Waxman D.J. (2007) Enhancement of intratumoral cyclophosphamide pharmacokinetics and antitumor activity in a P450 2B11-based cancer gene therapy model. *Cancer Gene Ther*. **14** (12): 935-944.

[165] Bistolas N., Wollenberger U., Jung C., Scheller F.W. (2005) Cytochrome P450 biosensors-a review. *Biosens Bioelectron*. **20** (12): 2408-2423.

[166] Schneider E., Clark D.S. (2013) Cytochrome P450 (CYP) enzymes and the development of CYP biosensors. *Biosens Bioelectron*. **39** (1): 1-13.

[167] Sakaki T., Yamamoto K., Ikushiro S. (2013) Possibility of application of cytochrome P450 to bioremediation of dioxins. *Biotechnol Appl Biochem*. **60** (1): 65-70.

[168] Kumar S., Jin M., Weemhoff J.L. (2012) Cytochrome P450-Mediated Phytoremediation using Transgenic Plants: A Need for Engineered Cytochrome P450 Enzymes. *J Pet Environ Biotechnol*. **3** (5): 1000127.

[169] Batard Y., LeRet M., Schalk M., Robineau T., Durst F., Werck-Reichhart D. (1998) Molecular cloning and functional expression in yeast of CYP76B1 a xenobiotic-inducible 7-ethoxycoumarin O-deethylase from Helianthus tuberosus. *Plant J.* **14** (1): 111-120.

[170] Siminszky B., Corbin F.T., Ward E.R., Fleischmann T.J., Dewey R.E. (1999) Expression of a soybean cytochrome P450 monooxygenase cDNA in yeast and tobacco enhances the metabolism of phenylurea herbicides. *Proc Natl Acad Sci USA*. **96** (4): 1750-1755.

[171] Sakaki T., Shinkyo R., Takita T., Ohta M., Inouye K. (2002) Biodegradation of polychlorinated dibenzo-p-dioxins by recombinant yeast expressing rat CYP1A subfamily. *Arch Biochem Biophys*. **401** (1): 91-98.

[172] Inui H., Ohkawa H. (2005) Herbicide resistance in transgenic plants with mammalian P450 monooxygenase genes. *Pest Manag Sci.* **61** (3): 286-291.

[173] Whittaker R.H., Margulis L. (1978) Protist classification and the kingdoms of organisms. *Biosystems*. **10** (1-2): 3-18.

[174] Galagan J.E., Henn M.R., Ma L.J., Cuomo C.A., Birren B. (2005) Genomics of the fungal kingdom: insights into eukaryotic biology. *Genome Res*. **15** (12): 1620-1631.

[175] Taylor J.W., Berbee M.L. (2006) Dating divergences in the Fungal Tree of Life: review and new analyses. *Mycologia*. **98** (6): 838-849.

[176] Lücking R., Huhndorf S., Pfister D.H., Plata E.R., Lumbsch H.T. (2009) Fungi evolved right on track. *Mycologia*. **101** (6): 810-822.

[177] Chen W., Lee M.K., Jefcoate C., Kim S.C., Chen F., Yu J.H. (2014) Fungal cytochrome P450 monooxygenases: their distribution structure functions family expansion and evolutionary origin. *Genome Biol Evol*. **6** (7): 1620-1634.

[178] Hawksworth D.L. (2001) The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycol Res*. **105** (12): 1422-1432.

[179] Blackwell M. (2011) The fungi: 1 2 3 ... 5.1 million species?. *Am J Bot*. **98** (3): 426 438.

[180] Gehrig H., Schüssler A., Kluge M. (1996) Geosiphon pyriforme a fungus forming endocytobiosis with Nostoc (cyanobacteria) is an ancestral member of the Glomales: evidence by SSU rRNA analysis. *J Mol Evol.* **43** (1): 71-81.

[181] Heckman D.S., Geiser D.M., Eidell B.R., Stauffer R.L., Kardos N.L., Hedges S.B. (2001) Molecular evidence for the early colonization of land by fungi and plants. *Science*. **293** (5532): 1129-1133.

[182] Trappe J.M. (1987) Phylogenetic and ecologic aspects of mycotrophy in the angiosperms from an evolutionary standpoint. in *Ecophysiology of VA mycrorrhizal plants (ed. G.R. Safir)*. CRC Press, Boca Raton, pp: 5–25.

[183] Zeigler R.S., Leong S.A., Teeng P.S. (1994) Rice Blast Disease. CAB International, Wallingford.

[184] Dasbach E.J., Davies G.M., Teutsch S.M. (2000) Burden of aspergillosis-related hospitalizations in the United States. *Clin Infect Dis*. **31** (6): 1524-1528.

[185] Gudlaugsson O., Gillespie S., Lee K., Vande Berg J., Hu J., Messer S., Herwaldt L., Pfaller M., Diekema D. (2003) Attributable mortality of nosocomial candidemia revisited. *Clin Infect Dis*. **37** (9): 1172-1177.

[186] Hoang L.M., Maguire J.A., Doyle P., Fyfe M., Roscoe D.L. (2004) Cryptococcus neoformans infections at Vancouver Hospital and Health Sciences Centre (1997 2002): epidemiology microbiology and histopathology. *J Med Microbiol.* **53** (9): 935-940.

[187] Chiller T.M., Galgiani J.N., Stevens D.A. (2003) Coccidioidomycosis. *Infect Dis Clin North Am.* **17** (1): 41-57.

[188] Georgopapadakou N.H. (1998) Antifungals: mechanism of action and resistance established and novel drugs. *Curr Opin Microbiol*. **1** (5): 547-557.

[189] Hoffmeister D., Keller N.P. (2007) Natural products of filamentous fungi: enzymes genes and their regulation. *Nat Prod Rep*. **24** (2): 393-416.

[190] Grigoriev I.V., Nikitin R., Haridas S., Kuo A., Ohm R., Otillar R., Riley R., Salamov A., Zhao X., Korzeniewski F., Smirnova T., Nordberg H., Dubchak I., Shabalov I. (2014) MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res*. **42**: 699-704.

[191] Syed K., Yadav J.S. (2012) P450 monooxygenases (P450ome) of the model white rot fungus Phanerochaete chrysosporium. *Crit Rev Microbiol*. **38** (4): 339-363.

[192] Ehrlich K.C., Chang P.K., Yu J., Cotty P.J. (2004) Aflatoxin biosynthesis cluster gene cypA is required for G aflatoxin formation. *Appl Environ Microbiol*. **70** (11): 6518-6524.

[193] Kimura M., Tokai T., Takahashi-Ando N., Ohsato S, Fujimura M. (2007) Molecular and genetic studies of fusarium trichothecene biosynthesis: pathways genes and evolution. *Biosci Biotechnol Biochem*. **71** (9): 2105-2123.

[194] Proctor R.H., Plattner R.D., Desjardins A.E., Busman M., Butchko R.A. (2006) Fumonisin production in the maize pathogen Fusarium verticillioides: genetic basis of naturally occurring chemical variation. *J Agric Food Chem*. **54** (6): 2424-2430.

[195] Schmale D.G., Munkvold G.P. (2009) Mycotoxins in crops: A threat to human and domestic animal health. *The Plant Health Instructor*. DOI: 10.1094/PHI-I-2009 0715-01.

[196] Leal G.A., Gomes L.H., Albuquerque P.S., Tavares F.C., Figueira A. (2010) Searching for Moniliophthora perniciosa pathogenicity genes. *Fungal Biol*. **114** (10): 842-854.

[197] Becher R., Wirsel S.G. (2012) Fungal cytochrome P450 sterol 14α-demethylase (CYP51) and azole resistance in plant and human pathogens. *Appl Microbiol Biotechnol*. **95** (4): 825-840.

[198] Goffeau A., Barrell B.G., Bussey H., Davis R.W., Dujon B., Feldmann H., Galibert F., Hoheisel J.D., Jacq C., Johnston M., Louis E.J., Mewes H.W., Murakami Y., Philippsen P., Tettelin H., Oliver S.G. (1996) Life with 6000 genes. *Science*. **274** (546): 563-567.

[199] Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D., Jaffe D., FitzHugh W., Ma L.J., Smirnov S., Purcell S., Rehman B., Elkins T., Engels R., Wang S., Nielsen C.B. Butler J., Endrizzi M., Qui D., *et al.* (2003) The genome sequence of the filamentous fungus Neurospora crassa. *Nature*. **422** (6934): 859 868.

[200] Sezutsu H., Le Goff G., Feyereisen R. (2013) Origins of P450 diversity. *Philos Trans R Soc Lond B Biol Sci*. **368** (1612): 20120428.

[201] Kelly D.E., Krasevec N., Mullins J., Nelson D.R. (2009) The CYPome (Cytochrome P450 complement) of Aspergillus nidulans. *Fungal Genet Biol*. **46**: S53-61.

[202] Shahid A.A., Rao A.Q., Bakhsh A., Husnain T. (2012) Entomopathogenic fungi as biological controllers: new insights into their virulence and pathogenicity. *Arch Biol Sci*. **64** (1): 21-42.

[203] Pedrini N., Zhang S., Juárez M.P., Keyhani N.O. (2010) Molecular characterization and expression analysis of a suite of cytochrome P450 enzymes implicated in insect hydrocarbon degradation in the entomopathogenic fungus Beauveria bassiana. *Microbiology*. **156** (8): 2549-2557.

[204] Blomquist G.J., Nelson D.R., De Renobales M. (1987) Chemistry biochemistry and physiology of insect cuticular lipids. *Arch Insect Biochem Physiol*. **6** (4): 227–265.

[205] Figueiras A.N.L., Girotti J.R., Mijailovsky S.J., Juarez M.P. (2009) Epicuticular lipids induce aggregation in Chagas disease vectors. *Parasit Vectors*. **2**: 8.

[206] Cho E.M., Liu L., Farmerie W., Keyhani N.O. (2006) EST analysis of cDNA libraries from the entomopathogenic fungus Beauveria (Cordyceps) bassiana. I. Evidence for stage-specific gene expression in aerial conidia in vitro blastospores and submerged conidia. *Microbiology*. **152** (9): 2843-2854.

[207] de Faria M.R., Wraight S.P. (2007) Mycoinsecticides and Mycoacaricides: A comprehensive list with worldwide coverage and international classification of formulation types. *Biological Control*. **43** (3): 237–256.

[208] Xiao G., Ying S.H., Zheng P., Wang Z.L., Zhang S., Xie X.Q., Shang Y., St Leger R.J., Zhao G.P., Wang C., Feng M.G. (2012) Genomic perspectives on the evolution of fungal entomopathogenicity in Beauveria bassiana. *Sci Rep*. **2**: 483.

[209] Malarvannan S., Murali P.D., Shanthakumar S.P., Prabavathy V.R., Nair S. (2010) Laboratory evaluation of the entomopathogenic fungi Beauveria bassiana against the Tobacco caterpillar Spodoptera litura Fabricius (Noctuidae: Lepidoptera). *Journal of Biopesticides.* **3** (1): 126-131.

[210] Howard A.F., N'guessan R., Koenraadt C.J., Asidi A., Farenhorst M., Akogbéto M., Thomas M.B., Knols B.G., Takken W. (2010) The entomopathogenic fungus Beauveria bassiana reduces instantaneous blood feeding in wild multi-insecticide-resistant Culex quinquefasciatus mosquitoes in Benin West Africa. *Parasit Vectors*. **3**: 87.

[211] Mnyone L.L., Ng'habi K.R., Mazigo H.D., Katakweba A.A., Lyimo I.N. (2012) Entomopathogenic fungi Metarhizium anisopliae and Beauveria bassiana reduce the survival of Xenopsylla brasiliensis larvae (Siphonaptera: Pulicidae). *Parasit Vectors*. **5**: 204.

[212] Bing L.A., Lewis L.C. (1993) Occurrence of the entomopathogen Beauveria bassiana (Balsamo) Vuillemin in different tillage regimes and in Zea mays L. and virulence towards Ostrinia nubilalis (Hübner). *Agr Ecosyst Environ*. **45** (1 2): 147 156.

[213] Posada F., Vega F.E. (2005) Establishment of the fungal entomopathogen Beauveria bassiana (Ascomycota: Hypocreales) as an endophyte in cocoa seedlings (Theobroma cacao). *Mycologia*. **97** (6): 1195-1200.

[214] Quesada-Moraga E., Landa B.B., Muñoz-Ledesma J., Jiménez-Diáz R.M., Santiago Alvarez C. (2006) Endophytic colonisation of opium poppy Papaver somniferum by an entomopathogenic Beauveria bassiana strain. *Mycopathologia*. **161** (5): 323-329.

[215] Fan Y., Fang W., Guo S., Pei X., Zhang Y., Xiao Y., Li D., Jin K., Bidochka M.J., Pei Y. (2007) Increased insect virulence in Beauveria bassiana strains overexpressing an engineered chitinase. *Appl Environ Microbiol*. **73** (1): 295-302.

[216] Joshi L., St Leger R.J., Bidochka M.J. (1995) Cloning of a cuticle-degrading protease from the entomopathogenic fungus Beauveria bassiana. *FEMS Microbiol Lett.* **125** (2-3): 211-217.

[217] Zhang S., Widemann E., Bernard G., Lesot A., Pinot F., Pedrini N., Keyhani N.O. (2012) CYP52X1 representing new cytochrome P450 subfamily displays fatty acid hydroxylase activity and contributes to virulence and growth on insect cuticular substrates in entomopathogenic fungus Beauveria bassiana. *J Biol Chem*. **287** (16): 13477-13486.

[218] Xu Y., Orozco R., Kithsiri Wijeratne E.M., Espinosa-Artiles P., Leslie Gunatilaka A.A., Patricia Stock S., Molnár I. (2009) Biosynthesis of the cyclooligomer depsipeptide bassianolide an insecticidal virulence factor of Beauveria bassiana. *Fungal Genet Biol.* **46** (5): 353-364.

[219] Grogan G.J., Holland H.L. (2000) The biocatalytic reactions of Beauveria spp. *J Mol Catal B: Enzym.* **9** (1-3): 1–32.

[220] van der Geize R., Dijkhuizen L. (2004) Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. *Curr Opin Microbiol*. **7** (3): 255-261.

[221] Bell K.S., Philp J.C., Aw D.W., Christofi N. (1998) The genus Rhodococcus. *J Appl Microbiol.* **85** (2): 195-210.

[222] Larkin M.J., Kulakov L.A., Allen C.C. (2005) Biodegradation and Rhodococcus-masters of catabolic versatility. *Curr Opin Biotechnol*. **16** (3): 282-290.

[223] McLeod M.P., Warren R.L., Hsiao W.W., Araki N., Myhre M., Fernandes C., Miyazawa D., Wong W., Lillquist A.L., Wang D., Dosanjh M., Hara H., Petrescu A., Morin R.D., Yang G., Stott J.M., Schein J.E., Shin H., Smailus D., *et al*. (2006) The complete genome of Rhodococcus sp. RHA1 provides insights into a catabolic powerhouse. *Proc Natl Acad Sci USA*. **103** (42): 15582-15587.

[224] Rosłoniec K.Z., Wilbrink M.H., Capyk J.K., Mohn W.W., Ostendorf M., van der Geize R., Dijkhuizen L., Eltis L.D. (2009) Cytochrome P450 125 (CYP125) catalyses C26-hydroxylation to initiate sterol side-chain degradation in Rhodococcus jostii RHA1. *Mol Microbiol*. **74** (5): 1031-1043.

[225] Rosłoniec K.Z. (2011) Steroid transformation by Rhodococcus strains and bacterial cytochrome P450 enzymes. Thesis: University of Groningen.

[226] Bonsor D., Butz S.F., Solomons J., Grant S., Fairlamb I.J., Fogg M.J., Grogan G. (2006) Ligation independent cloning (LIC) as a rapid route to families of recombinant biocatalysts from sequenced prokaryotic genomes. *Org Biomol Chem*. **4** (7): 1252-1260.

[227] Sabbadin F., Hyde R., Robin A., Hilgarth E.M., Delenne M., Flitsch S., Turner N., Grogan G., Bruce N.C. (2010) LICRED: a versatile drop-in vector for rapid generation of redox-self-sufficient cytochrome P450s. *Chembiochem*. **11** (7): 987 994.

[228] Pompon D., Louerat B., Bronine A., Urban P. (1996) Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol*. **272**: 51-64.

[229] Sambrook J., Russel D.W. (2001) *Molecular Cloning: A laboratory Manual*. Cold Spring Harbor Laboratory Press, York.

[230] Schiestl R.H., Gietz R.D. (1989) High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet*. **16** (5-6): 339-346.

[231] Hanahan D., Bloom F.R. (1996) Mechanisms of DNA transformation in *Escherichia coli and Salmonella: Cellular and Molecular Biology*. ASM Press, Washington D.C., pp: 2449-2459.

[232] Maniatis T., Fritsch E.F., Sambrook J. (1982) *Molecular cloning. A laboratory manual.* Cold Spring Harbor Laboratory, New York.

[233] Laemmli U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. **227** (5259): 680-685.

[234] Burnette W.N. (1981) Western blotting: electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal Biochem*. **112** (2): 195-203.

[235] Syed K., Mashele S.S. (2014) Comparative analysis of P450 signature motifs EXXR and CXG in the large and diverse kingdom of fungi: identification of evolutionarily conserved amino acid patterns characteristic of P450 family. *PLoS One.* **9** (4): e95616.

[236] Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J., Higgins D.G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*. **23** (21): 2947-2948.

[237] Felsenstein J. (1993) PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics. University of Washington, Seattle.

[238] Dereeper A., Guignon V., Blanc G., Audic S., Buffet S., Chevenet F., Dufayard J.F., Guindon S., Lefort V., Lescot M., Claverie J.M., Gascuel O. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res*. **36** (Web Server issue): W465-469.

[239] Imai M., Shimada H., Watanabe Y., Matsushima-Hibiya Y., Makino R., Koga H., Horiuchi T., Ishimura Y. (1989) Uncoupling of the cytochrome P-450cam monooxygenase reaction by a single mutation threonine-252 to alanine or valine: possible role of the hydroxy amino acid in oxygen activation. *Proc Natl Acad Sci USA*. **86** (20): 7823-7827.

[240] Andersen J.F., Tatsuta K., Gunji H., Ishiyama T., Hutchinson C.R. (1993) Substrate specificity of 6-deoxyerythronolide B hydroxylase a bacterial cytochrome P450 of erythromycin A biosynthesis. *Biochemistry*. **32** (8): 1905-1913.

[241] Slessor K.E., Farlow A.J., Cavaignac S.M., Stok J.E., De Voss J.J. (2011) Oxygen activation by P450(cin): Protein and substrate mutagenesis. *Arch Biochem Biophys*. **507** (1): 154-162.

[242] Rupasinghe S., Schuler M.A., Kagawa N.,Yuan H., Lei L., Zhao B., Kelly S.L., Waterman M.R., Lamb D.C. (2006) The cytochrome P450 gene family CYP157 does not contain EXXR in the K-helix reducing the absolute conserved P450 residues to a single cysteine. *FEBS Lett*. **580** (27): 6338-6342.

[243] Lin L., Fang W., Liao X., Wang F., Wei D., St Leger R.J. (2011) The MrCYP52 cytochrome P450 monoxygenase gene of Metarhizium robertsii is important for utilizing insect epicuticular hydrocarbons. *PLoS One*. **6** (12): e28984.

[244] Sakaki T., Shibata M. Yabusaki Y., Murakami H., Ohkawa H. (1989) Expression of bovine cytochrome P450c17 cDNA in Saccharomyces cerevisiae. *DNA*. **8** (6): 409 418.

[245] Shibata M., Sakaki T., Yabusaki Y., Murakami H., Ohkawa H. (1990) Genetically engineered P450 monooxygenases: construction of bovine P450c17/yeast reductase fused enzymes. *DNA Cell Biol*. **9** (1): 27-36.

[246] Hofmann K., Stoffel W. (1993) TMbase-a database of membrane spanning proteins segments. *J Biol Chem*. **347**: 166.

[247] Hirokawa T., Boon-Chieng S., Mitaku S. (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics*. **14**: 378-379.

[248] Mitaku S., Hirokawa T. (1999) Physicochemical factors for discriminating between soluble and membrane proteins: hydrophobicity of helical segments and protein length. *Protein Eng*. **12** (11):953-957.

[249] Mitaku S., Hirokawa T., Tsuji T. (2002) Amphiphilicity index of polar amino acids as an aid in the characterization of amino acid preference at membrane-water interfaces. *Bioinformatics*. **18** (4): 608-616.

[250] Le Grice S.F., Matzura H., Marcoli R., Iida S., Bickle T.A. (1982) The catabolite-sensitive promoter for the chloramphenicol acetyl transferase gene is preceded by two binding sites for the catabolite gene activator protein. *J Bacteriol.* **150** (1): 312 318.

[251] Mayhew M., da Silva A.C., Martin J., Erdjument-Bromage H., Tempst P., Hartl F.U. (1996) Protein folding in the central cavity of the GroEL-GroES chaperonin complex. *Nature*. **379** (6564): 420-426.

[252] Weis R., Winkler M., Schittmayer M., Kambourakis S., Vink M., Rozzell J.D., Gliedera A. (2009) A Diversified Library of Bacterial and Fungal Bifunctional Cytochrome P450 Enzymes for Drug Metabolite Synthesis. *Adv Synth Catal*. **351** (13): 2140-2146.

[253] van Beilen J.B., Holtackers R., Lüscher D., Bauer U., Witholt B., Duetz W.A. (2005) Biocatalytic production of perillyl alcohol from limonene by using a novel Mycobacterium sp. cytochrome P450 alkane hydroxylase expressed in Pseudomonas putida. *Appl Environ Microbiol*. **71** (4): 1737-1744.

[254] Omura T., Sato R. (1964) The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. *J Biol Chem*. **239**: 2370-2378.

[255] Fujita K., Kamataki T. (2002) Genetically engineered bacterial cells co-expressing human cytochrome P450 with NADPH-cytochrome P450 reductase: prediction of metabolism and toxicity of drugs in humans. *Drug Metab Pharmacokinet.* **17** (1): 1 22.

[256] Wiedmann M., Kurzchalia T.V., Hartmann E., Rapoport T.A. (1987) A signal sequence receptor in the endoplasmic reticulum membrane. *Nature*. **328** (6133): 830 833.

[257] Zelasko S., Palaria A., Das A. (2013) Optimizations to achieve high-level expression of cytochrome P450 proteins using Escherichia coli expression systems. *Protein Expr Purif*. **92** (1): 77-87.

[258] Barnes H.J., Arlotto M.P., Waterman M.R. (1991) Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in Escherichia coli. *Proc Natl Acad Sci USA*. **88** (13): 5597-5601.

[259] Sandhu P., Guo Z., Baba T., Martin M.V., Tukey R.H., Guengerich F.P. (1994) Expression of modified human cytochrome P450 1A2 in Escherichia coli: stabilization purification spectral characterization and catalytic activities of the enzyme. *Arch Biochem Biophys*. **309** (1): 168-177.

[260] Tang Z., Salamanca-Pinzón S.G., Wu Z.L., Xiao Y., Guengerich F.P. (2010) Human cytochrome P450 4F11: heterologous expression in bacteria purification and characterization of catalytic function. *Arch Biochem Biophys*. **494** (1): 86-93.

[261] Li Y.C., Chiang J.Y. (1991) The expression of a catalytically active cholesterol 7 alpha-hydroxylase cytochrome P450 in Escherichia coli. *J Biol Chem*. **266** (29): 19186-19191.

[262] Mast N., Andersson U., Nakayama K., Bjorkhem I., Pikuleva I.A. (2004) Expression of human cytochrome P450 46A1 in Escherichia coli: effects of N- and C-terminal modifications. *Arch Biochem Biophys*. **428** (1): 99-108.

[263] Nodate M., Kubota M., Misawa N. (2006) Functional expression system for cytochrome P450 genes using the reductase domain of self-sufficient P450Rhf from Rhodococcus sp. NCIMB 9784. *Appl Microbiol Biotechnol*. **71** (4): 455-462.

[264] Li S., Podust L.M., Sherman D.H. (2007) Engineering and analysis of a self-sufficient biosynthetic cytochrome P450 PikC fused to the RhfRED reductase domain. *J Am Chem Soc*. **129** (43): 12940-12941.

[265] Robin A., Köhler V., Jones A., Ali A., Kelly P.P., O'Reilly E., Turner N.J., Flitsch S.L. (2011) Chimeric self-sufficient P450cam-RhfRed biocatalysts with broad substrate scope. *Beilstein J Org Chem*. **7**: 1494-1498.

[266] Schückel J., Rylott E.L., Grogan G., Bruce N.C. (2012) A gene-fusion approach to enabling plant cytochromes p450 for biocatalysis. *Chembiochem*. **13** (18): 2758 2763.

[267] Studier F.W., Moffatt B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol*. **189** (1): 113-130.

[268] Stormo G.D., Schneider T.D., Gold L.M. (1982) Characterization of translational initiation sites in E. coli. *Nucleic Acids Res*. **10** (9): 2971-2996.

[269] Nishihara K., Kanemori M., Yanagi H., Yura T. (2000) Overexpression of trigger factor prevents aggregation of recombinant proteins in Escherichia coli. *Appl Environ Microbiol*. **66** (3): 884-8849.

[270] Yamasaki T., Izumi S., Ide H., Ohyama Y. (2004) Identification of a novel rat microsomal vitamin D3 25-hydroxylase. *J Biol Chem*. **279** (22): 22848-22856.

[271] Urban P., Werck-Reichhart D., Teutsch H.G., Durst F., Regnier S., Kazmaier M., Pompon D. (1994) Characterization of recombinant plant cinnamate 4-hydroxylase produced in yeast. Kinetic and spectral properties of the major plant P450 of the phenylpropanoid pathway. *Eur J Biochem*. **222** (3): 843-850.

[272] Duan H., Schuler M.A (2006) Heterologous expression and strategies for encapsulation of membrane-localized plant P450s. *Phytochem Rev*. **5** (2): 507-523.

[273] Huang F.C., Peter A., Schwab W. (2014) Expression and characterization of CYP52 genes involved in the biosynthesis of sophorolipid and alkane metabolism from Starmerella bombicola. *Appl Environ Microbiol*. **80** (2): 766-776.

[274] Pedrini N., Ortiz-Urquiza A., Huarte-Bonnet C., Zhang S., Keyhani N.O. (2013) Targeting of insect epicuticular lipids by the entomopathogenic fungus Beauveria bassiana: hydrocarbon oxidation within the context of a host-pathogen interaction. *Front Microbiol*. **4**: 24.

[275] Sharp P.M., Tuohy T.M., Mosurski K.R. (1986) Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Res*. **14** (13): 5125-5143.

[276] Zhang S.P., Zubay G., Goldman E. (1991) Low-usage codons in Escherichia coli, yeast, fruit fly, and primates. Gene. 105 (1): 61-72.

[277] Hehn A., Morant M., Werck-Reichhart D. (2002) Partial recoding of P450 and P450 reductase cDNAs for improved expression in yeast and plants. *Methods Enzymol.* **357**: 343-351.

[278] Batard Y., Hehn A., Nedelkina S., Schalk M., Pallett K., Schaller H., Werck-Reichhart D. (2000) Increasing expression of P450 and P450-reductase proteins from monocots in heterologous systems. *Arch Biochem Biophys*. **379** (1): 161-169.

[279] Kulig J.K., Spandolf C., Hyde R., Ruzzini A.C., Eltis L.D., Grönberg G., Hayes M.A., Grogan G. (2015) A P450 fusion library of heme domains from Rhodococcus jostii RHA1 and its evaluation for the biotransformation of drug molecules. *Bioorg Med Chem*. **23** (17): 5603-5609.

[280] Hargrove T.Y., Wawrzak Z., Lamb D.C., Guengerich F.P., Lepesheva G.I. (2015) Structure-Functional Characterization of Cytochrome P450 Sterol 14α-Demethylase (CYP51B) from Aspergillus fumigatus and Molecular Basis for the Development of Antifungal Drugs. *J Biol Chem* **290** (39): 23916-23934.

[281] Neeli R., Girvan H.M., Lawrence A., Warren M.J., Leys D., Scrutton N.S., Munro A.W. (2005) The dimeric form of flavocytochrome P450 BM3 is catalytically functional as a fatty acid hydroxylase. *FEBS Lett*. **579** (25): 5582-5588.

[282] Hlavica P. (2009) Assembly of non-natural electron transfer conduits in the cytochrome P450 system: a critical assessment and update of artificial redox constructs amenable to exploitation in biotechnological areas. *Biotechnol Adv*. **27** (2): 103-121.

[283] Zhang W., Liu Y., Yan J., Cao S., Bai F., Yang Y., Huang S., Yao L., Anzai Y., Kato F., Podust L.M., Sherman D.H., Li S. (2014) New reactions and products resulting from alternative interactions between the P450 enzyme and redox partners. *J Am Chem Soc*. **136** (9): 3640-3646.

[284] Adrio J.L., Demain A.L. (2014) Microbial enzymes: tools for biotechnological processes. Biomolecules. 4 (1): 117-139.

[285] McGovern P.E., Zhang J., Tang J., Zhang Z., Hall G.R., Moreau R.A., Nuñez A., Butrym E.D., Richards M.P., Wang C.S., Cheng G., Zhao Z., Wang C. (2004) Fermented beverages of pre- and proto-historic China. *Proc Natl Acad Sci USA*. **101** (51): 17593-17598.

[286] Aehle W. (2007) *Enzymes in Industry: Production and Applications. Third Edition*. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

[287] Behrendorff J.B., Huang W., Gillam E.M. (2015) Directed evolution of cytochrome P450 enzymes for biocatalysis: exploiting the catalytic versatility of enzymes with relaxed substrate specificity. *Biochem J*. **467** (1): 1-15.

[288] Wienkers L.C., Heath T.G. (2005) Predicting in vivo drug interactions from in vitro drug discovery data. *Nat Rev Drug Discov*. **4** (10): 825-833.

[289] Podust L.M., Sherman D.H. (2012) Diversity of P450 enzymes in the biosynthesis of natural products. *Nat Prod Rep*. **29** (10): 1251-1266.

[290] Guengerich F.P., Tang Z., Cheng Q., Salamanca-Pinzón S.G. (2011) Approaches to deorphanization of human and microbial cytochrome P450 enzymes. *Biochim Biophys Acta.* **1814** (1): 139-145.

[291] Vasina J.A., Baneyx F. (1996) Recombinant protein expression at low temperatures under the transcriptional control of the major Escherichia coli cold shock promoter cspA. *Appl Environ Microbiol*. **62** (4): 1444-1447.