

Discovery, Engineering and Application of Self-Sufficient P450s for Biocatalysis

A thesis submitted to the University of Manchester for the degree of
Doctor of Philosophy
in the Faculty of Science and Engineering

2018

Michele Tavanti
School of Chemistry

Table of Contents

Abbreviations	5
Amino Acid Abbreviations	7
Abstract: Discovery, Engineering and Application of Self-Sufficient P450s for Biocatalysis	8
Declaration and Copyright Statement.....	9
Acknowledgements.....	10
Chapter 1. Introduction.....	11
1.1 Oxidation Reactions Catalysed by Cytochrome P450s.....	11
1.2 Structural Features Associated with C-H Bond Activation	14
1.2.1 Thiolate Ligation.....	14
1.2.2 Proton Delivery	15
1.3 P450s for Industrial Organic Chemistry	16
1.4 Aims and Outline	20
1.5 References.....	21
Chapter 2. A Panel of New Thermostable CYP116B Self-Sufficient Cytochrome P450 Monooxygenases Catalyzing C-H Activation with Diverse Substrate Scope.....	23
2.1 Acknowledgements.....	23
2.2 Abstract.....	24
2.3 Introduction	24
2.4 Results and Discussion	27
2.4.1 Identification of RhF Homologues from Thermophiles.....	27
2.4.2 Expression, Purification and UV-visible Absorption Properties	29
2.4.3 Substrate Scope and Nucleotide Cofactor Preference.....	30
2.4.4 Thermal Stability of the P450 Panel.....	33
2.4.5 Whole Cell Biotransformations of Diclofenac.....	36
2.5 Conclusions	39
2.6 Subsequent Work	40
2.7 References.....	43
Chapter 3. The Crystal Structure of P450-TT Heme-Domain Provides the First Structural Insights into The Versatile Class VII P450s.....	46
3.1 Acknowledgements.....	46
3.2 Abstract.....	46
3.3 Introduction	47
3.4 Results and Discussion	48
3.4.1 Overall Structure	48
3.4.2 Active Site.....	50
3.4.3 Structural Comparison to Other Bacterial P450s	52
3.5 Subsequent Work	54
3.6 References.....	57

Chapter 4. One-pot Biocatalytic Double Oxidation of α-Isophorone for the Synthesis of Ketoisophorone.....	59
4.1 Acknowledgements.....	59
4.2 Abstract.....	60
4.3 Introduction	60
4.4 Results and Discussion	63
4.4.1 Screening and Improvement of P450cam-RhFRed Variants	63
4.4.2 Selection of HIP-oxidizing Cm-ADH10 and Structural Characterization.....	67
4.4.3 Assembling the System: Initial Unsuccessful Attempts	68
4.4.4 Parallel Reaction Optimization.....	70
4.4.5 Preparative-scale Double Oxidation of α -Isophorone to Ketoisophorone	73
4.5 Conclusions	75
4.6 References.....	76
Chapter 5. A Biocatalytic Cascade for the Amination of Unfunctionalized Cycloalkanes... 79	
5.1 Acknowledgements.....	79
5.2 Abstract.....	80
5.3 Introduction	80
5.4 Results and Discussion	81
5.5 References.....	86
Chapter 6. General Discussion and Conclusions	88
Chapter 7. Supporting Information.....	90
7.1 A Panel of New Thermostable CYP116B Self-Sufficient Cytochrome P450 Monooxygenases Catalysing C-H Activation with Diverse Substrate Scope.....	90
7.1.1 Experimental	90
7.1.1.1 Chemicals	90
7.1.1.2 Target Selection and Phylogenetic Analysis	90
7.1.1.3 Gene Cloning	91
7.1.1.4 Protein Expression and Purification	96
7.1.1.5 Cytochrome c Assay	100
7.1.1.6 Biotransformations	101
7.1.1.7 Assessment of Thermal Stability	103
7.1.1.8 Diclofenac Scale-Up.....	105
7.1.1.9 Determination of Coupling Efficiency	106
7.1.1.10 Analytics	106
7.1.1.11 Determination of Enantiomeric Excess	107
7.1.2 HPLC and GC-FID/MS data	108
7.1.3 References.....	125
7.2 The Crystal Structure of P450-TT Heme-Domain Provides the First Structural Insights into the Versatile Class VII P450s	126
7.2.1 Experimental	126
7.2.1.1 Materials	126
7.2.1.2 Molecular Biology Methods	126
7.2.1.3 Protein Expression and Purification	129
7.2.1.4 Expression and Screening of P450-AX Libraries in 96-well plates	131
7.2.1.5 Turnover Frequencies, Coupling Efficiency, Total Turnover Numbers, Whole-Cell Activity and Thermal Stability Determination	132
7.2.1.6 Crystallization, Data Collection and Structure Determination	133
7.2.1.7 Sequence Analysis	133
7.2.1.8 Structural Analysis.....	134
7.2.1.9 UV-visible Spectroscopic Properties of P450-TT Heme Domain	134

7.2.2	Supporting Figures and Tables	135
7.2.3	References.....	146
7.3	One-Pot Biocatalytic Double Oxidation of α-Isophorone for the Synthesis of Ketoisophorone	147
7.3.1	Experimental	147
7.3.1.1	Materials	147
7.3.1.2	Molecular Biology Methods	147
7.3.1.3	Protein Expression and Purification	148
7.3.1.4	Protein Crystallization, Structure Determination and Docking experiments	150
7.3.1.5	Clustering of Indigo-Positive Variants	151
7.3.1.6	Whole Cell Biotransformations with Library D P450cam-RhFRed Variants	151
7.3.1.7	Screening of ADHs	152
7.3.1.8	Scale Up of the Biotransformation of α -Isophorone with P450cam-RhFRed L244I-V247S for the Identification of 7	152
7.3.1.9	Preparation of (R)-6 and (S)-6	153
7.3.1.10	Analytical Scale Double Oxidation of α -Isophorone to Ketoisophorone.....	153
7.3.1.11	Preparative-Scale Double Oxidation of α -Isophorone to Ketoisophorone	154
7.3.1.12	Analytical Methods	155
7.3.2	Supplementary Figures and Tables	156
7.3.2.1	NMR and HRMS Spectra.....	161
7.3.3	References.....	164
7.4	A Biocatalytic Cascade for the Amination of Unfunctionalized Cycloalkanes	165
7.4.1	Experimental	165
7.4.1.1	Materials	165
7.4.1.2	Molecular Biology Methods	165
7.4.1.3	Protein Production and Purification.....	166
7.4.1.4	Determination of Volumetric Enzymatic Activities	168
7.4.1.5	Chemical Synthesis and Characterization	169
7.4.1.6	Spectroscopical Characterization	170
7.4.1.7	Biocatalytic Cascade for the Amination of Cycloalkanes.....	175
7.4.1.8	Analytics	177
7.4.2	Supporting Figures and Tables	178
7.4.3	GC-FID Chromatograms and GC-MS Traces	184
7.4.3.1	GC-FID Chromatograms.....	184
7.4.3.2	GC-MS Traces	195
7.4.4	References.....	198
Chapter 8. Other Contributions	199	
8.1	The Self-Sufficient P450 RhF Expressed in a Whole Cell System Selectively Catalyzes the 5-hydroxylation of Diclofenac.....	199
8.2	Cloning, Expression and Characterization of P450-Hal1 (CYP116B62) from <i>Halomonas</i> sp. NCIMB 172: a Self-Sufficient P450 with High Expression and Diverse Substrate Scope.....	200
8.3	Characterization of CYP102A25 from <i>Bacillus marmarensis</i> and CYP102A26 from <i>Pontibacillus halophilus</i>: P450 Homologues of BM3 with Preference Towards Hydroxylation of Medium-Chain Fatty Acids	201

Word Count: 43,694

Abbreviations

5-ALA	<i>5-aminolevulinic acid</i>
5-OH DCF	<i>5-hydroxydiclofenac</i>
ADH	<i>Alcohol Dehydrogenase</i>
AIM	<i>Autoinduction Medium</i>
AspRedAm	<i>Reductive Aminase from Aspergillus oryzae</i>
CboFDH	<i>Formate Dehydrogenase from Candida boidinii</i>
CDCl ₃	<i>Deuterated Chloroform</i>
Cm-ADH10	<i>ADH from Candida magnoliae</i>
Cpd 0	<i>Compound 0</i>
Cpd I	<i>Compound I</i>
Cpd II	<i>Compound II</i>
CPR	<i>NADPH-cytochrome P450 reductase</i>
CYPs	<i>Cytochrome P450s</i>
DCF	<i>Diclofenac</i>
DMSO	<i>Dimethyl Sulfoxide</i>
EtOAc	<i>Ethyl acetate</i>
FID	<i>Flame-Ionization Detector</i>
FMN	<i>Flavin Mononucleotide</i>
GC	<i>Gas Chromatography</i>
g _{cdw}	<i>Grams of Dry Cells</i>
GDH	<i>Glucose Dehydrogenase</i>
HIP	<i>4-hydroxy-α-isophorone</i>
HMIP	<i>3-(hydroxymethyl)-5,5-dimethylcyclohex-2-en-1-one</i>
HPLC	<i>High-pressure liquid chromatography</i>
HRMS	<i>High-Resolution Mass Spectrometry</i>
IMAC	<i>Immobilized metal ion affinity chromatography</i>
IPO	<i>2,3-epoxy-3,5,5-trimethyl-1-cyclohexanone</i>
<i>i</i> -PrOH	<i>Isopropanol</i>
IPTG	<i>β-D-1-thiogalactopyranoside</i>
KIP	<i>Ketoisophorone</i>
KPi	<i>Potassium Phosphate</i>
LB	<i>Luria-Bertani</i>
LC	<i>Liquid Chromatography</i>
MeCN	<i>Acetonitrile</i>
MeOH	<i>Methanol</i>
MS	<i>Mass-Spectrometry</i>
MTBE	<i>Methyl tert-butyl ether</i>
NAD(P)H	<i>Reduced Nicotinamide Adenine Dinucleotide (Phosphate)</i>

OTR	<i>Oxygen Transfer Rate</i>
P450-AT	<i>CYP116B65</i>
P450-AX	<i>CYP116B64</i>
P450-BM3	<i>CYP102A1</i>
P450cam	<i>CYP101A1</i>
P450-CreJ	<i>CYP288A2</i>
P450-JT	<i>CYP116B63</i>
P450-RhF	<i>CYP116B2</i>
P450s	<i>Cytochrome P450s</i>
P450-TB	<i>CYP116B29</i>
P450-TT	<i>CYP116B46</i>
PCR	<i>Polymerase-Chain Reaction</i>
Pdx	<i>Putidaredoxin</i>
PFOR	<i>Phthalate Family Oxygenase Reductase Domain</i>
RedAm	<i>Reductive Aminase</i>
RMSD	<i>Root-Mean Square Deviation</i>
ROS	<i>Reactive Oxygen Species</i>
rpm	<i>Revolutions per minute</i>
SEC	<i>Size-Exclusion Chromatography</i>
SSCR	<i>NADPH-dependent Carbonyl Reductase</i>
STY	<i>Space-Time Yield</i>
TB	<i>Terrific Broth</i>
THF	<i>Tetrahydrofuran</i>
TLC	<i>Thin-layer Chromatography</i>
TOF	<i>Turnover Frequency</i>
TTN	<i>Total Turnover Numbers</i>
vvm	<i>Volume of gas (or mixture of them) per unit volume of liquid medium per minute</i>
α -IP	<i>α-isophorone</i>
β -IP	<i>β-isophorone</i>

Amino Acid Abbreviations

A, Ala	Alanine	M, Met	Methionine
C, Cys	Cysteine	N, Asp	Asparagine
D, Asp	Aspartic Acid	P, Pro	Proline
E, Glu	Glutamic Acid	Q, Gln	Glutamine
F, Phe	Phenylalanine	R, Arg	Arginine
G, Gly	Glycine	S, Ser	Serine
H, His	Histidine	T, Thr	Threonine
I, Ile	Isoleucine	V, Val	Valine
K, Lys	Lysine	W, Trp	Tryptophan
L, Leu	Leucine	Y, Tyr	Tyrosine

Abstract: Discovery, Engineering and Application of Self-Sufficient P450s for Biocatalysis

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science and Engineering (School of Chemistry), 2018

Tavanti Michele

Cytochrome P450 monooxygenases (P450s) are a widespread class of enzymes involved in biosynthetic pathways and drug metabolism. They can catalyse a diverse range of challenging reactions, including C-H activations. Since these reactions often occur in a regio- and stereoselective fashion, harnessing the oxidative power of these enzymes for biocatalysis has attracted pharmaceutical and fine chemical industries. Despite their immense potential, the implementation of these enzymes as industrial biocatalysts is limited by some issues such as low expression levels, stability and substrate scope. Catalytically self-sufficient P450s, in which the heme and reductase domains are fused in a single polypeptide, are of particular interest for the development of chemical routes to high value-added compounds, as the need for identification and expression of separate redox partners is negated.

In an effort to enrich the number of P450s with valuable biocatalytic properties and encourage their utilization, this thesis describes the discovery, characterization, engineering and application of self-sufficient P450s.

Enzymes from thermophilic organisms were selected as starting point of our investigation, as thermostable proteins often possess increased stability under process conditions and mutational robustness. This initial search led to a panel of new and diverse class VII P450s displaying not only higher expression levels and thermal stability compared to their mesophilic counterpart, but also a remarkable substrate promiscuity. From this expanded panel of characterized enzymes, the first crystal structure of a class VII P450 was also determined, providing a valuable framework for future protein engineering.

Concurrently, P450-enabled biocatalytic cascades were also developed. Complications associated with the application of crude enzyme preparations and biocatalysts compatibility were addressed to demonstrate the applicability of these multi-enzyme systems on a preparative scale.

Declaration and Copyright Statement

Except where stated, the work referred to in the thesis has not been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

i) The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii) Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii) The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv) Further information on the conditions under which disclosure, publication and commercialization of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see <http://www.library.manchester.ac.uk/about/regulations/>) and in The University’s policy on Presentation of Theses

Acknowledgements

This work was accomplished within the European Union's Horizon 2020 Programme (ROBOX project) that is acknowledged for funding.

I am especially grateful to Professor Turner and Professor Flitsch for their invaluable support, trust and guidance throughout my doctoral studies. I feel honoured and privileged to have worked independently in such a constructive environment.

I must also thank Dr. Mark Corbett and Dr. Ian Rowles for managing our projects brilliantly, and Ms. Paula Tipton for ensuring things ran smoothly. Special thanks go to Dr. Colin W. Levy for his precious help with crystallography.

I will thank personally all the other people who contributed to this PhD, as I believe that a sincere hug is worth a thousand words. Nothing would have been possible without their help, especially in the darkest days. Moreover, they created a friendly and relaxed working environment which I believe strongly contributed to our achievements.

I greatly hope that this research group will keep doing exciting science, leading biocatalysis to a wider application through a "genes-to-kilo" philosophy. Moreover, I wish every present and past group member years of success, fulfilment and fun.

Chapter 1. Introduction

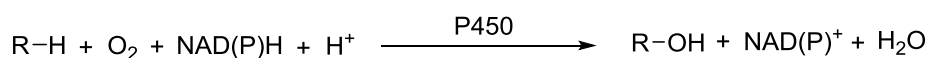
1.1 Oxidation Reactions Catalysed by Cytochrome P450s

Oxygen in the form of O_2 is an abundant component of Earth's atmosphere. Living organisms not only use molecular oxygen for oxidative phosphorylation, but also for the functionalization of organic molecules. Oxidizing enzymes employ oxygen as electron acceptor to carry out a variety of reactions involved in primary metabolism (i.e. reactions directly involved in cell growth), xenobiotic metabolism, synthesis of vitamins, hormones and other compounds.^[1] Despite its important roles in aerobic life, atmospheric O_2 is relatively inert towards most organic substrates. In fact, molecular oxygen is a triplet in its ground state (two unpaired electrons), whereas most organic substrates are in the singlet state (no unpaired electrons). Chemical reactions between triplet and singlet molecules are spin-forbidden and occur slowly.^[2] Thus, the electronic structure of O_2 prevents undesired reactions, protecting living organisms against deleterious oxidative stress. This required the evolution of enzymes capable of activating molecular oxygen in a controlled environment to carry out selective oxyfunctionalisation.

Most oxygen activating enzymes incorporate transition metals in order to overcome the barriers associated with the electronic structure of molecular oxygen.^[3,4] Cytochrome P450 monooxygenases (CYPs or P450s) are probably one of the best studied classes of metalloenzymes, characterized by the presence of a heme-iron prosthetic group at the active site. The vastness and diversity of P450 sequence space is well documented by the work of Prof. Nelson, which assigned nomenclature to more than 41,000 P450 genes, while sequencing projects have collected more than 300,000 CYP genes.^[5] In order to identify a specific P450, a nomenclature system has been devised based on sequence identity.^[6]

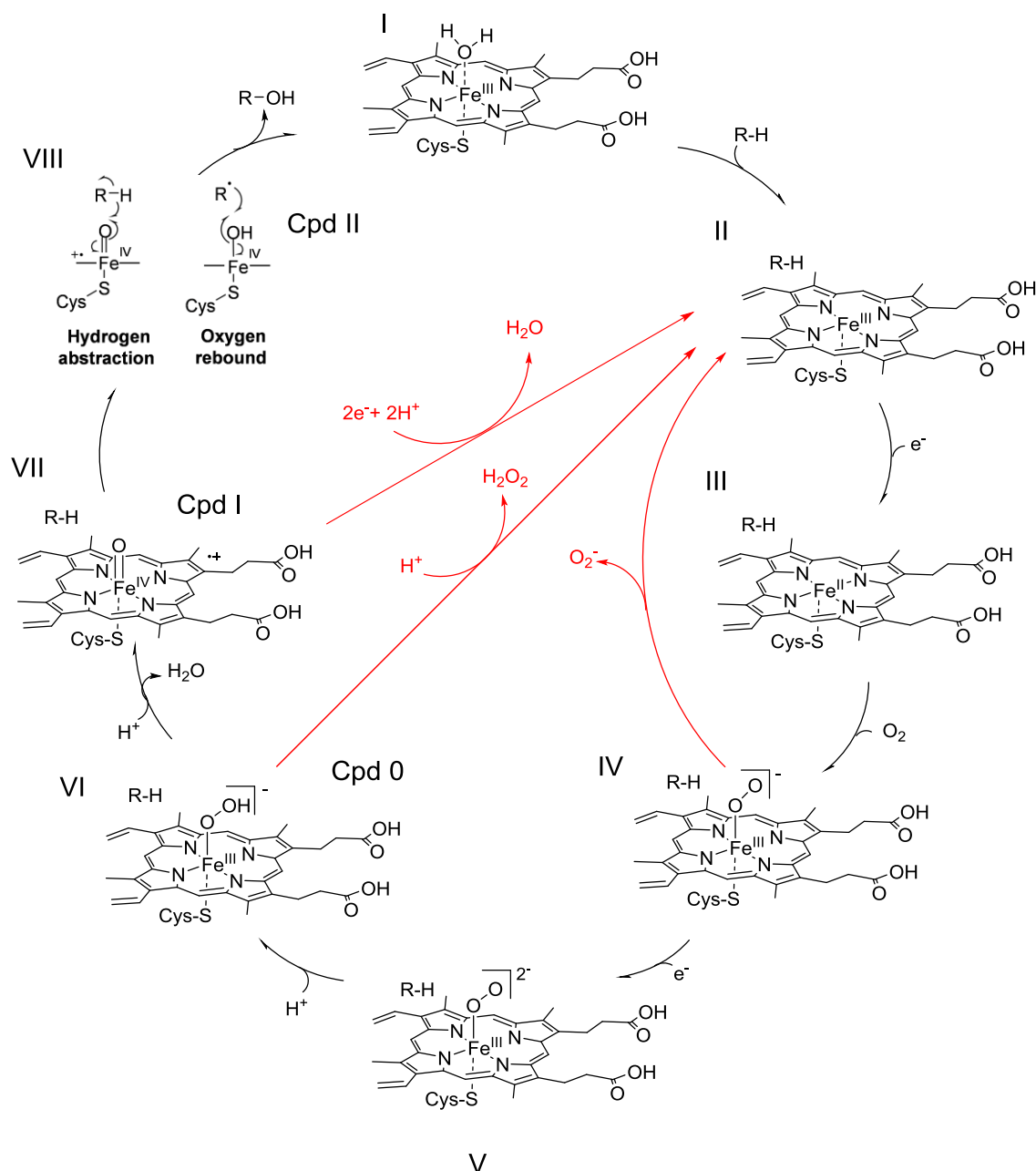
Cytochrome P450s are well known for their pivotal role in xenobiotic detoxification and biotransformation of endogenous compounds such as hormones and vitamins.^[7] These potent enzymes are capable of catalysing the hydroxylation of saturated C-H bonds, epoxidation of olefins, dealkylation reactions, aromatic hydroxylation and other less common reactions using molecular oxygen as oxidant.^[8] Overall, P450s insert one oxygen atom from

O₂ into the substrate (monooxygenation) and reduce the other one to water, using two electrons provided by NAD(P)H via redox partners (Scheme 1).^[9]



Scheme 1. Overall reaction scheme for a general P450-catalyzed hydroxylation.

The heme-iron is hexacoordinated in its resting state, with the proximal conserved cysteinate and water as axial ligands (I, Scheme 2). Ligand binding is a crucial step in the catalytic cycle, not only because it affects reaction selectivity, but also for its regulatory role in triggering the electron transfer from redox partners. In fact, axial water is displaced upon ligand binding, converting the enzyme from its resting, low-spin ferric form to the substrate-bound, high-spin ferric enzyme, giving rise to an upshift of redox potential (II, Scheme 2). This regulatory role of ligand binding is a clever design to avoid wastage of reducing equivalents in absence of the substrate (uncoupling of electrons to product formation, red branches, Scheme 2). Given the increased redox potential, the ferric intermediate is reduced by one electron supplied by a redox partner to its ferrous state, which then binds molecular oxygen (III and IV). This oxy-intermediate is reduced by another electron to give the ferric-peroxo intermediate (V) and subsequently protonated to Compound 0 (Cpd 0, VI). A second protonation step on the distal oxygen of the complex leads to departure of a water molecule and generation of the highly reactive porphyrin radical cation Compound I (Cpd I, VII). Generally, P450 reactions can be rationalized taking into account the ability of this intermediate to abstract hydrogen from the substrate to produce a carbon radical and the intermediate Compound II (Cpd II, VIII), followed by –OH insertion (oxygen rebound mechanism), product release and regeneration of the ferric enzyme.^[10] More complex P450 reactions may occur due to involvement of different intermediates in the catalytic cycle.^[8]



Scheme 2. The catalytic cycle of P450s. Uncoupling pathways are coloured in red.

As shown in Scheme 2 (red arrows), electrons shuttled by the redox partner can be wasted in the production of superoxide, hydrogen peroxide and water, instead of being used for substrate turnover. This so-called uncoupling limits the catalytic performance of P450s, increasing the cofactor demand and damaging the same enzyme through the generation of reactive oxygen species (ROS).^[11] Usually, coupling efficiency (ratio of the product molecules generated per NADPH molecules used) is close to 100% when P450s are presented with natural substrates,^[12] but in the presence of unnatural substrates or when mutations affecting

substrate binding are introduced, typical coupling efficiency values range from less than 10% to 40%.^[13]

The extent of the uncoupling of NAD(P)H oxidation from product formation depends on the nature of the substrate relative to P450 specificity, the efficiency of proton and electron delivery to reaction intermediates, water positioning into the active site and interaction with redox partners.^[14–16] The coupling of desolvation of the active site and redox potential shift upon substrate binding mentioned above is an elegant way through which the enzyme avoids uncoupling reactions.

1.2 Structural Features Associated with C-H Bond Activation

1.2.1 Thiolate Ligation

Despite major differences in primary structure, substrate scope and electron transfer systems, P450s display a highly conserved overall architecture (see Chapter 3 for further information).^[17] The ability of these versatile enzymes to catalyse the selective oxidation of a vast array of diverse substrates depends on the fine positioning of various structural elements (e.g., B'-helix, F-helix, G-helix and F/G loop) involved in substrate recognition and on the flexibility of this fold, with movements of backbone atoms up to 18 Å reported in the literature.^[18]

Unsurprisingly, the most conserved structural elements are found around the heme prosthetic group, where oxygen activation chemistry takes place. On the proximal side of the heme, a hydrophobic pocket comprises the absolutely conserved cysteinate axial ligand which is placed at hydrogen bonding distance with two backbone -NH groups. The electrostatic interaction between the Cys ligand and these amide protons stabilize heme-iron coordination and regulate the redox potential of the heme.^[19] Thiolate-ligation is one of the most fascinating and counterintuitive aspects of P450 chemistry. It is now clear that the critical function of the thiolate ligand is to afford a catalyst that can promote C-H activation chemistry while keeping the redox potential low enough to prevent oxidation of the protein framework.^[20] The driving force for C-H bond scission is given by the difference between the energy of the C-H bond broken and the strength of the O-H bond formed in Cpd II: a stronger

O–H bond increases the driving force for C–H bond scission.^[21] On the one hand, cysteinate ligation can reduce heme reduction potential due to the strong electron donor ability of the proximal ligand, but on the other hand, the “push effect” exerted by the proximal ligand increases the basicity of Cpd II (pKa~12).^[22] As the strength of the O–H bond formed in Cpd II after H-abstraction by Cpd I depends on the one-electron reduction potential of compound I and on the pKa of compound II, the increased basicity of Compound II afforded by thiolate ligation promotes C-H activation at a biologically viable Cpd I reduction potential.^[21]

1.2.2 Proton Delivery

During the P450 catalytic cycle, two electrons are shuttled from NAD(P)H to the heme via redox partners, while two water-derived protons are delivered to the iron-bound dioxygen. Structural biology studies on the model P450 CYP101A1 (camphor 5-hydroxylase, P450cam) provided much of our understanding about proton delivery in P450s. The first crystal structure of P450cam was quite puzzling, as it revealed a highly hydrophobic active site, lacking acid-base catalytic residues to facilitate proton transfer.^[23] However, sequence conservation analyses revealed a high degree of conservation of the signature sequence (Ala/Gly)GlyX(Glu/Asp)Thr, comprising a hydrogen-bond donor/acceptor (Thr252 in P450cam) preceded by an acidic residue (Asp251 in P450cam).^[15] This sequence motif is located on the long I-helix (~30 residues), making up the distal side of the heme and the active site. Indeed, it was later clarified that these amino acids are involved in delivering water-derived protons to the bound oxygen.^[24] In “closed” P450 structures, the threonine side-chain -OH is hydrogen-bonded to the backbone carbonyl oxygen of a preceding residue (Gly248 in P450cam) that would otherwise be involved in establishing regular α -helical contacts. This deviation from the helical geometry causes a widening of the I-helix that is important for water binding in the active site. The crystal structure of the oxy-P450cam reveals an even wider I-helix groove due to disruption of the unusual hydrogen bond between Thr252 and Gly248 which enables water molecules to move close to the heme.^[25]

In this scenario, a critical role is also played by Asp251: according to the most recent proposal, the interaction between P450cam and its redox partner (putidaredoxin, Pdx) would break Asp251 ion pairs with Arg186 and Lys178, allowing the acidic residue to rotate into the active site to deliver proton from the solvent to dioxygen.^[17,26] Additionally, mutational studies

suggest that Thr252 would function as hydrogen-bond acceptor to the hydroperoxy intermediate, promoting the second protonation step.^[24]

Although the acid-alcohol pair in the I-helix is a conserved feature in P450 sequences, exceptions exist. Crystal structure analysis of these enzymes revealed that substrate-assisted catalysis is exploited for proper positioning of water molecules close to the bound dioxygen.^[27,28]

1.3 P450s for Industrial Organic Chemistry

Cytochrome P450s have an impressive biotechnological potential, given their ability to functionalize inert C-H bonds in a regio- and stereoselective manner. Moreover, they can accept a vast variety of substrates and convert them to high added-value compounds using oxygen as oxidant under mild conditions. These desirable characteristics make them exceptional biocatalysts to use in green chemistry.^[29] In particular, these enzymes are of great interest to the pharmaceutical and fine chemical industries.^[30] For example, P450s could be exploited for the production of drugs and human drug metabolites, which are necessary for testing and approval of therapeutics under development.^[31] However, many limitations restrict the applicability of these enzymes for industrial biocatalysis (Table 1).^[32]

Table 1. Limitations associated with the application of P450s in biocatalysis.

	Limitations
Enzyme-related	Expression levels
	TOF
	Stability
	TTN
	Cofactor supply
	Need for redox partners
	Uncoupling
Process-related	Oxygen supply
	Substrate uptake/solubility
	Substrate/product inhibition
	Substrate/product toxicity

The notorious instability of these enzymes under process conditions, their multicomponent nature and the requirement of expensive reduced cofactors, means that the use of resting (metabolically active) whole-cells expressing the desired P450 gene is highly desirable.^[33] Even though resting whole-cells regenerate nicotinamide cofactors when provided with a carbon source, it has been shown that external supply of reduced cofactors and co-expression of cofactor regenerating enzymes (e.g., dehydrogenases) can improve product formation rates.^[34,35] This suggests that other cellular processes and uncoupling might compete for reducing equivalents with the P450 biotransformation. Uncoupling reactions have already been mentioned as a determinant of the poor biocatalytic performance of P450s. These abortive pathways not only compete for NAD(P)H consumption, but also contribute to protein/whole-cell instability, reducing the achievable total turnover numbers (TTN, moles of product(s) formed per mole of P450).^[11] Considering that P450s generally have low turnover frequency values (TOF, moles of product(s) formed per minute per mole of enzyme, typically lower than 1000 min^{-1})^[36] and poor expression levels, the extension of biocatalyst lifetime through high coupling efficiency is a prerequisite of crucial importance for the implementation of these enzymes in industrial processes.

Production of functional P450s in the widely employed *E. coli* is complicated by the incompatibility of the prokaryotic translational machinery with membrane signal modules of eukaryotic enzymes,^[37] insufficient endogenous heme synthesis^[38] and lack of endogenous redox partners.

Several P450 classes have been defined according to the electron transport system exploited to complete the catalytic cycle. For many years after the discovery of P450s it was believed that two major classes (class I and class II P450s) could account for the diversity of redox partner systems. Class I includes prokaryotic and mitochondrial (eukaryotic) enzymes: both systems employ an FAD-containing reductase (ferredoxin reductase) and an iron-sulfur protein (ferredoxin) to shuttle electrons from NAD(P)H to the P450. In the prokaryotic system all the components are cytoplasmic, whereas in the mitochondrial systems only the ferredoxin does not interact with the mitochondrial membrane. Class II P450s are mainly eukaryotic enzymes found in the endoplasmic reticulum, composed by a membranous P450 and a membrane-anchored NADPH-dependent FAD/FMN containing CPR (NADPH-cytochrome P450 reductase). As the number of characterized P450s increases, so does the

variety of electron transport chains. An array of different redox partner systems has been observed, with as much as ten classes reported to date.^[9] Importantly, fused systems have been described in which the redox partners are linked to the heme domain of the enzyme, creating multicomponent, self-sufficient systems. The first such system was discovered in *Bacillus megaterium* (CYP102A1, P450-BM3, classified as a class VIII P450 according to the electron transport chain),^[39] followed by a distinct system in *Rhodococcus* sp. NCIMB 9784 (CYP116B2, P450-RhF, class VII).^[40] In this context, the use of self-sufficient systems is very attractive, as the need for identification and production of separate electron transfer systems is obviated.

Considerable research has been devoted to tackling the above-mentioned limitations through protein and host engineering.^[30] However, process engineering is another valuable approach to address another set of parameters associated with P450 biocatalysis.

Monooxygenases require stoichiometric amounts of molecular oxygen as a substrate for the enzymatic reaction. However, the solubility of oxygen in water at 30°C and atmospheric pressure is around 0.24 mM and can be lower in the presence of salts and other compounds.^[41] Thus, oxygen is rapidly depleted at relevant enzyme concentrations. Moreover, oxygen will also be consumed in uncoupling reactions and for cell maintenance if a whole-cell reaction is carried out. In the latter case, oxygen is preferentially used for respiration instead of monooxygenation, given the higher oxygen affinity of the respiratory chain than monooxygenases.^[42] Therefore, oxygen supply is a central aspect for the development of a P450-based process and the rate at which oxygen is provided (OTR, oxygen transfer rate) will define the maximum rate of reaction (assuming that no other factors are limiting the reaction).

Finally, substrate and product related parameters need to be considered carefully as these compounds are often hydrophobic and their limited transfer across the cell membrane can limit the performance of the biocatalyst. Furthermore, toxic and inhibitory effects can have detrimental outcomes. Several permeabilization methods (physical, chemical and genetic modification) are available today to minimize the effect of mass-transfer limitations across the cell membrane.^[43] For example, co-expression of outer membrane proteins boosted

linear alkane hydroxylation activities up to 28-fold and enabled a two-fold improvement in limonene hydroxylation activity in recombinant *E.coli*.^[44,45]

The use of biphasic solvent systems has many advantages when there are issues associated with reactants and products: organic solvents can be employed to facilitate the transport of hydrophobic substrates to the aqueous phase^[46] and to establish a sink for toxic products.^[47] However, organic solvents must be chosen carefully given the potential negative effects on biocatalyst activity.^[48,49] Product inhibition/toxicity can be tackled through *in-situ* product removal,^[50] whereas substrate inhibition can be minimized by adopting substrate feeding strategies.^[51]

In conclusion, P450s are powerful biocatalysts but many inherent limitations are holding back their wide industrial implementation. To meet target values for relevant process metrics such as final product concentration, space-time yield and biocatalyst yield, a joint effort between chemists, biologists and chemical engineers is needed. Biocatalyst instability remains one of the major factors that limits the establishment of P450-based processes.^[30] Considering the low TOF values for these enzymes, combined with their poor stability, it has been suggested that the main target for P450-based process should be high value molecules, such as drugs, drug metabolites, flavour and fragrance ingredients, for example. If bulk chemicals would be chosen as targets, then significant biocatalyst improvements would be needed.^[52]

1.4 Aims and Outline

This PhD thesis is presented in Journal Format as most of the research carried out during the postgraduate programme constitutes a continuous body of work that has either been published or due to be submitted in peer-reviewed journals. The goal of the current thesis project was to address the limitations associated with P450 catalysis and harness their oxidative power to enable the synthesis of valuable compounds.

Chapter 2 describes our efforts to identify self-sufficient P450s from thermophilic organisms with higher expression levels, stability and activity compared to their “mesophilic” homologues. For this investigation, we focused on class VII P450s, given their self-sufficiency and substrate promiscuity. Importantly, our study revealed that CYP116B64 (P450-AX) is a good candidate for the preparative scale synthesis of 5-hydroxydiclofenac (5-OH DCF). As process development is a central aspect for the implementation of P450s on a larger scale, Section 2.6 describes the investigation of the whole-cell 5-hydroxylation of diclofenac (DCF) in the controlled environment offered by a stirred tank reactor with relevant mixing and oxygen supply.

Although the demonstrated P450-catalyzed production of 5-OH DCF reached the highest reported titres for a biotechnological process for selective oxyfunctionalization of DCF to 5-OH DCF,^[53] target values for relevant economic metrics have not been met yet.^[52] Protein engineering is a common approach employed to overcome limitations associated with natural enzymes. This process is greatly facilitated when the 3D structure of the enzyme (or of its homologs) is available. Chapter 3 deals with the determination of the first crystal structure of a class VII P450 (CYP116B46, P450-TT), which provided us with a guideline for semi-rational mutagenesis of P450-AX for improved activity towards DCF. Preliminary data on P450-AX engineering are given in Section 3.5.

Self-sufficient P450s were also employed in biocatalytic cascades to enable the synthesis of ketoisophorone *via* double allylic oxidation of α -isophorone (Chapter 4) and amines from cycloalkanes through a four-enzyme cascade (Chapter 5). An overall discussion of the project can be found in Chapter 6, whereas materials, methods and supporting information are reported in Chapter 7. Additional contributions are highlighted in Chapter 8.

1.5 References

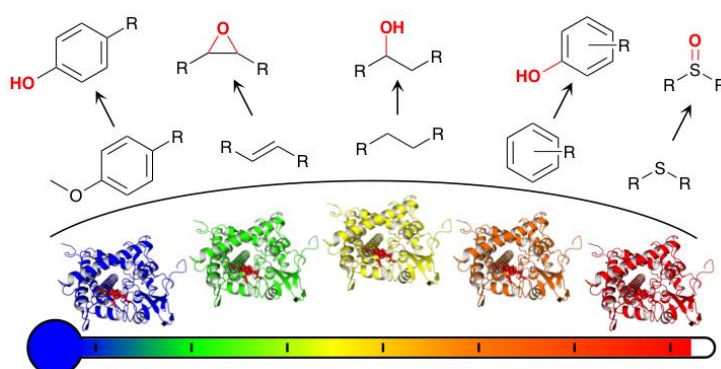
- [1] S. G. Burton, *Trends Biotechnol.* 2003, 21, 543–549.
- [2] D. F. Lewis, *J. Chem. Technol. Biotechnol.* 2002, 77, 1095–1100.
- [3] K. P. Jensen, U. Ryde, *J. Biol. Chem.* 2004, 279, 14561–14569.
- [4] M. M. Abu-Omar, A. Loaiza, N. Hontzeas, *Chem. Rev.* 2005, 105, 2227–2252.
- [5] D. R. Nelson, *Biochim. Biophys. Acta - Proteins Proteomics* 2018, 1866, 141–154.
- [6] D. R. Nelson, in *Methods Mol. Biol.*, Humana Press, Totowa, NJ, 2006, pp. 1–10.
- [7] P. R. Ortiz de Montellano, *Cytochrome P450: Structure, Mechanism and Biochemistry*, Springer Science & Business Media, 2015.
- [8] F. P. Guengerich, *Chem. Res. Toxicol.* 2001, 14, 611–650.
- [9] F. Hannemann, A. Bichet, K. M. Ewen, R. Bernhardt, *Biochim. Biophys. Acta - Gen. Subj.* 2007, 1770, 330–344.
- [10] A. W. Munro, H. M. Girvan, A. E. Mason, A. J. Dunford, K. J. McLean, *Trends Biochem. Sci.* 2013, 38, 140–150.
- [11] R. Fasan, M. M. Chen, N. C. Crook, F. H. Arnold, *Angew. Chemie* 2007, 119, 8566–8570.
- [12] M. J. Cryle, R. D. Espinoza, S. J. Smith, N. J. Matovic, J. J. De Voss, *Chem. Commun.* 2006, 2353.
- [13] P. Meinhold, M. W. Peters, A. Hartwick, A. R. Hernandez, F. H. Arnold, *Adv. Synth. Catal.* 2006, 348, 763–772.
- [14] P. J. Loida, S. G. Sligar, *Biochemistry* 1993, 32, 11530–11538.
- [15] I. G. Denisov, T. M. Makris, S. G. Sligar, I. Schlichting, *Chem. Rev.* 2005, 105, 2253–2277.
- [16] D. Holtmann, F. Hollmann, *ChemBioChem* 2016, 17, 1391–1398.
- [17] T. L. Poulos, *Chem. Rev.* 2014, 114, 3919–3962.
- [18] S.-Y. Park, K. Yamane, S. Adachi, Y. Shiro, K. E. Weiss, S. a Maves, S. G. Sligar, *J. Inorg. Biochem.* 2002, 91, 491–501.
- [19] S. Yoshioka, T. Tosha, S. Takahashi, K. Ishimori, H. Hori, I. Morishima, *J. Am. Chem. Soc.* 2002, 124, 14571–14579.
- [20] M. T. Green, *Curr. Opin. Chem. Biol.* 2009, 13, 84–88.
- [21] J. T. Groves, *Nat. Chem.* 2014, 6, 89–91.
- [22] T. H. Yosca, T. H. Yosca, J. Rittle, C. M. Krest, E. L. Onderko, A. Silakov, J. C. Calixto, R. K. Behan, M. T. Green, *Science* 2013, 342, 825–829.
- [23] T. L. Poulos, B. C. Finzel, I. C. Gunsalus, G. C. Wagner, J. Kraut, *J. Biol. Chem.* 1985, 260, 16122–16130.
- [24] S. Nagano, T. L. Poulos, *J. Biol. Chem.* 2005, 280, 31659–31663.
- [25] I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko, S. G. Sligars, et al., *Science* 2000, 287, 1615–1622.
- [26] S. Tripathi, H. Li, T. L. Poulos, *Science* 2013, 340, 1227–1230.
- [27] B. Zhao, F. P. Guengerich, M. Voehler, M. R. Waterman, *J. Biol. Chem.* 2005, 280, 42188–42197.
- [28] J. R. Cupp-Vickery, O. Han, C. R. Hutchinson, T. L. Poulos, *Nat. Struct. Biol.* 1996, 3, 632–637.
- [29] R. A. Sheldon, J. M. Woodley, *Chem. Rev.* 2017, 118, 801–838.
- [30] E. O’Reilly, V. Köhler, S. L. Flitsch, N. J. Turner, *Chem. Commun. (Camb)*. 2011, 47, 2490–501.

- [31] J. M. Caswell, M. O'Neill, S. J. C. Taylor, T. S. Moody, *Curr. Opin. Chem. Biol.* 2013, 17, 271–275.
- [32] M. K. Julsing, S. Cornelissen, B. Bühler, A. Schmid, *Curr. Opin. Chem. Biol.* 2008, 12, 177–186.
- [33] V. B. Urlacher, M. Girhard, *Trends Biotechnol.* 2012, 30, 26–36.
- [34] M. T. Lundemo, S. Notonier, G. Striedner, B. Hauer, J. M. Woodley, *Biotechnol. Prod. Process Eng.* 2016, 100, 1197–1208.
- [35] I. Kaluzna, T. Schmitges, H. Straatman, D. van Tegelen, M. Müller, M. Schürmann, D. Mink, *Org. Process Res. Dev.* 2016, 20.4, 814–819.
- [36] R. Bernhardt, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* 2014, 98, 6185–6203.
- [37] E. Leonard, M. A. G. Koffas, *Appl. Environ. Microbiol.* 2007, 73, 7246–7251.
- [38] I. N. Harnastai, A. A. Gilep, S. A. Usanov, *Protein Expr. Purif.* 2006, 46, 47–55.
- [39] L. O. Narhi, A. J. Fulco, *J. Biol. Chem.* 1986, 261, 7160–7169.
- [40] G. Roberts, G. Grogan, A. Greter, S. L. Flitsch, N. J. Turner, *J. Bacteriol.* 2002, 184, 3898–3908.
- [41] J. Brummund, M. Müller, T. Schmitges, I. Kaluzna, D. Mink, L. Hilterhaus, A. Liese, *J. Biotechnol.* 2016, 233, 143–150.
- [42] C. V. F. Baldwin, J. M. Woodley, *Biotechnol. Bioeng.* 2006, 95, 362–369.
- [43] R. R. Chen, *Appl. Microbiol. Biotechnol.* 2007, 74, 730–738.
- [44] M. K. Julsing, M. Schrewe, S. Cornelissen, I. Hermann, A. Schmid, B. Bühler, *Appl. Environ. Microbiol.* 2012, 78, 5724–5733.
- [45] S. Cornelissen, M. K. Julsing, J. Volmer, O. Riechert, A. Schmid, B. Bühler, *Biotechnol. Bioeng.* 2013, 110, 1282–1292.
- [46] A. Schmid, A. Kollmer, R. G. Mathys, B. Witholt, *Extremophiles* 1998, 2, 249–256.
- [47] H. Schewe, D. Holtmann, J. Schrader, *Appl. Microbiol. Biotechnol.* 2009, 83, 849–857.
- [48] T. S. Wong, F. H. Arnold, U. Schwaneberg, *Biotechnol. Bioeng.* 2004, 85, 351–358.
- [49] C. Laane, S. Boeren, K. Vos, C. Veeger, *Biotechnol. Bioeng.* 1987, 30, 81–87.
- [50] J. M. Woodley, M. Bisschops, A. J. J. Straathof, M. Ottens, *J. Chem. Technol. Biotechnol.* 2008, 83, 121–123.
- [51] I. Kaluzna, J. Brummund, M. Schuermann, *Chim. Oggi - Chem. Today* 2017, 35, 55–58.
- [52] M. T. Lundemo, J. M. Woodley, *Appl. Microbiol. Biotechnol.* 2015, 99, 2465–2483.
- [53] J. M. Klenk, L. Kontny, B. A. Nebel, B. Hauer, *bioRxiv* 2018, 315374.

Chapter 2. A Panel of New Thermostable CYP116B Self-Sufficient Cytochrome P450 Monooxygenases Catalyzing C-H Activation with Diverse Substrate Scope

Michele Tavanti,^[a] Joanne L. Porter,^[a] Selina Sabatini,^[a] Nicholas J. Turner^[a] and Sabine L. Flitsch^[a]

Published in *ChemCatChem*, **2018**, *10*, 1042–1051.



2.1 Acknowledgements

The research presented in this chapter was a collaborative effort between the doctoral candidate, Dr. Joanne L. Porter and Ms. Selina Sabatini under the supervision of Professor Nicholas J. Turner and Professor Sabine L. Flitsch. Dr. Joanne L. Porter and the doctoral candidate co-authored the manuscript, planned and carried out molecular biology and substrate screening experiments with the help of Ms. Selina Sabatini. The doctoral candidate performed thermal stability assays, cofactor preference assays and scale-up experiments. All the authors reviewed the text before submission.

The research presented in Section 2.6 was carried out during an external stay of the doctoral candidate at the Technical University of Denmark (DTU) with the assistance of Ms. Mafalda Dias Gomes and Mr. Rowan Malan Lindeque, under the supervision of Professor John M. Woodley.

^a Manchester Institute of Biotechnology (MIB), School of Chemistry, The University of Manchester, 131 Princess Street, M1 7DN, Manchester, United Kingdom

2.2 Abstract

The ability of cytochrome P450 monooxygenases to catalyse a wide variety of synthetically challenging C-H activation reactions makes them highly desirable biocatalysts both for the synthesis of chiral intermediates and for late stage functionalisations. However, P450s are plagued by issues associated with poor expression, solubility and stability. Catalytically self-sufficient P450s, in which the heme and reductase domains are fused in a single protein, obviate the need for additional redox partners and are attractive as biocatalysts. Here we present a panel of natural self-sufficient P450s from thermophilic organisms (CYP116B65 from *A. thermoflava*, CYP116B64 from *A. xiamenense*, CYP116B63 from *J. thermophila*, CYP116B29 from *T. bispora* and CYP116B46 from *T. thermophilus*). These P450s display enhanced expression and stability over their mesophilic homologues. Activity profiling of these enzymes revealed that each P450 displayed a different fingerprint in terms of substrate range and reactivity covering reactions as diverse as hydroxylation, demethylation, epoxidation and sulfoxidation. The productivity of the biotransformation of diclofenac to produce the 5-hydroxy metabolite rose by 42-fold using the thermostable P450-AX ($> 0.5 \text{ g L}^{-1} \text{ h}^{-1}$) compared to the previously reported P450-RhF system. In conclusion, we have generated a toolkit of thermostable self-sufficient P450 biocatalysts, with broad substrate range and reactivity.

2.3 Introduction

Cytochromes P450 (P450s) are a superfamily of heme-thiolate enzymes that are ubiquitous in Nature and play a primary role in drug metabolism and the biosynthesis of physiologically important compounds including steroids and vitamins.^[1] P450s have the ability to catalyse a wide range of synthetically challenging C-H activation reactions with high levels of regio- and stereoselectivity.^[2] The most common P450 catalysed reactions are monooxygenation reactions (e.g. hydroxylations and epoxidations) using a single oxygen atom from molecular oxygen, with the other oxygen atom reduced to water (Figure 1A). Two electrons supplied by NAD(P)H are shuttled following productive interaction with redox partners to enable molecular oxygen activation.^[3,4] As such, P450s have immense potential as biocatalysts due to the challenging nature and wide range of chemistry they can facilitate and present a green

alternative or partner to chemocatalysis.^[5] However despite these attractive features, the number of P450s in use industrially is limited.^[6] The key factors limiting their widespread application include low levels of expression, poor solubility and/or stability particularly in recombinant systems and narrow substrate scope.^[6-7]

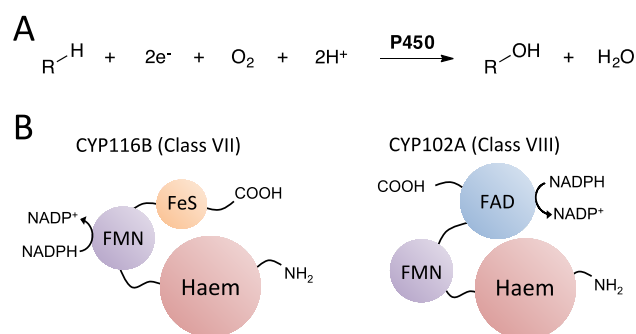


Figure 1. (A) P450-catalyzed oxidations, including C-H activation and (B) structural organization of CYP116B and CYP102A self-sufficient P450s.

There are two main approaches to overcome these restrictions (i) the use of protein engineering and (ii) the use of alternative naturally evolved enzymes from Nature's vast pool of diverse sequences. The use of natural over engineered enzymes is becoming an increasingly popular approach, especially in recent years with the advancements in DNA sequencing and the exponential increase in available sequences in protein and nucleotide databases.^[8] Given this increasing number of diverse, naturally evolved protein sequences available, there are now answers contained within these databases for many biocatalyst requirements. Moreover, the ease with which they can be accessed by gene synthesis outweighs the often-lengthy process of laboratory evolution or the sometimes fortuitous nature of rational engineering. Of particular interest for the development of biocatalytic processes is the use of enzymes from extremophiles (e.g. alkaliphiles, halophiles or thermophiles) that have already adapted to conditions in line with requirements for the target application.^[9] For example, enzymes originating from thermophilic organisms have been shown to display remarkable thermal and chemical stability in isolation and in recombinant systems.^[10] These enzymes are of immediate interest due to the crucial role stability plays when assessing the applicability of enzymes for industrial roles. Moreover, enhanced enzyme stability also creates an important platform from which to further engineer the catalytic properties of enzymes.^[11]

Previous studies have looked at the genomes of thermophilic organisms in order to identify more suitable biocatalysts. Nevertheless, the number of P450s from thermophilic microorganisms that have been studied so far is few, including CYP119A1 from *S. solfataricus*,^[12] CYP119A2 (P450st) from *S. tokodaii*,^[13] CYP154H1 from *T. fusca*^[14] and CYP175A1 from *T. thermophilus*.^[15] Given that required redox partners are also proteins, applications require not only thermostable heme proteins, but also the co-expression or addition of independent thermostable redox partners. In some cases the corresponding native thermophilic electron transport system has been identified, however their discovery is no simple task.^[12c,16] More conveniently, mesophilic reductase domains have been employed but limit the overall stability and coupling efficiency of the system to the stability of the most labile protein component.^[12a, 14] The self-sufficient nature of class VII (e.g. CYP116B2, P450-RhF) and class VIII (e.g. CYP102A1, P450-BM3) P450s containing the heme and reductase domains on a single polypeptide strand makes them desirable biocatalysts, as the need for identification and separate expression of redox partners is negated (Figure 1B). Isolated from the mesophilic bacterium *Bacillus megaterium*, the P450-BM3 fatty acid hydroxylase is the best studied self-sufficient P450.^[17] The subsequently discovered class VII self-sufficient enzymes include the mesophilic P450-RhF from *Rhodococcus* sp. NCIMB 9784,^[18] which displays a varied range of activities,^[19] including hydroxylation of the widely used anti-inflammatory drug diclofenac.^[20] P450-RhF has a unique FMN and 2Fe-2S ferredoxin containing phthalate dioxygenase-type reductase domain fused to the C-terminus of the P450 to create a monomeric 85-kDa self-sufficient enzyme.^[18a, 21] The reductase domain has been exploited to create artificial fusion proteins with various independent heme domains. These chimeric constructs have been used for benzylic oxidation reactions and the production of drugs and drug metabolites.^[22] For example, the RhF reductase domain was fused to the engineered P450_{Prava} and utilized in the single step fermentative production of the cholesterol lowering drug pravastatin.^[23] However, the stability of both the native and artificial constructs is poor, with previous results showing that upon purification the native enzyme is largely inactive due to instability of the FMN-FeS reductase domain.^[20]

The idea of combining thermostability and catalytic self-sufficiency to create a robust biocatalyst is very attractive. Recently a fungal BM3 homologue from *M. thermophila* has been isolated and characterized. However CO-binding studies monitoring conversion from P450 to the inactive P420 form with increasing temperature gave a T_{50} value of only 43 °C,^[24]

which is comparable to the mesophilic BM3.^[25] Alternatively, engineering studies have been directed towards creating thermostable self-sufficient enzymes. For example, thermostable P450 chimeras have been generated by exchange of the BM3 reductase domain with the more stable one from the homologue CYP102A3 or through structure-guided SCHEMA recombination of the heme domains of CYP102A homologues.^[26] In another study the heme domain of CYP175A1 was linked to its native redox partners to give an artificial fusion protein displaying 50 % residual activity after incubation at 80°C for 10 minutes.^[27]

Here, our aim was to identify natural P450s from thermophilic origins that could outperform their mesophilic counterparts in terms of expression, solubility and stability in recombinant systems. In particular, we sought self-sufficient P450s to avoid the search for redox partners with matching thermostability and to simplify protein production. Thus, the thermal stability and catalytic properties of five class VII P450s have been investigated and compared to that of P450-RhF. Given that these multi-domain catalysts have evolved in thermostable organisms, it was also interesting to ask if the observed thermostability would be the same for all subdomains, or if one of the domains consistently displayed higher stability in these five systems.

2.4 Results and Discussion

2.4.1 Identification of RhF Homologues from Thermophiles

Initially BLAST searches of nucleotide and protein databases were carried out using P450-BM3 or P450-RhF as the query sequence to identify homologues from thermophilic bacteria. Several BM3 homologues were identified from thermophilic organisms, however most were fungal (Table S1, section 7.1.1.2). Due to the interesting range of activity displayed by P450-RhF, as outlined above, we selected five RhF homologues from thermophilic bacteria for characterization: CYP116B65 (P450-AT) from *Amycolatopsis thermoflava*, CYP116B64 (P450-AX) from *Albidovulum xiamenense*, CYP116B63 (P450-JT) from *Jhaorihella thermophila*, CYP116B29 (P450-TB) from *Thermobispora bispora* and CYP116B46 (P450-TT) from *Tepidiphilus thermophilus* (Table 1). The amino acid sequences were aligned and the amino acid compositions were assessed (Figure S1, S2 and Table S2, section 7.1.1.3). It has been commonly observed that thermophilic enzymes have an increased number of hydrophobic

and aromatic interactions as well as more H-bonding and ionic interactions.^[28] As such, the amino acid composition of thermophiles can show trends towards larger non-polar, aromatic and charged amino acids at the expense of small non-polar and polar uncharged residues.^[10c, 29] Here general trends could be observed with fewer Ala and an increase in Ile, Phe and Arg residues with increasing thermal stability, though it is difficult to rationalize without structural context. A phylogenetic tree was constructed with the five new sequences and the five previously characterized class VII CYP116B enzymes (Figure 2). Preliminary data exist for several proposed class VII enzymes,^[30] however the number of fully characterized class VII enzymes is very limited with only five previously published (CYP116B1 from *C. metallidurans*, CYP116B3 from *R. ruber*, CYP116B4 from *L. aggregata*, CYP116B5 from *A. radioresistens* and the previously mentioned CYP116B2 RhF).^[18b, 31] The most closely related to RhF is P450-AT with 75 % sequence identity, followed by TB with 69 % identity. In all cases there is closer homology between the heme domains than the reductase domains.

Table 1. Sequence information for P450-RhF and the identified homologues AT, AX, JT, TB and TT

P450	MW (kDa)	CYP number	Organism	Cultivation temperature ^[a]	Source
RhF	85.3	CYP116B2	<i>Rhodococcus sp.</i> NCIMB strain 9784	25 °C	Soil
AT	84.9	CYP116B65	<i>Amycolatopsis thermoflava</i> DSM strain 44574	45 °C	Soil, China
AX	86.0	CYP116B64	<i>Albidovulum xiamenense</i> DSM strain 24422	55 °C	Terrestrial hot spring, China
JT	86.5	CYP116B63	<i>Jhaorihella thermophila</i> DSM strain 23413	28 °C	Coastal hot spring, Taiwan
TB	87.2	CYP116B29	<i>Thermobispora bispora</i> DSM strain 43833	45 °C	Decaying manure
TT	89.4	CYP116B46	<i>Tepidiphilus thermophilus</i> DSM strain 27220	50 °C	Terrestrial hot spring, India

^[a] Cultivation conditions for microorganisms were obtained from the NCIMB or DSMZ catalogues.

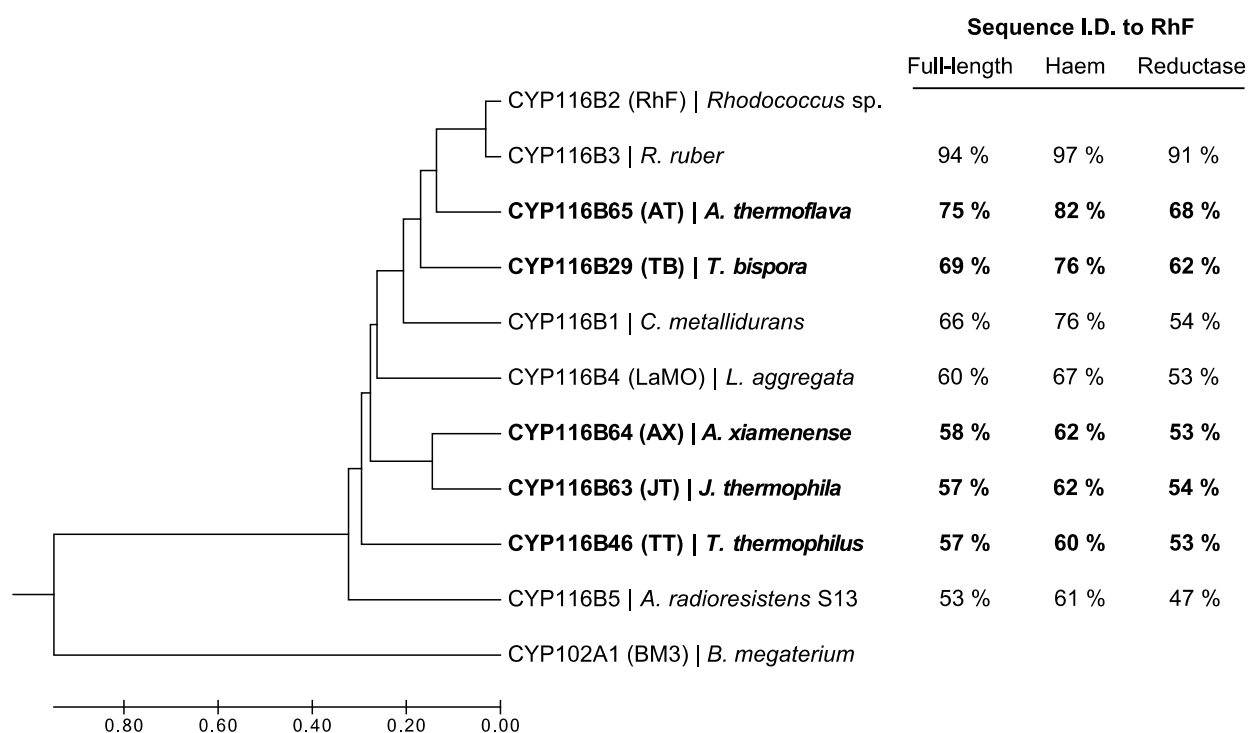


Figure 2. Phylogenetic tree with new and previously reported Class VII P450s as well as the Class VIII BM3. The evolutionary analyses were conducted in MEGA7 with sequences aligned using ClustalW and evolutionary history inferred using the UPGMA method. The evolutionary distances were calculated using the Poisson correction method and are in the units of number of amino acids substitutions per site.

2.4.2 Expression, Purification and UV-visible Absorption Properties

The genes encoding P450 AT, AX, JT, TB and TT were cloned from genomic DNA into pET28a and expressed in *Escherichia coli* BL21(DE3) with cleavable polyhistidine tags. SDS-PAGE analysis showed overexpression of the correct size constructs (AT = 87.0, AX = 88.2, JT = 88.7, TB = 90.8 and TT = 93.0 kDa) with good levels of soluble protein (Figure S4, section 7.1.1.4). The P450 concentration was measured in the corresponding cell lysates with the new panel of P450s showing up to 5.6-fold better expression and stability during the lysis process compared to RhF (TT = 315.5 and RhF = 56.1 nmol_{P450} g_{cdw}⁻¹) (Table S4, section 7.1.1.4). The new panel of enzymes were purified by IMAC (immobilized metal ion affinity chromatography) following sonication of whole cells expressing the respective P450. The purified samples were subsequently analysed by SDS- PAGE and showed purification of the full-length enzymes (Figure S5, section 7.1.1.4).

The UV-visible absorption features of the newly expressed and IMAC purified P450s were characteristic of typical P450 hemoproteins (Figure S6, section 7.1.1.4). The oxidized form of the purified proteins gave spectra with the absorption maximum of the main Soret band at 419 nm and the α and β bands located at 567 and 533 nm respectively for AT. Similarly, the main Soret band and the α and β bands were located at 422, 567 and 536 nm for AX, 420, 568 and 530 nm for JT, 420, 570 and 532 nm for TB and 420, 568 and 537 nm for TT. Reduction with sodium dithionite causes a slight shift in the main Soret band (< 420 nm) and for it to diminish in intensity. The addition of CO to the dithionite-reduced samples causes the expected shift in absorption to 449-451 nm (449, 451, 450, 449 and 449 nm for AT, AX, JT, TB and TT respectively) upon formation of an Fe(II)-CO complex, which is where the P450 name derives.

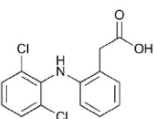
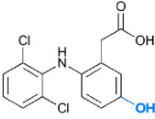
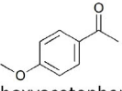
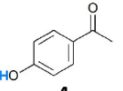
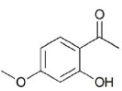
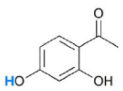
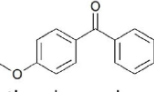
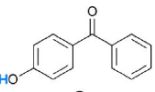
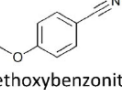
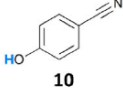
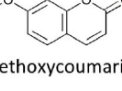
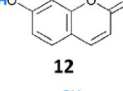
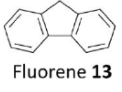
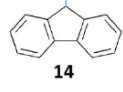







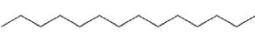
2.4.3 Substrate Scope and Nucleotide Cofactor Preference

Prior to investigation of the substrate scope the cofactor preference of the class VII P450s AT, AX, JT, TB and TT was determined by monitoring the reduction of cytochrome *c* as the electron acceptor in steady-state reactions. Both NADH and NADPH were tested as electron donors but there was a strong preference observed for NADPH over NADH in all cases (Table S6, section 7.1.1.5). This is consistent with the previously characterized Class VII enzymes CYP116B1, B2, B3 and B4 all of which also showed preference toward NADPH.^[18a, 31a, 31c, 31d] The observed k_{cat} value of 5.6 ± 0.7 , 8.3 ± 0.7 and 40.5 ± 0.8 s⁻¹ for NADPH as electron donor by TB, AX and TT is in the same range as those previously determined for CYP116B1-3 ($k_{\text{cat}} = 2.5 - 39$ s⁻¹), however AT and JT exhibit higher rates with k_{cat} values of 69 ± 2 and 70 ± 2 s⁻¹ respectively.

P450-RhF displays a notable degree of substrate promiscuity, being able to catalyse not only hydroxylation reactions but also the *O*-dealkylation, sulfoxidation and epoxidation of a diverse range of compounds.^[19] We assembled twelve such compounds creating a test set with which to explore the substrate scope of P450-AT, AX, JT, TB and TT. Reactions were run using soluble cell lysates (the equivalent of 200 mg mL⁻¹ wet cell weight) (P450 concentrations determined by CO binding are shown in Table S8, section 7.1.1.6) for 20 h with NADP⁺ and a cofactor recycling system consisting of glucose and glucose dehydrogenase (GDH) to facilitate production of the more expensive NADPH. Table 2 shows the total substrate conversions

catalysed by each enzyme including RhF. In most cases the major product could be identified by comparison with an authentic standard. The number and approximate distribution of unidentified products have also been indicated in order to provide a general fingerprint^[32] of the respective enzyme activities (Table 2). A range of activities were observed including demethylation, epoxidation, hydroxylation, and sulphoxidation. Previously P450-RhF has been shown to possess good activity towards diclofenac (**1**) and produce the human metabolite 5-hydroxydiclofenac (**2**).^[20] In this study P450-AT, AX, JT and TT all displayed better conversions (> 85 %) to the hydroxylated product than RhF (63 %) under the screening conditions. Furthermore, in terms of total substrate conversion, RhF was outperformed in all cases by two or more of the new class VII P450s from this study, with the exception of just α -isophorone (**17**) and tetradecane (**22**) where no activity was observed with any P450. Demethylation was observed for P450 biotransformations with 4-methoxyacetophenone (**3**), 2-hydroxy-4-methoxyacetophenone (**5**), 4-methoxybenzophenone (**7**), 4-methoxybenzotrile (**9**) and 7-methoxycoumarin (**11**) and the demethylated compound was the major product in most cases. Hydroxylation was observed with fluorene (**13**) to give 9-fluorenol (**14**) as the major product. In addition, benzylic C-H activation of ethylbenzene (**15**) resulted in 35 % conversion by AT, 8 % by JT and 13 % by TB, to produce (*S*)-1-phenylethanol (**16**) (90 \pm 10 % *ee*) (Table S10, section 7.1.1.11). Furthermore, the sulfoxidation capability of AT, AX, JT, TB and TT was demonstrated with methyl phenyl sulphide (**18**) to give conversions of 59, 34, 58, 58 and 21 % respectively to the corresponding sulfoxide **19** (predominantly (*S*)-methyl phenyl sulfoxide). Epoxidation activity was also demonstrated using styrene (**20**) as substrate with conversions of 10-20 % to the corresponding epoxide **21** by AT, JT and TB.

Table 2. Substrate conversions and product fingerprint from Class VII P450s

Substrate	Product	Conversions (%) ^[a]					
		RhF	AT	AX	JT	TB	TT
 Diclofenac 1	 2	63	>99	>99	>99	37	86
 4-Methoxyacetophenone 3	 4	5	>99	11	28	24	32
 2-Hydroxy-4-methoxyacetophenone 5	 6	56	>99	10	41	95	26
 4-Methoxybenzophenone 7	 8	13	>99	9	12	29	10
 4-Methoxybenzonnitrile 9	 10	7	56	6	10	12	7
 7-Methoxycoumarin 11	 12	22	95	<5	34	55	37
 Fluorene 13	 14	10	29	12	9	43	6
 Ethylbenzene 15	 16	0	35	0	8	13	0
 α -Isophorone 17	n/a	0	0	0	0	0	0
 Methyl phenyl sulphide 18	 19	8	59	34	58	58	21
 Styrene 20	 21	0	16	6	15	15	<5
 Tetradecane 22	n/a	0	0	0	0	0	0

^[a] Reactions were carried out using soluble cell lysates. Products fingerprints are indicated in the charts under the respective conversion value with blue indicating a product identified through comparison with an authentic standard and grey indicating unknown products (products that accounted for less than 10% of the final reaction are not indicated in the charts).

2.4.4 Thermal Stability of the P450 Panel

Previous literature on thermostable P450s placed the major focus on the stability of the heme domain at high temperatures. However, P450 chemistry depends also on the effective supply of reducing equivalents from redox partners, the stability of which needs to match the heme domain. As we were interested in the overall construct stability, the thermal stability of the P450s was measured in two ways, (i) determining the residual CO binding by formation of the Fe(II)-CO complex at 450 nm and (ii) by monitoring the residual activity toward 7-methoxycoumarin (**11**). The plots showing P450 concentration determined by CO binding were obtained using cell-free lysates and include P450-RhF for comparison (Figure 3A). Each of the P450s from this study displayed greater thermal stability than RhF, which was encouraging given the previous issues with stability of P450-RhF.^[20] P450 AT and JT showed modest improvements from RhF, while AX, TB and TT showed significant improvements with increases in T_{50} of greater than 15 °C. The residual activity experiments shown in Figure 3B were conducted using purified protein and demonstrate not only thermal stability but also confirm the self-sufficient nature of these P450s. RhF is not stable during routine purification and is therefore not included for comparison.^[20] Similarly to the data shown in Figure 3A, the P450s exhibiting the greatest stability determined by residual activity were TB, AX and TT. These operating temperatures are largely consistent with the growth temperatures of their native hosts (JT < AT = TB < TT < AX) (Table 1).

Values for T_{50} were extracted by fitting the data to a Hill type equation (section 7.1.1.7) and are compared in Figure 3C. In most cases there are small discrepancies between the observed T_{50} values for the residual CO binding and the residual activity, but for AT, AX, JT and TT these differences are less than 3 °C. Differences are expected due to the different experimental setup and the use of lysate for the CO binding compared to the use of purified protein with a defined buffer composition used for the residual activity. However, in the case of P450-TB (and AT to a lesser extent) the T_{50} values determined by each method differ significantly. The T_{50} determined by CO binding (TB: $T_{50} = 59.3 \pm 0.7$ °C) is greater than the T_{50} determined by residual activity (TB: $T_{50} = 53.1 \pm 0.3$ °C) by 6.2 ± 1 °C. The source of difference can likely be attributed to a less stable reductase domain for TB since the CO binding is concerned purely with the integrity of the heme domain, while the residual activity requires

function of the entire construct from electron transport through the FMN-FeS domain to substrate turnover in the heme domain.

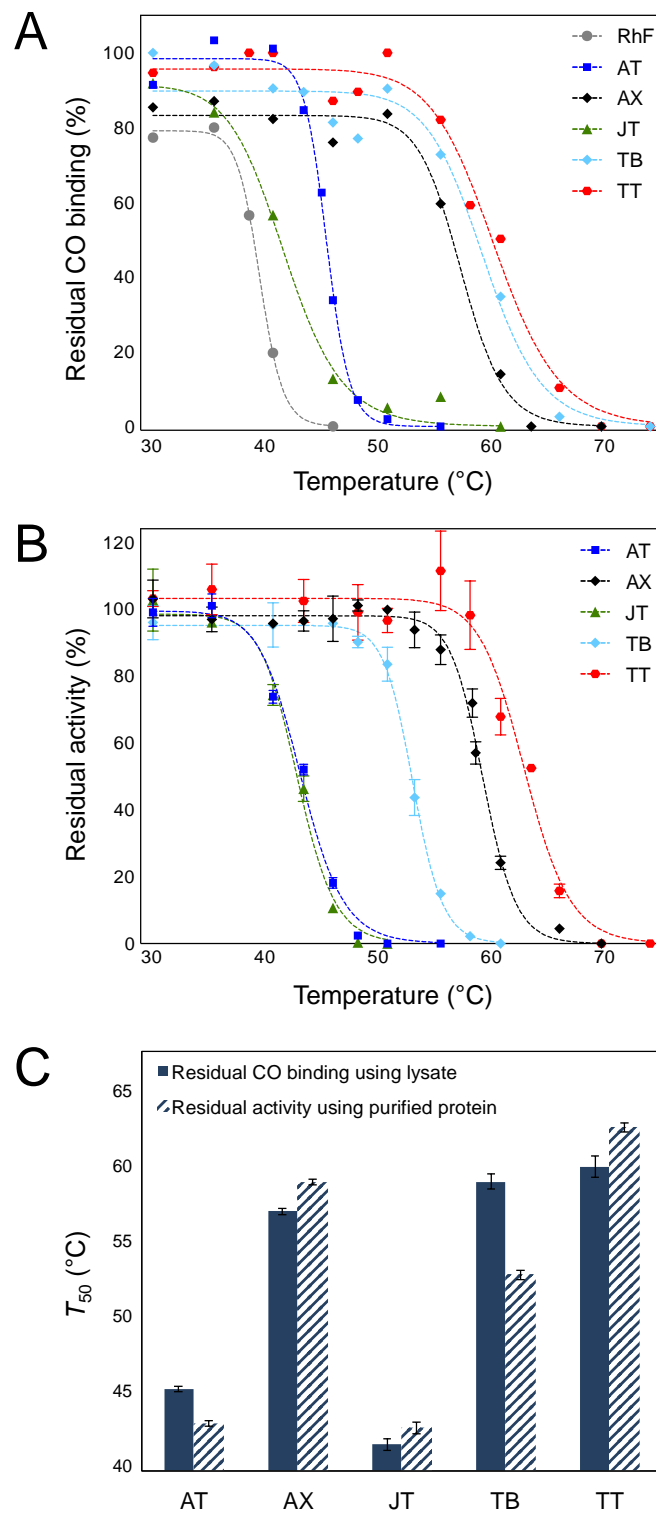


Figure 3. Thermal stability of P450s assessed by (A) the residual CO binding following incubation of soluble cell lysates at elevated temperature for 15 min and (B) the residual activity toward 7-methoxycoumarin following incubation at elevated temperature 15 min. (C) The corresponding T_{50} values determined from either method.

To ascertain the relative stability of the FMN/2Fe–2S reductase domains compared to their partner heme domains, residual electron transfer was measured using cytochrome *c* as electron acceptor following incubation at the individual T_{50} determined by CO binding (where residual activity > 50 % would suggest a more stable reductase domain compared to the corresponding heme domain and residual activity < 50 % would suggest a less stable reductase domain) (Figure 4). Electron transfer activity of the FMN/2Fe-2S phthalate dioxygenase reductase from *Pseudomonas cepacia* with cytochrome *c* is mediated by the 2Fe-2S domain.^[33] If a similar sequence of electron transfer steps NADPH→FMN→2Fe-2S→cytochrome *c* can be postulated for the phthalate dioxygenase-type reductase domain of class VII P450s, then the measured residual electron transfer activity would be a reasonable estimate of the stability of the functional reductase domain.

The residual electron transfer values were in accordance with the residual activities observed for 7-methoxycoumarin (**11**). Under these experimental conditions the AX, JT, and TT reductase domains showed equal or greater stability than their fused heme domains (similarly to the results in Figure 3C), while as expected the TB and AT reductase domains were mostly inactive following incubation. This indicates a comparatively less stable reductase than heme domain and suggests that electron transport is indeed the limiting factor influencing substrate turnover at higher temperatures. This agrees with previous reports that the reductase domain of these self-sufficient enzymes may be responsible for reduced stability of the whole construct.^[20, 26a]

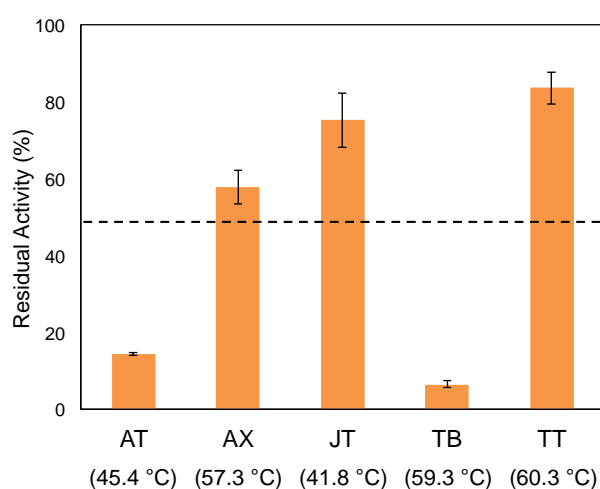


Figure 4. Residual electron transfer monitored by cytochrome *c* reduction following heat treatment at the T_{50} calculated by CO binding, where (---) shows the 50 % residual CO binding at the indicated temperature.

The P450 displaying the highest stability through either experimental method was TT. To assess not only maximum temperature stability but also the stability at moderate temperatures over longer periods, half-life assays were conducted at 50 and 58 °C (Figure S8, section 7.1.1.7). TT displayed a half-life of approximately 9.5 h with incubation at 50 °C, while the half-life was reduced to approximately 2 h at 58 °C. In addition, further studies were conducted investigating the kinetic parameters of P450-TT. Having already established the preference for NADPH and determined the kinetic properties at 25 °C (Table S6, section 7.1.1.5), we wanted to see if conducting the reaction at higher temperature yielded an increase in electron transfer rate (Table S7, section 7.1.1.7). At 50 °C using NADPH as substrate with cytochrome *c* as electron acceptor the k_{cat} value was $80 \pm 2 \text{ s}^{-1}$ which is approximately double the rate when the reaction was performed at the lower temperature of 25 °C ($k_{\text{cat}} = 40.4 \pm 0.8 \text{ s}^{-1}$).

2.4.5 Whole Cell Biotransformations of Diclofenac

Diclofenac (**1**) is a widely used non-steroidal anti-inflammatory drug that along with its 4'-hydroxy and 5-hydroxy (**2**) metabolites is persistent in the environment due to low biodegradability. The P450-RhF catalysed hydroxylation of diclofenac (**1**) in an *E. coli* whole cell system has been previously reported for the regioselective production of 5-hydroxydiclofenac (**2**).^[20] The production levels of P450-RhF were comparable to those of P450-BM3 and good product yields (0.29 g L^{-1}) were achieved using a substrate feeding strategy. However, when high concentrations of substrate were added at the beginning of the reaction in a single batch, very low conversions were observed. Here our initial screening experiments (as discussed above) suggested that the present thermostable enzymes would be significantly better biocatalysts for this reaction (Table 2). To investigate the suitability of these P450s for the production of 5-hydroxydiclofenac (**2**) whole cell biotransformations were conducted (full information regarding P450 content as determined by CO-difference spectroscopy on the corresponding cell lysates can be found in Table S8, section 7.1.1.6). Diclofenac (1, 2, 10 or 20 mM final concentrations) was added in a single addition to start the reactions, which were carried out at both 20 and 40 °C (Figure 5). As a baseline for our investigations, we observed that P450-RhF catalysed the transformation of diclofenac (**1**) to

produce 5-hydroxydiclofenac (**2**) (~ 0.5 mM) with 1 and 2 mM substrate loading at 20 °C but not at higher substrate concentrations. Furthermore, for the whole cell biotransformations conducted at 40 °C with P450-RhF (Figure 5B), there was no appreciable amount of product detected for any of the RhF catalysed reactions regardless of initial substrate concentration. When the temperature was raised to 40 °C and when using high substrate loading (10 and 20 mM), AX, JT, TB and TT all outperform RhF in terms of final product titre. P450-AX also outperforms RhF under its optimum conditions of 20 °C using 1 or 2 mM diclofenac and in fact performed significantly better than not just RhF but also AT, JT, TB and TT under all of the tested conditions. It is interesting to note that increasing the temperature of the AX catalysed production of 5-hydroxydiclofenac (**2**) from 20 to 40 °C with a 10 mM substrate loading led to a 100 % increase in product formation.

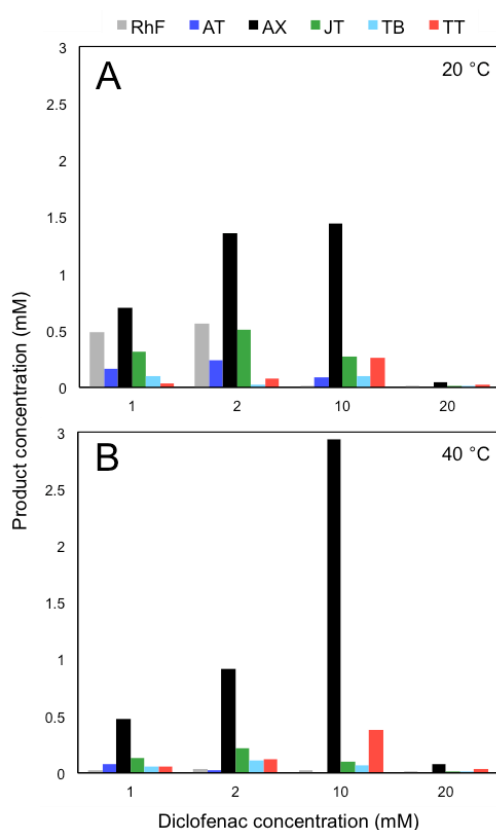


Figure 5. Whole cell biotransformations with 1, 2, 10 and 20 mM diclofenac conducted at (A) 20 °C and (B) 40 °C.

The scale-up of P450 based reactions at gram scale is still a challenge.^[34] The number of reported preparative and gram scale reactions is low,^[35] this is especially apparent in comparison to other enzyme systems. We aimed to assess scalability of the P450-AX whole

cell biotransformation of diclofenac for comparison with the previous scale-up work with P450-RhF.^[20] Accordingly, the reaction was scaled-up linearly from 0.5 to 100 mL with 10 mM diclofenac, 200 mg mL⁻¹ wet cell load (corresponding to 6.2 μM P450 final concentration) and conducted at 40 °C as before. The reaction was monitored over time (Figure S9, section 7.1.1.8) and after 5-10 h a final product concentration of 5.0 mM (1.6 g L⁻¹) was obtained, corresponding to a conversion of 50 %, with some substrate and/or product loss to the biomass. Production of the 5-hydroxydiclofenac metabolite (**2**) was confirmed by ¹H NMR spectroscopy, which gave spectral data consistent with reported values.^[20] The whole cell system showed good scalability with a greater conversion in the scaled-up biotransformation than for the initial small-scale reactions (2.9 mM). Product formation plateaued after approximately 5 h (Figure S9, section 7.1.1.8) at 50 % conversion despite the use of a thermally stable P450 (AX: $T_{50} = 59.3 \pm 0.2$ °C). This could be due to the use of whole cells rather than purified protein where mesophilic cellular enzymes are relied upon for cofactor recycling. Nevertheless, without any optimization, the AX system showed more robustness with higher substrate loading and an impressive 42-fold improvement in productivity when compared to the previously reported metrics for the P450-RhF catalysed whole cell system (Table 3),^[20] making it a more suitable system for the production of 5-hydroxydiclofenac (**2**).

Table 3. Comparison of the RhF and AX catalysed production of 5-hydroxydiclofenac

	P450-RhF ^[a]	P450-AX
Substrate loading (delivery time)	1 mM (0 h) 0.3 mM (8 h) 0.3 mM (23 h)	10 mM (0 h)
Average productivity (g L ⁻¹ h ⁻¹)	0.012	0.51
Product yield (g L ⁻¹)	0.29	1.6
Maximum specific activity (mg _{product} h ⁻¹ mg _{P450} ⁻¹)	0.07	1.18
TTN	382	800

^[a] Values determined from previously published results^[21]

Uncoupling in P450 reactions is another important parameter to consider as it influences reaction rate and results in the formation of reactive oxygen species (ROS), which can affect

the stability and performance of the whole system.^[34b] ROS are side products of P450 reactions occurring due to inefficient coupling of electrons from NADPH to formation of, in this case, the hydroxylation product. The modified human P450 CYP2C9 catalyses the oxidation of diclofenac with a coupling efficiency of ~ 30 %.^[36] Here the coupling efficiency of the AX catalysed hydroxylation of diclofenac is 51 ± 3 %. In comparison with the self-sufficient BM3 with its native substrates, where coupling is typically > 90 %, ^[36] this value is slightly modest and could reduce periods of whole cell activity.^[38] However for other P450s and with non-physiological substrates, such high coupling efficiency is rare.

2.5 Conclusions

Since our first discovery of P450-RhF,^[18b] this class of self-sufficient P450 monooxygenases has found wide applications and a number of homologues have been identified.^[31] Here, we have presented a panel of five naturally self-sufficient P450s from thermophilic microorganisms. P450-AT (CYP116B65) from *A. thermoflava*, AX (CYP116B64) from *A. xiamenense*, JT (CYP116B63) from *J. thermophila*, TB (CYP116B29) from *T. bispora* and TT (CYP116B46) from *T. thermophilus* are novel catalytically self-sufficient class VII P450 monooxygenases. The need for identification and production of corresponding thermostable redox partners and the commonly encountered issues with stability of P450-PFOR constructs, such as that of P450-RhF, have been successfully overcome by using these self-sufficient thermostable enzymes. AT, AX, JT, TB and TT can be expressed reliably in a recombinant *E. coli* system and have shown considerable thermal stability in comparison with P450-RhF. In particular, P450-AX and TT showed high levels of soluble expression and excellent robustness with T_{50} values for residual activity of 59.3 ± 0.2 and 62.9 ± 0.3 °C respectively. TT maintained activity at 50 °C with a half-life of over 9 h and also displayed a significantly (100 %) enhanced electron transfer rate at higher temperatures. Furthermore, assessment of thermal stability has highlighted that stability of the corresponding heme and reductase domains has not necessarily evolved equivalently. The reductase domains of P450-TB and AT may limit activity of the overall fusion enzyme at high temperatures, however JT, AX and TT show excellent stability over the entire multi-domain construct. Activity profiling has shown these new class VII P450s to have different substrate fingerprints and to be capable of catalysing oxidation reactions, including aromatic hydroxylation, demethylation, epoxidation and sulfoxidation. In

addition, whole cells expressing P450-AX have been used for the biotransformation of diclofenac to the human metabolite 5-hydroxydiclofenac at preparative scale. The stability, reproducibility and higher product yields compared to the previously reported RhF system make P450-AX an ideal candidate for the regioselective production of 5-hydroxydiclofenac (**2**), which is required for toxicology studies and as a standard for environmental analysis and is difficult to access chemically. This panel of robust P450s with differing substrate scope and product profiles is an important addition to the toolbox of biocatalysts available for biooxidations using molecular oxygen.

2.6 Subsequent Work

In order to assess the oxygen demand of the scaled-up synthesis of 5-hydroxydiclofenac (5-OH DCF), batch reactions were carried out in small-scale stirred tank reactors using a whole-cell biocatalyst expressing P450-AX. Each reactor contained a gas sparger, two Rushton turbines, as well as mechanisms for temperature and pH control. The reaction volume was 100 mL and the stirring speed and aeration rate were varied.

As outlined in section 1.3, oxygen limitation may arise from a number of phenomena when using a P450-based whole-cell biocatalyst (limiting OTR, cell maintenance and uncoupling). Firstly, the oxygen consumption rate (OCR) of cells without the target substrate was determined (Figure 6). Linear extrapolation to higher cell concentrations shows that the OCR at a cell concentration of 123 g_{cww}/L (the designated cell concentration for subsequent scale-up experiments) was estimated to be approximately 80 mmol O₂ L⁻¹ h⁻¹ at 37°C (the designated process temperature for subsequent scale-up experiments). In the employed experimental setup, the maximum oxygen transfer rate that can be achieved is approximately 21 mmol O₂ L⁻¹ h⁻¹ sparging air and 103 mmol O₂ L⁻¹ h⁻¹ sparging pure oxygen (at atmospheric pressure and 37 °C). From these estimations, the reaction will always be limited by the oxygen transfer rate when using air, as the OCR of cells without the substrate is higher than the achievable OTR obtained when sparging air.

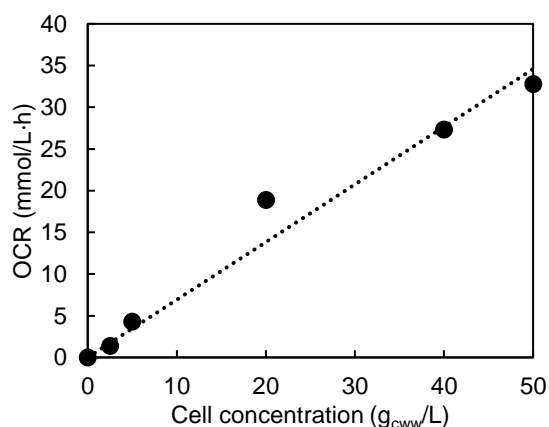


Figure 6. Oxygen consumption rate (OCR) of cells when sparging pure oxygen at atmospheric pressure. The diffusion of oxygen from the liquid to the air was experimentally determined and subtracted to all data points. Higher cell concentrations were not tested due to the sensitivity of the oxygen probe used.

Next, batch reactions were carried out sparging mixtures of oxygen and nitrogen at various ratios to evaluate whether the reaction was oxygen limited. In this work, each ratio is represented by the concentration of oxygen in the gas feed (%). For the first batch reaction, air was supplied (~0.216 mM O₂ solubility at 1 atm and 37 °C) and in the second batch pure oxygen (~1.03 mM O₂ solubility at 1 atm and 37 °C) was sparged into the bioreactor. From the results obtained (Figure 7), it can be concluded that the reaction is oxygen limited when air is used for sparging, since the use of pure oxygen resulted in a faster reaction rate.

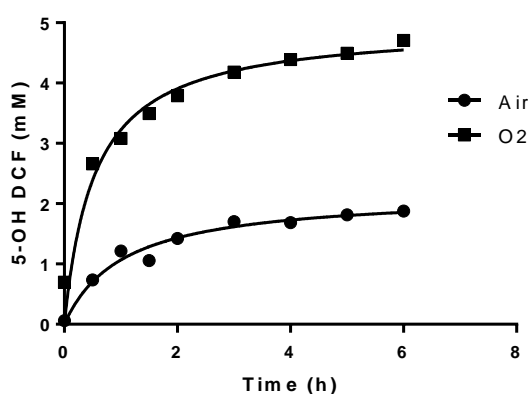


Figure 7. Oxidation of 5 mM of diclofenac sparging air (aeration rate 0.4 vvm) and pure oxygen (aeration rate 0.2 vvm). Reaction set-up: 0.2 M KPi buffer pH 8, 123 mg/mL cells wet weight, 10 mg/mL glucose, stirring speed of 200 rpm, 20 μ L Antifoam 204, 37°C and 100 mL reaction volume.

Table 4 shows the process metrics calculated from the results of these batch reactions. Based on the productivity, the oxygen consumption for product formation (when sparging pure

oxygen) was estimated to be approximately 2 mmol O₂ L⁻¹ h⁻¹. Additionally, since P450-AX has a coupling efficiency of approximately 50 % (section 2.4.5), it can be assumed that a further 2 mmol O₂ L⁻¹ h⁻¹ was converted to ROS. Recalling that the OCR was estimated to be approximately 80 mmol O₂ L⁻¹ h⁻¹, it would appear that the majority of oxygen provided to the reactor is employed for cell maintenance. When pure oxygen was supplied to the reaction, both the space-time yield (STY) and biocatalyst yield were improved more than 2-fold, indicating that the enzyme is not used efficiently when sparging air. At constant agitation and aeration rates, oxygen sparging results in a 5-fold increase in oxygen transfer rate, compared to air sparging. In this case, the reaction rate increased, but to a lower extent. Therefore, it appears that the reaction is no longer oxygen limited when sparging with pure oxygen, as a 5-fold increase in reaction rate was not observed. If the reaction rate had increased with the same factor, it would show that the reaction kinetics are dependent on oxygen concentration.

Table 4. Comparison between the calculated process metrics obtained from experiments carried out with two different oxygen supply methods.

Process Metric	Gas feed	
	Air	100% O ₂
Reaction conversion (%)	38	94
Product concentration (g/L)	0.6	1.5
STY (g L ⁻¹ h ⁻¹)	0.24	0.62
Biocatalyst yield (g _{product} /g _{cww})	0.0048	0.012

It is known that the relative concentration of oxygen in the gas feed can affect the lifetime of the whole-cell biocatalyst.^[38] Our results suggest that productivity might not be compromised if a lower oxygen concentration is used, which is not only desirable for instability issues but also for economic and safety reasons. To test this hypothesis, initial rate experiments were carried out at different oxygen concentrations, as shown in Figure 8. The reaction rate was found to increase as the oxygen concentration in the feed gas was increased to 75%, followed by a drop in reaction rate at 100% oxygen concentration. Therefore, a lower concentration of oxygen in the gas feed should be employed to maximize the performance of the biocatalyst.

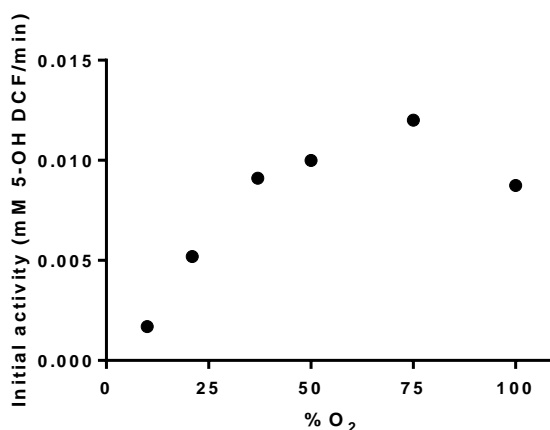


Figure 8. Effect of oxygen concentration on initial reaction rates. Different oxygen concentrations were obtained by mixing nitrogen and oxygen accordingly to the values in the graph. Solubility of oxygen at 37 °C and atmospheric pressure: 10% - 0.103 mM; 21% - 0.216 mM; 37% - 0.381 mM; 50% - 0.515 mM; 75 % - 0.773 mM; 100% - 1.03 mM.

In summary, this follow-up study on the preparative scale synthesis of 5-OH DCF by whole-cell expressing P450-AX showed that biocatalyst-related limitations and process parameters still need to be optimized to reach practical values for the relevant economic metrics. From the biological perspective, improvements in protein expression levels and coupling efficiency are needed for the efficient utilization of reducing equivalents and oxygen at low cell concentration. In terms of process engineering, oxygen supply optimization could further improve the whole-cell hydroxylation of diclofenac.

2.7 References

- [1] a) R. Bernhardt, *J. Biotechnol.* 2006, 124, 128-145; b) D. R. Nelson, *Human genomics* 2009, 4, 59-65.
- [2] a) S. T. Jung, R. Lauchli, F. H. Arnold, *Curr. Opin. Biotechnol.* 2011, 22, 809-817; b) R. Fasan, *ACS Catalysis* 2012, 2, 647-666.
- [3] F. Hannemann, A. Bichet, K. M. Ewen, R. Bernhardt, *Biochim. Biophys. Acta, Gen. Subjects* 2007, 1770, 330-344.
- [4] a) F. P. Guengerich, *J. Biochem. Mol. Toxicol.* 2007, 21, 163-168; b) F. P. Guengerich, A. W. Munro, *J. Biol. Chem.* 2013, 288, 17065-17073; c) E. M. Isin, F. P. Guengerich, *Biochim. Biophys. Acta* 2007, 1770, 314-329.
- [5] a) E. M. Carreira, N. J. Turner, M. Höning, P. Sondermann, *Angew. Chem. Int. Ed.* 2017; b) B. M. Nestl, S. C. Hammer, B. A. Nebel, B. Hauer, *Angew. Chem.* 2014, 53; c) D. J. Pollard, J. M. Woodley, *Trends Biotechnol.* 2007, 25, 66-73; d) J. L. Porter, R. A. Rusli, D. L. Ollis, *ChemBioChem* 2016, 17, 197-203; e) N. J. Turner, E. O'Reilly, *Nature Chemical Biology* 2013, 9, 285-288.
- [6] H. M. Girvan, A. W. Munro, *Curr. Opin. Chem. Biol.* 2016, 31, 136-145.
- [7] a) R. Bernhardt, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* 2014, 98; b) E. O'Reilly, V.

- Kohler, S. L. Flitsch, N. J. Turner, *Chem. Commun.* 2011, 47, 2490-2501.
- [8] S. Goodwin, J. D. McPherson, W. R. McCombie, *Nat. Rev. Genet.* 2016, 17, 333-351.
- [9] B. Van den Burg, *Curr. Opin. Microbiol.* 2003, 6, 213-218.
- [10] a) J. A. Littlechild, *Frontiers in bioengineering and biotechnology* 2015, 3, 161-161; b) P. Turner, G. Mamo, E. N. Karlsson, *Microb. Cell Fact.* 2007, 6, 9-9; c) C. Vieille, G. J. Zeikus, C. Vieille, *Microbiol. Mol. Biol. Rev.* 2001, 65, 1-43.
- [11] a) J. D. Bloom, S. T. Labthavikul, C. R. Otey, F. H. Arnold, *PNAS* 2006, 103, 5869-5874; b) N. Tokuriki, D. S. Tawfik, *Curr. Opin. Struct. Biol.* 2009, 19, 596-604; c) N. Tokuriki, D. S. Tawfik, *Science* 2009, 203-208.
- [12] a) L. S. Koo, R. A. Tschirret-Guth, W. E. Straub, P. Moënne-Loccoz, T. M. Loehr, P. R. Ortiz De Montellano, *J. Biol. Chem.* 2000, 275, 14112-14123; b) S.-Y. Park, K. Yamane, S.-i. Adachi, Y. Shiro, K. E. Weiss, S. a. Maves, S. G. Sligar, *J. Inorg. Biochem.* 2002, 91, 491-501; c) A. V. Puchkaev, P. R. Ortiz De Montellano, *Arch. Biochem. Biophys.* 2005, 434, 169-177.
- [13] a) Y. Oku, A. Ohtaki, S. Kamitori, N. Nakamura, M. Yohda, H. Ohno, Y. Kawarabayasi, *J. Inorg. Biochem.* 2004, 98, 1194-1199; b) A. V. Puchkaev, T. Wakagi, P. R. Ortiz de Montellano, *JACS* 2002, 124, 12682-12683.
- [14] A. Schallmeyer, G. Den Besten, I. G. P. Teune, R. F. Kembaren, D. B. Janssen, *Appl. Microbiol. Biotechnol.* 2011, 89, 1475-1485.
- [15] a) R. K. Behera, S. Mazumdar, *Int. J. Biol. Macromol.* 2010, 46, 412-418; b) J. K. Yano, F. Blasco, H. Li, R. D. Schmid, A. Henne, T. L. Poulos, *J. Biol. Chem.* 2003, 278, 608-616.
- [16] T. Mandai, S. Fujiwara, S. Imaoka, *FEBS J.* 2009, 276, 2416-2429.
- [17] A. W. Munro, D. G. Leys, K. J. McLean, K. R. Marshall, T. W. B. Ost, S. Daff, C. S. Miles, S. K. Chapman, D. A. Lysek, C. C. Moser, C. C. Page, P. L. Dutton, *Trends Biochem. Sci* 2002, 27, 250-257.
- [18] a) G. A. Roberts, A. Celik, D. J. B. Hunter, T. W. B. Ost, J. H. White, S. K. Chapman, N. J. Turner, S. L. Flitsch, *J. Biol. Chem.* 2003, 278, 48914-48920; b) G. A. Roberts, G. Grogan, A. Greter, S. L. Flitsch, N. J. Turner, *J. Bacteriol.* 2002, 184, 3898-3908.
- [19] a) A. Celik, G. A. Roberts, J. H. White, S. K. Chapman, N. J. Turner, S. L. Flitsch, *Chem. Commun.* 2006, 4492-4494; b) E. O'Reilly, M. Corbett, S. Hussain, P. P. Kelly, D. Richardson, S. L. Flitsch, N. J. Turner, *Catal. Sci. Technol.* 2013, 3, 1490-1492.
- [20] J. M. Klenk, B. A. Nebel, J. L. Porter, J. K. Kulig, S. A. Hussain, S. M. Richter, M. Tavanti, N. J. Turner, M. A. Hayes, B. Hauer, S. L. Flitsch, *Biotechnol. J.* 2017, 12, 1600520.
- [21] D. J. B. Hunter, G. A. Roberts, T. W. B. Ost, J. H. White, S. Mueller, N. J. Turner, S. L. Flitsch, S. K. Chapman, *FEBS Lett.* 2005, 579, 2215-2220.
- [22] a) A. Eichler, Ł. Gricman, S. Herter, P. P. Kelly, N. J. Turner, J. Pleiss, S. L. Flitsch, *ChemBioChem* 2016, 17, 426-432; b) P. P. Kelly, A. Eichler, S. Herter, D. C. Kranz, N. J. Turner, S. L. Flitsch, *Beilstein J. of Org. Chem.* 2015, 11, 1713-1720; c) J. K. Kulig, C. Spandolf, R. Hyde, A. C. Ruzzini, L. D. Eltis, G. Grönberg, M. A. Hayes, G. Grogan, *Bioorg. Med. Chem.* 2015, 23, 5603-5609.
- [23] K. J. McLean, M. Hans, B. Meijrink, W. B. van Scheppingen, A. Vollebregt, K. L. Tee, J.-M. van der Laan, D. Leys, A. W. Munro, M. A. van den Berg, *PNAS* 2015, 112, 2847-2852.
- [24] G. J. Baker, H. M. Girvan, S. Matthews, K. J. McLean, M. Golovanova, T. N. Waltham, S. E. J. Rigby, D. R. Nelson, R. T. Blankley, A. W. Munro, *ACS Omega* 2017, 2, 4705-4724.
- [25] A. W. Munro, J. G. Lindsay, J. R. Coggins, S. M. Kelly, N. C. Price, *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* 1996, 1296, 127-137.
- [26] a) S. Eiben, H. Bartelmäs, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* 2007, 75, 1055-1061; b) Y. Li, D. A. Drummond, A. M. Sawayama, C. D. Snow, J. D. Bloom, F. H. Arnold, *Nat. Biotechnol.* 2007, 25, 1051-1056.
- [27] T. Mandai, S. Fujiwara, S. Imaoka, *Biochem. Biophys. Res. Commun.* 2009, 384, 61-65.

- [28] K. L. Harris, R. E. S. Thomson, S. J. Strohmaier, Y. Gumulya, E. M. J. Gillam, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 2017, 1866, 97-115.
- [29] C. R. Nishida, P. R. Ortiz de Montellano, *Biochem. Biophys. Res. Commun.* 2005, 338, 437-445.
- [30] R.-J. Li, J.-H. Xu, Y. Yuecai, N. Wirth, J. Ren, B.-B. Zeng, H. Yu, *New J. Chem.* 2016, 40, 8928-8934.
- [31] a) L. Liu, R. D. Schmid, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* 2006, 72, 876-882; b) D. Minerdi, S. J. Sadeghi, G. Di Nardo, F. Rua, S. Castrignanò, P. Allegra, G. Gilardi, *Mol. Microbiol.* 2015, 95, 539-554; c) A. J. Warman, J. W. Robinson, D. Luciakova, A. D. Lawrence, K. R. Marshall, M. J. Warren, M. R. Cheesman, S. E. J. Rigby, A. W. Munro, K. J. McLean, *FEBS J.* 2012, 279, 1675-1693; d) Y.-C. Yin, H.-L. Yu, Z.-J. Luan, R.-J. Li, P.-F. Ouyang, J. Liu, J.-H. Xu, *ChemBioChem* 2014, 15, 2443-2449.
- [32] a) C. von Bühler, P. Le-Huu, V. B. Urlacher, *ChemBioChem* 2013, 14, 2189-2198; b) K. Zhang, S. El Damaty, R. Fasan, *JACS* 2011, 133, 3242-3245.
- [33] G. T. Gassner, D. P. Ballou, *Biochemistry* 1995, 34, 13460-13471. [34] a) J. Brummund, M. Müller, T. Schmitges, I. Kaluzna, D. Mink, L. Hilterhaus, A. Liese, *J. Biotechnol.* 2016, 233, 143-150; b) M. T. Lundemo, J. M. Woodley, *Appl. Microbiol. Biotechnol.* 2015, 99.
- [35] a) I. Kaluzna, T. Schmitges, H. Straatman, D. van Tegelen, M. Müller, M. Schürmann, D. Mink, *Org. Process Res. Dev.* 2016, 20, 814-819; b) K. Zhang, B. M. Shafer, M. D. Demars, 2nd, H. A. Stern, R. Fasan, *J. Am. Chem. Soc.* 2012, 134, 18695-18704; c) V. Tyagi, H. Alwaseem, K. M. O'Dwyer, J. Ponder, Q. Y. Li, C. T. Jordan, R. Fasan, *Biorg. Med. Chem.* 2016, 24, 3876-3886; d) P. Bracco, D. B. Janssen, A. Schallmeyer, *Microbial Cell Factories* 2013, 12, 95-95.
- [36] V. R. Dodhia, A. Fantuzzi, G. Gilardi, *J. Biol. Inorg. Chem.* 2006, 11, 903-916.
- [37] a) M. J. Cryle, R. D. Espinoza, S. J. Smith, N. J. Matovic, J. J. De Voss, *Chem Commun (Camb)* 2006, 2353-2355; b) M. A. Noble, C. S. Miles, S. K. Chapman, D. A. Lysek, A. C. MacKay, G. A. Reid, R. P. Hanzlik, A. W. Munro, *Biochem. J* 1999, 339, 371-379.
- [38] R. Fasan, M. M. Chen, N. C. Crook, F. H. Arnold, *Angew. Chem.* 2007, 119, 8566 – 8570.

Chapter 3. The Crystal Structure of P450-TT Heme-Domain Provides the First Structural Insights into The Versatile Class VII P450s

Michele Tavanti,^[a] Joanne L. Porter,^[a] Colin W. Levy,^[a] J. Rubén Gómez Castellanos,^[b] Sabine L. Flitsch^[a] and Nicholas J. Turner^[a]

Published in Biochemical and Biophysical Research Communications, **2018**, 501, 846-850.

3.1 Acknowledgements

The research presented in this chapter was a collaborative effort between the doctoral candidate, Dr. Joanne L. Porter, Dr. Colin W. Levy and Dr. J. Rubén Gómez Castellanos, under the supervision of Professor Nicholas J. Turner and Professor Sabine L. Flitsch. The doctoral candidate, Dr. Joanne L. Porter and Dr. Colin W. Levy planned the experiments. The doctoral candidate carried out expression trials, protein purification and molecular biology experiments. Dr. Colin W. Levy assisted the doctoral candidate in the set up of crystallization trials, collected synchrotron X-ray data and determined the crystal structure. The doctoral candidate analysed the crystal structure and wrote the manuscript. All the authors reviewed the text before submission.

3.2 Abstract

The first crystal structure of a class VII P450, CYP116B46 from *Tepidiphilus thermophilus*, has been solved at 1.9 Å resolution. The structure reveals overall conservation of the P450-fold and a water conduit around the I-helix. Active site residues have been identified and sequence comparisons have been made with other class VII enzymes. A structure similarity search demonstrated that the P450-TT structure is similar to enzymes capable of oxy-functionalization of fatty acids, terpenes, macrolides, steroids and statins. The insight gained from solving this structure will provide a guideline for future engineering and modelling studies on this catalytically promiscuous class of enzymes.

^a Manchester Institute of Biotechnology (MIB), School of Chemistry, The University of Manchester, 131 Princess Street, M1 7DN, Manchester, United Kingdom.

^b Department of Biology and Biotechnology "Lazzaro Spallanzani", University of Pavia, Via Ferrata 9, 27100, Pavia, Italy.

3.3 Introduction

Cytochrome P450 monooxygenases (P450s or CYPs) are a widespread class of heme-containing enzymes capable of catalysing a diverse array of oxidation reactions with important roles in biosynthetic pathways and xenobiotic metabolism.^[1] Since these reactions often occur with high regio- and stereoselectivity, P450s have recently been targeted for synthetic applications.^[2] The most common and synthetically attractive reaction catalysed by these enzymes is the monooxygenation of unactivated C-H bonds under mild conditions using molecular oxygen as the sole oxidant. In order to carry out this reaction, two NAD(P)H-derived electrons are shuttled by redox partners for reductive activation of dioxygen. Two protonation steps are necessary to generate the ferryl compound I which represents the catalyst for oxygen insertion.^[3]

Cytochrome P450s can be classified according to the electron transport system exploited to complete their catalytic cycle.^[4] Most of these P450 systems require separate redox proteins to function. However, P450 classes have been discovered in which the redox partners are fused to the heme domain of the enzyme, creating multicomponent, self-sufficient systems. These are attractive biocatalysts, as the need for identification and expression of separate redox partners is obviated.

Structural data available for P450 enzymes led to the conclusion that an overall-P450 fold is conserved and large conformational rearrangements are sometimes observed upon ligand binding.^[5] Even though structural-based standard numbering systems have certainly improved the quality of P450 sequence alignments,^[6] the process of accurately converting sequence information into a 3D structure is complicated by the low sequence identity among P450s belonging to different families (10-30 %) and by the diversity within the active site.^[7]

Class VII P450s represent a unique class of fused enzymes in which electrons are transferred by a phthalate family oxygenase-like reductase domain (PFOR)- containing a FMN and a 2Fe–2S cluster.^[8] These self-sufficient enzymes have recently received attention for their remarkable substrate promiscuity, with further efforts made to engineer them through directed evolution for higher dealkylation activity, to improve the catalytic performance in the hydroxylation of tetralin derivatives or to carry out anti-Markovnikov oxidation of

styrenes.^[9–11] Moreover, benefitting from the rapid expansion of available CYP sequences, we have recently reported the discovery of a panel of new class VII P450s with desirable biocatalytic properties such as good expression levels, thermal stability and diverse substrate scope.^[12,13] In particular, CYP116B46 from *T. thermophilus* displayed not only better expression levels but also enhanced thermal stability when compared to the previously characterized CYP116B2 (P450-RhF).

Since the discovery of the first class VII P450-RhF in 2002, several more have been reported and thoroughly characterized biochemically.^[14–19] However, as yet no physiological role has been determined for these enzymes with the exception of only CYP116B5.^[18] Moreover, X-ray crystal structures for the heme domain of these proteins are not available to date. In order to provide a link between the sequence and the 3D structure of this valuable class of enzymes, we have solved the first X-ray structure of the heme domain of a class VII P450. As thermally stable enzymes tend to generate protein crystals better suited for X-ray diffraction,^[20] we focused on CYP116B46 (UniProt A0A0K6ITW2) to get a first insight into the structure of this class of self-sufficient enzymes.

3.4 Results and Discussion

3.4.1 Overall Structure

The substrate-free crystal structure of CYP116B46 (P450-TT) was determined at 1.9 Å resolution from residue 28 to 443 by molecular replacement using the recently solved structure of CYP288A2 (PDB 5GWE) as search model.^[21]

Overall, P450-TT adopts the triangular P450 fold composed of 19 α -helices and 10 β -strands divided into 3 sheets embedding the heme (Figure 1A, the structural nomenclature of Hasemann *et al.* was adopted where possible and secondary structural elements annotated with ESPrint).^[22] Close inspection of the protein surface does not reveal an open tunnel leading to the heme prosthetic group. Analysis of the structure using the CAVER 3.0 tool enabled identification of a putative tunnel ~ 14 Å long, with a ~ 9 Å wide entrance and a bottleneck of ~ 1.7 Å (Figure 1B). This tunnel is mainly lined by hydrophobic residues (Table S2, section 7.2.2) but its opening appears too small to accommodate organic molecules (for comparison, methane radius is 1.9 Å).^[23] Unsurprisingly, initial ligand docking experiments into P450-TT structure did not place the tested substrate (diclofenac) into the active site.

However, it was possible to obtain docking poses compatible with P450 chemistry after molecular dynamics simulations (data not shown). Therefore, opening motions might be needed to accommodate the diverse range of substrates accepted by this class VII P450.^[12]

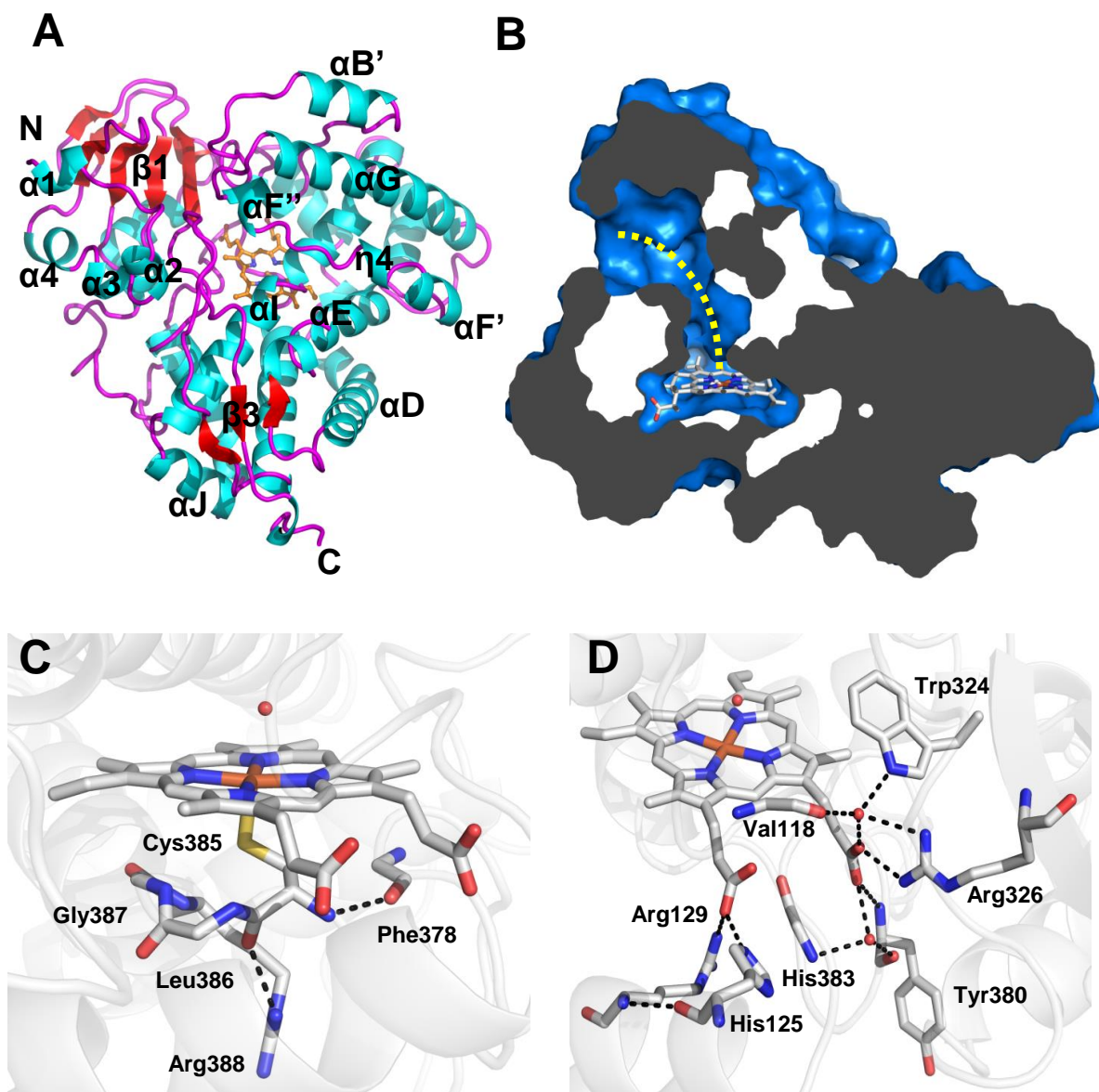


Figure 1. A) P450-TT overall fold with helices represented in cyan, β -strands in red and coils in magenta. B) Surface representation of P450-TT crystal structure showing a narrow tunnel (yellow dotted line) leading to the heme pocket. C) The cysteine-pocket with the proximal ligand Cys385 interacting with the terminal guanidino nitrogen of Arg388, and the carbonyl oxygen of Phe378. D) Heme-propionate interactions with surrounding residues. The buried prosthetic group is represented with sticks.

The heme-iron is hexacoordinated, with Cys385 acting as the proximal ligand and water as the sixth ligand (Figure 1C). The proximal ligand is part of the so-called “cysteine-pocket”, which includes Phe378, Gly379, Tyr380, Gly381, Gln384, Leu386, Gly387 and Arg388. The K-helix bears the sequence ³⁰⁷IPXAXEECLR³¹⁶ in which we can recognize the well described EXXR motif followed by the conserved Arg369 in the so-called “meander” (the loop between K'-helix and the cysteine-pocket).^[22]

The heme propionates are either coordinated to the protein chain directly through charged residues or indirectly through water, allowing the insertion of the heme cofactor in the interior of the molecule (Figure 1D). In particular, the A-ring propionate interacts with N^ε of His125, and the Arg129 side chain, while the D-ring propionate hydrogen bonds with the main chain amide group of Tyr380, the carbonyl group of the same residue and the main chain amide group of His383 through a water molecule, the side chain of Arg326 and finally with Val118, Trp324 and Arg326 through a second water molecule.

As in other P450 structures, the region of the I-helix between Ala274 and Thr279 shows a disruption of the regular α -helix hydrogen-bonding pattern (Figure S3, section 7.2.2).^[24] Moreover an extensive chain of water molecules hydrogen bonding with Val273, Ala274, Ala275 (on the “back” of the I-helix) and Glu277 can be observed, the latter being the farthest from the heme and closest in distance to the bulk solvent. A similar water channel connecting the active site to the bulk solvent has been observed in other P450 structures.^[25]

3.4.2 Active Site

The X-ray structure of P450-TT reveals that the active site is buried within the protein core, with 18 residues organized in what could be described as a tiered system above the heme (Figure 2).^[26] The six residues of tier 1 (Val118, Ala274, Thr278, Pro320, Val321 and Trp324) are located just above the heme, with Ala274 and Thr278 involved in the I-helix distortion and Trp324 interacting with the D-ring propionate. Four residues form tier 2 (Val91, Thr116, Ala270 and Phe424) and lie in a ring above the tier 1 residues. Ala270 is the only residue located in the I-helix, while the other residues come from neighboring loops. Leu92, Phe206, Val273 and Glu277 comprise tier 3, which includes the conserved acidic residue (Glu277). Finally, tier 4 is entirely composed by residues from the F''-helix (His201, Thr202, Asn204,

Ala205) and delineate a lid over the active site. Overall, the cone-shaped active site is dominated by apolar and aromatic residues, with charged residues only populating the upper tiers.

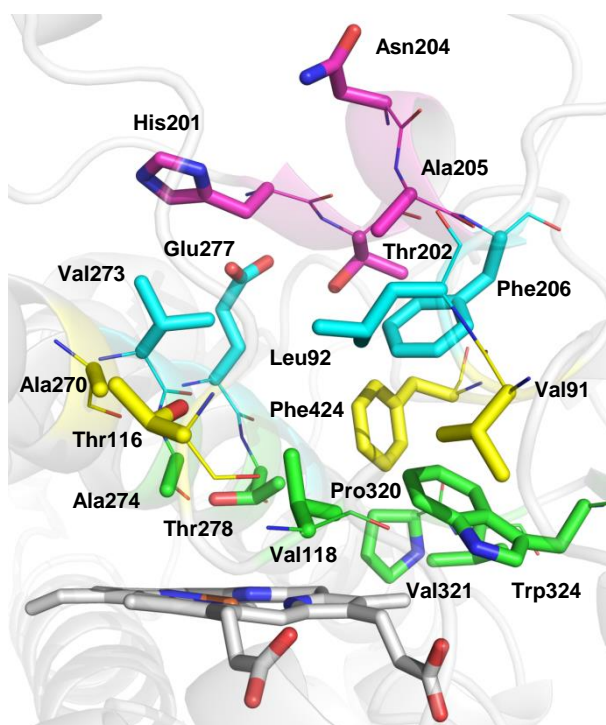


Figure 2. The active site of P450-TT. Four tiers of residues are shown with different colors: green (tier 1), yellow (tier 2), cyan (tier 3) and magenta (tier 4). The heme prosthetic group is also shown.

Guided by the insights gained from the first class VII P450 crystal structure, we proceeded to analyze local sequence similarities with other class VII P450s. The full heme-domain alignment of characterized class VII P450s can be found in section 7.2.2 (Figure S4). Active site residues and substrate channels are crucial factors determining P450 substrate-selectivity, identification of these residues is important for the purposes of drug design and for tailoring enzymes for biocatalysis. In order to carry out a sequence conservation analysis at putative substrate recognition sites, 96 class VII sequences (subfamily K9GPR5, including uncharacterized proteins) were analyzed systematically using the 3DM™ system. This type of analysis is particularly useful when consensus mutagenesis is employed for protein engineering.^[27,28]

As shown in Table S3 (section 7.2.2), amino acids found in the P450-TT active site are well conserved throughout class VII P450s (occurrence higher than 87 %), the exception being Val321 (70 % conserved), which is often substituted by isoleucine. The same systematic analysis was carried out to evaluate the conservation of residues lining the putative substrate access channel. Most of these residues are conserved throughout class VII P450s (Table S2, section 7.2.2). Notable exceptions are observed between helix-B' and helix-C (Phe113, Asn114) and in the I-helix (Ile271 and Ile272).

At this stage, the absence of general trends in active site diversity does not allow to draw general conclusions on substrate scope and product distributions for these P450s.

3.4.3 Structural Comparison to Other Bacterial P450s

A DALI structural similarity search was carried out in order to gain more functional insights based on previously solved P450 crystal structures.^[29] As shown in Table S4 (section 7.2.2), P450-TT showed the highest structural similarity to P450-CreJ (PDB 5XJN, root-mean square deviation (RMSD) = 1.9 Å), an enzyme involved in the *p*-cresol biodegradation pathway *via* phosphorylated intermediates in *Corynebacterium glutamicum*.^[21] As discussed in section 3.3, the P450 fold is conserved despite low degrees of sequence identity (30% sequence identity) between CYP families (Figure S5, A, section 7.2.2). The structural superposition is poor at both N- and C-termini and most significant deviations in the structural alignment can be observed for α B' and neighbouring loops and the whole region from α F' to the N-terminal region of α G (Figure S5, B). These regions are generally recognized as substrate recognition sites in P450s.^[30] Moreover, critical residues for substrate binding in P450-CreJ are not found in the corresponding position of P450-TT structure (Figure S6, section 7.2.2).

Similarly, the superposition of P450-TT and P450cam structures (PDB 2CPP, RMSD = 2.6 Å)^[31] reveals a similar core structure and structural variability for regions involved in substrate recognition (Figure S7, section 7.2.2). For example, the B'-helix, which houses P450cam residues Tyr96 and Phe87 involved in substrate access and binding,^[32] is significantly shifted away from the heme in P450-TT structure. Consequently, residues located in α B' are not found close to the heme. Moreover, the G-helix in P450-TT structure is longer than the one observed in P450cam and contributes to the shortening of the F-G loop.

Finally, additional similar structures include P450s acting on fatty acids, bulky substrates, such as vitamin D3, steroids, terpenes, statins, cholesterol and macrolides (see Table S4, section 7.2.2).

In summary, crystallization and structure determination of the CYP116B46 (P450-TT) heme domain revealed that the features associated with the P450-fold are well conserved, including the cysteine-pocket, propionate coordination and the I-helix hydrogen-bonding pattern disruption. A comparative analysis of the P450-TT structure demonstrated structural similarity to P450s involved in the oxidation of fatty acids and bulky substrates such as steroids, terpenes, statins, cholesterol and macrolides. As such, future investigations into substrate scope of class VII P450s might be extended to bulkier substrates, such as those accepted by human P450s. Eventually, this could reveal that the versatile framework of these biocatalysts can be exploited to generate enzymes with human-like activities. To conclude, the crystal structure of CYP116B46 provides the first glimpse into the structure of class VII P450s and provides a framework to inform engineering studies and improve our ability to build 3D models for this class of enzymes.

3.5 Subsequent Work

The resolution of the first class VII P450 structure provided us with a framework for the engineering P450-AX for diclofenac hydroxylation (Chapter 2). A model of P450-AX heme-domain structure was generated from P450-TT structure (PDB ID: 6GII) as template using SWISS-MODEL^[33] (Figure 3). Residues predicted to be in the active site were targeted to produce 14 saturation mutagenesis libraries using NNK codon degeneracy (Table S5, section 7.2.2). In order not to impair monooxygenase catalysis, residues likely involved in oxygen activation chemistry were not considered (Ala263, Thr267 and Glu266).

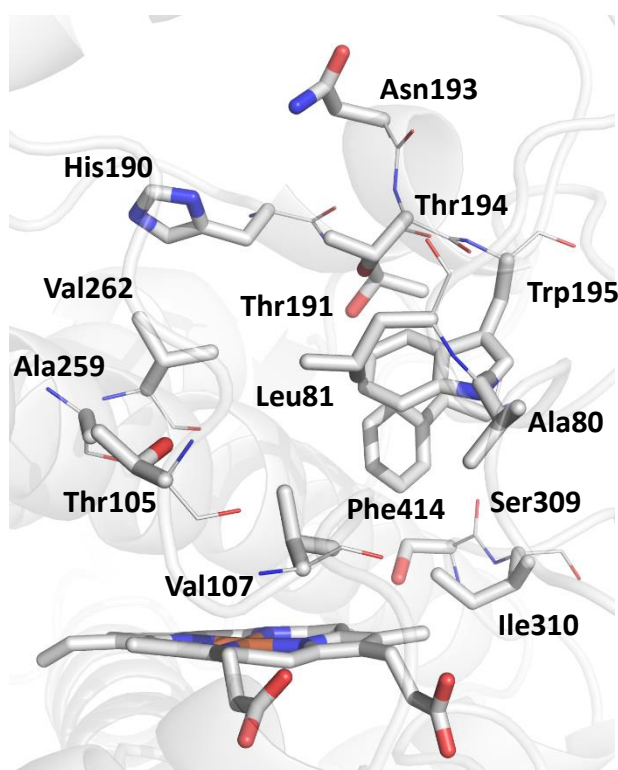


Figure 3. P450-AX active site model generated with SWISS-MODEL. Residues targeted for site-saturation mutagenesis are indicated. The heme prosthetic group is also shown.

As the use of a NADPH depletion assay for screening purposes may lead to the selection of mutants with higher NADPH consumption rates that are uncoupled to product formation, a fast HPLC screening method (2 minutes) was developed to analyse whole-cell biotransformations carried out in a 96-deep well format. The first mutagenesis round led to the selection and characterization of 8 variants in terms of relevant biocatalytic parameters

(Table 1). Improved turnover frequencies were observed, with TOF values reaching up to 37.1 min⁻¹ for L81I, corresponding to more than fivefold increase compared to the parent enzyme. All the selected variants supported higher TTN than wild-type P450-AX and displayed a more efficient utilization of NADPH, with more than 29000 turnovers with diclofenac for variant L81G and 87.4% coupling for variant H190T. Thermal stability assessment and CO-difference spectroscopy on lysates revealed that most of the introduced mutations had a slight negative effect on thermal stability and protein expression levels, respectively (Table S6, section 7.2.2).

The simple combination of beneficial mutations originating from different libraries is a viable approach,^[34] albeit potentially capable of masking epistatic interactions.^[35] Epistasis has a profound effect on protein evolution, making the effects of mutations not always additive, but rather dependent on the genetic background.^[36] As such, a combinatorial library was generated by overlap extension PCR employing degenerate primers designed to include the beneficial mutations listed on Table 1 along with the corresponding wild-type residues. Because of limitations associated with the design of degenerate codons,^[37] the produced library led to the introduction of mutations not observed during the first round of active-site testing (see section 7.2.1.2.2). The double mutants L81S/N193E, L81V/N193A and L81G/N193A were selected for further characterization. Residues L81 and N193 are predicted to be positioned in the upper part of the active site, with N193 forming a lid above the heme. Unsurprisingly, these residues are part of regions generally recognized to be highly flexible and involved in substrate recognition.^[30] In particular, L81 is predicted to be located in a loop between α B and α B', while N193 is predicted to be part of a short helical element preceding the F-G loop.

Interestingly, L81S was not identified as a beneficial mutation during our first active site testing. This might have been due to inherent screening variability or because the single mutant was not active enough. For our curiosity, P450-AX L81S was generated by site-directed mutagenesis to investigate diclofenac hydroxylation activity. As shown in Table S7 (section 7.2.2), modest improvements in product formation rate (TOF=13.1 min⁻¹) and *in-vivo* activity (0.9 vs 0.7 $\mu\text{mol}_{\text{product}} \text{min}^{-1} \text{g}_{\text{cww}}^{-1}$ for L81S and wild type, respectively) were measured. Therefore, it is likely that moderate improvements combined with screening variability did not lead to the selection of this mutant during the first engineering round. However, introduction of mutation N193E in a L81S background led to significant improvements in both

in-vivo and *in-vitro* reaction rates. Notably, these activity-enhancement effects could have not been predicted from data on single point mutants only, as these do not account for the contribution of epistasis in shaping protein fitness landscapes.

Table 1. Total turnovers, turnover frequencies and coupling efficiencies for 5-OH DCF production by selected P450-AX variants.

		TOF (min ⁻¹)	Coupling (%)	TTN
	WT	6.7	48.4	7725
Active-site testing	N193V	14.5	62.4	8550
	N193A	20.8	70.1	10634
	H190T	21.2	87.4	15950
	A259G	26.5	68.7	17375
	L81I	37.1	69.5	26375
	L81V	43.4	58.9	26463
	N193E	28.8	67.9	26500
	L81G	36.5	68.0	29375
Recombination	L81S/N193E	61.1	61.3	7475
	L81V/N193A	56.3	71.5	11425
	L81G/N193A	59.2	51.7	36750

Mean values of at least three independent experiments. Errors estimated to be at most $\pm 10\%$.

Coming to the *in-vitro* investigation of diclofenac hydroxylation by the selected variants originating from the recombination library, the purified proteins showed improved turnover frequency values (TOF=56.3-61.1 min⁻¹, Table 1), although none of the variants obtained by recombination displayed higher coupling efficiency than H190T (87.4%). Nevertheless, variant L81G/N193A supported more than 36000 turnovers with diclofenac, corresponding to almost a fivefold increase in TTN compared to the parental enzyme. Moreover, *in-vivo* activity was enhanced by the mutations introduced following recombination (2.4-2.9 $\mu\text{mol}_{\text{product}} \text{min}^{-1} \text{g}_{\text{cww}}^{-1}$, Table S8, section 7.2.2), which is desirable, considering the envisioned application of these enzymes in whole-cell format. However, the early introduced mutations led to larger improvements in the measured parameters compared to mutations introduced after recombination. These diminishing returns are often encountered during protein engineering campaigns. In particular, higher coupling efficiency (68%) and about fourfold improved TOF and TTN values were measured for P450-AX L81G compared to wild type P450-AX, whereas P450-AX L81G/N193A displayed reduced coupling efficiency (51.7%) and less than a twofold increase in TOF and TTN values compared to P450-AX L81G. Mutations away from the active-

site or stabilizing mutations are sometimes needed to “pull out” from unfavourable trajectories encountered while exploring protein fitness landscapes.^[36]

To test the robustness of the best mutant L81G/N193A under more industrially relevant conditions, bioreactor experiments have been planned.

3.6 References

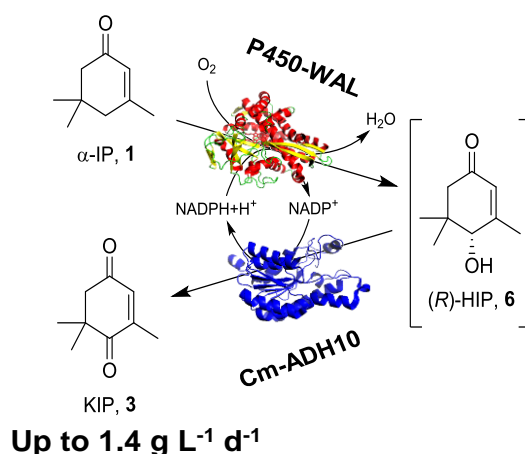
- [1] F. P. Guengerich, *Chem. Res. Toxicol.* 2001, 14, 611–650.
- [2] V. B. Urlacher, M. Girhard, *Trends Biotechnol.* 2012, 30, 26–36.
- [3] A. W. Munro, H. M. Girvan, A. E. Mason, A. J. Dunford, K. J. McLean, *Trends Biochem. Sci.* 2013, 38, 140–150.
- [4] F. Hannemann, A. Bichet, K. M. Ewen, R. Bernhardt, *Biochim. Biophys. Acta - Gen. Subj.* 2007, 1770, 330–344.
- [5] P. R. Ortiz de Montellano, *Cytochrome P450: Structure, Mechanism and Biochemistry*, Springer Science & Business Media, 2015.
- [6] Ł. Gricman, C. Vogel, J. Pleiss, *Proteins Struct. Funct. Bioinform.* 2015, 83, 1593–1603.
- [7] T. C. Pochapsky, S. Kazanis, M. Dang, *Antioxid. Redox Signal.* 2010, 13, 1273–1296.
- [8] G. Roberts, G. Grogan, A. Greter, S. L. Flitsch, N. J. Turner, *J. Bacteriol.* 2002, 184, 3898–3908.
- [9] S. C. Hammer, G. Kubik, E. Watkins, S. Huang, H. Minges, F. H. Arnold, *Science* 2017, 358, 215–218.
- [10] L. Liu, R. D. Schmid, V. B. Urlacher, *Biotechnol. Lett.* 2010, 32, 841–845.
- [11] R.-J. Li, J.-H. Xu, Q. Chen, J. Zhao, A.-T. Li, H.-L. Yu, *ChemCatChem* 2018, 10, 2962–2968.
- [12] M. Tavanti, J. L. Porter, S. Sabatini, N. J. Turner, S. L. Flitsch, *ChemCatChem* 2018, 10, 1042–1051.
- [13] J. L. Porter, S. Sabatini, J. Manning, M. Tavanti, J. L. Galman, N. J. Turner, S. L. Flitsch, *Enzyme Microb. Technol.* 2018, 113, 1–8.
- [14] D. J. B. Hunter, G. A. Roberts, T. W. B. Ost, J. H. White, S. Müller, N. J. Turner, S. L. Flitsch, S. K. Chapman, *FEBS Lett.* 2005, 579, 2215–2220.
- [15] Y. C. Yin, H. L. Yu, Z. J. Luan, R. J. Li, P. F. Ouyang, J. Liu, J. H. Xu, *ChemBioChem* 2014, 15, 2443–2449.
- [16] J.-D. Zhang, A.-T. Li, Y. Yang, J.-H. Xu, *Appl. Microbiol. Biotechnol.* 2010, 85, 615–624.
- [17] A. J. Warman, J. W. Robinson, D. Luciakova, A. D. Lawrence, K. R. Marshall, M. J. Warren, M. R. Cheesman, S. E. J. Rigby, A. W. Munro, K. J. McLean, *FEBS J.* 2012, 279, 1675–1693.
- [18] D. Minerdi, S. J. Sadeghi, G. Di Nardo, F. Rua, S. Castrignanò, P. Allegra, G. Gilardi, *Mol. Microbiol.* 2015, 95, 539–554.
- [19] L. Liu, R. D. Schmid, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* 2006, 72, 876–882.
- [20] J. A. McIntosh, T. Heel, A. R. Buller, L. Chio, F. H. Arnold, *J. Am. Chem. Soc.* 2015, 137, 13861–13865.
- [21] L. Du, S. Dong, X. Zhang, C. Jiang, J. Chen, L. Yao, X. X. Wang, X. Wan, X. Liu, X. X. Wang, et al., *Proc. Natl. Acad. Sci. U. S. A.* 2017, 114, 201702317.

- [22] C. A. Hasemann, R. G. Kurumbail, S. S. Boddupalli, J. A. Peterson, J. Deisenhofer, *Structure* 1995, 2, 41–62.
- [23] P. J. Ábek, J. Florián, V. Martínek, *Phys. Chem. Chem. Phys.* 2016, 18, 30344–30356.
- [24] T. M. Makris, K. Von Koenig, I. Schlichting, S. G. Sligar, *Biochemistry* 2007, 46, 14129–14140.
- [25] Y. T. Lee, R. F. Wilson, I. Rupniewski, D. B. Goodin, *Biochemistry* 2010, 49, 3412–3419.
- [26] P. J. Loida, S. G. Sligar, *Biochemistry* 1993, 32, 11530–11538.
- [27] M. Thompson, N. J. Turner, M. Thompson, *ChemCatChem* 2017, 9, 3833–3836.
- [28] G. Saab-Rincón, H. Alwaseem, V. Guzmán-Luna, L. Olvera, R. Fasan, *ChemBioChem* 2017, 19, 622–632.
- [29] L. Holm, P. Rosenström, *Nucleic Acids Res.* 2010, 38, 545–549.
- [30] O. Gotoh, *J. Biol. Chem.* 1992, 267, 83–90.
- [31] I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko, S. G. Sligar, *Science* 2000, 287, 1615–1622.
- [32] T. L. Poulos, B. C. Finzel, A. J. Howard, *J. Mol. Biol.* 1987, 195, 687–700.
- [33] L. Bordoli, F. Kiefer, K. Arnold, P. Benkert, J. Battey, T. Schwede, *Nat. Protoc.* 2009, 4, 1–13.
- [34] M. T. Reetz, J. D. Carballeira, J. Peyralans, H. Höbenreich, A. Maichele, A. Vogel, *Chem.Eur. J.* 2006, 12, 6031–6038.
- [35] M. T. Reetz, J. Sanchis, *ChemBioChem* 2008, 9, 2260–2267.
- [36] M. Goldsmith, D. S. Tawfik, *Curr. Opin. Struct. Biol.* 2017, 47, 140–150.
- [37] S. Kille, C. G. Acevedo-Rocha, L. P. Parra, Z. G. Zhang, D. J. Opperman, M. T. Reetz, J.P. Acevedo, *ACS Synth. Biol.* 2013, 2, 83–92.

Chapter 4. One-pot Biocatalytic Double Oxidation of α -Isophorone for the Synthesis of Ketoisophorone

Michele Tavanti,^[a] Fabio Parmeggiani,^[a] J. Rubén Gómez Castellanos,^[b] Andrea Mattevi^[b]
and Nicholas J. Turner^[a]

Published in *ChemCatChem*, **2017**, *9*, 3338-3348.



4.1 Acknowledgements

The research presented in this chapter was a collaborative effort between the doctoral candidate, Dr. Fabio Parmeggiani and Dr. J. Rubén Gómez Castellanos, under the supervision of Professor Andrea Mattevi and Professor Nicholas J. Turner. The doctoral candidate performed molecular biology experiments, biotransformations and scale-up experiments. Dr. Fabio Parmeggiani synthesized analytical standards, carried out NMR experiments and wrote the corresponding sections in this chapter. Dr. J. Rubén Gómez Castellanos determined the crystal structure of Cm-ADH10, carried out docking experiments and wrote the corresponding section in the text. The doctoral candidate wrote the rest of the manuscript. All the authors reviewed the text before submission.

^a Manchester Institute of Biotechnology (MIB), School of Chemistry, The University of Manchester, 131 Princess Street, M1 7DN, Manchester, United Kingdom.

^b Department of Biology and Biotechnology "Lazzaro Spallanzani", University of Pavia, Via Ferrata 9, 27100, Pavia, Italy

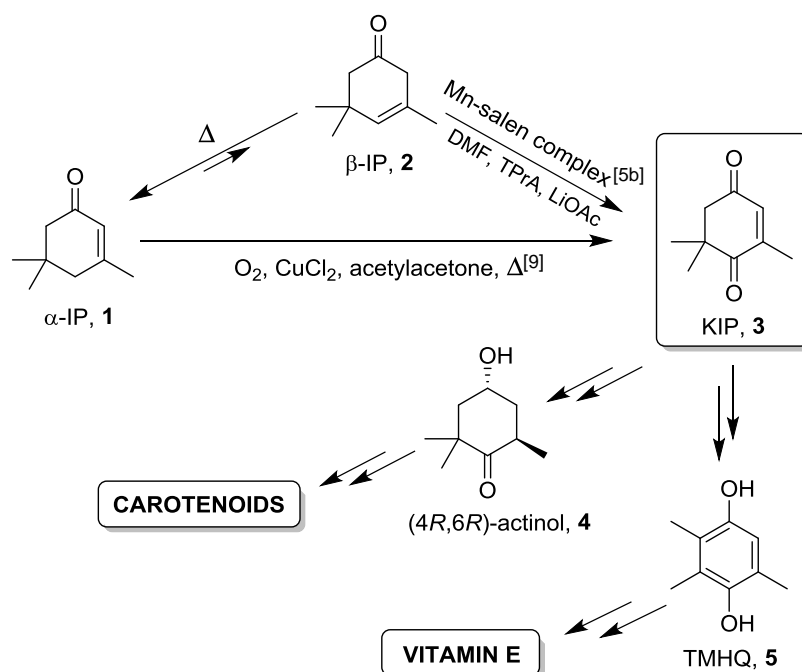
4.2 Abstract

The chemical synthesis of ketoisophorone, a valuable building block of vitamins and pharmaceuticals, suffers from a number of drawbacks in terms of reaction conditions and selectivity. Herein, the first biocatalytic one-pot double oxidation of the readily available α -isophorone to ketoisophorone is described. Variants of the self-sufficient P450cam-RhFRed with improved activity have been identified to carry out the first step of the designed cascade (regio- and enantioselective allylic oxidation of α -isophorone to 4-hydroxy- α -isophorone). For the second step, the screening of a broad panel of alcohol dehydrogenases (ADHs) led to the identification of Cm-ADH10 from *Candida magnoliae*. The crystal structure of Cm-ADH10 was solved and docking experiments confirmed the preferred position and geometry of the substrate for catalysis. The synthesis of ketoisophorone was demonstrated both as a one-pot two-step process and as a cascade process employing designer cells co-expressing the two biocatalysts, with a productivity of up to 1.4 g L⁻¹ d⁻¹.

4.3 Introduction

Monoxygenated terpenoids are relevant target compounds for the synthesis of active pharmaceutical ingredients (APIs), fragrances and nutritional supplements.^[1] For instance, ketoisophorone (2,6,6-trimethylcyclohex-2-ene-1,4-dione, KIP, **3**), is one of the key intermediates to access useful building blocks for the synthesis of carotenoids and α -tocopherol (vitamin E),^[2] as shown in Scheme 1. For instance, the biocatalytic conversion of KIP to the chiral intermediate (4*R*,6*R*)-4-hydroxy-2,2,6-trimethylcyclohexanone ((4*R*,6*R*)-actinol, **4**), which is an important precursor of several carotenoids, has been recently demonstrated in a one-pot system employing an engineered *Candida macedoniensis* old yellow enzyme (CmOYE) and *Corynebacterium aquaticum* (6*R*)-levodione reductase.^[3] Moreover, KIP can be used as a precursor of trimethylhydroquinone (TMHQ, **5**), the key building block in the synthesis of vitamin E.^[4] Commercially, KIP is produced from the readily available 3,5,5-trimethyl-2-cyclohexen-1-one (α -isophorone, α -IP, **1**) by isomerization to 3,5,5-trimethyl-3-cyclohexen-1-one (β -isophorone, β -IP, **2**), followed by oxidation to KIP.^[5a-e] However, the isomerization is a challenging process and requires high temperatures.^[6] Therefore, the direct oxidation of α -IP has become an attractive alternative route to KIP.

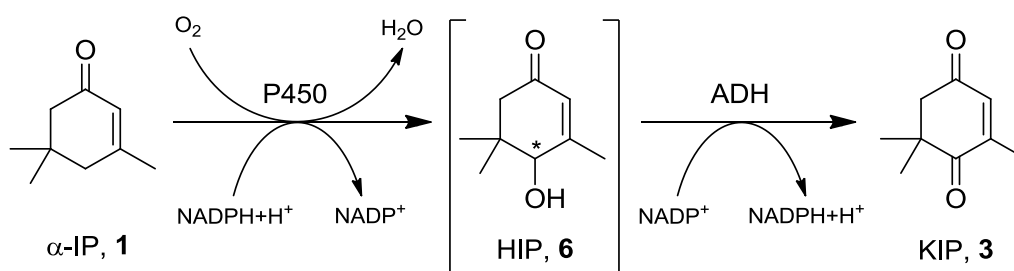
However, this selective allylic oxidation proved to be a challenging task and currently such transformation requires toxic heavy metal catalysts,^[7] generates undesired by-products^[8] and/or requires high temperatures and organic solvents.^[9]



Scheme 1. Current syntheses of KIP from α -IP by chemical oxidation, a gateway to the production of high-value building blocks for the industry of carotenoids and vitamins.

In general, the regioselective oxidation of non-activated hydrocarbons under mild reaction conditions remains a “dream reaction” in organic chemistry.^[10] With this regard, a biotechnological approach offers a green alternative to traditional chemistry, since biocatalytic reactions generally feature exquisite selectivity and are carried out under mild conditions.^[11] Cytochrome P450s (CYPs or P450s) are a diverse superfamily of heme oxidoreductases capable of performing many oxidative reactions, most notably the insertion of oxygen into a chemically inert C-H bond, using oxygen as a benign oxidant and releasing water as by-product.^[12a,b,c,d,e,f] During this catalytic cycle, oxygen activation is enabled by two electrons supplied by NAD(P)H and shuttled by redox partners.^[13a,b,c,d] Moreover, the outstanding diversity of P450s and their wide distribution in all kingdoms of life imply that their substrate scope and catalytic versatility could be much superior to other enzymes.^[14] Overall, P450s should be considered as valuable catalysts in the enzymatic toolbox of the organic chemist, even if major limitations have to be tackled to implement them in

biocatalytic processes.^[15a,b,c,d] Kaluzna *et al.* have recently reported the production of 4-hydroxy- α -isophorone (HIP, **6**, Scheme 2) from α -IP to a pilot-scale using the well-known P450-BM3 from *Bacillus megaterium* (CYP102A1),^[16] paving the way for a sustainable P450-based KIP production. On the other hand, several reports deal with the application of P450s to catalyse the initial C-H activation in one-pot multi-enzymatic transformations: examples are the oxyfunctionalization of linear hydrocarbons (both *in vitro* and *in vivo*),^[17a,b,c] the production of cycloalkanones or lactones from the corresponding cycloalkanes,^[18a,b,c] the synthesis of the grapefruit flavour (+)-nootkatone,^[19] the conversion of cyclohexane to cyclohexanediols,^[20] and the stereoselective benzylic amination of ethylbenzene derivatives.^[21] Therefore, considering these biocatalytic strategies, we designed a cascade system to access KIP from α -IP *via* a double allylic oxidation (Scheme 2), employing a P450 in the first step to afford HIP, followed by further oxidation by an alcohol dehydrogenase (ADH).



Scheme 2. Proposed one-pot biocatalytic system from α -IP to KIP.

Herein we report the development and optimization of the envisioned cascade using an improved variant of the chimeric self-sufficient P450cam-RhFRed,^[22] consisting of the heme domain of the well-known CYP101A1 (P450cam, camphor 5-monooxygenase, UniProt P00183) from *Pseudomonas putida* fused to the NADPH-dependent reductase domain (RhFRed) of CYP116B2 from *Rhodococcus* sp. (P450RhF, UniProt Q8KU27),^[23] in combination with Cm-ADH10 (GenBank AGA42262.1) from *Candida magnoliae*.^[24] Along with P450-BM3, P450cam is one of the most studied P450s,^[25a,b,c,d,e] and the plethora of information available has been exploited to engineer this enzyme mainly targeting active site hot spots Y96, F87, L244, and V247 in order to generate biocatalysts for terpene oxidation,^[26a,b,c] to accept small alkanes^[27a,b,c] or industrial pollutants such as large polycyclic aromatic hydrocarbons^[28] and polychlorinated benzenes.^[29] Our group has previously reported directed evolution methods to redesign the P450cam active site^[30a,b,c] and recently Kelly *et al.* have developed a colony-

based screening based on indigo production from indole which was exploited as a first visual screen to identify self-sufficient P450cam-RhFRed variants with altered promiscuous activities.^[31] It was shown that this activity correlates well with that towards structurally different substrates such as ethylbenzenes,^[32] corroborating the role of neutral mutations in protein evolution.^[33a,b,c,d,e,f] Following the same line of thought, we have screened the developed libraries to find variants for the first allylic oxidation of α -IP to HIP. For the second step, a broad panel of commercially available ADHs was evaluated, leading to the identification of Cm-ADH10 as a suitable candidate.

4.4 Results and Discussion

4.4.1 Screening and Improvement of P450cam-RhFRed Variants

The 96 in-house P450cam-RhFRed variants were pooled according to the mutated residues generating five pooled libraries (A, B, C, D, EFG, section 7.3.1.5) of roughly equal size to screen for α -IP hydroxylation activity in whole-cell biotransformations. In general, clustering strategies allow to reduce the screening effort, avoiding the need for screen on a single clone level.^[34] In the following step, positive clusters may be broken down to the single clone level in order to identify mutations improving the desired activity. Figure 1 shows the measured biocatalytic performance expressed in terms of the total turnover numbers (TTN), that is the ratio between total product concentration and P450 concentration, calculated over 24 h. In fact, the variability of P450 expression in *E. coli* may lead to misjudging the observed yields,^[35] thus, a fast, whole-cell assay was implemented to determine P450 concentration which was then used to normalize enzyme activity.^[31] Finally, the enantiomeric excess values (*ee*) were also determined by comparison with synthesized standards. The P450-catalyzed single oxidation of α -IP may lead to (at least) three major products (Figure 1): HIP (**6**), the regioisomeric allylic oxidation product (3-(hydroxymethyl)-5,5-dimethylcyclohex-2-en-1-one, HMIP, **7**) and isophorone oxide (2,3-epoxy-3,5,5-trimethyl-1-cyclohexanone, IPO, **8**). Product **7** was identified after scale-up, purification and NMR characterization (section 7.3.1.8). The formation of isophorone oxide was not observed and (*R*)-HIP was the preferred product (Figure 3b), with library D (mutations at residues L244-V247) showing the highest TTNs in the desired α -IP allylic oxidation. Notably, no activity was detected for mutants in pool C (mutations at residues M184-T185).

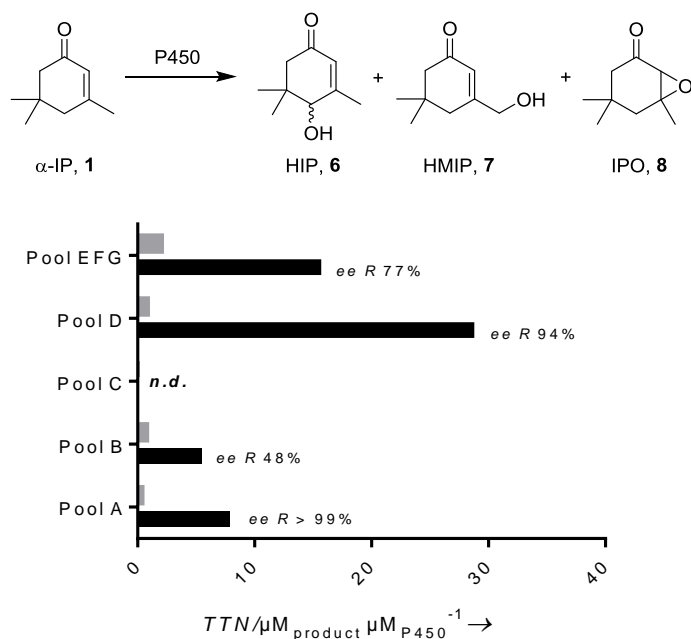


Figure 1. Possible products of the P450 oxidation of α -IP and comparison of TTNs for HIP (black bars) or HMIP (grey bars) obtained for P450cam-RhFRed pooled variants. HIP ee values are also given.

These initial results drove the next step of the investigation towards the single-clone level analysis of mutants at positions L244-V247. Residues L244 and V247 are positioned over the heme, with L244 located very close to the heme prosthetic group, and V247 in the upper section of the active site.^[36] These positions have been often targeted in conjunction with F87-Y96 in order to affect active site volume and manipulate the regio- and enantioselectivity of the oxidation of C-H bonds of unnatural substrates, as discussed in the introduction. The indigo screening adopted previously using P450cam-RhFRed-Y96F as template (indicated as parent)^[31] led to the selection of mutants which carried mainly apolar sidechains at position 247 (the only exceptions being V247N and V247S), while more variability was introduced at position 244, with polar residues (asparagine, serine, cysteine) and a positive charged residue such as histidine. Analysis of biotransformation products (Figure 2) demonstrated that positions 244 and 247 greatly affect TTNs of α -IP. Interestingly, the majority of mutants bearing the bulky aromatic phenylalanine residue show not only reduced TTNs but also a decrease in the regioselectivity of the reaction, with the formation of HIP and HMIP in approximately equal amounts by mutants L244Y-V247F and L244I-V247F. Given the observed positive effect of the substitution of L244 with small or apolar residues, the mutation L244A

was also introduced, considering the effect of the L244A mutation on substrate acceptance, regio- and enantioselectivity.^[37] This mutation was studied either in presence of the wild-type V247 or the substitution V247L, which occurred in many positive hits. The two mutants L244A and L244A-V247L showed the best TTN values (94 ± 9 and 83 ± 11 , respectively) and were selected for further engineering. Similarly to Bell *et al.*,^[26b] we reasoned that the performance of the biocatalyst could be enhanced by improving the enzyme-substrate fit, i.e., by pushing the substrate closer to the heme. With this in mind, the bulky tryptophan was introduced in place of phenylalanine at position 87, known to be involved in substrate access and recognition.^[27b,29] As shown in Figure 2, the introduction of F87W in the L244A-V247L background led to more than a 6-fold improvement of TTNs values with respect to the L244A-V247L mutant (as full conversion was obtained with this variant, the value might be underestimated), with no HMIP detected and an *ee* of 99% (*R*)-HIP, suggesting a better orienting effect of this substrate towards the highly reactive iron-oxo species. Finally, the best improved variant (Y96F-F87W-L244A-V247L, now termed P450-WAL) was selected for subsequent optimization experiments for the designed cascade reaction.

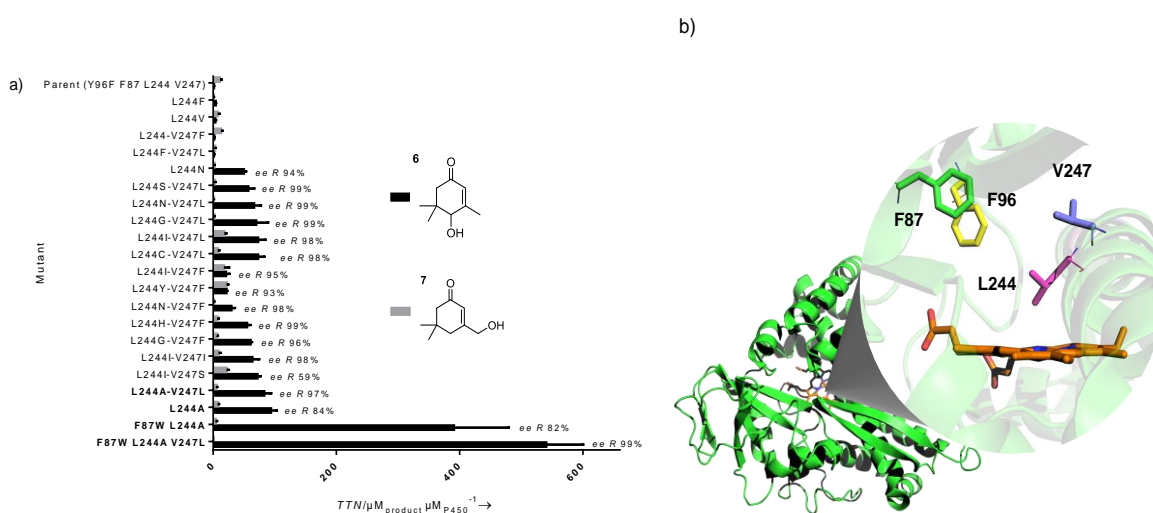


Figure 2. (a) Comparison of TTNs of α -IP to HIP (black bars) or α -IP to HMIP (grey bars) obtained for library D. HIP *ee* values are given, except where the conversions were too low. Variants generated by site-directed mutagenesis are indicated in bold. (b) Active site of P450cam variant Y96F (generated with PyMOL from PDB 2CPP^[25e]) showing the four residues targeted in this study: F87 (green), F96 (yellow), L244 (magenta) and V247 (blue). The heme prosthetic group is represented in orange.

The docking of α -IP into the modelled P450-WAL active site revealed several conformations of which three show an ideal positioning of C4 towards the heme iron (Figure 3). The mutated residues (Y96F-F87W-L244A-V247L) form a pocket that position the abstracted hydrogen of α -IP at a distance of less than 3.0 Å from the iron-linked oxygen, with an angle formed between these two atoms and the iron atom between 109°-149°, and ΔG^0 ranging from -6.73 kJ mol⁻¹ (a, Figure 3) to -6.46 kJ mol⁻¹ (c, Figure 3), satisfying the requirements for hydroxylation reactions.^[38] It is very likely that the mutations introduced can reshape the active site to position α -IP towards its oxidation into (R)-HIP: in these arrangements, the L244A mutation might create space to accommodate the dimethyl groups of C1 of α -IP, while the V247L mutation might actually push the substrate closer to and in line with the heme oxygen. In the conformations shown in Figure 3a and Figure 3b, it is also possible to identify a hydrogen bond between α -IP and W87, which might justify the lower ΔG^0 when compared with Figure 3c. It is interesting to note that the docking of the P450cam Y96F mutant indicates that the configurations favoured by α -IP would not yield the oxidation at C4 (Figure S1, section 7.3.2).

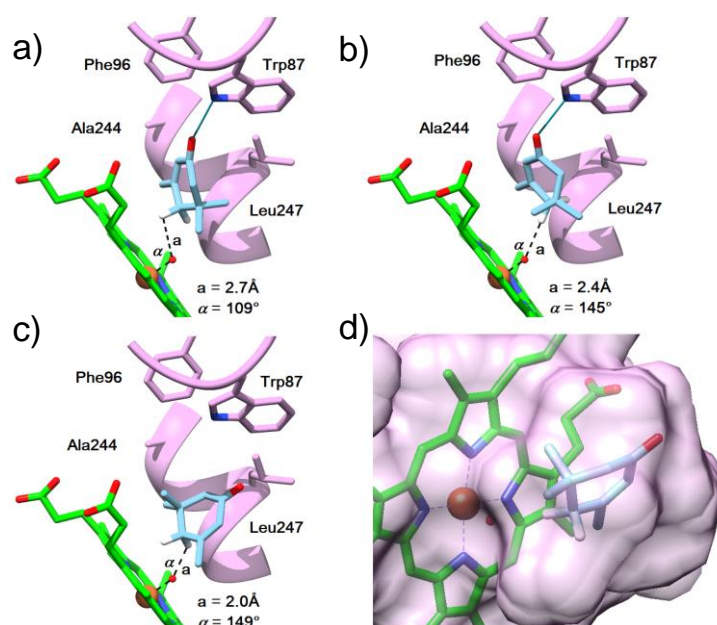


Figure 3. Docking of α -IP in the modelled active site of P450-WAL. In the catalytic site, the mutated residues W87, F96, A244 and L247 are shown in purple with, α -IP in cyan and the heme group in green. (d) The surface of the pocket created by the mutated residues in the active site shows that C4 faces the heme group for the oxidation into 4(R)-HIP.

4.4.2 Selection of HIP-oxidizing Cm-ADH10 and Structural Characterization

A screening kit of 116 wild-type ADHs from very diverse organisms and with a broad range of accepted substrates was provided by c-LEcta (c-LEcta GmbH, Leipzig, Germany). In order to identify suitable biocatalysts for the oxidation of (*R*)-HIP, the panel was screened against this substrate by HPLC. Several enzymes (>80% of the panel) showed almost no activity, likely because of the steric hindrance of the substrate, however, among the positive hits, two were selected for having the highest activity: Cm-ADH10 from *Candida magnoliae* (GenBank AGA42262.1) and the NADPH-dependent carbonyl reductase (SSCR) from *Sporobolomyces salmonicolor* (UniProt Q9UUN9).^[37a,b,c] Notably, both ADHs accept NADP(H) as cofactor, which is desirable to create a self-sufficient cascade with respect to the cofactor.^[40] The corresponding wild-type genes were cloned into a pET28a vector to carry out expression trials, which revealed good expression levels for Cm-ADH10, as opposed to SSCR (Figure S2, section 7.3.2). For this reason, we selected the former for subsequent characterization and combination with the P450. To understand better the catalytic properties of Cm-ADH10, we solved the crystal structure of its complex with NADP⁺ to a resolution of 1.6 Å (PDB 5MLN, Figure S3 and Figure S4, section 7.3.2). After soaking the crystal with a solution of (*R*)-HIP (10 mM), no substrate was bound to the crystal. However, after docking (performed with AutoDock Vina),^[41] the most favoured calculation for the bound ligand (*R*)-HIP confirms that the cavity can accommodate the substrate at a distance and geometry that would favour the transfer of a hydride from the chiral carbon on the 4-(*R*)-hydroxy moiety to C4N of NADP⁺ (Figure 4). In this position, the 4-hydroxy group would be in the right geometry to interact with the catalytically-important residues S144 and Y157, while the rest of the molecule of 4(*R*)-HIP is possibly stabilized by interactions with residues H149 and Y189 in the entrance of the active site. Coincidentally, this area in the crystal structure is occupied by a tetrad of water molecules that have been reported to be bound to the Tyr-OH and Lys side chain, thus mimicking substrate and ribose hydroxyl group positions.^[42] Previous studies support this concept, where Y157 would function as the catalytic base, whereas S144 would stabilize the substrate and K161 would form hydrogen bonds with the nicotinamide ribose moiety and lower the pK_a of the Tyr-OH to promote proton transfer.^[43] Since the docking does not reveal

any direct interaction with this (or other) residues in the active site, it is very likely that the Cm-ADH10 activity arises from its spacious active site that would allow the binding of (*R*)-HIP in an induced-fit mechanism.

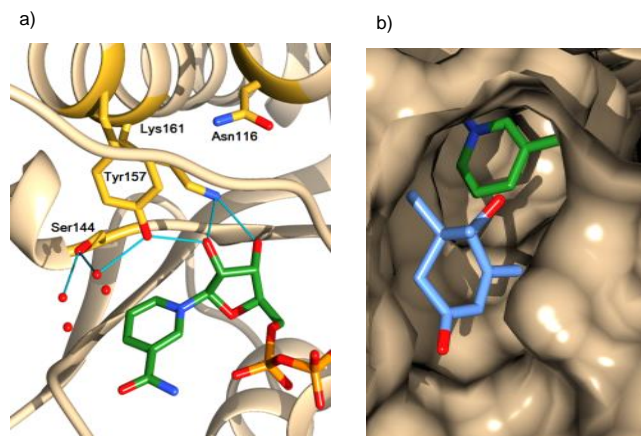


Figure 4. The crystal structure of Cm-ADH10. (a) In the catalytic site four water molecules are within close distance (3.5 Å) of C4N of NADP⁺ (green), where a substrate would be located for hydride transfer; Tyr157 and Lys161 (yellow, along with other important residues) interact with nicotinamide ribose to orient the C4N carbon towards the substrate, and Ser144 would stabilize the substrate. (b) (*R*)-HIP (light blue) docked in the active site cavity with the chiral carbon pointing towards C4N of NADP⁺.

Preliminary biotransformations were carried out with varying (*R*)-HIP concentrations (10-100 mM) and a constant NADP⁺ concentration (0.25 mM), using 1 mg mL⁻¹ purified Cm-ADH10 and 0.5% v/v ethyl acetoacetate as co-substrate for cofactor regeneration by the same ADH acting as a dual-functional enzyme (see below for further details about co-substrate screening). With this setup, conversions ranged from 92% with 10 mM substrate, down to 48% with 100 mM substrate (see Figure S5, section 7.3.2). Thus, Cm-ADH10 lends itself to being applied in the designed bi-enzymatic cascade.

4.4.3 Assembling the System: Initial Unsuccessful Attempts

Initially, we tried to combine the two selected biocatalysts in a one-pot, two-step format, adding 1 mg mL⁻¹ of purified Cm-ADH10, cofactor and co-substrate (see above) directly to the supernatant of the whole-cell P450 reaction carried out at 20°C with 200 mg mL⁻¹ wet cell load in 50 mM sodium phosphate buffer pH 7.2, 100 mM KCl (indicated as standard biotransformation buffer) and 10 mg mL⁻¹ glucose for cofactor regeneration by *E. coli*

metabolism. Unexpectedly, we obtained very low conversion values for the ADH catalysed reaction, starting from just 10 mM α -IP in the first step. We reasoned that potential inhibitors for the second step could originate from the growth of the whole-cell biocatalysts in the M9 minimal medium supplemented with glucose. Considering that growth media can influence the performance of biocatalysts,^[44] and that glucose is one of the most acidogenic carbon sources for enterobacteria,^[45] the pH of the supernatant resulting from the first oxidation step carried out in phosphate buffer with different concentration was measured (Figure 5). Surprisingly, the pH dropped by 1.5 units when the standard biotransformation buffer was used, and even with a 300 mM potassium phosphate (KPi) buffer the pH decrease was significant. Nevertheless, by increasing the concentration of the buffer employed in the first step, higher conversions were observed in the second one. Eventually, more than a 2-fold increase in HIP conversion was observed when the P450 catalysed reaction carried out in standard biotransformation buffer was titrated to pH 8.0 before the addition of Cm-ADH10. A similar trend was observed when cells grown in M9 medium were simply resuspended in buffer without any addition of substrate or glucose, suggesting that the observed pH decrease is linked to the metabolism of cells grown to $OD_{600} \sim 5.0-6.5$. Even if no attempt was made to gain a better understanding of the phenomenon and investigating the effect of washing the biocatalyst, the latter experiment highlighted the need for a careful optimization of each individual step before combination.

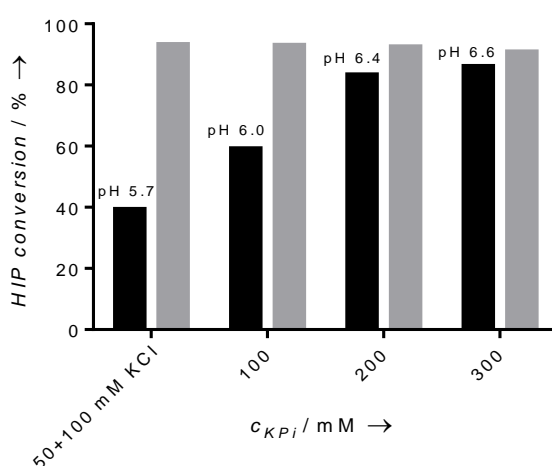


Figure 5. Effect of the pH reached after the P450-catalyzed oxidation of α -IP (10 mM) on Cm-ADH10 performance. KPi buffer concentration used in the first step is given on the x-axis. Black bars: HIP conversion after direct addition of Cm-ADH10 (the pH measured after the first step is indicated above). Grey bars: HIP conversion when the supernatant of the first reaction was titrated to pH 8.0 before Cm-ADH10 addition. Detailed reaction conditions are given in the text.

4.4.4 Parallel Reaction Optimization

Next, we proceeded with the parallel optimization of the two oxidative steps with respect to buffer concentration, pH and temperature. For the P450 allylic oxidation step, KP_i buffer (pH 8.0) was chosen for the unique K^+ binding site of P450cam, which displays higher stability and superior camphor binding in presence of K^+ ions.^[46] The effect of buffer concentration was examined, along with the temperature optimum. As shown in Figure 6a, conversion improved with increasing buffer concentration up to 200-300 mM, and by comparing biotransformations carried out in 50 mM KP_i without KCl, with KCl or in 100 mM KP_i , it can be concluded that buffer capacity seems to have a greater effect on conversion than K^+ concentration. Eventually, 200 mM KP_i buffer was chosen to investigate the effect of temperature on conversion, finding the optimum at 28°C (Figure 6b).

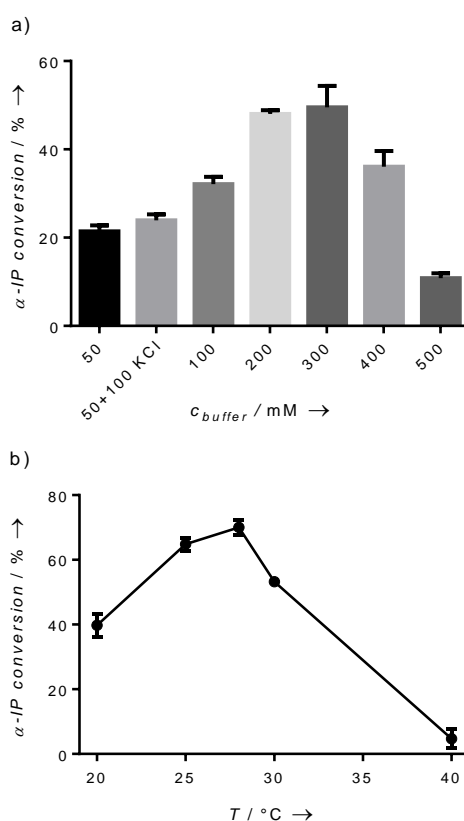


Figure 6. Optimization of the allylic oxidation of α -IP catalysed by P450-WAL. (a) Effect of buffer concentration. Reaction set-up: 200 mg mL^{-1} wet cells resuspended in KP_i buffer pH 8.0 (concentration given on the x-axis), 10 mg mL^{-1} glucose, 20 mM α -IP, 2% DMSO, 20°C, 1 mL final volume in deep-well plate, 24 h reaction time. (b) Effect of temperature. Reaction set-up: same as above, with the exceptions being the buffer adopted (200 mM KP_i , pH 8.0) and the substrate concentration (15 mM).

With respect to the Cm-ADH10 alcohol oxidation, different co-substrates were selected and tested in order to exploit the biocatalyst as a dual-functional enzyme capable of performing the whole substrate oxidation-cofactor regeneration cycle (Figure 7a).^[47] As potential effective co-substrates we tested acetone (**9a**), as well as activated ketones such as chloroacetone (**9b**), ethyl acetoacetate (**9c**) and ethyl levulinate (**9d**), since their reduction (and hence cofactor regeneration) is favoured by an additional intramolecular hydrogen bond between the newly formed alcohol functional group and the electronegative moiety.^[48] Chloroacetone and ethyl acetoacetate proved to be superior and, even if the latter performed better during the first six hours of reaction, the former pushed the conversion to 84% with 40 mM substrate after 24 h (vs. 67% with ethyl acetoacetate). Therefore, chloroacetone was employed as hydrogen acceptor in all the subsequent optimization steps. The pH effect on Figure 7b partially explains our initial unsuccessful attempts to combine the two enzymes: in fact, enzyme activity drops below pH 6.0, with an optimum pH range between 7.0 and 9.0. Moreover, the buffer system employed in the P450 catalysed reaction (200 mM KP_i) displayed perfect compatibility with the second step (see Figure S6, section 7.3.2). Finally, we turned our attention to temperature optimization, finding the highest conversion at 40°C (Figure 7c). With these optimized conditions in hand, we moved on to the envisaged bi-enzymatic cascade concept to carry out scale-up experiments for product isolation and characterization.

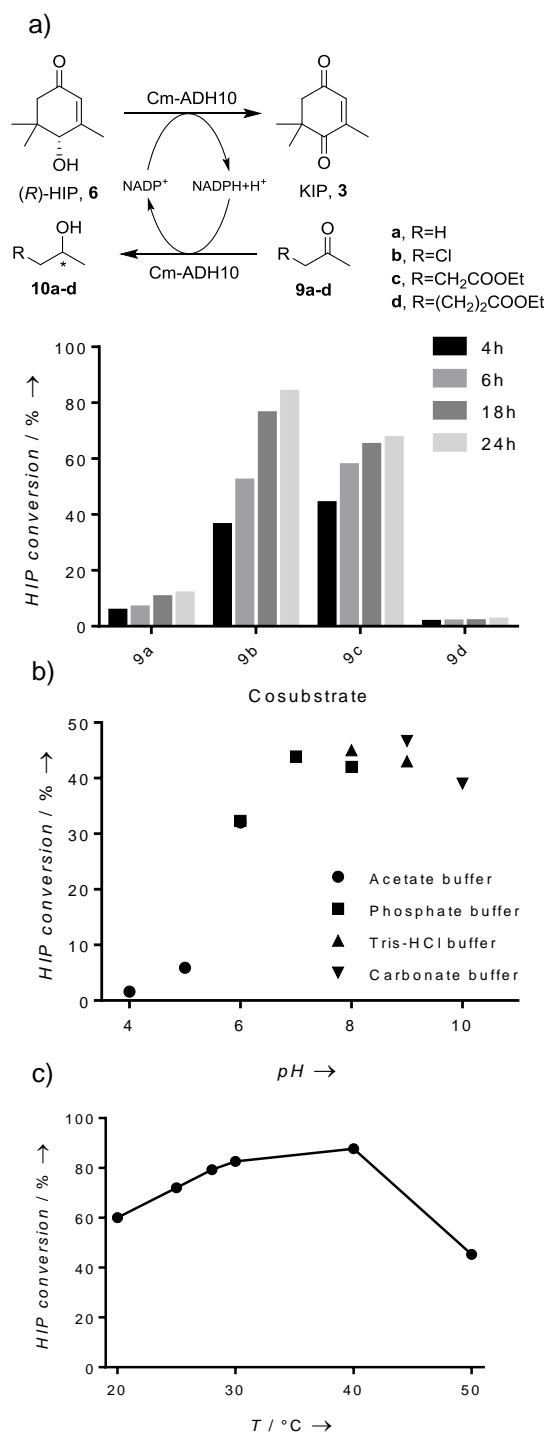


Figure 7. Optimization of the oxidation of (*R*)-HIP catalysed by Cm-ADH10. (a) Effect of different co-substrates. Reaction set up: 100 mM KP_i buffer pH 7.5, 1 mg mL⁻¹ Cm-ADH10, 0.25 mM NADP⁺, 0.5% v/v cosubstrate, 40 mM (*R*)-HIP, 30°C. (b) Effect of pH. Reaction set up: 100 mM indicated buffer, 1 mg mL⁻¹ Cm-ADH10, 0.25 mM NADP⁺, 0.5% v/v chloroacetone, 40 mM (*R*)-HIP, 30°C, 4 h. (c) Effect of temperature. Reaction set up: 200 mM KP_i buffer pH 8.0, 1 mg mL⁻¹ Cm-ADH10, 0.25 mM NADP⁺, 0.5% v/v cosubstrate, 40 mM (*R*)-HIP, 24 h.

4.4.5 Preparative-scale Double Oxidation of α -Isophorone to Ketoisophorone

In order to simplify the entire process further, the oxidation of (*R*)-HIP was accomplished using a concentrated cell-free extract (CFE, corresponding to 1 mg/mL final protein concentration) of *E. coli* overexpressing Cm-ADH10, thus avoiding the expensive protein purification step. The course of the reaction under the optimized conditions was followed over a total reaction time of 50 h, showing that the P450 catalysed allylic oxidation of 10 mM α -IP reached 94 \pm 2% conversion in 18 h, whereas the second step was much faster, reaching complete conversion over 6 h (Figure 8). Remarkably, the addition of NADP⁺ was found to be unnecessary to achieve the highest conversion in the second step, further reducing process costs. Similarly to Kaluzna *et al.*^[16], some degree of overoxidation of HIP to KIP was observed during the first step, probably resulting in the subsequent reduction of the C=C bond of KIP by endogenous ene-reductases (<5% by-product).^[18c]

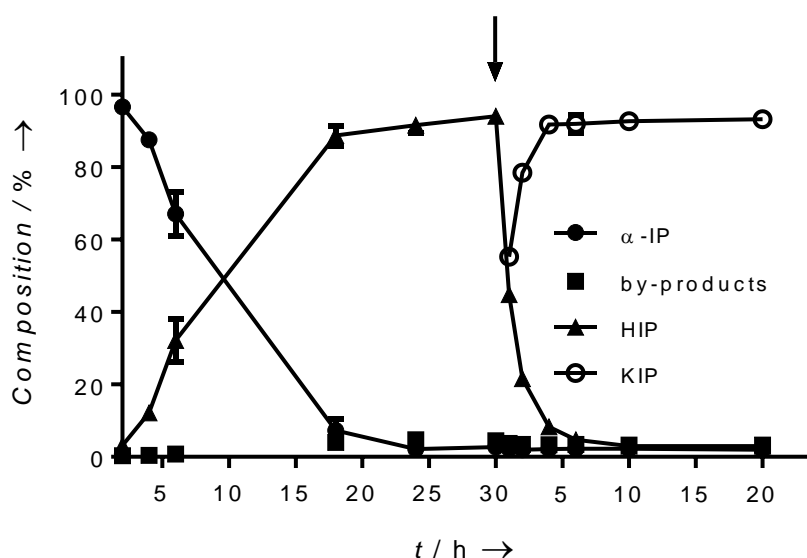
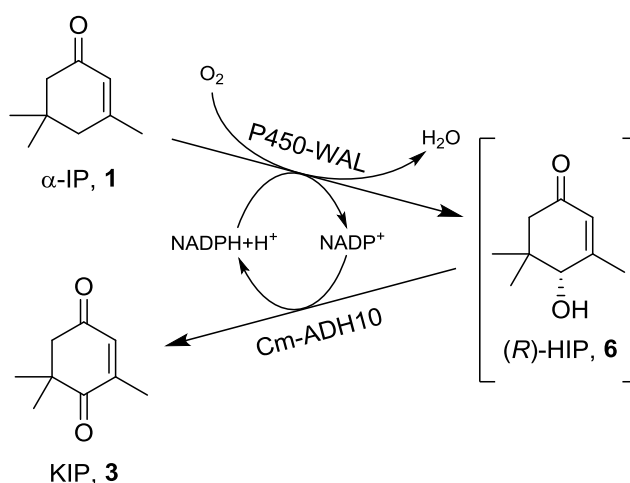


Figure 8. Time-course experiment for the one-pot two-step double allylic oxidation of α -IP. Reaction set up: 200 mg mL⁻¹ wet cells resuspended in 200 mM KP_i buffer pH 8.0, 10 mg mL⁻¹ glucose, 10 mM α -IP, 2% DMSO (first step, 28°C), then addition of 20% v/v Cm-ADH10 cell-free extract and 1.6 eq. of chloroacetone to the supernatant of the first reaction (second step, 40°C). The arrow indicates the time of addition.

Given these results, we have also attempted to co-express the P450-WAL variant with Cm-ADH10 in the same host with a two-vector system to carry out the whole-cell double oxidation of α -IP (Scheme 3). After 24 h, the conversion reached an average value of 65 \pm 2%,

approximately 30% lower than for the one-pot two-step process carried out at screening scale (Figure S7b, section 7.3.2). However, the co-expression of the two enzymes and cofactor preferences make this reaction redox-balanced, avoiding addition of a CFE and co-substrate. Lower conversion can likely be linked to the metabolic burden associated with the co-expression of multiple biocatalysts:^[49] in fact, the fraction of active P450 as determined by CO difference spectroscopy was reduced 3-fold in the “designer” microorganism (6.0 μM for the two step process vs. 1.9 μM for the whole-cell process), which accounted for the residual α -IP. Based on the quantified products after 24 h, a productivity of 1.4 $\text{g L}^{-1} \text{d}^{-1}$ and 1.1 $\text{g L}^{-1} \text{d}^{-1}$ could be calculated for the two-step and the whole-cell process, respectively. However, the one-step process displayed higher TTN for the P450 step when compared to the two-step process (3421 vs. 1567), possibly indicating a more effective cofactor recycling system and/or a positive effect resulting from the removal of the intermediate HIP.^[18c]



Scheme 3. Whole-cell double oxidation cascade of α -IP to KIP using *E. coli* co-expressing P450-WAL and Cm-ADH10.

The scale-up of P450-catalysed whole-cell reactions is notoriously a challenging task, due to biological and process limitations.^[15a] In order to evaluate the scalability of our one-pot two-step double allylic oxidation of α -IP, we carried out a linear scale-up (100 mg α -IP, 65 mL total volume). After 24 h, a good scalability was demonstrated, with 80% substrate converted, an excellent 87% recovery and a total isolated yield of 56% after extraction with ethyl acetate and column chromatography (see Figure S7a, section 7.3.2).

The calculated amount of HIP produced per biomass (biocatalyst yield) after the scale-up of the P450-catalysed allylic hydroxylation ($0.006 \text{ g}_{\text{product}}/\text{g}_{\text{cww}}$) is lower than the corresponding value obtained using P450-BM3 at high cell density ($0.016 \text{ g}_{\text{product}}/\text{g}_{\text{cww}}$, 100 L scale-up)^[16] or the three-component P450cam system ($0.045 \text{ g}_{\text{product}}/\text{g}_{\text{cww}}$, 200 mL reaction volume).^[50] However, a direct comparison between the system reported herein and similar systems for HIP production is complicated by differences in scale, redox partner systems employed, cell loading and different process optimization strategies adopted.

4.5 Conclusions

The direct oxidation of inexpensive α -isophorone to ketoisophorone is a sought-after chemical transformation, in that the latter has found many applications in the synthesis of carotenoids. As each chemical process leading to KIP is plagued by many drawbacks, here we demonstrated the first biocatalytic one pot double oxidation of α -isophorone to ketoisophorone by combining an improved variant of the chimeric self-sufficient P450cam-RhFRed and Cm-ADH10 from *Candida magnoliae*. A panel of previously developed P450 variants has been screened and rationally re-designed to improve the enzyme-substrate fit, generating data that could constitute a valid starting point for modelling experiments. Cm-ADH10 was identified as a suitable catalyst for the second step and its crystal structure is provided. Although no substrate is bound, further docking experiments may shed light on key residues involved in the selectivity towards HIP and confirm the preferred position and geometry of the substrate for catalysis. This will enable further intensification of the process. The need for careful and stepwise reaction optimization when designing cascades has been highlighted once again by our initial unsatisfactory results. Eventually, bottlenecks have been identified and the designed cascade has been demonstrated in two different systems: a one-pot, two-step process with Cm-ADH10 acting as a dual-functional enzyme for direct cofactor regeneration (using just 1.6 eq. of chloroacetone) or a one-pot, whole-cell cascade process employing a “designer cell” to co-express the two biocatalysts for a redox-neutral reaction. Despite a lower productivity, the latter system deserves attention in that it allows an easier and faster preparation of the biocatalyst.

4.6 References

- [1] S. Schulz, M. Girhard, V. B. Urlacher, *ChemCatChem* 2012, 4, 1889–1895.
- [2] M. Eggersdorfer, D. Laudert, U. Letinois, T. McClymont, J. Medlock, T. Netscher, W. Bonrath, *Angew. Chem. Int. Ed.* 2012, 51, 12960–12990.
- [3] S. Horita, M. Kataoka, N. Kitamura, T. Nakagawa, T. Miyakawa, J. Ohtsuka, K. Nagata, S. Shimizu, M. Tanokura, *ChemBioChem* 2015, 16, 440–445.
- [4] W. Bonrath, M. Eggersdorfer, T. Netscher, *Catal. Today* 2007, 121, 45–57.
- [5] a) W. Bonrath, T. Netscher, *Appl. Catal. A Gen.* 2005, 280, 55–73; b) M. Klatt, T. Muller, R. Bockstiegel, *Process for Preparing Oxoisophorone*, 2001, US 6,300,521 B1. c) J. Becker, U. Hochstrasser, W. Skorianetz, *Process for the Preparation of Oxophorones*, 1976, 3,944,620; d) R. Hahn, U. Gora, K. Huthmacher, F. Hubner, *Process for the Preparation of Ketoisophorone*, 1996, EP 0 808 816 B1; e) J. Mao, N. Li, H. Li, X. Hu, *J. Mol. Catal. A Chem.* 2006, 258, 178–184.
- [6] H. Bellut, *Procedure for the Production of Beta-Isophorone from Alpha-Isophorone*, 1994, US 4,845,303.
- [7] E. F. Murphy, A. Baiker, *J. Mol. Catal. A Chem.* 2002, 179, 233–241.
- [8] C. Wang, G. Wang, J. Mao, Z. Yao, H. Li, *Catal. Commun.* 2010, 11, 758–762.
- [9] W. Zhong, L. Mao, Q. Xu, Z. Fu, G. Zou, Y. Li, D. Yin, H. Luo, S. R. Kirk, *Appl. Catal. A Gen.* 2014, 486, 193–200.
- [10] H. Gröger, *Angew. Chemie - Int. Ed.* 2014, 53, 3067–3069.
- [11] J. Brummund, M. Müller, T. Schmitges, I. Kaluzna, D. Mink, L. Hilterhaus, A. Liese, *J. Biotechnol.* 2016, 233, 143–150.
- [12] a) P. R. Ortiz de Montellano, *Cytochrome P450: Structure, Mechanism and Biochemistry*, 2015; b) T. L. Poulos, *Chem. Rev.* 2014, 3919–3962. c) A. W. Munro, H. M. Girvan, A. E. Mason, A. J. Dunford, K. J. McLean, *Trends Biochem. Sci.* 2013, 38, 140–150; d) H. Renault, J. E. Bassard, B. Hamberger, D. Werck-Reichhart, *Curr Opin Plant Biol* 2014, 19, 27–34; e) J. A. McIntosh, C. C. Farwell, F. H. Arnold, *Curr. Opin. Chem. Biol.* 2014, 19, 126–134; f) G.-D. Roiban, M. T. Reetz, *Chem. Commun.* 2008, 51, 2208–2224.
- [13] a) F. Hannemann, A. Bichet, K. M. Ewen, R. Bernhardt, *Biochim. Biophys. Acta - Gen. Subj.* 2007, 1770, 330–344. b) D. Hamdane, H. Zhang, P. Hollenberg, *Photosynth. Res.* 2008, 98, 657–666. c) A. W. Munro, H. M. Girvan, K. J. McLean, *Nat. Prod. Rep.* 2007, 24, 585–609. d) T. M. Makris, I. Denisov, I. Schlichting, S. G. Sligar, in *Cytochrome P450 Structure Mechanism and Biochemistry Third Ed.*, 2005, pp. 149–182.
- [14] H. M. Girvan, A. W. Munro, *Curr. Opin. Chem. Biol.* 2016, 31, 136–145.
- [15] a) M. T. Lundemo, J. M. Woodley, *Appl. Microbiol. Biotechnol.* 2015, 99, 2465–2483; b) V. B. Urlacher, M. Girhard, *Trends Biotechnol.* 2012, 30, 26–36. c) R. Bernhardt, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* 2014, 98, 6185–6203; d) E. O'Reilly, V. Köhler, S. L. Flitsch, N. J. Turner, *Chem. Commun. (Camb)*. 2011, 47, 2490–501.
- [16] I. Kaluzna, T. Schmitges, H. Straatman, D. van Tegelen, M. Müller, M. Schürmann, D. Mink, *Org. Process Res. Dev.* 2016, 20.4, 814–819.
- [17] a) C. A. Müller, B. Akkapurathu, T. Winkler, S. Staudt, W. Hummel, H. Gröger, U. Schwaneberg, *Adv. Synth. Catal.* 2013, 355, 1787–1798; b) C. A. Müller, A. Dennig, T. Welters, T. Winkler, A. J. Ruff, W. Hummel, H. Gröger, U. Schwaneberg, *J. Biotechnol.* 2014, 191, 196–204; c) R. Agudo, M. T. Reetz, *Chem. Commun. (Camb)*. 2013, 49, 10914–6.
- [18] a) S. Staudt, E. Burda, C. Giese, C. A. Müller, J. Marienhagen, U. Schwaneberg, W. Hummel, K. Drauz, H. Gröger, *Angew. Chemie - Int. Ed.* 2013, 52, 2359–2363; b) A. Pennec, F. Hollmann, M. S. Smit, D. J. Opperman, *ChemCatChem* 2015, 7, 236–239; c) C. A. Müller, A. M. Weingartner, A. Dennig, A. J. Ruff, H. Gröger, U. Schwaneberg, *J. Ind. Microbiol. Biotechnol.* 2016, 43.12, 1641–1646.
- [19] S. Schulz, M. Girhard, S. K. Gaßmeyer, V. D. Jäger, D. Schwarze, A. Vogel, V. B. Urlacher, *ChemCatChem* 2015, 7, 601–604.
- [20] A. Li, A. Ilie, Z. Sun, R. Lonsdale, J.-H. Xu, M. T. Reetz, *Angew. Chemie* 2016, 1–5.

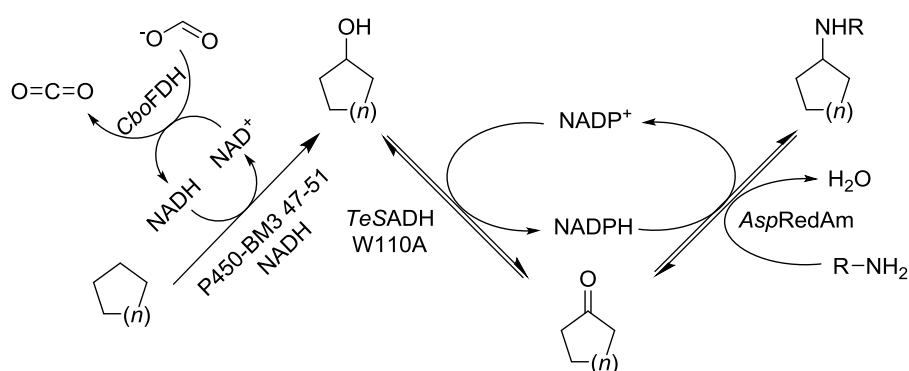
- [21] P. Both, H. Busch, P. P. Kelly, F. G. Mutti, N. J. Turner, S. L. Flitsch, *Angew. Chem. Int. Ed.* 2015, 1511–1513.
- [22] A. Robin, V. Kohler, A. Jones, A. Ali, P. P. Kelly, E. O'Reilly, N. J. Turner, S. L. Flitsch, *Beilstein J. Org. Chem.* 2011, 7, 1494–1498.
- [23] E. O'Reilly, M. Corbett, S. Hussain, P. P. Kelly, D. Richardson, S. L. Flitsch, N. J. Turner, *Catal. Sci. Technol.* 2013, 3, 1490–1492.
- [24] A. Gupta, E. Tschentscher, M. Bobkova, *Process for the Enantioselective Enzymatic Reduction of Secodione Derivatives*, 2012, US 8,323,936.
- [25] a) A. M. Colthart, D. R. Tietz, Y. Ni, J. L. Friedman, M. Dang, T. C. Pochapsky, *Sci. Rep.* 2016, 6, 22035; b) F. P. Guengerich, M. R. Waterman, M. Egli, *Trends Pharmacol. Sci.* 2016, 37, 625–640; c) S. A. Hollingsworth, T. L. Poulos, *Protein Sci.* 2015, 24, 49–57; d) I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko, S. G. Sligar, *Science*, 2000, 287, 1615–1622; e) T. L. Poulos, B. C. Finzel, A. J. Howard, *J. Mol. Biol.* 1987, 195, 687–700.
- [26] a) S. G. Bell, R. J. Sowden, L. L. Wong, *Chem. Commun.* 2001, 635–636; b) S. G. Bell, X. Chen, R. J. Sowden, F. Xu, J. N. Williams, L.-L. Wong, Z. Rao, *J. Am. Chem. Soc.* 2003, 125, 705–714; c) R. J. Sowden, S. Yasmin, N. H. Rees, S. G. Bell, L.-L. Wong, *Org. Biomol. Chem.* 2005, 3, 57–64.
- [27] a) J. A. Stevenson, J. K. Bearpark, L. L. Wong, *New J. Chem.* 1998, 22, 551–552; b) S. G. Bell, J.-A. Stevenson, H. D. Boyd, S. Campbell, A. D. Riddle, E. L. Orton, L.-L. Wong, *Chem. Commun. (Camb)*. 2002, 490–491; c) F. Xu, S. G. Bell, J. Lednik, A. Insley, Z. Rao, L. L. Wong, *Angew. Chemie - Int. Ed.* 2005, 44, 4029–4032.
- [28] C. F. Harford-Cross, A. B. Carmichael, F. K. Allan, P. A. England, D. A. Rouch, L.-L. Wong, *Protein Eng.* 2000, 13, 121–128.
- [29] J. P. Jones, E. J. O'Hare, L. L. Wong, *Eur. J. Biochem.* 2001, 268, 1460–1467.
- [30] a) R. E. Speight, F. E. Hancock, C. Winkel, H. S. Bevinakatti, M. Sarkar, S. L. Flitsch, N. J. Turner, *Tetrahedron Asymmetry* 2004, 15, 2829–2831; b) A. Çelik, R. E. Speight, N. J. Turner, A. Celik, R. E. Speight, N. J. Turner, *Chem. Commun.* 2005, 3652–4; c) S. K. Manna, S. Mazumdar, *Dalton Trans.* 2010, 39, 3115–3123.
- [31] P. P. Kelly, A. Eichler, S. Herter, D. C. Kranz, N. J. Turner, S. L. Flitsch, *Beilstein J. Org. Chem.* 2015, 11, 1713–1720.
- [32] A. Eichler, L. Gricman, S. Herter, P. Kelly, N. Turner, J. Pleiss, S. Flitsch, *Chembiochem* 2015, 17, 426–432.
- [33] a) A. Currin, N. Swainston, P. J. Day, D. B. Kell, *Chem. Soc. Rev.* 2015, 44, 1172–239; b) P. A. Romero, F. H. Arnold, *Nat. Rev. Mol. Cell Biol.* 2009, 10, 866–76; c) J. D. Bloom, P. a Romero, Z. Lu, F. H. Arnold, *Biol. Direct* 2007, 2, 17; d) J. D. Bloom, S. T. Labthavikul, C. R. Otey, F. H. Arnold, *Proc. Natl. Acad. Sci. U. S. A.* 2006, 103, 5869–5874; e) R. D. Gupta, D. S. Tawfik, *Nat. Methods*, 2008, 5, 939–942; f) Kaltenbach, Miriam, and Nobuhiko Tokuriki, *Generation of effective libraries by neutral drift. Directed Evolution Library Creation: Methods and Protocols*, 2014, 69–81.
- [34] G. Hoffmann, K. Bönsch, T. Greiner-Stöffele, M. Ballschmiter, *Protein Eng. Des. Sel.* 2011, 24, 439–446.
- [35] M. J. Weissenborn, S. Notonier, S.-L. Lang, K. B. Otte, S. Herter, N. J. Turner, S. L. Flitsch, B. Hauer, *Chem. Commun.* 2016, 52, 6158–6161.
- [36] P. J. Loida, S. G. Sligar, *Biochemistry* 1993, 32, 11530–11538.
- [37] J. J. De Voss, O. Sibbesen, Z. Zhang, P. R. Ortiz de Montellano, *J. Am. Chem. Soc.* 1997, 119, 5489–5498.
- [38] R. Lonsdale, J. N. Harvey, A. J. Mulholland, *J. Phys. Chem. B* 2010, 114, 1156–1162.
- [39] a) K. Kita, T. Fukura, K. I. Nakase, K. Okamoto, H. Yanase, M. Kataoka, S. Shimizu, *Appl. Environ. Microbiol.* 1999, 65, 5207–5211; b) K. Kita, K. I. Nakase, H. Yanase, M. Kataoka, S. Shimizu, *J. Mol. Catal. - B Enzym.* 1999, 6, 305–313; c) S. Kamitori, A. Iguchi, A. Ohtaki, M. Yamada, K. Kita, *J. Mol. Biol.* 2005, 352, 551–558.
- [40] S. Kara, A. Bornadel, R. Hatti-Kaul, F. Hollmann, *ChemCatChem* 2015, 7, 2442–2445.
- [41] O. Trott, A. J. Olson, *J. Comput. Chem.* 2010, 31, 2967–2970.

- [42] C. Filling, K. D. Berndt, J. Benach, S. Knapp, T. Prozorovski, E. Nordling, R. Ladenstein, H. Jörnvall, U. Oppermann, *J. Biol. Chem.* 2002, 277, 25677–25684.
- [43] H. Jörnvall, B. Persson, M. Krook, S. Atrian, R. González-Duarte, J. Jeffery, D. Ghosh, *Biochemistry* 1995, 34, 6003–13.
- [44] H. Ramesh, P. Zajkoska, M. Rebroš, J. M. Woodley, *Enzyme Microb. Technol.* 2016, 83, 7–13.
- [45] F. C. Neidhardt, P. L. Bloch, D. F. Smith, *J. Bacteriol.* 1974, 119, 736–747.
- [46] B. OuYang, S. S. Pochapsky, G. M. Pagani, T. C. Pochapsky, *Biochemistry* 2006, 45, 14379–14388.
- [47] W. Kroutil, H. Mang, K. Edegger, K. Faber, *Adv. Synth. Catal.* 2004, 346, 125–142.
- [48] I. Lavandera, A. Kern, V. Resch, B. Ferreira-Silva, A. Glieder, W. M. F. Fabian, S. De Wildeman, W. Kroutil, *Org. Lett.* 2008, 10, 2155–2158.
- [49] T. Bayer, S. Milker, T. Wiesinger, F. Rudroff, M. D. Mihovilovic, *Adv. Synth. Catal.* 2015, 357, 1587–1618.
- [50] S. Dezvarei, J. H. Z. Lee, S. G. Bell, *Enzyme Microb. Technol.* 2018, 111, 29–37.

Chapter 5. A Biocatalytic Cascade for the Amination of Unfunctionalized Cycloalkanes

Michele Tavanti,^[a] Juan Mangas-Sanchez,^[a] Sarah L. Montgomery,^[a] Matthew P. Thompson^[a]
and Nicholas J. Turner^[a]

Published in *Org. Biomol. Chem.*, **2017**, *15*, 9790-9793.



5.1 Acknowledgements

The research presented in this chapter was a collaborative effort between the doctoral candidate, Dr. Juan Mangas-Sanchez, Ms. Sarah L. Montgomery and Dr. Matthew P. Thompson, under the supervision of Professor Nicholas J. Turner. The doctoral candidate performed molecular biology experiments, biotransformations and scale-up experiments. Dr. Juan Mangas-Sanchez and Ms. Sarah L. Montgomery synthesized analytical standards, performed NMR experiments and wrote the corresponding sections in this chapter. The doctoral candidate wrote the rest of the manuscript. All the authors reviewed the text before submission.

^a Manchester Institute of Biotechnology (MIB), School of Chemistry, The University of Manchester, 131 Princess Street, M1 7DN, Manchester, United Kingdom

5.2 Abstract

Here we describe a one-pot, four-enzyme, cascade involving a cytochrome P450 monooxygenase, a formate dehydrogenase, an alcohol dehydrogenase and a reductive aminase for the synthesis of secondary amines from cycloalkanes. Amine product concentrations of up to 19.6 mM were achieved. The preparative scale amination of cyclohexane was also demonstrated with a space-time yield of 2 g L⁻¹ d⁻¹.

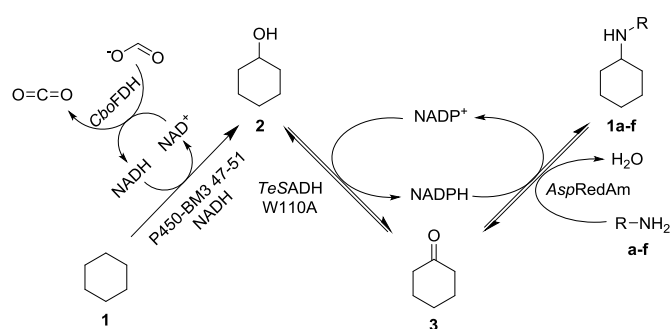
5.3 Introduction

A large number of biologically active molecules contain nitrogen functionalities,^[1,2] and consequently there is demand for the development of efficient synthetic routes to access these valuable chemicals. Biocatalytic routes have been developed for the synthesis of chiral amines using transaminases,^[3,4] ammonia lyases,^[5,6] amine dehydrogenases,^[7] monoamine oxidases^[8,9] and imine reductases.^[10–12] Recently, the toolbox of enzymes for amine synthesis has been expanded by the discovery of an NADP(H)-dependent reductive aminase (RedAm) from *Aspergillus oryzae* (AspRedAm). This enzyme is capable of catalysing not only the reduction of preformed imines, but also the coupling of amines and ketones followed by the subsequent reduction of the imine intermediate.^[13] All of these enzymes require electrophilic centres to insert nitrogen atoms, whereas a more elegant route would be the direct replacement of an unactivated C-H bond with a C-N bond.

In organic chemistry, C-H activation is typically enabled by metal complexes and requires the presence of auxiliary groups.^[14–20] Rare enzymatic counterparts have been found in nature to carry out this attractive chemistry,^[21] and only recently heme-proteins have been engineered to perform C(sp³)-H amination reactions. This biocatalytic amination reaction was achieved through the introduction of mutations at residues essential for the native mono-oxygenation activity of a cytochrome P450 or through the installation of synthetic metal complexes in place of the natural heme.^[22–27] Even though these studies have extended the potential for biocatalytic C-H functionalization, remarkably these novel catalysts either rely on pre-functionalized molecules possessing both the nitrogen functionality and the target C-H bond (intramolecular C-H amination), or on pre-activated aminating agents (*e.g.*, tosyl azides).

Oxidative enzymes have previously been successfully combined with amination biocatalysts to functionalize C(sp³)-H bonds,^[28,29] alkenes^[30] or to convert alcohols to amines.^[31,32] Moreover, biocatalytic redox-neutral hydrogen-borrowing processes have been demonstrated for the conversion of alcohols into the corresponding primary amines.^[33,34,35] Thus far, these studies are limited to the synthesis of primary amines due to the nature and specificity of the amination biocatalyst. However, recently we have demonstrated the synthesis of aliphatic and aromatic secondary amines starting from primary and secondary alcohols *via* hydrogen-borrowing employing a single alcohol dehydrogenase (ADH) and the newly characterized *AspRedAm*.^[36]

Herein, we report a new biocatalytic cascade for the conversion of cycloalkanes to secondary amines employing a cytochrome P450 mono-oxygenase for the first oxidation step followed by amination using an ADH in combination with a RedAm (Scheme 1). As a model substrate, we selected cyclohexane **1** for P450-catalyzed hydroxylation to the corresponding alcohol cyclohexanol **2** followed by ADH mediated oxidation to cyclohexanone **3** and subsequent amination with *AspRedAm*.



Scheme 1. Biocatalytic cascade for the amination of cycloalkanes.

5.4 Results and Discussion

For the initial C-H activation step, we selected the P450-BM3 mutant R47L/Y51F which was previously shown to have high activity towards cyclohexane.^[37] In particular, it was observed that the hydroxylation of cyclohexane improved when the reaction was carried out in a biphasic system, with the substrate acting as co-solvent.^[38] In order to uncouple the cofactor dependence of the two steps and avoid competition between the P450 and the RedAm for

reducing equivalents, the two amino acid substitutions R966D and W1046S were introduced in the reductase domain of the self-sufficient P450-BM3 to increase its affinity towards NADH.^[38,39] Cofactor regeneration for the oxyfunctionalisation step was performed by the formate dehydrogenase from *Candida boidinii* (CboFDH).^[40] Finally, for the amination step *via* hydrogen-borrowing we selected a variant of the ADH from *Thermoanaerobacter ethanolicus* (TeSADH W110A)^[41] and AspRedAm, as previously reported by our group.^[36]

In an attempt to exploit all the selected enzymes in a one-step process, we initially investigated the effect of amines **a–f** at different substrate concentrations on the performance of the P450, expressed as residual total turnover numbers (TTNs, Figure 1). The negative effect on P450 performance was most pronounced at 250 mM amine, with only small primary amines, such as ammonia **a** and methylamine **b**, being tolerated. At 50 mM amine, P450 hydroxylation activity was still significantly affected, especially with propargylamine **c**.

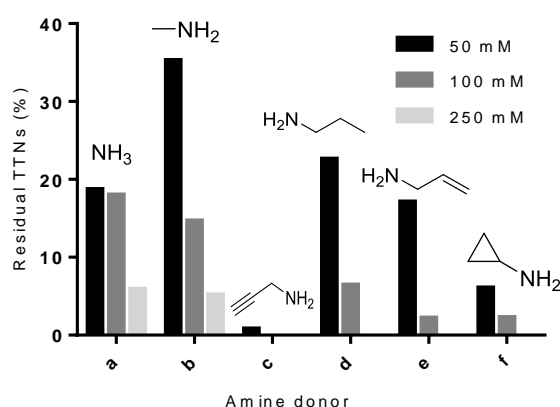


Figure 1. Effect of various amine donors selected for the amination cascade on the performance of the P450. See section 7.4.1.7.1 for full experimental conditions.

In the light of these initial results, we incubated crude enzyme preparations (the exception being AspRedAm, used in purified form) with **1** in the presence of 1 mM NAD⁺ and 1 mM NADP⁺, 250 mM sodium formate and 250 mM methylamine, but unfortunately no amine formation was detected. We speculated that the cofactors required for the hydrogen-borrowing step (NADP⁺/NADPH) were intercepted by endogenous *E. coli* enzymes, leading to imbalanced cofactor regeneration.^[33,34] Thus, the amount of lysate added for the

hydroxylation step was reduced and, gratifyingly, conditions were found to carry out the reaction as a one-step process (4.5 mM *N*-methylcyclohexylamine **1b** produced, Table S1, section 7.4.2). Next, we decided to extend the one-step cascade for the amination of **1** using a panel of six amine nucleophiles at different loadings (250, 100, 50, 25 and 10 mM, as reported in Table S1), guided by our initial investigation on the effect of the selected amines on hydroxylation activity (Figure 2 and Table S1). The outcome of the amination cascade was observed to be a trade-off between the inhibitory effect of the amine donor on the P450 and the reactivity of the same amine nucleophile in the *AspRedAm*-catalyzed reductive amination.^[13] 3.6 mM *N*-propargylcyclohexylamine **1c** and 2.1 mM *N*-cyclopropylcyclohexylamine **1f** were produced when propargylamine **c** and cyclopropylamine **f** were used as reacting partners in the one-step process, with 0.7 mM cyclohexanone **3** left. Interestingly, the highest product concentration was observed with **e** (8.3 mM), although a considerable amount of **2** (4.7 mM) and **3** (5.9 mM) remained unreacted. Moreover, 9.6 mM cyclohexanone **3** accumulated with just 1 mM NADP⁺ when **a** was employed as the amine donor, although only a low concentration of cyclohexylamine **1a** was produced (0.2 mM, Table S2). These results suggest that NADP⁺ regeneration might also be supported by the P450 (which can still accept the phosphorylated cofactor)^[38] and/or by other enzymes in the lysate.

To test the latter hypothesis, experiments with purified proteins were performed using **e** as the amine donor, and almost total conversion of **3** to **1e** was achieved (92%). This result indicates that endogenous enzymes in *E.coli* might compete with *AspRedAm* for reducing equivalents (Figure S1, section 7.4.2). Taking into account some of the inherent limitations associated with a one-step approach and the application of purified enzymes, we decided to explore the potential of the one-pot cascade as a two-step process with crude enzyme preparations. This approach made it possible to increase the enzyme loading in the first step without impacting on the presence of the amine donor for the second step (Table S2, section 7.4.2). With this setup, production of *N*-propargylcyclohexylamine **1c** reached 19.6 mM, and good results were also obtained with allylamine **e** and cyclopropylamine **f** (15.4 and 17.1 mM product, respectively), with **3** accounting for 19%, 35% and 16% of total products formed (Figure 2 and Table S2, section 7.4.2). Less pronounced improvements were achieved when amines **a**, **b** and **d** were employed in the two-step process, with 50%-80% products represented by **3**.

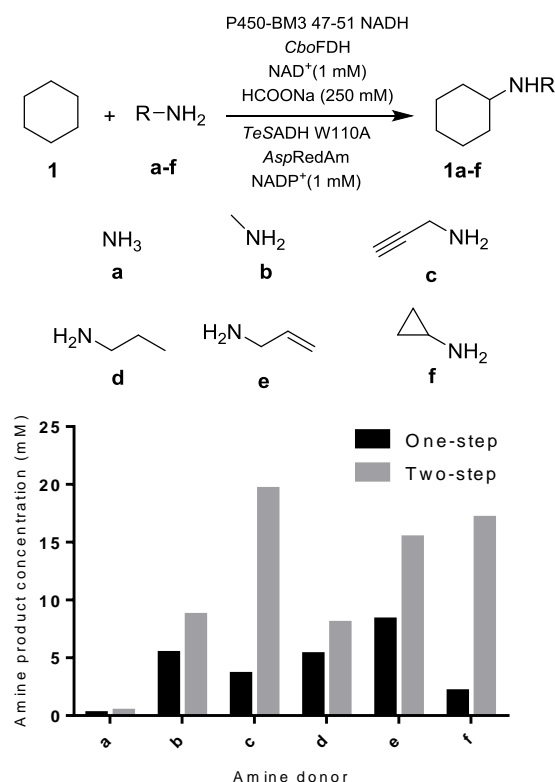


Figure 2. One-step and two-step cascade for the amination of cyclohexane using a panel of primary amines. See section 7.4.1.7.3 for full experimental conditions.

The P450 variant employed in this study can accept a variety of cycloalkanes^[37] and, similarly, TeSADH has already been employed with a range of cyclic alcohols.^[39] AspRedAm can also catalyse the reductive amination of cyclopentanone **8**, albeit less efficiently than that of **3**.^[13] Thus, we explored the possibility of expanding the scope of the cascade to other cycloalkanes (Figure 3 and Table S3, section 7.4.2). When employing amine donors **c**, **e** and **f**, product formation was detected when starting from cyclopentane **4** and cycloheptane **5**. For cyclooctane **10**, even though product formation was detected when AspRedAm was incubated with cyclooctanone **11** and **e** (40 % conversion, Figure S2, section 7.4.2), unfortunately, the multi-step amination of **10** was unsuccessful.

When **c** was reacted with **4**, 10.6 mM *N*-propargylcyclopentylamine **4c** was produced (50% of total products formed), whereas **e** and **f** gave only 2.2 mM *N*-allylcyclopentylamine **4e** and 3.0 mM *N*-cyclopropylcyclopentylamine **4f**, with the accumulation of the intermediate oxidized products, which represented 85 and 79% of total products formed, respectively. Modest

product concentrations and an excess cycloheptanone **9** (61% total products formed) were obtained when the bulkier cycloheptane **5** was used, with 3.3 mM *N*-propargylcycloheptylamine **5c** produced with **c** and no product detected with both **e** and **f**. We speculate that the reactivity of *AspRedAm* towards specific carbonyl acceptors and amine partners influences the outcome of the amination cascade both in terms of yield and conversion of the intermediate ketone to the amine.

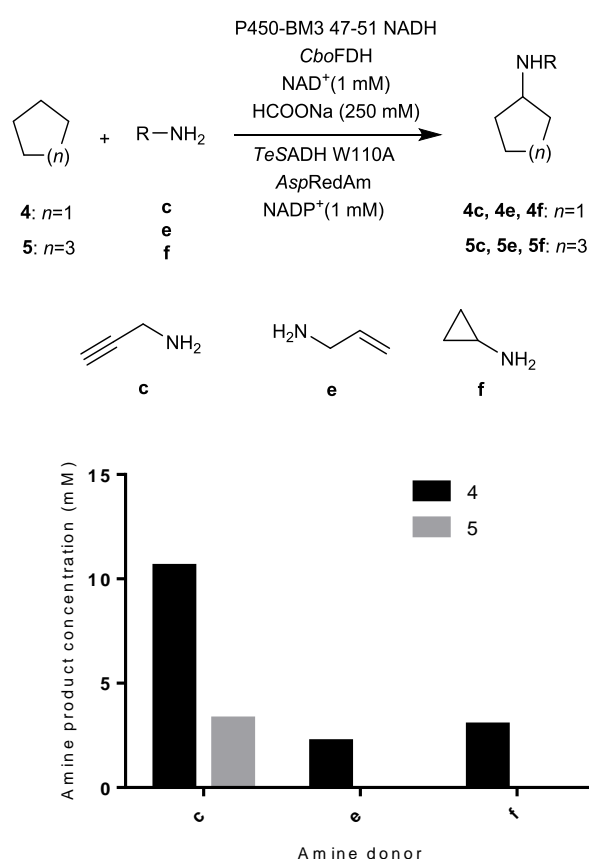


Figure 3. Two-step cascade for the amination of cyclopentane and cycloheptane. See section 7.4.1.7.3 for full experimental conditions.

To examine the scalability of the process, we selected the two-step cascade for the amination of cyclohexane with **c** to form **1c**. A 2-fold improvement in TTN for the hydroxylation step was found when the reaction was carried out in a flask with orbital shaking instead of a round-bottom flask with magnetic stirring (5707 vs. 2976), probably due to a better mass-transfer and/or oxygenation (Figure S3, section 7.4.2).^[38] In a similar work, higher but comparable TTN

values (9620) were obtained when the NADPH-dependent P450-BM3 R47L/Y51F was employed for cyclohexane hydroxylation.^[38] Even if it is difficult to identify the reasons behind the higher observed TTN, the better performance of the biocatalyst may be linked to the extended reaction time achieved through the addition of stabilising agents.

Following the first C-H activation stage, the RedAm-catalysed amination step yielded 50 mg of the final product **1c** giving a space-time yield of 2 g L⁻¹ d⁻¹ (Figure S4, section 7.4.2).

In summary, we have demonstrated a biocatalytic cascade for the preparation of a set of secondary amines from cycloalkanes. Unlike chemical approaches for the addition of amines across aliphatic C-H bonds, the presented cascade is carried out in aqueous buffer, does not produce any toxic by-products and does not require additional deprotection steps to obtain the final amine product. The plethora of information available for oxidoreductases, combined with the broad substrate tolerance of reductive aminases, will allow for a larger application of the presented cascade to access a variety of different amines starting from unfunctionalized compounds.

5.5 References

- [1] D. Ghislieri, N. J. Turner, *Top. Catal.* 2014, 57, 284–300.
- [2] T. C. Nugent, M. El-Shazly, *Adv. Synth. Catal.* 2010, 352, 753–819.
- [3] J. L. Galman, I. Slabu, N. J. Weise, C. Iglesias, F. Parmeggiani, R. C. Lloyd, N. J. Turner, *Green Chem.* 2017, 19, 361–366.
- [4] C. K. Savile, J. M. Janey, E. C. Mundorff, J. C. Moore, S. Tam, W. R. Jarvis, J. C. Colbeck, A. Krebber, F. J. Fleitz, J. Brands, et al., *Science*, 2010, 329, 305–309.
- [5] F. Parmeggiani, S. L. Lovelock, N. J. Weise, S. T. Ahmed, N. J. Turner, *Angew. Chemie* 2015, 54, 4608–4611.
- [6] N. J. Weise, F. Parmeggiani, S. T. Ahmed, N. J. Turner, *J. Am. Chem. Soc.* 2015, 137, 12977–12983.
- [7] T. Knaus, W. Böhmer, F. G. Mutti, *Green Chem.* 2017, 19, 453–463.
- [8] D. Ghislieri, A. P. Green, M. Pontini, S. C. Willies, I. Rowles, A. Frank, G. Grogan, N. J. Turner, *J. Am. Chem. Soc.* 2013, 135, 10863–10869.
- [9] R. S. Heath, M. Pontini, B. Bechi, N. J. Turner, *ChemCatChem* 2014, 6, 996–1002.
- [10] Z. Maugeri, D. Rother, *J. Biotechnol.* 2017, 258, 167–170.
- [11] P. Matzel, M. Gand, M. Höhne, *Green Chem.* 2017, 19, 385–389.
- [12] J. Mangas-Sanchez, S. P. France, S. L. Montgomery, G. A. Aleku, H. Man, M. Sharma, J. I. Ramsden, G. Grogan, N. J. Turner, *Curr. Opin. Chem. Biol.* 2017, 37, 19–25.
- [13] G. A. Aleku, S. P. France, H. Man, J. Mangas-Sanchez, S. L. Montgomery, M. Sharma, F. Leipold, S. Hussain, G. Grogan, N. J. Turner, *Nat. Chem.* 2017, 9, 961–969.

- [14] C. Liang, F. Collet, F. Robert-Peillard, P. Mu, R. H. Dodd, P. Dauban, I. De Chimie, A. V. De Terrasse, P. Müller, R. H. Dodd, et al., *J. Am. Chem. Soc.* 2008, 130, 343–350.
- [15] D. N. Zalatan, J. Du Bois, *Top. Curr. Chem.* 2010, 292, 347–378.
- [16] M.-L. Louillat, F. W. Patureau, *Chem. Soc. Rev.* 2014, 43, 901–910.
- [17] K. M. Waltz, J. F. Hartwig, *Science*. 1997, 277, 211–213.
- [18] J. L. Roizen, M. E. Harvey, J. D. U. Bois, *Acc. Chem. Res.* 2012, 45, 911–922.
- [19] D. N. Zalatan, J. Du Bois, J. Du Bois, *J. Am. Chem. Soc.* 2008, 130, 9220–9221.
- [20] J. A. Labinger, J. E. Bercaw, *Nature* 2002, 417, 507–514.
- [21] P. A. Frey, O. T. Magnusson, *Chem. Rev.* 2003, 103, 2129–2148.
- [22] C. C. Farwell, R. K. Zhang, J. A. McIntosh, T. K. Hyster, F. H. Arnold, *ACS Cent. Sci.* 2015, 1, 89–93.
- [23] P. Dydio, H. M. Key, H. Hayashi, D. S. Clark, J. F. Hartwig, *J. Am. Chem. Soc.* 2017, 139, 1750–1753.
- [24] J. A. McIntosh, P. S. Coelho, C. C. Farwell, Z. J. Wang, J. C. Lewis, T. R. Brown, F. H. Arnold, *Angew. Chemie - Int. Ed.* 2013, 52, 9309–9312.
- [25] R. Singh, J. N. Kolev, P. A. Sutura, R. Fasan, *ACS Catal.* 2015, 5, 1685–1691.
- [26] C. K. Prier, R. K. Zhang, A. R. Buller, S. Brinkmann-Chen, F. H. Arnold, *Nat. Chem.* 2017, 9, 629–634.
- [27] P. R. Ortiz de Montellano, *Cytochrome P450: Structure, Mechanism, and Biochemistry*, Springer Science & Business Media, 2015.
- [28] M. Schrewe, N. Ladkau, B. Bühler, A. Schmid, *Adv. Synth. Catal.* 2013, 355, 1693–1697.
- [29] P. Both, H. Busch, P. P. Kelly, F. G. Mutti, N. J. Turner, S. L. Flitsch, *Angew. Chemie - Int. Ed.* 2015, 55, 1511–1513.
- [30] S. Wu, Y. Zhou, T. Wang, H.-P. Too, D. I. C. Wang, Z. Li, *Nat. Commun.* 2016, 7, 11917.
- [31] M. Pickl, M. Fuchs, S. M. Glueck, K. Faber, *ChemCatChem* 2015, 7, 3121–3124.
- [32] L. Martínez-montero, V. Gotor, V. Gotor-Fernández, I. Lavandera, *Green Chem.* 2017, 19, 474–480.
- [33] K. Tauber, M. Fuchs, J. H. Sattler, J. Pitzer, D. Pressnitz, D. Koszelewski, K. Faber, J. Pfeffer, T. Haas, W. Kroutil, *Chem. - A Eur. J.* 2013, 19, 4030–4035.
- [34] F. G. Mutti, T. Knaus, N. S. Scrutton, M. Breuer, N. J. Turner, *Science*. 2015, 349, 1525–1529.
- [35] M. Thompson, N. J. Turner, M. Thompson, *ChemCatChem* 2017, 9, 3833–3836.
- [36] S. L. Montgomery, J. Mangas-Sanchez, M. P. Thompson, G. A. Aleku, B. Dominguez, N. J. Turner, *Angew. Chemie - Int. Ed.* 2017, 56, 10491–10494.
- [37] A. Pennec, C. L. Jacobs, D. J. Opperman, M. S. Smit, *Adv. Synth. Catal.* 2014, 357, 118–130.
- [38] S. C. Maurer, K. Kühnel, L. A. Kaysser, S. Eiben, R. D. Schmid, V. B. Urlacher, *Adv. Synth. Catal.* 2005, 347, 1090–1098.
- [39] A. A. Pennec, F. Hollmann, M. S. Smit, D. J. Opperman, *ChemCatChem* 2015, 7, 236–239.
- [40] H. Slusarczyk, S. Felber, M. R. Kula, M. Pohl, *Eur. J. Biochem.* 2000, 267, 1280–1289.
- [41] M. M. Musa, R. S. Phillips, M. Laivenieks, C. Vieille, M. Takahashi, S. M. Hamdan, *Org. Biomol. Chem.* 2013, 11, 2911.

Chapter 6. General Discussion and Conclusions

The development of methods for C-H functionalization is a matter of central importance for the chemical industry. Broad substrate scope, safe, sustainable and selective synthetic technologies are greatly required. In this context, biocatalysis with P450 enzymes is extremely attractive. Despite decades of research in academia, the major limitations associated with P450 catalysis still need to be overcome.

Cytochrome P450-BM3 is one of the best characterized P450s and represents an ideal candidate for biotechnological applications, given its self-sufficiency, fast reaction rates and wealth of literature available. However, the self-sufficient class VII P450s are gaining increasing attention for their broader substrate scope compared to wild-type P450-BM3. At the beginning of these doctoral studies, only five class VII P450s had been characterized and thus our primary goal was to add diversity to this class of enzymes. Considering the issues associated with P450 stability, genomes of thermophilic organisms were investigated and five new class VII P450s were added to the panel of enzymes available (Chapter 2). Interestingly, homologues from thermophilic organisms were also found when P450-BM3 was used as a query sequence for the initial BLAST searches (section 7.1.1.2), indicating that there is still room for adding naturally thermostable enzymes to the array of class VIII enzymes reported. The scalability of the whole-cell hydroxylation of diclofenac by the newly discovered P450-AX (section 2.4.5) indicates that genomic data available in databases should be fully exploited to get access to new or improved enzymatic activities. However, protein and process engineering are still needed to speed-up the implementation of P450s for biocatalytic processes. As a matter of fact, although P450-AX proved to be a promising candidate for the preparative scale synthesis of 5-hydroxydiclofenac, process metrics are still far from being met (J. Lima-Ramos, P. Tufvesson, J. M. Woodley, *Green Process. Synth.* 2014, 3, 195–213). Investigation of oxygen limitations led to modest improvements of relevant metrics and hence protein engineering is needed to increase biocatalyst activity (section 2.6).

The determination of the first class VII P450 structure provides a precious guideline for future and ongoing mutagenesis studies (Chapter 3 and section 3.5). The structure itself does not add much to our knowledge on P450 enzymes, but, along with previous research on class VII P450s, it is expected to promote a wider use and engineering of these enzymes towards industrial applications.

The challenging chemistry carried out by P450s means that there is great potential to use these enzymes to enable biocatalytic cascades starting from cheap materials. However, additional complications must be considered when combining enzymes in one-pot, such as compatibility problems. The realization of the one-pot two-step biocatalytic synthesis of ketoisophorone was initially hindered by the acidification of the medium employed for the growth of the P450 biocatalyst (section 4.4.3). Eventually, bottlenecks were overcome and the cascade was demonstrated both as a one-pot two-step process and as an *in-vivo* cascade, with a productivity of up to 1.4 g L⁻¹ d⁻¹. Looking towards industrial implementation of the process, it needs to be considered that target process metrics for ketoisophorone synthesis are expected to be higher than for drug metabolites. Again, improvements are especially needed for the P450 reaction in terms of protein expression, activity and substrate loadings. Considering the great diversity of ADHs available from various suppliers, it is likely that a more extensive enzyme screening might reveal better biocatalysts than Cm-ADH10 for the oxidation of 4-hydroxy-isophorone to ketoisophorone. Moreover, cosubstrate selection should be inspired by an overall economic evaluation of the process.

Unsurprisingly, the combination of four enzymes for the amination of unfunctionalized cycloalkanes proved to be even more challenging (Chapter 5). Here limitations such as the inhibition of the P450 by the amine donor used in the reductive amination step and disruption of the internal cofactor balancing by endogenous *E. coli* enzymes, reduced the overall performance of the one-pot reaction. Similar complications had to be addressed in a comparable cascade for the synthesis of cyclohexylamine published after our work (H. L. Yu, T. Li, F. F. Chen, X. J. Luo, A. Li, C. Yang, G. W. Zheng, J. H. Xu, *Metab. Eng.* 2018, 47, 184–189). In principle, the cascade presented in this thesis could be adapted to generate valuable chiral amines or could be adapted to include a P450-alcohol oxidase step to generate a set of different amine products.

Biocatalytic cascades are still in their infancy and their industrial application will certainly be complicated by their multistep nature. Nonetheless, enzymatic cascades remain a very interesting approach to build-up complex molecules from cheap starting materials in one-pot. As shown in this work, a case-by-case optimization might be needed along with fine-tuning of the selected biocatalysts for the specific task to be carried out in the multi-step process.

Chapter 7. Supporting Information

7.1 A Panel of New Thermostable CYP116B Self-Sufficient Cytochrome P450 Monooxygenases Catalysing C-H Activation with Diverse Substrate Scope

7.1.1 Experimental

7.1.1.1 Chemicals

Solvents, chemicals and carbon monoxide for CO difference spectroscopy were obtained from Sigma-Aldrich (Dorset, UK) unless otherwise specified. Gases for GC-FID analysis were purchased from BOC gases (Guildford, UK) and competent cells and enzymes for molecular biology were purchased from New England Biolabs (Hitchin, UK).

7.1.1.2 Target Selection and Phylogenetic Analysis

In order to identify suitable class VII or class VIII enzymes, P450 RhF or P450 BM3 were used as the query sequence to perform BLAST searches with BLASTP, PSI-BLAST and tBLASTn functions at the NCBI website (<http://www.ncbi.nlm.nih.gov/BLAST/>) using standard parameters.

Table S1. Sequence information for BM3 homologues identified from thermophilic organisms.

<i>NCBI amino acid/nucleotide accession</i>	<i>Sequence I.D. to BM3</i>	<i>Organism</i>	<i>Cultivation temperature</i>	<i>Source</i>
ACX65034.1	60 %	<i>Geobacillus</i> sp. Strain Y412MC10	50 °C	Hot spring, Yellowstone national park, USA
XM_003653681	35 %	<i>Thielavia terrestris</i> NRRL 8126	37-45 °C	Soil
XM_003663599	35 %	<i>Myceliophthora thermophila</i> ATCC 42464	37-45 °C	Paper compost
XM_002381902	34 %	<i>Aspergillus flavus</i> NRRL 3357	25-28 °C	Peanut cotyledons, USA

Note: cultivation conditions for microorganisms were obtained from the ATCC catalogue or from literature^[1]

Sequence alignments were performed using Clustal Omega. The evolutionary analyses were conducted in MEGA7 with sequences aligned using ClustalW within the MEGA7 software and evolutionary history was inferred using the UPGMA method.^[2] The evolutionary distances were calculated using the Poisson correction method and are in units of number of amino acids substitutions per site. Molecular weights and amino acid compositions of P450 RhF and its homologues were calculated using the ProtParam program on the ExpASy proteomic server (<http://web.expasy.org/protparam/>).

7.1.1.3 Gene Cloning

Genomic DNA for the identified thermophilic microorganisms was obtained from DSMZ (Braunschweig, Germany). Gene targets were amplified using typical PCR conditions with the primers listed in Table S3. PCR products were gel purified and cloned into the pET28a vector using the In-Fusion[®] PCR Cloning kit from Clontech (California, USA) following the manufacturer's instructions. Primer synthesis and DNA sequencing were performed by Eurofins Genomics (Ebersberg, Germany).

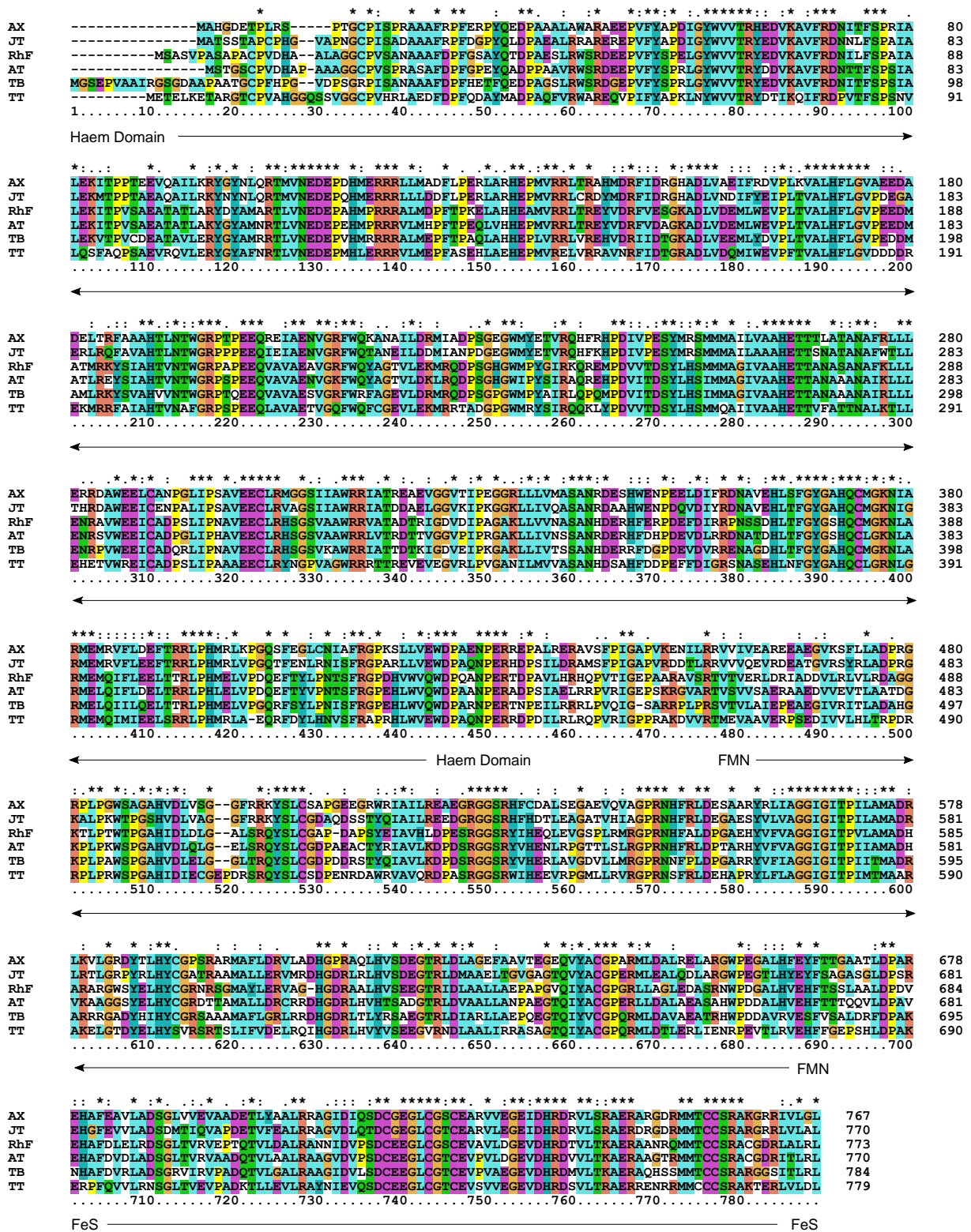


Figure S1. Sequence alignment of the full-length P450 RhF with homologues AT, AX, JT, TB and TT using Clustal Omega.

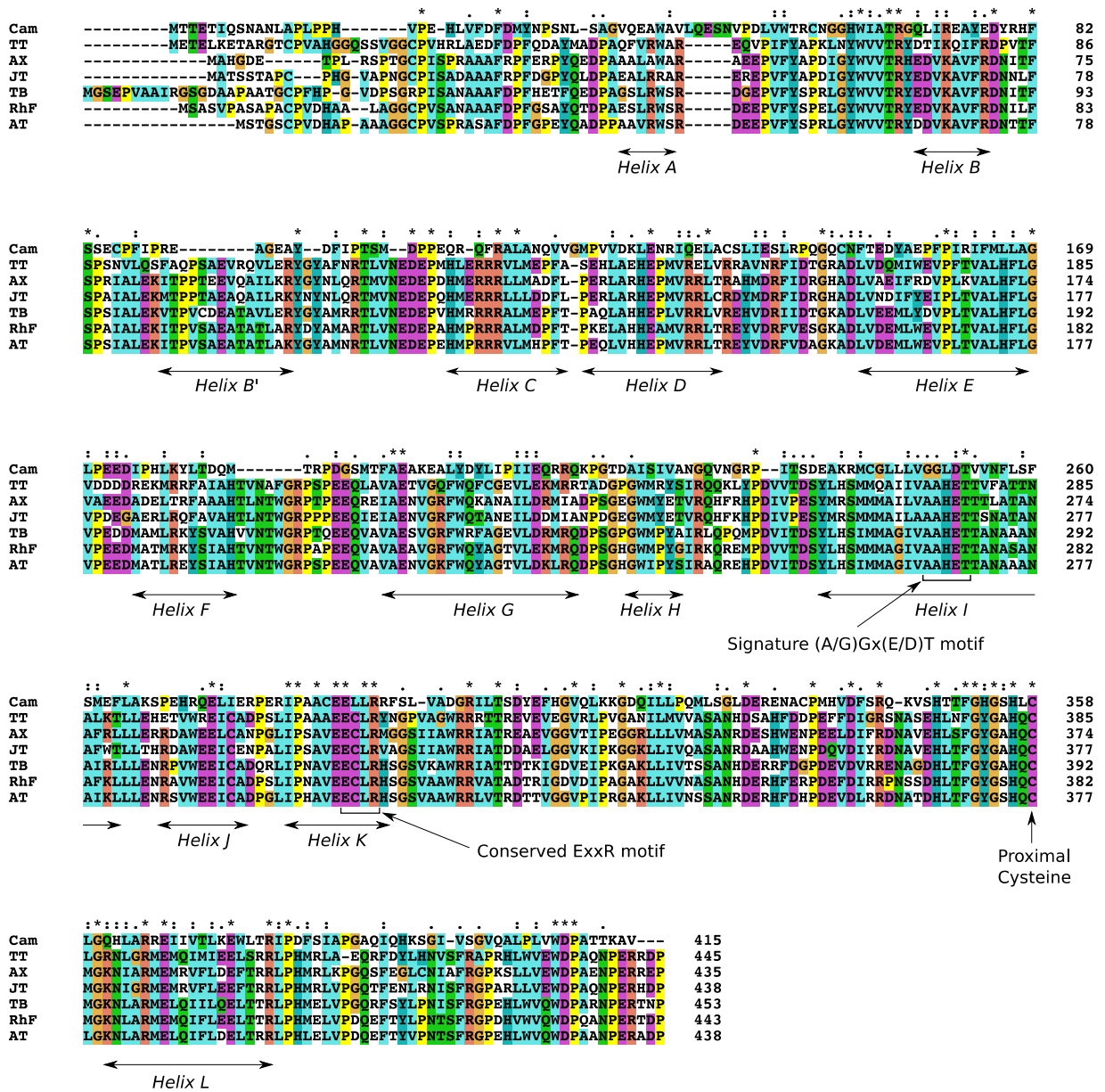


Figure S2. Sequence alignment of the heme domains with the class I P450cam (CYP101A1) with secondary structural features and signature P450 motifs highlighted based on the crystal structure of P450cam.^[3]

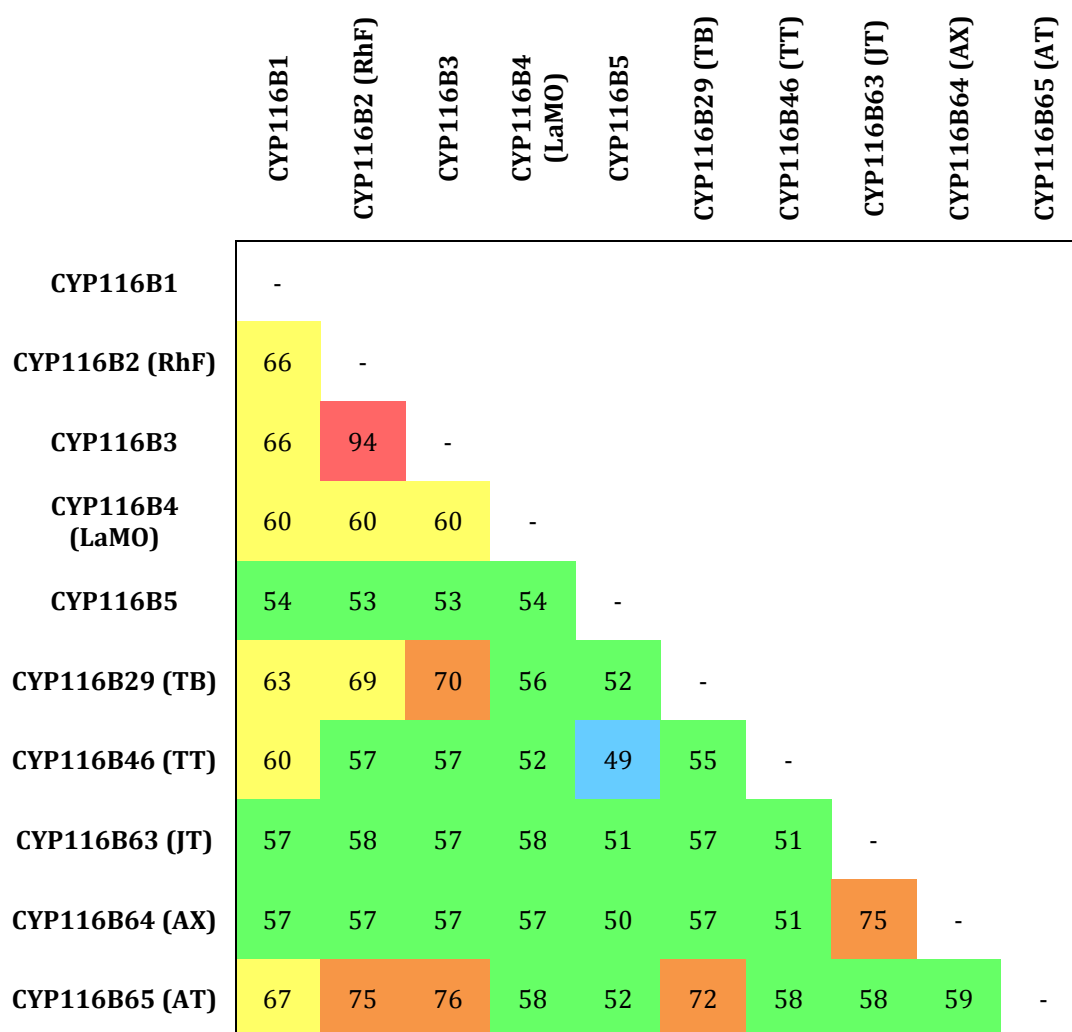
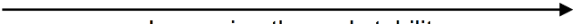


Figure S3. Heat map highlighting the % sequence identities between class VII CYP116B enzymes including the newly characterized panel of AT, AX, JT, TB and TT. Where red is > 90 %, dark orange is 80-90 %, orange is 70-80 %, yellow is 60-70 %, green is 50-60 % and blue is < 50 % sequence identity.

Table S2. Amino acid composition (%) of P450 RhF and homologues JT, AT, TB, AX and TT

Amino acid		RhF	JT	AT	TB	AX	TT	
Non polar, aliphatic	Glycine	G	6.5	6.5	7.5	7.7	7.4	5.4
	Alanine	A	12.4	11.3	10.8	12.3	10.5	8.1 ←
	Proline	P	6.6	6.9	5.8	5.7	6.9	6.2
	Valine	V	7.8	8.7	6.1	6.0	7.8	8.7
	Leucine	L	8.9	8.8	8.8	9.0	8.4	8.3
	Isoleucine	I	2.8	3.0	4.3	4.4	4.7	3.9 →
	Methionine	M	3.0	1.8	3.2	3.0	2.8	3.0
Polar, uncharged	Serine	S	5.3	4.9	3.5	3.8	4.7	4.7
	Threonine	T	5.0	6.4	4.9	3.8	4.8	4.5
	Cysteine	C	1.8	2.1	1.8	1.8	1.7	1.9
	Asparagine	N	2.6	2.2	2.9	2.1	2.2	2.8
	Glutamine	Q	2.1	2.1	2.6	1.6	2.7	3.5
Aromatic	Phenylalanine	F	2.6	2.3	3.2	3.8	2.7	3.7 →
	Tyrosine	Y	2.7	2.6	2.7	2.0	2.4	2.4
	Tryptophan	W	1.7	1.6	1.4	1.6	1.4	1.5
Basic	Arginine	R	8.2	8.3	10.5	11.5	10.5	11.4 →
	Lysine	K	1.6	1.9	1.4	1.7	1.7	1.5
	Histidine	H	4.0	4.3	3.4	3.1	3.3	3.5
Acidic	Aspartic acid	D	6.9	7.4	7.1	5.5	6.9	5.9
	Glutamic acid	E	7.6	6.9	7.8	9.8	6.6	9.0



 Increasing thermal stability

Table S3. Primers sequences used for cloning genes from genomic DNA.

Target	Template		Primer sequence (5' -> 3')	Restriction enzymes
AT	DSM 44574	Forward	CGCGCGGCAGCCATATGAGCACAGGGTCGTGCC	NdeI/XhoI
		Reverse	GGTGGTGGTGCTCGATCAGAGCCGAGGGTGATGC	
AX	DSM 24422	Forward	CGCGCGGCAGCCATATGGCCACGGTGACGAAACC	NdeI/XhoI
		Reverse	GGTGGTGGTGCTCGATCAGAGCCCGAGCAGC	
JT	DSM 23413	Forward	CGCGCGGCAGCCATATGGCCACCGCTCCACCG	NdeI/XhoI
		Reverse	GGTGGTGGTGCTCGATCAGAGCGCCAGCACCAGC	
TB	DSM 43833	Forward	AATGGGTCGCGGATCCATGGGCTCAGAGCCGG	BamHI/SalI
		Reverse	CCGCAAGCTTGTCTGACTCAGAGCGCAGGGTGATG	
TT	DSM 27220	Forward	AATGGGTCGCGGATCCATGGAACTGAACTCAAGGAAACGG	BamHI/SalI
		Reverse	CCGCAAGCTTGTCTGACTTACAGATCGAGTACCAGACGCTCG	

Note: 5' overhangs designed for In-Fusion cloning are shown in red text. The templates used for amplification and the restriction enzymes employed to cut the receiving vector are also indicated.

7.1.1.4 Protein Expression and Purification

Chemically competent *Escherichia coli* BL21(DE3) cells were transformed by heat shock with a pET28a plasmid bearing the P450 gene sequences according to manufacturer instructions. Transformants were plated on LB agar supplemented with kanamycin (50 µg/ml) and grown at 37 °C for 16 h. Single colonies were used to inoculate starter cultures of LB medium (8 mL) supplemented with antibiotic which were grown overnight. The starter cultures were used to inoculate 2 L baffled flasks containing minimal media (800 mL) consisting of 1x M9 salts, 0.4 % glucose, 0.05 % FeCl₃, 1 mM MgSO₄ and 1 mM CaCl₂. Cultures were grown at 37 °C to an OD₆₀₀ of 0.6 - 0.8 then cooled to the expression temperature of 20 °C and induced with β-D-1-thiogalactopyranoside (IPTG) (0.4 mM) and supplemented with 5-aminolevulinic acid (5-ALA). Protein expression was carried out at 20 °C with shaking at 200 rpm for 20 h. SDS-PAGE analysis was employed to confirm protein production. Cells were harvested by centrifugation (2831 g, 20 min, 4 °C) using a Beckman Coulter Avanti J-E centrifuge, cell pellets were washed with 0.2 M potassium phosphate buffer (KPi) and stored at -20 °C.

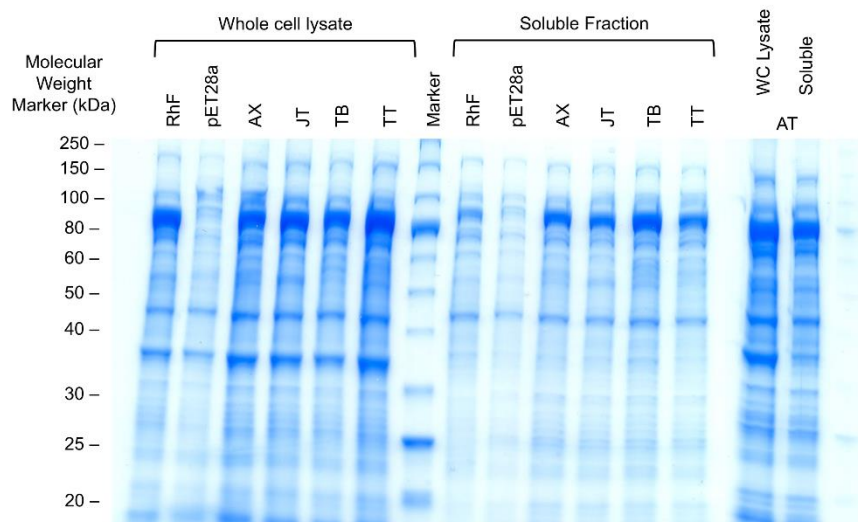


Figure S4. SDS-PAGE of the whole cell lysate and soluble fractions for cells expressing P450s

Table S4. Expression comparison between RhF and the new thermophilic P450 panel.

<i>P450</i>	<i>nmol_{P450} g_{cdw}⁻¹</i>	<i>Fold improvement (compared to RhF)</i>
<i>RhF</i>	56.1	
<i>AT</i>	223.5	4.0
<i>AX</i>	151.1	2.7
<i>JT</i>	61.0	1.1
<i>TB</i>	144.2	2.6
<i>TT</i>	315.5	5.6

Protein purification was done by immobilized metal ion affinity chromatography (IMAC) and all steps were carried out at 4 °C. Cells were resuspended in 90% buffer A (0.1 M Tris, 0.3 M NaCl, pH 8) and 10% buffer B (0.1 M Tris, 0.3 M NaCl, 0.3 M imidazole, pH 8) to a wet cell concentration of 200 mg/mL. Cells were lysed by ultrasonication (Bandelin Sonopuls sonicator) with 20 cycles of 15 s on and 45 s off, 40 % amplitude. Cell debris was removed by ultracentrifugation (48384 g, 30 min, 4 °C). The supernatant was filtered through a 0.20 µm filter then loaded onto a pre-equilibrated 5 mL HisTrap FF column (GE Healthcare). The target protein was eluted following series of step-wise washes of 30 mM and 60 mM imidazole and then eluted with 300 mM imidazole on a ÄKTA Pure system (GE Healthcare) with a flow rate

of 2 mL min⁻¹ as described in Table S5. Enzyme elution was monitored at 280 nm and appropriate fractions were combined concentrated using a 30,000 molecular weight cut-off filter (Vivaspin column, GE Healthcare) and desalted in 0.1 M Tris, pH 8. SDS-PAGE analysis of the purified samples showed purification of the full-length constructs (Figure S5). Determination of P450 concentration was performed on a Cary 50 UV/Visible spectrophotometer (Agilent Technologies, Santa Clara, USA) according to the protocol by Omura and Sato,^[4] using an extinction coefficient of $\epsilon_{450-490} = 91 \text{ mM}^{-1} \text{ cm}^{-1}$ for the optical difference spectrum between the reduced-CO bound and reduced spectra. Absorbance spectra of the purified proteins are shown in Figure S6. Typically, purified enzymes were diluted to a final concentration of 2 mg mL⁻¹ and aliquots were snap-frozen in liquid nitrogen until further use (no more than 2 weeks).

Table S5. ÄKTA Pure conditions for protein purification.

Step	% Buffer B	Column volumes
Wash 1	10	5
Wash 2	20	5
Elution	100	10
Re-equilibration	0	5

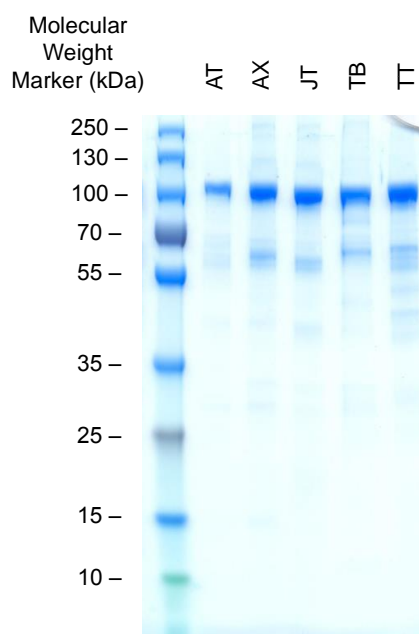


Figure S5. SDS-PAGE of P450s following IMAC purification

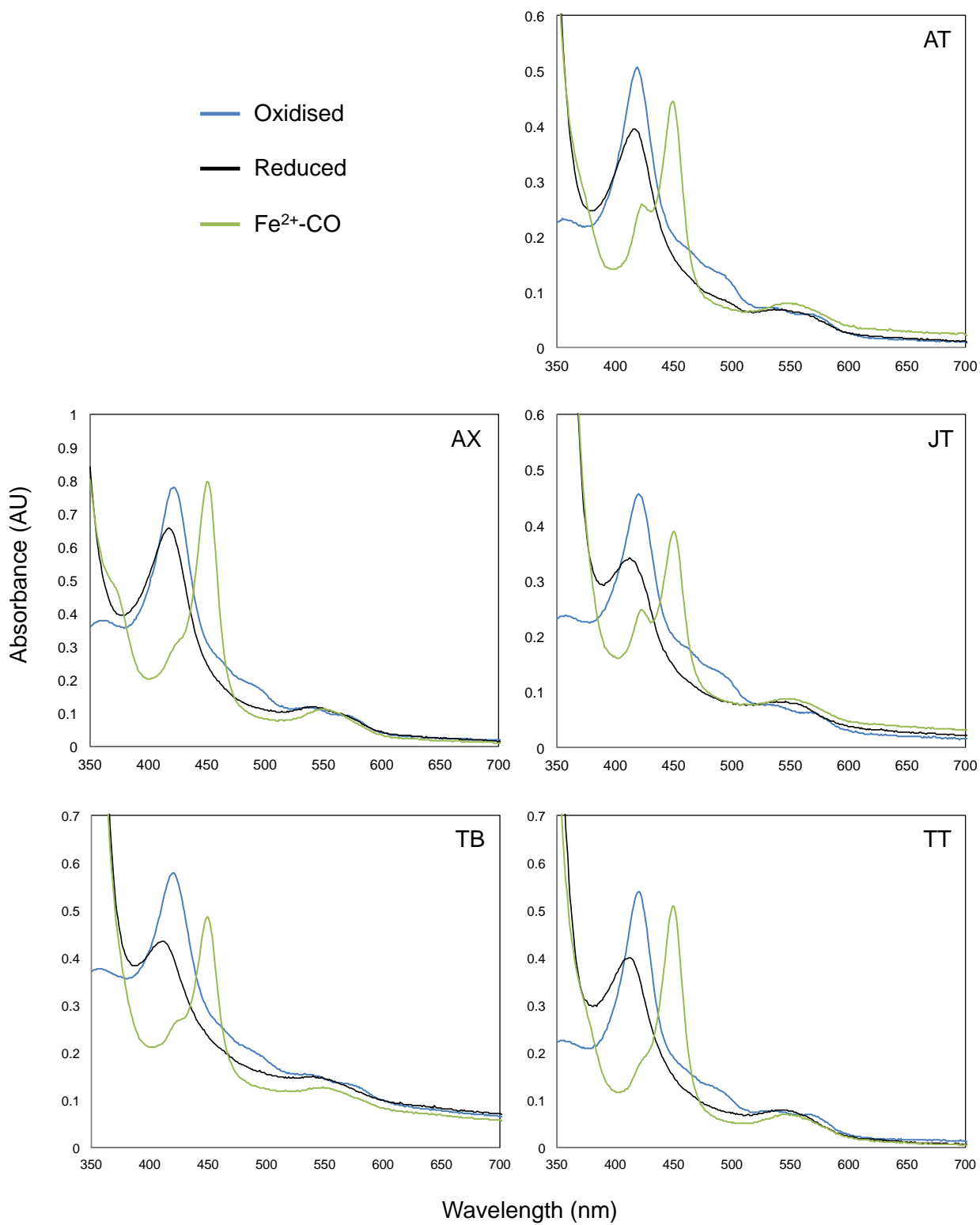


Figure S6. Absorption spectra of purified P450s and their ferrous-carbon monoxide bound complexes

7.1.1.5 Cytochrome *c* Assay

Cofactor preference was determined using a standard cytochrome *c* assay.^[5] Briefly, electron transfer from the reductase domain to the external electron acceptor (50 μM final concentration) was carried out in 0.1 M Tris, pH 8, with 9 nM enzyme and cofactor concentrations varying across the appropriate ranges (0.5 - 300 μM for NADPH and 0.06 - 2.4 mM for NADH). The change in absorbance was monitored at 550 nm ($\Delta\epsilon_{550} = 28,000 \text{ M}^{-1} \text{ cm}^{-1}$) employing a Cary 60 UV-visible spectrophotometer. Measurements were performed in duplicate (NADH) or triplicate (NADPH) at 25 °C (Table S6) or 50 °C (Table S7). K_M and k_{cat} values were determined by fitting the data to the Michaelis-Menten equation by nonlinear regression of the initial cytochrome *c* reduction rate using GraphPad Prism 6 and the error is the standard deviation.

Table S6. Kinetic properties of AT, AX, JT, TB and TB using NADH and NADPH as reductants and cytochrome *c* as electron acceptor with temperature control at 25 °C

P450	Reductant	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($\text{M}^{-1} \text{s}^{-1}$)	Cofactor Preference (Ratio k_{cat}/K_M NADPH/NADH)
AT	NADH	1232 \pm 100	20.4 \pm 0.8	1.7 \pm 0.2 $\times 10^4$	495
	NADPH	8.4 \pm 0.7	69 \pm 2	8.2 \pm 0.7 $\times 10^6$	
AX	NADH	213 \pm 15	5.7 \pm 0.1	2.7 \pm 0.2 $\times 10^4$	282
	NADPH	1.1 \pm 0.1	8.3 \pm 0.1	7.5 \pm 0.7 $\times 10^6$	
JT	NADH	640 \pm 50	71 \pm 2	1.1 \pm 0.1 $\times 10^5$	117
	NADPH	5.4 \pm 0.4	70 \pm 2	1.3 \pm 0.1 $\times 10^7$	
TB	NADH	2.7 \pm 0.3	4.3 \pm 0.1	1.6 \pm 0.1 $\times 10^6$	n/a
	NADPH	< 1	5.6 \pm 0.7	n/a	
TT	NADH	482 \pm 50	28.8 \pm 0.1	6.0 \pm 0.6 $\times 10^4$	270
	NADPH	2.5 \pm 0.2	40.4 \pm 0.8	1.6 \pm 0.1 $\times 10^7$	

Table S7. Kinetic parameters for TT using NADPH as reductant and cytochrome *c* as electron acceptor with temperature control at 50 °C

P450	Reductant	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($\text{M}^{-1} \text{s}^{-1}$)
TT	NADPH	3.3 \pm 0.2	80 \pm 2	3.2 \pm 0.3 $\times 10^7$

7.1.1.6 Biotransformations

The lysate reactions displayed in Table 2 were conducted with a final lysate concentration equivalent to 200 mg mL⁻¹ wet cells, corresponding absorbance spectra and P450 concentrations of the lysates are given in Figure S7 and Table S8 respectively. Reactions were run in 0.2 M KPi + 0.4 % glycerol, pH 8 with D-glucose (10 mg mL⁻¹), glucose dehydrogenase (1 mg mL⁻¹) (CDX-901, Codexis) and NADP⁺ (1 mg mL⁻¹) to a final reaction volume of 400 μL. Substrates were added from 50 mM stocks in DMSO (2 % v/v final concentration). Reactions were typically carried out at 20 °C with 200 rpm orbital shaking, unless otherwise specified.

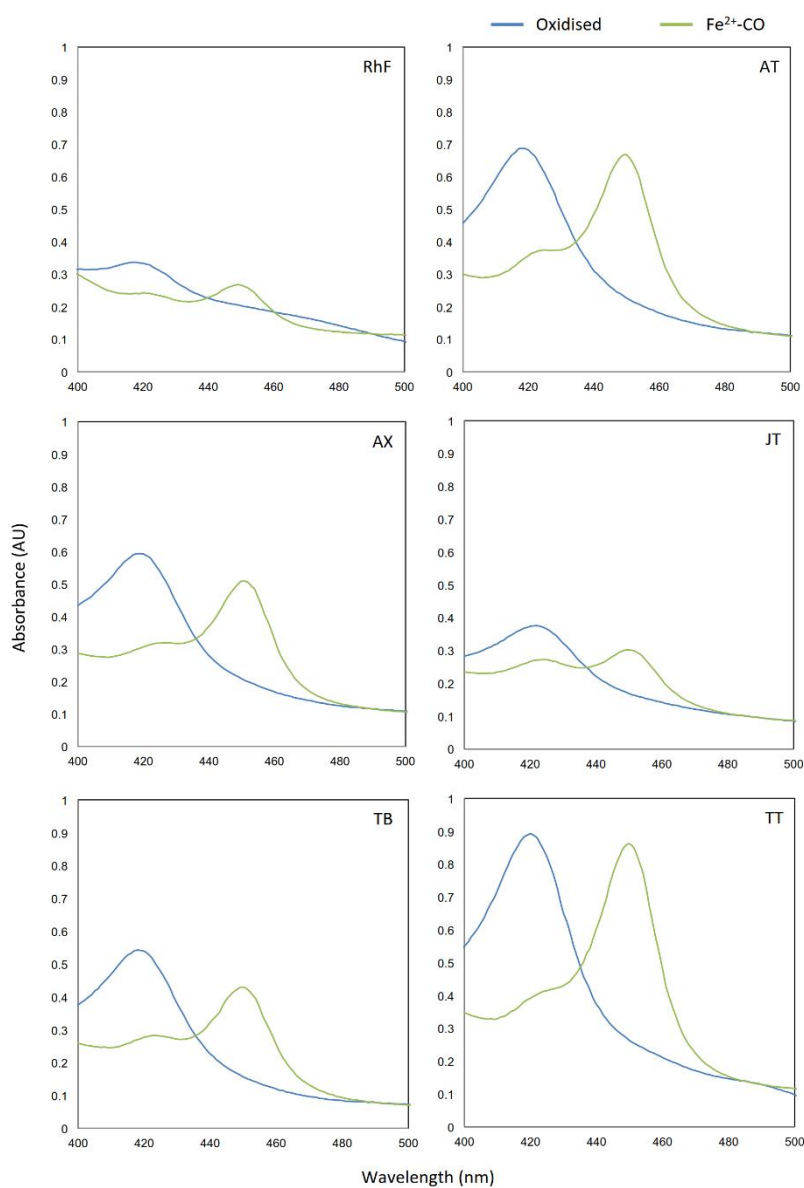


Figure S7. Absorption spectra of the cytochrome P450 soluble cell lysates and their ferrous-carbon monoxide bound complexes.

Table S8. P450 concentrations from cell lysates

P450	P450 Concentration in Lysates (μM)
RhF	1.8
AT	7.2
AX	4.8
JT	2.0
TB	4.6
TT	10.0

Note: P450 concentrations are from the lysates used in the biotransformations in Table 2, this is equivalent to the lysate obtained from 200 mg mL^{-1} wet cells.

Whole cell biotransformations with diclofenac were carried out with $440 \mu\text{L}$ of re-suspended cells (200 mg mL^{-1} final concentration) and the addition of glucose ($50 \mu\text{L}$ of 100 mg mL^{-1}). Substrate was added ($10 \mu\text{L}$) from an appropriate stock solution in DMSO (10 % v/v final concentration) to give the required substrate concentration. The reactions were conducted at $20 \text{ }^\circ\text{C}$ in deep 48-well plates with gas permeable seal or at $40 \text{ }^\circ\text{C}$ in 2 mL tubes, both with shaking at 180 rpm for 24 h.

HPLC was used to analyze biotransformations with diclofenac (**1**), 4-methoxyacetophenone (**3**), 2-hydroxy-4-methoxyacetophenone (**5**), 4-methoxybenzophenone (**7**), 4-methoxybenzotrile (**9**), 7-methoxycoumarin (**11**) and fluorene (**13**) substrates. GC-FID or GC-MS was used to analyze biotransformations with ethylbenzene (**15**), α -isophorone (**17**), methyl phenyl sulfide (**18**), styrene (**20**) and tetradecane (**22**) as substrates.

For HPLC analysis, equal volume of MeCN was added to the biotransformation. The sample was mixed and centrifuged. The supernatant was filtered for HPLC analysis. For GC analysis, the same volume of methyl tert-butyl ether (MTBE) was added to the biotransformation mixture, the organic layer was dried with MgSO_4 and filtered for GC analysis.

7.1.1.7 Assessment of Thermal Stability

Thermal stability was assessed in two ways; the first was by the residual 450 nm CO-binding band in soluble cell lysates and the second was by residual activities of purified proteins towards 7-methoxycoumarin (**11**) following incubation at elevated temperatures for 15 minutes. Clarified cell free extracts were produced by sonication of cells resuspended in biotransformation buffer (0.2 M KPi + 0.4 % glycerol, pH 8) to 200 mg/mL wet cells using a Soniprep 150 ultrasonicator (MSE, London, UK), employing 20 cycles of 20 s on and 20 s off. Cell debris was removed as described previously (section 7.1.1.4). Aliquots of the supernatant (50 μ L) obtained after centrifugation were dispensed in 0.2 mL PCR tubes and incubated at various temperatures using a Mastercycler Gradient thermal cycler (Eppendorf). The residual 450 nm CO-binding peak was determined as described in the above section (section 7.1.1.4) by comparing P450 concentrations in non-incubated (retained at 4 $^{\circ}$ C) and incubated samples. The data were fitted to a Hill type equation (Eq. 1)^[6] to extract T_{50} values, where T_{50} is defined as the temperature at which 50 % of CO-binding is retained after 15 min heat treatment. Single measurements were taken (due to the volume and time required for each measurement) and the error is the fitting error.

$$y = y_0 - (y_0 \cdot T^h) / (T_{50}^h + T^h) \quad \text{Eq. 1}$$

Similarly, residual activities of purified proteins were determined by comparing demethylation activities of not incubated and incubated samples towards 7-methoxycoumarin. Product formation (7-hydroxycoumarin) was monitored directly by fluorescence spectroscopy using a microtiter plate reader (Infinite M200 Pro, Tecan, Männedorf, Switzerland). Assays were carried out in black 96-well plates (FluoroNunc) with 0.3 mM NADPH, 1 mM 7-methoxycoumatin and 1.8 μ M purified protein (final concentrations). The fluorescence intensity was recorded employing $\lambda_{\text{ex}} = 397$ nm; $\lambda_{\text{em}} = 466$ nm and a manual gain of 45. Slopes were calculated using the Magellan evaluation software and data from triplicate measurements were analyzed using GraphPad Prism 6 and fit Eq. 1 to extract T_{50} values, where T_{50} is defined as the temperature at which 50 % of the initial activity is retained after 15 min heat treatment. The error is the standard deviation between the three data sets.

Stability of the reductase domains was assessed by heating purified protein to the T_{50} of the heme domain (calculated by CO binding), then measuring the residual activity of the reductase domain using cytochrome *c* as the electron acceptor. Samples were incubated at the appropriate temperature (AT at 45.4, AX at 57.3, JT at 41.8, TB at 59.3 and TT at 60.3 °C) then assays were conducted with the addition of 50 μ M cytochrome *c* and 32 μ M NADPH to a total volume of 200 μ L. Cytochrome *c* reduction was monitored at 550 nm on a microtiter plate reader at 25 °C. Data were obtained in triplicate and the error is the standard deviation. Residual activity above 50 % indicates the reductase domain is more stable than its fused heme domain, results below 50 % indicate a less stable reductase domain under these experimental conditions. The half-life of P450-TT (Figure S8) was determined at 50 and 58 °C by measuring the residual activity toward 7-methoxycoumarin as above at various time points during incubation.

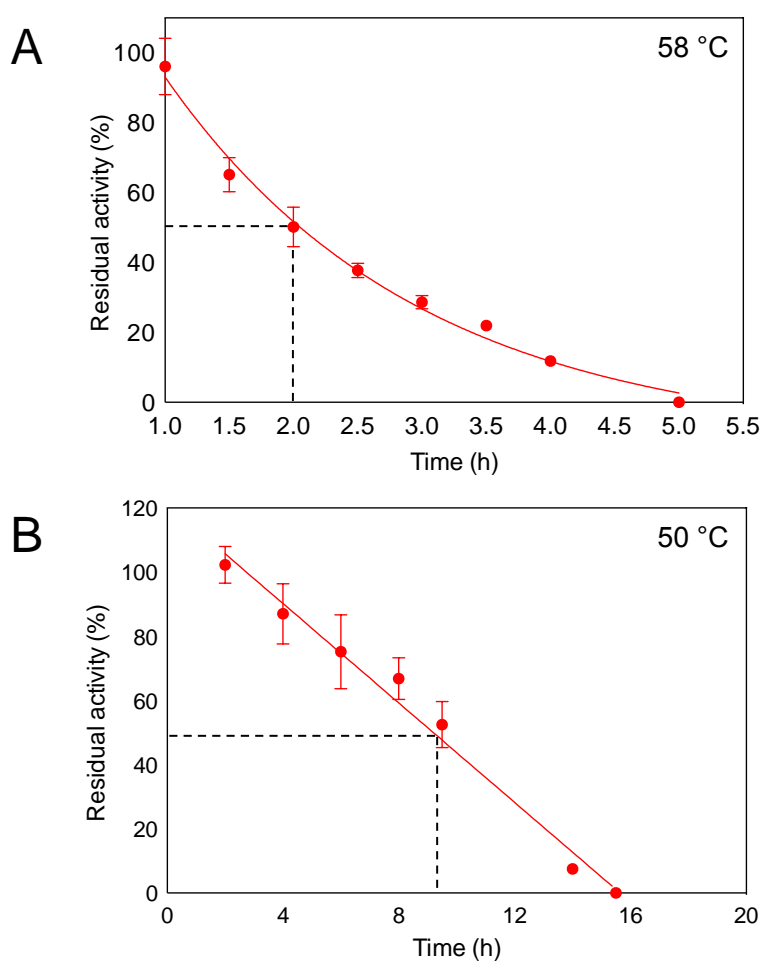


Figure S8. Half-life of P450 TT at (a) 58 degrees and (b) 50 degrees measured by monitoring activity toward 7-methoxycoumarin over time. Data were obtained in triplicate and the error is the standard error.

7.1.1.8 Diclofenac Scale-Up

Whole-cell biotransformations of diclofenac were scaled up linearly to 300 mg substrate (10 mM final concentration, 200 mg mL⁻¹ wet cell weight, 10 mg mL⁻¹ glucose, 100 mL final volume) in order to characterize the product and follow the reaction over time (Figure S9). Reactions were carried out at 40 °C with 150 rpm orbital shaking in a 500 mL sealed baffled flask. The product 5-hydroxydiclofenac was isolated by flash-chromatography as previously reported.^[7]

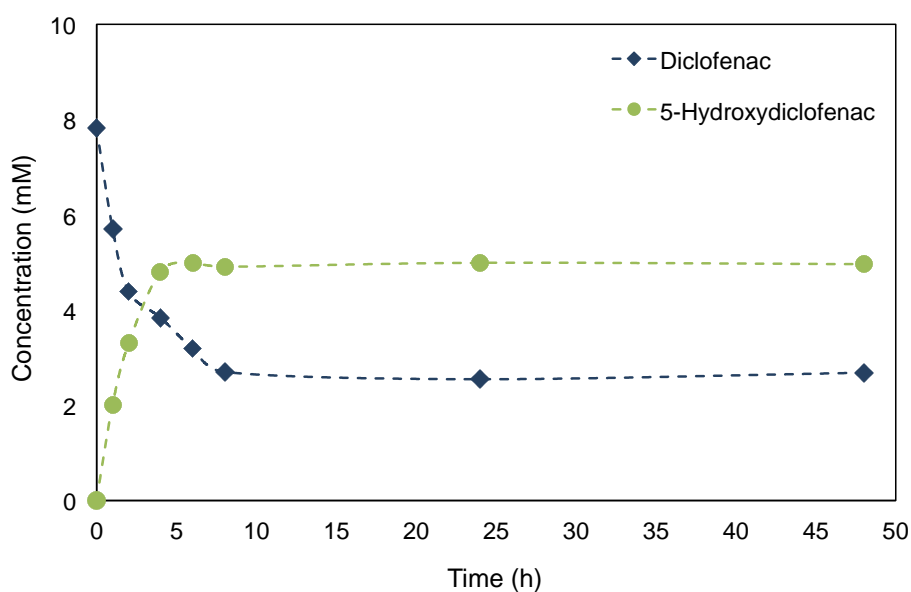


Figure S9. Time course of scaled-up hydroxylation of diclofenac at 40 °C using whole cells expressing P450 AX.

¹H NMR spectroscopy of the purified product was performed on a Bruker Avance 400 spectrometer operating at 400 MHz; ¹H NMR (400 MHz CD₃OD, TMS): δ ppm 3.64 (s, 2H; CH₂), 6.31 (d, J = 8.6 Hz; 1H; 3-H), 6.49 (dd, J = 8.6, 2.8 Hz, 1H; 4-H), 6.65 (d, J = 2.6 Hz, 1H; 6-H), 6.91 (t, J = 8.1 Hz, 1H; 4'-H) and 7.29 (d, J = 8.0 Hz, 2H; 3' and 5'-H).

7.1.1.9 Determination of Coupling Efficiency

The coupling efficiency the P450-AX catalysed conversion of diclofenac with NADPH cofactor utilization was determined by comparison of NADPH oxidation with formation of the 5-hydroxydiclofenac (**2**) product. Reactions were performed in triplicate with 0.2 mM NADPH, 0.25 mM diclofenac (**1**) and 1 μ M purified P450-AX. NADPH oxidation was measured to completion on a microtiter plate reader (Tecan) then reactions were quenched with MeCN and analyzed by HPLC to determine formation of **2**. The error is the standard deviation.

7.1.1.10 Analytics

Reversed-phase HPLC was carried out on an Agilent System (Santa Clara, CA, USA) equipped with a G4225A degasser, G1311A quaternary pump, a G1313A autosampler unit, a G1315C diode array detector and a G1316A column compartment with temperature control. Analyses were carried out using a Phenomenex KINETEX[®] column (C18, 5 μ m particle size, 4.6 mm diameter x 250 mm) with a gradient solvent system of H₂O + 0.1% TFA (A)/ MeCN + 0.1% TFA (B) at a flow of 1 or 0.5 mL min⁻¹ as indicated in Table S9 at 25 °C. Injection volume was 20 or 2 μ L for 1 mM and 10 mM biotransformations, respectively. Chromatograms were monitored at 210, 240, 250, 280 and 320 nm. Conversion values were calculated based on substrate depletion measured by comparison with control reactions set up with cells carrying an empty pET28a vector.

Table S9. HPLC/ LCMS gradient elution conditions

<i>Time (min)</i>	<i>% B</i>
0	15
20	90
20.01	100
22.01	100
22.10	15
34	15

GC-FID analysis was performed on an Agilent 6850 GC (Agilent, Santa Clara, CA, USA) with a flame ionization detector equipped with an Agilent HP-1 column (30 m x 0.32 mm x 0.25 μ m), with helium carrier gas. The inlet temperature was set at 200 °C, detector temperature at 250 °C and pressure maintained at 6.8 psi. For analysis of 14, 17 and 19 the column oven was set to 50 °C and increased to 200 °C at a rate of 10 °C min⁻¹ and held isothermally for 1 min. For analysis of 16, the oven temperature was increased to 220 °C and held isothermally for 2 min from the same starting temperature. Conversions were calculated based on the relative peak areas of the product and remaining substrate.

GC-MS analysis was performed on a 7890B GC system coupled to a 5977B MS detector (Agilent Technologies) and was used to determine the major product from biotransformations with 2-hydroxy-4-methoxyacetophenone (**5**). An Agilent HP-1 MS column (30 m x 0.32 mm x 0.25 μ m) with helium carrier gas at a flow rate of 2 mL min⁻¹ was used for separation. The inlet temperature was set to 270 °C, with the column oven set to 50 °C and held for 1 min then increase at a rate of 10 °C min⁻¹.

7.1.1.11 Determination of Enantiomeric Excess

Biotransformations with ethyl benzene (**15**) and methyl phenyl sulfide (**18**) were analyzed on a β -DEX-325 chiral column (Supelco) while styrene (**20**) biotransformations were analyzed on a CP-ChiraSil-DEX CB column (Agilent). Commercial standards of (*R*)-1-phenylethanol, (*S*)-1-phenylethanol, (*R*)-styrene oxide and (*S*)-styrene oxide were used to determine the stereochemistry of the respective biotransformation products. Racemic methyl phenyl sulfoxide was used to establish if separation of enantiomers could be achieved, then AT, AX, JT, TB and TT biotransformations were compared to RhF, which has previously been determined to be (*S*) selective.^[8]

For GC analysis of 1-phenylethanol (**16**) and methyl phenyl sulfoxide (**19**) the column oven was set to 130 °C and held isocratic for 20 min then increased by 2 °C min⁻¹ to 150 °C then further increased to 200 °C at a rate of 20 °C min⁻¹. For analysis of styrene oxide (**21**) the oven was set to 100 and increased to 120 °C at a rate of 10 °C min⁻¹, held for 2 min then increased to 180 °C at a rate of 10 °C min⁻¹.

Table S10. Enantiomeric excess determined for products **16**, **19** and **21**

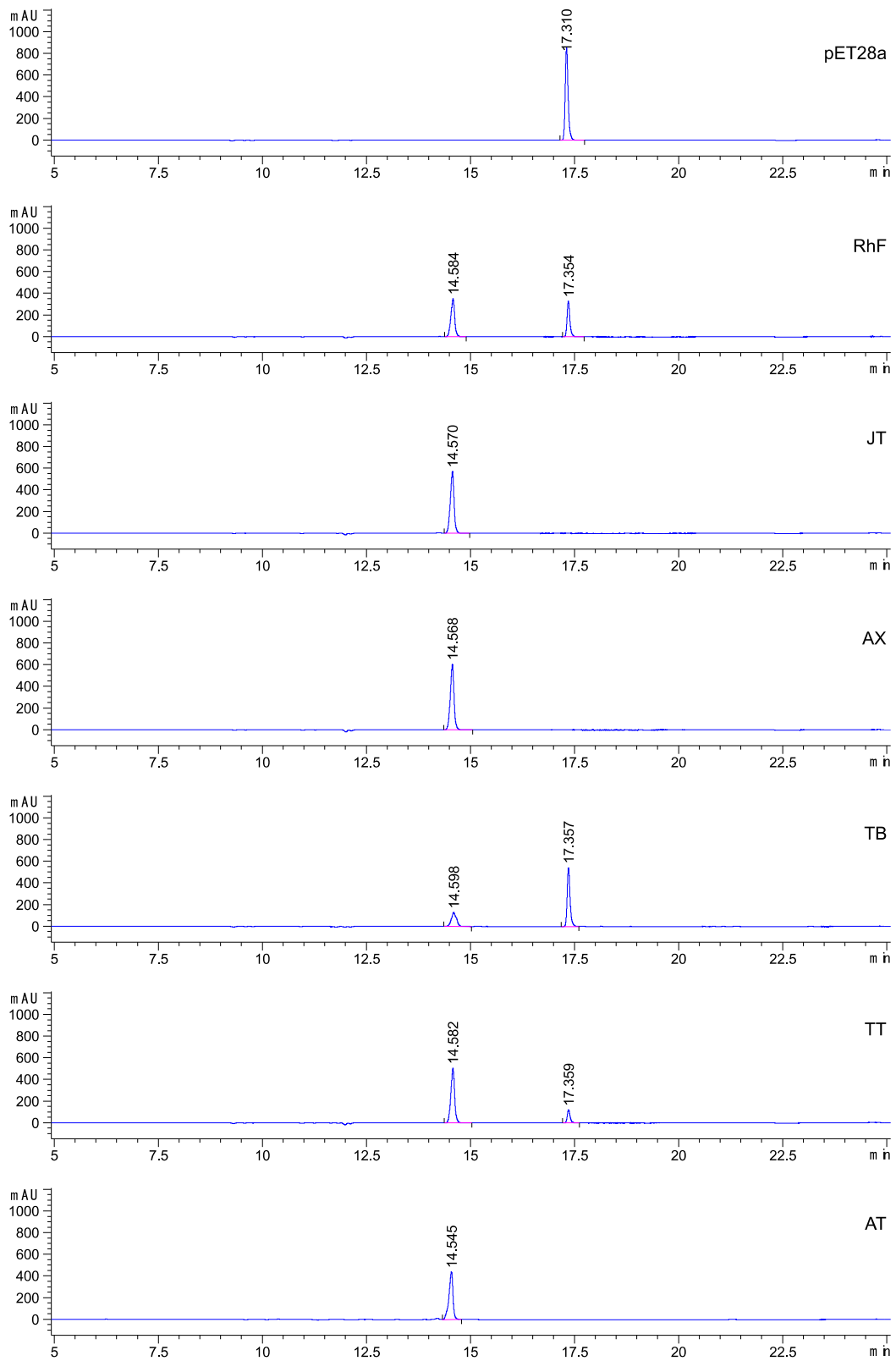
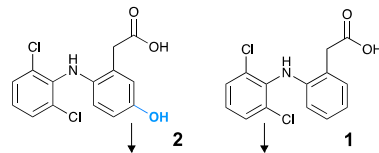
	1-Phenylethanol (16)	<i>ee</i> (%) Methyl phenyl sulfoxide (19) ^a	Styrene oxide (21)
RhF ^b	n/a	89	n/a
AT	90 ± 10 (<i>S</i>)	> 90 (<i>S</i>)	<i>rac</i>
AX	n/a	> 90 (<i>S</i>)	<i>rac</i>
JT	90 ± 10 (<i>S</i>)	85 ± 10 (<i>S</i>)	<i>rac</i>
TB	90 ± 10 (<i>S</i>)	80 ± 10 (<i>S</i>)	<i>rac</i>
TT	n/a	>80	n/a

^a Full baseline resolution of racemic standard could not be achieved.

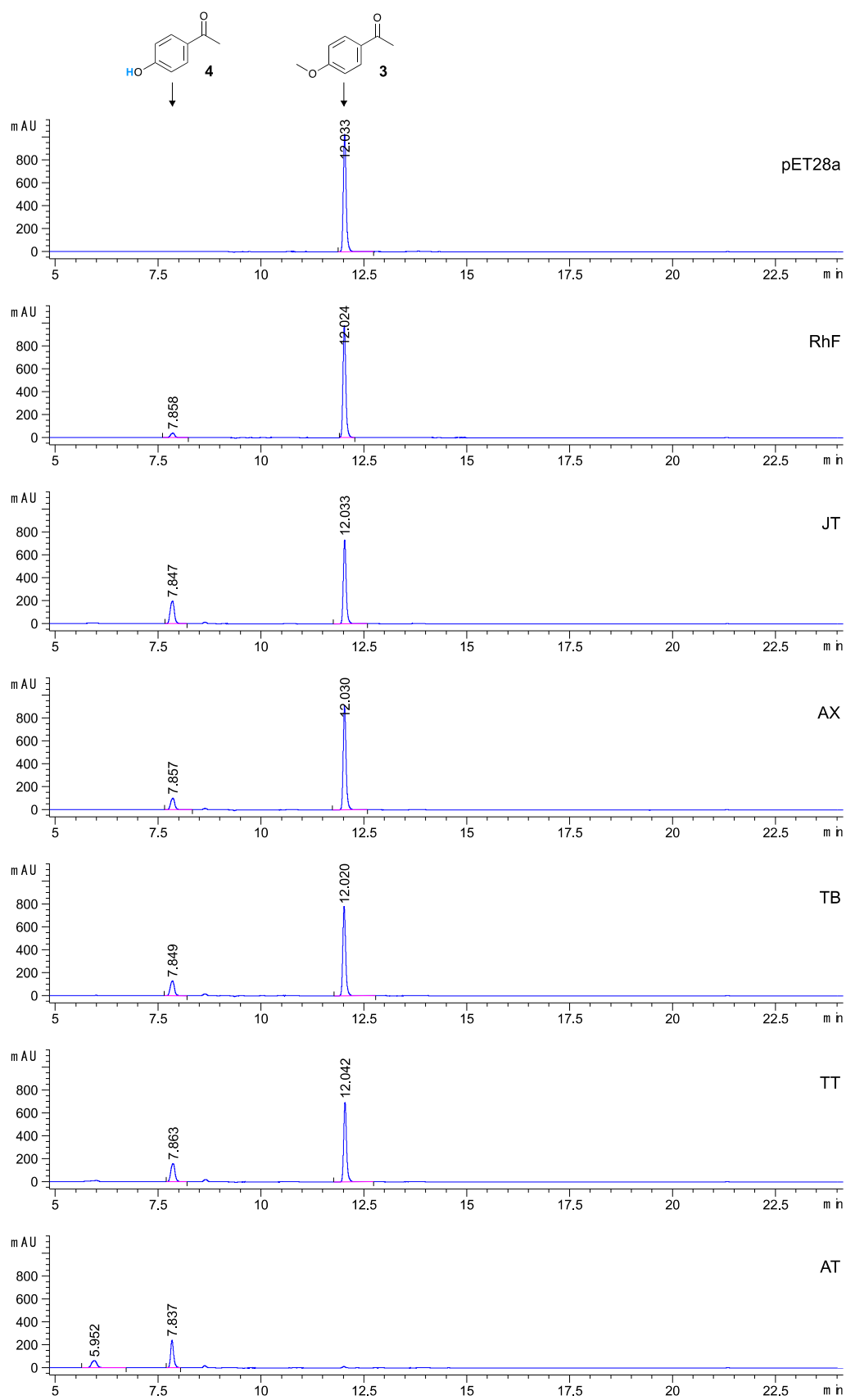
^b Values were obtained from literature^[8] and were used to assign the stereochemistry of the AT, AX, JT, TB and TT methyl phenyl sulfoxide products.

7.1.2 HPLC and GC-FID/MS data

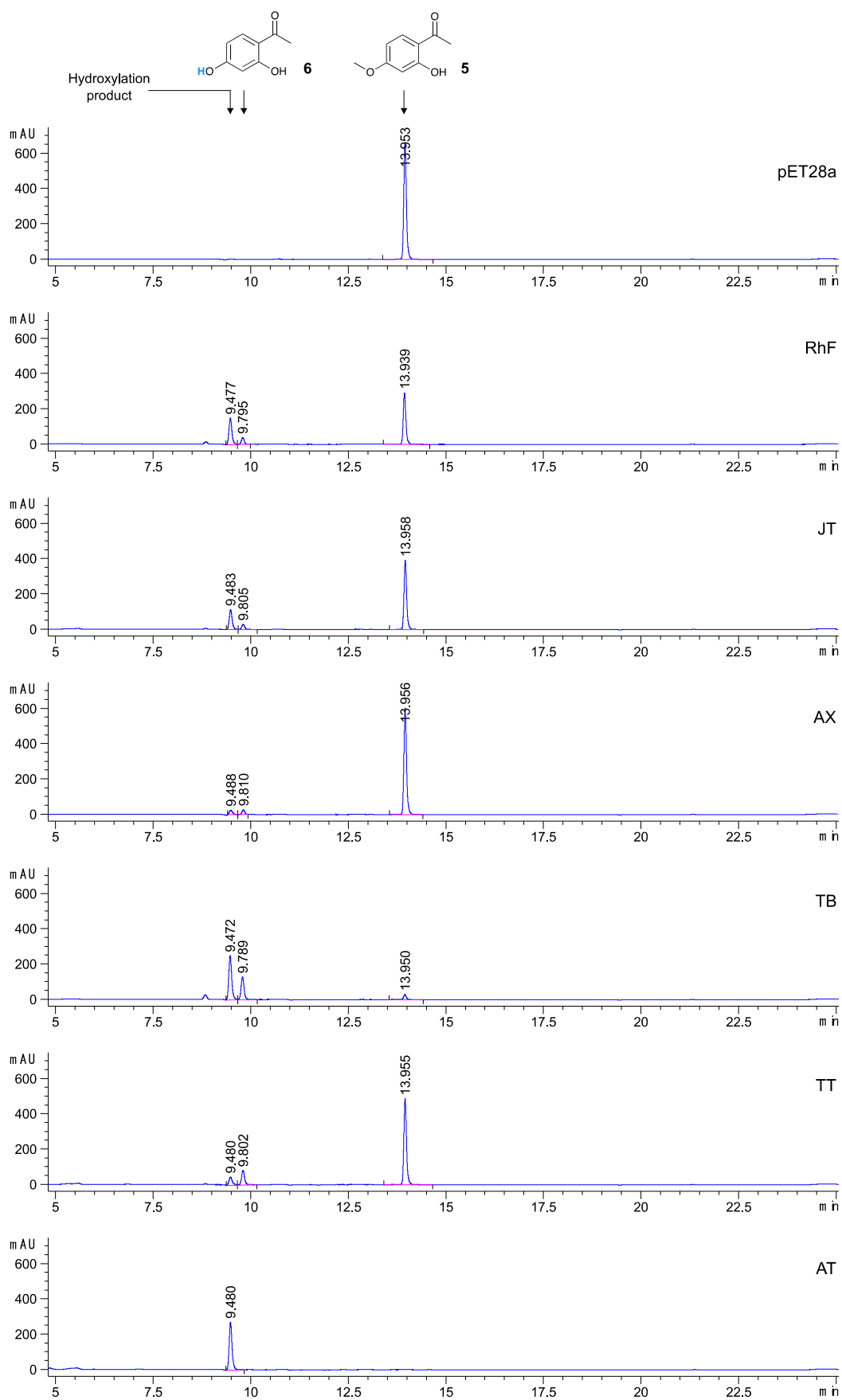
The HPLC, GC traces and MS data obtained from the P450 activity profiling screen are provided below.



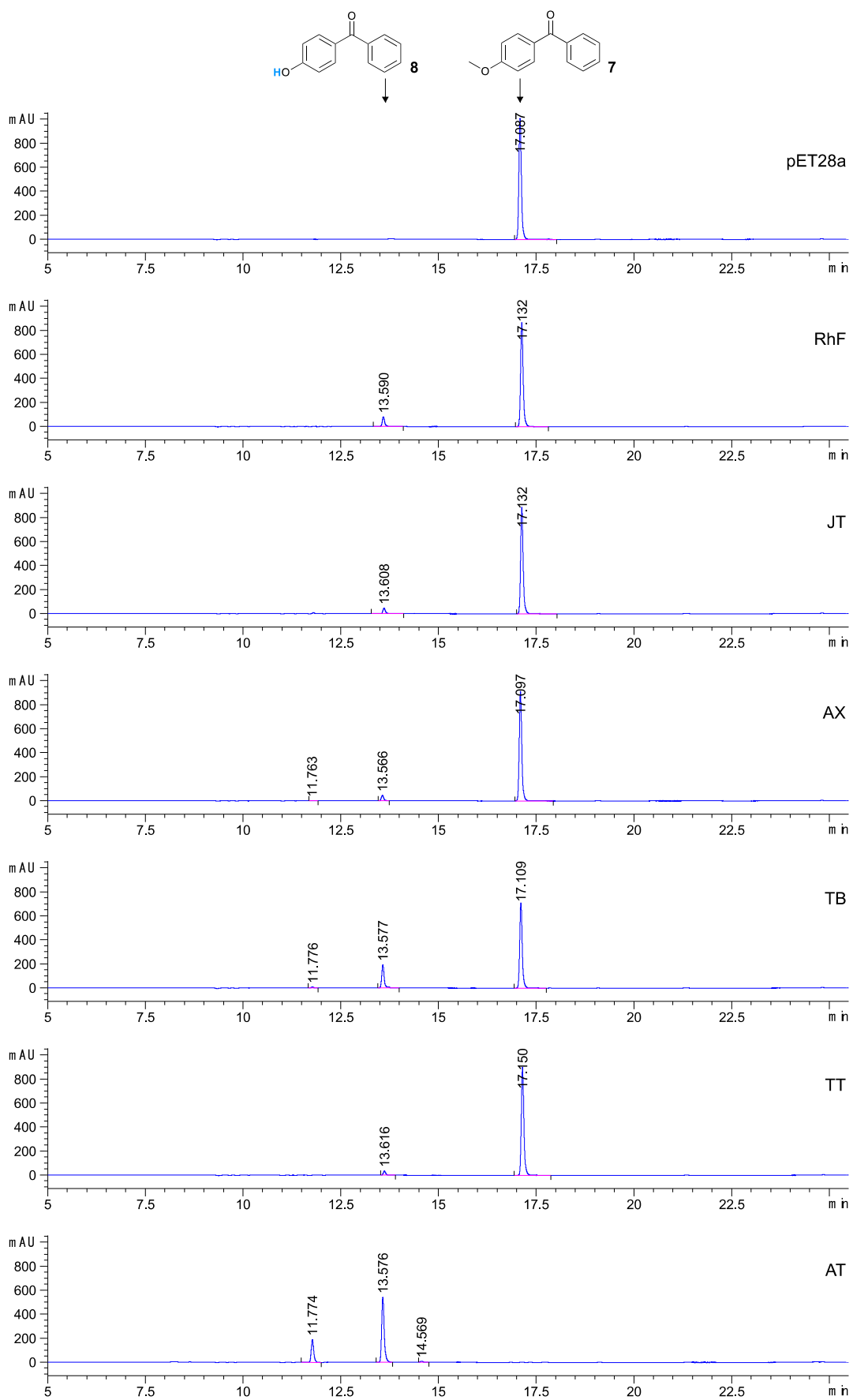
HPLC traces for the P450 catalysed biotransformation of diclofenac (1)



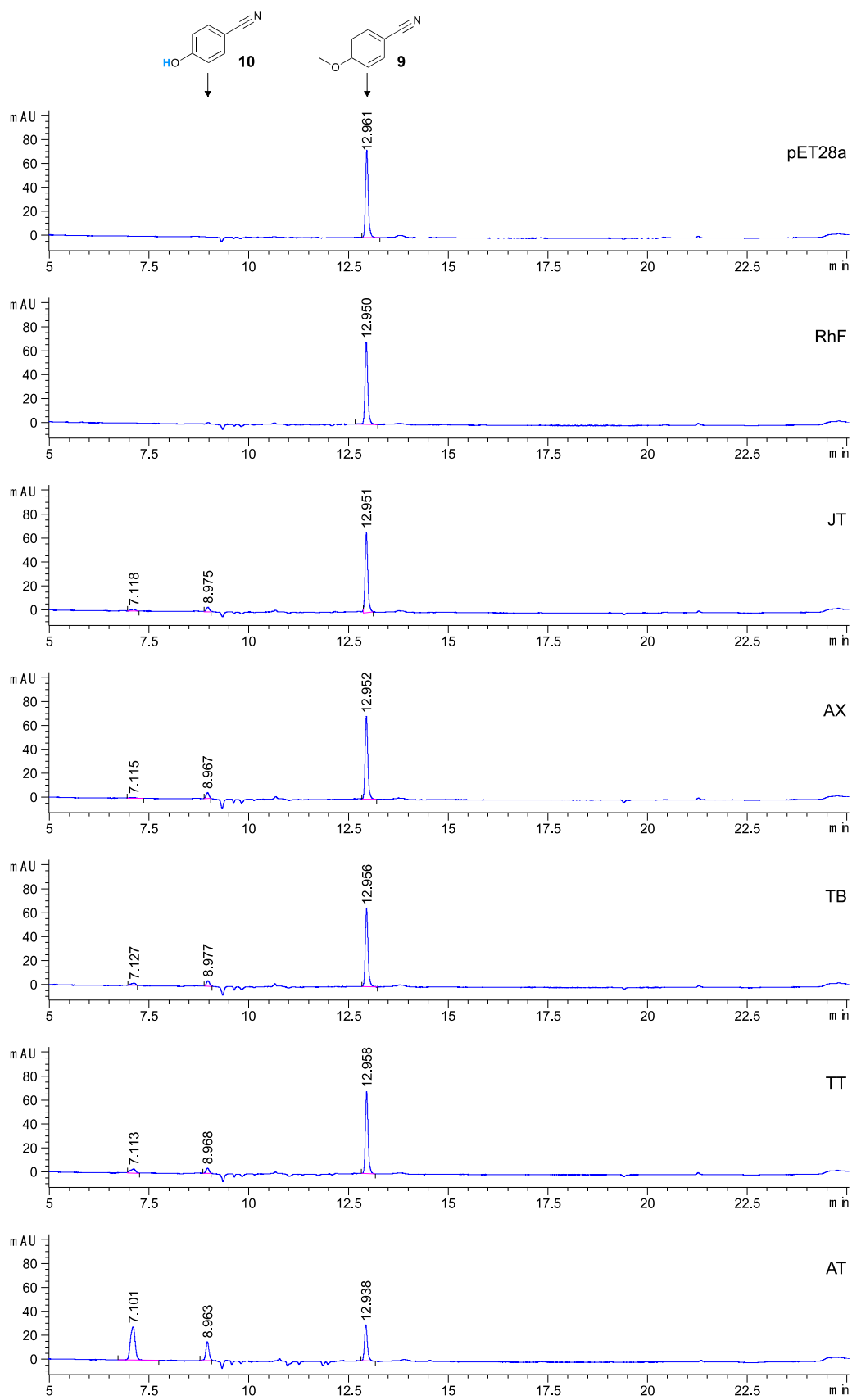
HPLC traces for the P450 catalysed biotransformation of 4-methoxyacetophenone (**3**)



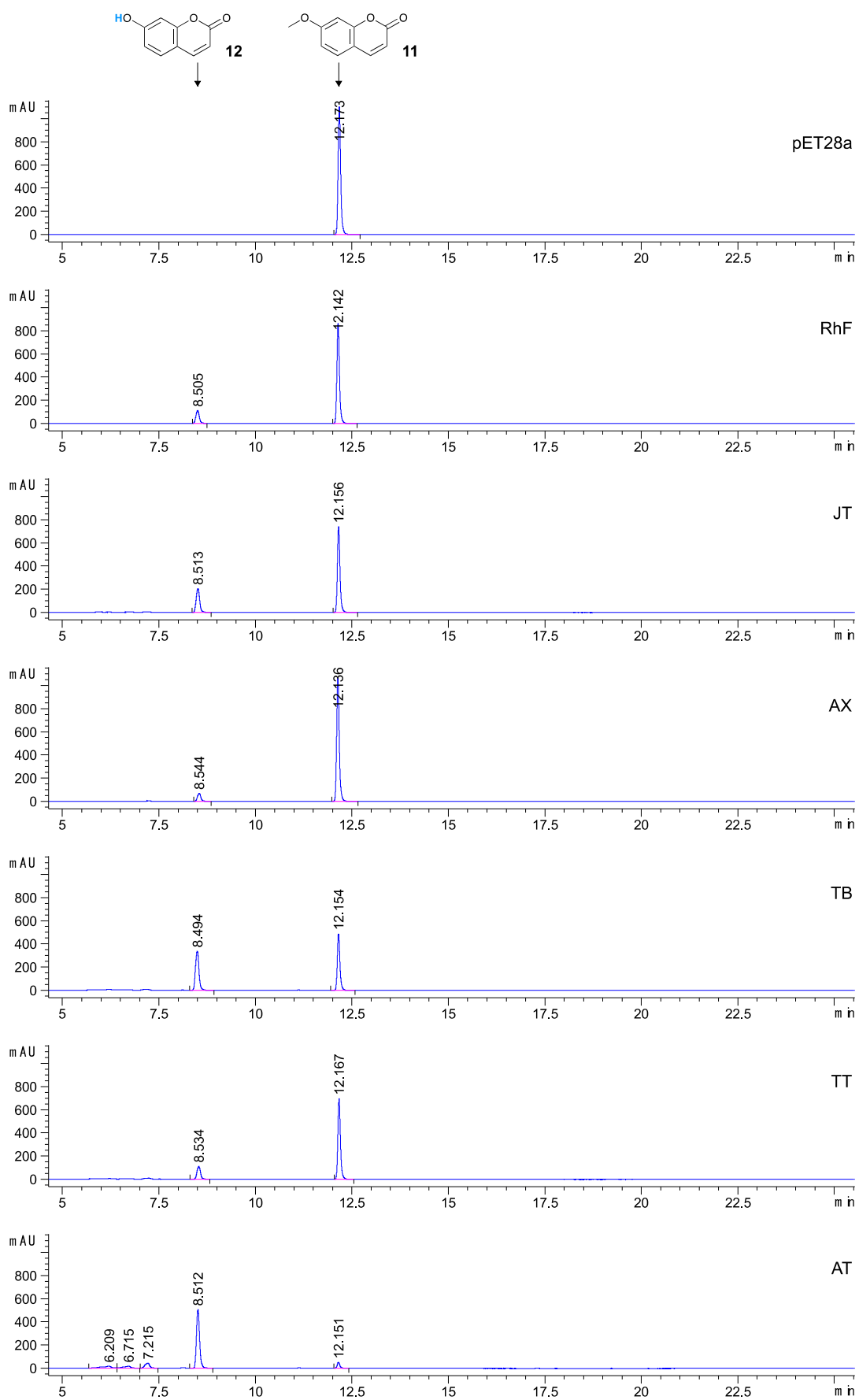
HPLC traces for the P450 catalysed biotransformation of 2-hydroxy-4-methoxyacetophenone (**5**)



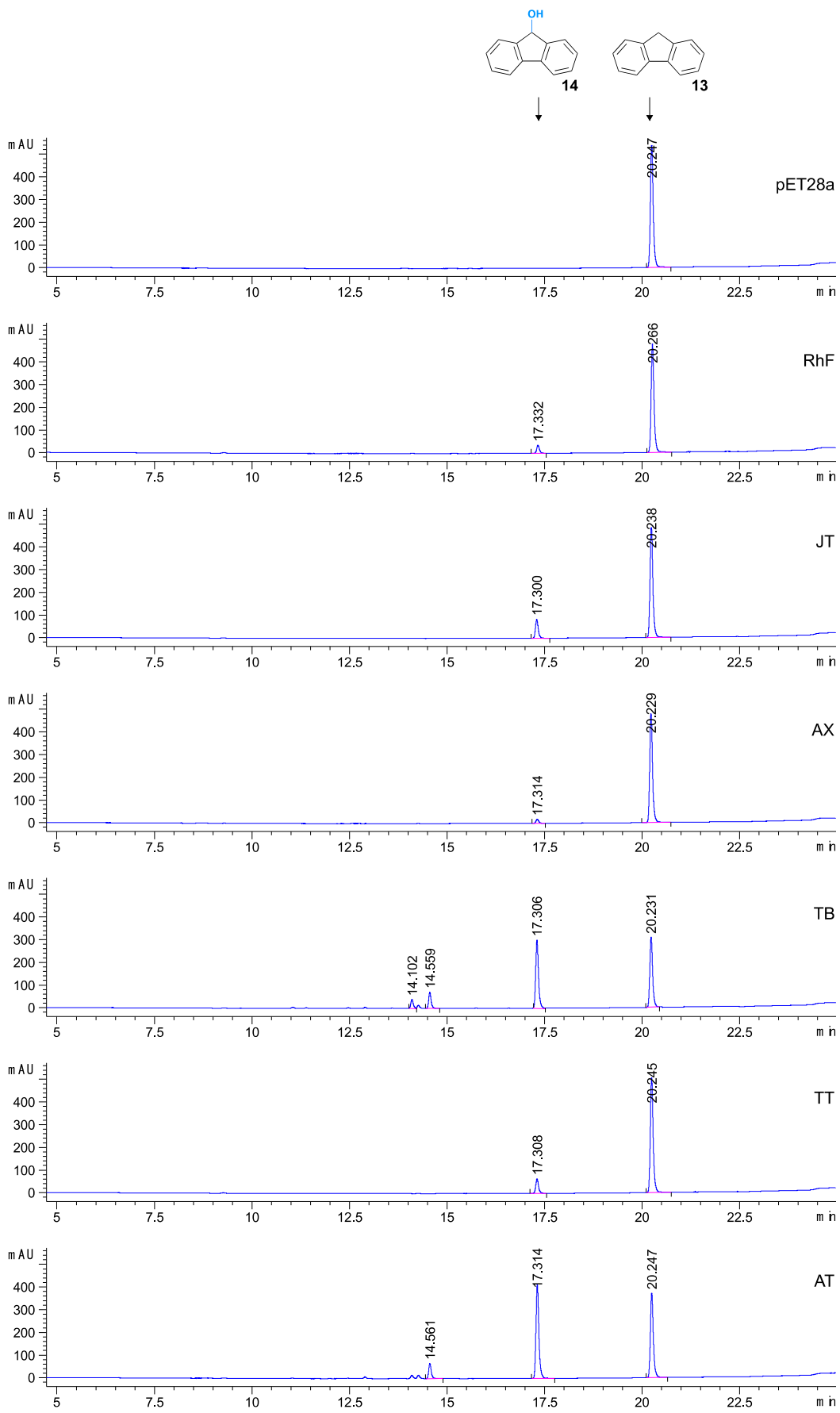
HPLC traces for the P450 catalysed biotransformation of 4-methoxybenzophenone (7)



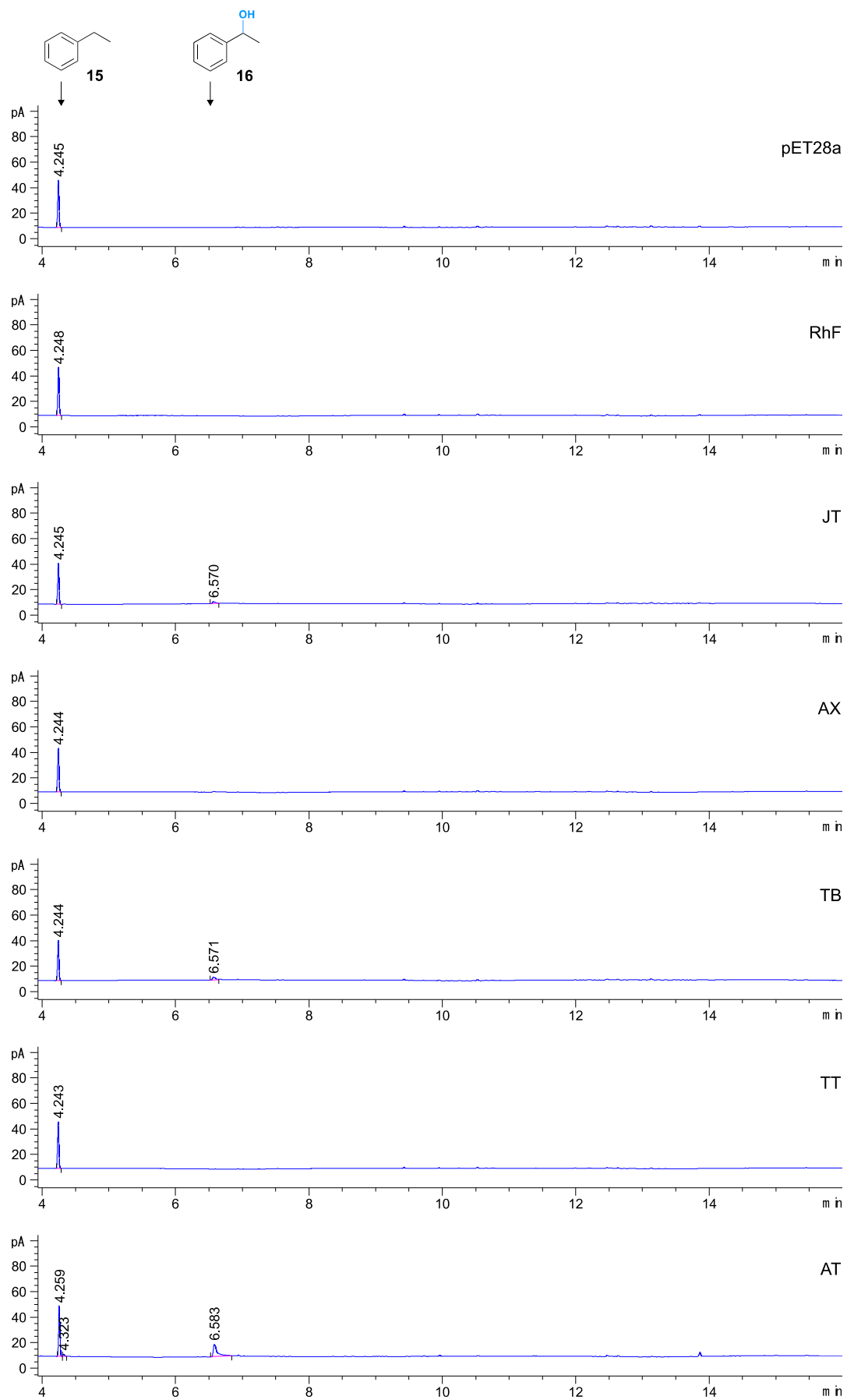
HPLC traces for the P450 catalysed biotransformation of 4-methoxybenzonitrile (9)



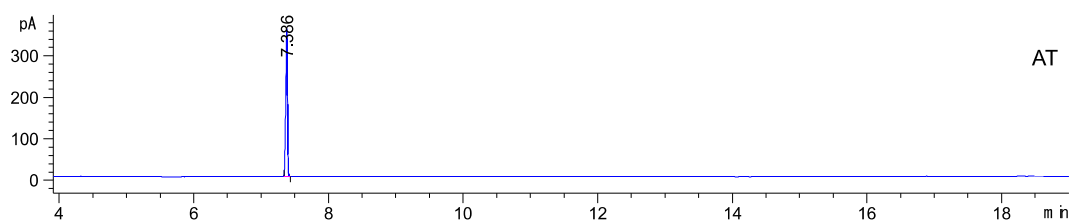
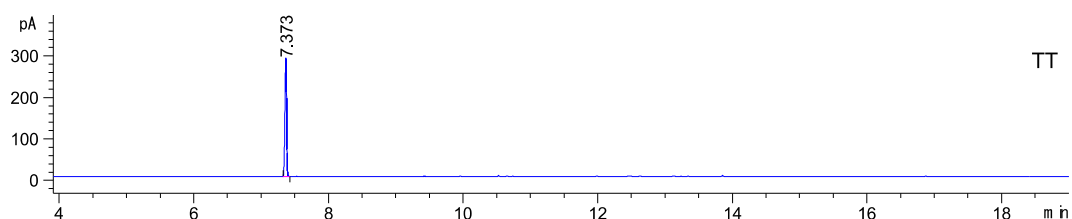
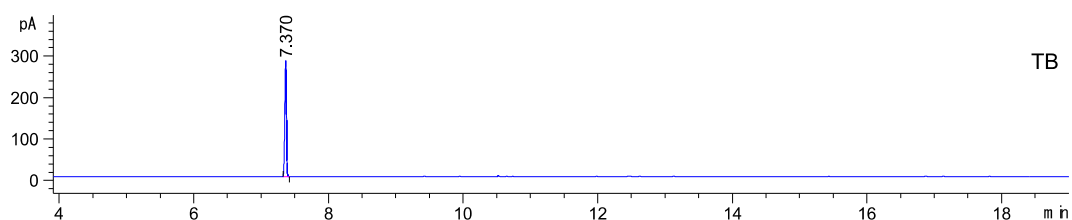
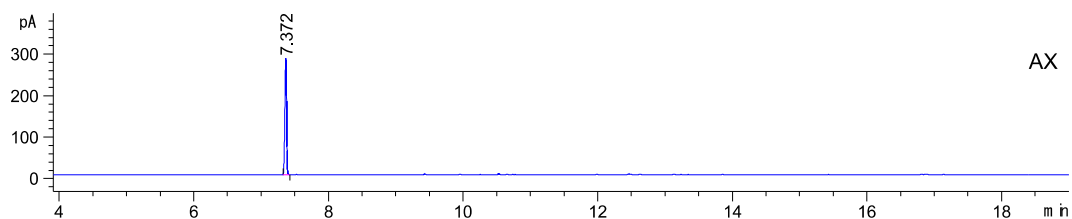
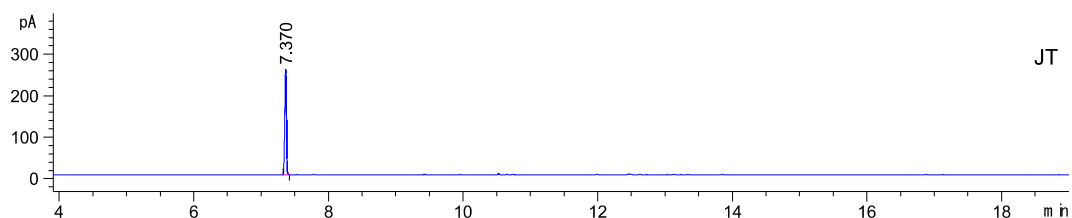
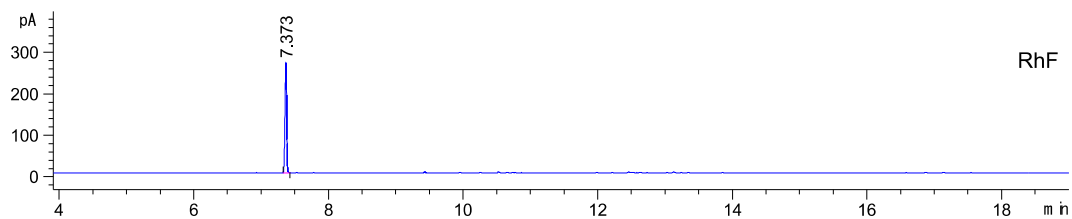
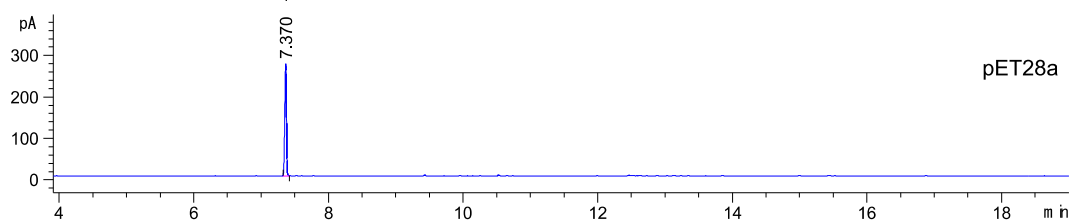
HPLC traces for the P450 catalysed biotransformation of 7-methoxycoumarin (**11**)



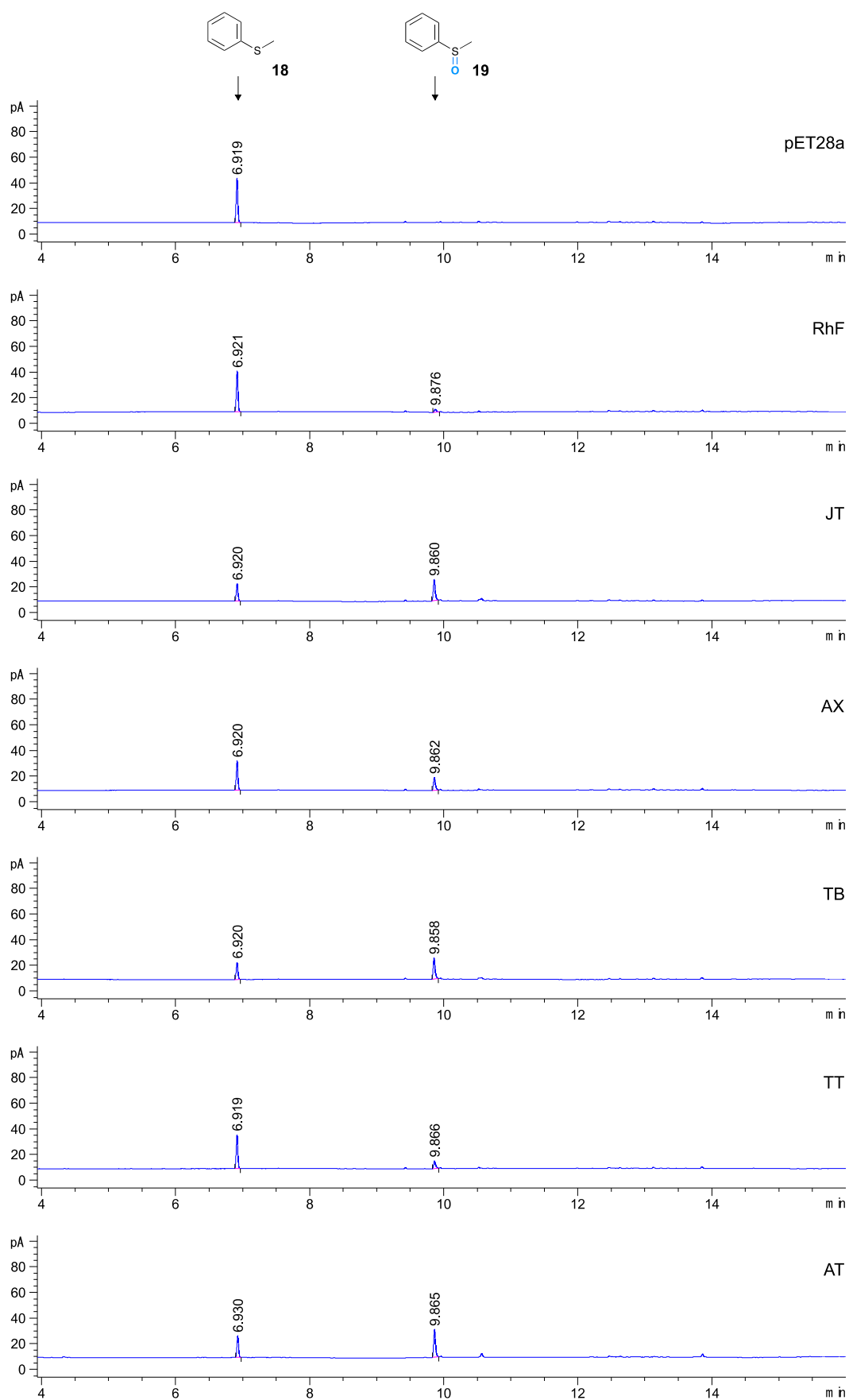
HPLC traces for the P450 catalyzed biotransformation of fluorene (**13**)



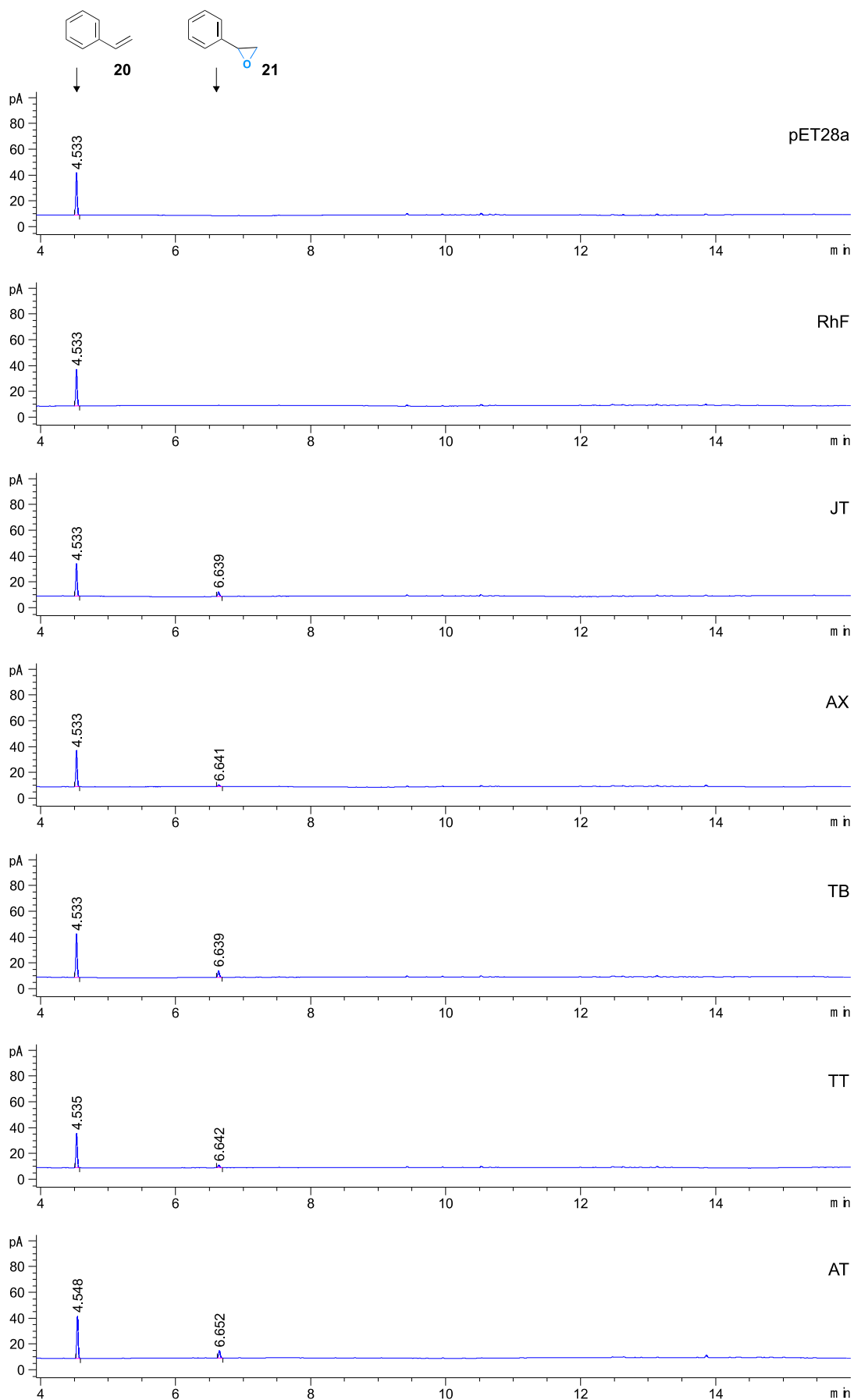
GC-FID traces for the P450 catalysed biotransformation of ethylbenzene (**15**)



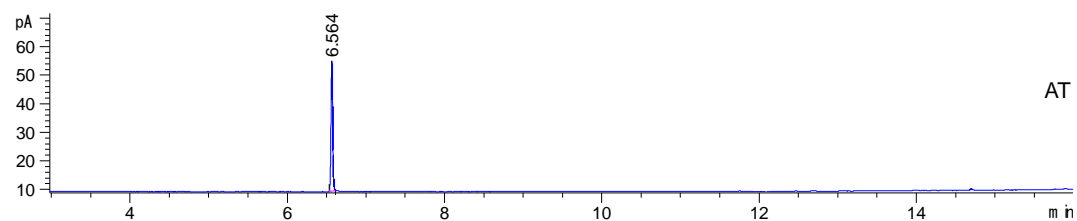
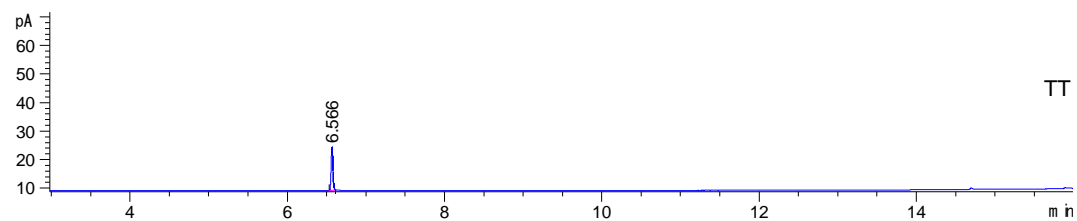
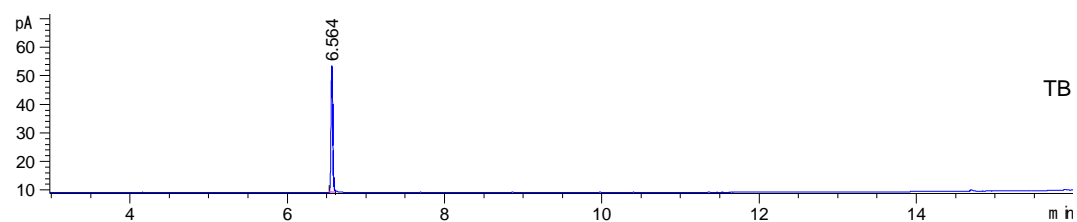
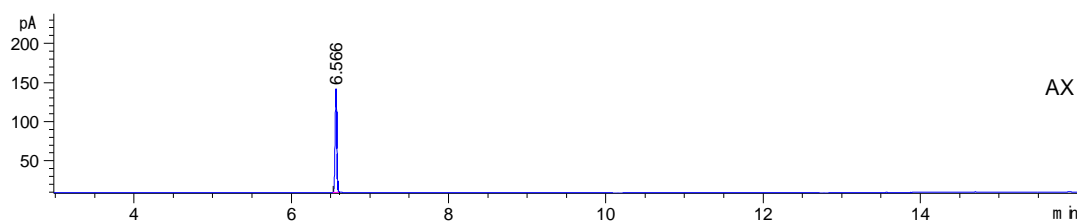
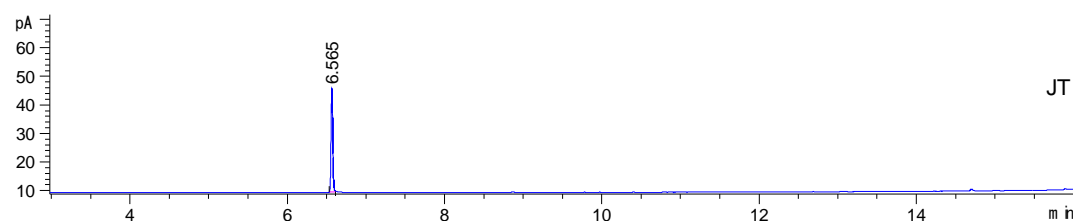
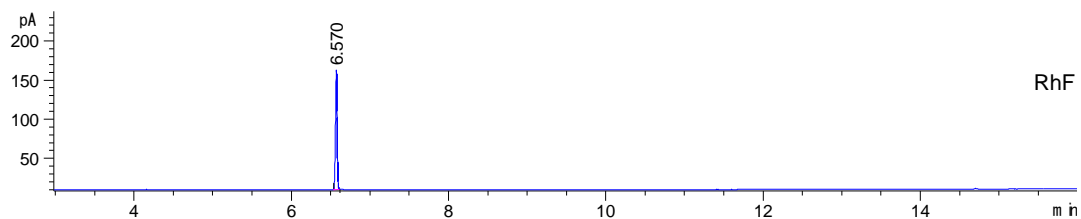
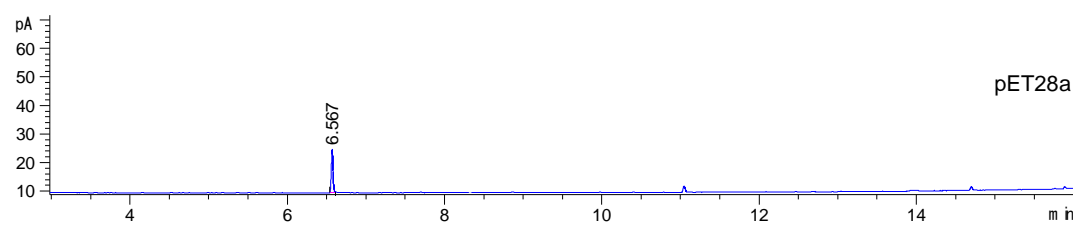
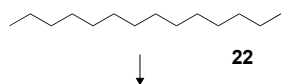
GC-FID traces for the P450 catalysed biotransformation of α -isophorone (**17**)



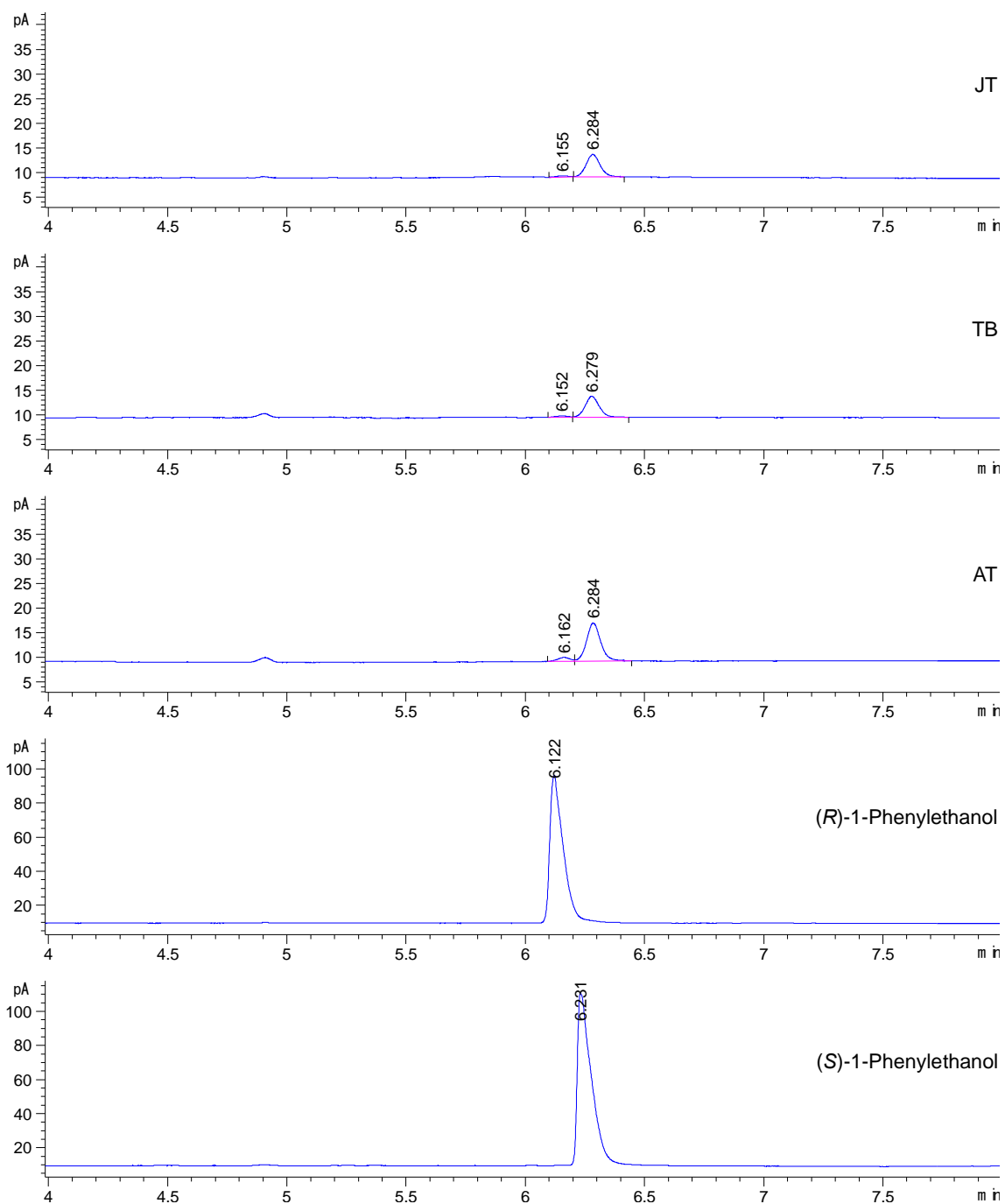
GC-FID traces for the P450 catalyzed biotransformation of methyl phenyl sulfide (**18**)



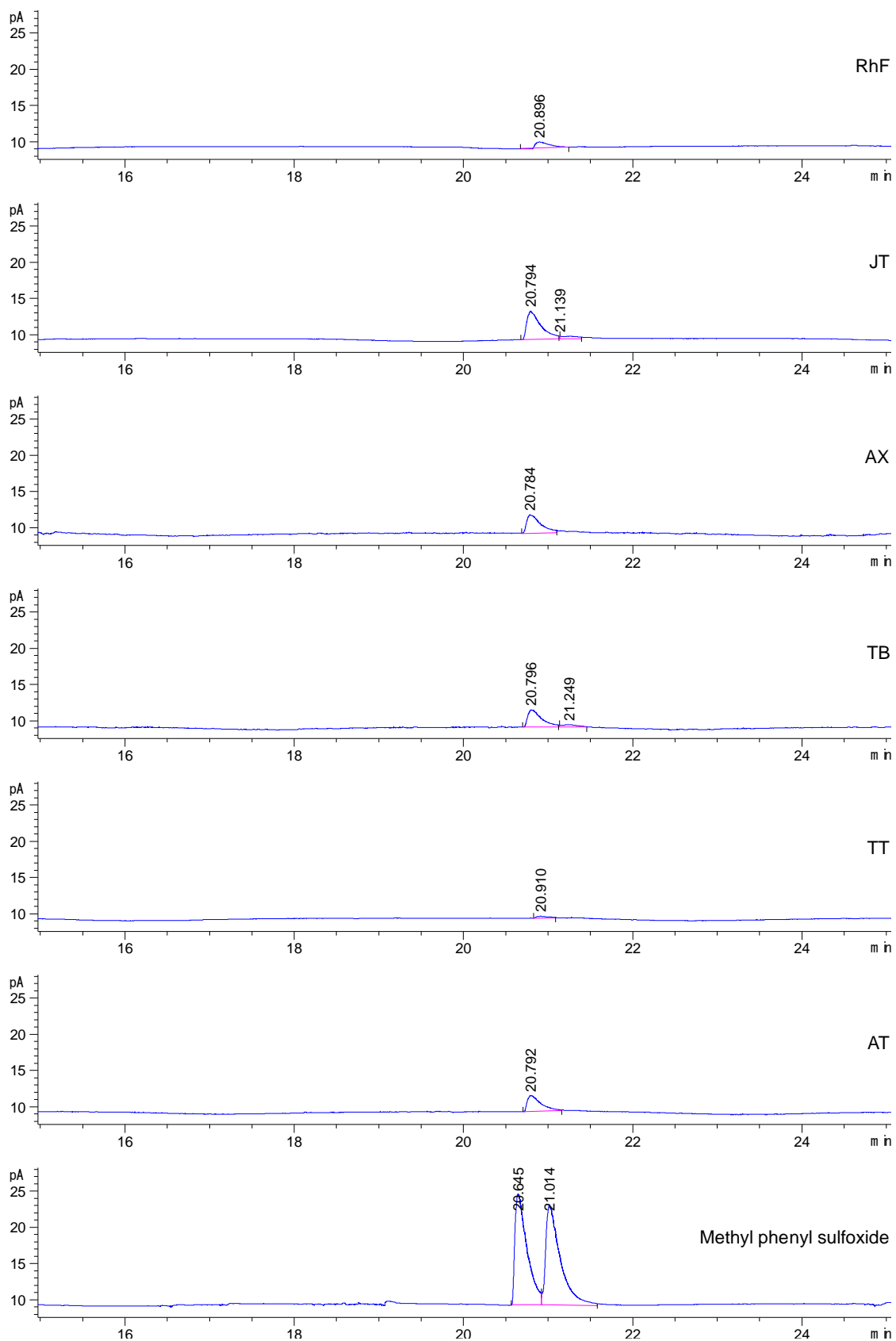
GC-FID traces for the P450 catalysed biotransformation of styrene (**20**)



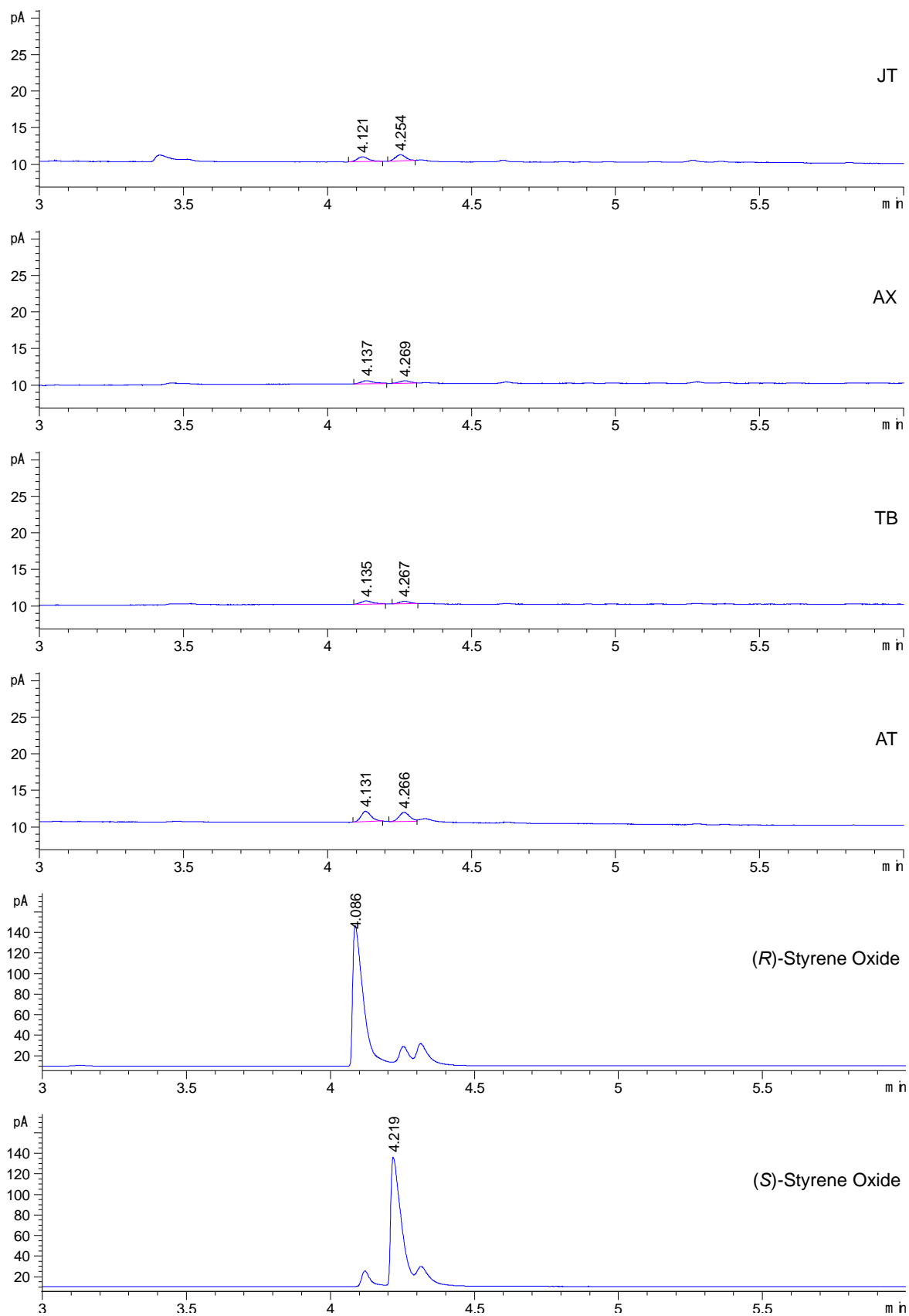
GC-FID traces for the P450 catalysed biotransformation of tetradecane (22)



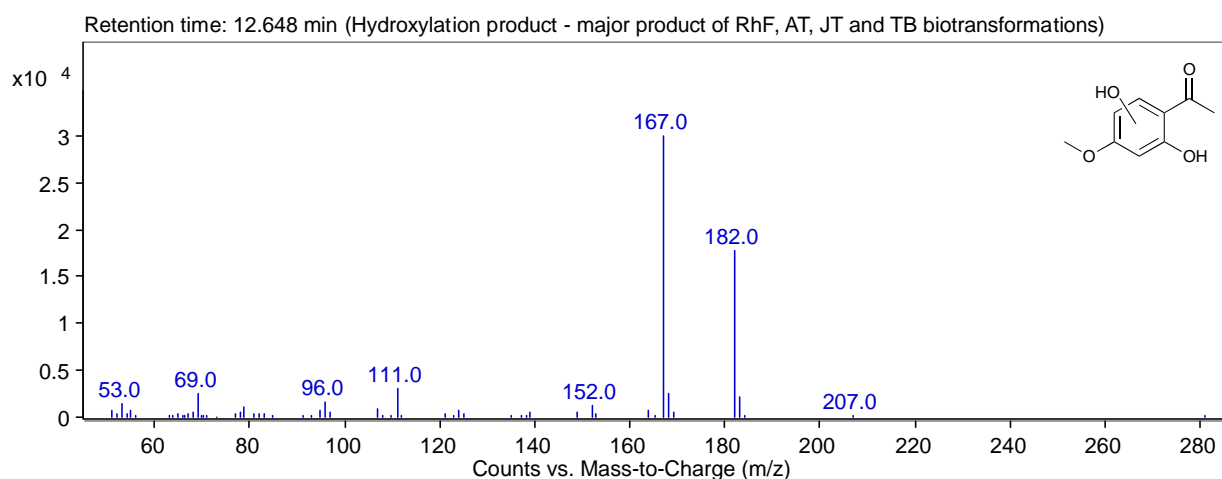
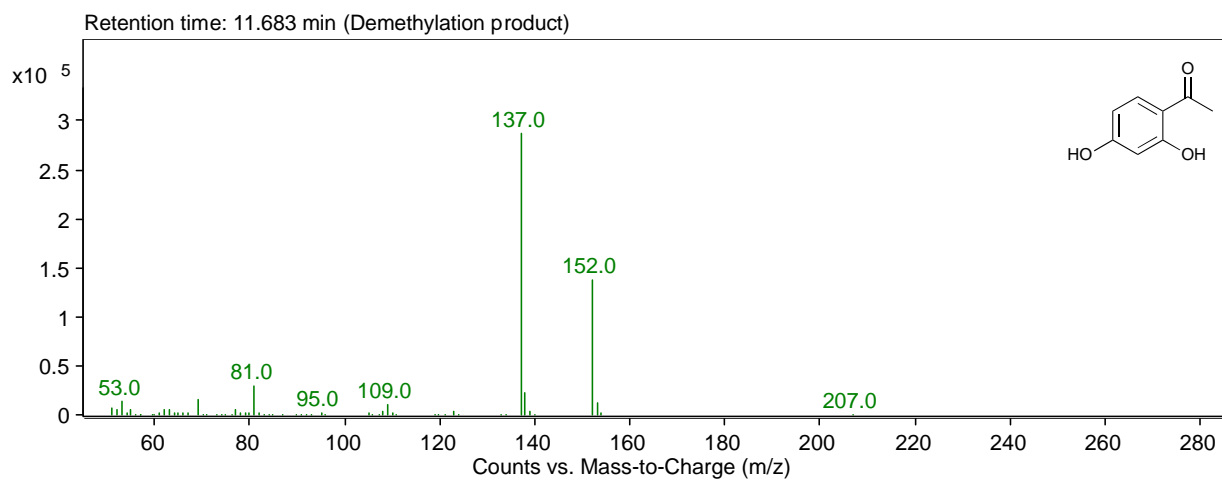
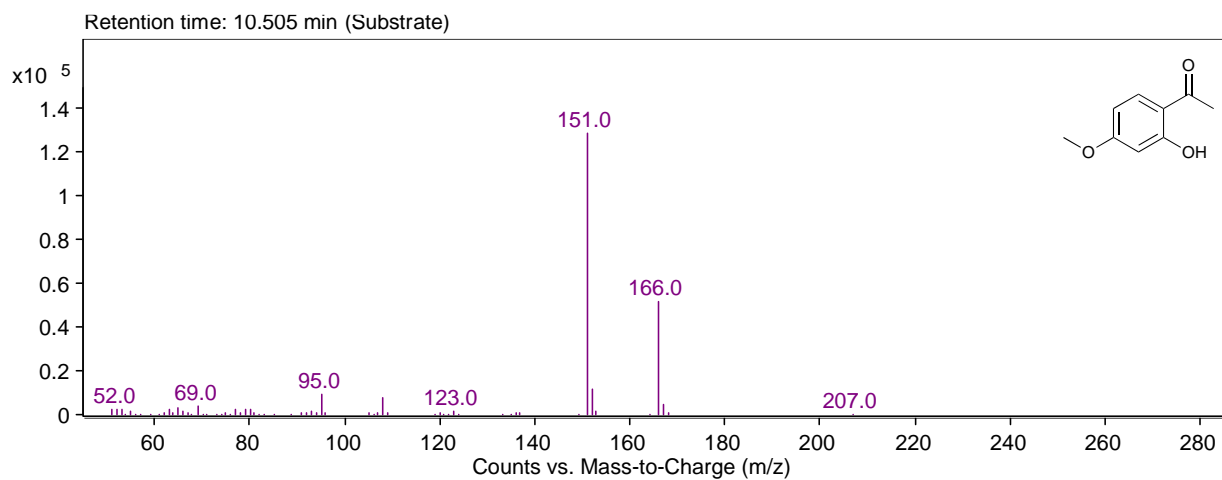
GC-FID traces for the assessment of enantiomeric excess of the P450 catalysed production of 1-phenylethanol (**16**)



GC-FID traces for the assessment of enantiomeric excess of the P450 catalysed production of methyl phenyl sulfoxide (**19**)



GC-FID traces for the assessment of enantiomeric excess of the P450 catalysed production of styrene oxide (**21**)



Mass spectra obtained from GC-MS analysis of P450 biotransformations with 2-hydroxy-4-methoxyacetophenone (**5**).

7.1.3 References

- [1] D. A. Mead, S. Lucas, A. Copeland, A. Lapidus, J.-F. Cheng, D. C. Bruce, L. A. Goodwin, S. Pitluck, O. Chertkov, X. Zhang, et al., *Stand. Genomic Sci.* 2012, 6, 381–400.
- [2] S. Kumar, G. Stecher, K. Tamura, *Mol. Biol. Evol.* 2016, 33, 1870–1874.
- [3] T. L. Poulos, B. C. Finzel, I. C. Gunsalus, G. C. Wagner, J. Kraut, *J. Biol. Chem.* 1985, 260, 16122–16130.
- [4] T. Omura, R. Sato, *J. Biol. Chem.* 1964, 239, 2370–2378.
- [5] M. C. U. Gustafsson, O. Roitel, K. R. Marshall, M. A. Noble, S. K. Chapman, A. Pessegueiro, A. J. Fulco, M. R. Cheesman, C. Von Wachenfeldt, A. W. Munro, *Biochemistry* 2004, 43, 5474–5487.
- [6] B. J. Stevenson, J.-W. Liu, D. L. Ollis, *Biochemistry* 2008, 47, 3013–3025.
- [7] J. M. Klenk, B. A. Nebel, J. L. Porter, J. K. Kulig, S. A. Hussain, S. M. Richter, M. Tavanti, N. J. Turner, M. A. Hayes, B. Hauer, et al., *Biotechnol. J.* 2017, 12, 1600520.
- [8] E. O'Reilly, M. Corbett, S. Hussain, P. P. Kelly, D. Richardson, S. L. Flitsch, N. J. Turner, *Catal. Sci. Technol* 2013, 3, 1490–1492.

7.2 The Crystal Structure of P450-TT Heme-Domain Provides the First Structural Insights into the Versatile Class VII P450s

7.2.1 Experimental

7.2.1.1 Materials

Chemicals, solvents and carbon monoxide for CO difference spectroscopy were obtained from Sigma-Aldrich (Dorset, UK). Chemically competent cells and enzymes for molecular biology were purchased from New England Biolabs (Hitchin, UK). M9 Minimal Salts Base (5X) were purchased from Formedium (Hunstanton, UK).

7.2.1.2 Molecular Biology Methods

7.2.1.2.1 Generation of Constructs for Crystallography Trials

A pET28a vector encoding the full-length *N*-terminal polyhistidine tagged CYP116B46 (P450-TT) from *Tepidiphilus thermophilus* was obtained from our previous work.^[1] Inverse PCR was carried out to obtain a pET28a vector encoding the *N*-terminal polyhistidine tagged P450-TT heme domain. Primer synthesis and DNA sequencing were performed by Eurofins Genomics (Ebersberg, Germany). Genes and primers employed in this work are given below.

Primer	5'->3' sequence
TT full for, Tm=59°C	TAAGTCGACAAGCTTGCGGC
TT full rev, Tm=60°C	GGGATCGCGCCGTTTCAG

> P450-TT 6XHis heme domain

```
ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATG
ACTGGTGGACAGCAAATGGGTCGCGGATCCATGGAACTGAACTCAAGGAAACGGCACGCGGGACCTGTCC
CGTCGCCCATGGTGGGCAGAGCTCTGTGGGCGGCTGCCCGTGCATCGTCTGGCAGAGGACTTCGATCCATT
CCAAGACGCTTATATGGCCGATCCGGCGCAATTCGTTTCGTTGGGCACGCGAGCAGGTACCCATTTTCTACGCA
CCAAAGCTCAACTATTGGGTGGTGACACGCTACGACACGATCAAACAGATCTTCCGCGATCCGGTCACCTTCA
GCCCATCCAACGTCCTGCAATCCTTCGCCAGCCTTCTGCCGAAGTGCGCCAGGTACTIONGAGCGTTACGGCTAC
GCTTTCAACCGCACCCCTGGTCAATGAAGACGAGCCGATGCACCTGGAGCGGCGGCGCGTGTGATGGAACCC
TTCGCCTCCGAGCACCTGGCGGAACACGAGCCGATGGTGCGCGAGCTGGTGCGCCGCGCCGTTAATCGCTTC
ATCGACACGGGGCGAGCCGACCTCGTCGATCAAATGATATGGGAAGTTCCTTTACCGTCGCCCTGCACTTTC
TCGGAGTGGATGACGACGACCGGGAGAAGATGCGCCGTTTCGCCATCGCCACACCCGTGAATGCTTTCGGTC
```

GGCCTTCTCCCGAAGAACAGCTCGCTGTGGCCGAGACGGTAGGGCAGTTCTGGCAGTTTTGTGGGGAAGTGC
 TTGAGAAGATGCGGCGCACTGCTGATGGTCCTGGCTGGATGCGCTATTCGATCCGCCAGCAAAGCTCTATCC
 TGACGTCGTCACCGATTTCGTATCTGCACTCGATGATGCAGGCGATCATCGTCGCGGCCACGAGACGACGGTG
 TTTGCCACCACTAACGCCCTCAAGACTCTGCTGGAGCACGAAACCGTCTGGCGCGAGATCTGCGCCGATCCTT
 CCCTCATCCCAGCGGCGGCCGAAGAATGCCTGCGCTACAACGGGCGGTGGCGGGGTGGAGGCGCCGACC
 ACGCGGGAGGTGGAAGTGGAGGGAGTGCGCCTGCCGGTCGGGGCAAACATTCTCATGGTCGTCGCCTCGGC
 CAACCACGATTCCGCGCATTTCGACGATCCAGAGTTCTTCGACATTGGCCGAGCAACGCCTCCGAGCATCTCA
 ACTTTGGCTATGGGGCGCACCAAGTGCCTCGGCAGGAACCTCGGTGCGATGGAAATGCAGATCATGATCGAGG
 AACTCTCCCGCCGCTTGCCGCACATGCGCCTAGCCGAGCAGCGGTTGACTACCTGCACAACGTCTCCTCCGC
 GCACCGCGTCATCTGTGGGTGCAATGGGATCCAGCACAGAATCCTGAACGGCGCGATCCC

7.2.1.2.2 P450-AX Library Generation

P450-AX libraries were generated by MEGAWHOP cloning^[2] using the cloned wild-type gene as template.^[1] Custom DNA oligo were prepared by Eurofins Genomics. Ordered primer sequences are given below:

Ala80NNKfw	CTTCTCCCCGCGAATCANNKCTCGAAAAGATCACCC
Leu81NNKfw	CTTCTCCCCGCGAATCGCGNNKGAAAAGATCACCC
Thr105NNKfw	CAATCTCCAGCGCANNKATGGTGAACGAGGACG
Val107NNKfw	CCAGCGCACCATGNNKAACGAGGACGAGC
His190NNKfw	TTCGCCGCCGCCANNKACGCTCAACACC
Thr191NNKfw	TTCGCCGCCGCCANNKCTCAACACCTGG
Asn193NNKfw	CACACGCTCANNKACCTGGGGCC
Thr194NNKfw	CACACGCTCAACNNKTGGGGCCGGCCGACG
Trp195NNKfw	CACACGCTCAACACCNNKGGCCGGCCGACG
Ala259NNKfw	GCGCTCGATGATGATGNNKATTCTCGTCGCCG
Val262NNKfw	GATGGCGATTCTCANNKGCCGCGCATGAGACC
Ser309NNKfw	CGCATGGGCGGCANNKATCATCGCCTGG
Ile310NNKfw	CGCATGGGCGGCTCGNNKATCGCCTGGCGG
Phe414NNKfw	GCAACATCGCCANNKCGCGGCCCAAAGAGC
80/81 rev	GTGATGTTGTCGCGGAAGACGGC

Thr105rev	TAGCCGTAGCGCTTCAGGATCGC
Val107rev	AGATTGTAGCCGTAGCGCTTCAGG
190/191rev	GCGGGTAAGCTCATCGGCATCTTCC
193/194/195rev	GGCGGCGGCGAAGCGGGTAAGC
Ala259rev	ATGTAGCTCTCGGGCACGATGTCGG
Val262rev	ATCATCGAGCGCATGTAGCTCTCGG
309/310rev	CAGGCATTCTCCACCGCG
Phe414rev	ACAGCCCCTCGAAGCTCTGACC
AX mid rev	TGGTGAATACTCGAAGTGCAACGCC
Leu81VKC fw	CCCCGGAATCGCGVKCGAAAAGATCACCC
His190Asn193rev	GGCGGCGGCGAAGCGGGTAAG
His190Asn193 fw	CTTACCCGCTTCGCCGCCMMCACGCTCRHMACCTGGGGCCGCGCCGA CG
Ala259GSA rev	ATGCGCGGCGACGAGAATTSCCATCATCATCGAGCGC

Primer	Position	Degenerate codon(s)	Encoded amino acids
Leu81VKC fw	81	VKC	R, G, I, L, S, V
His190Asn193 fw	190/193	MMC/RHM	N, H, P, T/A, N, D, E, I, K, T, V
Ala259GSA rev	259	GSA	A, G

For the first round of protein engineering, 14 megaprimers were synthesized by PCR reactions, using primer 309/310rev and the corresponding mutagenic primer for positions 80, 81, 105 and 107, primer Phe414rev and the corresponding mutagenic primer for positions 190, 191, 193, 194, 195, 259, 262 and finally primer AX mid rev and the corresponding mutagenic primer for positions 309, 310 and 414. Primers were designed to have similar melting temperatures and to generate amplicons of ~1 kb. PCR reactions were carried out in thin-walled 200 µL PCR tubes (final volume of 50 µL) employing Phusion DNA Polymerase and the provided GC-buffer following the manufacturer's instructions. For each single NNK library three PCRs were carried out employing a different annealing temperature (65, 68, 72°C), in order to try to get the

highest quality libraries.^[3] PCR products were then mixed, purified using a PCR clean up kit (Qiagen) and used as primers in the subsequent whole-plasmid reaction, following the original MEGAWHOP protocol. Template was digested following *DpnI* treatment for 6 h, the mixture purified once again using a PCR clean up kit and 5 μ L transformed into NEB-5-alpha competent cells according to manufacturer instructions. All colonies were scratched from agar plates and plasmid DNA isolated to assess library quality.^[3]

For the construction of the recombination library, degenerate codons were designed with SwiftLib (<http://rosettadesign.med.unc.edu/SwiftLib/>) to include the beneficial mutations identified at four positions and wild-type amino acids at those positions. Mutations were introduced by overlap extension PCR. First, mutations at position 81 were introduced using primers Leu81VKC fw and His190Asn193rev. The overlapping fragment bearing mutations at positions 190, 193 and 259 was produced using primers His190Asn193 fw and Ala259GSA rev. PCR products were purified (QIAquick Gel Extraction Kit) and the overlap extension PCR carried out employing 50 ng of each overlapping fragment and Leu81VKC fw/Ala259GSA rev as external primers (final reaction volume of 50 μ L). The produced megaprimer was purified (QIAquick Gel Extraction Kit) and the recombination library generated following the original MEGAWHOP protocol using 0.5 μ g megaprimer, 50 ng plasmid template and a final reaction volume of 50 μ L. All the subsequent steps were carried out as described above.

7.2.1.3 Protein Expression and Purification

Protein expression was accomplished as reported previously.^[1] Protein purification for crystallization trials was carried out resuspending cells in 50 mM HEPES buffer, 300 mM NaCl, 20 mM imidazole, pH 8.0 (buffer A) to a wet cell concentration of 200 mg/mL. The crude cell extract was prepared by ultrasonication (Bandelin Sonopuls sonicator) with 20 cycles of 15 s on and 45 s off, 40 % amplitude. Cell debris was removed by ultracentrifugation (48384 g, 30 min, 4 °C). P450 concentration was determined using a Cary 50 UV/Visible spectrophotometer (Agilent Technologies, Santa Clara, USA) according to the protocol by Omura and Sato,^[4] using an extinction coefficient of $\epsilon_{450-490} = 91 \text{ mM}^{-1} \text{ cm}^{-1}$ for the optical difference spectrum between the reduceCO bound/reduced spectra.

Immobilized metal ion affinity chromatography (IMAC) was employed as first purification step using Ni-agarose resin (Qiagen) in a gravity flow column following the manufacturer's instructions. To ensure maximal removal of protein impurities, the ratio of clarified supernatant to resin was adjusted so that the amount of P450 added exceeded the resin capacity (50 mg/mL) by 10 %. After 1 h incubation at 4 °C on a rotating mixer, the protein bound resin was loaded in a gravity flow column and washed with 20 CV (column volumes) buffer A before elution with 5 CV of 50 mM HEPES buffer, 300 mM NaCl, 300 mM imidazole, pH 8.0. Red coloured fractions containing purified P450 were collected, concentrated and desalted in 50 mM HEPES pH 8 (buffer C) using a PD-10 column (GE Healthcare) following the manufacturer's instructions.

Anion-exchange chromatography was carried out using a 6-ml RESOURCE Q column (GE Healthcare) on a ÄKTA Pure system (GE Healthcare) operated at 4 mL/min (4 °C). After an initial water wash (5 CV), the column was charged with 50 mM HEPES, 1 M NaCl, pH 8 (buffer D, 5 CV), equilibrated in buffer C (5 CV) and the sample loaded. A first column wash with buffer C (10 CV) was followed by elution with increasing concentration of buffer D (from 0 to 50 % over 20 CV) and a final wash (buffer D, 5 CV). Enzyme elution was monitored at 280 nm and red fractions displaying at least 50 % maximal absorbance of the main peak at 280 nm were collected for size-exclusion chromatography (SEC).

Enzyme solutions were concentrated to 1 mL using a 30,000 molecular weight cut-off filter (Vivaspin column, GE Healthcare) and loaded onto a HiLoad 16/600 Superdex 200 pg pre-equilibrated with 50 mM HEPES, 300 mM NaCl, pH 8. Size-exclusion chromatography was performed at 1 mL/min (4 °C) on a ÄKTA Pure system. Enzyme elution was monitored at 280 nm and red fractions displaying at least 50 % maximal absorbance of the main peak at 280 nm were analyzed by SDS-PAGE analysis (Figure S1), collected, concentrated and desalted in buffer C using a PD-10 column. Before carrying out crystallization trials, protein solutions were concentrated to 20 mg/mL, as determined by NanoDrop (Thermo Scientific), using $\epsilon_{280} = 67755 \text{ M}^{-1} \text{ cm}^{-1}$ and stored overnight at 4 °C. Typically, 4 mg of purified P450-TT were obtained per liter of culture following this procedure.

P450-AX variants were purified following a protocol similar to that described previously.^[1] IMAC was carried out using Ni-agarose resin (Qiagen) in a gravity flow column, employing 2

mL stationary phase per 800 mL broth (~9 g of wet cells) following the manufacturer's instructions. Enzyme solutions were concentrated and desalted in 0.1 M Tris, pH 8 using a PD-10 column (GE Healthcare). Protein aliquots were snap-frozen in liquid nitrogen and kept at -80°C overnight prior to characterization.

7.2.1.4 Expression and Screening of P450-AX Libraries in 96-well plates

Chemically competent BL21 (DE3) cells were transformed with the desired library (0.5 µL, 50-100 ng/µL DNA concentration) following the manufacturer's instructions and 30 µL of the resulting solution spread on a large petri dish containing LB agar supplemented with 50 µg/mL kanamycin for selection. Individual colonies were picked with sterile toothpicks and grown overnight in 96-deep well plates covered with a gas-permeable membrane (Breathe Easier, Sigma-Aldrich) into 150 µL LB supplemented with 50 µg/mL kanamycin. After overnight growth, expression cultures were inoculated with 40 µL of preculture into 960 µL of M9 prepared as previously described.^[1] Glycerol stocks were prepared by adding 40 µL of 21% sterile glycerol to the remainder of the preculture before transferring at -80°C. Cell growth was carried out in an Infors Multitron Pro for 2h 20 min at 30°C, 950 rpm and 75% relative humidity. When the OD₆₀₀ reached 0.8-1, gene expression was induced (0.4 mM IPTG and 0.5 mM 5-aminolevulinic acid, final concentration) and growth continued for 20 h.

Afterwards, cells were harvested by centrifugation (4000 rpm, 20 min, 4 °C) using a Beckman Coulter Avanti J-E centrifuge (JS-5.3 rotor) and cell pellets resuspended in 160 µL of 0.2 M potassium phosphate buffer (KPi), pH 8. A diclofenac/glucose mixture was added to the cell suspension (40 µL, 3 mM diclofenac, 10 mg/mL glucose, 5% DMSO, final concentration), and plates were incubated at 30°C in an Infors Multitron Pro incubator for 40 minutes.

Reactions were quenched with 300 µL methanol with 0.1% trifluoroacetic acid and spun down (4000 rpm, 30 min, 4 °C). Supernatants were filtered into Verex™ vials (Qsert) using 1 mL filter tips and subjected to HPLC analysis. For the first round of protein engineering 1302 clones were screened to ensure 95% coverage. The best variants of each library were selected and re-screened in triplicates.

For the second round of protein engineering, cell pellets were resuspended in 320 µL of 0.2 M KPi buffer and volumes adjusted accordingly to keep the final concentration of components

unchanged. A total of 1192 clones were screened to ensure ~85% coverage (the oversampling required for 85% coverage was calculated using CASTER 2.0). The best variants were selected and re-screened in triplicates.

Reversed-phase HPLC was carried out on an Agilent System (Santa Clara, CA, USA) equipped with a G4225A degasser, G1311A quaternary pump, a G1313A autosampler unit, a G1315C diode array detector and a G1316A column compartment with temperature set at 25°C. Analyses were carried out using a Thermo Scientific™ HyPURITY™ column (C18, 3 µm particle size, 2.1 mm diameter x 50 mm length) operating the system in isocratic mode (65% methanol with 0.1% TFA and 35% water with 0.1% TFA) at a flow of 0.5 mL/min for 2 minutes. Injection volume was 0.5 µL and chromatograms were monitored at 280 nm.

7.2.1.5 Turnover Frequencies, Coupling Efficiency, Total Turnover Numbers, Whole-Cell Activity and Thermal Stability Determination

Initial rates of 5-hydroxydiclofenac formation and NADPH depletion were determined in 200 µL 0.1 M KPi buffer, pH 8, 500 nM purified enzyme, 0.3 mM NADPH, 2.5 mM diclofenac and 2.5% DMSO (final concentrations). Assays were carried out in 96-well plates using a microtiter plate reader (Infinite M200 Pro, Tecan, Männedorf, Switzerland). Purified protein was added to an aliquoted mixture containing buffer/substrate and the reaction started after addition of the reduced cofactor. The mixture was rapidly shaken by pipetting up and down and NADPH depletion followed at 340 nm. Reactions were quenched after 4-8 minutes by adding 150 µL of methanol with 0.1% TFA to 150 µL of reaction mixture. Initial rates of NADPH oxidation were determined using an extinction coefficient of 6.22 mM⁻¹ cm⁻¹ at 340 nm for NADPH and a pathlength of 0.55 cm. Product formation rates were determined by HPLC analysis. The coupling efficiency is defined as the ratio between product formation rates and NADPH oxidation rates in presence of the substrate.

Total turnovers were determined in 0.1 M KPi buffer, pH 8, 25 nM purified enzyme, 1 mg/mL glucose dehydrogenase (CDX-901, Codexis), 1 mg/mL NADP⁺, 10 mg/mL D-glucose, 2.5 mM diclofenac and 2.5% DMSO (final concentrations). Reactions were carried out in 2 mL Eppendorf tubes (400 µL final volume) at 20°C and 250 rpm (New Brunswick™ Innova® 42R incubator) for 17-20 h. For HPLC analysis, an equal volume of methanol with 0.1% TFA was

added to the biotransformation, the sample vortexed and centrifuged for 5-10 minutes. The supernatant was filtered before HPLC analysis.

Whole-cell biotransformations for *in-vivo* activity determination were performed in 0.2 M KPi pH 8 with 40 mg/mL resuspended cells, 10 mg/mL glucose, 5 mM diclofenac and 10% DMSO (final concentrations). Reactions were carried out in 2 mL Eppendorf tubes (400 μ L final volume) at 37°C and 250 rpm (New Brunswick™ Innova® 44R incubator) for 30 minutes. Reactions were worked-up and analysed as reported above.

Finally, thermal stability was assessed as previously described^[1] by measuring residual activities of purified proteins towards 7-methoxycoumarin after incubation at high temperatures.

7.2.1.6 Crystallization, Data Collection and Structure Determination

Crystals of CYP116B46 (P450-TT) were obtained by mixing 200 nL of protein with 200 nL of condition A12 (0.1 M MES 6.5 22 % v/v PEG Smear Broad) from the BCS screen (Molecular Dimensions Ltd., Newmarket, UK). Drops were incubated at 4 °C for 24 hours prior to inspection. Single crystals suitable for data collection were cryo protected with the addition of 15% Peg 200 prior to plunge cooling in liquid nitrogen.

Data were collected from a single cryo protected crystal of P450-TT at beamline i04-1 (Diamond Light Source). All data were indexed, scaled and subsequently integrated with Xia2. Structure determination was initially performed by molecular replacement in Phaser using a search model derived from the previously solved cytochrome P450 structure (5GWE). A combination of automated and manual rebuilding and refinement in Phenix and COOT were used to produce a complete model. Validation with both Molprobity and PDB_REDO were integrated into the iterative rebuild process. Complete data collection and refinement statistics are available in Table S1. The atomic coordinates and structure factors of P450-TT structure have been deposited in the RCSB Protein Data Bank under the accession code 6GII.

7.2.1.7 Sequence Analysis

Sequence alignments were performed with ClustalW.^[5] The output files were then used to run ESPrint (for the characterized class VII P450s).^[6] Conservation analyses were carried out

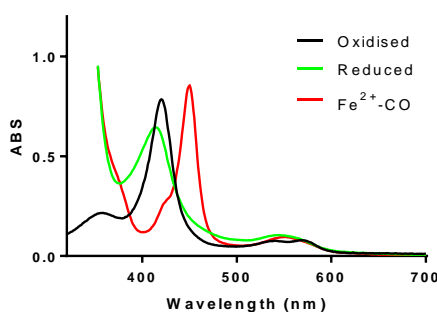
using the 3DM™ system (subfamily K9GPR5, <https://www.bio-product.nl/>).^[7] The 3DM™ system provides multiple sequence alignments of protein superfamilies based on multiple structure alignments, thereby generating a 3D numbering scheme (3DM standard numbering). When a 3DM standard numbering could not be assigned, class VII sequences were aligned using the CYPED database.^[8] Amino acid composition of characterized class VII P450 heme domains were computed starting from ClustalW alignments using MEGA7.^[9]

7.2.1.8 Structural Analysis

Structures were visualized and figures generated with PyMOL (the PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). The CAVER plug-in in PyMOL^[10] was employed using standard parameters and a probe radius of 1.4 Å to identify paths leading to the buried active site. Structural superposition were performed in COOT^[11] and the output files used to generate structure-based sequence alignments in Chimera (Match-> Align, residue-residue cutoff distance of 5 Å).

7.2.1.9 UV-visible Spectroscopic Properties of P450-TT Heme Domain

The figure below shows the UV-visible spectroscopic analysis for the P450-TT heme domain. The ferric heme domain has a Soret maximum at 419 nm, with α - and β -bands at ~570 and ~540 nm, respectively. Moreover, the heme delta band is visible at ~357 nm. Upon sodium-dithionite reduction, the intensity of the main Soret band is reduced and shifted at ~414 nm, while a single spectral feature is visible at ~550 nm. In the CO-bound spectrum, the expected shift of the Soret band at 450 nm is observed, indicating that cysteine thiolate is the heme-iron proximal ligand.



UV-visible spectroscopic analysis of P450-TT heme domain showing its oxidized (black line), sodium dithionite-reduced (green line) and CO-bound (red line) spectra.

7.2.2 Supporting Figures and Tables

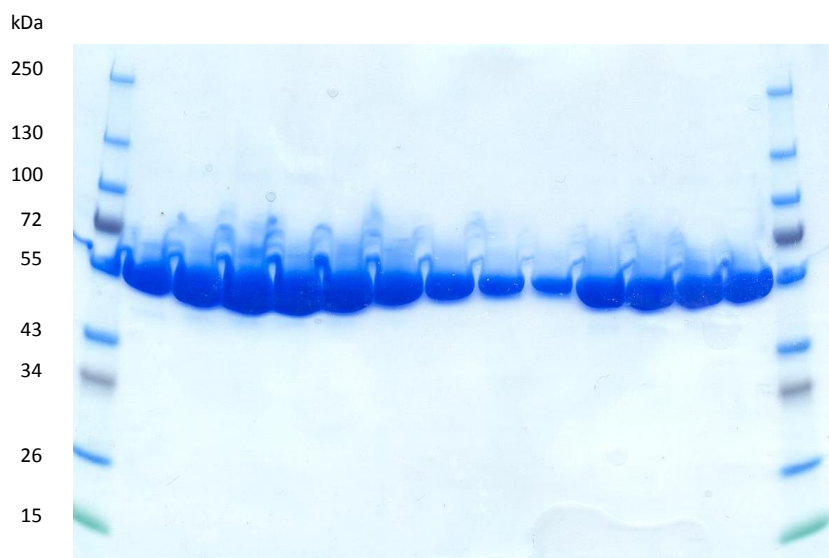


Figure S1. SDS-PAGE analysis of P450-TT selected fractions following a three-step purification protocol.

Extensive crystallization trials were also carried out after a two-step protein purification protocol (in which the intermediate anion-exchange chromatography step was omitted). Despite the high purity of the protein preparation (Figure) the following crystallization trials yielded crystals not suitable for X-ray diffraction. Thus, a three-step protein purification was implemented.

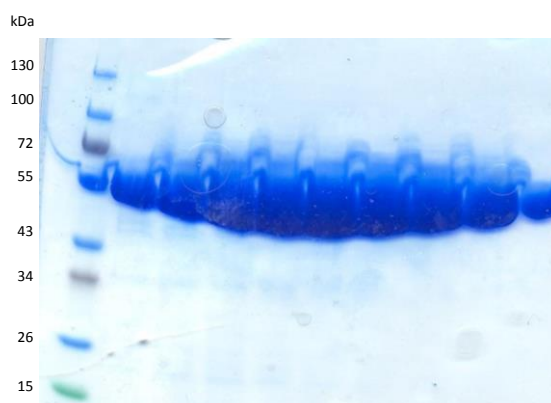


Figure S2. SDS-PAGE analysis of P450-TT selected fractions following a two-step purification protocol.

Table S1. Crystallographic Data and Refinement Statistics.

Data collection		Refinement	
Wavelength (Å)	0.9159	Reflections used in refinement	32914 (3124)
Space group	P 2 ₁ 2 ₁ 2 ₁	Reflections used for R-free	1639 (174)
Unit-cell parameters		R-work	0.1787 (0.3145)
<i>a</i> (Å)	53.2916	R-free	0.2290 (0.3166)
<i>b</i> (Å)	74.3841	Number of non-hydrogen atoms	4127
<i>c</i> (Å)	105.013	Macromolecules	3571
α (°)	90	Ligands	43
β (°)	90	Solvent	513
γ (°)	90	Protein residues	416
Resolution range	43.32-1.9 (1.968-1.9)	RMS deviation	
Total reflections	153818 (13406)	Bond lengths (Å)	0.006
Unique reflections	33286 (3242)	Bond angles (°)	1.07
Multiplicity	4.6 (4.1)	Ramachandran plot analysis (%)	
Completeness (%)	97.91 (95.01)	Ramachandran favored	97.34
$\langle I/\sigma(I) \rangle$	7.79 (2.71)	Ramachandran allowed	2.66
R-merge	0.1256 (0.4703)	Ramachandran outliers	0
R-meas	0.1423 (0.5426)	Rotamer outliers	2.38
R-pim	0.06539 (0.2647)	Average B-factor	17.25
CC1/2	0.989 (0.69)	Macromolecules	15.88
CC*	0.997 (0.904)	Ligands	8.91
		Solvent	27.52

Table S2. Conservation analysis of amino acid residues lining the putative substrate access channel. The first column indicates the amino acid occurring at the alignment position in P450-TT sequence. Numbering refers to the alignment in Figure S4. Active site residues are indicated by the yellow background. Where exact numbers are not given, a 3DM standard position could not be retrieved. Approximate figures are given after inspection of a class VII sequence alignment generated through the CYPED database.

Amino acid	Alignment position	Non polar, aliphatic						Polar, uncharged					Aromatic			Basic			Acid						
		Gly	Ala	Pro	Val	Leu	Met	Ile	Ser	Thr	Cys	Asn	Gln	Phe	Trp	Tyr	His	Lys	Arg	Asp	Glu				
Val	91	> 96%																							
Leu	92	> 96%																							
Tyr	111	No standard numbering available. Position occupied by either Tyr or Phe																							
Phe	113	2.3%							25.0%	72.7%						<0.1%									
Asn	114	19.0%	9.5%							2.1%						5.3%	1.0%	55.8%						4.2%	3.1%
Arg	115							4.2%						2.1%	3.2%							90.5%			
Thr	116	2.1%							1.0%	3.2%								1.0%	92.7%						
Leu	117							1.1%	88.4%	9.5%	1.1%														
Val	118	96.8%												3.2%											
Asn	119	2.1%												94.7%						3.2%					
Phe	206	> 96%																							
Val	221							91.7%	4.2%								1.0%	2.1%							
Ser	266	1.0%	2.1%												92.7%	1.0%	3.1%								
Met	267							1.0%	1.0%	91.7%	5.2%						1.0%								
Gln	269							92.7%					3.1%			3.1%			1.0%						
Ala	270	90.6%											7.3%			2.1%									
Ile	271	77.1%	1.0%							1.0%	14.6%								1.0%	1.0%			3.1%		
Ile	272							1.0%	17.7%	7.3%	72.9%	1.0%													
Val	273	2.1%						89.6%					5.2%			3.1%									
Ala	274	100%																							
Ala	275	5.2%	91.7%											3.1%											
Thr	278												100%												
Val	321	69.8%						2.1%			27.1%					1.0%									
Trp	324	100%																							

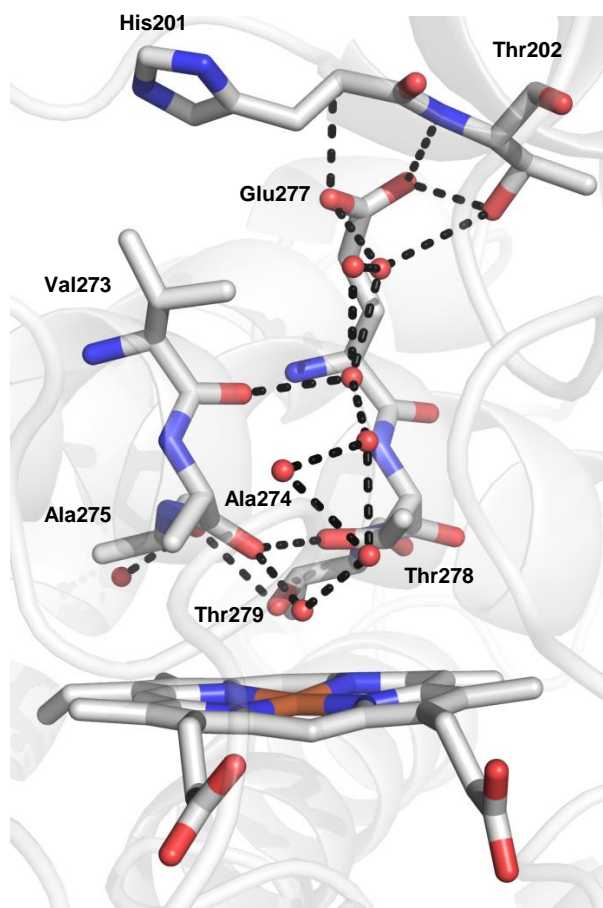


Figure S3. I-helix water channel. A water conduit can be seen running from Glu277 to Ala274. Residues interacting with Glu277 (His201 and Thr202) are also shown. Interactions are represented with black dashed lines, whereas water molecules are shown as red spheres.

CYP11B46 $\alpha 1$
1 10 20 30
CYP116B46METELKETARGT **C**PVAH.....GGSSVVG **G**CPVHRL**A**ED**D**DP**F**Q
CYP116B64MAHGDE...TFLRSPT **G**CPISPR**A**AA**R**PF**E**
CYP116B29 MGSEFVAAIRGSGDAAPAAATG **C**PFHP.....GV.DPS **G**RPISAN**A**AA**R**DP**F**H
CYP116B65MSTGS **C**PVDH.....APA.AAG **G**CPVSPR**A**SA**D**DP**F**G
CYP116B63MATSSTAPCPHGAVPN **G**CPISAD**A**AA**R**PF**D**
CYP116B2MSASVPASAPA **C**PVDH.....AA..LAG **G**CPVSN**A**AA**R**DP**F**G
CYP116B62MTTPARQPDGMAEG **C**PFHQPAQATSHGTSQTTLSPT **G**CPVSR**D**AA**R**DP**F**G
CYP116B1MPQTNAPASSGS **C**PIDH.....SALRAPN **G**CPIS**H**QAA**R**DP**F**G
CYP116B3MSASVP..ASA **C**PVDH.....AA..LAG **G**CPVSN**A**AA**R**DP**F**G
CYP116B4 MER.....TANPADVPAGGKSS **E**GKAG.....TPPAABA **Q**CPFS**K**MA**D**DA**F**A
CYP116B5MNVKVPMHGQ **C**SFHA..... **L**E**L**Q**R**K**G**KN**D**DM**F**A

CYP116B46 $\alpha 2$ $\alpha 3$ $\alpha 4$ $\beta 1-1$ $\beta 1-2$ αB $\beta 2-3$
40 50 60 70 80 90
CYP116B46 **D**AY**M**AD**D**PA**Q**F**V**RV**A**RE**Q**V**I**F**I**Y**A**FK**L**NY**W**V**V**TR**Y**DI**K**IK**I**FR**D**EV**T**FS**P**SV**L**Q**S**FA**Q**PS
CYP116B64 **R**F**Y**Q**D**DP**A**AL**A**W**A**RA**E**DP**V**F**Y**AD**D**IG**Y**W**V**TR**H**ED**V**K**A**V**F**RD**N**IT**F**SP**R**I**A**LE**K**IT**P**T
CYP116B29 **E**T**F**Q**D**DP**A**GS**L**RW**S**RD**Q**EP**V**F**Y**SP**R**LG**Y**W**V**TR**Y**ED**V**K**A**V**F**RD**N**IT**F**SP**S**I**A**LE**K**IT**V**VC
CYP116B65 **F**E**Y**Q**D**DP**A**AR**W**SR**D**Q**E**P**V**F**Y**SP**R**LG**Y**W**V**TR**Y**DD**V**K**A**V**F**RD**N**IT**F**SP**S**I**A**LE**K**IT**V**VS
CYP116B63 **G**P**Y**Q**D**DP**A**EL**R**SR**A**ER**D**EP**V**F**Y**AD**D**IG**Y**W**V**TR**Y**ED**V**K**A**V**F**RD**N**IT**F**SP**A**I**A**LE**K**IT**P**T
CYP116B2 **S**AY**Q**D**D**PA**E**SL**R**WR**S**RD**Q**EP**V**F**Y**SP**R**LG**Y**W**V**TR**Y**ED**V**K**A**V**F**RD**N**IT**F**SP**A**I**A**LE**K**IT**V**VS
CYP116B62 **A**SY**Q**D**D**PA**E**SL**R**WR**S**RD**Q**EP**V**F**Y**SP**K**LG**Y**W**V**SR**Y**EV**K**A**V**F**R**D**N**IT**F**SP**A**I**A**LE**K**IT**P**AP
CYP116B61 **D**GY**Q**D**D**PE**Y**V**R**W**S**RA**Q**EP**V**F**Y**SP**K**LG**Y**W**V**TR**Y**DD**I**K**A**I**F**RD**N**IT**F**SP**S**I**A**LE**K**IT**P**T
CYP116B3 **P**AY**Q**D**D**PA**E**SL**R**WR**S**RD**E**EP**V**F**Y**SP**E**LG**Y**W**V**TR**Y**ED**V**K**A**V**F**RD**N**IT**F**SP**A**I**A**LE**K**IT**V**VS
CYP116B4 **G**F**Y**Q**D**DP**A**EL**R**WR**S**RD**Q**EP**V**F**Y**SP**N**LG**Y**W**V**SR**Y**DD**I**K**A**V**F**RD**N**IT**F**SP**R**I**A**LE**K**IT**P**AT
CYP116B5 **V**P**Y**Q**D**DP**A**LA**L**KE**F**RA**Q**EP**I**F**I**F**S**E**A**MG**Y**W**I**V**T**RY**E**D**V**K**A**I**F**RD**E**I**T**FS**A**C**N**A**L**E**K**IT**P**S**C**

CYP116B46 $\alpha B'$ $\eta 1$ αC $\eta 2$ $\eta 3$ αD
100 110 120 130 140 150
CYP116B46 **A**EV**R**Q**V**LE**R**Y**G**Y**A**R**T**L**V**NE**D**E**P**H**I**ER**R**R**V**LM**E**P**F**AS**E**H**L**A**H**E**P**M**V**R**R**I**V**R**R**A**V**N**R**F
CYP116B64 **E**EV**Q**AI**L**K**R**Y**G**Y**N**L**Q**R**T**L**V**NE**D**E**P**H**M**ER**R**RL**L**MA**D**FL**P**ER**L**A**H**E**P**M**V**R**R**I**V**TR**A**H**M**D**R**F
CYP116B29 **D**E**A**T**A**V**L**E**R**Y**G**Y**A**R**T**L**V**NE**D**E**P**H**M**ER**R**RL**A**LM**E**PF**T**PA**Q****L**A**H**E**P**L**V**R**R**I**V**TR**H**V**D**R**I**
CYP116B65 **A**E**A**T**A**L**L**A**K**Y**G**Y**A**M**R**T**L**V**N**E**D**E**P**H**M**ER**R**RL**V**LM**H**PF**T**E**Q**L**V**H**H**E**P**M**V**R**R**I**V**T**R**E**Y**V**D**R**F**
CYP116B63 **A**E**A**Q**A**I**L**R**K**Y**N**Y**N**L**Q**R**T**L**V**NE**D**E**P**H**M**ER**R**RL**L**L**D**DF**L**P**E**R**L**A**H**E**P**M**V**R**R**I**V**C**R**D**Y**M**D**R**F**
CYP116B2 **A**E**A**T**A**L**L**A**R**Y**D**Y**A**M**A**R**T**L**V**NE**D**E**P**H**M**ER**R**RL**A**LM**D**PF**T**PE**K**E**L**A**H**E**A**M**V**R**R**I**V**T**R**E**Y**V**D**R**F**
CYP116B62 **P**E**A**V**R**I**L**E**S**Y**G**F**A**M**R**R**T**L**V**NE**D**E**P**H**M**ER**R**RL**L**M**E**AF**L**P**E**N**L**E**K**H**E**A**M**V**R**R**I**V**R**S**Y**M**D**H**F**
CYP116B61 **E**A**A**N**A**V**L**A**S**Y**G**Y**A**M**R**T**L**V**N**E**D**E**P**H**M**ER**R**RL**A**LM**E**PF**T**PA**E****L**A**H**E**P**M**V**R**R**I**V**T**R**E**Y**V**D**R**F**
CYP116B3 **E**A**T**A**T**L**A**R**Y**D**Y**A**M**A**R**T**L**V**N**E**D**E**P**H**M**ER**R**RL**A**LM**D**PF**T**PE**K**E**L**A**H**E**A**M**V**R**R**I**V**T**R**E**Y**V**D**R**F**
CYP116B4 **P**E**A**M**E**V**L**K**S**Y**G**Y**A**M**R**T**L**V**N**E**D**E**P**H**M**ER**R**RL**A**LM**H**FL**P**DN**L**E**A**R**Q**E**M**V**R**R**I**V**T**R**E**K**I**D**A**T
CYP116B5 **P**E**A**L**K**I**L**E**K**Y**G**Y**M**R**T**L**V**NE**D**E**P**H**M**ER**R**RL**A**LM**D**AF**T**PE**Q**N**L**E**F**H**O**H**E**V**R**R**I**V**T**R**R**K**R**V**D**G**F**

CYP116B46 $\beta 3-1$ $\alpha E'$ αE $\alpha F'$ $\eta 4$ $\alpha F''$ αG
160 170 180 190 200 210
CYP116B46 **I**D**T**GR**A**D**L**V**D**Q**M**I**W**E**V**P**F**T**V**A**I**H**F**LG**V**D**D**D**R**E**K**M**R**R**F**A**I**A**H**T**V**N**A**F**G**R**P**S**E**E**Q**L**A**V**A**E
CYP116B64 **I**D**R**G**H**A**D**L**V**A**E**I**F**R**D**V**P**L**K**V**A**I**H**FL**G**V**A**E**E**D**A**DE**L**T**R**F**A**A**H**L**I**N**T**W**G**R**P**E**E**Q**E**I**A**E
CYP116B29 **I**D**T**G**K**A**D**L**V**E**E**M**L**Y**D**V**P**L**T**V**A**I**H**FL**G**V**P**E**D**D**M**A**M**L**R**K**Y**S**V**A**H**V**V**N**T**W**G**R**P**E**E**Q**V**A**V**A**E**
CYP116B65 **V**D**A**G**K**A**D**L**V**E**E**M**L**W**E**V**P**L**T**V**A**I**H**FL**G**V**P**E**D**D**M**A**L**R**E**Y**S**I**A**H**T**V**N**T**W**G**R**P**S**E**E**Q**V**A**V**A**E**
CYP116B63 **I**D**R**G**H**A**D**L**V**N**D**I**F**Y**E**I**P**L**T**V**A**I**H**FL**G**V**P**D**E**G**A**E**R**L**R**Q**F**A**V**A**H**L**I**N**T**W**G**R**P**E**E**Q**I**A**E**
CYP116B2 **V**ES**G**K**A**D**L**V**E**E**M**L**W**E**V**P**L**T**V**A**I**H**FL**G**V**P**E**E**D**M**A**M**R**K**Y**S**I**A**H**T**V**N**T**W**G**R**P**A**E**E**Q**V**A**V**A**E
CYP116B62 **I**D**K**G**R**A**D**L**V**E**E**M**P**R**E**I**P**M**T**V**A**I**H**FL**G**V**P**S**E**D**A**K**E**L**R**K**F**S**V**A**H**L**I**N**T**W**G**R**P**E**E**Q**K**I**A**E
CYP116B61 **I**D**T**GR**A**D**L**V**E**E**M**L**W**E**V**P**L**T**V**A**I**H**FL**G**V**P**E**E**D**M**L**L**R**Q**Y**S**I**A**H**T**V**N**T**W**G**R**P**E**E**Q**V**A**V**A**E**
CYP116B3 **V**ES**G**K**A**D**L**V**E**E**M**L**W**E**V**P**L**T**V**A**I**H**FL**G**V**P**E**E**D**M**A**M**R**K**Y**S**I**A**H**T**V**N**T**W**G**R**P**A**E**E**Q**V**A**V**A**E
CYP116B4 **I**D**S**G**R**V**D**L**V**E**A**M**L**Y**E**V**P**L**N**V**A**I**H**FL**G**V**P**E**D**D**I**A**I**L**K**N**F**S**V**A**H**S**V**N**T**W**G**K**P**T**E**E**Q**V**A**I**A**E
CYP116B5 **I**Y**K**G**R**A**D**L**V**Q**E**M**L**W**E**I**P**L**M**V**A**I**H**FL**G**V**P**E**D**D**M**Q**E**L**R**K**F**S**V**A**H**T**V**N**T**W**G**R**P**E**E**Q**L**E**V**A**E**

CYP116B46 αG αH αI
220 230 240 250 260 270
CYP116B46 **T**VG**Q**F**W**Q**F**CG**E**V**L**E**K**M**R**R**T**A**D**CG**W**M**R**YS**I**R**Q**K**L**Y**P**D**V**VT**D**S**V**L**H**S**M**M**A**I**V**A**A**H**E**T**T**
CYP116B64 **N**VGR**F**W**Q**K**A**N**A**I**L**D**R**M**I**AD**P**SG**W**M**Y**ET**V**R**Q**H**F**R**H**P**D**I**V**PS**Y**M**R**S**M**M**A**I**V**A**A**H**E**T**T**
CYP116B29 **S**VGR**F**W**R**F**A**GE**V**L**D**R**M**R**D**DP**S**CG**W**M**P**Y**A**I**R**L**Q**P**M**PD**V**IT**D**S**V**L**H**S**I**M**M**A**I**V**A**A**H**E**T**
CYP116B65 **N**VG**K**F**W**Q**Y**AG**T**V**L**D**K**L**R**DD**P**SG**H**W**P**Y**S**I**R**A**Q**R**E**H**P**D**V**IT**D**S**V**L**H**S**I**M**M**A**I**V**A**A**H**E**T**
CYP116B63 **N**VGR**F**W**Q**T**A**NE**I**L**D**D**M**I**A**NP**D**CG**W**M**Y**ET**V**R**Q**H**F**K**H**P**D**I**V**PS**Y**M**R**S**M**M**A**I**V**A**A**H**E**T**T**
CYP116B2 **A**VGR**F**W**Q**Y**A**GT**V**L**E**K**M**R**D**DP**S**CG**W**M**P**Y**G**I**R**K**Q**R**E**M**P**D**V**VT**D**S**V**L**H**S**M**M**A**I**V**A**A**H**E**T**T**
CYP116B62 **D**VG**Q**F**W**K**T**A**T**I**L**D**R**M**R**ED**P**T**CE**W**M**Y**D**S**I**R**M**H**N**Q**H**P**EV**V**P**ES**V**L**R**S**M**M**A**I**V**A**A**H**E**T**T**
CYP116B61 **A**VGR**F**W**Q**L**A**GR**I**L**D**K**M**R**D**DP**S**CG**W**M**Q**Y**G**L**R**K**Q**R**E**L**P**EV**VT**D**S**V**L**H**S**M**M**A**I**V**A**A**H**E**T**
CYP116B3 **A**VGR**F**W**Q**Y**A**GT**V**L**E**K**M**R**D**DP**S**CG**W**M**P**Y**G**I**R**M**Q**Q**M**PD**V**VT**D**S**V**L**H**S**M**M**A**I**V**A**A**H**E**T**T**
CYP116B4 **D**VG**Q**F**W**N**Y**A**G**K**I**E**K**M**R**KE**P**D**CT**W**M**H**E**T**I**R**K**N**A**EM**P**D**I**V**P**D**S**V**H**S**M**M**A**I**V**A**A**H**E**T**T**
CYP116B5 **C**VG**Q**F**W**E**Y**SG**R**V**L**E**K**M**K**N**N**PE**CR**W**M**Y**D**M**L**A**K**N**R**V**M**PE**V**VT**D**N**V**L**H**S**M**M**A**I**V**A**A**H**E**T**T**

CYP116B46 αJ $\eta 5$ αK $\beta 1-4$ $\beta 2-1$ $\beta 2-2$
280 290 300 310 320 330
CYP116B46 **V**F**A**T**N**A**L**K**T**I**L**L**E**H**T**V**R**E**I**C**A**D**P**S**I**P**A**A**A**E**E**C**L**R**Y**NG**P**V**A**Q**W**R**R**I**T**R**E**V**E**V**G**V**R**I
CYP116B64 **I**L**A**T**A**N**A**F**R**I**L**L**E**R**R**D**A**NE**L**C**A**NP**GL**I**P**S**A**V**E**E**C**L**R**M**G**S**I**A**W**R**R**I**A**T**R**E**A**V**G**V**T**I
CYP116B29 **A**N**A**A**N**A**I**R**L**L**E**N**R**SV**BE**I**C**AD**Q**R**L**I**P**N**A**V**E**E**C**L**R**H**G**S**V**A**W**R**R**I**A**T**T**K**I**G**D**V**E**I
CYP116B65 **A**N**A**A**N**A**I**K**L**L**L**E**N**R**S**V**BE**I**C**AD**P**S**L**I**P**N**A**V**E**E**C**L**R**H**G**S**V**A**A**W**R**R**I**L**V**T**R**D**T**T**V**G**G**M**S**L
CYP116B63 **S**N**A**T**A**N**A**F**R**I**L**L**T**H**R**D**A**NE**I**C**EN**P**AL**I**P**S**A**V**E**E**C**L**R**V**G**S**I**A**W**R**R**I**A**T**D**A**E**L**G**G**V**I
CYP116B2 **A**N**A**S**A**N**A**F**K**L**L**L**E**N**R**A**V**BE**I**C**AD**P**S**L**I**P**N**A**V**E**E**C**L**R**H**G**S**V**A**A**W**R**R**I**A**T**D**A**T**R**I**G**D**V**I**
CYP116B62 **A**F**A**T**N**A**F**R**I**L**L**S**N**R**A**S**W**RE**I**C**EN**P**AL**I**P**S**A**I**E**E**C**L**R**A**G**S**V**A**W**R**R**I**A**T**D**A**D**V**G**G**V**I
CYP116B61 **A**N**A**S**A**N**A**I**K**L**L**L**Q**H**P**D**V**W**RE**I**C**ED**P****AL**I**P**N**A**V**E**E**C**L**R**H**G**S**V**A**A**W**R**R**I**L**V**T**R**D**T**T**V**G**G**M**S**L
CYP116B3 **A**N**A**S**A**N**A**F**K**L**L**L**E**N**R**SV**BE**I**C**AD**P**S**L**I**P**N**A**V**E**E**C**L**R**H**G**S**V**A**A**W**R**R**I**V**A**T**I**D**T**R**I**G**D**V**I**
CYP116B4 **S**L**A**S**A**G**M**F**K**I**L**L**T**H**R**Q**A**W**OD**I**C**ED**P**S**L**I**P**N**A**V**E**E**C**L**R**Y**G**S**I**V**A**W**R**R**O**A**T**A**T**R**I**G**D**V**I**
CYP116B5 **A**L**A**S**A**N**A**L**K**L**L**L**A**D**R**K**V**K**K**I**C**D**NP**Q**L**I**P**N**A**V**E**E**C**L**R**H**G**S**V**A**W**R**R**O**V**T**I**E**S**E**V**S**G**V**K**I

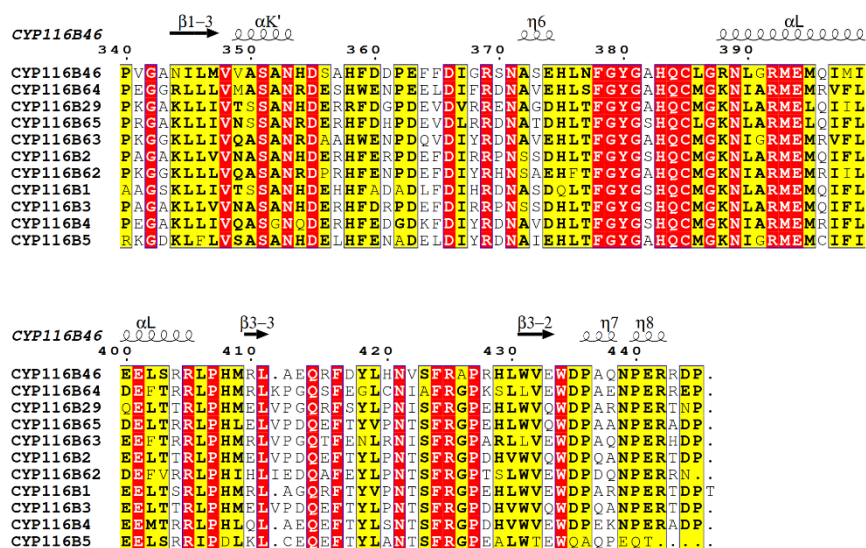


Figure S4. Sequence alignment of characterized class VII P450s. Secondary structural elements are named after Hasemann *et al.*^[12] α -helices and β -strands are indicated by the corresponding letters and η indicated 3_{10} helices. Columns with invariant residues are highlighted by the red background whereas columns highlighted in yellow indicate equivalent residues (based on physico-chemical properties, cut-off=70%). The figure was generated and secondary structural elements annotated with ESPrint.

Table S3. Conservation analysis of active site amino acid residues between class VII P450s. The first column indicates the amino acid occurring at the alignment position in P450-TT sequence. Numbering refers to the alignment in Figure S4. Where exact numbers are not given, a 3DM standard position could not be retrieved. Approximate figures are given after inspection of a class VII sequence alignment generated through the CYPED database.

Alignment position	Non polar, aliphatic							Polar, uncharged					Aromatic			Basic			Acid	
	Gly	Ala	Pro	Val	Leu	Met	Ile	Ser	Thr	Cys	Asn	Gln	Phe	Trp	Tyr	His	Lys	Arg	Asp	Glu
Val 91	> 96%																			
Leu 92				> 96%																
Thr 116	2.1%			1.0%		3.2%		1.0%	92.7%											
Val 118	96.8%											3.2%								
His 201				1.0%	1.0%							1.0%	3.1%		89.6%		4.2%			
Thr 202	1.0%	5.2%		1.0%	1.0%	2.1%		2.1%	87.5%											
Asn 204				1.0%	2.1%								4.2%	3.1%						
Ala 205								> 96%												
Phe 206													> 96%							
Ala 270		90.6%						7.3%			2.1%									
Val 273		2.1%		89.6%									5.2%	3.1%						
Ala 274		100%																		
Glu 277																			100%	
Thr 278									100%											
Pro 320			1%					99.0%												
Val 321				69.8%	2.1%		27.1%		1.0%											
Trp 324													100%							
Phe 424												100%								

Table S4. Structures homologous to P450-TT as identified by DALI search. Top ranking sequences have been selected setting a DALI Z-score cutoff equal to $n/10 - 4$ (n = where n is the number of residues in the query structure).^[13]

Protein	PDB code	Z	rmsd	%id	Function or accepted substrates
CYP288A2 (CreJ)	5xjn-A	50.8	1.9	31	<i>p</i> -cresol biodegradation pathway
CYP107H1 (P450 _{Biol})	3eje-D	46.6	2.1	25	Pimelic acid synthesis through oxidative degradation of ACP-bound fatty acids (biotin biosynthetic pathway)
CYP109A2	5ofq-A	46	1.9	23	Vitamin D3
CYP109E1	5l91-A	45.3	2	23	Steroids, terpenes, statins
CYP142	2xkr-A	45.1	2.2	25	Cholesterol
P450 PksS	4yzr-A	45.1	2	23	Bacillaene metabolism
CalO2	3buj-A	44.8	2.2	24	Putative orsellinic acid (anticipated to be ACP-bound) oxidase (calicheamicin biosynthetic pathway)
CYP107L2	5cje-A	44.5	2.1	28	n.d.
CYP107L1 (PikC)	2bvj-B	44.4	2	28	Macrolides (pikromycin/methymycin biosynthetic pathway)
CYP109B1	4rm4-A	44.3	1.9	26	Testosterone, terpenes, compactin, fatty acids

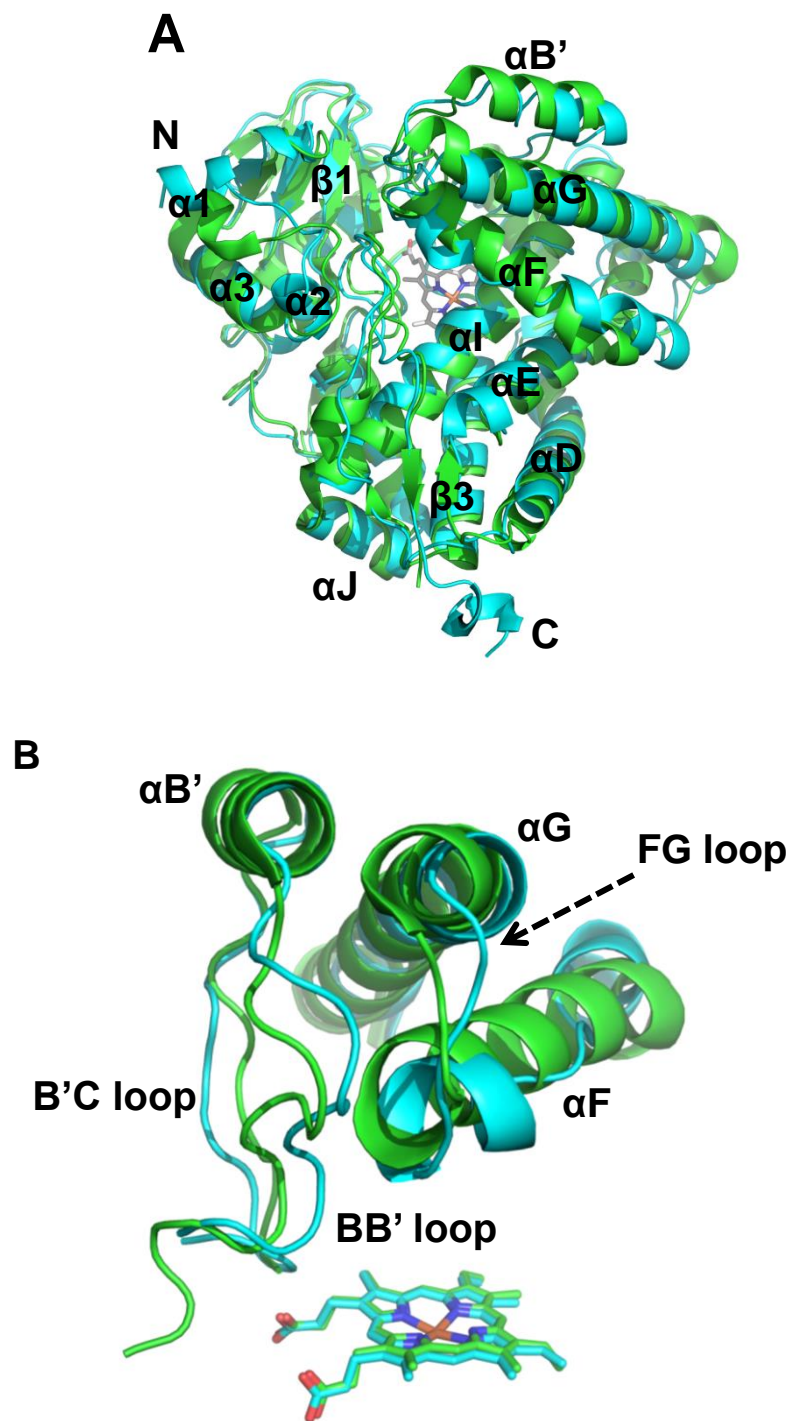


Figure S5. Secondary structure matching between P450-TT (cyan) and P450-CreJ (PDB 5XJN, chain A, green). A) Overall superposition. B) Detailed view of regions showing significant deviation in the structural alignment. The heme is represented as sticks.

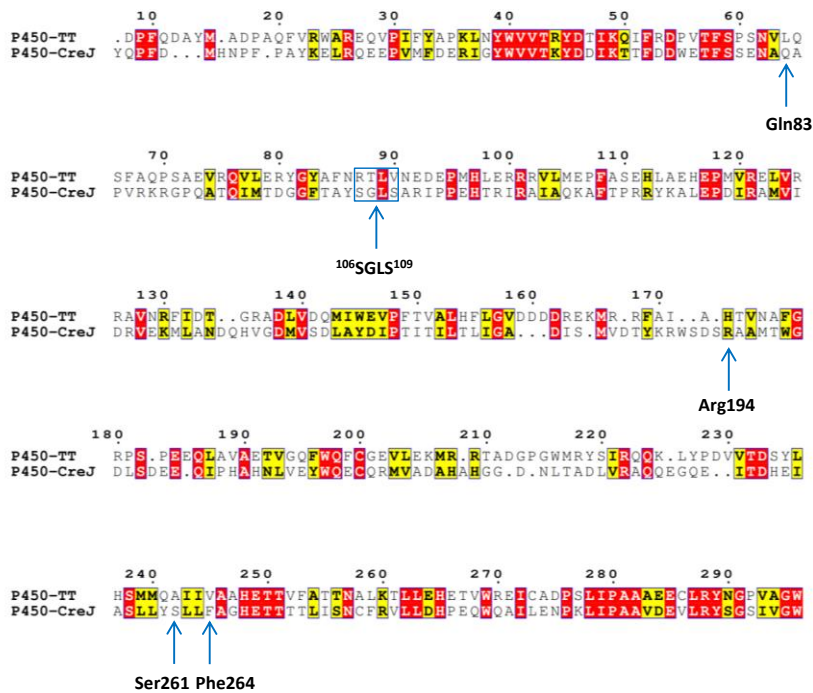


Figure S6. Structure-based sequence alignment between P450-TT and P450-CreJ. Critical residues for phosphorylated *p*-cresol binding according to [14].

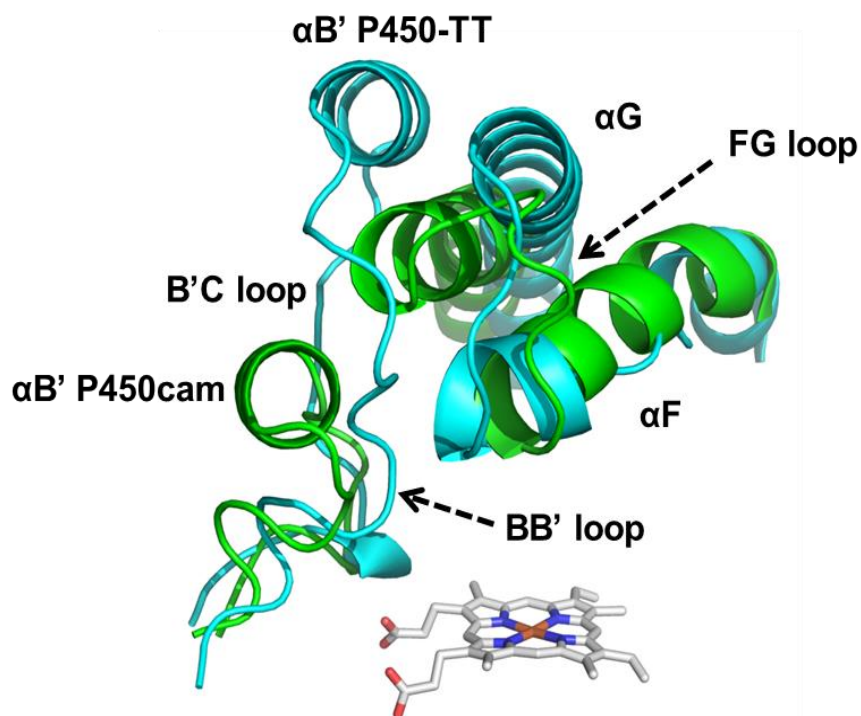


Figure S7. Secondary structure matching between P450-TT (cyan) and P450cam (PDB 2CPP, green).

Table S5. P450-AX putative active site residues targeted for semi-rational mutagenesis. The corresponding residues in the P450-TT active site are indicated.

P450-TT	P450-AX
Tier1	
Val118	Val107
Pro320	Ser309
Val321	Ile310
Tier2	
Val91	Ala80
Thr116	Thr105
Ala270	Ala259
Phe424	Phe414
Tier3	
Leu92	Leu81
Phe206	Trp195
Val273	Val262
Tier4	
His201	His190
Thr202	Thr191
Asn204	Asn193
Ala205	Thr194

Table S6. Thermal stability of wild-type and evolved P450-AX assessed by residual activity towards 7-methoxycoumarin after incubation at elevated temperatures and comparison of protein expression levels as determined by CO-difference spectroscopy of the corresponding cell lysates obtained by sonication. The error is the standard error.

		T_{50} (°C)	P450 content (nmol _{P450} g _{cww} ⁻¹)
	WT	57.7±0.2	27
Active-site testing	N193V	55.7±0.3	20
	N193A	60±0.3	19
	H190T	54.6±0.2	28
	A259G	55.1±0.3	23
	L81I	55.8±0.2	24
	L81V	55.9±0.2	27
	N193E	57.4±0.2	19
	L81G	56.3±0.4	20
Recombination	L81S/N193E	55.5±1.3	22
	L81V/N193A	56.4±0.2	22
	L81G/N193A	55.3±0.2	20

Table S7. Total turnovers, turnover frequency, coupling efficiency and *in-vivo* activity of P450-AX L81S.

	TOF (min ⁻¹)	Coupling (%)	TTN	Whole-cell activity ($\mu\text{mol}_{\text{product}} \text{min}^{-1} \text{g}_{\text{cww}}^{-1}$)	P450 content (nmol _{P450} g _{cww} ⁻¹)
WT	6.7	48.4	7725	0.7	27
L81S	13.1	65.7	13283	0.9	39

Table S8. *In-vivo* activity of selected variants originating from the recombination library.

	Whole-cell activity ($\mu\text{mol}_{\text{product}} \text{min}^{-1} \text{g}_{\text{cww}}^{-1}$)
L81S/N193E	2.4
L81V/N193A	2.5
L81G/N193A	2.9

7.2.3 References

- [1] M. Tavanti, J. L. Porter, S. Sabatini, N. J. Turner, S. L. Flitsch, *ChemCatChem* 2018, 10, 1042–1051.
- [2] K. Miyazaki, *MEGAWHOP Cloning: A Method of Creating Random Mutagenesis Libraries via Megaprimer PCR of Whole Plasmids*, Elsevier Inc., 2011.
- [3] S. Kille, C. G. Acevedo-Rocha, L. P. Parra, Z. G. Zhang, D. J. Opperman, M. T. Reetz, J. P. Acevedo, *ACS Synth. Biol.* 2013, 2, 83–92.
- [4] R. Omura, T. Sato, *J. Biol. Chem.* 1964, 239, 2370–2378.
- [5] F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, et al., *Mol. Syst. Biol.* 2011, 7, 1–6.
- [6] X. Robert, P. Gouet, *Nucleic Acids Res.* 2014, 42, 320–324.
- [7] R. K. Kuipers, H. J. Joosten, W. J. H. Van Berkel, N. G. H. Leferink, E. Rooijen, E. Ittmann, F. Van Zimmeren, H. Jochens, U. Bornscheuer, G. Vriend, et al., *Proteins Struct. Funct. Bioinf.* 2010, 78, 2101–2113.
- [8] M. Fischer, M. Knoll, D. Sirim, F. Wagner, S. Funke, J. Pleiss, *Bioinformatics* 2007, 23, 2015–2017.
- [9] S. Kumar, G. Stecher, K. Tamura, *Mol. Biol. Evol.* 2016, 33, 1870–1874.
- [10] M. Petrek, M. Otyepka, P. Banás, P. Kosinová, J. Koca, J. Damborský, *BMC Bioinformatics* 2006, 7, 316.
- [11] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 2010, 66, 486–501.
- [12] C. A. Hasemann, R. G. Kurumbail, S. S. Boddupalli, J. A. Peterson, J. Deisenhofer, *Structure* 1995, 2, 41–62.
- [13] L. Holm, S. Kääriäinen, P. Rosenström, A. Schenkel, *Bioinformatics* 2008, 24, 2780–2781.
- [14] L. Du, S. Dong, X. Zhang, C. Jiang, J. Chen, L. Yao, X. X. Wang, X. Wan, X. Liu, X. X. Wang, et al., *Proc. Natl. Acad. Sci. U. S. A.* 2017, 114, 201702317.

7.3 One-Pot Biocatalytic Double Oxidation of α -Isophorone for the Synthesis of Ketoisophorone

7.3.1 Experimental

7.3.1.1 Materials

Chemicals, solvents, and carbon monoxide for CO difference spectroscopy were obtained from Sigma-Aldrich (Poole, Dorset, UK). Gases for GC/FID analysis were purchased from BOC gases (Guildford, UK). (*R*)-HIP was kindly provided by DSM (DSM Ahead R&D B.V.-Innovative Synthesis, Urmonderbaan 22, NL-6167 RD Geleen, The Netherlands). Competent cells and enzymes for molecular biology were purchased from New England Biolabs (NEB). M9 minimal salts (5X) were purchased from Sigma-Aldrich, reconstituted by stirring the recommended amount of powder in water and sterilized by autoclaving. 40% glucose (w/v), antibiotics 1000X, 1 M MgSO₄, 1 M CaCl₂, and 25% (w/v) FeCl₃ were prepared in dH₂O and filter sterilized through a 0.2 μ m syringe filter.

7.3.1.2 Molecular Biology Methods

Primers synthesis and DNA sequencing were performed by Eurofins Genomics. Ordered primer sequences are given below (mismatching bases are given in red).

Primer	5'->3' sequence
F87W for, T _m =57°C	GAGTGCCCGTGGATCCCTCGTGAAGC
F87W rev, T _m =57°C	GCTGGAAAAGTGCGGTAATC
L244A for, T _m =59°C	AGGATGTGTGGCGCGTTACTGGTCGGC
L244A rev, T _m =61°C	CTTGGCTTCGTCACCTGGTGATCG
L244A-V247L for, T _m =63°C	AGGATGTGTGGCGCGTTACTGCTCGGCGGCCTGGATAC
AvrII rev, T _m =65°C	TAGTCTCCTAGGTCAGAGTCGAGGGCCAGCc
AfIII rev, T _m =65°C	TCGTCTCTTAAAGTCAGAGTCGAGGGCCAGCC
T7 for modified, T _m =62°C	TAATACGACTCACTATAGGGAGACCACAACGG
NdeI ADH10 for, T _m =63°C	ACGTAGCATATGACGACTACTTCAAACGCGCTTGTC
NcoI CmADH for, T _m =63°C	ATGCTACCATGGGGATGACGACTACTTCAAACGCGCTTGTC
AfIII CmADH rev, T _m =60°C	ATGTATCTTAAGCAATCAAGCCATTGTCGACCAC
XhoI ADH10 rev, T _m =60°C	ACGTCACTCGAGTTAAGCAATCAAGCCATTGTCGACCAC
AvrII ADH10 rev, T _m =60°C	AGTCAGCCTAGGTTAAGCAATCAAGCCATTGTCGACCAC
NdeI Sporo for, T _m =60°C	CTGGATCATATGGCCAAAATCGATAATGCCGTG
XhoI Sporo rev, T _m =61°C	TGCATGCTCGAGTTAGGCTGTTTCGCTACCAACCAGG

Inverse PCR reactions were carried out using Eppendorf Mastercycler Gradient thermal cyclers according to NEB guidelines, followed by *DpnI* digestion before carrying out ligation reaction for 1 h at 25°C with T4 DNA ligase and polynucleotide kinase, according to manufacturer's instructions. NEB 5-alpha competent *E. coli* (high efficiency) were then transformed according to the manufacturer's instruction and the sequence verified by plasmid sequencing. Expression plasmids were generated by standard restriction cloning.

P450cam-RhFRed site-directed mutants were made starting from variants in the previously developed libraries using the appropriate primers reported in the table. An *AvrII* restriction site was added to P450cam-RhFRed by PCR using the primers "T7 for modified" and "AvrII rev" and the sequence cloned in pCDF-1b using *NcoI* and *AvrII* restriction sites.

For the two-vector strategy for coexpression (or ADHs expression trials), SSCR and Cm-ADH10 were cloned into the pET28a vector, using *NdeI* and *XhoI* restriction sites that were introduced by PCR using primers "NdeI Sporo for" and "XhoI Sporo rev" or "NdeI ADH10 for" and "XhoI ADH10 rev", respectively.

Plasmid carrying genes encoding for the selected ADHs were kindly provided by c-LEcta GmbH, Leipzig, Germany.

7.3.1.3 Protein Expression and Purification

Plasmids (pET-14b, pCDF-1b, pET-28a) carrying genes encoding P450cam-RhFRed variants were stored at 4°C. Chemically competent *E. coli* BL21 (DE3) cells were transformed by heat shock according to manufacturer's instructions. Transformants were plated on LB agar with added antibiotic (150 µg/ml ampicillin or 50 µg/ml for spectinomycin and kanamycin) and grown at 37°C for 16 hours. Single colonies were picked from plates to prepare starter cultures in LB medium supplemented with antibiotic. After 16 hours, expression cultures were inoculated 1/100 using LB starter cultures and, to guarantee proper aeration, cultur volume was no more than 25% of the total conical flask volume. P450 expression in minimal medium was carried out according to Kelly *et al.*^[1]: 1X M9 salts solution was supplemented with 0.4% glucose, 0.05% FeCl₃, 1 mM MgSO₄, 1 mM CaCl₂ and cells were grown with vigorous shaking (200 rpm) at 37 °C until an OD₆₀₀ of 0.8 was reached. At this stage, isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.4 mM) was added to induce protein expression, along with 5-

aminolevulinic acid hydrochloride (5-ALA, 0.5 mM) supplementation. Protein expression was carried out at 20 °C with shaking at 200 rpm for 20 h and used immediately. The use of complex media (e.g., LB, TB and auto-induction media) resulted in the expression of the protein in the insoluble fraction or soluble inactive protein (data not shown).

Similarly, production of Cm-ADH10 and SSCR was carried out following a protocol similar to that for P450cam-RhFRed expression, the exception being the medium used (TB instead of M9) and the need for 5-ALA supplementation (not added). SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was employed to confirm protein expression. Recombinant Cm-ADH10 was obtained by inoculating 1 L of Terrific broth containing 50 µg/mL kanamycin with transformed cell *E. coli* BL21 (DE3) and incubated until an OD₆₀₀ of 0.6; protein production was induced with 0.4 mM IPTG at 20 °C and production was kept for 16 h. Cells were harvested by centrifugation (2500 *g*, 20 min, 4 °C) and kept at -80 °C until purification. Cells from 1 L of culture were suspended in 50 mL of 20 mM Tris·HCl, pH 7.5. The crude extract was prepared by sonication for 5 min, 70% amplitude, 5 s on/off. Cell debris was removed by ultracentrifugation at 31,000 *g* for 45 min, 4 °C. The supernatant was filtered using a 0.45 µm filter. The cell-free extract was applied to a 5 mL HisTrap FF column using an ÄKTA Pure (GE Healthcare) at 1 ml/min at 4 °C and a 20 mM imidazole solution in 20 mM Tris·HCl buffer, pH 7.5. The column was washed with 4 – 6 CV of a 50 mM imidazole solution in 20 mM Tris·HCl buffer, pH 7.5. Pure enzyme was eluted using 3 CV of a 500 mM imidazole solution in 20 mM Tris·HCl buffer, pH 7.5. Enzyme was concentrated and excess of imidazole was removed with an Amicon® Ultra concentrator (10,000 NMWL; Millipore) at 4000 *g*. Typically, the enzyme was diluted to 10 mg/mL in 50 mM KPi buffer (pH 7.5) to carry out biotransformation trials and kinetic measurements whereas, for crystallographic studies, concentrated aliquots (1 – 1.5 mL) were treated with thrombin (0.5 U, 4 °C, overnight shaking) and the cleaved 6x-His Tag was removed by gel filtration in a Superdex 200 10/30 GL column. Pure protein, as judged by UV chromatogram and SDS-PAGE, was concentrated to 77 mg/mL and stored at -80 °C.

7.3.1.4 Protein Crystallization, Structure Determination and Docking experiments

7.3.1.4.1 Crystallization and X-ray Data Collection

Native Cm-ADH10 was crystallized at 293 K using the sitting-drop vapor diffusion technique at 20 °C. Equal volumes of 12 mg/mL Cm-ADH10 in 20 mM Tris-HCl at pH 7.5 and reservoir solution were mixed. The reservoir solution contained 40% polyethylene glycol (PEG) 200 in 0.1 M MES at pH 6.5. Initial conditions were screened using the JSCG, PEG and ammonium sulfate screens (Qiagen) in 96-well sitting drop trays. X-Ray diffraction data were collected at the ID23-2 beamline of the European Synchrotron Radiation Facility in Grenoble, France (ESRF).

7.3.1.4.2 Structure Determination

The images were integrated and scaled using MOSFLM.^[2] Intensities were merged and converted to amplitudes with Aimless^[3] and other software of the CCP4 Suite (Table S1).^[4] The structure was solved with MOLREP^[5] and the coordinates of the dehydrogenase/reductase from *Sinorhizobium meliloti* 1021 (PDB: 3V2G) as the search model. Phases were improved by density modification and the model was rebuilt with ARP/WARP,^[6] COOT^[7] and REFMAC5^[8] were employed to carry out alternating cycles of further model building and refinement. Figures were created with Chimera.^[9]

7.3.1.4.3 Homology Modelling

Homology modelling was performed using SWISS-MODEL.^[10] as template, the map of P450cam bound to camphor (PDB 1DZ9)^[11] with molecular oxygen in the active site was employed indicating the mutations Y96F and Y96F-F87W-L244A-V247L to obtain models for P450cam-Y96F and P450-WAL, respectively. The obtained models excluded all cofactors, solvent and metals from the original model; therefore, these were superimposed on the models from 1DZ9 using COOT,^[12,13] and the resulting models were used for docking experiments.

7.3.1.4.4 Docking Experiments

SwissDock^[14,15] was used to run ligand-receptor docking calculation of α -IP on the homology models for P450camY96F and P450-WAL. AutoDock Vina^[16] was used to run ligand-receptor

docking calculation of (*R*)-HIP on Cm-ADH10. Structures were prepared with DockPrep^[17] in Chimera. The best models yielded a score of -5.2 kcal/mol, a rmsd (l.b.) of 1.224 and a rmsd (u.b.) of 3.056 in the first experiment, and -5.6 kcal/mol, a rmsd (l.b.) of 0.0 and a rmsd (u.b.) of 0.0 for the second experiment.

7.3.1.5 Clustering of Indigo-Positive Variants

Plasmids (pET-14b) encoding P450cam-RhFRed mutants generated in a previous work^[1] were purified from bacterial glycerol stocks. Variants were produced targeting 7 pairs of residues by saturation mutagenesis using NDT codon degeneracy. This resulted in the production of seven libraries: A (F87-F96), B (F98-T101), C (M184-T185), D (L244-V247), E (G248-T252), F (V295-D297) and G (I395-V396). In order to cluster these variants, 100 ng of each plasmid purified were added to one of five pools according to the mutations introduced, the exceptions being variants from library G248-T252, V295-D297 and I395-V396 that were pooled together because the reduced size of these libraries (giving five pooled libraries in total). These plasmids preparations were then transformed into *E. coli* NEB-5-alpha competent cells and plasmids purified again for transformation into *E. coli* BL21 (DE3) for small-scale expression (125 mL). During subsequent screenings, transformation efficiencies of each pool member were not determined. Whole cell biotransformations were carried out as described in the section below. Estimated P450 content ranged from 5 to 9 μ M.

7.3.1.6 Whole Cell Biotransformations with Library D P450cam-RhFRed Variants

Expression cultures were centrifuged (2500 *g*, 20 min, 4 °C) and the pellet resuspended to 230 mg/mL (wet weight) in the appropriate biotransformation buffer (50 mM sodium phosphate buffer, pH= 7.2, 100 mM KCl). Determination of P450 concentrations in whole cells was performed on a plate reader (Tecan Infinite 200 series, Männedorf, CH) according to Kelly *et al.*^[1] and for cell lysates on a Cary 50 UV/Visible spectrophotometer (Agilent) according to the protocol by Omura and Sato.^[18] Biotransformations were carried out in 48-deep well plates on a 1 mL scale, with 880 μ L of re-suspended cells (final P450 concentration ranging from 1 to 4 μ M), 100 μ L of 100 mg/mL glucose and the appropriate final concentration of substrate typically added from concentrated stocks in DMSO (2% v/v final concentration). Plates were sealed with a gas permeable membrane and reactions carried out with 250 rpm orbital shaking for 24 h. Reaction mixtures were extracted with 1 mL of methyl-tert-butyl ether

(MTBE) with 1 mM decane (as internal standard for quantitative analysis) vortexed for 30 s and centrifuged 5 minutes before removing the organic layer, which was then transferred to fresh tubes containing MgSO₄. Finally, 250 μ L of the MTBE extract was taken for GC analysis and diluted when necessary. In order to determine the *ee*%, the samples were also analyzed by chiral normal phase HPLC.

7.3.1.7 Screening of ADHs

A panel of alcohol dehydrogenases enzymes as freeze-dried cell free extracts in 96-deep well plates was kindly supplied by c-LEcta GmbH (Leipzig, Germany). For a first screening, each freeze-dried extract was resuspended in 50 μ L of 50 mM sodium phosphate buffer pH = 7.2, 100 mM KCl, and reaction carried out on a 500 μ L scale with both 10 mM (*R*)- and (*S*)-HIP, 0.5 mM of both NAD⁺ and NADP⁺, and acetone 5% v/v (30 °C). Then, best hits were chosen for a further screen on a 500 μ L scale with 10 mM (*R*)-HIP, 0.5 mM NADP⁺ and acetone 5% v/v. Reaction mixtures were extracted with methyl-tert-butyl ether (MTBE), vortexed for 30 s and centrifuged 5 minutes before removing the organic layer, which was then transferred to fresh tubes containing MgSO₄. Finally, 250 μ L of the MTBE extract was taken for chiral normal phase HPLC analysis.

7.3.1.8 Scale Up of the Biotransformation of α -Isophorone with P450cam-RhFRed L244I-V247S for the Identification of **7**

In order to characterize α -IP oxidation products, P450cam-RhFRed L244I-V247S was chosen for its good activity profile and reduced selectivity. To a solution of α -IP (2 mM), glucose (10 mg/mL) and KCl (100 mM) in NaP_i buffer (50 mM, pH 7.2, final volume 200 mL), wet *E. coli* BL21(DE3) cells producing P450cam-RhFRed L244I-V247S were added (200 mg/mL). The mixture was incubated at 20°C, 250 rpm for 24 h, then centrifuged to remove the cells. The reaction mixture was acidified to pH \sim 3.0 with 3 M HCl and centrifuged (2500 *g*, 20 min, 4°C). The supernatant was extracted with EtOAc (2 x 200 mL), the organic layer was dried over MgSO₄ and evaporated, to afford an orange residue. Silica gel chromatography purification (cyclohexane with increasing amounts of EtOAc) afforded **6** and **7** of sufficient purity. Compound **7** was identified by comparison of its ¹H NMR spectrum with literature data.^[19] ¹H NMR (400 MHz, CDCl₃): δ = 6.15 (s, 1H), 4.23 (s, 2H), 2.27 (s, 2H), 2.14 (s, 2H), 1.05 (s, 6H).

7.3.1.9 Preparation of (R)-6 and (S)-6

The two enantiomers of **6** were prepared according to Hennig *et al.*^[20]

Benzeneruthenium(II) chloride dimer (10.5 mg) and (1*S*,2*R*)-2-amino-1,2-diphenylethanol (17.7 mg) were dissolved in *i*-PrOH (5 mL) and the red solution stirred for 30 min at 80°C. Afterwards, additional *i*-PrOH (36.5 mL) was added and the solution cooled to 28°C. Then, ketoisophorone (631 mg) and NaOH solution (2 mL, 0.1 M in *i*-PrOH) were added to give a darker solution and the reaction followed by TLC until completion (~ 3 h). The reaction mixture was then filtered through celite and the filtrate evaporated to afford a black oily residue. The product was purified by silica gel chromatography, eluting with cyclohexane containing increasing amounts of EtOAc. Evaporation of the solvent under reduced pressure afforded (S)-HIP as a greenish oil (253.4 mg, 40% isolated yield). $[\alpha]_{\text{D}}^{22} = -82$ (c = 1, MeOH) vs. *lit.* ^[20] $[\alpha]_{\text{D}}^{22} = -105.8$ (c = 1, MeOH).

For the preparation of (R)-**6**, the same procedure was followed, but (1*R*,2*S*)-2-amino-1,2-diphenylethanol was used (285.1 mg, 45% isolated yield). $[\alpha]_{\text{D}}^{22} = +101$ (c = 1, MeOH) vs. *lit.* ^[20] $[\alpha]_{\text{D}}^{22} = +105.9$ (c = 1, MeOH).

NMR data for both samples are consistent with previously reported values. ¹H NMR (400 MHz, CDCl₃): $\delta = 5.87$ (s, 1H), 4.04 (s, 1H), 2.31 (m_{AB}, 2H), 2.05 (s, 3H), 1.08 (s, 3H), 1.02 (s, 3H); ¹³C NMR (100 MHz, CDCl₃): $\delta = 198.7$ 160.6, 126.6, 77.4, 49.1, 38.6, 27.1, 21.6, 21.3.

7.3.1.10 Analytical Scale Double Oxidation of α -Isophorone to Ketoisophorone

Method A (one-pot, two-step). Step 1. To a solution of α -IP (10-20 mM) and glucose (10 mg/mL) in the appropriate buffer (50-500 mM, pH 8.0, final volume 1 mL, 48-deep well plates), wet *E. coli* BL21(DE3) harboring a pCDF-1b vector carrying the gene encoding P450-WAL were added (200 mg/mL). The mixture was incubated at 20-40°C, 250 rpm for 18-24 h, then centrifuged to remove the cells. Step 2. Either the supernatant of step 1 or a solution of **6** (10-100 mM) in the appropriate buffer (100-500 mM, pH 4.0-10.0) was used for the ADH oxidation step. Cm-ADH10 was added either as a 10x concentrated cell lysate of *E. coli* BL21(DE3) (~1 mg/mL final protein conc.) or as purified protein (1 mg/mL final conc.) along with the appropriate cosubstrate (0.5 % v/v) and NADP⁺ (0.25 mM final conc.). The mixture

was incubated at 20-50°C, 180 rpm for 6-24 h. Both steps were followed by taking small samples for GC analysis after extraction. When measurements were accomplished in triplicates, mean values are given. The error bar represents the standard deviation.

Method B (whole cell cascade). *E. coli* BL21 (DE3) competent cells were co-transformed with plasmids bearing the sequences encoding P450-WAL (pCDF-1b) and Cm-ADH10 (pET28a) and plated on LB agar containing 25 µg/mL of kanamycin and spectinomycin. After 16 h incubation at 37°C, single colonies were picked and grown at 37°C in M9 medium supplemented with antibiotics. Protein production was carried out similarly to P450cam-RhFRed variants, including 25 µg/mL of kanamycin and spectinomycin. The whole-cell double oxidation of α -isophorone was carried out in KP_i buffer (200 mM, pH 8.0, final volume 1 mL, 48-deep well plates) with α -IP (10 mM), glucose (10 mg/mL) and whole cells (200 mg/mL cell wet weight). The mixture was incubated at 28°C with shaking (200 rpm) for 24 h. Samples were analyzed by GC after extraction. When measurements were accomplished in triplicates, mean values are given. The error bar represents the standard deviation.

7.3.1.11 Preparative-Scale Double Oxidation of α -Isophorone to Ketoisophorone

To a solution of α -IP (10 mM) and glucose (10 mg/mL) in KP_i buffer (200 mM, pH 8.0, final volume 65 mL), wet *E. coli* BL21(DE3) cells producing P450-WAL were added (200 mg/mL). The mixture was incubated at 28°C, 150 rpm for 18 h, then centrifuged to remove the cells. A 10x concentrated cell lysate of *E. coli* BL21(DE3) producing Cm-ADH10 (~1 mg/mL final protein conc.) and chloroacetone (1.6 equiv.) were added to the supernatant. The mixture was incubated at 40°C, 150 rpm for 6 h. Both steps were followed by taking small samples for GC analysis. After completion of the reaction, the mixture was extracted with EtOAc (3 x 30 mL). The organic phase was dried over MgSO₄ and evaporated under reduced pressure to afford a yellowish crude product that was submitted to column chromatography purification (petroleum ether 40-60°C with increasing amounts of EtOAc). KIP was obtained as a colourless oil (46 mg, 56% isolated yield, >97% purity by ¹H NMR). ¹H NMR (400 MHz, CDCl₃): δ = 6.55 (s, 1H), 2.71 (s, 2H), 2.00 (s, 3H), 1.24 (s, 6H); ¹³C NMR (100 MHz, CDCl₃): δ = 203.6, 197.9, 149.1, 137.2, 52.0, 45.3, 26.2, 17.0; HRMS (ESI+): m/z for C₉H₁₃O₂⁺ [M+H]⁺ calcd. 153.0910, found 153.0897.

7.3.1.12 Analytical Methods

NMR spectra were recorded on a Bruker Avance 400 spectrometer (400.1 MHz for ^1H and 100.6 MHz for ^{13}C) in deuterated chloroform at 298 K using tetramethylsilane (TMS) as an internal standard. Chemical shifts are reported as δ in parts per million (ppm) and are calibrated against residual solvent signal.

Chiral normal phase HPLC analysis was carried out on an Agilent system equipped with a G4225A degasser, G1311A quaternary pump, a G1329A well plate autosampler unit, a G1315B diode array detector and a G1316A thermostated column compartment. Separation of the enantiomers of HIP was carried out using a CHIRALPAK[®] AS-H column (5 μm particle size, 4.6 mm diameter x 250 mm; Daicel Chemical Industries Ltd.) operating in isocratic mode with 80% hexane and 20% isopropanol for 18 min at 25°C. Injection volume was 5 μL and chromatograms were monitored at 254 nm. Retention times were as follows: KIP 9.5 min, (S)-HIP 10.9 min, α -IP 12.4, (R)-HIP 13.3 min.

GC analysis was performed on an Agilent 6850 system with a flame ionization detector (FID) equipped with an Agilent HP-1 column (30 m x 0.32 mm x 0.25 μm). Samples were suitably diluted and 2 μL were injected at a split ratio 10:1. The inlet temperature was set at 200°C, the detector temperature at 250°C and the pressure maintained at 6.8 psi. The following method was applied: 50°C; 10°C/min to 220°C, hold 2 min. The corresponding retention times were: α -IP 7.4 min, KIP 7.6, possible KIP-reduction by-product 7.8 min and HIP 9.7 min. Calibration curves of decane vs **6** or **3** were constructed in order to calculate TTNs (total turnover numbers) or conversion values.

7.3.2 Supplementary Figures and Tables

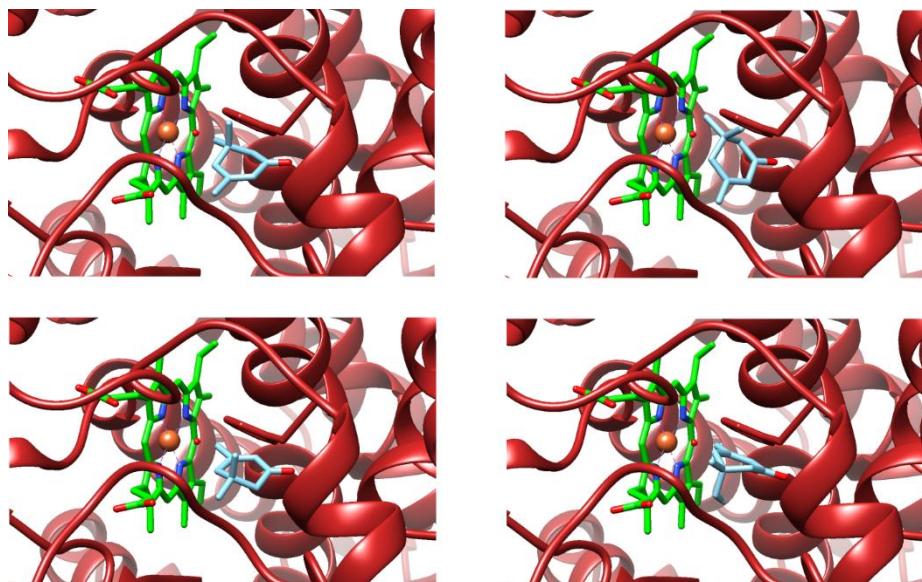


Figure S1. Docking of α -IP in the modelled active site of P450cam Y96F: F96 side chain is represented as stick, with α -IP in cyan and the heme group in green. The docking shows the most energetically favoured conformations for α -IP. However, C4 of α -IP is never within catalytic distance and angle from the heme oxygen.

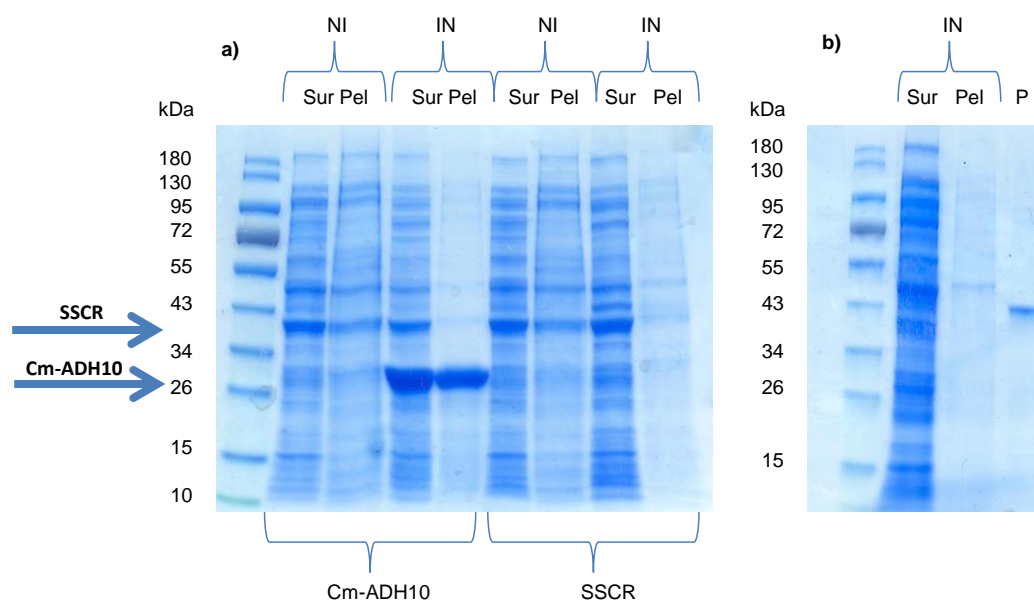


Figure S2. SDS-PAGE showing lysate (Lys), soluble (Sur) and insoluble (Pel) fractions resulting from (a) Cm-ADH10 (expected size 27 kDa) and SSCR (expected size 39 kDa) expression trials and (b) purification of SSCR (P lane, purified protein). NI = total cell extract from not-induced cells, IN = total cell extract from IPTG-induced cells. The arrows indicate the expected band corresponding to Cm-ADH10 or SSCR. The ProtParam tool (<http://web.expasy.org/protparam/>) was employed to compute protein molecular weights.

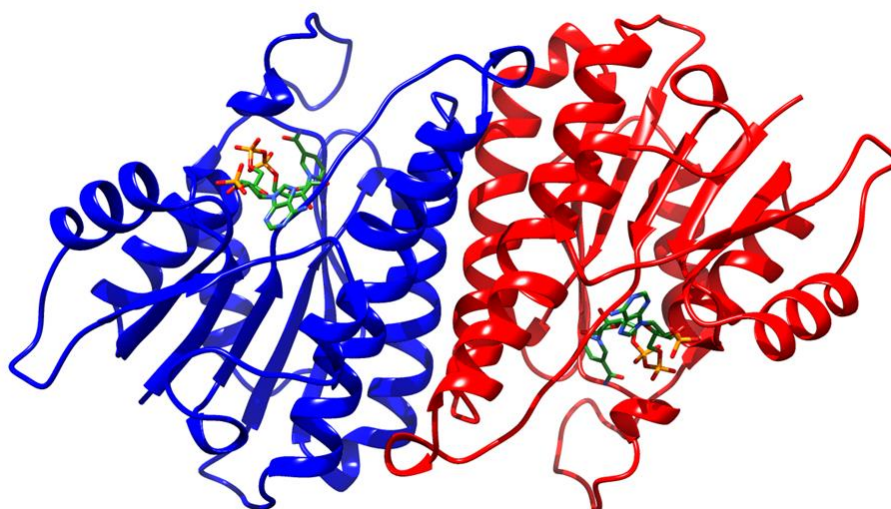


Figure S3. Cartoon representation of the Cm-ADH10 dimer in the asymmetric unit with bound NADP⁺ (green). Two monomers are shown in red and blue.

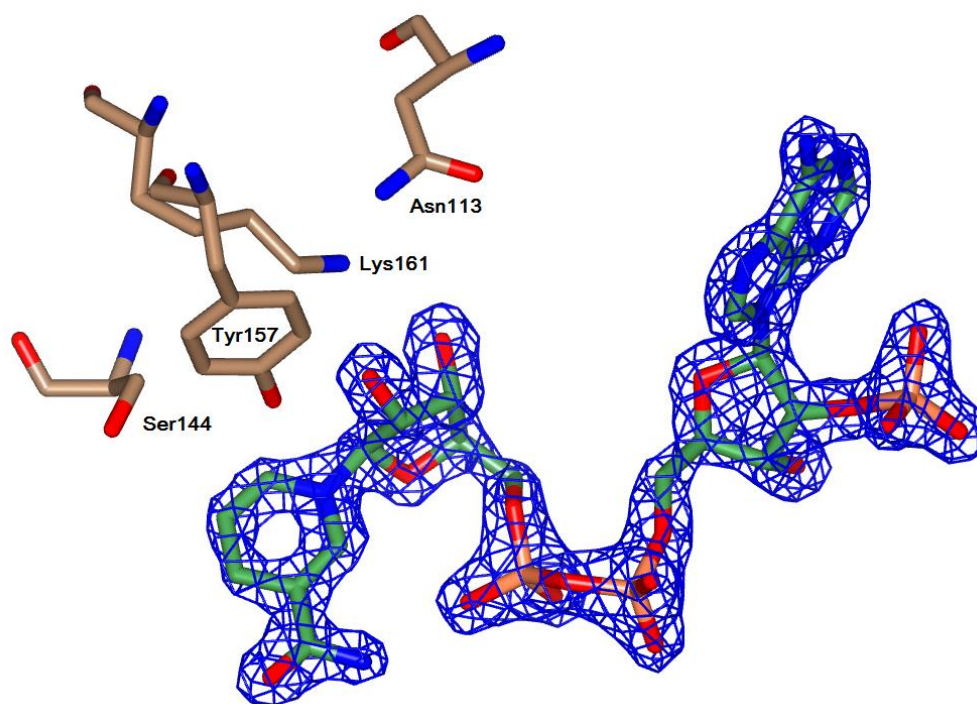


Figure S4. Representation of the NADP⁺ (green) electron density at 1 σ in the Cm-ADH10 monomer, surrounded by the active site tetrad (grey residues).

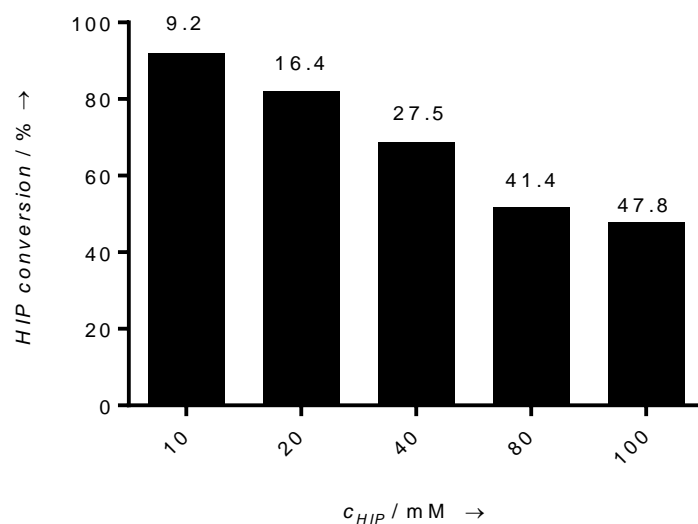


Figure S5. Conversion values obtained with different HIP concentration. Reaction set-up: 1 mg/mL purified Cm-ADH10, 0.25 mM NADP⁺ and 0.5% v/v ethyl acetoacetate as cosubstrate for cofactor regeneration (30°C, 24 h). Product concentrations (mM) are given above each bar.

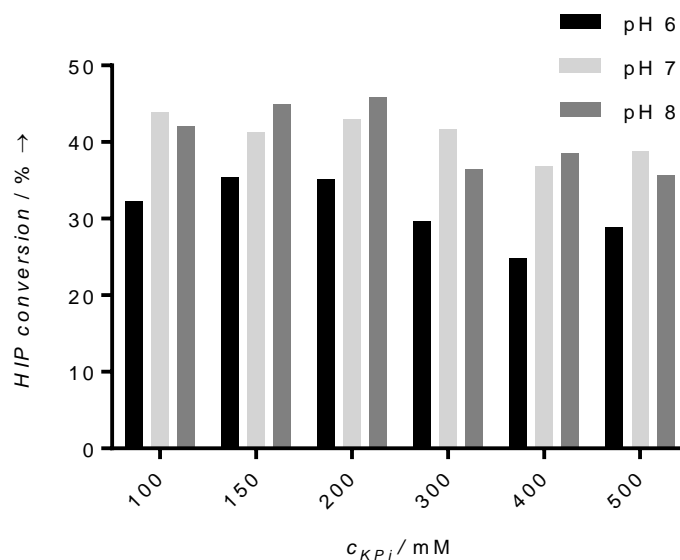


Figure S6. Effect of buffer concentration and pH on Cm-ADH10 activity. Reaction set-up: 1 mg/mL purified Cm-ADH10, 0.25 mM NADP⁺, 40 mM HIP and 0.5% v/v ethyl acetoacetate as cosubstrate for cofactor regeneration (30°C, 4.5 h).

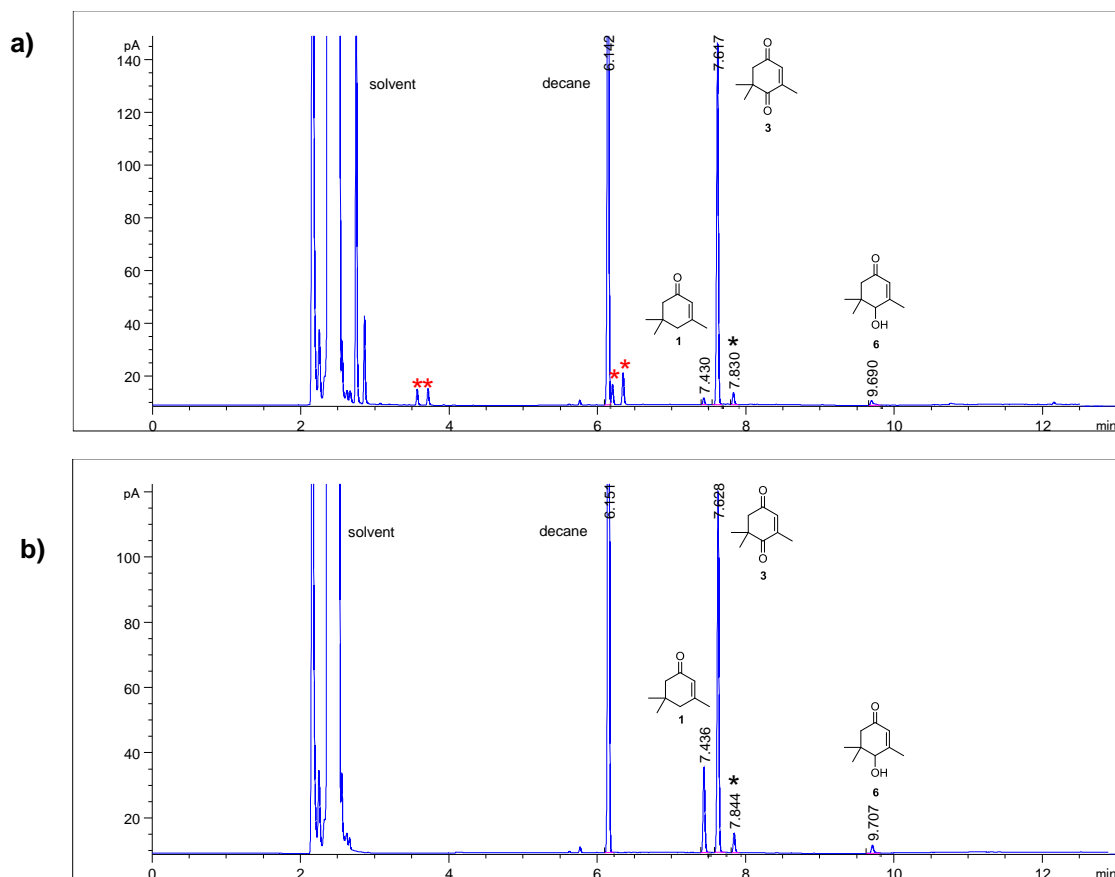


Figure S7. (a) Scale up experiment for product isolation and characterization: GC trace showing successful conversion of **1** to **3** over 24 h. A possible by-product resulting from KIP C=C bond reduction during the first step is indicated by the black asterisk. Red asterisks indicate other unidentified impurities; (b) GC trace for the whole-cell catalyzed double oxidation of **1**. Lower conversion may be attributed to residual starting material not oxidized by P450-WAL. A possible by-product resulting from KIP C=C bond reduction during the first step is indicated by the black asterisk.

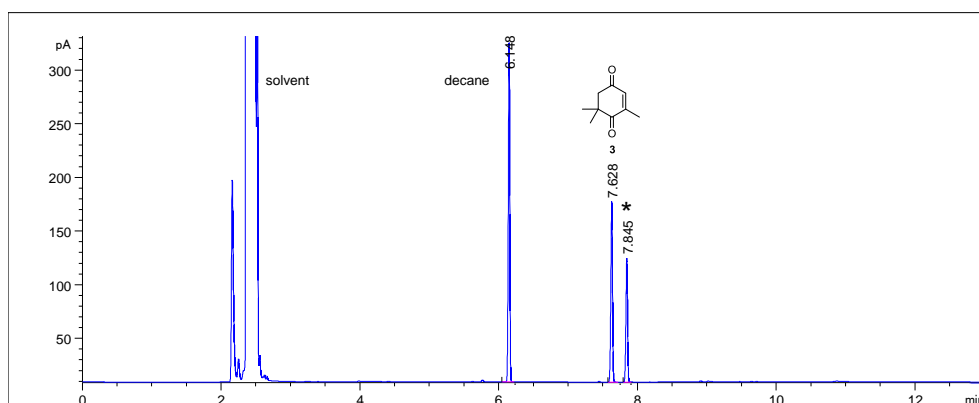
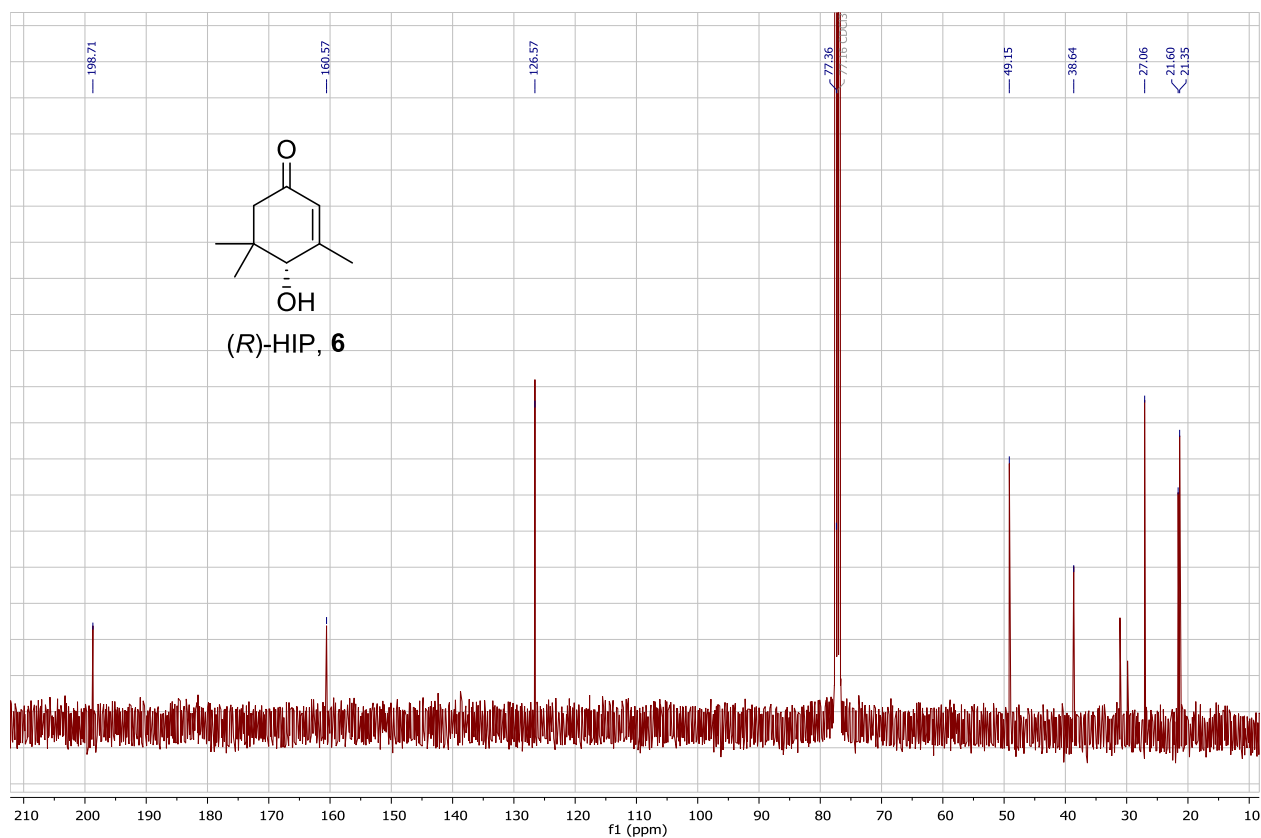
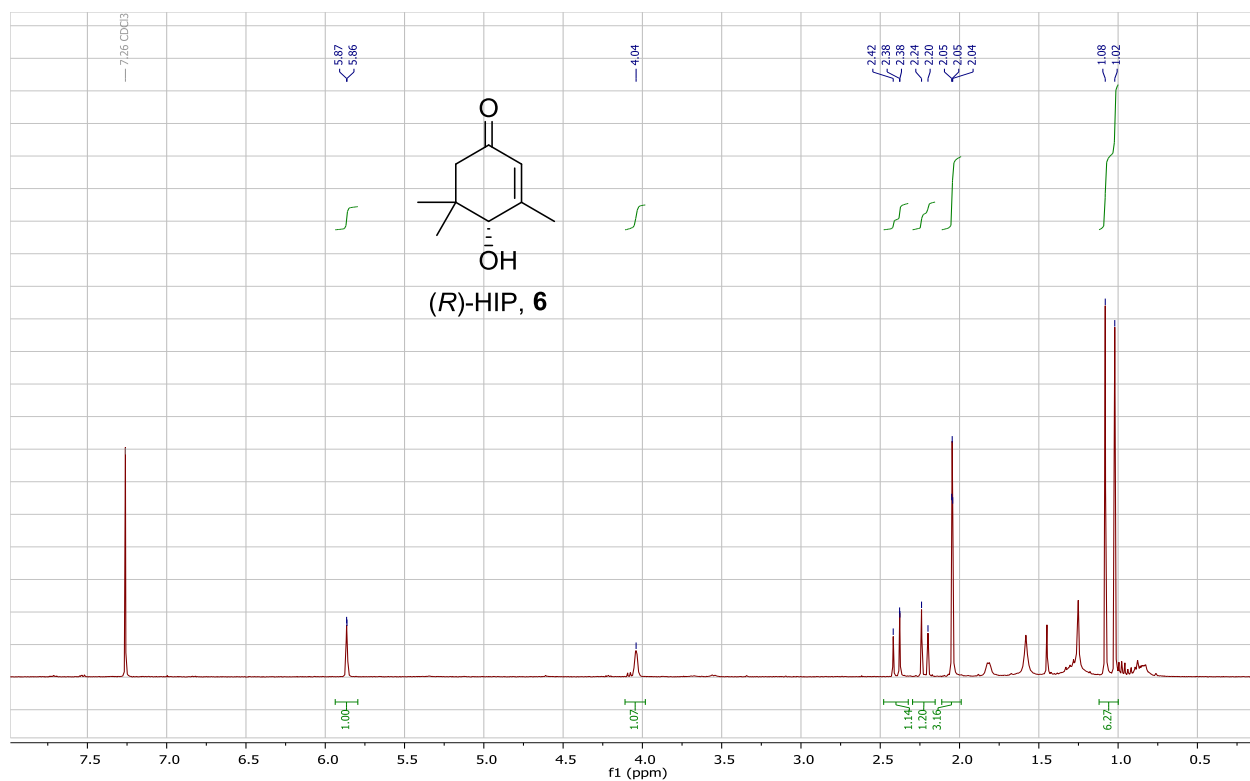


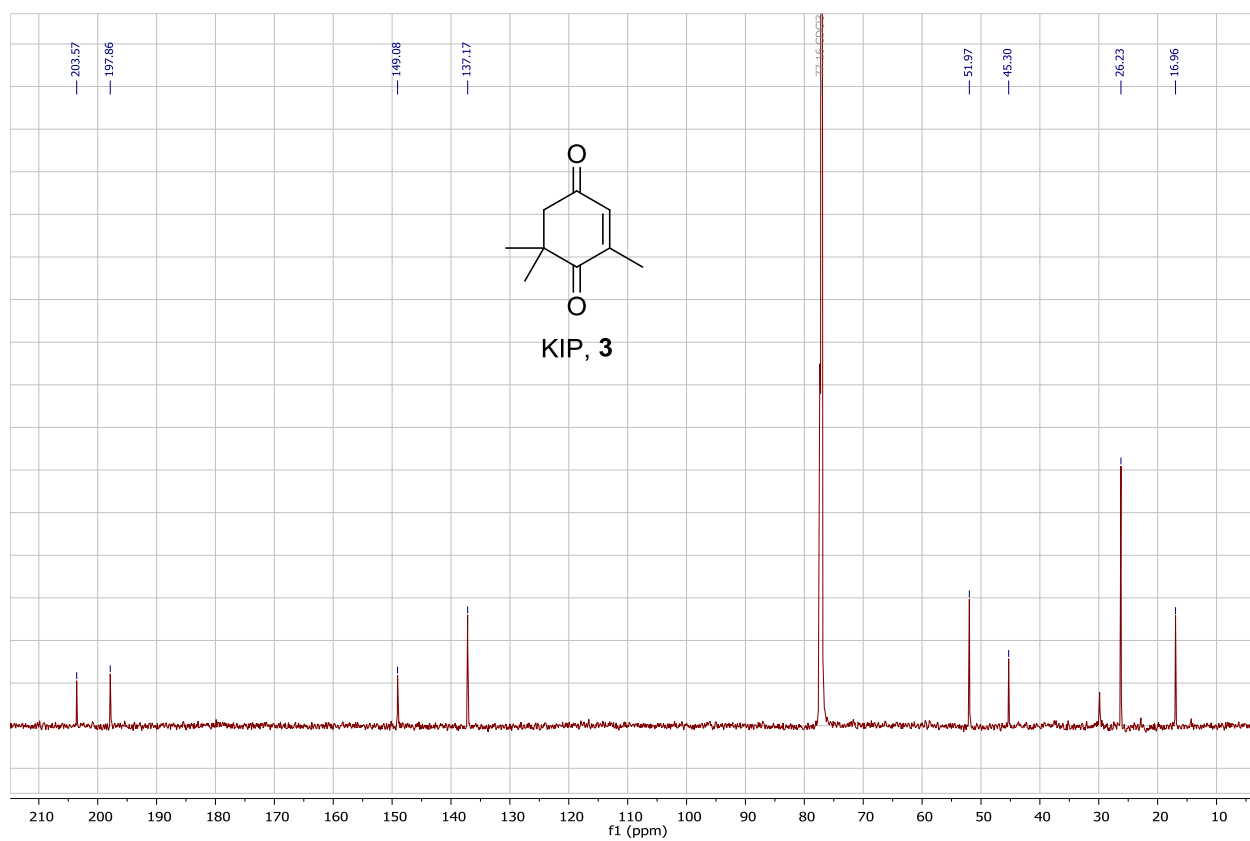
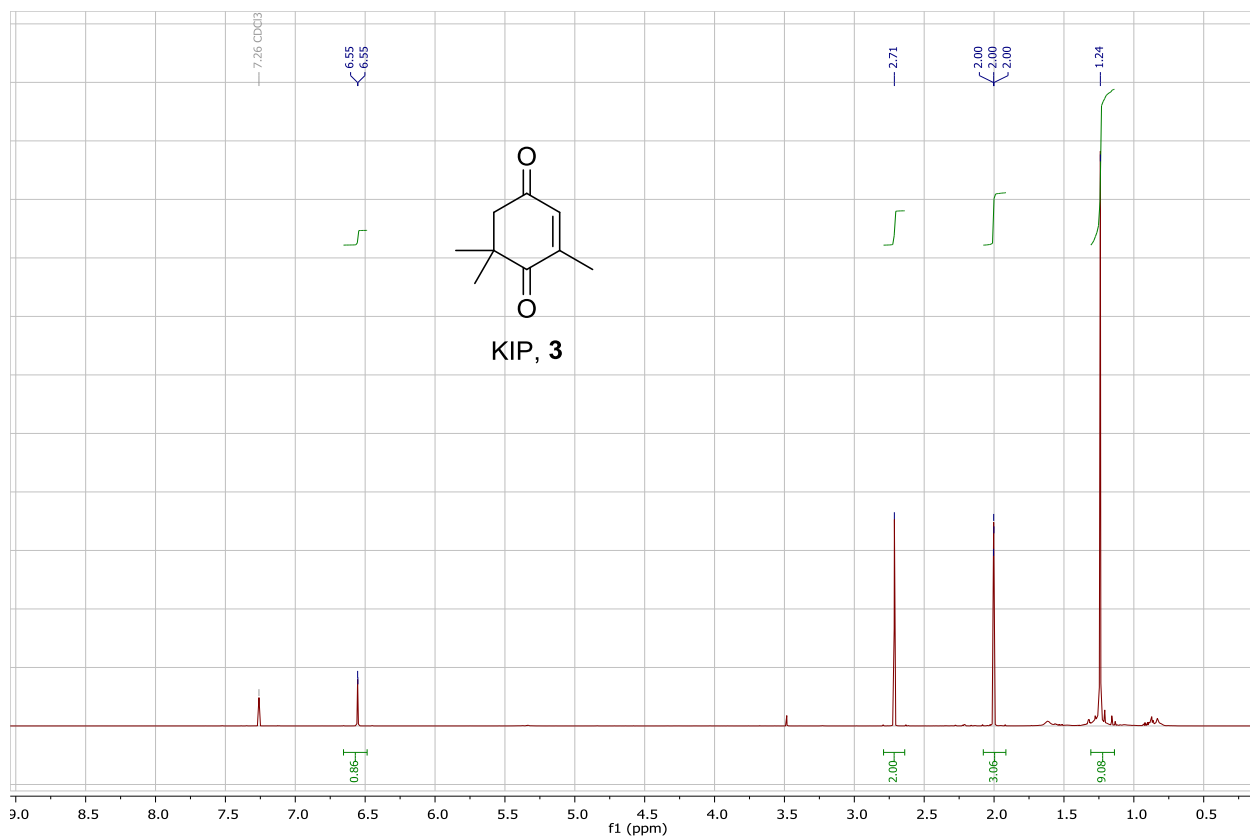
Figure S8. GC analysis of the biotransformation extract resulting from incubation of KIP with *E. coli* BL21(DE3) transformed with an empty pET-28a vector. The black asterisk indicates the by-product resulting from endogenous enzymatic activities.

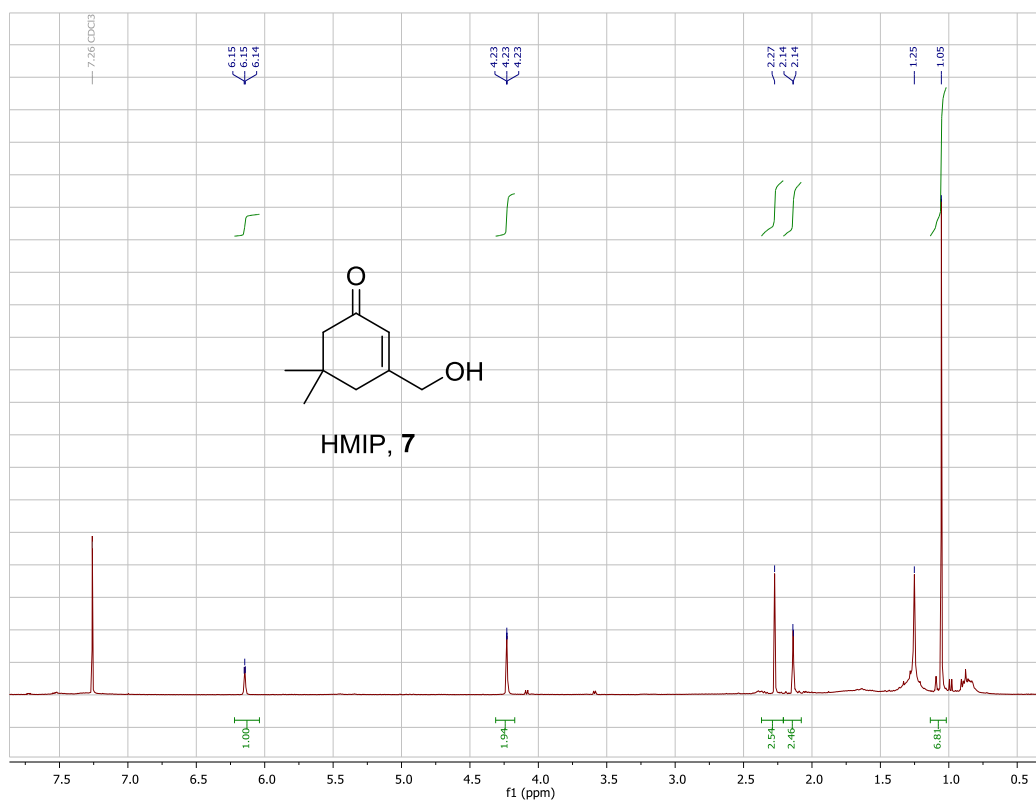
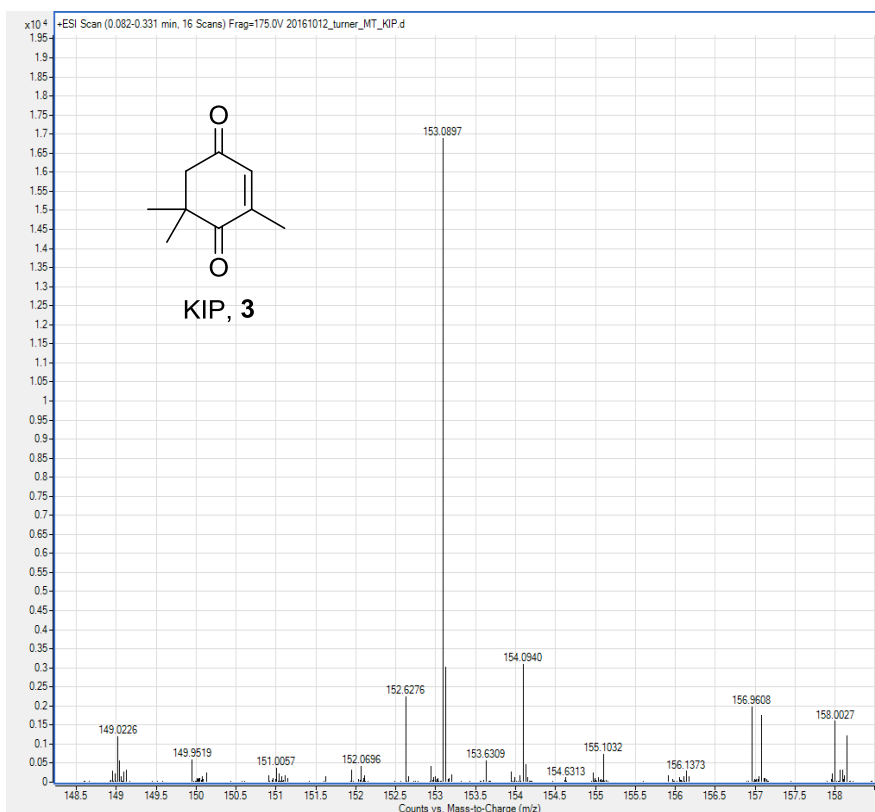
Table S1. Data collection and refinement statistics for the crystal structure of Cm-ADH10 binary complex with NADP⁺ (1.6 Å resolution). Statistics for the highest-resolution shell are shown in parentheses. Data were obtained with PHENIX.^[21]

PDB Code	5MLN
Wavelength (Å)	0.8729
Resolution range (Å)	40.3 - 1.6 (1.65 - 1.60)
Space group	P 2 ₁ 22 ₁
Unit cell (Å)	74.178 80.425 80.52 90 90 90
Unique reflections	65173 (6407)
Multiplicity	8.1 (8.2)
Completeness (%)	100 (99)
Mean I/sigma(I)	21.12 (2.85)
R-merge	0.058 (0.779)
CC_{1/2}	1 (0.767)
Reflections used in refinement	65169 (6407)
Reflections used for R-free	3247 (305)
R-work	0.147 (0.2248)
R-free	0.18 (0.2658)
Number of non-hydrogen atoms	4033
RMS(bonds) (Å)	0.025
RMS(angles) (°)	2.3
Ramachandran favored (%)	97
Ramachandran allowed (%)	2.5
Ramachandran outliers (%)	0.42
Rotamer outliers (%)	2.9
Average B-factor	23.0
macromolecules	22.0
ligands	25.8
solvent	40.3

7.3.2.1 NMR and HRMS Spectra



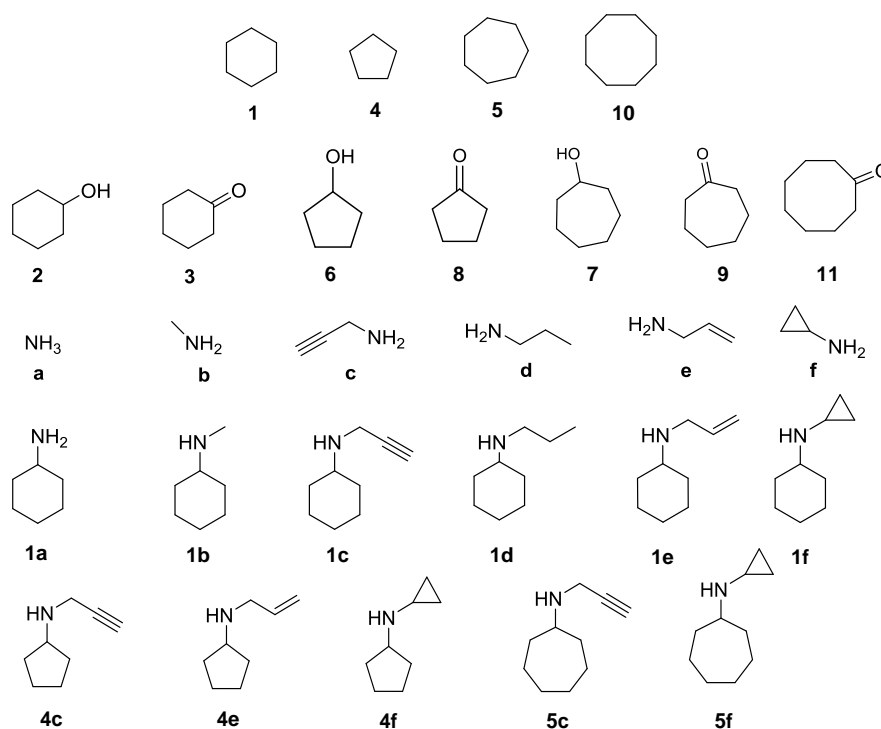




7.3.3 References

- [1] P. P. Kelly, A. Eichler, S. Herter, D. C. Kranz, N. J. Turner, S. L. Flitsch, *Beilstein J. Org. Chem.* 2015, 11, 1713–1720.
- [2] T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, A. G. W. Leslie, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 2011, 67, 271–281.
- [3] P. R. Evans, G. N. Murshudov, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 2013, 69, 1204–1214.
- [4] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, et al., *Acta Crystallogr. Sect. D Biol. Crystallogr.* 2011, 67, 235–242.
- [5] A. Vagin, A. Teplyakov, *J. Appl. Crystallogr.* 1997, 30, 1022–1025.
- [6] G. Langer, S. X. Cohen, V. S. Lamzin, A. Perrakis, *Nat. Protoc.* 2008, 3, 1171–9.
- [7] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 2010, 66, 486–501.
- [8] G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 2011, 67, 355–367.
- [9] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* 2004, 25, 1605–1612.
- [10] P. Benkert, M. Biasini, T. Schwede, *Bioinformatics* 2011, 27, 343–350.
- [11] I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko, S. G. Sligar, *Scienc* 2000, 287, 1615–1622.
- [12] K. Arnold, L. Bordoli, T. Schwede, *Bioinformatics* 2006, 22, 195–201.
- [13] M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. G. Cassarino, M. Bertoni, L. Bordoli, et al., *Nucleic Acids Res.* 2014, 42, 252–258.
- [14] A. Grosdidier, V. Zoete, O. Michielin, *J. Comput. Chem.* 2011, 32, 2149–2159.
- [15] A. Grosdidier, V. Zoete, O. Michielin, *Nucleic Acids Res.* 2011, 39, 270–277.
- [16] O. Trott, A. J. Olson, *J. Comput. Chem.* 2010, 31, 2967–2970.
- [17] P. Lang, S. Brozell, S. Mukherjee, E. Pettersen, *RNA* 2009, 1219–1230.
- [18] R. Omura, Tsuneo; Sato, *J. Biol. Chem.* 1964, 239, 2370–2378.
- [19] Y. Mikami, Y. Fukunaga, M. Arita, Y. Obi, T. Kisaki, *Agric. Biol. Chem.* 1981, 45, 791 – 793.
- [20] M. Hennig, K. Pu, M. Scalone, *Tetrahedron: Asymmetry*, 2000, 11, 1849–1858.
- [21] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, et al., *Acta Crystallogr. Sect. D Biol. Crystallogr.* 2010, 66, 213–221.

7.4 A Biocatalytic Cascade for the Amination of Unfunctionalized Cycloalkanes



List of substrates and products mentioned in the text.

7.4.1 Experimental

7.4.1.1 Materials

Solvents, commercially available chemicals and carbon monoxide for CO difference spectroscopy were obtained from Sigma-Aldrich (Poole, Dorset, UK). Gases for GC-FID analysis were purchased from BOC gases (Guildford, UK). Chemically competent cells and enzymes for molecular biology were purchased from New England Biolabs (Hitchin, UK). Terrific Broth Base autoinduction medium including trace elements (TB-AIM) was purchased from Formedium (Hunstanton, UK).

7.4.1.2 Molecular Biology Methods

Custom primer synthesis and plasmid DNA sequencing were performed by Eurofins Genomics. Primers employed in this work are given below (mismatching bases are given in red).

Primer	5'->3' sequence
R47L Y51F for, Tm=66°C	AGGCGCCTGGTCTGGTAACGCGCTTCTTATCAAGTCAGCGTC
R47L Y51F rev, Tm=63°C	CGAATTTAAAGATTTCTCCTAATTCATCCGCAATTTTCATCAAAGC
R966D for, Tm=60°C	CCGCTTTTTCTGACATGCCAAATCAGCCGAAAAC
R966D rev, Tm=64°C	TATGAAGCGTAATGATGCCTTCGCTTTGGG
W1046S for, Tm= 59°C	CAAAAGACGTGTGGCTGGGTAACCTCG
W1046S rev, Tm= 61°C	CGTATCGGCCTTTTCTTCTAGCTGC

Target mutations were introduced by inverse PCR using Eppendorf Mastercycler Gradient thermal cyclers, with buffers and enzymes supplied in the Phusion DNA Polymerase kit (NEB). Reactions (50 µL) were carried out in thin-walled 200 µL PCR tubes following the manufacturer's instructions. Next, template DNA was removed by a 2 h digestion with *DpnI* followed by PCR purification (QIAquick PCR Purification Kit). Ligation reactions were performed for 1 h at 25°C with T4 DNA ligase and polynucleotide kinase. NEB 5-alpha competent *E. coli* (high efficiency) were then transformed according to manufacturer's instructions, single colonies were picked and grown in 5 mL Luria-Bertani medium (LB) containing 50 µg/mL kanamycin for 16 h. Plasmid DNA was isolated using a mini-prep kit (Qiagen) and the sequence verified by plasmid sequencing.

7.4.1.3 Protein Production and Purification

Chemically competent *E. coli* BL21 (DE3) were transformed by heat shock with a pET28a vector (or pET28b for *TeSADH* W110A) encoding the desired *N*-terminal polyhistidine tagged enzyme under a T7 promoter. Transformants were grown on LB agar with 50 µg/ml kanamycin at 37°C for 16 hours. Starter cultures were prepared in LB medium with 50 µg/ml kanamycin at 37°C for 16 hours by picking single colonies from agar plates.

7.4.1.3.1 P450-BM3

Expression cultures (800 mL, TB-AIM) were inoculated with 8 mL starter culture and cells grown at 37°C 200 rpm until OD₆₀₀ = 0.8 was reached. At this stage, 5-Aminolevulinic acid hydrochloride (5-ALA, 0.5 mM) was added and the growth continued at 20°C. After 20 h, cells were harvested by centrifugation (2500 *g*, 20 min, 4 °C) and kept at -20 °C until further use.

For protein purification, cells were resuspended in 90% buffer A (Tris-HCl 0.1 M, 0.3 M NaCl, pH 8) and 10% buffer B (Tris-HCl 0.1 M, 0.3 M NaCl, 0.3 M imidazole, pH 8) to a final concentration of 200 mg/mL cell wet weight. Cells were lysed by ultrasonication with a Bandelin Sonopuls sonicator (20 cycles, 15 s on and 45 s off, 40 % amplitude) and the supernatant obtained by ultracentrifugation (48384 *g*, 30 min, 4 °C). Filtration of the supernatant was performed with a 0.45 µm filter before loading onto a pre-equilibrated 5 mL HisTrap FF column (GE Healthcare). The target protein was purified using an ÄKTA Pure system (GE Healthcare) operated at 2 mL min⁻¹ following a series of steps reported in the table below:

Step	% B	Column volumes
Wash 1	10	5
Wash 2	20	5
Elution	100	10
Re-equilibration	0	5

Protein elution was followed at 280 nm, fractions were collected and desalted in 0.2 M Tris-HCl, pH 8 with a 30,000 molecular weight cut-off filter (Vivaspin column, GE Healthcare). Protein concentration was either determined by CO-difference spectroscopy (P450) following the method of Omura and Sato^[1] or spectrophotometrically (*TeSADH* W110A: $\epsilon=25600 \text{ M}^{-1} \text{ cm}^{-1}$; *CboFDH* $\epsilon=51465 \text{ M}^{-1} \text{ cm}^{-1}$; *AspRedAm* $\epsilon= 24410 \text{ M}^{-1} \text{ cm}^{-1}$, wavelength $\lambda=280 \text{ nm}$). Typically, proteins were stored at -80°C as stocks at 10 mg mL⁻¹ protein concentration.

When crude enzyme preparations were used to carry out biotransformation, cells were resuspended in 0.2 M Tris-HCl, pH 8 to 200 mg mL⁻¹ cell wet weight and lysed as described above. The supernatant obtained after ultracentrifugation was either used without further purification or concentrated until the desired protein concentration/enzymatic activity was attained.

7.4.1.3.2 *TeSADH W110A* and *CboFDH*

Expression cultures (400 mL, TB-AIM) were inoculated with 4 mL starter culture and cells grown at 37°C 250 rpm for 24 h. Cell harvesting, storage and purification were performed as reported above.

For lyophilization of *TeSADH W110A* cell free extracts (CFE), cells were resuspended in 50 mM potassium phosphate buffer, 0.5 M NaCl, pH 7.5 to a final concentration of 300 mg/mL cell wet weight and lysed by ultrasonication as described above. The supernatant obtained by ultracentrifugation was frozen in liquid nitrogen before freeze-drying for two days (Heto Powerdry LL 1500, Thermo Electron Corporation).

7.4.1.3.3 *AspRedAm*

Expression cultures (600 mL, 2-YT broth) were inoculated with 6 mL starter culture and cells grown at 37°C 250 rpm until $OD_{600} = 0.8$ was reached. At this stage, isopropyl β -D-1-thiogalactopyranoside (IPTG, 0.4 mM) was added to induce protein expression and the growth continued at 20°C. Cell harvesting, storage and purification were performed as reported above.

7.4.1.4 Determination of Volumetric Enzymatic Activities

Volumetric activities were determined by monitoring NAD(P)H formation using a microtiter plate reader (Infinite M200 Pro, Tecan, Männedorf, Switzerland) under the following conditions: Tris-HCl 0.2 M, pH 8, 1 mM NAD^+ (*CboFDH*) or $NADP^+$ (*TeSADH W110A*), 250 mM formate (*CboFDH*) or 30 mM cyclohexanol (final 4% v/v DMSO for *TeSADH W110A*), 20 μ L CFE, 200 μ L final volume, 22°C. Formation of the reduced cofactor was followed for 1 minute, and the volumetric units ($U\ mL^{-1}$, defined as the quantity of enzyme that reduces 1 μ mol of oxidized cofactor in 1 min) were calculated from slopes of NAD(P)H formation calculated over 10-30 s using $\epsilon_{NADH} = 6.22\ mM^{-1}\ cm^{-1}$ and a pathlength of 0.55 cm.

7.4.1.5 Chemical Synthesis and Characterization

7.4.1.5.1 Synthetic procedures

Compounds **1a-b** were purchased from commercial suppliers. Preparation and spectroscopic characterization for compounds **1c-f** are reported in Aleku *et al.*^[2]

7.4.1.5.2 Reductive amination procedure for the preparation of amine products 4c-d and 5c

To a solution of the corresponding ketone (2.0 mmol) in dry THF (5 mL) under N₂ were added the corresponding amine (2.2 mmol), sodium triacetoxyborohydride (0.636 g, 3.0 mmol) and glacial acetic acid (0.114 mL, 2.0 mmol). The reaction was stirred for 16 hours at 20°C under N₂ then carefully quenched by addition of 1 M aqueous HCl (10 mL). Ethyl acetate (EtOAc, 10 mL) was added and the phases separated. The aqueous phase was extracted with a further portion of EtOAc (10 mL). The aqueous phase was then basified to pH 12 by addition of 5 M NaOH. The product was extracted into EtOAc (2 x 20 mL). The organic phase was dried over anhydrous MgSO₄ and the solvent removed under reduced pressure to afford the corresponding amine products with no further purification required.

7.4.1.5.3 Reductive Amination Procedure for the Preparation of Amine Product **4f**

To a stirred flask of cyclopentanone (0.2 g, 2.4 mmol) in dry methanol (MeOH, 10 mL) over 4 Å molecular sieves under nitrogen was added cyclopropylamine (0.332 mL, 4.8 mmol). The reaction was stirred at room temperature overnight, after which NaBH₄ (0.181 g, 4.8 mmol) was added over 10 minutes. The mixture was stirred for a further 2 hours, then the solvent was removed under reduced pressure. The resulting slurry was resuspended in EtOAc (20 mL), filtered then extracted twice with 1 M HCl. The aqueous phase was then basified (pH 12) with 5 M NaOH and extracted three times with EtOAc. The organic phases were combined, dried over MgSO₄ and the solvent removed under reduced pressure to afford the title compound as a yellow oil.

7.4.1.5.4 Typical Procedure for RedAm-catalysed Reductive Amination

A 500 μL reaction mixture contained 20 mM D-glucose, 0.5 mg mL^{-1} GDH (Codexis, CDX-901), 0.5 mM NADP^+ , 1 mg mL^{-1} purified RedAm, 5 mM ketone, 100 mM amine nucleophile (in Tris-HCl buffer adjusted to pH 9) and 2 % v/v DMSO. Reactions were incubated at 30 $^{\circ}\text{C}$ with 250 rpm shaking for 24 h, after which they were quenched by the addition of 30 μL of 10 M NaOH and extracted twice with 500 μL methyl *tert*-butyl ether (MTBE). The organic fractions were combined and dried over anhydrous MgSO_4 and analyzed by GC-FID.

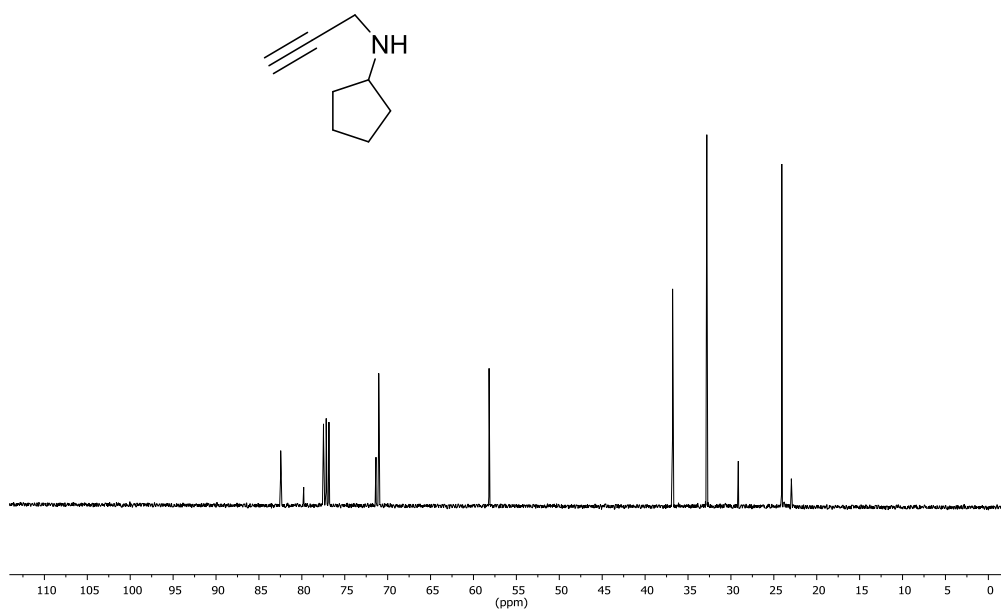
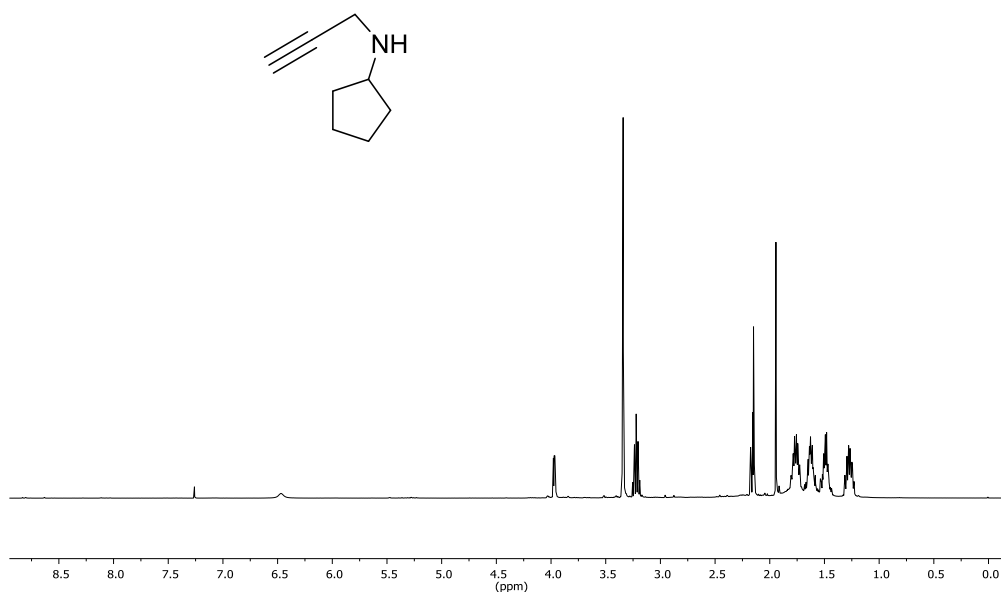
7.4.1.6 Spectroscopical Characterization

Spectra from ^1H and ^{13}C NMR runs were recorded on a Bruker Avance 400 instrument (400 MHz for ^1H and 100 MHz for ^{13}C) in CDCl_3 using residual protic solvent as an internal standard. Reported chemical shifts (δ) (in parts per million (ppm)) are relative to the residual protic solvent signal (CHCl_3 in CDCl_3 , $^1\text{H} = 7.26$; $^{13}\text{C} = 77.0$).

High-resolution mass spectrometry (HRMS) was recorded using a Waters LCT time-of-flight mass spectrometer, connected to a Waters Alliance LC (Waters, Milford, MA, USA). Data were processed with Waters Masslynx software.

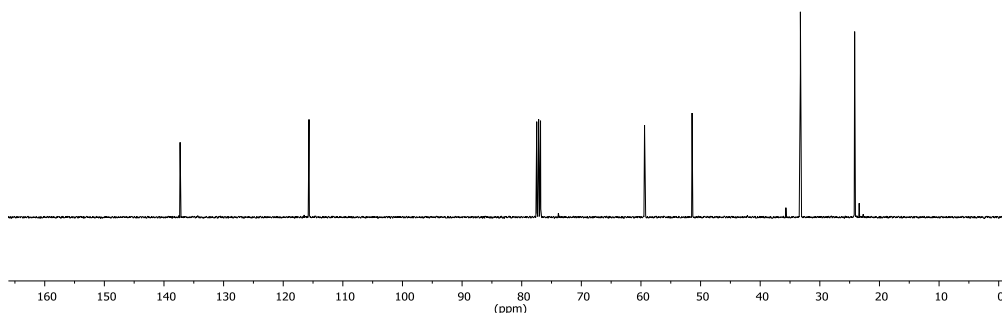
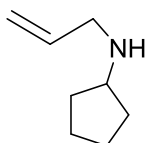
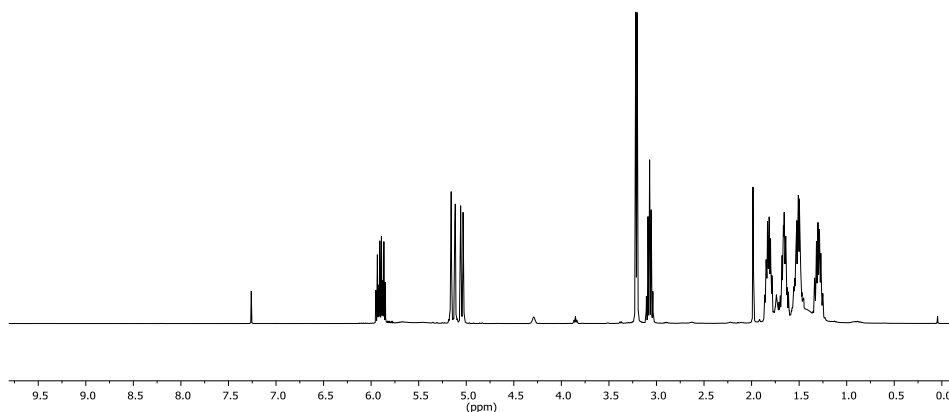
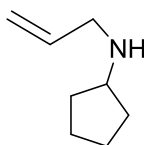
7.4.1.6.1 N-propargylcyclopentylamine **4c**

Brown oil isolated, 12% yield. $^1\text{H NMR}$ δ_{H} (400 MHz, CDCl_3) 3.41 (d, $J = 2.5$ Hz, 2H), 3.29 (quint, $J = 11.2$ Hz, 1H), 2.20 (t, $J = 2.5$ Hz, 1H), 1.80 – 1.44 (m, 9H). $^{13}\text{C NMR}$ δ_{C} (100 MHz CDCl_3) 82.46 (C), 70.98 (CH), 58.18 (CH), 36.79 (CH_2), 32.80 (CH_2), 24.06 (CH_2). HRMS calcd. for $\text{C}_8\text{H}_{14}\text{N}^+$ 124.1121 $[\text{M}+\text{H}]^+$, found 124.1099.



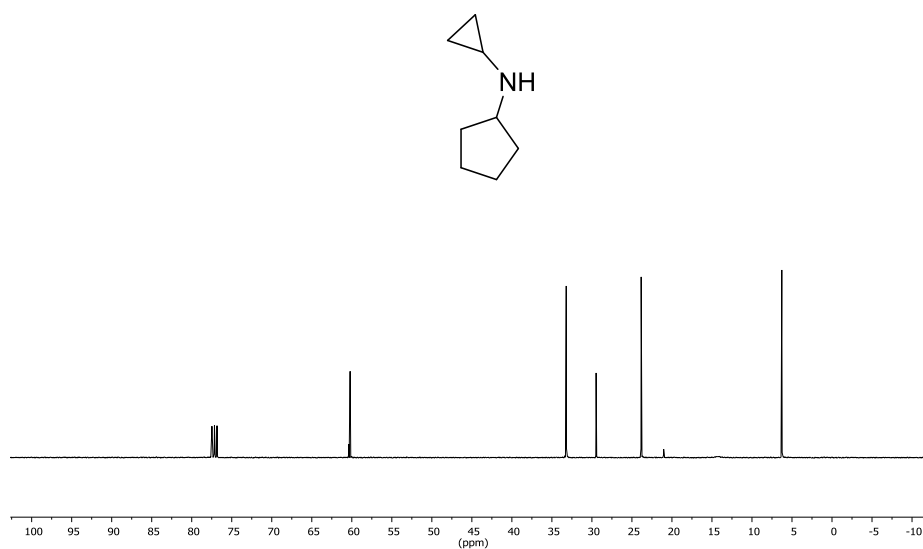
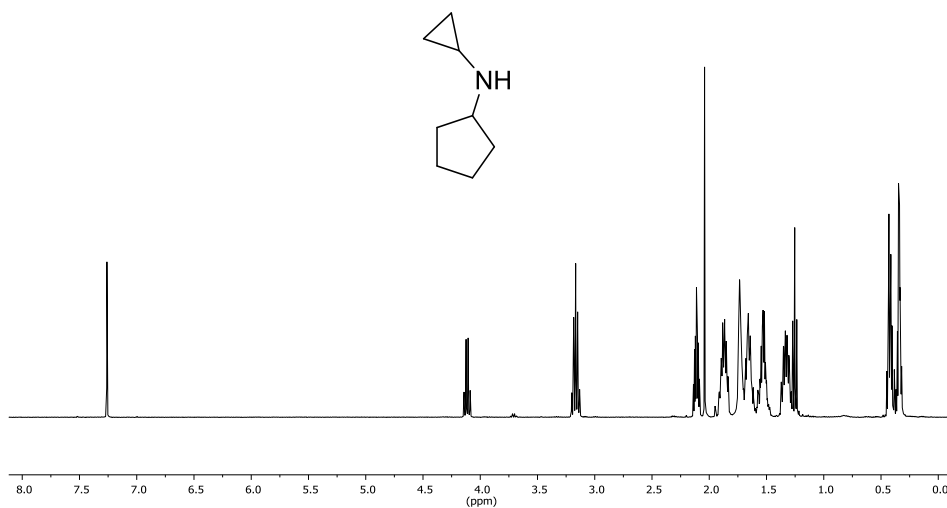
7.4.1.6.2 N-allylcyclopentylamine **4e**

Yellow oil isolated, 58% yield. $^1\text{H NMR}$ δ_{H} (400 MHz, CDCl_3) 6.01 – 5.70 (m, 1H), 5.27 – 4.96 (m, 2H), 4.07 – 3.65 (m, 1H), 3.22 (dt, $J = 6.1, 1.4$ Hz, 1H), 3.08 (p, $J = 6.8$ Hz, 1H), 1.92 – 1.73 (m, 2H), 1.77 – 1.57 (m, 2H), 1.59 – 1.44 (m, 2H), 1.31 (ddd, $J = 13.3, 7.6, 4.0$ Hz, 2H).. $^{13}\text{C NMR}$ δ_{C} (100 MHz CDCl_3) 137.3 (CH), 115.7 (CH₂), 59.4 (CH), 51.4 (CH₂), 42.2 (CH₂), 33.3 (CH₂), 24.2 (CH₂), 23.4 (CH₂). **HRMS** calcd. for $\text{C}_8\text{H}_{16}\text{N}^+$ 126.1277 $[\text{M}+\text{H}]^+$, found 126.1265.



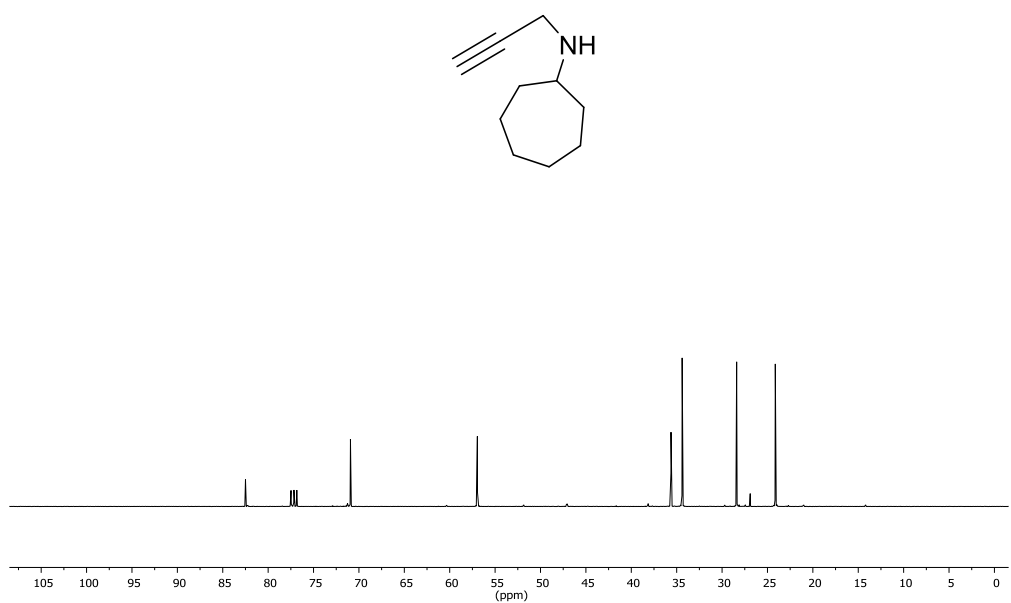
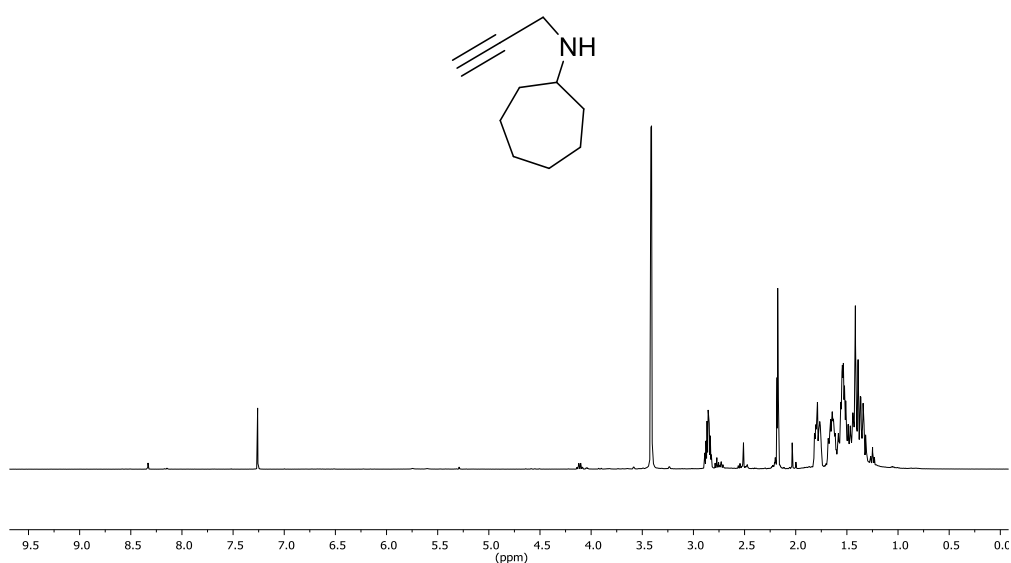
7.4.1.6.3 N-cyclopropylcyclopentylamine **4f**

Yellow oil isolated, 64% yield. $^1\text{H NMR}$ δ_{H} (400 MHz, CDCl_3) 3.02 (quin, $J = 6.9$ Hz, 1H), 1.99 – 1.94 (m, 1H), 1.76 – 1.69 (m, 2H), 1.61 – 1.56 (s, 1H), 1.55 – 1.45 (m, H), 1.43 – 1.33 (m, 2H), 1.23 – 1.15 (m, 2H), 0.32 – 0.17 (m, 4H). $^{13}\text{C NMR}$ δ_{C} (100 MHz CDCl_3) 60.13 (CH), 33.12 (CH_2), 29.39 (CH), 23.76 (CH_2), 6.16 (CH_2). **HRMS** calcd. for $\text{C}_8\text{H}_{16}\text{N}^+$ 126.1204 $[\text{M}+\text{H}]^+$, found 126.1281.



7.4.1.6.4 N-propargylcycloheptylamine **5c**

Orange oil isolated, 76% yield. $^1\text{H NMR}$ δ_{H} (400 MHz, CDCl_3) 3.43 (d, $J = 2.5$ Hz, 2H), 2.87 (quint, $J = 10.5, 4.1$ Hz, 1H), 2.20 (t, $J = 2.5$ Hz, 1H), 1.82 – 1.79 (s, 1H), 1.67 – 1.53 (m, 12H). $^{13}\text{C NMR}$ δ_{C} (100 MHz CDCl_3) 82.40 (C), 70.98 (CH), 56.97 (CH), 35.59 (CH_2), 34.35 (CH_2), 28.36 (CH_2), 24.13 (CH_2). **HRMS** calcd. for $\text{C}_{10}\text{H}_{18}\text{N}^+$ 152.1434 $[\text{M}+\text{H}]^+$, found 152.1400.



7.4.1.7 Biocatalytic Cascade for the Amination of Cycloalkanes

7.4.1.7.1 Amine Inhibition Studies

In order to investigate the effect of amines **a–f** at different loadings on the performance of the P450, the hydroxylation step was carried out in 28 mL glass vials, with 6 μM P450, 0.3 U mL^{-1} *CboFDH* (both added as concentrated CFE), 1 mM NAD^+ , 250 mM sodium formate, various concentrations of the amine nucleophile and finally 125 μL cycloalkane added neat. Reaction volumes were made up to 500 μL with Tris-HCl 0.2 M, pH 8 and vials incubated in a horizontal position at 20°C with 200 rpm orbital shaking for 24 h. Reactions were extracted in 500 μL MTBE (2 x 250 μL) with 2 mM decane as internal standard. The organic fractions were dried over anhydrous MgSO_4 and analyzed by GC-FID as reported in section 7.4.1.8. The relative performance was expressed as residual total turnover numbers (TTNs, ratio between total product concentration and P450 concentration, calculated over 24 h), that is the ratio between TTNs in the presence of the amine and TTNs when the amine nucleophile was not added to the reaction mixture.

7.4.1.7.2 One-step Process

The one-step cascade was carried out in 28 mL glass vials, with 1.8 μM P450, 0.16 U mL^{-1} *CboFDH* (both added as freshly lysed CFE), 0.3 U mL^{-1} *TeSADH* W110A (added as rehydrated lyophilized CFE), 2 mg mL^{-1} purified *AspRedAm*, 1 mM of both NAD^+ and NADP^+ , 250 mM sodium formate, various concentrations of the amine nucleophile (as reported in Table S1) and finally 125 μL cyclohexane added neat.

Reaction volumes were made up to 500 μL with Tris-HCl 0.2 M, pH 8 and vials incubated in a horizontal position at 20°C with 200 rpm orbital shaking for 24 h. Reactions were quenched by addition of NaOH (5 M, 100 μL) and extracted with 500 μL MTBE (2 x 250 μL) with 2 mM decane as internal standard. The organic fractions were dried over anhydrous MgSO_4 and analysed by GC-FID. When ammonia was employed as amine nucleophile, samples were derivatized with acetic anhydride and an excess of triethylamine at room temperature prior to analysis.

7.4.1.7.3 Two-step Process

The two-step biocatalytic cascade was carried out in 28 mL glass vials, with 6 μM P450, 0.3 U mL^{-1} *Cbo*FDH (both added as concentrated CFE), 1 mM NAD^+ , 250 mM sodium formate and finally 125 μL cycloalkane added neat. Reaction volumes were made up to 500 μL with Tris-HCl 0.2 M, pH 8 and vials incubated in a horizontal position at 20°C with 200 rpm orbital shaking for 24 h. Afterwards, the components for the second step were added as follows: 0.2 U mL^{-1} *Te*SADH W110A (added as rehydrated lyophilized CFE), 2 mg mL^{-1} purified *Asp*RedAm, 1 mM NADP^+ and 100 mM of the appropriate amine nucleophile. Reaction volumes were made up to 750 μL and vials incubated at 30°C with 200 rpm shaking.

When purified proteins were used, reactions were set-up similarly with 6 μM P450, 1 U mL^{-1} *Cbo*FDH, 2 U mL^{-1} *Te*SADH W110A, 2 mg mL^{-1} *Asp*RedAm and 25 mM allylamine. The amount of cyclohexane added was reduced to 20 μL in order to decrease possible inhibitory effects on enzymatic performance while preserving P450 activity.^[3]

After 24 h, reactions were quenched by addition of NaOH (5 M, 150 μL) and extracted with 750 μL MTBE (2 x 375 μL) with 2 mM decane as internal standard. When cyclopentane was used as starting material, 4 mM cyclohexane instead of 2 mM decane was used as internal standard. The organic fractions were dried over anhydrous MgSO_4 and analyzed by GC-FID as reported in section 7.4.1.8. When ammonia was employed as amine nucleophile, samples were derivatized with acetic anhydride and an excess of triethylamine at room temperature.

7.4.1.7.4 Preparative-scale Production of *N*-propargylcyclohexylamine **1c** by Two-Step Amination Cascade with Cyclohexane and Propargylamine **c**

The two-step amination cascade was scaled-up linearly to a final volume of 30 mL. The first step was carried out on a 20 mL scale using Tris-HCl 0.2 M, pH 8 containing 6 μM P450, 0.3 U mL^{-1} *Cbo*FDH (both added as concentrated CFE), 1 mM NAD^+ , 250 mM sodium formate and finally 5 mL cyclohexane added neat. Reactions were performed either employing a 500 mL round-bottom flask with magnetic stirring (200 rpm) or a 250 mL conical flask incubated at 18°C with 150 rpm orbital shaking. The latter proved to be superior in terms of cyclohexanol **2** produced after 19 h (Figure S3). Similarly to small-scale reactions, the second step was performed with 0.2 U mL^{-1} *Te*SADH W110A (added as rehydrated lyophilized CFE), 2 mg mL^{-1}

purified AspRedAm, 1 mM NADP⁺ and 100 mM propargylamine **c**. The reaction was incubated at 30°C and 150 rpm orbital shaking for 2 hours. The solution was then basified with 5 M NaOH (1 mL) and the product extracted into MTBE (2 x 30 mL). The organics were separated by centrifugation (4°C, 2,831 *g*, 2 minutes), collected, dried over anhydrous MgSO₄ and the solvent removed under reduced pressure to give a yellow oil. The latter was dissolved in 10 mL dry diethyl and the product (50 mg) was obtained as yellow hydrochloride salt after addition of 2 M ethereal HCl (300 µL) and filtration. Further details can be found in section 7.4.3.1.

NMR and HRMS characterization data were in accordance with previously reported data.^[2,4]

7.4.1.8 Analytcs

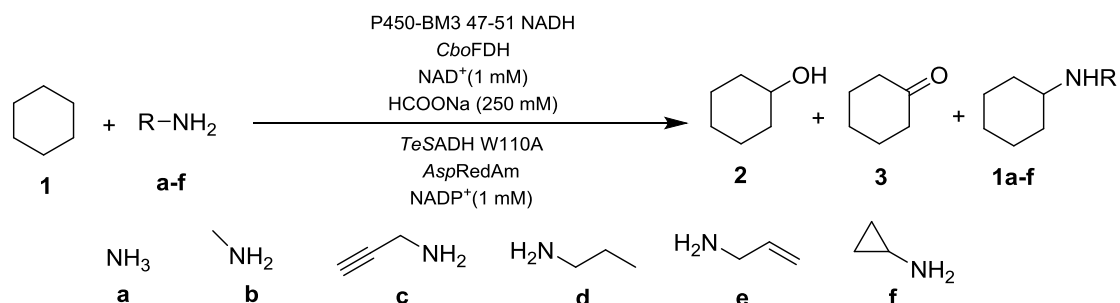
GC analysis was performed on an Agilent 6850 system (Agilent, Santa Clara, CA, USA) with a flame ionization detector (FID) equipped with a CP-ChiraSil-DEX CB column (Agilent, Santa Clara, CA, USA with 0.25 µm film thickness and 0.25 mm internal diameter. After sample preparation, 2 µL were injected at a split ratio 100:1. The inlet temperature was set at 200°C, the detector temperature at 250°C and the pressure maintained at 6.8 psi, using helium as carrier gas. The following method was applied: 50°C; 10°C min⁻¹ to 200°C; hold 2 min.

Yields were determined using the appropriate calibration curve internal standard vs cyclopentanone **8**, cyclohexanone **3** and cycloheptanone **9**. Moreover, calibration curves were also constructed for *N*-propargylcyclohexylamine **1c** and *N*-cyclopropylcyclohexylamine **1f**.

GC-MS analysis was performed on an Agilent 7890B/5977B series gas chromatograph/mass selective detector equipped (electron impact positive mode) with an Agilent HP-1 MS column (0.32 mm internal diameter, 30 m length, 0.25 µm film thickness). Helium was used as the carrier gas at flow rate of 1.6 ml min⁻¹ and the injection temperature was 270°C. Injection volume was 2 µl and samples were injected at a split ratio of 10:1. The following method was applied: 50°C; 10°C min⁻¹ to 200°C; hold 2 min.

7.4.2 Supporting Figures and Tables

Table S1. One-step amination of cyclohexane at different amine loadings. Values reported in the text and in Table S2 are written in bold.



Amine donor	Amine donor concentration (mM)	Product concentration (mM)		
		Cyclohexanol	Cyclohexanone	Amine
a	250	1.5	11.6	0.2
	100	2.3	9.6	0.2
	50	4.0	9.7	0.1
b	250	1.8	5.1	4.5
	100	3.6	6.9	5.4
	50	6.0	8.2	4.7
c	25	1.4	0.7	3.6
	10	4.7	2.7	4.1
	50	6.7	9.3	5.3
d	25	6.9	9.5	2.5
	10	15.2	14.2	1.8
	50	1.9	3.4	6.1
e	25	4.7	5.9	8.3
	10	9.4	9.3	6.2
	50	0.2	0.2	1.7
f	25	0.5	0.7	2.1
	10	1.0	1.9	1.6

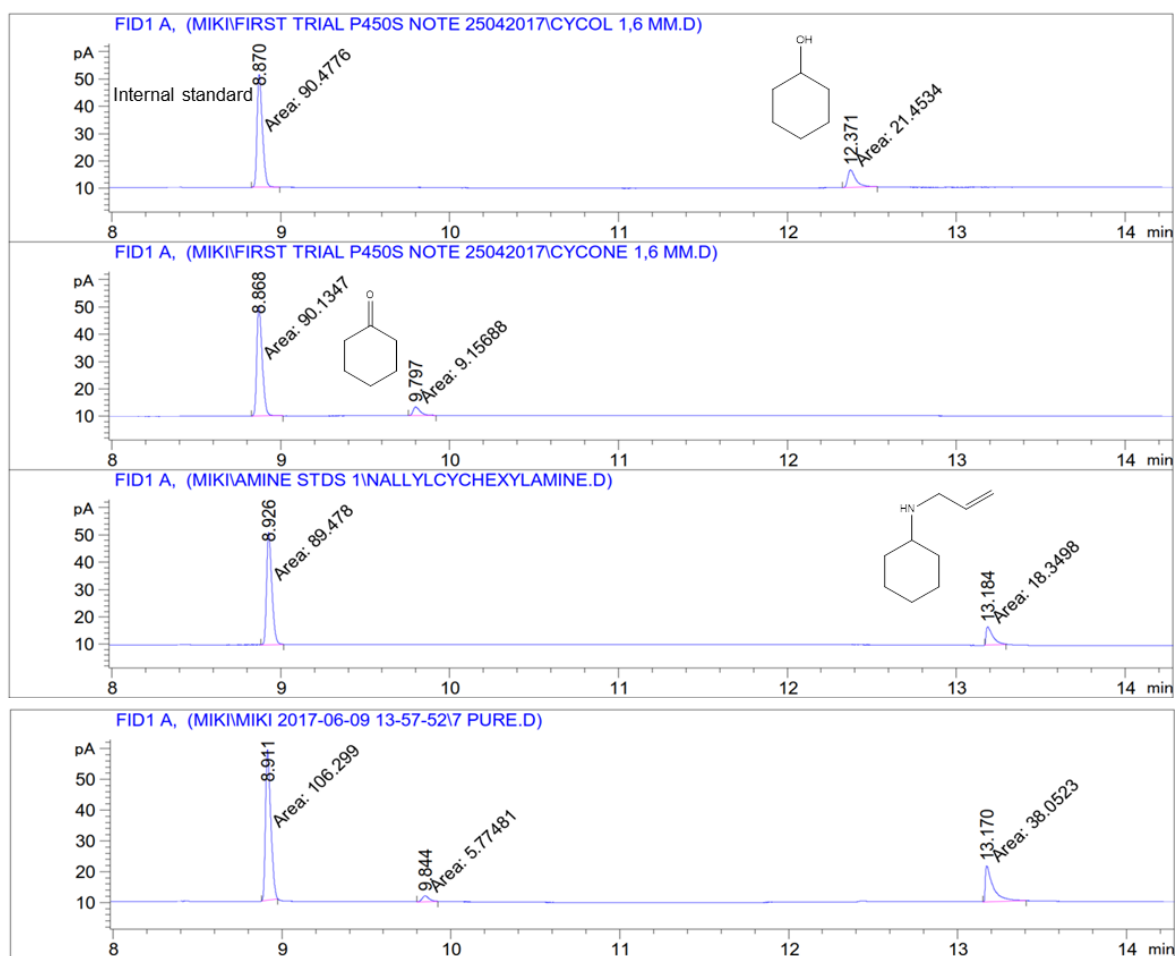
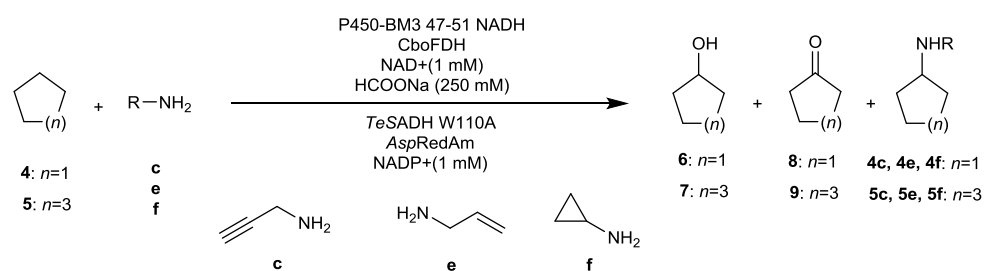


Figure S1. GC-FID chromatogram for the one-pot cascade for the amination of cyclohexane **1** with allylamine **e** employing purified proteins. Standards (top, second and third trace) and two-step process (fourth trace) are shown. Yields of 1.2 mM amine **1e** can be calculated with only a trace amount of cyclohexanone **3** detected. The poor yield of amine product led us to employ crude enzyme preparations, instead of purified proteins.

Table S2. Comparison between the one-step and the two-step amination of cyclohexane.

	Amine donor	Amine donor	Product concentration (mM)		
		concentration (mM)	Alcohol	Ketone	Amine
One-step	a	100	2.3	9.6	0.2
	b	100	3.6	6.9	5.4
	c	25	1.4	0.7	3.6
	d	50	6.7	9.3	5.3
	e	25	4.7	5.9	8.3
	f	25	0.5	0.7	2.1
Two-step	a	100	3.4	15.2	0.4
	b	100	0.9	10.1	8.7
	c	100	2.0	5.0	19.6
	d	100	0.3	9.5	8.0
	e	100	0.7	8.7	15.4
	f	100	3.3	3.9	17.1

Table S3. Two-step amination of cyclopentane and cycloheptane with amines **c**, **e** and **f**.

Cycloalkane	Amine donor	Product concentration (mM)		
		Alcohol	Ketone	Amine
Cyclopentane	c	7.3	3.2	10.6
	e	7.1	5.0	2.2
	f	8.1	3.7	3.0
Cycloheptane	c	0.0	5.1	3.3

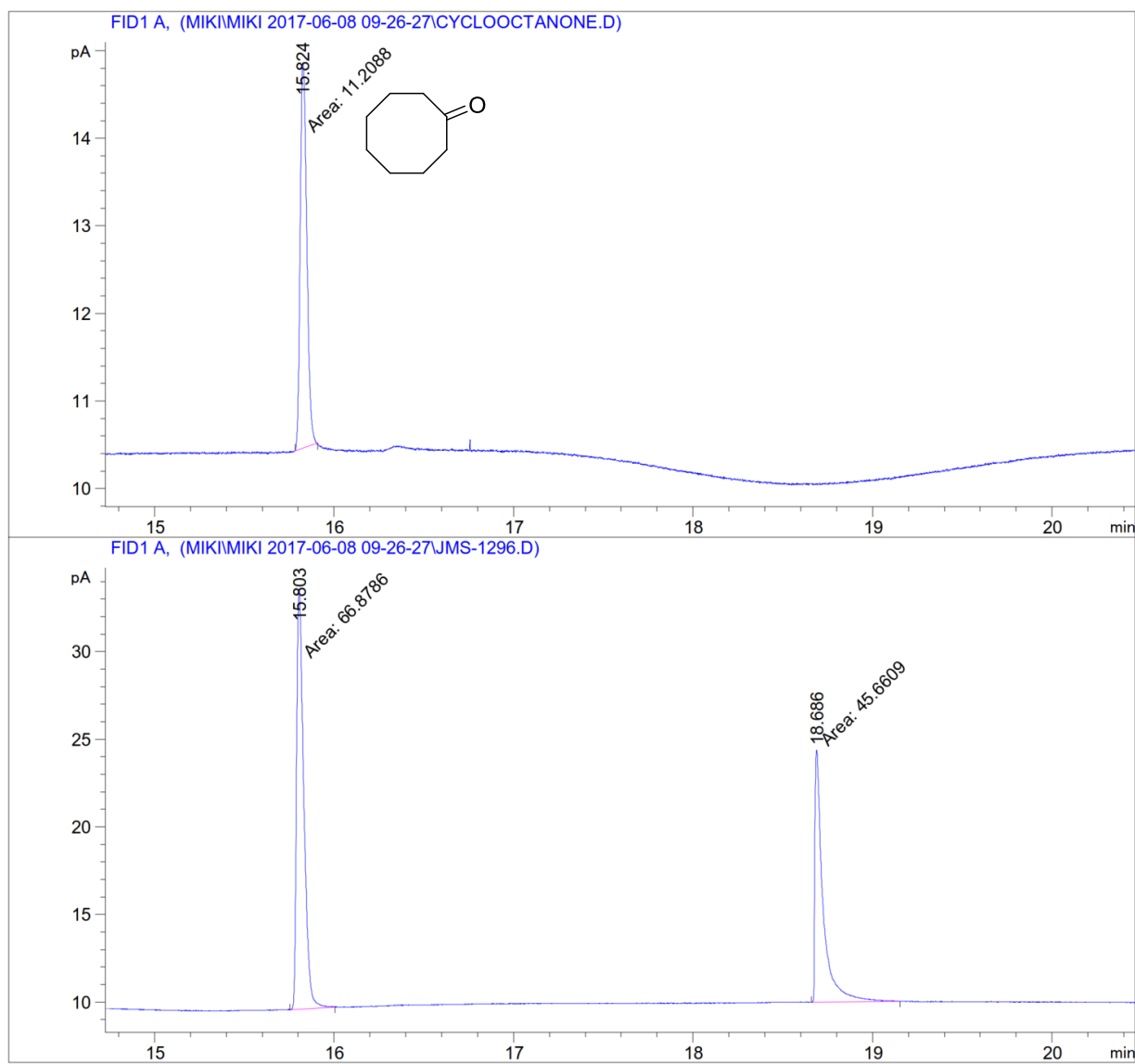


Figure S2. GC-FID chromatogram for the reductive amination of cyclooctanone with allylamine **e** employing *AspRedAm*. Standard (top) and biotransformation (second trace) are shown. A putative product peak in the biotransformation trace (40% conversion) suggests that *AspRedAm* is active in the reductive amination of cyclooctanone. Further details are given in section 7.4.3.2.

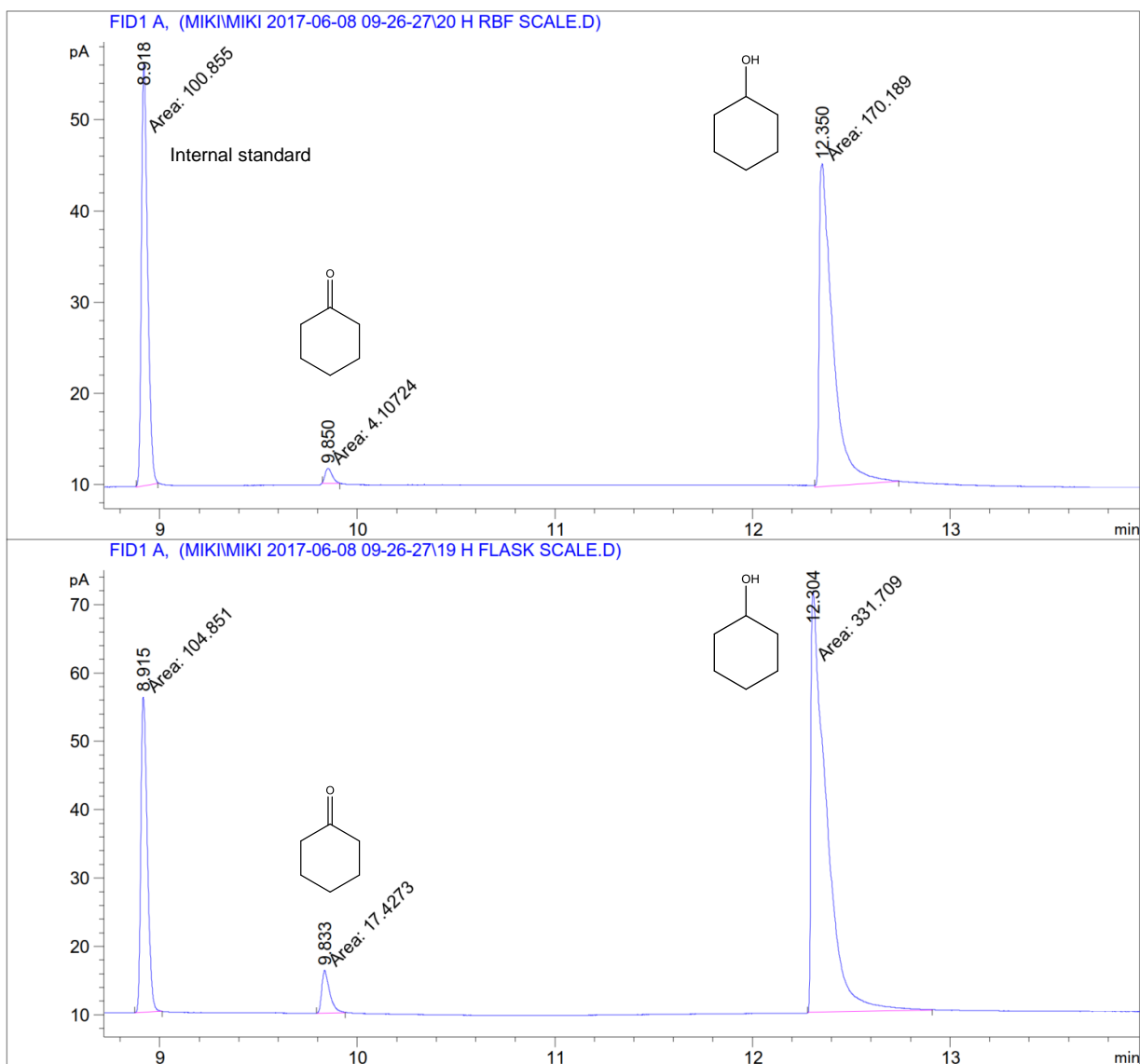


Figure S3. GC-FID chromatogram for the first step of the preparative scale amination of cyclohexane **1** showing cyclohexanol **2** production after 19-20 h. Changing the reaction vessel from a stirred round bottom flask (top trace) to a shaken conical flask (bottom trace), almost doubled the amount of cyclohexanol **2** produced after 19-20 h. Additionally, cyclohexanol **2** might be overoxidized by the P450, as suggested by cyclohexanone **3** formation.

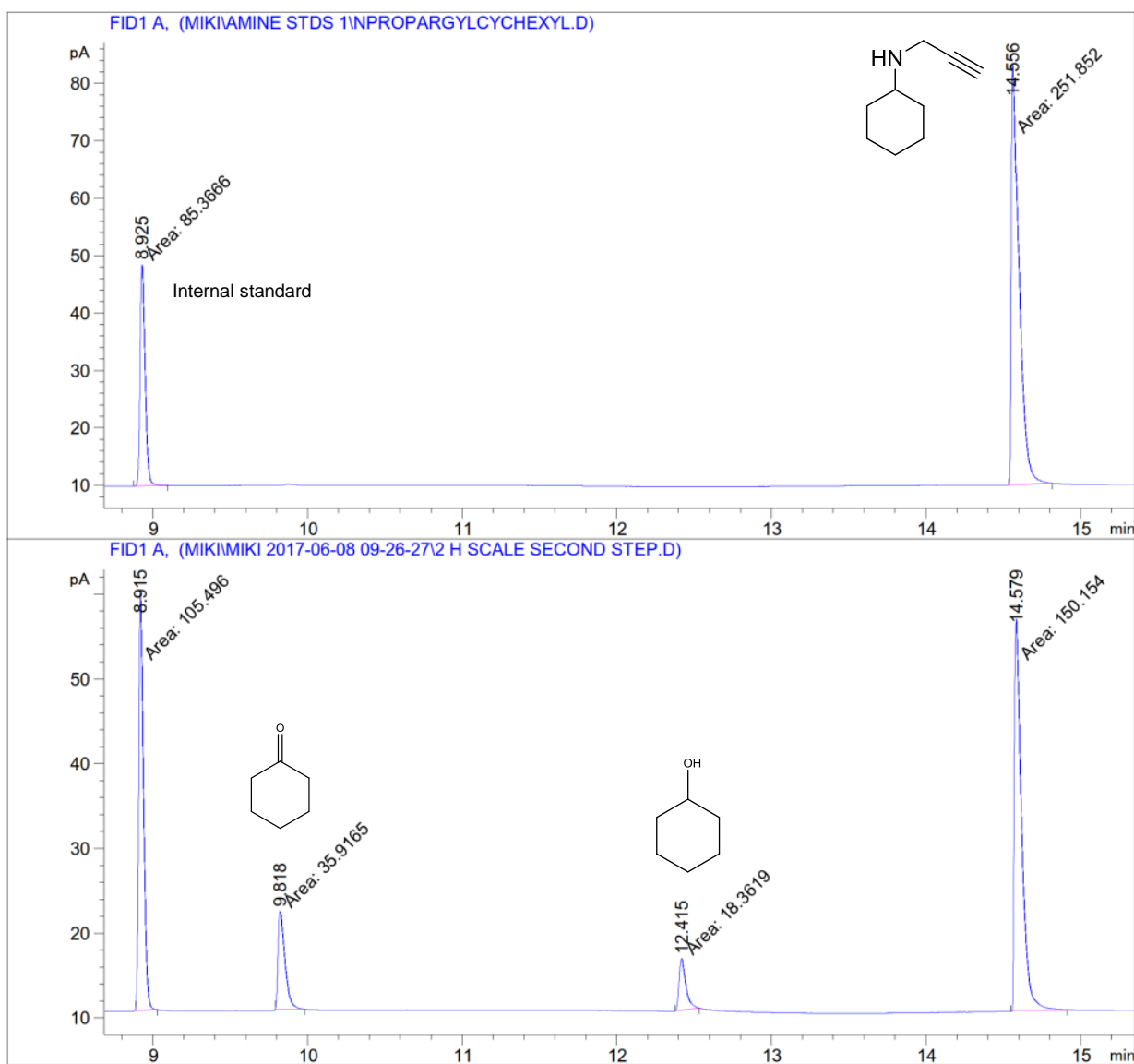


Figure S4. GC-FID chromatogram for the second step of the preparative scale amination of cyclohexane **1** with propargylamine **c** showing product standard (top trace) and biotransformation after 2 h.

7.4.3 GC-FID Chromatograms and GC-MS Traces

7.4.3.1 GC-FID Chromatograms

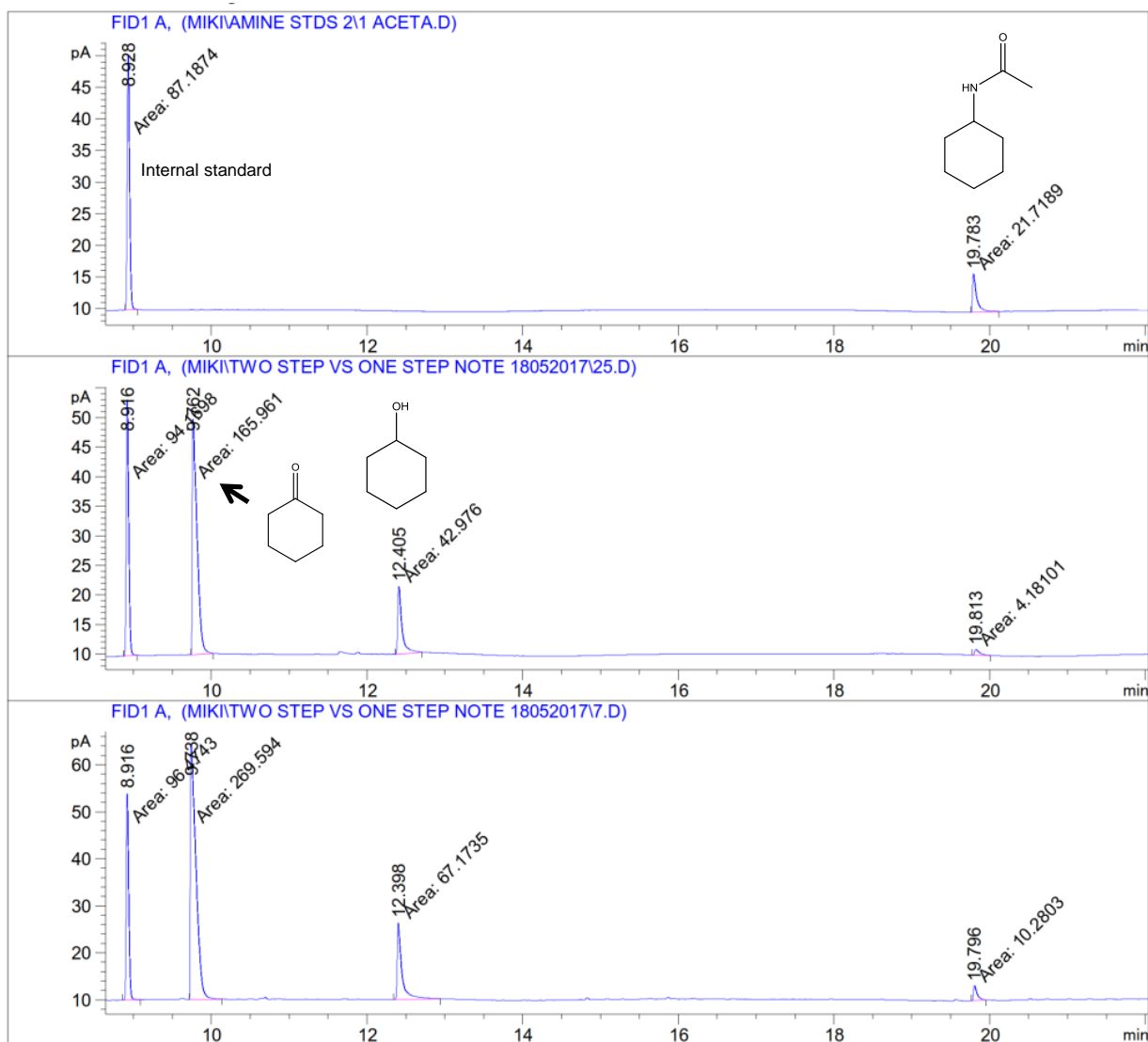


Figure S5. GC-FID chromatogram for the biocatalytic cascade for the amination of cyclohexane **1** with ammonia **a**. Amine standard (top trace), one-step cascade (middle trace) and two-step cascade (bottom trace) are shown. Samples were derivatized with acetic anhydride and an excess of triethylamine at room temperature prior to analysis.

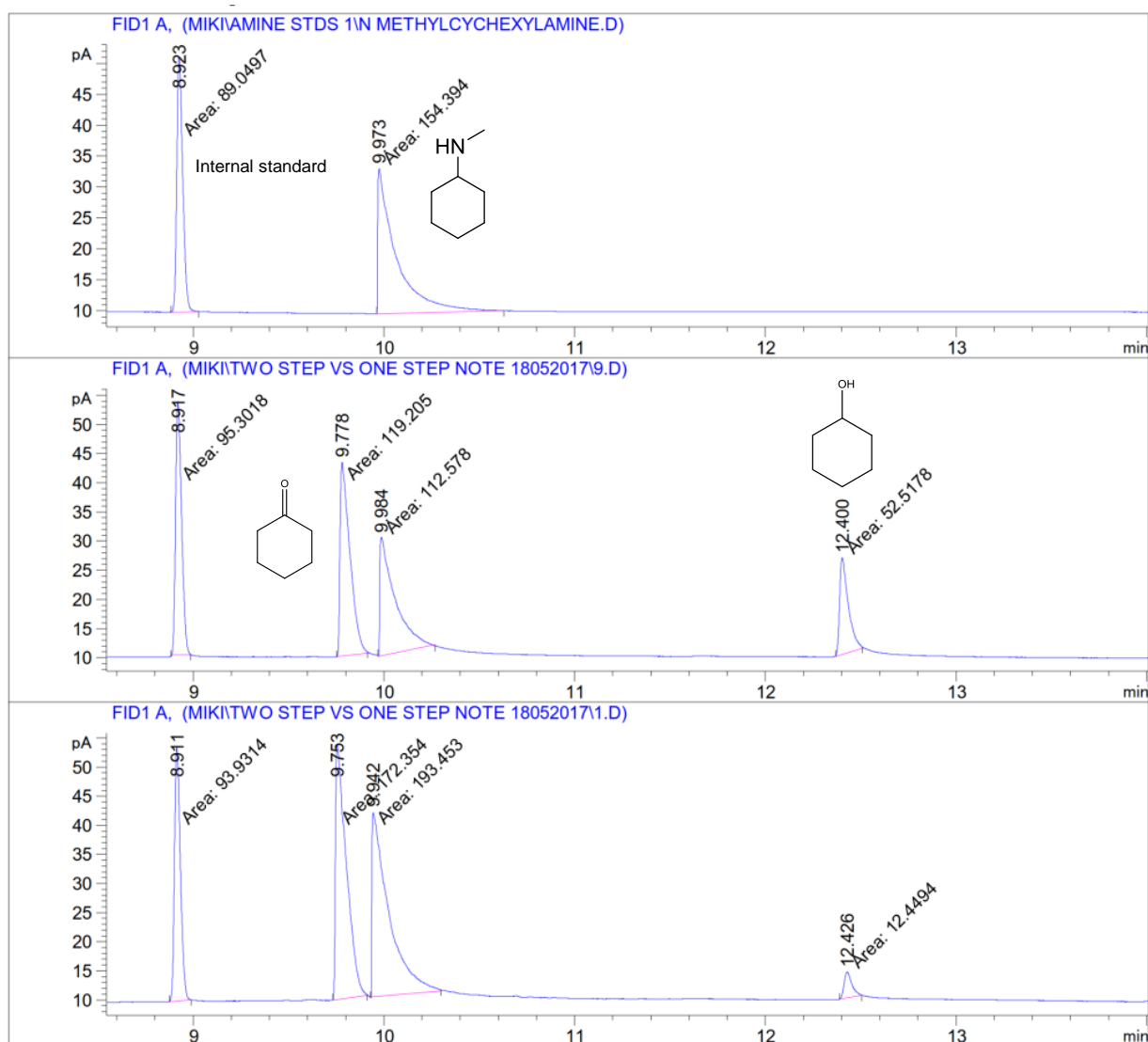


Figure S6. GC-FID chromatogram for the biocatalytic cascade for the amination of cyclohexane **1** with methylamine **b**. Amine standard (top trace), one-step cascade (middle trace) and two-step cascade (bottom trace) are shown.

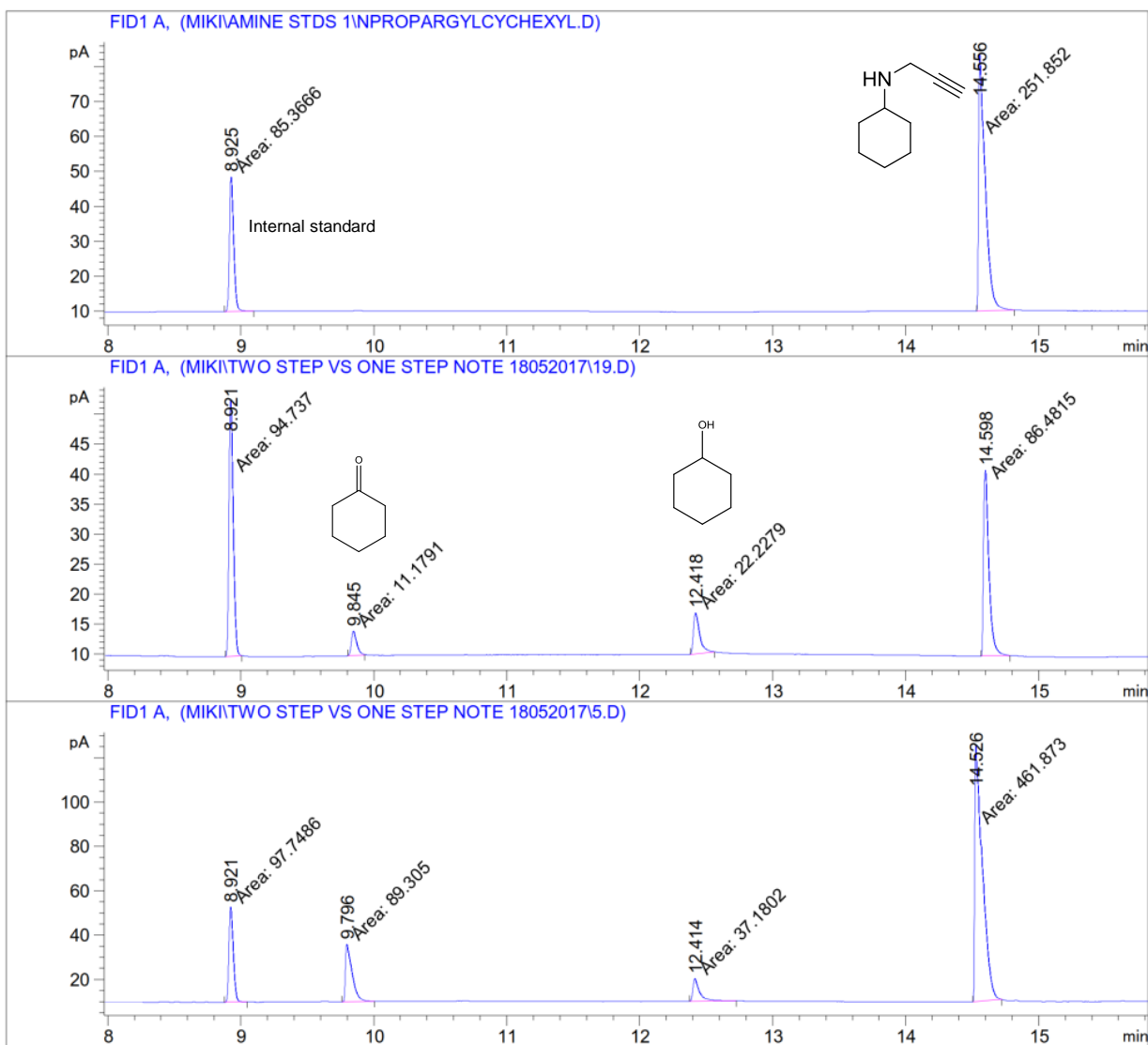


Figure S7. GC-FID chromatogram for the biocatalytic cascade for the amination of cyclohexane **1** with propargylamine **c**. Amine standard (top trace), one-step cascade (middle trace) and two-step cascade (bottom trace) are shown.

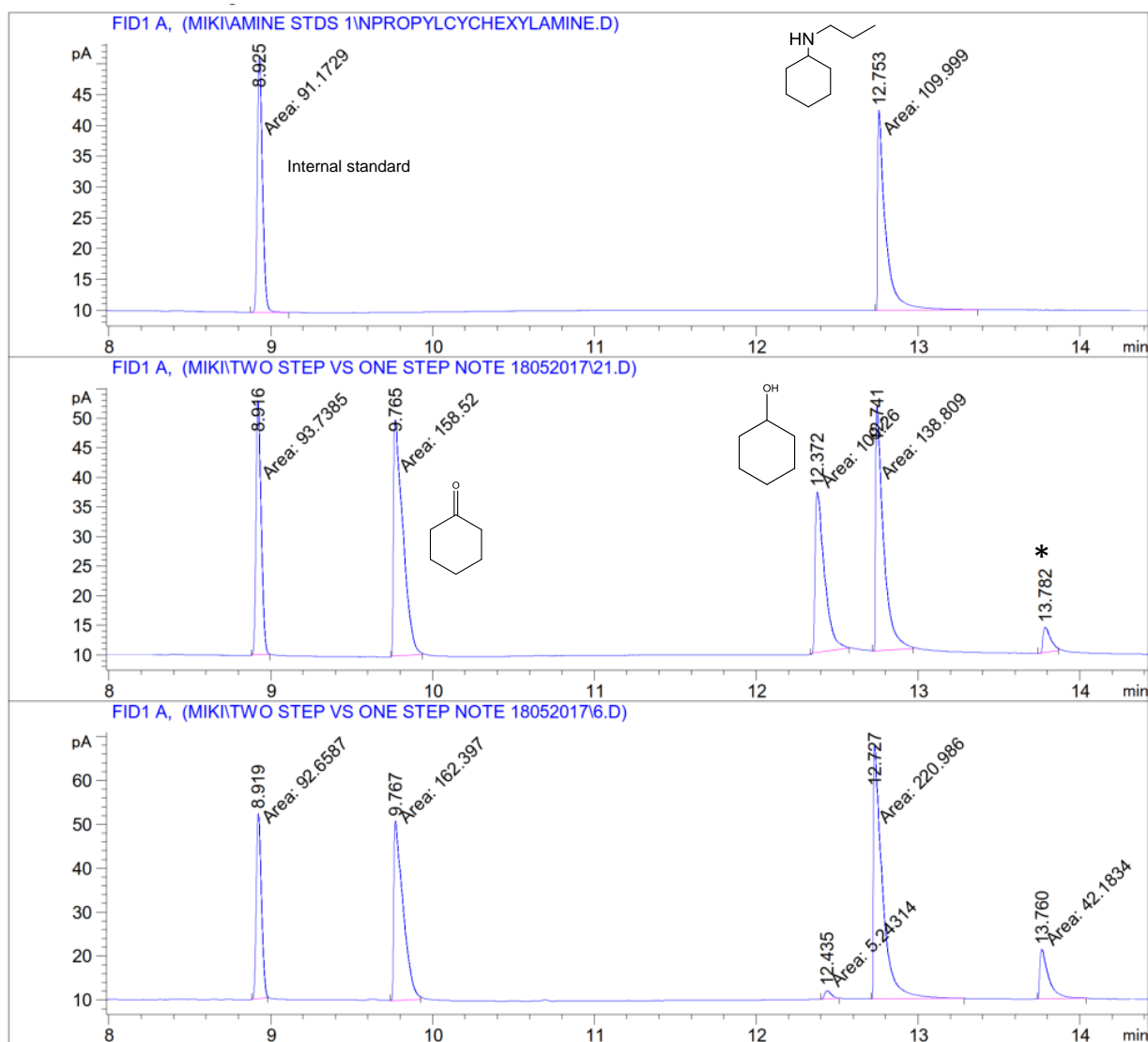


Figure S8. GC-FID chromatogram for the biocatalytic cascade for the amination of cyclohexane **1** with propylamine **d**. Amine standard (top trace), one-step cascade (middle trace) and two-step cascade (bottom trace) are shown. An asterisk indicates a putative imine peak. Further details are given in section 7.4.3.2.

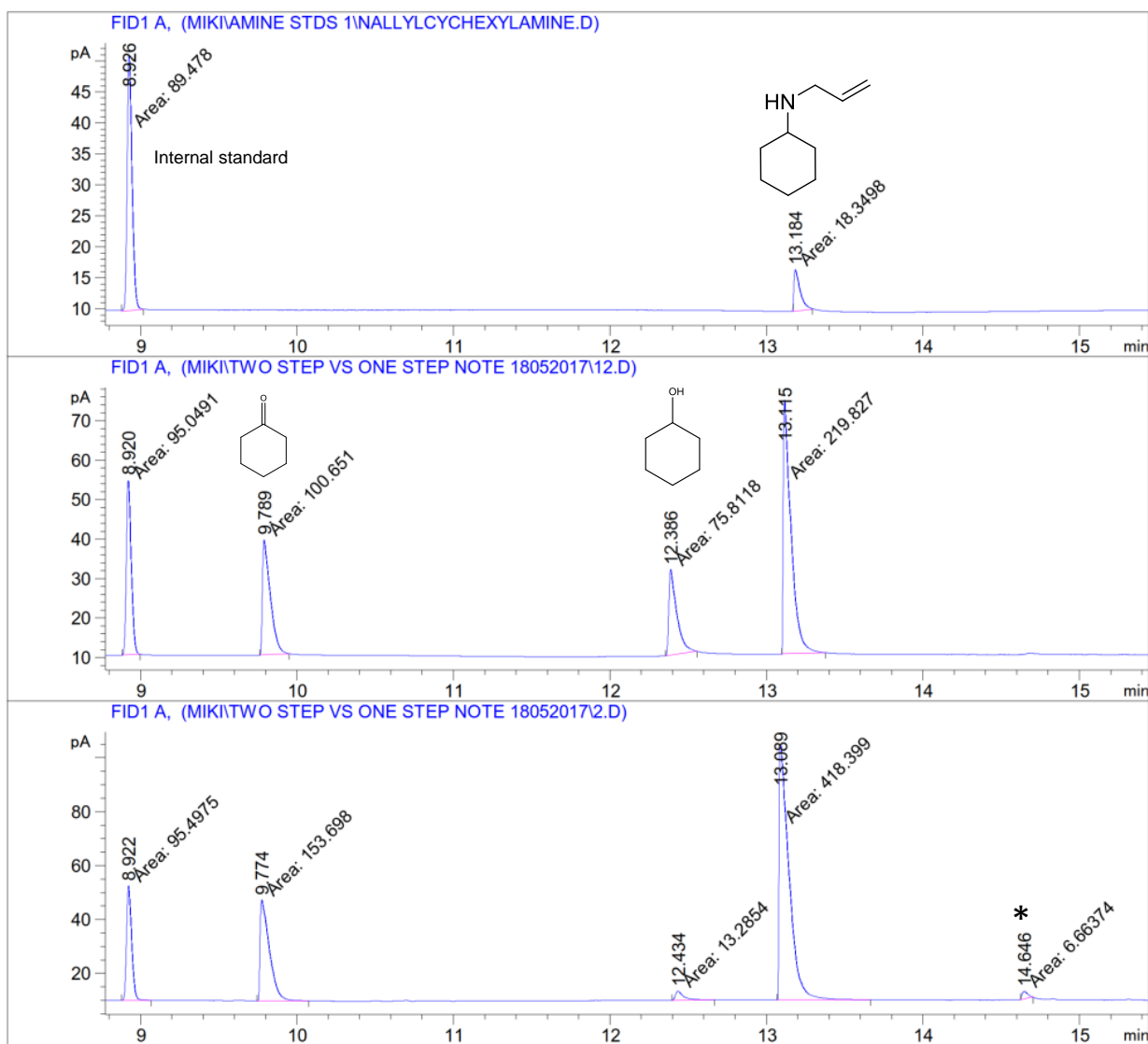


Figure S9. GC-FID chromatogram for the biocatalytic cascade for the amination of cyclohexane **1** with allylamine **e**. Amine standard (top trace), one-step cascade (middle trace) and two-step cascade (bottom trace) are shown. An asterisk indicates a putative imine peak. Further details are given in section 7.4.3.2.

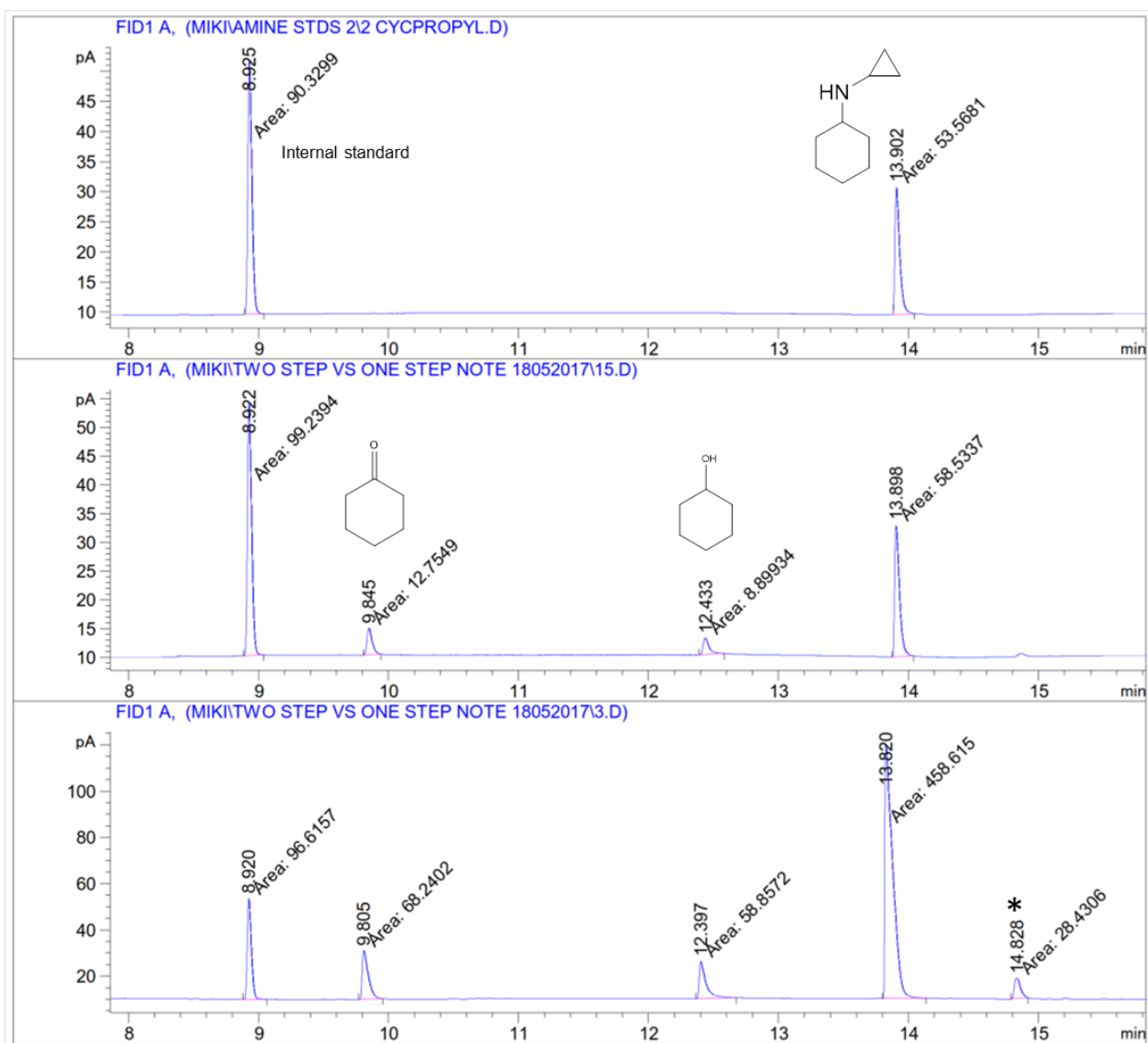


Figure S10. GC-FID chromatogram for the biocatalytic cascade for the amination of cyclohexane **1** with cyclopropylamine **f**. Amine standard (top trace), one-step cascade (middle trace) and two-step cascade (bottom trace) are shown. An asterisk indicates a putative imine peak. Further details are given in section 7.4.3.2.

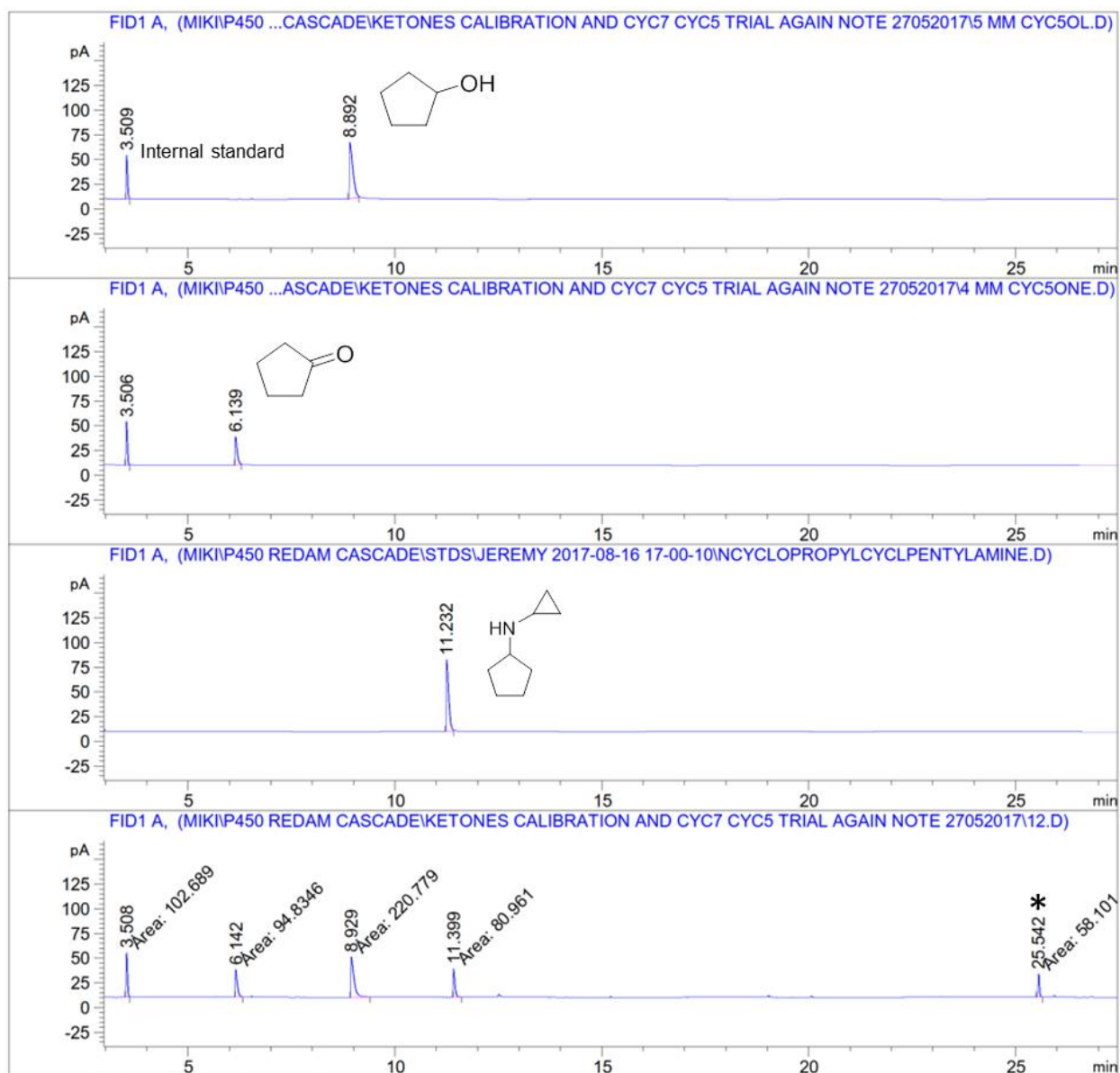


Figure S11. GC-FID chromatogram for the biocatalytic cascade for the amination of cyclopentane **4** with cyclopropylamine **f**. Standards (top, second and third trace) and two-step cascade (fourth trace) are shown. An asterisk indicates an unidentified by-product that is not detected after amination of cyclopentanol *via* hydrogen-borrowing (see GC trace below). Further details are given in Section 7.4.3.2.

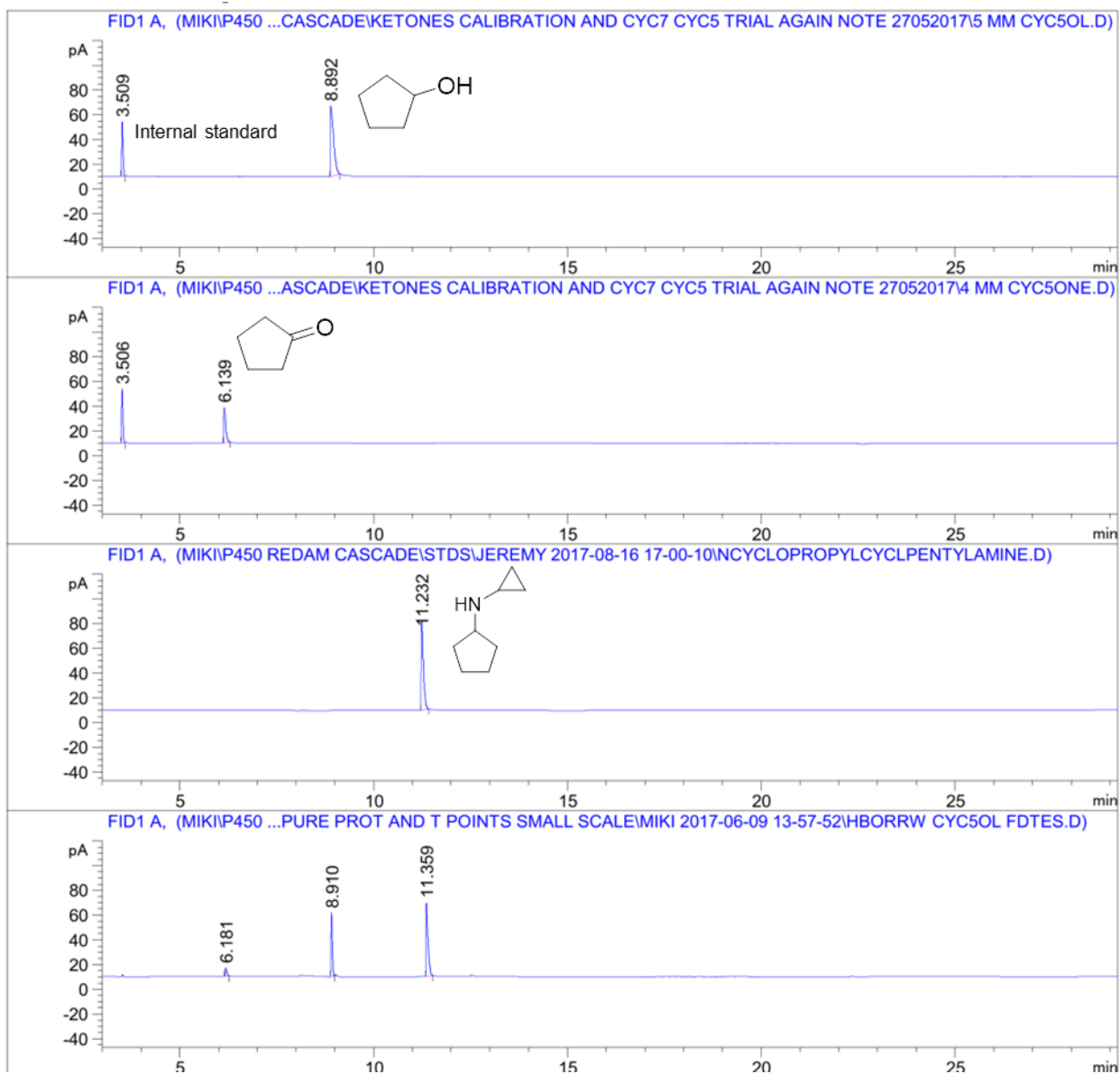


Figure S11. GC-FID chromatogram for *AspRedAm*/*TeSADH* W110A-catalyzed amination of cyclopentanol **4** with cyclopropylamine **f**, showing standards (top, second and third trace) and biotransformation (fourth trace).

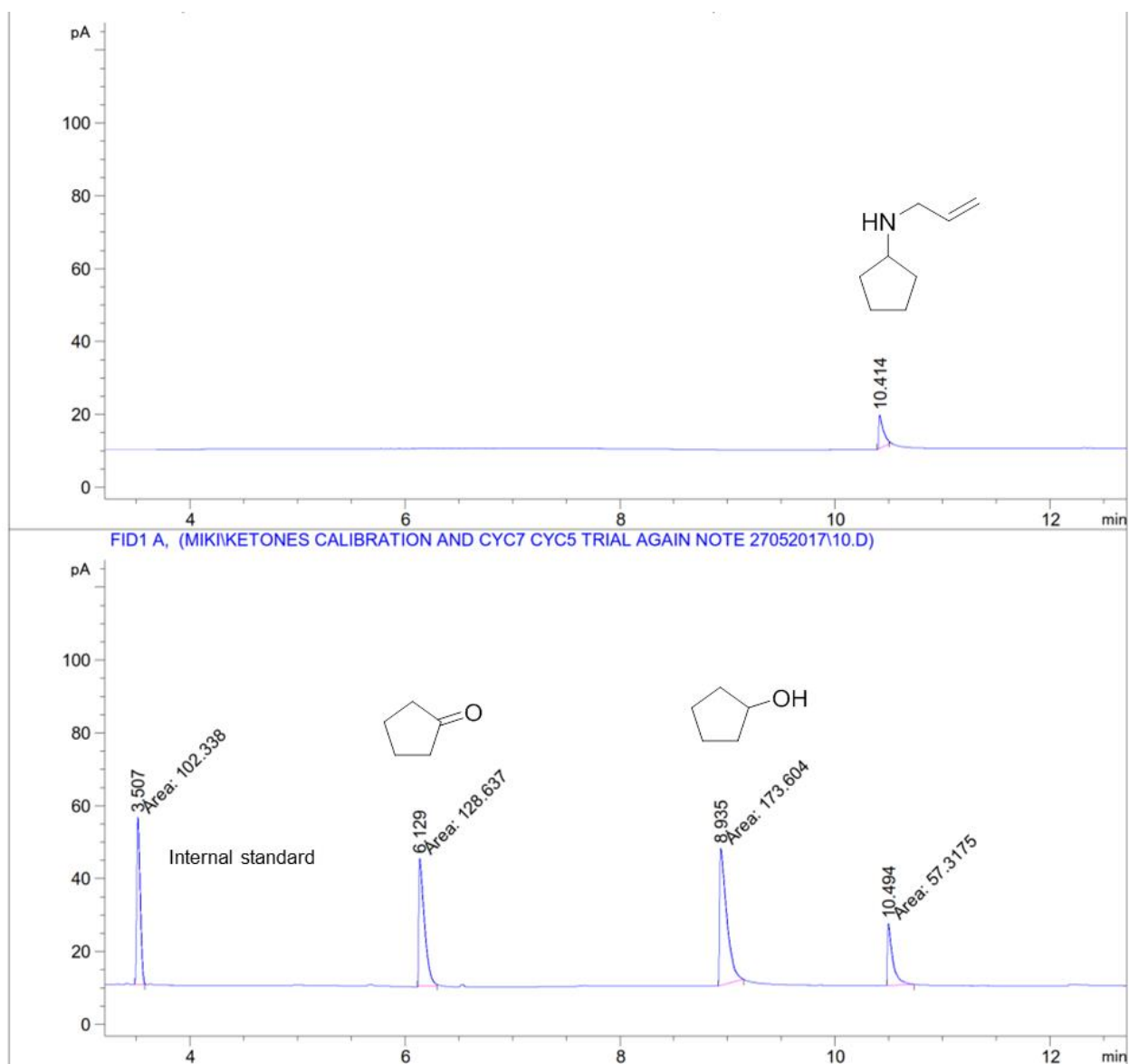


Figure S12. GC-FID chromatogram for the biocatalytic cascade for the amination of cyclopentane **4** with allylamine **e**. Amine standard (top trace) and two-step cascade (bottom trace) are shown.

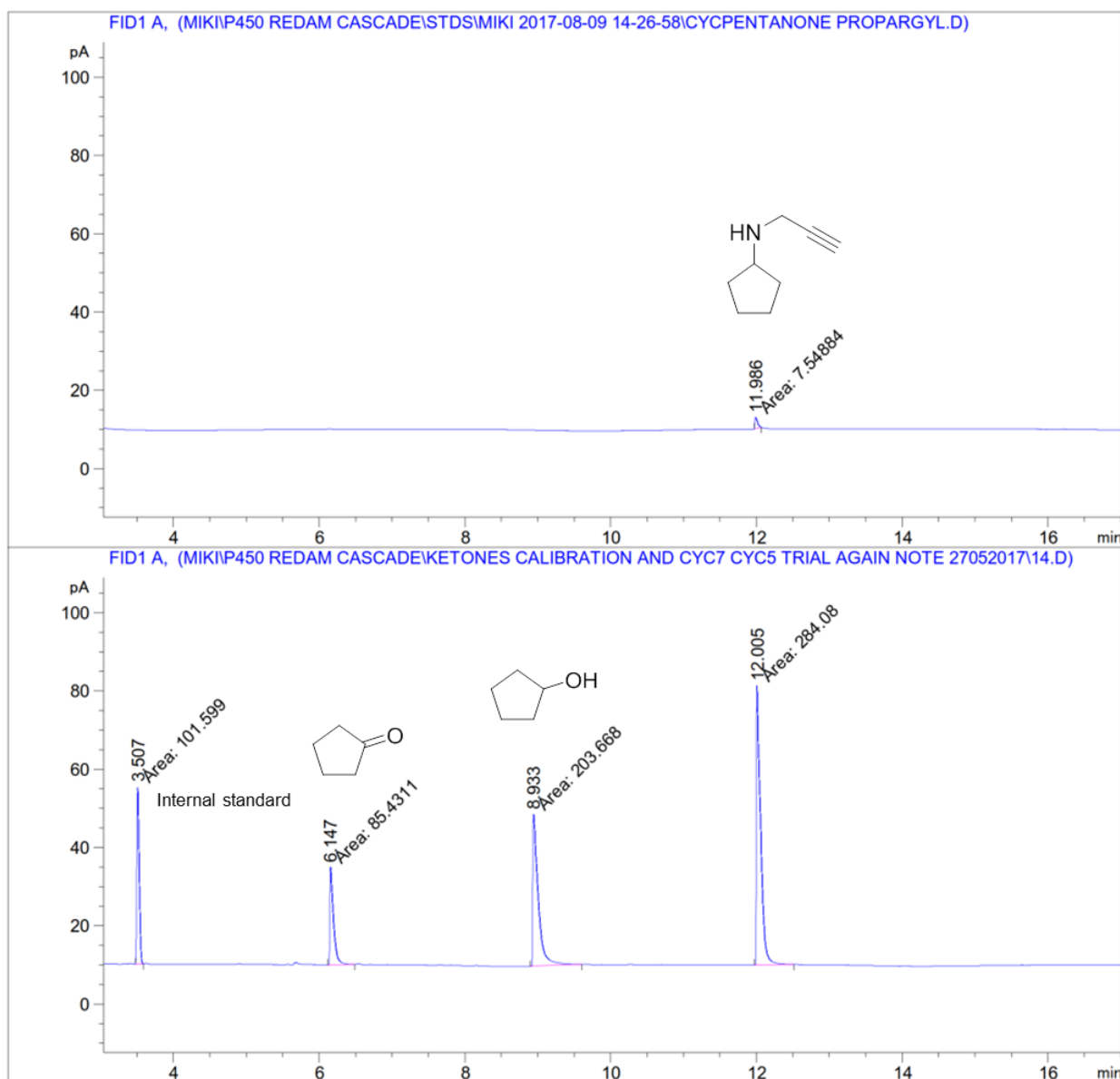


Figure S13. GC-FID chromatogram for the biocatalytic cascade for the amination of cyclopentane **4** with propargylamine **c**. Amine standard (top trace) and two-step cascade (bottom trace) are shown.

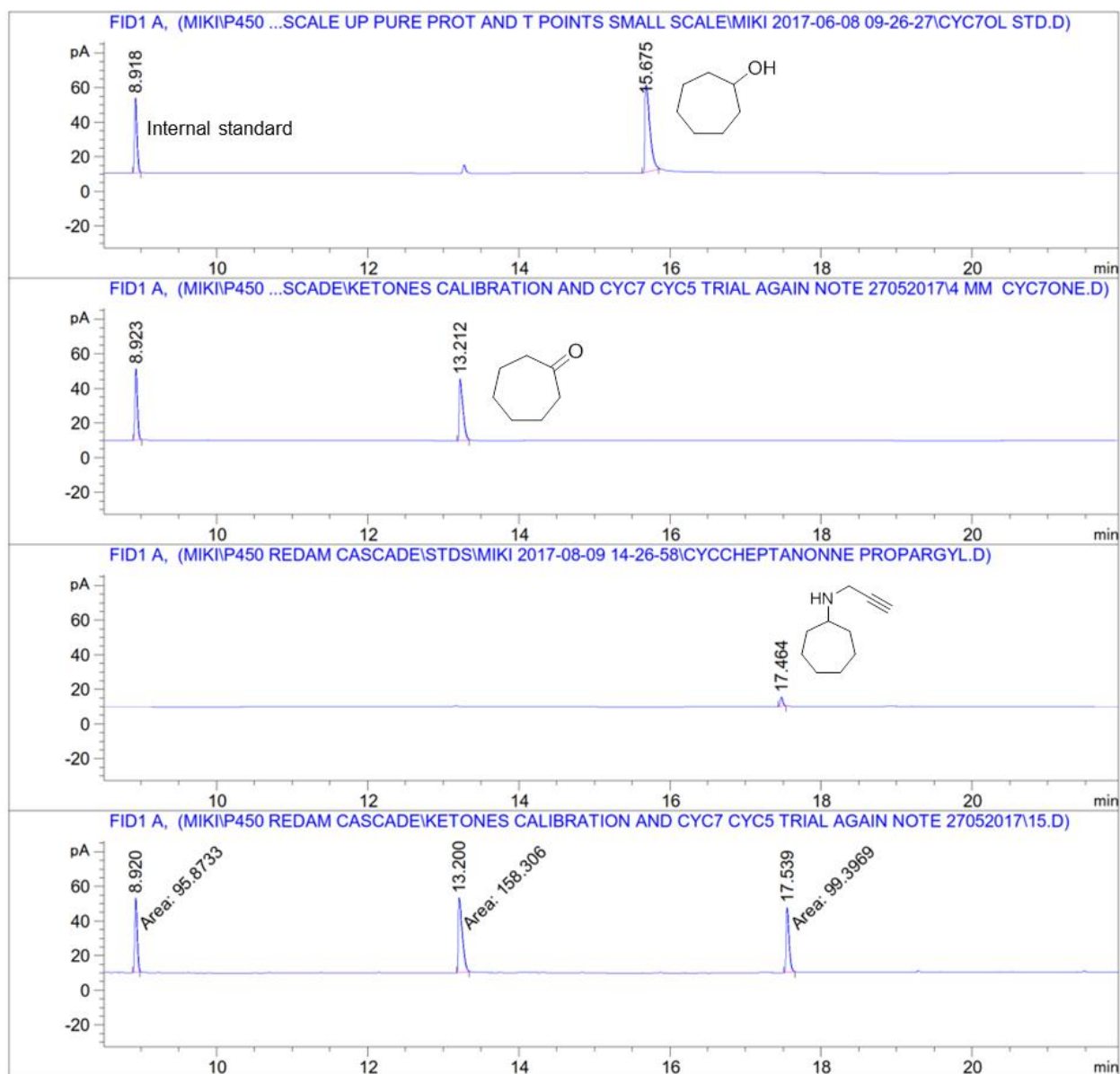


Figure S14. GC-FID chromatogram for the biocatalytic cascade for the amination of cycloheptane **6** with propargylamine **c**. Standards (top, second and third trace) and two-step cascade (fourth trace) are shown.

7.4.3.2 GC-MS Traces

7.4.3.2.1 By-product Analysis

In order to gain a better understanding on the nature of a minor product (indicated with an asterisk in the GC traces above) found in the amination of cyclohexane **1** with amines **d**, **e** and **f**, GC-MS analysis was carried out. Given the fragmentation patterns shown below, we suppose that the additional peak is an imine that is not reduced by the reductive aminase. Unfortunately, it was not possible to detect/resolve the additional peak (approximately 1%) for the reaction carried out with allylamine **e**.

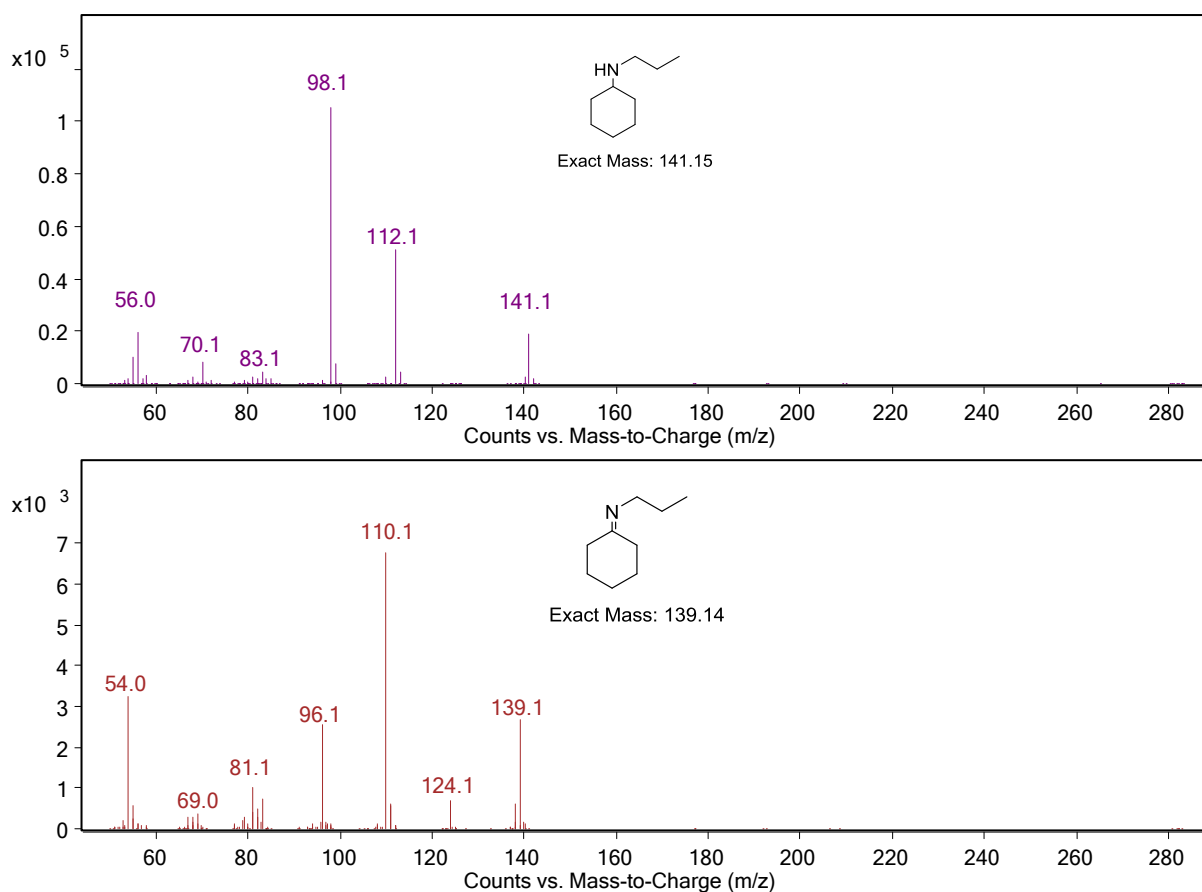


Figure S15. GC-MS analysis: amination of cyclohexane **1** with propylamine **d**. The MS trace for the amine product (top trace) is compared with the MS trace for the additional peak found in GC traces.

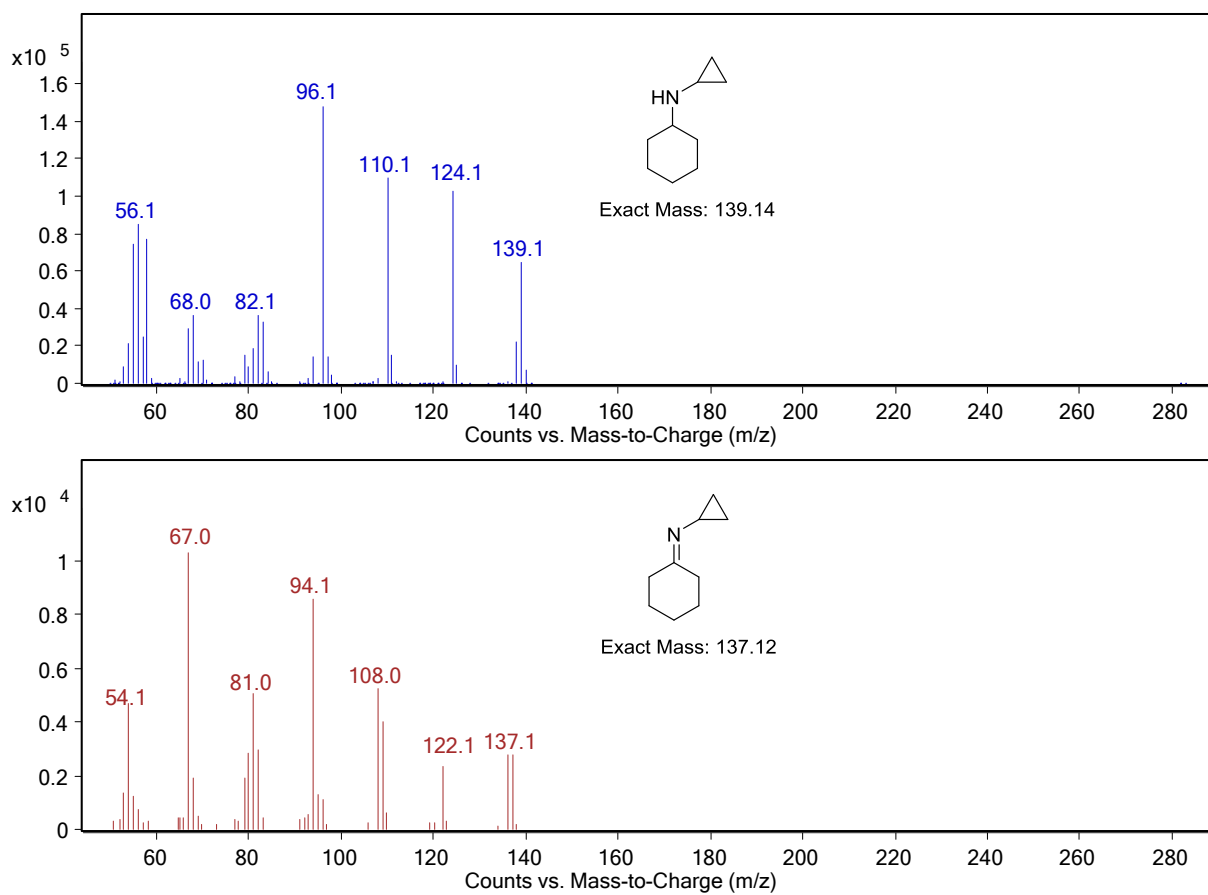


Figure S16. GC-MS analysis: amination of cyclohexane **1** with cyclopropylamine **f**. The MS trace for the amine product (top trace) is compared with the MS trace for the additional peak found in GC traces.

Finally, the amination of cyclopentane **4** with cyclopropylamine **f** was analysed by GC-MS for the identification of the by-product peak observed in biotransformation extracts. The similar fragmentation pattern suggests that the two compounds might be related, although further experiments should be done in order to identify this by-product.

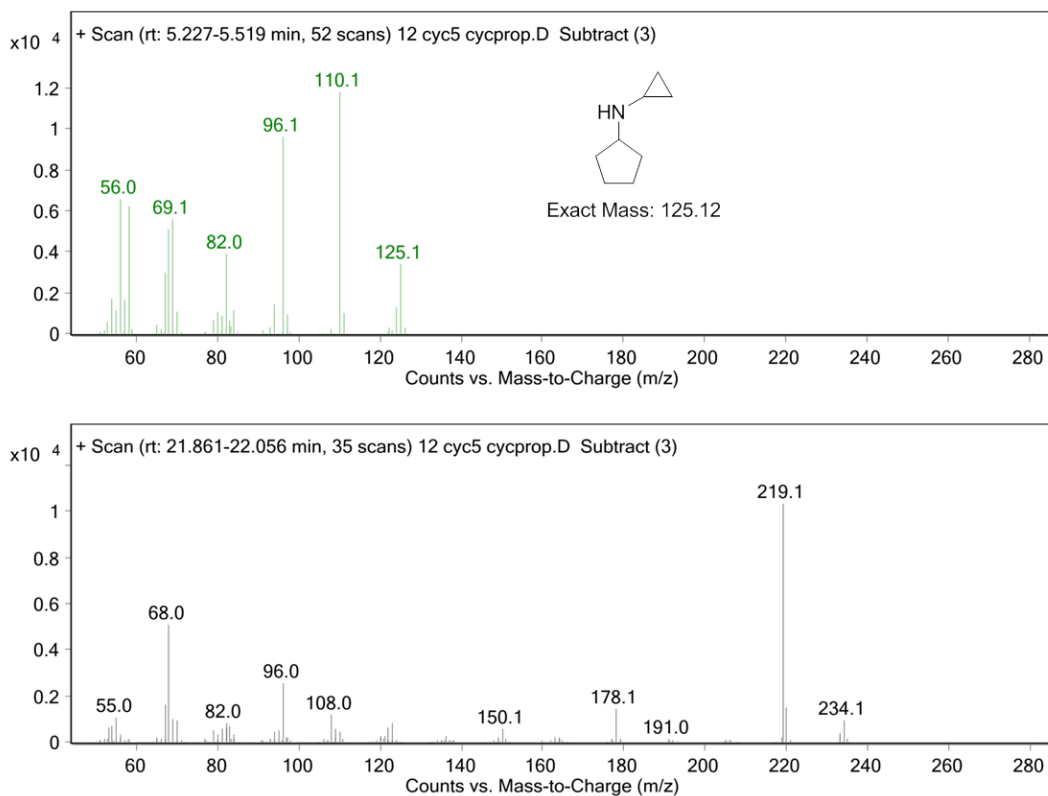


Figure S17. GC-MS analysis: Amination of cyclopentane **4** with cyclopropylamine **f**. The MS trace for the amine product (top trace) is compared with the MS trace for the additional peak found in GC traces (bottom trace).

7.4.3.2.2 Reductive Amination of Cyclooctanone with Allylamine e

GC-MS analysis suggests that *AspRedAm* is capable of catalyzing reductive amination of cyclooctanone with allylamine **e**.

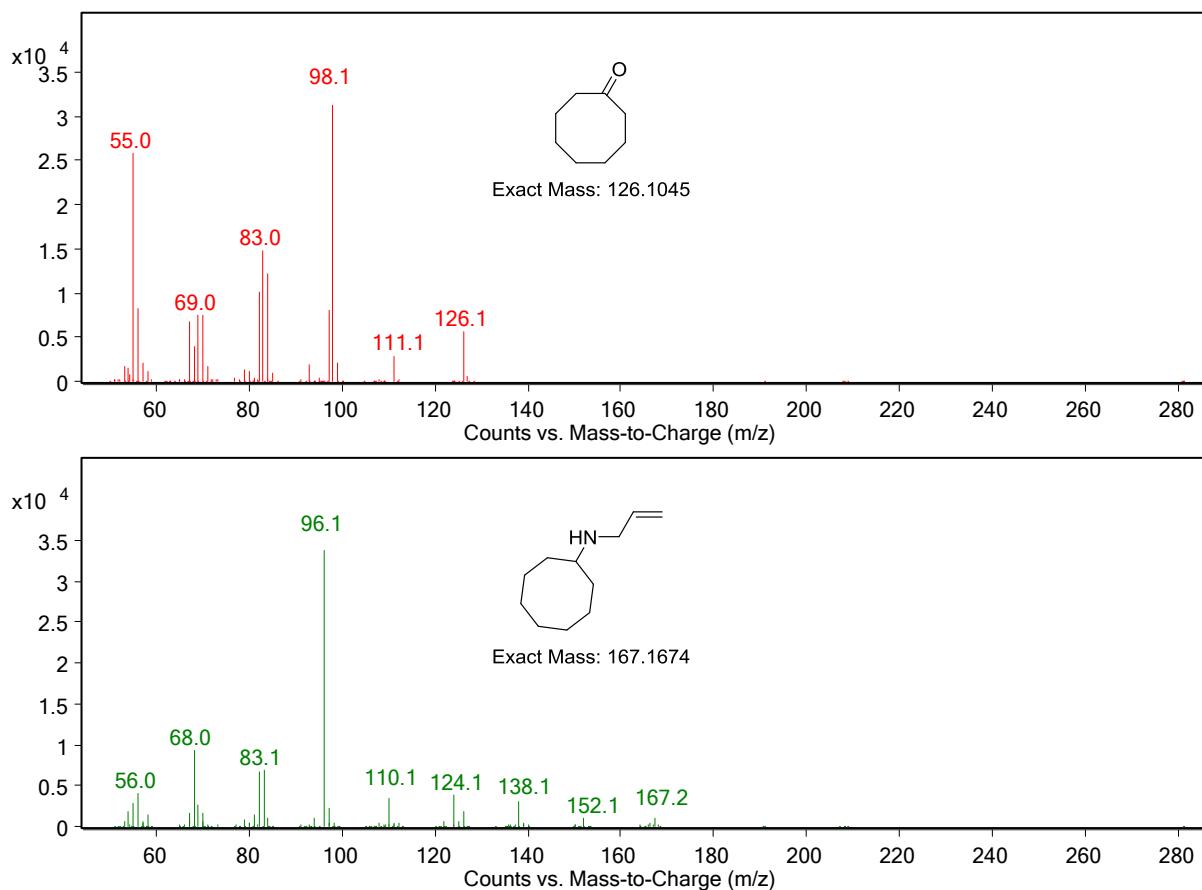


Figure S18. GC-MS analysis: amination of cyclooctanone with allylamine **e**. The MS trace for the substrate (top trace) is compared with the MS trace for the additional peak found in GC traces.

7.4.4 References

- [1] R. Omura, Tsuneo; Sato, J. Biol. Chem. 1964, 239, 2370–2378.
- [2] G. A. Aleku, S. P. France, H. Man, J. Mangas-Sanchez, S. L. Montgomery, M. Sharma, F. Leipold, S. Hussain, G. Grogan, N. J. Turner, Nat. Chem. 2017, 9, 961–969.
- [3] A. A. Penneç, F. Hollmann, M. S. Smit, D. J. Opperman, ChemCatChem 2015, 7, 236–239.
- [4] S. L. Montgomery, J. Mangas-Sanchez, M. P. Thompson, G. A. Aleku, B. Dominguez, N. J. Turner, Angew. Chemie - Int. Ed. 2017, 56, 10491–10494.

Chapter 8. Other Contributions

8.1 The Self-Sufficient P450 RhF Expressed in a Whole Cell System Selectively Catalyzes the 5-hydroxylation of Diclofenac

Jan M. Klenk^a, Bernd A. Nebel^a, Joanne L. Porter^b, Justyna K. Kulig^{c,d}, Shaneela A. Hussain^b, Sven M. Richter^a, Michele Tavanti^b, Nicholas J. Turner^b, Martin A. Hayes^d, Bernhard Hauer^a and Sabine L. Flitsch^b

Abstract

P450 monooxygenases are able to catalyse the highly regio- and stereoselective oxidations of many organic molecules. However, the scale-up of such bio-oxidations remains challenging due to the often-low activity, level of expression and stability of P450 biocatalysts. Despite these challenges they are increasingly desirable as recombinant biocatalysts, particularly for the production of drug metabolites. Diclofenac is a widely used anti-inflammatory drug that is persistent in the environment along with the 4'- and 5-hydroxy metabolites. Here we have used the self-sufficient P450 RhF (CYP116B2) from *Rhodococcus* sp. in a whole cell system to reproducibly catalyse the highly regioselective oxidation of diclofenac to 5-hydroxydiclofenac. The product is a human metabolite and as such is an important standard for environmental and toxicological analysis. Furthermore, access to significant quantities of 5-hydroxydiclofenac has allowed us to demonstrate further oxidative degradation to the toxic quinoneimine product. Our studies demonstrate the potential for gram-scale production of human drug metabolites through recombinant whole cell biocatalysis.

Acknowledgements

The research presented in this publication was largely completed prior to the beginning of the current project. The doctoral candidate screened a panel of P450s for activity towards diclofenac and carried out experiments with purified P450-RhF.

^a Institute of Technical Biochemistry, University of Stuttgart, Stuttgart, Germany

^b Manchester Institute of Biotechnology (MIB), School of Chemistry, The University of Manchester, Manchester, United Kingdom

^c Present address: Crop Science Division, Bayer AG, Monheim am Rhein, Germany

^d Cardiovascular and Metabolic Diseases DMPK, Innovative Medicines and Early Development Biotech Unit, AstraZeneca, Mölndal, Sweden

8.2 Cloning, Expression and Characterization of P450-Hal1 (CYP116B62) from *Halomonas* sp. NCIMB 172: a Self-Sufficient P450 with High Expression and Diverse Substrate Scope

Joanne L. Porter,^[a] Selina Sabatini,^[a] Jack Manning,^[a] Michele Tavanti,^[a] James L. Galman,^[a] Nicholas J. Turner^[a] and Sabine L. Flitsch^[a]

Abstract

Cytochrome P450 monooxygenases are able to catalyse a range of synthetically challenging reactions ranging from hydroxylation and demethylation to sulfoxidation and epoxidation. As such they have great potential for biocatalytic applications but are underutilized due to often-poor expression, stability and solubility in recombinant bacterial hosts. The use of self-sufficient P450s with fused heme and reductase domains has already contributed heavily to improving catalytic efficiency and simplifying an otherwise more complex multi-component system of P450 and redox partners. Herein, we present a new addition to the class VII family with the cloning, sequencing and characterization of the self-sufficient CYP116B62 Hal1 from *Halomonas* sp. NCIMB 172, the genome of which has not yet been sequenced. Hal1 exhibits high levels of expression in a recombinant *E. coli* host and can be utilized from cell lysate or used in purified form. Hal1 favours NADPH as electron donor and displays a diverse range of activities including hydroxylation, demethylation and sulfoxidation. These properties make Hal1 suitable for future biocatalytic applications or as a template for optimization through engineering.

Acknowledgements

The research presented in this publication overlapped nicely with the efforts reported in Chapter 2. The doctoral candidate assisted the set-up of experimental assays and reviewed the manuscript as a whole.

^a Manchester Institute of Biotechnology (MIB), School of Chemistry, The University of Manchester, 131 Princess Street, M1 7DN, Manchester, United Kingdom

8.3 Characterization of CYP102A25 from *Bacillus marmarensis* and CYP102A26 from *Pontibacillus halophilus*: P450 Homologues of BM3 with Preference Towards Hydroxylation of Medium-Chain Fatty Acids

Joanne L. Porter,^[a] Jack Manning,^[a] Selina Sabatini,^[a] Michele Tavanti,^[a] Nicholas J. Turner^[a] and Sabine L. Flitsch^[a]

Abstract

Cytochrome P450 monooxygenases are highly desired biocatalysts owing to their ability to catalyse a wide variety of chemically challenging C–H activation reactions. The CYP102A subfamily of enzymes are natural catalytically self-sufficient proteins consisting of a haem and FMN-FAD reductase domain fused in a single-component system. They catalyse the oxygenation of saturated and unsaturated fatty acids to produce primarily ω -1, ω -2 and ω -3 hydroxy acids. These monooxygenases have potential applications in biotechnology; however, their substrate range is still limited and there is a continued need to add diversity to this class of biocatalysts. Herein, we present the characterization of two new members of this class of enzymes, CYP102A25 (BMar) from *Bacillus marmarensis* and CYP102A26 (PHal) from *Pontibacillus halophilus*, both of which express readily in a recombinant bacterial host. BMar exhibits the highest activity toward myristic acid and shows moderate activity towards unsaturated fatty acids. PHal exhibits broader activity towards mid-chain-saturated (C14–C18) and unsaturated fatty acids. Furthermore, PHal shows good regioselectivity for the hydroxylation of myristic acid, targeting the ω -2 position for C–H activation.

Acknowledgements

Similarly to the previous publication, the doctoral candidate assisted the set-up of experimental assays and reviewed the manuscript as a whole.

^a Manchester Institute of Biotechnology (MIB), School of Chemistry, The University of Manchester, 131 Princess Street, M1 7DN, Manchester, United Kingdom