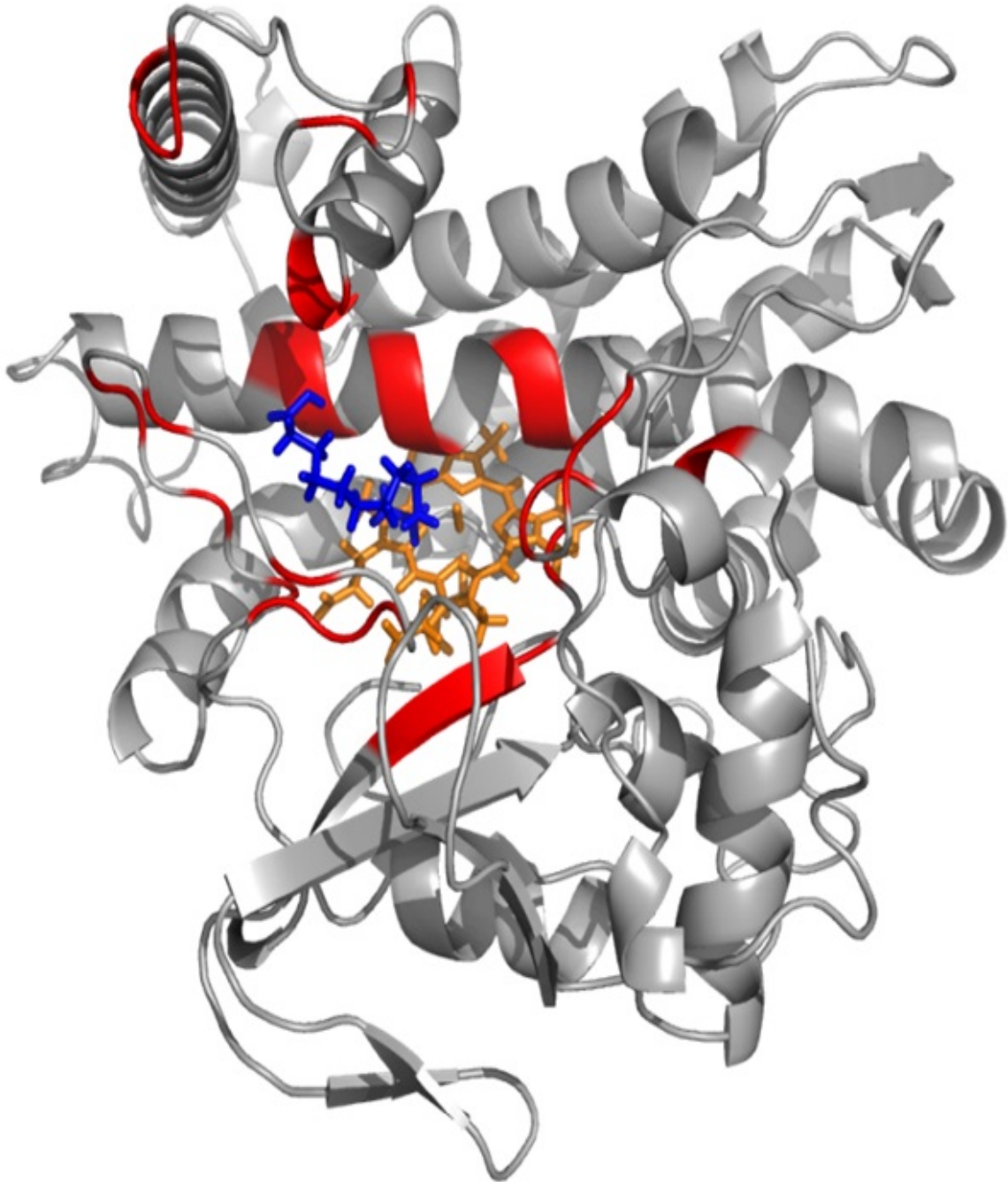


# Development of highly efficient CYP153A-catalysed terminal hydroxylation of fatty acids



**Sandra Notonier**

Institute of Technical Biochemistry, University of Stuttgart



# **Development of highly efficient CYP153A-catalysed terminal hydroxylation of fatty acids**

## **Entwicklung einer effizienten terminalen Hydroxylierung von Fettsäuren mittels CYP153A Enzymen**

An approved thesis presented to the Faculty of Energy Technology, Process Engineering and Biological Engineering of the University of Stuttgart in fulfilment of the requirements for the Degree of Doctor in Natural Sciences (Dr. rer. nat.)

submitted by

**Sandra Notonier**

From Marseille, France

Main examiner:	Prof. Dr. Bernhard Hauer
Co-examiner:	Prof. Dr. Karl-Heinrich Engesser
President of the jury:	Prof. Dr.-Ing. Ralf Takors
Thesis defense date:	10.12.2015

Institute of Technical Biochemistry, University of Stuttgart, Germany

**2016**

## **Cover**

Homology model of the heme domain of CYP153A from *Marinobacter aquaeolei*. Residues at the substrate entrance tunnel and within the active site (red) targeted for mutagenesis after docking of the substrate dodecanoic acid (blue) closed to the prosthetic heme (orange). PyMOL image from Dr. Łukasz Gricman.

## Declaration of authorship

I hereby declare that the present thesis entitled "Development of highly efficient CYP153A-catalysed terminal hydroxylation of fatty acids" is the result of my own work, that all sources used or quoted have been indicated, and that I have not used any illegitimate means. I further declare that I have not submitted this thesis for a degree in some form or another.

## Erklärung über die Eigenständigkeit der Dissertation

Ich versichere, dass ich die vorliegende Arbeit mit dem Titel „Entwicklung einer effizienten terminalen Hydroxylierung von Fettsäuren mittels CYP153A Enzymen“ selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe; aus fremden Quellen entnommene Passagen und Gedanken sind als solche kenntlich gemacht. Des Weiteren bestätige ich ausdrücklich, dass die hier vorgelegte Dissertation nicht in gleicher oder ähnlicher Form bei einer anderen Institution zur Erlangung eines akademischen Grades eingereicht wurde.

Signature / Unterschrift: Sandra Notonier

Place and date / Ort und Datum: Stuttgart, 31.01.2016



## Acknowledgments

My sincere gratitude goes to my advisor, Prof. Dr. Bernhard Hauer for his supervision, help and support during my PhD project. Thank you very much for giving me the opportunity of being part of the challenging P450 field and for the fruitful discussions and guidance. I am grateful to Dr. Bettina Nestl for her supervision, great and innovative ideas, nice discussions and permanent support and encouragement all through this research project. I am thankful to Prof. John Woodley for welcoming me in his work team at the Technical University of Denmark (DTU) to study P450-catalysed reactions through process development. I wish to thank Prof. Dr. Karl-Heinrich Engesser and Prof. Dr.-Ing. Ralf Takors for advices and help and I am deeply grateful to Dr. Andreas Schirmer, for introducing me to the P450 world and giving me the opportunity to gain experience and knowledge in a biotech-based company.

Being part of the Marie Curie P4FIFTY network was a rewarding experience. I would like to thank Dr. Marie Lundemo under the supervision of Prof. John Woodley from DTU for the nice collaboration during two years of my PhD project. I also wish to thank each member of the P4FIFTY network team: the principal investigators, Prof. Neil Bruce and my mentor Dr. Gideon Grogan from the University of York, Prof. Nicholas Turner and Prof. Sabine Flitsch from the University of Manchester, Prof. Jürgen Pleiss from the University of Stuttgart, Prof. Dick Janssen and Dr. Andy-Mark Thunnissen from the University of Groningen, Prof. Danièle Werck-Reichhart from CNRS, University of Strasbourg, Prof. Birger Lindberg Møller from the University of Copenhagen, Prof. John Woodley from DTU, Prof. Rita Bernhardt from Saarland University, Dr. Jürgen Riegler from Lonza company and Dr. Martin Hayes from AstraZenaca company. Special thanks to Margaret for the great project administration and the pleasant discussions during meetings and conferences.

I would like to thank each fellow from this P4FIFTY project for the nice atmosphere and the fruitful discussions on P450s during annual meetings, excursions, conferences, collaborations and laboratory rotation along these three years: Maria, Claudia, Anja, Slavimora, Nina, Ilona, Tina, Martina, Krutika, Marie, Flòra, Michael and Justyna. Special thanks to Łukasz for the pleasant talks during our journeys to P4FIFTY meetings and conferences and for the nice collaboration, his help and great support in the modelling, docking simulation and computational work.

I would like to thank the P450 team from the ITB for the nice atmosphere and nice discussions: previous PhD students Dr. Sumire Honda Malca and Dr. Daniel Scheps for their patience and training regarding P450s experiments, Dr. Martin Weissenborn, Sara Hoffmann and Dominique

Darimont for the nice teamwork along three years, the innovative ideas and successful exchanges regarding P450 challenges. I also would like to thank Maike Kuschel, Alexander Wittmann and Simone Prömel for their experimental support during my project. Special thanks to Sarah-Luise Lang for the successful lab work on the assay and Dr. Susanne Herter for the collaboration on the galactose oxidase assay.

I wish to thank Dr. Sandra Facey for her precious help during the second part of my PhD project, the encouragement and guidance, Dr. Janosch Klebensberger and Dr. Bernd Nebel for the help at different stages of my research. I also would like to thank Sven Richter for his precious help during process development studies at the ITB.

I wish thank to all members of the ITB, special mention to Dr. Stephan Hammer, Sara Hoffmann, Dr. Juliane Stahmer, Dr. Dennis Wetzl, Dr. Ingrid Chastinet and Dr. Per-Olof Syrén for the help, new ideas and fruitful discussions regarding work and for the great moments spent outside the ITB. I also thank Melanie A., Mihaela, Silke, Sven B., Jennifer, Patrick, Rebecca, Jörg, Julia, Silvia, Max, Christine, Thorsten, Lars, Marko, Tobias, Jan, Lisa Kontny, Nico, Lisa, Maike, Sebastian, Konrad, Sabrina, Ondrej, Philipp S., Jens, Leonie, Philipp T., Sandra V., Quy, Constantin, Matthias, Björn, Melanie M., Mina, Chris, Christian and Sven H.



## List of publications

**Part of the work presented here has already been published or is currently in preparation for publication:**

- I.** “Process limitations of a whole-cell P450 catalyzed reaction using a CYP153A-CPR fusion construct expressed in *Escherichia coli*.” M. T. Lundemo§, S. Notonier§, G. Striedner, B. Hauer & J. M. Woodley\*. *Applied Microbiology and Biotechnology* 100, 1197–1208 (2016).  
§ These authors contributed equally to this work.

My contribution to the work presented in this paper is described in the *Chapter 1* of the dissertation.

- II.** “Semi-rational protein engineering of CYP153A<sub>M.aq.</sub>-CPR for efficient terminal hydroxylation of short to long chain fatty acids.” S. Notonier, Ł. Gricman, J. Pleiss, B. Hauer\*. *Manuscript in preparation*

The manuscript in preparation refers to my work described in the *Chapter 2* of the dissertation.

- III.** “High-throughput whole-cell screening assay for terminal hydroxylation of fatty acids and alkanes by P450s.” M. J. Weissenborn,§ S. Notonier,§ S.-L. Lang, K. B. Otte, S. Herter, N. J. Turner, S. L. Flitsch and B. Hauer\*. *Communication in preparation*  
§ These authors contributed equally to this work.

Part of the work presented in this communication was based on my results described in the *Chapter 3* of the dissertation.

**Publication which has not been included in this work:**

- IV.** “The Impact of Linker Length on P450 Fusion Constructs: Activity, Stability and Coupling.” S. M. Hoffmann,‡ M. J. Weissenborn,‡ Ł. Gricman, S. Notonier, J. Pleiss, B. Hauer\*. *Submitted to ACS Chemical Biology* (2015)  
‡ These authors contributed equally to this work.

The work presented in this PhD thesis was carried out between July 2012 and July 2015, supervised by Prof. Dr. Bernhard Hauer, at the Institute of Technical Biochemistry, University of Stuttgart, Germany.

The research leading to these results has received funding from the People Programme (P4FIFTY, Marie Curie Actions) of the European Union's 7<sup>th</sup> Framework Programme (FP7/2007-2013) under REA Grant Agreement 289217.



## Table of contents

<b>Declaration of authorship</b> .....	<b>1</b>
<b>Acknowledgments</b> .....	<b>3</b>
<b>List of publications</b> .....	<b>5</b>
<b>Table of contents</b> .....	<b>7</b>
<b>Abbreviations</b> .....	<b>11</b>
<b>Abstract</b> .....	<b>15</b>
<b>Zusammenfassung</b> .....	<b>17</b>
<b>Introduction</b> .....	<b>21</b>
<b>1. Biotechnology: the green alternative to fossil energies</b> .....	<b>21</b>
1.1. Enzymes at the heart of a sustainable environment.....	21
1.2. Renewable energies .....	22
1.2.1. Biofuels.....	22
1.2.2. Bio-based chemicals, flavours and fragrance market.....	22
1.2.3. Terminal hydroxylated fatty acids: synthesis and application .....	25
<b>2. Cytochrome P450 monooxygenase enzymes</b> .....	<b>26</b>
2.1. Nomenclature .....	26
2.2. Reactions and mechanism .....	27
2.2.1. Variety of reactions.....	27
2.2.2. Mechanism.....	27
2.3. Structure and diversity .....	29
2.4. Fusion proteins .....	30
2.5. Microbial and fungal P450s.....	31
2.6. CYP153A family and generation of a fusion construct .....	32
<b>3. Cytochrome P450s as industrial tools</b> .....	<b>34</b>
<b>Aim of the project</b> .....	<b>37</b>
<b>Material and Method</b> .....	<b>39</b>
<b>1. Chemicals, media and bacterial strains</b> .....	<b>39</b>
<b>2. Plasmids and constructs</b> .....	<b>40</b>
2.1. Plasmids and strains .....	40
2.2. Mutant libraries .....	40
2.3. Quick Change protocol .....	43
<b>3. Cell cultivation, protein expression and whole resting cell preparation</b> .....	<b>43</b>
3.1. CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> at small scale .....	43
3.2. CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> in shaking flasks .....	44

3.3. CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> in 1 L fermenter .....	44
3.4. Biocatalytic investigations with resting cells .....	44
3.4.1. 24 deep-well plates .....	44
3.4.2. 100 mL Erlenmeyer flasks .....	45
3.4.3. 1 L stirring tank reactor .....	45
3.5. Galactose oxidase expression.....	45
<b>4. Protein purification.....</b>	<b>45</b>
4.1. Purification of CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> .....	45
4.2. Purification of the galactose oxidase.....	46
4.3. <i>In vitro</i> performance and kinetics parameters .....	46
<b>5. Substrate and product inhibition studies .....</b>	<b>47</b>
5.1. <i>In vitro</i> reactions.....	47
5.2. <i>In vivo</i> biotransformations in shaking flasks .....	47
5.3. <i>In vivo</i> reactions in bioreactors .....	47
<b>6. Cofactor limitations investigations .....</b>	<b>48</b>
<b>7. Transport limitation studies.....</b>	<b>48</b>
<b>8. Determination of P450 concentration .....</b>	<b>48</b>
<b>9. Cell dry weight .....</b>	<b>49</b>
<b>10. Extraction procedure and GC-FID method.....</b>	<b>49</b>
<b>11. Galactose oxidase screening assay .....</b>	<b>49</b>
<b>Results.....</b>	<b>51</b>
<i>Chapter 1: Identification of bottlenecks in CPR2<sub>mut</sub>-catalysed reactions: how to address the limitations and improve whole-cell biocatalysis.....</i>	<i>51</i>
<b>1. Substrate inhibition .....</b>	<b>51</b>
1.1. Enzymatic reactions.....	51
1.2. Whole resting cells in shaking flasks.....	51
1.3. <i>In vivo</i> biotransformations in fermenter .....	52
<b>2. Product inhibition .....</b>	<b>53</b>
2.1. <i>In vitro</i> reactions.....	53
2.2. <i>In vivo</i> reactions .....	54
<b>3. Cofactor regeneration limited.....</b>	<b>55</b>
<b>4. Transport limitation .....</b>	<b>56</b>
<b>5. P450 instability.....</b>	<b>57</b>
<b>6. Feeding strategies.....</b>	<b>57</b>
<b>7. Standard conditions required to improve a process .....</b>	<b>58</b>
7.1. Protein expression rate.....	59

7.2. Bacterial strain .....	60
7.3. Biotransformation efficiency.....	60
<b>Chapter 2: Generation of mutant libraries and screening for improved activity.....</b>	<b>63</b>
<b>1. Library generation.....</b>	<b>63</b>
1.1. Homology model.....	63
1.2. Positions and residues selected for mutations .....	64
1.3. Substrate entrance tunnel .....	64
1.3.1. Degenerated codon.....	64
1.3.2. Selection through <i>in vivo</i> tests .....	66
<b>2. <i>In vitro</i> activity tests.....</b>	<b>66</b>
<b>3. Improved variants .....</b>	<b>68</b>
3.1. Characterization <i>in vitro</i> .....	68
3.1.1. <i>In vitro</i> activity comparison.....	68
3.1.2. Substrate/product inhibition.....	69
3.1.3. Performance over time.....	71
3.2. <i>In vivo</i> validation.....	71
<b>Chapter 3: Establishment of a whole-cell screening assay for identification of improved</b>	
<b>    P450-activity .....</b>	<b>73</b>
<b>1. Galactose oxidase assay .....</b>	<b>73</b>
<b>2. Establishment of the assay .....</b>	<b>74</b>
2.1. Demonstration in a buffer system.....	74
2.2. Deactivation of the supernatant.....	75
2.3. Detection limit .....	76
2.3.1. Increased concentrations of $\omega$ -OHC <sub>12:0</sub> .....	76
2.3.2. Addition of increased concentrations of C <sub>12:0</sub> .....	76
<b>3. Selectivity of the galactose oxidase .....</b>	<b>77</b>
<b>4. Screening of mutant libraries.....</b>	<b>78</b>
4.1. GOase vs. GC-FID .....	78
4.2. Protein content correlation.....	79
<b>Discussion.....</b>	<b>83</b>
<b>    Chapter 1: Identification of bottlenecks in CPR2<sub>mut</sub>-catalysed reactions: how to address</b>	
<b>        the limitations and improve whole-cell biocatalysis.....</b>	<b>83</b>
<b>    Chapter 2: Generation of mutant libraries and screening for improved activity.....</b>	<b>89</b>
<b>    Chapter 3: Establishment of a whole-cell screening assay for identification of improved</b>	
<b>        P450-activity .....</b>	<b>95</b>
<b>Conclusion and outlook.....</b>	<b>99</b>

<b>References</b> .....	<b>103</b>
<b>Appendix</b> .....	<b>115</b>
A1. Gene CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> .....	115
A2. Mutant library .....	116
A3. Cofactor analogue .....	118
A4. Solid feeding strategy.....	118
A5. Control reactions (GOase assay).....	119
A6. GOase assay vs. GC-FID .....	119
<b>Curriculum vitae</b> .....	<b>123</b>

## Abbreviations

3xGGS	three times “glycine-glycine-serine” in the linker sequence
(Gly-Gly-Ser) <sub>n</sub>	number of “glycine-glycine-serine” in the linker sequence
3-HPA	3-hydroxypropionic acid
8ZY-4LAC	auto-induction medium for galactose oxidase expression
× g	gravitational acceleration
°C	degrees Celsius
μL	microliter
μm	micrometer
μM	micromolar
α,ω-DC <sub>12:0</sub>	1,12-dicarboxylic dodecanoic acid
ω-OHC <sub>12:0</sub>	12-hydroxydodecanoic acid
ω-1-OHC <sub>12:0</sub>	11-hydroxydodecanoic acid
ω-2-OHC <sub>12:0</sub>	10-hydroxydodecanoic acid
ω-3-OHC <sub>12:0</sub>	9-hydroxydodecanoic acid
ω-OHFA	terminal hydroxy fatty acid
ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)
AlkB	alkane-1-monooxygenase
AlkM	long-chain alkane oxygenase from <i>Acinetobacter species</i>
<i>B. megaterium</i>	<i>Bacillus megaterium</i>
BDO	1,4-butanediol
BEHP	bis(2-ethylhexyl)phthalate
bp	base pair
BMO	butane monooxygenase
BSTFA	<i>N,O</i> -bis(trimethylsilyl)trifluoroacetamide containing 1 % trimethylchlorosilane
C <sub>8:0</sub>	octanoic acid
C <sub>12:0</sub>	dodecanoic acid
C <sub>14:0</sub>	tetradecanoic acid
C <sub>16:0</sub>	hexadecanoic acid
cdw	cell dry weight
cdm	cell dry mass
CO	carbon monoxide gas
CPR	cytochrome P450 reductase

CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub>	CYP153A monooxygenase from <i>Marinobacter aquaeolei</i> VT8 fused to reductase domain of CYP102A1 from <i>Bacillus megaterium</i>
CPR2 <sub>mut</sub>	CYP153A monooxygenase from <i>Marinobacter aquaeolei</i> VT8 fused to reductase domain of CYP102A1 from <i>Bacillus megaterium</i> , mutant G307A, 3xGGS
cww	cell wet weight
CYP	cytochrome P450 monooxygenase
CYP153A <sub>M.aq.</sub>	CYP153A monooxygenase from <i>Marinobacter aquaeolei</i> VT8
DCA	dicarboxylic acid
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxyribonucleotide mix
<i>E. coli</i>	<i>Escherichia coli</i>
FAD	flavin adenine dinucleotide
FeS	ferredoxin
FMN	flavin mononucleotide
g	gram
g <sub>cdw</sub>	pro gram cell dry weight
g <sub>cww</sub>	pro gram cell wet weight
GC-FID	gas chromatography coupled to flame ionization detector
G6P	glucose-6-phosphate
G6PDH	glucose-6-phosphate dehydrogenase
GOase	galactose oxidase
GOase <sub>M3-5</sub>	galactose oxidase mutant M3-5
h	hour
HPLC	High-performance liquid chromatography
HRP	horseradish peroxidase
IC <sub>50</sub>	concentration of inhibitor
IPTG	isopropyl- $\beta$ -D-thiogalactopyranoside
kbp	kilobase pair
$k_{cat}$	catalytic constant: turnover number
kDa	kilo Dalton
$K_m$	kinetics constant: affinity of the enzyme for a substrate
$K_s$	constant of substrate inhibition
L	liter



LadA	long-chain alkane oxygenase from <i>Geobacillus thermodentrificans</i>
LB	lysogeny (Luria-Bertani) broth
<i>M. aquaeolei</i>	<i>Marinobacter aquaeolei</i>
MD	molecular dynamics
mg	milligram
min	minute
ml	millilitre
mM	millimolar
MMO	methane monooxygenase
NAD(P)/NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
NACMH	nicotinamide-3,4-chloro-methoxy-phenyl-hydride
ng	nanogram
nm	nanometer
OD <sub>600</sub>	optical density measured at 600 nm
OD <sub>420</sub>	optical density measured at 420 nm
OHFA	hydroxylated fatty acid
P450	cytochrome P450 monooxygenase
P450 <sub>BM3</sub>	CYP102A1 from <i>Bacillus megaterium</i>
P450 <sub>cam</sub>	P450 from <i>Pseudomonas putida</i>
PBS	phosphate-buffered saline
PDO	1,3-propanediol
PES	polyethersulfone
PLA	polylactic acid
PHA	polyhydroxyalkanoate
PHB	polyhydroxybutyrate
PMO	propane monooxygenase
pmol	picomol
<i>P. putida</i>	<i>Pseudomonas putida</i>
PCR	polymerase chain reaction
QM	quantum mechanical
P450 <sub>Rhf</sub>	P450 from <i>Rhodococcus specie</i>
ROs	reactive oxygen species
rpm	revolution per minute
sec	second
<i>sp.</i>	specie
TB	terrific broth

$V_{\max}$	catalytic constant, maximal velocity value
vs.	versus
v/v	volume per volume
vvm	volume per volume per minute

## Abstract

Terminally hydroxylated fatty acids ( $\omega$ -OHFA) are of great interest to industry: in the area of high end polymers, in fine chemicals, in the cosmetic and fragrance industry. The activation and oxidation of C-H bonds, is however still chemically challenging. It requires rough conditions, high pressure, high temperature and the processes still suffer from poor selectivity. Alternatively, the functionalization of carbon atoms can be achieved by using versatile cytochrome P450 monooxygenases with high regio- and stereoselectivity (P450s or CYPs). These heme containing proteins can be found in many organisms, they are involved in a diverse range of reactions and they are able to catalyse the conversion of a large panel of substrates. CYP153A from *Marinobacter aquaeolei* (CYP153A<sub>M.aq.</sub>) constitutes a promising catalyst for the oxidation of non-activated carbon atoms due to its high regioselectivity in the hydroxylation of different small to medium chain alkanes, fatty acids and primary alcohols as well as its efficient expression in *Escherichia. coli* at high yields. For bacterial whole-cell applications, the enzyme was further engineered as a protein chimera by fusing the reductase domain of P450<sub>BM3</sub>, from *Bacillus megaterium*, to the heme domain of CYP153A<sub>M.aq.</sub> (CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub>). The mutant G307A (position located in the binding pocket) was further identified as improved variant towards medium chain-length fatty acids regarding activity, regioselectivity and coupling efficiency.

To perform efficient whole-cell biotransformations and therefore to increase the yield of bioconversion of fatty acids into  $\omega$ -OHFA, the system requires the identification of present bottlenecks. The approaches to characterize and to optimise the up-scaling process involved testing different feeding strategies, evaluation of substrate/product inhibition, transport limitations estimation, cofactor availability evaluation and biocatalyst stability investigations. Such limitations often associated to the P450 biocatalysts and/or related to whole-cell biotransformations were evaluated in this study, using the fusion constructs CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> towards the model substrate dodecanoic acid (C<sub>12:0</sub>) *in vitro* and *in vivo*.

Substrate and product inhibition were described as challenges that need to be addressed; inhibitory effects have been observed on the reaction rate at the CYP level and within whole cells. The application of a solid substrate strategy was shown to be a promising option to improve the process. The approach of a feeding of the solid C<sub>12:0</sub> in 1 L fermenter enabled a continuous dissolving of the substrate along a reaction course with whole-cells leading to an increase of the rate of reaction and an increase in the yield of bioconversion of C<sub>12:0</sub> into 12-hydroxydodecanoic acid ( $\omega$ -OHC<sub>12:0</sub>). The application of the solid substrate showed 2.8-fold

improvement in comparison to a classic batch (substrate dissolved with co-solvent). Despite a well-developed whole-cell P450 catalyst, the low activity and the poor stability of the artificial fusion construct are the main identified limitations to reach sufficient biocatalyst yield and space-time yield to achieve an economically feasible process.

Strategies to overcome existing bottlenecks regarding low activity involved the generation of mutant libraries to screen and to characterize improved versions of the current chimera biocatalyst. Positions located in close proximity to the active site of the protein and found at the substrate entrance channel were selected. Amino acid residues were chosen after creation of a homology model based on sequence alignments and docking studies. Site-directed mutagenesis was applied to the positions in the active site and the use of a degenerated codon was the preferred method to generate diversity for the positions part of the substrate entrance tunnel. Improved enzyme variants were identified after a whole-cell screening procedure (for the positions located in the substrate access tunnel) and were further deeply characterized *in vitro*. Indeed, different chain lengths of fatty acids were tested (C<sub>8:0</sub>, C<sub>12:0</sub> and C<sub>16:0</sub>) and the rate of reaction and the efficiency of biotransformation were evaluated compared to the wild type enzyme. The final validation of the improved variants was estimated *in vivo* for an industrialization perspective. Consecutive mutagenesis steps and screening of variants led to the creation of one double mutant containing one mutation within the binding pocket and one at the substrate entrance channel. The improved variant showed enhanced activity towards short to long chain-length fatty acids (C<sub>8:0</sub>, C<sub>12:0</sub> and C<sub>16:0</sub>) *in vitro*, with 14-fold, 2-fold and 4.8-fold improvement, respectively, compared to the wild type activity. A greater resistance to higher concentrations of substrate/product was observed with this variant as well.

In parallel to the mutant libraries generation, a whole-cell colorimetric assay was also established, employing the galactose oxidase enzyme, able to oxidize  $\omega$ -OHC<sub>12:0</sub> (product from CYP153A<sub>M.aq.</sub>) and to generate hydrogen peroxide. Coupling a horseradish peroxidase and ABTS (oxidation of the dye, monitored at 420 nm) demonstrates the activity of the galactose oxidase in the presence of the  $\omega$ -OHFA and therefore displays the activity of P450 variants. The assay was established in a buffer system with purified enzymes before the demonstration of the principle with the supernatant from whole-cell based-reactions. In order to validate the successful concept, whole resting cells biocatalysis were performed with selected mutants. The identification of improved variants via the galactose oxidase assay was correlated to the P450 content in whole-cells for an accurate comparison based on the biocatalyst concentration and not only based on the cell density. The results obtained from whole-cell based-reactions with such assay were confirmed by *in vitro* tests performed with a concentration of enzyme normalized beforehand.

## Zusammenfassung

Terminal hydroxylierte Fettsäuren ( $\omega$ -OHFA) sind für die Industrie von großem Interesse: im Bereich der high-end Polymere, in der Feinchemie, in der Kosmetik- und Duftstoffindustrie. Die Aktivierung und Oxidation von C-H Bindungen gilt noch immer als chemisch herausfordernd. Hierfür werden harsche Reaktionsbedingungen, hohe Drücke sowie hohe Temperaturen benötigt und die Verfahren sind noch immer durch niedrige Selektivitäten beeinträchtigt. Alternativ kann die Funktionalisierung von Kohlenstoffatomen durch die Verwendung von vielseitigen Cytochrom P450 Monooxygenasen mit hohen Regio- und Stereoselektivitäten erreicht werden. Diese Häm-enthaltenden Proteine können in vielen Organismen gefunden werden, sie sind an einer diversen Reihe von Reaktionen beteiligt und sie sind fähig, eine Vielzahl von Substraten umzusetzen. Sowohl aufgrund der hohen Regioselektivität bei der Hydroxylierung von verschiedenen kurz- bis mittelkettigen Alkanen, Fettsäuren und primären Alkoholen, als auch der effizienten Expression in *E.coli* in hohen Titern, stellt CYP153A aus *Marinobacter aquaeolei* (CYP153A<sub>M.aq.</sub>) einen vielversprechenden Katalysator für die Oxidation von nicht-aktivierten Kohlenstoffatomen dar. Durch die Fusion der Reduktasedomäne von P450<sub>BM3</sub> aus *Bacillus megaterium* an die Häm-domäne von CYP153A<sub>M.aq.</sub> (CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub>) wurde das Enzym für die bakterielle Ganzzell Katalyse zu einem Protein-Chimär weiterentwickelt. Bezüglich Aktivität, Regioselektivität und Couplingeffizienz wurde die Mutante G307A (in der Bindetasche befindlich) als verbesserte Variante gegenüber mittelkettigen Fettsäuren identifiziert.

Um effiziente Ganzzell Biotransformationen durchführen zu können und somit die Ausbeute der Umsetzung von Fettsäuren zu  $\omega$ -OHFA zu steigern, ist es notwendig vorhandene Engpässe des Systems zu identifizieren. Zu den Ansätzen, die verfolgt wurden, um den Up-Scaling Prozess zu charakterisieren und zu optimieren, gehörten verschiedene Fütterungsstrategien, die Evaluation von Substrat/Produkt-Inhibitionen, die Abschätzung von Transportlimitationen, die Überprüfung der Kofaktor-Verfügbarkeit sowie die Untersuchung der Stabilität des Biokatalysators. Solche Engpässe, die häufig mit dem P450 Biokatalysator und/oder Ganzzell Biotransformationen verbunden werden, wurden in dieser Studie evaluiert, indem die Fusionskonstrukte CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> mit dem Modellsubstrat Dodekansäure (C<sub>12:0</sub>) *in vitro* und *in vivo* verwendet worden sind.

Substrat- und Produktinhibition wurden als Herausforderungen beschrieben, bei denen Handlungsbedarf besteht; inhibitorische Wirkungen wurden auf die Reaktionsraten der Biokatalysatoren selbst und innerhalb ganzer Zellen beobachtet. Es konnte gezeigt werden, dass die Verwendung des Substrates als Feststoff eine vielversprechende Option zur Verbesserung

des Verfahrens darstellt. Der Ansatz der Fütterung von fester  $C_{12:0}$  im 1 L Fermenter ermöglichte die kontinuierliche Lösung des Substrats während eines Reaktionsverlaufs mit ganzen Zellen, was zu einer Erhöhung der Reaktionsrate und des Umsatzes bei der Biotransformation von  $C_{12:0}$  zu 12-hydroxydodekansäure ( $\omega$ -OHC $_{12:0}$ ) führte. Die Verwendung des Substrats als Feststoff ergab eine 2.8-fache Verbesserung im Vergleich zu einem klassischen Ansatz, in dem das Substrat vorher in Co-Lösungsmittel gelöst wurde. Trotz des gut entwickelten Ganzzell P450 Katalysators erweisen sich die geringe Aktivität und die unzureichende Stabilität des artifiziellen Fusionskonstruktes als die Hauptlimitationen zur Erreichung einer ausreichenden Biokatalysator-Menge und Raum-Zeit-Ausbeute, um ein wirtschaftlich durchführbares Verfahren zu ermöglichen.

Einige der Strategien zur Eliminierung vorhandener Aktivitäts-Engpässe beinhaltete die Erstellung von Mutantenbibliotheken, um optimierte Varianten des momentan vorhandenen chimären Biokatalysators zu screenen und zu charakterisieren. Es wurden Positionen nahe des aktiven Zentrums und am Substrateingangskanal ausgewählt. Nach der Erstellung eines Homologiemodells, basierend auf Sequenzalignments und Dockingstudien wurden Aminosäurereste ausgewählt. Bei den Positionen in der aktiven Tasche wurde ortsgerichtete Mutagenese angewandt, wobei die Verwendung von degenerierten Codons die präferierte Methode war, um Diversität an den Positionen des Substrateingangs zu generieren. Verbesserte Enzymvarianten wurden nach einem Ganzzell-Screening identifiziert (für die Positionen im Substrateingang) und wurden im Weiteren *in vitro* umfassend charakterisiert. Verschieden lange Fettsäuren ( $C_{8:0}$ ,  $C_{12:0}$  und  $C_{16:0}$ ), die Reaktionsraten und die Effizienz der Biotransformationen wurden im Vergleich zum Wildtyp evaluiert. Eine abschließende Validierung der verbesserten Varianten wurde im Hinblick auf eine industrielle Anwendung *in vivo* durchgeführt. Weitere Mutageneserunden und das Screening der dadurch erzeugten Varianten führte zu einer Doppelmutante, die eine Mutation innerhalb der Bindetasche und die zweite am Substrateingang beinhaltet. Die verbesserte Variante zeigte *in vitro* gesteigerte Aktivität gegenüber kurz- zu langkettigen Fettsäuren ( $C_{8:0}$ ,  $C_{12:0}$  und  $C_{16:0}$ ) mit einer 14-, 2-beziehungsweise 4.8-fachen Steigerung im Vergleich zur Aktivität des Wildtyps. Diese verbesserte Variante zeigte erhöhte Aktivität gegenüber kurz- bis langkettigen Fettsäuren *in vitro* und eine erhöhte Resistenz gegenüber höheren Substrat- und Produktkonzentrationen.

Parallel zur Generierung der Mutantenbibliothek wurde unter der Verwendung von Galaktoseoxidase, die  $\omega$ -OHC $_{12:0}$  (Produkt von CYP153A $_{M.aq.}$ ) oxidieren kann und dabei Wasserstoffperoxid bildet, ein kolorimetrischer Ganzzellassay entwickelt. Indem eine Meerrettichperoxidase und das Färbemittel ABTS gemeinsam verwendet werden, kann die

Oxidation des Farbstoffes bei einer Wellenlänge von 420 nm verfolgt werden. Dadurch kann die Aktivität der Galaktoseoxidase im Vorhandensein von  $\omega$ -OHFA dargestellt werden, was der Aktivität der P450 Varianten entspricht. Der Assay wurde zunächst für ein Puffersystem mit aufgereinigtem Enzym entwickelt und anschließend wurde gezeigt, dass anstatt des aufgereinigten Enzyms der Überstand von Ganzzell basierenden Reaktionen verwendet werden kann. Um das erfolgreiche Assay-Konzept zu bestätigen wurden Ganzzell Umsätze mit ruhenden Zellen, die ausgewählte Mutanten enthielten, durchgeführt. Die Identifikation verbesserter Varianten durch den Galaktoseoxidase Assay wurde zum P450 Gehalt in den ganzen Zellen korreliert, um einen richtigen Vergleich zu ermöglichen, der auf der Konzentration des Biokatalysators und nicht nur auf der Zelldichte basiert. Die Ergebnisse der Ganzzell-basierten Reaktionen mit diesem Assay wurden durch *in vitro* Untersuchungen bestätigt, die mit einer Enzymmenge durchgeführt wurde, die der zuvor normalisierten entsprach.





## Introduction

### 1. Biotechnology: the green alternative to fossil energies

#### 1.1. Enzymes at the heart of a sustainable environment

*“Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs”.*<sup>[1]</sup> As defined by the World Commission on Environment and Development's, the concept of sustainability include an economic development, a social development and the protection of the environment.<sup>[2,3]</sup> It aims at improving our management of renewable energies since the decline of fossil energies, also responsible for the climate change, global warming concerns and degradation of our natural ecosystems. Compared to conventional production, sustainable systems would require less energy, less material and less emission of greenhouse gas and other pollutants.<sup>[4]</sup> As clean alternatives e.g. wind power, solar energy, hydropower, biofuels, heat-biomass, have emerged and have been developed since 2004 but only a few percentages are at the moment operated compared to the consumption of fossil energies.<sup>[5]</sup>

Biotechnology and bioremediation approaches using enzymes are presently extensively developed to replace these overexploited fossil energies and to provide a greener environment.<sup>[6]</sup> Enzymes are powerful biocatalysts necessary in all living systems that have the ability to catalyse important biochemical reactions and often possessing a large substrate spectrum.<sup>[7]</sup> Found in every organism, some enzymes can bear extreme temperatures, extreme pH conditions or the presence of organic solvents that make these proteins perfect candidates for industrial development.<sup>[8]</sup> Bioremediation part of the emerging biotechnologies can be defined as a strategy to depollute the environment including decontamination of soil by microorganisms, plants or animal, and water waste management.<sup>[9,10]</sup> Numerous of industrial pollutants in air, water or soils can be naturally removed by enzymes.<sup>[11]</sup> For example, the ability of biocatalysts isolated from plants to neutralize class of explosives, remaining in soils after decade of military activities, has been demonstrated.<sup>[12]</sup> Toluene dioxygenases isolated from *Pseudomonas putida* are involved in the first step for the degradation of toluene.<sup>[13]</sup> Hydrolases and oxygenases are part of the two most important classes of enzymes to catalyse the hydrolysis of a large range of pesticides.<sup>[14]</sup> This was exemplified by the successful covalent immobilization of an organophosphate degrading enzyme used for pesticides depletion.<sup>[15]</sup> Biotechnology also comprises the fields of genetic engineering, protein engineering, and metabolic engineering and associated to the development of industrial production, the area of bioengineering.<sup>[4]</sup> Enzymes have become essential for industrial biotechnology, in synthetic biology and bioprocesses to

catalyse a large number of reactions.<sup>[16]</sup> Advances in DNA sequencing, protein tailoring and new design and engineering of metabolic pathway have contributed to establish effective biocatalysts.<sup>[16]</sup> Enzymes and microbes provide services, in the laundry area for instance and new sustainable processes for food, pharmaceuticals, cosmetics and bulk chemicals and also have several applications as analytical and diagnostic tools.<sup>[17]</sup> One success story, recently achieved through metabolic engineering of microbial pathway combined with synthetic biology, is the production of artemisinic acid.<sup>[18]</sup> Artemisinin possess antimalarial characteristics, and the production of 25 g L<sup>-1</sup> was accomplished through biosynthetic routes leading to the development of a new competitive antimalarial treatment.<sup>[19]</sup>

## **1.2. Renewable energies**

### **1.2.1. Biofuels**

While petroleum dominates, the utilization of biological routes is a growing approach to generate renewable energies as biofuels.<sup>[20]</sup> Besides the production of biogas by methanogenic bacteria, bioethanol and biodiesel are extensively studied and developed. Bioethanol is obtained after fermentation of simple sugar raw material or from starch whereas biodiesel contains a mixture of fatty acid methyl ester and fatty acids ethyl ester.<sup>[21]</sup> The most efficient biofuels production would use the sugarcane-based bioethanol, reducing greenhouse gases by 90 %.<sup>[22]</sup> A major drawback from this kind of production of biofuel is the dependency on raw materials as sugars, starch or crops in competition with food markets.<sup>[21]</sup> The controversy associated to the first-generation of biofuels led to the research and the investigation to produce a second generation using cellulosic feedstocks.<sup>[23,24]</sup> The utilization of cellulosic and agricultural waste for the production of bio-based chemical or biofuels still suffers from the difficulty of converting complex sugars into microbial fermentative sugars.<sup>[25]</sup> The production of bioethanol using this technology offers many advantages but more research need to be focused on feedstock preparation and fermentation methods need to be performed to make the process economically feasible.<sup>[26]</sup> Improvements have already been made in this area, by engineering enzymes and microorganisms producing cellulases to hydrolyse complex sugars and also producing catalysts needed from ethanol production.<sup>[27]</sup> Since a few years, the third generation of biofuels arises with the development of strategies for the cultivation of microalgae generating large amount of fats, proteins and carbohydrates via the photosynthesis phenomenon.<sup>[28,29]</sup>

### **1.2.2. Bio-based chemicals, flavours and fragrance market**

In addition to biofuels production, biotechnologies have great applications in the chemical industry, vast market including basic chemicals, specialty chemicals, cosmetics and

care products and life science products (Table 1).<sup>[30]</sup> Bulk chemicals, also termed as commodity chemicals (like surfactants, detergents or pesticides) and fine chemicals, corresponding to intermediates used in drugs, food additives, or flavours and fragrances areas, used to derive almost exclusively from petrochemical feedstocks.<sup>[31]</sup> Whole-cell biocatalysis and the application of enzymes can contribute in this field in the production of compounds enantiomerically pure with a high chemo-, regio- and stereoselectivity.<sup>[32]</sup> Biotechnology companies develop nowadays fermentation processes to drive the production of sugar-derived chemicals with the identical properties, functionalities and performance to replace these petrochemicals products.<sup>[33]</sup> By 2020 the biofuels and bio-based bulk chemicals, plastics and bioprocessing enzymes market is estimated at \$ 95 billion.<sup>[20]</sup>

Lactic acid, citric acid or succinic acid produced by fermentation routes, have significantly attracted attention for their application in the polymers area.<sup>[34-36]</sup> Calysta Energy company has developed for example a process to directly convert methane into commodity chemicals through fermentation and Nature Works used the methane fermentation to produce lactic acid.<sup>[33]</sup> Genetically modified yeast allowed the production of *trans*- $\beta$ -farnesene by Amyris and a yeast-based technology from Reverdia was developed to generate bio-succinic acid.<sup>[37,38]</sup> Fermentative methods provide polyhydroxybutyrate (PHB), used for biodegradable plastics and 1,3-propanediol (PDO) valuable biopolymer produced in *E. coli* and developed by Genencor and DuPont in collaboration.<sup>[39,40]</sup> A successful partnership between BASF, Cargill and Novozymes companies led to the production of bio-acrylic from 3-hydroxypropionic acid (3-HPA) and the commercialisation of 1,4-butanediol (BDO) by DuPont/Tate & Lyle was announced subsequently to the development of the fermentative technology by Genomatica.<sup>[41,42]</sup> Solvay is another firm, which uses the co-product glycerol, produced during the biodiesel fabrication, to generate at industrial scale an important chemical precursor for plastics, coating and adhesives, the epichlorohydrin.<sup>[43]</sup> Other corporations as Renewable Energy Life Science Group provide engineered strains for the production of alkanes, fatty acids and alcohols via synthetic biology applied at the metabolic pathway level and Codexis rapidly became a leading biocatalyst developer company.<sup>[44-46]</sup> The following table is a non-exhaustive list of chemicals and bio-based polymers produced in a fermentative manner (associated to their industrial applications) by companies involved in the development of the technology process and/or implicated in the commercialisation of the products.

**Table 1.** Chemical market, products and applications

Molecules	Applications	Companies	References
Isobutanol	Solvents, paints, biofuels, renewable fuels	Gevo	[47]
Poly(lactic acid) (PLA) Polyethylene Poly(hydroxyalkanoate) (PHA)	Plastics, food packaging	NatureWorks LLC Braskem, Dow Chemical, Metabolix, SyntheZyme	[48-50]
BiOH Polyols	Coatings, adhesives, elastomers	Cargill	[51]
Isoprene	Rubbers, adhesives	Genencor, Amyris	[52]
Isosorbide	Polyesters, surfactants, pharmaceuticals	Roquette, ADM, Cargill	[53]
1,4-Butanediol (BDO)	Spandex, povidone-iodine, drugs	Genomatica	[42]
1,3-Propanediol (PDO)	Cosmetics, personal care, home cleaning products	Genencor and DuPont Tate & Lyle	[40]
Succinic acid	Flavouring, dyes, perfumes, lacquers	Myriant, BioAmber, DSM	[53]
Isobutene	Polymers	Global Bioenergies	[54]
3-Hydroxypropionic acid (3-HPA)	Polymers production, food preservatives, probiotics	BASF, Cargill, Novozyme	[41]
Adipic acid	Nylon intermediates	Verdezyne	[55]
Epichlorohydrin	Coatings, adhesives, electronics	Solvay	[43]
<i>trans</i> - $\beta$ -farnesene	Building blocks, solvents, vitamins	Amyris	[37]

### 1.2.3. Terminal hydroxylated fatty acids: synthesis and application

The synthesis of valuable bioplastics and the preparation of high end polymers can also be initiated with terminal hydroxylated fatty acids ( $\omega$ -OHFA) and  $\alpha,\omega$ -dicarboxylic acids ( $\alpha,\omega$ -DCA);<sup>[56,57]</sup> oils and fats being the most abundant renewable raw material to provide chemicals.<sup>[58]</sup> They may serve as building blocks of poly- $\omega$ -OHFA, similar to PHA, polyesters and polyamides biodegradable material.<sup>[59]</sup> Besides the great interest for the chemical industry, these compounds are used in the cosmetics area for flavours and fragrances by companies like Givaudan or Firmenich in addition to applications for the production of anticancer agents and polyketides antibiotics.<sup>[60,61]</sup>

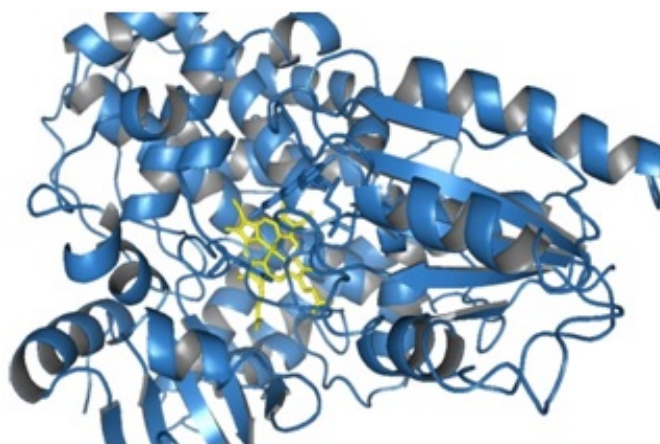
Chemically-based processes require expensive and toxic metal catalysts, rough conditions as high pressure and high temperature and the selective functionalization of carbon atoms is difficult to achieve.<sup>[62]</sup> Alternatively to chemical routes, the oxygenase enzymes can catalyse the introduction of one molecule of oxygen into non-functionalized carbon chain with high regio- and stereoselectivity for a large panel of substrates and various applications under mild conditions.<sup>[63]</sup> Oxygenation reactions are often catalysed by peroxidases, flavin-dependent monooxygenases (Baeyer-Villiger reactions), non-heme oxygenases and heme-containing cytochrome P450 monooxygenases.<sup>[64]</sup> Short chain-length alkanes can be hydroxylated by methane monooxygenases (MMO), propane monooxygenases (PMO) and butane monooxygenases (BMO), referred as binuclear metallic di-iron or di-copper alkane monooxygenases, while medium chain-length alkanes can be oxidized by the integral membrane non-heme iron monooxygenase, pAH1 and alkB from *Pseudomonas species*.<sup>[65,66]</sup> Alkane oxygenase from *Pseudomonas fluorescens* CHAO, AlkM from *Acinetobacter sp.* and LadA from *Geobacillus thermodentrificans* are classified as long chain-length alkane oxygenases.<sup>[67-69]</sup> As an alternative to the aforementioned catalysts, cytochrome P450 monooxygenase enzymes have been shown to hydroxylate a large spectrum of substrates including alkanes and fatty acids.<sup>[70]</sup>

The successful synthesis of terminal hydroxylated fatty acids via the biological routes reports strain engineering and cultivation of yeasts. *Candida* and *Yarrowia sp.* can grow on alkanes and fatty acids and are able to produce and to accumulate  $\omega$ -OHFA and  $\alpha,\omega$ -DCA via CYP52 family, cytochrome P450 monooxygenase enzymes.<sup>[71,72]</sup> Once the pathways for product depletion, substrate oxidation and product over-oxidation were blocked, the engineered *Candida tropicalis* has been shown to generate 174 g L<sup>-1</sup> of  $\omega$ -OHFA and 6 g L<sup>-1</sup> of  $\alpha,\omega$ -DCA from 200 g L<sup>-1</sup> methyl tetradecanoate after six days of biotransformation.<sup>[59]</sup> However the production at large scale of biosynthetic  $\omega$ -OHFA and  $\alpha,\omega$ -DCA via yeast is still challenging, due to low space-time yield and the requirement of suitable bioprocess implementation.<sup>[73]</sup>

## 2. Cytochrome P450 monooxygenase enzymes

### 2.1. Nomenclature

Cytochrome P450 monooxygenase enzymes (termed CYPs or P450s), originally discovered in rat liver microsomes, are from all kingdoms and are since their discovery extensively studied and characterized.<sup>[74,75]</sup> These hemoproteins display a characteristic maximum absorption peak of the reduced carbon monoxide (CO) gas-bound complex at 450 nm, from which derives their name and this unusual spectral property is still used to estimate the protein content.<sup>[76,77]</sup> The iron protoporphyrin IX centre coordinated to a cysteine thiolate into a hydrophobic binding pocket is responsible for the characteristic Soret absorption of the Fe<sup>II</sup>-CO complex at OD<sub>450</sub> (Figure 1).<sup>[78]</sup> A correct folded and therefore active P450 leads to the absorption peak at 450 nm, whereas a non-active enzyme will exhibit a maximum peak at 420 nm, resulting from a low and inefficient protein expression.<sup>[79]</sup> The replacement of the cysteine thiolate by a histidine ligand, demonstrated with the eukaryotic CYP1A2 and P450<sub>cam</sub> (from *Pseudomonas putida*), led to an inactive protein. This study emphasised the requirement of the cysteine thiolate group for functional P450s.<sup>[80-82]</sup>



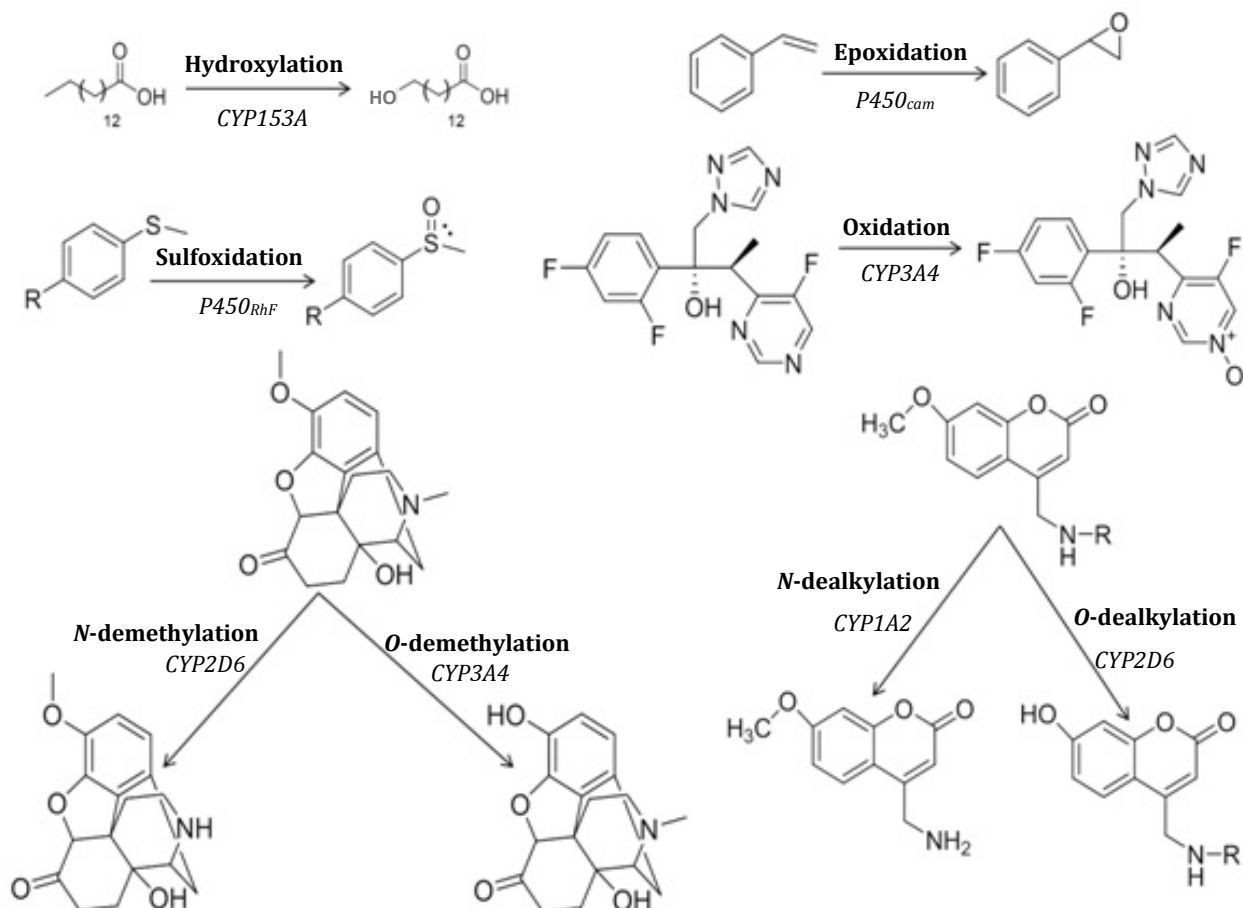
**Figure 1.** Visualization of the crystal structure of P450<sub>BM3</sub> (from CYP102A1 *Bacillus megaterium*), containing the iron heme in the centre of the active site, highlighted in yellow colour (Protein Data Bank code 1BU7).

CYPs are classified into families and subfamilies depending on the percentage of similarity: > 40 % and > 55 % similarity respectively.<sup>[83]</sup> The number of P450 sequences continuously increases, approaching nowadays 50 000 and the structures and sequences information are available online, incorporated into several P450 databases. As examples, “CYPED” from the ITB, University of Stuttgart, “Fungal Cytochrome P450 Database” from Seoul National University or “The cytochrome P450 homepage” from David Nelson, provide structures and sequences information to facilitate P450 studies and protein engineering.<sup>[84-87]</sup>

## 2.2. Reactions and mechanism

### 2.2.1. Variety of reactions

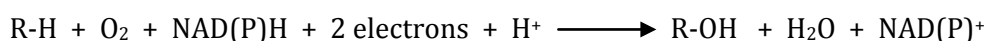
These versatile enzymes can incorporate an oxygen atom into a wide variety of substrates including alkanes, fatty acids, terpenes, steroids, prostaglandins, aromatics or xenobiotics compounds and therefore they can catalyse a large range of reactions such as hydroxylation, dealkylation, epoxidation or reduction reactions (Figure 2).<sup>[70]</sup>



**Figure 2.** Hydroxylation, epoxidation, sulfoxidation, oxidation, *O/N*-demethylation and *O/N*-dealkylation are P450-catalysed reactions exemplified by several CYP families.<sup>[88–93]</sup>

### 2.2.2. Mechanism

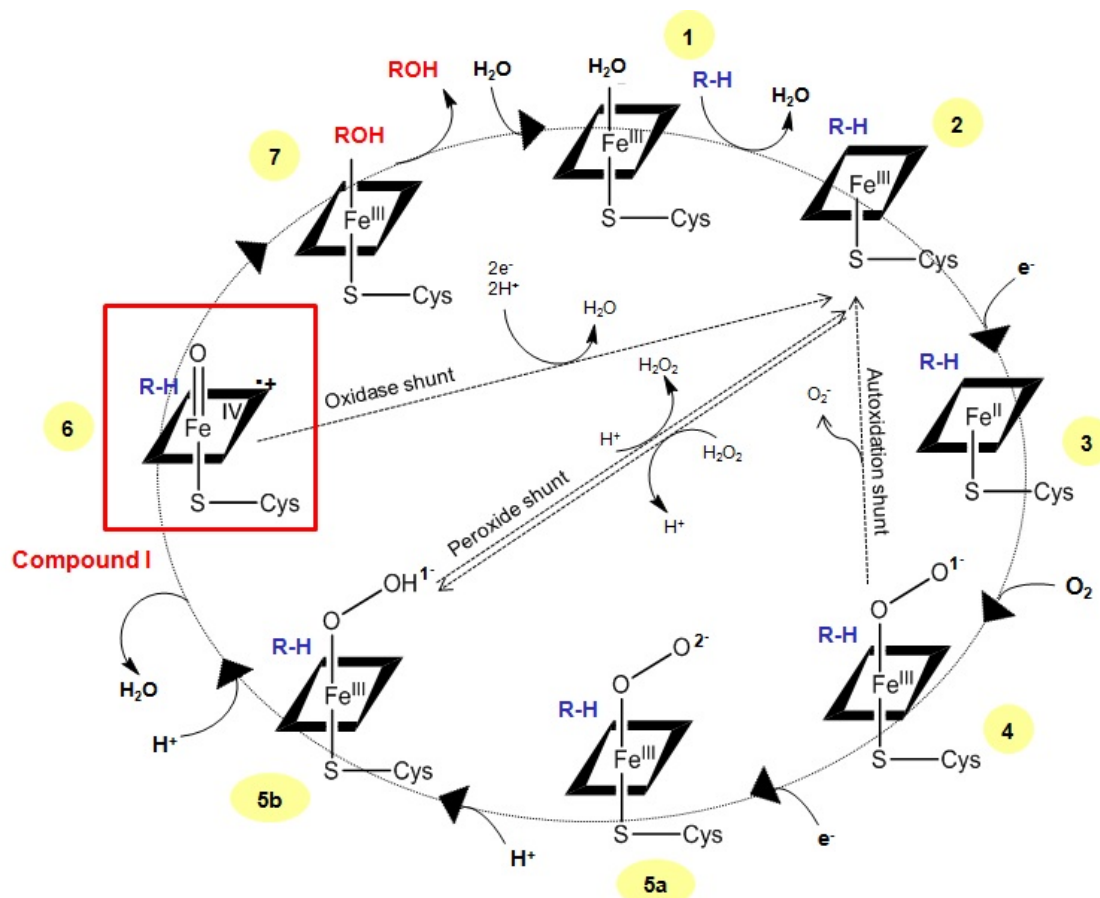
The classic hydroxylation reaction can be summarized by the reductive scission of the dioxygen bond, leading to the transfer of one molecule of oxygen to the substrate and the release of one molecule of water:



The “catalytic wheel” mechanism was proposed early 70’s exemplified with P450<sub>cam</sub> (Figure 3).<sup>[94–96]</sup> The substrate binds Fe<sup>III</sup> (1) and displaces a molecule of water (2), leading to the transfer of electron and the reduction into Fe<sup>II</sup> (3). An oxy-ferrous intermediate is formed by the binding of Fe<sup>II</sup> to dioxygen (4) and a second electron reduces the iron-peroxo complex generating an iron-hydroperoxo intermediate (5a). This intermediate is cleaved, releasing one molecule of water (5b) and forming the iron-oxo ferryl specie, termed compound I (6). The activated oxygen atom of the compound I oxidizes the substrate (7) and the product is displayed by a water molecule (1). However three potential side reactions, termed uncoupling events, can occur:

- The autoxidation shunt forming a superoxide anion from the oxidation of the intermediate oxy-ferrous;
- The peroxide shunt pathway leads to the formation of hydrogen peroxide when the peroxide or hydroperoxide anion dissociates from the iron;
- The oxidase shunt generates water with the reduction of the activated oxygen of the compound I instead of the oxidation of the substrate bound.

The formation of the reactive species instead of oxidized substrate cannot always be circumvented. This rapidly leads to the degradation of the prosthetic heme centre.<sup>[97]</sup>



**Figure 3.** “Catalytic wheel” of P450s, representation adapted from *Denisov et al., 2005*.<sup>[98]</sup>

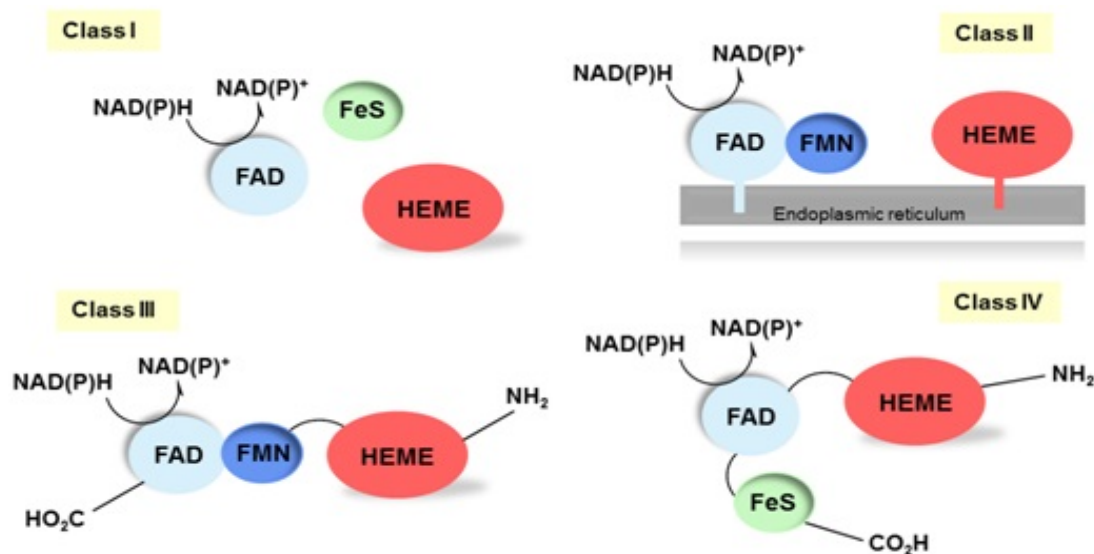


### 2.3. Structure and diversity

P450<sub>cam</sub>, extensively studied for the hydroxylation of substrate camphor, was the first P450 crystal structure published in 1987 (Protein Data Bank code 2CCP).<sup>[99]</sup> This structure had constituted a reference until the resolution of the structure of P450<sub>BM3</sub>, the most catalytically efficient P450 known so far with a  $k_{cat}$  of 17 100 min<sup>-1</sup> in the presence of arachidonic substrate (Protein Data Bank code 2HPD).<sup>[100-103]</sup>

Despite versatile families and diverse P450s discovered, a general structure is attributed to these enzymes: A-L  $\alpha$ -helices, 1-5  $\beta$ -strands, a meander loop and the prosthetic group.<sup>[84]</sup> This overall fold of P450s is maintained with conserved structures closed to the heme (termed helices I and L), while large variations are observed for distal regions from the heme controlling substrate specificity.<sup>[80]</sup>

P450<sub>BM3</sub> was a model to gain understanding on the reductase-P450 interactions for the transfer of electrons required for catalysis.<sup>[104]</sup> The cofactor NAD(P)H delivers two electrons in the reductase domain usually transferred via redox partners to the heme domain where the reaction takes place after binding of the substrate. Several topologies of P450 have been reported and led to a classification of 10 classes depending on the rearrangement of the reductase components (four main classes represented in Figure 4).<sup>[105]</sup> Most of the bacterial P450s are class I, three-component complex in which an iron-sulphur protein (ferredoxin) shuttles the electrons from the ferredoxin reductase to the heme. Eukaryotic P450s are generally membrane-bound proteins (class II), incorporated in the endoplasmic reticulum and receive electrons from a single polypeptide chain cytochrome P450 reductase (CPR) composed of flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN). Class III and class IV refer to “self-sufficient” P450s, natural soluble fused constructs found in bacteria and fungi. The respective best-known examples are P450<sub>BM3</sub> containing a CPR and P450<sub>Rhf</sub> from *Rhodococcus* sp., where the FMN-containing reductase is fused to a ferredoxin-like centre and to the P450 heme.<sup>[106]</sup>



**Figure 4.** Four main classes of P450s. Class I, bacterial three components system; Class II, mammalian P450s (fused structure, membrane-bound); Class III, bacterial fused domain found in P450<sub>BM3</sub>; Class IV, fusion system in *Rhodococcus* sp. Figure modified from O'Reilly et al., 2011.<sup>[107]</sup>

## 2.4. Fusion proteins

P450<sub>BM3</sub> is a natural soluble fusion biocatalyst in a single 119 kDa polypeptide chain containing CYP and reductase domain, which had served as model tool to understand the redox partner correlations and the membrane-bound P450s system.<sup>[80]</sup> Since the discovery of P450<sub>BM3</sub> enzyme in 1980s, a large number of self-sufficient chimeras enzymes had been generated leading to enhanced expression due to the single polypeptide in a soluble form with a higher catalytic activity and stability as well as efficient intramolecular electron transfer.<sup>[108]</sup> The number of engineered P450s has increased and many examples relate fusions between a class I or II P450 heme domain of interest and the CPR, reductase domain of P450<sub>BM3</sub>.<sup>[73,109]</sup> More recently P450<sub>RhF</sub> redox partners were shown to be suitable for the generation of efficient P450 chimera as well.<sup>[110]</sup> The artificial construct P450<sub>cam-RhF</sub> effectively hydroxylated (+)-camphor into 5-exo-hydroxycamphor and CYP153A13a SK2 from *Alcanivorax borkumensis* fused to the RhFred reductase successfully performed the hydroxylation of *n*-octane into 1-octanol.<sup>[111,112]</sup>

Several examples also show the impact of the modification of the linker between the reductase domain and the P450 domain, playing an essential structural role for an efficient electron transfer.<sup>[111,113]</sup> The variation in the length of the linker connecting the heme domain of CYP153A from *Marinobacter aquaeolei* (CYP153A<sub>M.aq.</sub>) and the reductase domain of CYP116B (*Rhodococcus* sp.) by extension of two amino acids, showed improved activity, higher stability

and increased coupling efficiency of the best variant. It was demonstrated that the fusion construct P450<sub>cam-RhFRed</sub> was 20-fold more active towards (+)-camphor to produce 5-exo-hydrocamphor during *in vivo* experiments after the introduction of a small polypeptide chain between the P450 heme domain and the redox partners. A flexible linker (Gly-Gly-Ser)<sub>n</sub> has been extensively used for engineering fusion proteins as it was shown to enhance the stability between the two fused domains.<sup>[73,114]</sup>

## 2.5. Microbial and fungal P450s

Many human P450s play role in the metabolism of endogenous steroids and xenobiotic compounds of great interest for the pharmaceutical industry and display a broad substrate scope, but they still suffer from low reaction turnover and reduced selectivity.<sup>[107,115]</sup> Numerous of bacterial P450s have nevertheless also the ability to metabolize steroids as demonstrated in *Bacillus sp.* studies or CYPs from *Streptomyces sp.* driving the effort to generate “human like” P450 activities.<sup>[107,116,117]</sup> Many microbial and fungal P450s can also have applications for natural product biotransformations and bioremediation properties.<sup>[80]</sup> Fungal P450s from yeast are class II associated proteins and while fungi are useful for bioremediation like the commercially available white-rot *Phanerochaete chrysosporium* for pollutants degradation, other yeast strains can utilize and transform alkanes and fatty acids found in soil as described previously.<sup>[59,118,119]</sup> As an alternative to the yeast platform, other P450s enable the activation of alkanes and/or fatty acids as well to produce terminal hydroxylated compounds; the greatest number remaining for reported CYPs from the prokaryotic kingdom (Table 2).

**Table 2.** Native or engineered P450s involved in the activation of alkanes and/or fatty acids

P450s	Species or strains	Characteristics	Substrates range	Ref
Mammalian CYP4	Rat, rabbit, mouse, human	Membrane bound FAD/FMN reductase NADPH dependent	C <sub>7</sub> -C <sub>10</sub> alkanes, fatty acids	[120]
Eukaryotic CYP52	<i>Candida tropicalis</i> <i>Candida maltose</i> <i>Candida apicola</i> <i>Yarrowia lipolytica</i>	Membrane bound FAD/FMN reductase NADPH dependent	C <sub>10</sub> -C <sub>16</sub> alkanes, fatty acids	[59,71]
Prokaryotic engineered CYP101	<i>Pseudomonas putida</i>	Soluble protein Ferredoxin/ Ferredoxin reductase NADH dependent	C <sub>3</sub> -C <sub>10</sub> alkanes	[121]

Table 2. (continued)

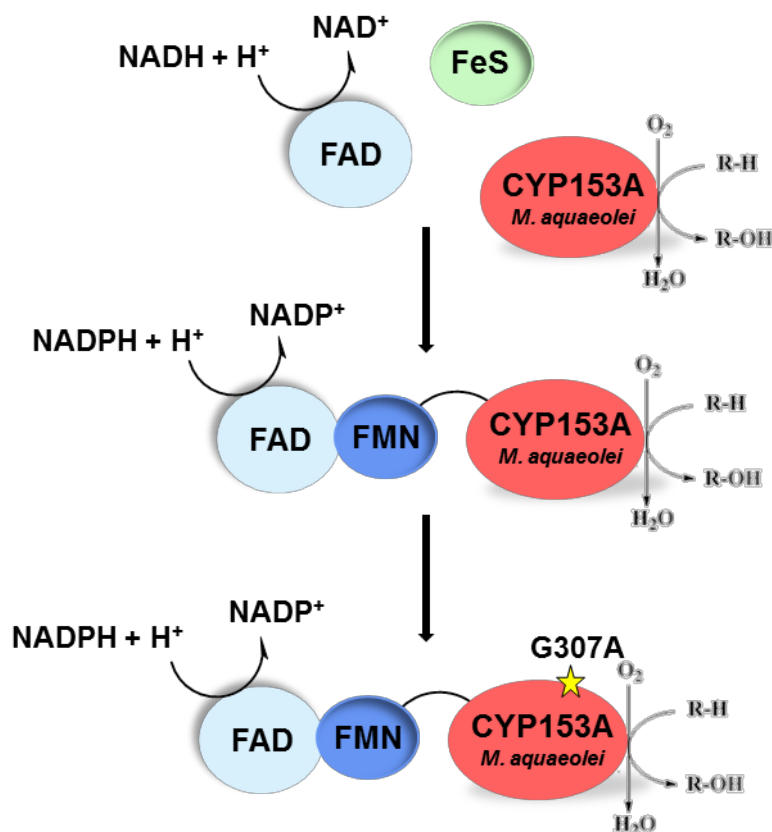
P450s	Species or strains	Characteristics	Substrates range	Ref
Prokaryotic engineered CYP102	<i>Bacillus megaterium</i>	Soluble protein	C <sub>1</sub> alkanes	[122,123]
	<i>Bacillus subtilis</i>	FAD/FMN reductase NAD(P)H dependent	C <sub>8</sub> -C <sub>12</sub> alkanes, fatty acids	
Prokaryotic CYP109B1	<i>Bacillus subtilis</i>	Soluble protein FAD/FMN reductase NAD(P)H dependent	C <sub>10</sub> -C <sub>14</sub> alcohols C <sub>8</sub> -C <sub>18</sub> fatty acids	[124]
Putative CYPs	<i>Rhodococcus rhodochrous</i> ATCC 10907 ( <i>Corynebacterium sp.</i> 7E1C/ <i>Gordonia rubripertincus</i> 7E1C)	Hypothetical P450s NAD(P)H dependent	C <sub>8</sub> , C <sub>10</sub> fatty acids, C <sub>5</sub> -C <sub>13</sub> alkane, C <sub>6</sub> -C <sub>10</sub> , C <sub>12</sub> , C <sub>14</sub> , C <sub>17</sub> , C <sub>18</sub> alkene	[125,126]
Prokaryotic CYP124	<i>Mycobacterium tuberculosis</i>	Soluble protein Ferredoxin/ Ferredoxin reductase NADH dependent	C <sub>12</sub> -C <sub>6</sub> methyl- branched fatty acids	[127]
Prokaryotic engineered CYP153	<i>Acinetobacter sp.</i> <i>Alcanivorax borkumensis</i> SK2 <i>Polaromonas sp.</i> <i>Mycobacterium sp.</i>	Soluble protein Ferredoxin/ Ferredoxin reductase NAD(P)H dependent	C <sub>4</sub> -C <sub>16</sub> alkanes, alkenes, 1-alcohols, C <sub>8</sub> -C <sub>16</sub> fatty acids	[88,128]
	<i>Mycobacterium sp. HXN-1500</i>		C <sub>3</sub> -C <sub>8</sub> alkanes	
	<i>Marinobacter aquaeolei</i>		C <sub>8</sub> -C <sub>16</sub> fatty acids and alkanes	

## 2.6. CYP153A family and generation of a fusion construct

The bacterial CYP153A family has been reported to hydroxylate a large variety of compounds in addition to the fatty acids such as alkanes, terpenes and primary alcohols.<sup>[88,128-130]</sup> The natural substrate, linear C<sub>5</sub> to C<sub>12</sub> alkanes, can be catalysed into the corresponding alcohols by this CYP family.<sup>[112]</sup> P450s from CYP153 family are three-component protein system class I, where a ferredoxin transfers the electrons from the reductase domain to the heme. The

first member of this family characterized was a CYP from *Acinetobacter sp.* before the isolation of CYP153A from *Mycobacterium sp.* converting primary alcohols and  $\alpha,\omega$ -diols.<sup>[88,131–134]</sup> CYP124A1 from *Mycobacterium tuberculosis* has been reported as methyl-branched fatty acid  $\omega$ -hydroxylase but was shown poorly active and  $\omega$ -specific towards linear fatty acids.<sup>[127]</sup> However CYP153A6 HXN-1500 from the same specie *Mycobacterium* has been successfully engineered and expressed in *E. coli* and *P. putida* to catalyse the production of 6 g L<sup>-1</sup> of (*S*)-perillyl alcohol from (*S*)-limonene in a two-phase system and with a decrease of side product by changing the host.<sup>[129]</sup> *E. coli* was also used as heterologous host to express the same engineered enzyme for the biocatalysis of *n*-octanol leading to 8.7 g L<sup>-1</sup> of 1-octanol after 24 h and 330 mg L<sup>-1</sup> of 1-octanol was formed in *P. putida* production strain.<sup>[135,136]</sup> At smaller scale biotransformation, the artificial fusion construct CYP153A13<sub>RhF</sub> was shown to produce 61 mg of 1-octanol g<sub>cdw</sub><sup>-1</sup>.<sup>[130]</sup>

Within the CYP153 family, CYP153A16 from *Mycobacterium marinum* and CYP153A from *Marinobacter aquaeolei* (CYP153A<sub>M.aq.</sub>) have been described as fatty acids  $\omega$ -hydroxylases, while CYP153A from *Polaromonas* was shown to be an alkane  $\omega$ -hydroxylase.<sup>[88]</sup> Investigations on CYP153A<sub>M.aq.</sub> have shown a preference of this biocatalyst for medium and long chain-length of fatty acids.<sup>[128]</sup> An artificial fusion construct was further engineered to generate a self-sufficient catalytic P450 consisting in the heme domain of CYP153A<sub>M.aq.</sub> fused to the reductase domain CPR of P450<sub>BM3</sub>.<sup>[73]</sup> A small focused mutant library was generated from the chimera protein, targeting improved activity for the terminal hydroxylation of primary alcohols and fatty acids and/or a shift in the substrate range. Among nine positions selected with specific residues chosen for substitutions after rational design studies, the single mutant G307A, position located in the binding pocket was identified as improved variant and the linker length was further modified, containing 3xGGS (termed CPR2<sub>mut</sub>). This engineered enzyme displayed a higher selectivity (> 95 %), an increased activity and a higher coupling efficiency compared to the wild type biocatalyst for the terminal hydroxylation of fatty acids in the range of C<sub>12:0</sub>-C<sub>14:0</sub> (Figure 5). The enhanced variant was expressed in *E. coli* for the biotransformation of dodecanoic acid methyl ester into 12-hydroxydodecanoic acid methyl ester, reaching a concentration of 4 g L<sup>-1</sup> in 28 h in 1 L fermenter in a two-phase system.<sup>[73]</sup> Furthermore, a bacterial *E. coli* strain was engineered to enhance the transport of longer chain fatty acids ( $\Delta$ *fadD* and *fadL* overexpressed) and harbouring the natural operon of CYP153A<sub>M.aq.</sub> to produce of 2.4 g L<sup>-1</sup> of  $\omega$ -hydroxy palmitic acid in shake flasks starting from 2.6 g L<sup>-1</sup> of palmitic acid substrate.<sup>[137]</sup>



**Figure 5.** Successive construction of the protein chimera CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> (G307A mutant, position located in the binding pocket, 3xGGs linker length) named CPR2<sub>mut</sub>. The heme domain from *Marinobacter aquaeolei* was fused to the reductase domain CPR of P450<sub>BM3</sub> and subsequently a mutant (G307A, 3xGGs) was generated showing improved selectivity and activity for the terminal hydroxylation of fatty acids.

### 3. Cytochrome P450s as industrial tools

Besides the extensive research for the bulk and fine chemicals, CYPs are remarkable useful biocatalysts for the pharmaceutical industry and for the production of drug metabolites.<sup>[116]</sup> Several synthetic steps for the production of antibiotics require P450s as the production of derivatives of erythromycin, which involves CYP107 from *Saccharopolyspora erythraea* or Taxol drug using plant P450s.<sup>[138,139]</sup> Novartis generated different strains enabling the expression of 14 different recombinant human P450s associated to the human CPR reductase domain for the production of drug metabolites. Sanofi company produces at large scale the anti-malaria drug artemisinin acid compound, successfully achieved through strain engineering, metabolic engineering and synthetic pathways (as previously introduced in section 1.1). This was done with the integration of CYP71AV1 isolated from *Artemisia annua*.<sup>[19]</sup> CYPs are also involved in the conversion of vitamin D3 into an active form, CYP1051A from *Streptomyces*

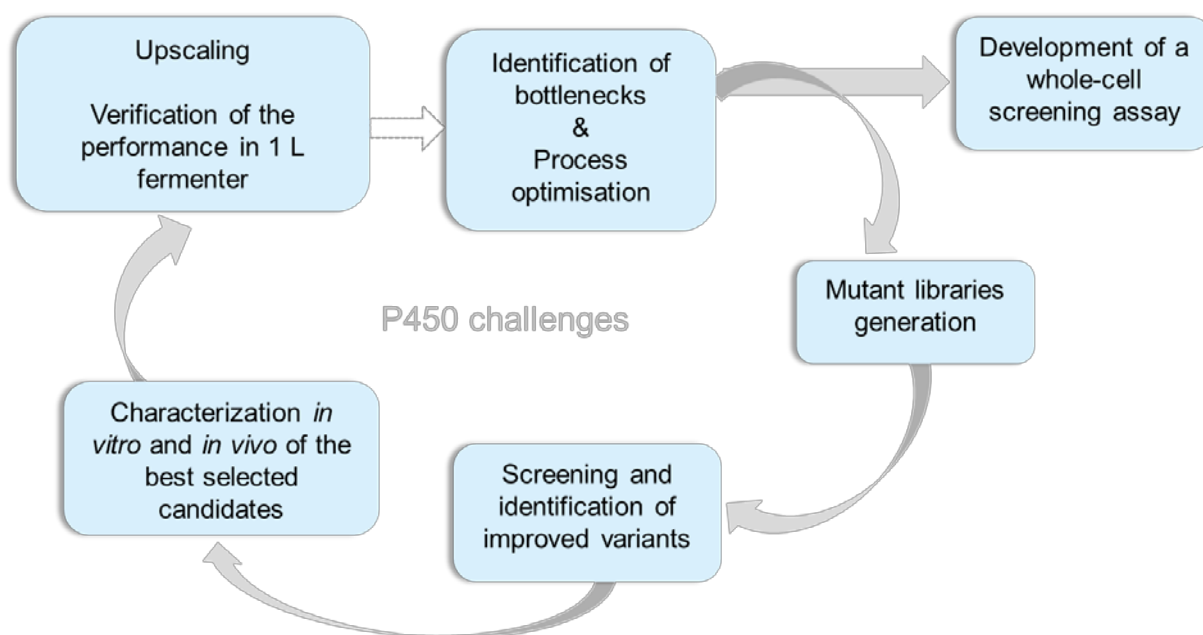
*griseolus* was shown to generate more than 2 mg L<sup>-1</sup> of 1 $\alpha$ 25-dihydroxyvitamine D3 in the engineered strain *Streptomyces lividans*. The production of hydrocortisone is a well-established process through the bioconversion of compound Reichstein S steroid by CYP from *Curvularia sp.* and provastatin drug, used for prevention of heart diseases, can be produced via the heterologous expression of eukaryotic P450s in *E. coli*.<sup>[140,141]</sup> A mutant of P450<sub>BM3</sub> (F87V) was shown to generate epoxyeicosatrienoic acid and leukotoxin B from arachidonic acid and linoleic acid, respectively.<sup>[142]</sup> In addition to pharmaceutical applications, dyes production via CYPs has been demonstrated: production of flowers with new colours can have application in the industry of horticulture.<sup>[143]</sup> Despite the high potential of CYP-catalysed reactions for industrial purposes, the system is hindered by several limitations at the biocatalyst level. Challenges are associated to low activities, need of redox partners, limiting electron transfer and uncoupling events.<sup>[116]</sup> In addition, a major restrictive factor is the P450 instability.<sup>[107]</sup> Reported challenges related to the host, regarding whole-cell biocatalysis, include substrate/product toxicity and inhibition, low substrate uptake, side product formation and limited oxygen transfer.<sup>[144–146]</sup> Therefore, extensive research is nowadays performed to understand these complex versatile P450s and to overcome the identified limitations for future industrial accomplishments.





## Aim of the project

The terminal hydroxylation of fatty acids via cytochrome P450 monooxygenase enzyme CPR2<sub>mut</sub> has been well investigated in the past *in vitro* and *in vivo*. However the process still suffers from limitations preventing from any possibility of industrialization and production at large scale of hydroxylated compounds. Three different aims have been attributed to this study to understand and to overcome such limitations, using as model CYP153A from *Marinobacter aquaeolei*.



**Figure 6.** Targets and strategies to improve the production of terminal hydroxylated fatty acids. The identification of limitations and improvements of the whole-cell process constitute the starting point of this study. The generation of mutant libraries, in parallel with the development of a screening assay were performed before the identification and the validation of enhanced mutants.

- ❖ According to previous studies, the formation of  $\omega$ -OHFA stops after 8 h during *E. coli* whole resting cell biotransformations using CPR2<sub>mut</sub>.<sup>[73]</sup> The potential limitations to this system, related to isolated P450s and correlated to P450 whole-cell catalysis include:
  - substrate/product inhibition
  - limited cofactor regeneration
  - limitation in the substrate transport
  - low P450 activity and stability

Besides the investigations on the limitations, solutions to address the identified challenges and upscaling strategies have to be implemented to reach an industrial suitable process.

- ❖ One alternative to overcome limitations often observed for monooxygenase enzymes involve protein engineering. Therefore, the second aim from this study includes the generation of mutant libraries deriving from CYP153A<sub>M.aq.</sub> wild type and the identification of improved variants regarding the activity for the terminal hydroxylation of fatty acids. Further enzyme characterizations have to be performed on the selected improved mutants and a validation of the performance should be achieved at larger scale during *in vivo* studies.
  
- ❖ To enable the screening of a large number of variants, a screening assay needs to be developed. The assay to established should be based on the use of the galactose oxidase enzyme originally evolved for oxidizing primary alcohols but shown to oxidize hydroxylated fatty acids as well.<sup>[147]</sup> A whole-cell based assay would enable the screening of large-size mutant libraries in a reduced time.

## Material and Method

### 1. Chemicals, media and bacterial strains

Yeast extract, agar and tryptone were purchased from Roth (Germany). Dodecanoic acid ( $C_{12:0}$ ) and 12-hydroxydodecanoic ( $\omega$ -OHC $_{12:0}$ ) and all other chemicals were acquired from Sigma-Aldrich (Germany). *E. coli* strain BL21 (DE3) was purchased from Novagen (Madison, WI, USA) and the *E. coli* BL21 ( $\Delta$ *fadD*), knock-out strain for the beta-oxidation pathway was engineered at the ITB, University of Stuttgart, Germany. *E. coli* strains HMS174 (DE3) was obtained from the Striedner lab (BOKU, Vienna, Austria) and *E. coli* DH5 $\alpha$  purchased from Invitrogen (Darmstadt, Germany) were transformed with *Dpn I*-treated PCR product.

- TB medium: 12 g L<sup>-1</sup> tryptone, 24 g L<sup>-1</sup> yeast extract, 4 mL of glycerol and water to a final volume of 900 mL before sterilization. Addition of 100 mL of 10xTB salt (720 mM of K<sub>2</sub>HPO<sub>4</sub> and 170 mM of KH<sub>2</sub>PO<sub>4</sub>).
- LB medium: 10 g L<sup>-1</sup> of tryptone, 5 g L<sup>-1</sup> of yeast extract, 10 g L<sup>-1</sup> of NaCl and water to a final volume of 1 L before sterilization. Addition of 15 g L<sup>-1</sup> of agar for LB agar medium.
- Auto-induction medium for galactose oxidase expression (8ZY-4LAC):<sup>[148]</sup> 40 g L<sup>-1</sup> yeast extract, 80 g L<sup>-1</sup> tryptone, 50xLAC (glycerol 25 %, glucose 2.5 % and  $\alpha$ -lactose monohydrate 10 %), 20xNPS ((NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 0.5 M, KH<sub>2</sub>PO<sub>4</sub> 1 M, Na<sub>2</sub>HPO<sub>4</sub> 1 M), 20xNPSC (NH<sub>4</sub>Cl 1 M, Na<sub>2</sub>SO<sub>4</sub> 0.5 M, KH<sub>2</sub>PO<sub>4</sub> 0.5 M, Na<sub>2</sub>HPO<sub>4</sub> 0.5 M, trace elements (FeSO<sub>4</sub>·7H<sub>2</sub>O (3.6 mM), ZnSO<sub>4</sub>·7H<sub>2</sub>O (30 mM), CuSO<sub>4</sub>·7H<sub>2</sub>O (1.5 mM), MnSO<sub>4</sub>·4H<sub>2</sub>O (0.7 mM), Na<sub>2</sub>B<sub>4</sub>O<sub>7</sub>·10H<sub>2</sub>O (0.3 mM) and (NH<sub>4</sub>)<sub>2</sub>MoO<sub>7</sub>·4H<sub>2</sub>O (0.04 mM)).
- Enriched minimum medium for fermentation during upscaling studies has been prepared as described elsewhere:<sup>[149]</sup> 3 g L<sup>-1</sup> KH<sub>2</sub>PO<sub>4</sub>, 6 g L<sup>-1</sup>, K<sub>2</sub>HPO<sub>4</sub>·3H<sub>2</sub>O; based on the theoretical cell dry mass (cdm) 0.15 g cdm<sup>-1</sup> of yeast extract, 0.25 g sodium citrate, 0.10 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 0.02 g CaCl<sub>2</sub>·2H<sub>2</sub>O, 4 mg CuCl<sub>2</sub>·2H<sub>2</sub>O, 3.2 mg ZnSO<sub>4</sub>·7H<sub>2</sub>O, 3 g glucose, 50  $\mu$ L trace elements (prepared in 5 N HCl : 40 g L<sup>-1</sup> FeSO<sub>4</sub>·7H<sub>2</sub>O, 10 g L<sup>-1</sup> MnSO<sub>4</sub>·H<sub>2</sub>O, 10 g L<sup>-1</sup> AlCl<sub>3</sub>·6H<sub>2</sub>O, 4 g L<sup>-1</sup> CoCl<sub>2</sub>, 2 g L<sup>-1</sup> ZnSO<sub>4</sub>·7H<sub>2</sub>O, 2 g L<sup>-1</sup> Na<sub>2</sub>MoO<sub>7</sub>·2H<sub>2</sub>O, 1 g L<sup>-1</sup> CuCl<sub>2</sub>·2H<sub>2</sub>O, and 0.5 g L<sup>-1</sup> H<sub>3</sub>BO<sub>3</sub>).

## 2. Plasmids and constructs

### 2.1. Plasmids and strains

Plasmid pET-28a(+) (Novagen, Madison, WI, USA) and L-rhamnose inducible plasmid pJOE4782.1 (kindly provided by Dr. Josef Altenbuchner) with the fusion construct CPR2<sub>mut</sub> inserted, were generated at the ITB, University of Stuttgart, Germany.<sup>[150]</sup> pET-30a vector containing the galactose oxidase enzyme mutant (GOase<sub>M3-5</sub>) gene was kindly provided by Prof. Sabine Flitsch and Prof. Nicholas Turner, University of Manchester.<sup>[147]</sup> The different *E. coli* strains and plasmids used in this study are listed in the Table 3.

**Table 3.** Plasmids constructed and used bacterial strains

Plasmids and strains
<i>E. coli</i> BL21 (DE3)_pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub>
<i>E. coli</i> BL21 (DE3)_pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (G307A)
<i>E. coli</i> BL21 (DE3)_pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (variant libraries)
<i>E. coli</i> BL21 (DE3)_pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> +(3xGGS)
<i>E. coli</i> BL21 (DE3)_pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (G307A)+(3xGGS) (termed CPR2 <sub>mut</sub> )
<i>E. coli</i> BL21 ( <i>ΔfadD</i> )_pJOE4782.1- CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (G307A)+(3xGGS) (termed CPR2 <sub>mut</sub> )
<i>E. coli</i> HMS174 (DE3)_pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (G307A)+(3xGGS) (termed CPR2 <sub>mut</sub> )
<i>E. coli</i> BL21 (DE3)_pET30a-GOase <sub>M3-5</sub>

### 2.2. Mutant libraries

The plasmid pET-28a(+) harbouring the artificial construct CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> was mutated on several positions using the Quick Change method in order to generate mutant libraries at the entrance of the substrate tunnel and within the active site (primers designed in Table 4 and Table 5, Appendix A1 and A2). The positions were selected after creation of a homology model based on multiple CYP sequence alignments and docking simulations (bioinformatics work provided by Łukasz Gricman, ITB, University of Stuttgart).<sup>[84,151,152]</sup>

**Table 4.** Designed primers containing degenerated codon (substrate entrance tunnel positions) based on CASTER 2.0 tool.<sup>[153]</sup>

Positions	Primers	Sequences (5'--> 3')
R77	Forward Reverse	ccg ttt tta tac <u>SYG</u> cag ggt cag tgg cgc gcg cca ctg acc ctg <u>CRS</u> gta taa aaa cgg
D134	Forward Reverse	ccg caa atc att ctc ggt <u>GKS</u> cct ccg gag ggg ctg tgc cga cag ccc ctc cgg agg <u>SMC</u> acc gag aat gat ttg cgc
S140	Forward Reverse	cct ccg gag ggg ctg <u>YKK</u> gtg gaa atg ttc gaa cat ttc cac <u>MMR</u> cag ccc ctc cgg agg
S233	Forward Reverse	ggt gca gca <u>SSB</u> gcc acc ggc ggg gag ttt aaa ctc ccc gcc ggt ggc <u>VSS</u> tgc tgc acc
T235	Forward Reverse	ggt gca gca tgc gcc <u>SKS</u> ggc ggg gag ttt aaa ctc ccc gcc <u>SMS</u> ggc cga tgc tgc acc
S453	Forward Reverse	cgg gtg cag <u>GSS</u> aac ttc gtg cgg ggc tat ata gcc ccg cac gaa gtt <u>SSC</u> ctg cac ccg

**Table 5.** Designed primers to generate mutant libraries via single point mutation

Variants	Primers	Sequences (5'--> 3')
Q129P	Forward Reverse	gcc gag ccg <u>CCG</u> atc att ctc ggt gac cct agg gtc acc gag aat gat <u>CGG</u> cgg ctc ggc
I131P	Forward Reverse	ccg caa atc <u>CCG</u> ctc ggt gac cct ccg cgg agg gtc acc gag <u>CGG</u> gat ttg cgg
V141G	Forward Reverse	gag ggg ctg tgc <u>GGC</u> gaa atg ttc ata gcg cgc tat gaa cat ttc <u>GCC</u> cga cag ccc ctc
M143G	Forward Reverse	ggg ctg tgc gtg gaa <u>GGC</u> ttc ata gcg atg cat cgc tat gaa <u>GCC</u> ttc cac cga cag ccc
M143S	Forward Reverse	ggg ctg tgc gtg gaa <u>AGC</u> ttc ata gcg cgc tat gaa <u>GCT</u> ttc cac cga cag ccc
I145V	Forward Reverse	gtg gaa atg ttc <u>GTG</u> gcg atg gat ccg ccg cgg cgg atc cat cgc <u>CAC</u> gaa cat ttc cac
I145L	Forward Reverse	gaa atg ttc <u>CTG</u> gcg atg gat ccg ccg cgg cgg atc cat cgc <u>CAG</u> gaa cat ttc
M228L	Forward Reverse	gag tgg tgc gac aga <u>CTG</u> gca ggt gca tgc acc tgc <u>CAG</u> tct gtc cga cca ctc
A229F	Forward Reverse	gac aga atg <u>TTT</u> ggt gca gca tgc cga tgc tgc acc <u>AAA</u> cat tct gtc
A229L	Forward Reverse	gac aga atg <u>CTG</u> ggt gca gca tgc cga tgc tgc acc acc <u>CAG</u> cat tct gtc

Table 5. (continued)

Variants	Primers	Sequences (5'--> 3')
L301A	Forward Reverse	gag ttt atc ggt aat <u>GCG</u> acg ctg ctc ata tat gag cag cgt <u>CGC</u> att acc gat aaa ctc
L301V	Forward Reverse	gag ttt atc ggt aat <u>GTG</u> acg ctg ctc ata tat gag cag cgt <u>CAC</u> att acc gat aaa ctc
T302L	Forward Reverse	ggt aat ttg <u>CTG</u> ctg ctc ata gtc ggc ggc gcc gcc gac tat gag cag <u>CAG</u> caa att acc
T302I	Forward Reverse	ggt aat ttg <u>ATT</u> ctg ctc ata gtc ggc ggc gcc gcc gac tat gag cag <u>AAT</u> caa att acc
L303D	Forward Reverse	ggt aat ttg acg <u>GAT</u> ctc ata gtc ggc ggc gcc gcc gac tat gag <u>ATC</u> cgt caa att acc
L303T	Forward Reverse	ggt aat ttg acg <u>ACC</u> ctc ata gtc ggc ggc gcc gcc gac tat gag <u>GGT</u> cgt caa att acc
V306F	Forward Reverse	ctg ctc ata <u>TTT</u> ggc ggc aac gat acg cgt atc gtt gcc gcc <u>AAA</u> tat gag cag
V306I	Forward Reverse	ctg ctc ata <u>ATT</u> ggc ggc aac gat acg acg cgt cgt atc gtt gcc gcc <u>AAT</u> tat gag cag
G307R	Forward Reverse	ctg ctc ata gtc <u>CGT</u> ggc aac gat acg acg cgt cgt atc gtt gcc <u>ACG</u> gac tat gag cag
D310E	Forward Reverse	ctc ata gtc ggc ggc aac <u>GAA</u> acg acg cgc aac gtt gcg cgt cgt <u>TTC</u> gtt gcc gcc gac tat gag
Y356F	Forward Reverse	caa acg ccg ctg gcc <u>TTT</u> atg cgc gcg cat <u>AAA</u> ggc cag cgg cgt ttg
Y356L	Forward Reverse	caa acg ccg ctg gcc <u>CTG</u> atg cgc gcg cat <u>GAC</u> ggc cag cgg cgt ttg
M357L	Forward Reverse	gcc tat <u>CTG</u> cgc cga atc gcc aag ctt ggc gat tcg gcg <u>CAG</u> ata ggc
R358P	Forward Reverse	gcc tat atg <u>CCG</u> cga atc gcc aag cag ctg ctt ggc gat tcg <u>CGG</u> cat ata ggc
F455V	Forward Reverse	cgg gtg cag tcc aac <u>GTG</u> gtg cgg ggc tat ata gcc ccg cac <u>CAC</u> gtt gga ctg cac ccg

Besides single point mutations, five double mutants and one triple mutant (combination of the best mutants described) were generated by Quick Change technique according to above mentioned designed primers: G307A/S2140R; G307A/S140L; G307A/S233G; G307A/T302I; S233G/T302I; G307A/S233G/T302I.

### 2.3. Quick Change protocol

**Table 6.** Quick Change procedure and PCR programme

Components	Concentrations/ volumes	PCR Program
DNA polymerase buffer +MgSO <sub>4</sub> 10X	1X / 5 µL	Denaturation: 95 °C, 2 min - Denaturation: 95 °C, 30 sec - Annealing: 55 °C, 60 sec - Extension: 72 °C, 4 min (30 sec/kb of plasmid length) - Final extension: 72 °C, 7 min - Storage at 8 °C
Forward Primer	10 pmol / 1 µL	
Reverse Primer	10 pmol / 1 µL	
Deoxyribonucleotide mix (dNTPs)	50 mM / 1 µL	
Plasmid template	100 ng / 2 µL	
<i>Pfu</i> Ultra II polymerase enzyme	1 µL	
Sterile water	Up to 50 µL	

Following the PCR reaction, the mixtures were treated for 2 hours digestion with 1 µL of restriction enzyme *Dpn I* at 37 °C (Fermentas) to remove the parental plasmid. The treated PCR mixture was purified and concentrated by a ZymoClean DNA concentrator kit and transformed into *E. coli* DH5α.

## 3. Cell cultivation, protein expression and whole resting cell preparation

### 3.1. CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> at small scale

An initial seed culture from overnight LB media (from a fresh transformant of *E. coli*), was used to inoculate 2 mL of TB media (containing 30 µg mL<sup>-1</sup> kanamycin) in a 24 deep-well plate for cell cultivation at 37 °C and 180 rpm on a plate shaker (Edmund Bühler GmbH, Tübingen, Germany). The protein was induced when OD<sub>600</sub> reached 0.8-1 by the addition of 0.5 mM of 5-aminolevulinic acid, 0.5 mM FeCl<sub>3</sub> and 0.1 mM IPTG. After protein expression overnight (16 h-20 h) at 25 °C, 180 rpm, the cells were harvested (4000 rpm, 20 min, 4 °C) and washed one time with 2 mL of 100 mM potassium phosphate buffer pH 7.4. The final whole resting cell resuspension corresponded to 100 g<sub>cww</sub> L<sup>-1</sup> in 100 mM potassium phosphate buffer pH 7.4 for immediate use.

### 3.2. CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> in shaking flasks

An initial seed culture (one single colony from *E. coli* retransformed) in LB media was used to inoculate 1 L or 2 L Erlenmeyer shaking flasks using TB media (30  $\mu\text{g mL}^{-1}$  kanamycin) for cell cultivation and protein expression aerobically in shaking incubators (Multitron, Infors HT, Bottmingen, Switzerland). Cells were grown at 37 °C and 180 rpm until OD<sub>600</sub> of 0.7-1 was reached before induction and protein expression using 0.5 mM of 5-aminolevulinic acid and 0.1 mM of IPTG. Protein expression was performed for 16 h-20 h at 25 °C and 180 rpm, and the cells were harvested the following day (4000 rpm, 20 min, 4 °C), washed two times with 100 mL of 100 mM potassium buffer pH 7.4 and resuspended to reach a final cell concentration of 100 g<sub>cww</sub> L<sup>-1</sup>.

### 3.3. CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> in 1 L fermenter

An initial seed culture from -80 °C glycerol stock or fresh transformant in LB media (grown overnight at 37 °C, 180 rpm) was used to inoculate 1 L fermenter (Infors AG, Bottmingen, Switzerland). The fermentation was performed by glucose limited fed-batch with a growth rate of 0.1 h<sup>-1</sup>.<sup>[149]</sup> Iris software controlled the fermentation process and the feed started with the depletion of the carbon source from the batch media (pO<sub>2</sub> spike indication). The protein expression was induced by 0.5  $\mu\text{mol IPTG per g cell dry weight (cdw)}$  and the temperature was decreased from 37 °C to 30 °C, pH 7.2, controlled by addition of 14 % NH<sub>4</sub>OH and pO<sub>2</sub> set to 30 % was controlled by agitation. Foam was controlled by addition of 0.5 mL antifoam 204 (Sigma-Aldrich) per liter media. The following day, the cells were harvested by centrifugation (4000 rpm, 25 min, 4 °C), washed one time and resuspended in 100 mM potassium phosphate buffer pH 7.4, supplemented with 20 mM glucose and 1 % glycerol to a final concentration of 100 g<sub>cww</sub> L<sup>-1</sup> and stored under gently shaking in 4 °C until further use.

### 3.4. Biocatalytic investigations with resting cells

#### 3.4.1. 24 deep-well plates

A two-hour biotransformation was carried out in a final volume of 1 mL in 24 deep-well plate or in 2 mL Eppendorf tube, containing 50 g<sub>cww</sub> L<sup>-1</sup> resting cells in 100 mM potassium phosphate buffer pH 7.4, 2 mM substrate fatty acid provided (40 mM stock, 5 % DMSO (v/v) end concentration) and performed at 25 °C under 500 rpm stirring. The reaction was stopped after two hours, the whole resting cells were centrifuged (7000 rpm, 3 minutes, 4 °C) and the supernatant was stored at -20 °C until further analysis.



### 3.4.2. 100 mL Erlenmeyer flasks

In 100 mL Erlenmeyer shaking flasks, 50 g<sub>cww</sub> L<sup>-1</sup> resting cells in 100 mM potassium phosphate buffer pH 7.4, were used for whole-cell biotransformation studies, supplemented with 0.5 % glucose (1 M stock, (v/v)) and 1 % glycerol (v/v) at the start, after 4 h and 6 h or 8 h. The resting cell biocatalytic activities were performed at 25 °C and 180 rpm in shaking incubators tested for different concentrations of substrate C<sub>12:0</sub> (40 mM stock, 5 % DMSO (v/v) end concentration).

### 3.4.3. 1 L stirring tank reactor

The biotransformations investigations were performed with 300 mL of final volume of 50 g<sub>cww</sub> L<sup>-1</sup> of cell culture in a 1 L fermenter (resuspended in 100 mM potassium phosphate buffer pH 7.4). Several concentrations of fatty acids (5 % DMSO (v/v) end concentration) were provided and the reaction was carried out at 30 °C under stirring. Glucose 0.5 % (1 M stock, (v/v)) and 1 % glycerol (v/v) were added to the whole resting cell at the beginning, after 4 h and 8 h, pO<sub>2</sub>, temperature and pH (10 % H<sub>3</sub>PO<sub>4</sub> and 5 M NaOH) were controlled with setpoint of 30 % (via Iris software), 30 °C and 7.4, respectively, and aeration was set to 3.75 vvm.

## 3.5. Galactose oxidase expression

A seed culture of *E. coli* BL21 (DE3) transformed with GOase<sub>M3-5</sub> was grown overnight in 5 mL LB containing 30 µg mL<sup>-1</sup> kanamycin (from 30 mg mL<sup>-1</sup> stock solution) and 500 µL from the overnight culture was used to inoculate 250 mL of an auto-induction medium (8ZY-4LAC) containing kanamycin (30 µg mL<sup>-1</sup>) in 2 L baffled Erlenmeyer flask.<sup>[148]</sup> Cells were grown at 26 °C and 250 rpm for 60 h in a shaking flasks incubator. After protein expression, the cells were centrifuged for 20 min at 6000 rpm at 4 °C and resuspended in 50 mM sodium phosphate buffer before protein purification.

## 4. Protein purification

### 4.1. Purification of CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub>

The cells were centrifuged (4000 rpm, 20 min, 4 °C), then 1 g of cells was resuspended in 5 mL of 50 mM TRIS-HCl buffer pH 7.4 and disrupted using an EmulsiFlex-C5 (Avestin, Mannheim, Germany) or by sonication on ice (3x2 min, 1 min interval, 40 % amplitude, 0.35 sec cycle time (Branson Sonifier 250, Schwäbisch Gmünd, Germany)). The cell lysate was centrifuged to remove any cell debris (17 000 rpm, 45 min, 4 °C) and the supernatant containing

the soluble P450 protein was recovered. The protein was purified using an Äkta system (Healthcare Biosciences, Uppsala, Sweden), with a weak anion exchange column (30 mL) packed with Toyopearl DEAE 650M (Tosoh, Stuttgart, Germany) at a maximum flow rate of 10 mL min<sup>-1</sup>. The binding buffer consisted of 50 mM TRIS-HCl buffer pH 7.4 and for elution a step gradient of 50 mM TRIS-HCl containing 1 M NaCl was used. The protein elution was detected by a characteristic absorbance at 418 nm, in addition to the total protein detection at 280 nm<sup>[154]</sup> and eluted at a concentration of 250 mM NaCl. After purification, the protein was concentrated (by centrifugation 8000 rpm, 4 °C) using 100 kDa cut-off Vivaspin tubes (Sartorius, Goettingen, Germany) and aliquots were stored at -20 °C.

#### 4.2. Purification of the galactose oxidase

1 g of cell pellet was resuspended in 5 mL of phosphate-buffered saline (PBS) binding buffer and was lysed by EmulsiFlex (Avestin, Mannheim, Germany) and centrifuged again for 45 min at 4 °C and 17 000 rpm. The supernatant was filtered and diluted with the same volume of PBS binding buffer. A single step affinity purification was performed by the ÄKTA purifier 10 (GE Healthcare Biosciences, Uppsala, Sweden) using a Strep-Tag-II column. The protein was eluted with PBS buffer containing 2.5 mM of desthiobiotin. After purification, the protein samples were dialyzed into 100 mM potassium phosphate buffer pH 7, containing copper-sulphate (1 mM) for protein loading for 24 h at 4 °C. The next day the protein samples were dialyzed into 100 mM potassium phosphate buffer pH 7.4, for 24 h at 4 °C to remove the copper and for storage. Following the dialysis steps, the samples were concentrated using viva spin tubes (Sartorius, Goettingen, Germany) and by centrifugation at 5000 rpm, during 30 minutes.

#### 4.3. *In vitro* performance and kinetics parameters

Identified mutants from the mutant libraries were tested for improved activity *in vitro* towards C<sub>8:0</sub>, C<sub>12:0</sub> and C<sub>16:0</sub> (1 mM of fatty acid from 20 mM stock, 5 % DMSO (v/v) end concentration). The reactions were performed in 100 mM potassium phosphate buffer pH 7.4, for 2 hours at 700 rpm stirring, 30 °C with 0.5 μM P450 (lysate) and the reactions were stopped by the addition of 30 μL of HCl and stored at -20 °C. Additional characterizations regarding kinetic constants were determined for the best mutants based on initial rates (2 minutes reaction) and the activity was tested along 2 h performance with sampling after 5 minutes, 15 minutes, 30 minutes, 1 h and 2 h.

## 5. Substrate and product inhibition studies

### 5.1. *In vitro* reactions

Substrate and product inhibition were investigated using purified enzyme CPR2<sub>mut</sub> (0.5  $\mu$ M and 1  $\mu$ M) and lysate from the best selected variants from the generated libraries (0.5  $\mu$ M) in a final volume of 300  $\mu$ L in Eppendorf tubes containing 100 mM potassium phosphate buffer pH 7.4, 1 mM NADPH cofactor and a cofactor regeneration system (5 mM glucose-6-phosphate, 1 mM MgCl<sub>2</sub>, 12 U mL<sup>-1</sup> glucose-6-phosphate dehydrogenase from *Leuconostoc mesenteroides*). For substrate inhibition studies the concentrations of C<sub>12:0</sub> were varied: 0.1, 0.2, 0.25, 0.5, 1, 1.3, 2 and 2.5 mM were tested. The product inhibition experiments were performed with 1 mM of C<sub>12:0</sub> added at the start of the reaction and 0, 0.1, 0.2, 0.5, 1 or 1.5 mM of  $\omega$ -OHC<sub>12:0</sub> supplemented as well. The substrate and product were dissolved in DMSO prior to addition and the final DMSO concentration was kept at 5 % (v/v) in the substrate inhibition study and 10 % (v/v) in the product inhibition study, assuming one-phase system. The biotransformations were performed for 2 minutes at 30 °C, 700 rpm stirring, starting with the addition of 1 mM NADPH.

### 5.2. *In vivo* biotransformations in shaking flasks

*In vivo* substrate and product inhibition studies were performed similarly to the *in vitro* studies with CPR2<sub>mut</sub> only. Biotransformations with 50 g<sub>cww</sub> L<sup>-1</sup> resting whole-cells in 100 mM potassium phosphate buffer pH 7.4, in 100 mL Erlenmeyer shaking flasks were used and various initial substrate concentrations for substrate inhibition studies (0.5, 1, 1.2, 2, 2.5 and 3 mM of C<sub>12:0</sub>) and a constant substrate concentration (1 mM C<sub>12:0</sub>) with various amounts of product added at time zero were applied for product inhibition studies (0, 0.5, 1, 2.5, 5 mM of  $\omega$ -OHC<sub>12:0</sub>). Each fatty acid was dissolved in DMSO (5 % or 10 % (v/v)). Inhibition profiles were estimated from initial reaction rates.

### 5.3. *In vivo* reactions in bioreactors

Substrate inhibition was also investigated at larger scale for 24 h, in 1 L fermenter, using 300 mL of final volume containing 50 g<sub>cww</sub> L<sup>-1</sup> of whole resting cells in 100 mM potassium phosphate buffer pH 7.4, and three different concentrations of C<sub>12:0</sub> were tested: 5 mM, 10 mM and 15 mM (5 % DMSO (v/v) end concentration).

## 6. Cofactor limitations investigations

Cofactor limitation was investigated by addition of external NADPH or an artificial cofactor, nicotinamide-3,4-chloro-methoxy-phenyl-hydrate (NACMH, prepared by Dr. Sebastian Löw, ITB, University of Stuttgart (Appendix A3)). They were provided in stoichiometric concentrations 1.5 mM or 3 mM (from 10 mM stock) to 50 g<sub>cww</sub> L<sup>-1</sup> whole resting cells in 100 mM potassium phosphate buffer pH 7.4. This was evaluated for 8 h and the reaction was initiated by the addition of 1.5 mM of C<sub>12:0</sub> (5 % DMSO (v/v) end concentration).

## 7. Transport limitation studies

Potential transport limitation of the substrate across the cell membrane was examined by application of different permeabilization methods (as mechanical and chemical techniques) prior to the biocatalytic reaction. 10 mL of 100 g<sub>cww</sub> L<sup>-1</sup> of cells were spun down and the pellet frozen in -20 °C overnight. Acetone treated cells were incubated with 5 % acetone (v/v) during 2 minutes while vortexing. Sonication was performed during 1x2 minutes (amplitude 60 %, 0.5 sec cycles) (UP400 S; Hielscher Ultrasonic GmbH, Teltow, Germany). Subsequently to the permeabilization technique (acetone treatment and sonication) cells were spun down and the pellet was resuspended in 100 mM potassium phosphate buffer pH 7.4, to reach a cell density of 100 g<sub>cww</sub> L<sup>-1</sup>.

## 8. Determination of P450 concentration

The P450 concentration from purified proteins, from cell lysates and from whole-cells was determined by CO differential spectral assay.<sup>[76,77]</sup> The samples from fermentations and biotransformations were stored as cell pellets at -20 °C until analysis or directly measured. Small fractions were resuspended in 2 mL 100 mM potassium phosphate buffer pH 7.4 and a spatula of sodium hydrosulfite was added to the resuspension before incubation on ice for 30 min (in the case of whole cells). The resuspension was split into two cuvettes, of which one was treated with carbon monoxide gas (CO) for a few seconds. Differential spectra were immediately measured between 400 nm and 500 nm on spectrophotometer (Ultrospec 3100 pro, Amersham Biosciences, Freiburg, Germany) or on plate reader (POLARstar Omega BMG Labtech, Ortenberg, Germany). The samples from whole-cells were kept 1 hour at 4 °C before spectrophotometric measure. As an alternative to the sodium dithionite, P450 spectra were recorded after addition of 1 mM NADPH, 200 µL of fatty acid and subsequently addition of CO gas.

## 9. Cell dry weight

Pellet samples of 1 mL from cell cultivations and biotransformations were kept frozen at -20 °C until analysis; thawed and resuspended in 1 mL 100 mM potassium phosphate buffer pH 7.4 and filtered through pre-weighed 0.22 µm PES membrane filter (Frisenette, Knebel, Denmark) by application of vacuum. Prior and after application of the sample, the filters were weighted and dried in a microwave and left to equilibrate in a desiccator.

## 10. Extraction procedure and GC-FID method

Samples were extracted twice with the same volume of methyl-*ter*-butylether (MTBE), the organic phases were pooled and evaporated (using GeneVac EZ 2 Plus, Ipswich England). The internal standard decanoic acid was used for the quantification of C<sub>8:0</sub> and C<sub>12:0</sub>, whereas pentanoic acid was used in the case of C<sub>16:0</sub>. The samples were resuspended in equal volume of 45 µL of *N,O*-bis(trimethylsilyl)trifluoroacetamide containing 1 % trimethylchlorosilane (BSTFA) and MTBE for 30 min derivatization at 70 °C. Reaction progress was monitored by GC-FID (Shimadzu, Japan) using a Elite-5 column (PerkinElmer, Waltham, MA, USA) and hydrogen as carrier gas. For analysis of fatty acids, the column oven was set at 100 °C for 2 min for C<sub>8:0</sub>, 140 °C for 2 min for C<sub>12:0</sub> and 200 °C for 1 min, for C<sub>16:0</sub>, raised to 200 °C, 250 °C at a rate of 10 °C min<sup>-1</sup>, for 1 min for C<sub>8:0</sub> and C<sub>12:0</sub>, respectively and to 300 °C at a rate of 8.50 °C min<sup>-1</sup> for 1 min in the case of C<sub>16:0</sub>, held isotherm for 1 min, and then raised to 300 °C at 65 °C min<sup>-1</sup> for C<sub>8:0</sub>, to 310 °C at 65 °C min<sup>-1</sup> for C<sub>12:0</sub> and C<sub>16:0</sub>.

## 11. Galactose oxidase screening assay

The colorimetric galactose oxidase (GOase) assay was monitored at 420 nm following the oxidation of the dye ABTS. Prior to the assay, the samples (supernatants from whole resting cell biotransformations) were heated at 90 °C for 30 minutes in order to ensure the incapacity of potential metabolic activity during the screening assay. The screening assay was performed in 96-well plates (Greiner Bio-One GmbH, Frickenhausen, Germany). To each well was supplemented: 118 µL of the heated supernatant, 1 µL the commercially available horseradish peroxidase (HRP, 900 units µL<sup>-1</sup> stock), 8 µL 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS dye from 10 mM stock), 10 µL purified GOase (from stock solution 3.5-8 µM mL<sup>-1</sup>) and 100 mM potassium phosphate buffer pH 7.4 for a final volume of 200 µL. The increase of absorption of the oxidized ABTS was monitored on plate reader λ = 420 nm (POLARstar Omega BMG Labtech, Ortenberg, Germany) for several hours (3-6 hours).



## Results

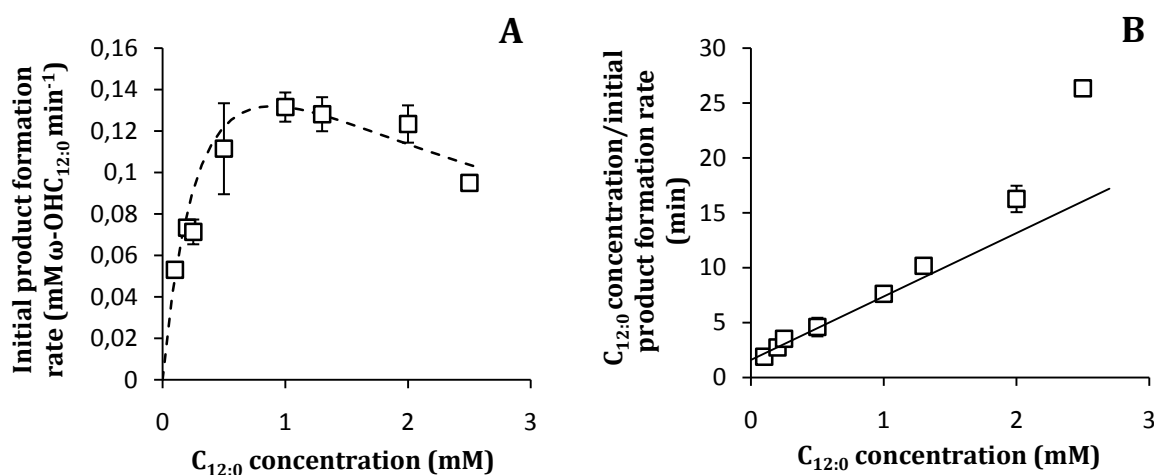
### Chapter 1: Identification of bottlenecks in CPR2<sub>mut</sub>-catalysed reactions: how to address the limitations and improve whole-cell biocatalysis

A large part of this chapter was recently published in the journal *Applied Microbiology and Biotechnology* (publication I).

#### 1. Substrate inhibition

##### 1.1. Enzymatic reactions

*In vitro* reactions using purified CPR2<sub>mut</sub> have been performed to determine potential substrate inhibition (Figure 7). Increasing concentrations of C<sub>12:0</sub> in the range of [0.1-2.5 mM] have been assayed in a two-minute reaction, using 1  $\mu$ M of P450 enzyme. Above 1 mM concentration we can observe an inhibitory effect of the substrate applied resulting in a decrease in the rate of reaction. The reciprocal plot also suggests the substrate inhibition, high substrate concentrations (higher than 1 mM) deviate from the regression line (solid line).<sup>[155]</sup> The potential kinetics constants determined are  $K_M$  0.4,  $V_{max}$  0.25 and  $K_S$  2.



**Figure 7.** Substrate dependence of initial product formation rate catalysed by 1  $\mu$ M of purified CPR2<sub>mut</sub> (A). Trendline is indicated with a dashed line. Reciprocal plot for detection of substrate inhibition (B). Inhibitory effects are observed for concentrations of C<sub>12:0</sub> above 1 mM.

##### 1.2. Whole resting cells in shaking flasks

Substrate inhibition was also tested *in vivo*, using increased concentrations of C<sub>12:0</sub> in the range of [0.5-3 mM] applied to *E. coli* whole resting cells, CPR2<sub>mut</sub> expressed (Table 7). As

previously observed *in vitro*, the initial rate of reaction (evaluated after 2 h) decreased in the presence of large quantities of substrate provided (beyond 2 mM concentration).

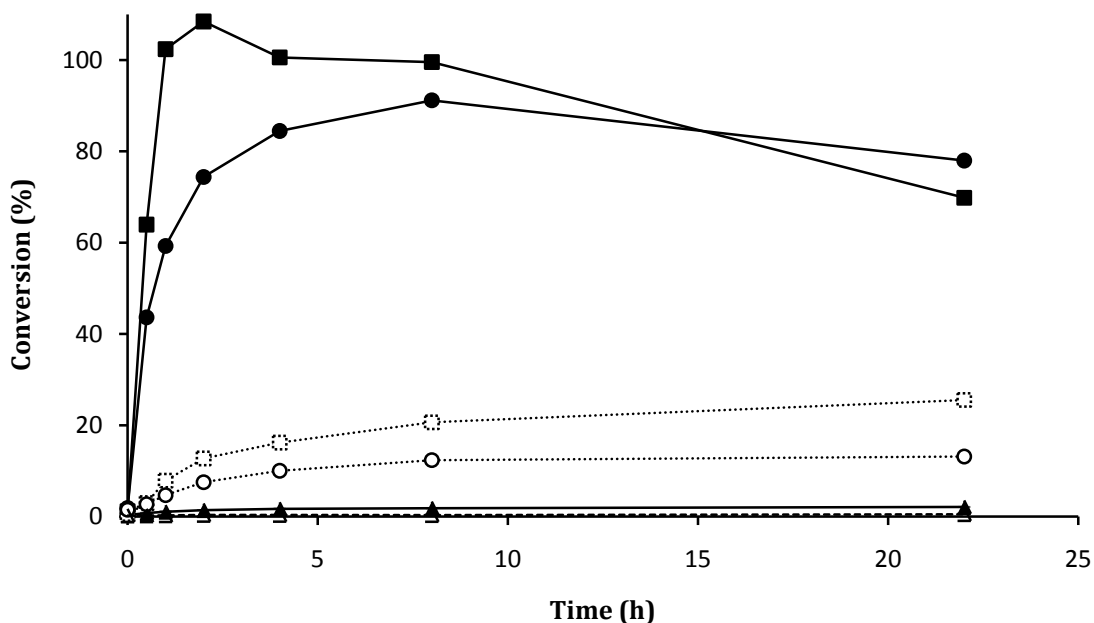
**Table 7.** Evaluation of the initial rate of product formed along a whole resting cell biotransformation for increased concentrations of C<sub>12:0</sub> provided at the start of the reaction (5 % DMSO (v/v))

C <sub>12:0</sub> provided at the start of the reaction (mM)	Initial product formation rate (mM $\omega$ -OHC <sub>12:0</sub> h <sup>-1</sup> )
0.5	0.29 ± 0.07
1	0.31 ± 0.09
1.2	0.33 ± 0.09
<b>2</b>	<b>0.39 ± 0.04</b>
2.5	0.32 ± 0.19
3	0.28 ± 0.02

### 1.3. *In vivo* biotransformations in fermenter

The substrate inhibition observed at the enzyme level and along whole-cell biotransformations in shaking flasks was tested *in vivo* at larger scale as well (Figure 8). 5, 10 or 15 mM of substrate C<sub>12:0</sub> were provided as starting material to *E. coli* resting cells in 1 L bioreactor and the biotransformation was performed for 22 h. The fastest rate of reaction can be obtained for 5 mM of C<sub>12:0</sub> supplemented: 95 % of the substrate is converted after 1 h only and the highest conversion of substrate can be obtained after 2 hours of biotransformation. In the case of 10 mM of initial concentration of C<sub>12:0</sub>, the initial rates were equivalent to those observed for 5 mM, however the rate of reaction was shown to slow down after 50 % of the substrate consumed (corresponding to 5 mM) and no complete conversion could be accomplished, suggesting an inhibitory effect of the product  $\omega$ -OHC<sub>12:0</sub>. The application of 15 mM of substrate displayed a complete inhibitory effect as no product was formed along the biotransformation. For 5 mM and 10 mM of substrate, after 8 h of biotransformation, the concentration of the product accumulated decreased over time and this disappearance can be partially explained by an over-oxidation of  $\omega$ -OHC<sub>12:0</sub> into  $\alpha,\omega$ -dicarboxylic acid ( $\alpha,\omega$ -DC<sub>12:0</sub>).





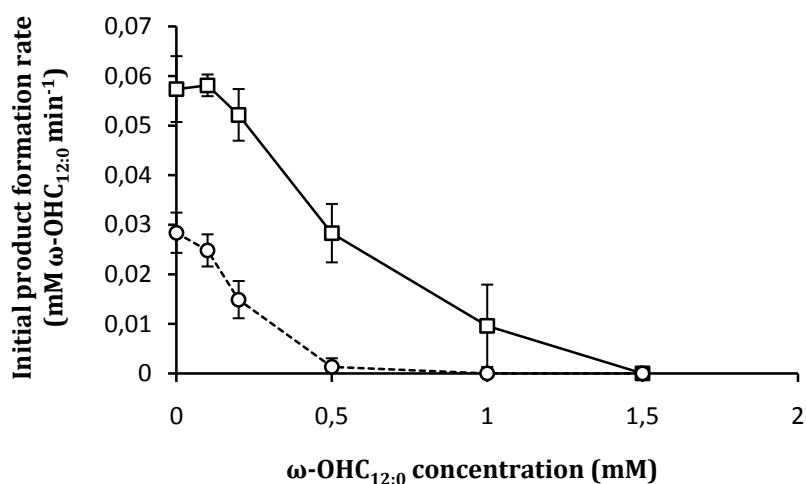
**Figure 8.** Conversion of  $C_{12:0}$  by  $CPR2_{mut}$  into  $\omega\text{-OHC}_{12:0}$  and over-oxidation to the corresponding  $\alpha,\omega\text{-DC}_{12:0}$  in bioreactor with increased substrate concentrations provided (5 mM (boxes), 10 mM (circles) and 15 mM (triangles)). Solid lines represent  $\omega\text{-OHC}_{12:0}$  and dashed lines represent dicarboxylic acid. Experimental conditions:  $9 \text{ g}_{cdw} \text{ L}^{-1}$ , 100 mM potassium phosphate buffer pH 7.4. The substrate was dissolved in DMSO and applied using a final DMSO concentration of 5 % (v/v).

## 2. Product inhibition

As the results from the substrate inhibition investigations at different scales suggest an inhibitory effect of  $\omega\text{-OHC}_{12:0}$ , similarly experiments have been performed using a constant concentration of substrate and variations in the concentration of the product  $\omega\text{-OHC}_{12:0}$  provided.

### 2.1. *In vitro* reactions

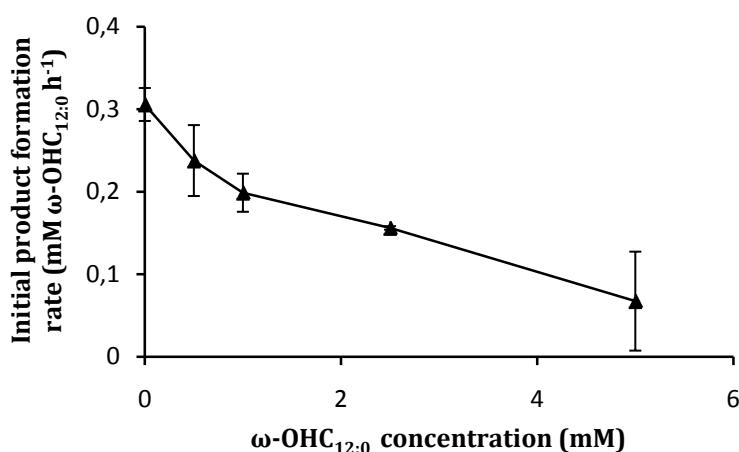
Enzymatic reactions have been performed with  $0.5 \mu\text{M}$  and  $1 \mu\text{M}$  of purified  $CPR2_{mut}$ , using a constant concentration of substrate of 1 mM, and initial concentrations of  $\omega\text{-OHC}_{12:0}$  comprised between [0-1.5 mM] for a two-minute reaction course (Figure 9). The rate of reaction slows down and no activity can be detected at 1 mM and 1.5 mM of  $\omega\text{-OHC}_{12:0}$  provided for the two concentrations of enzyme  $0.5 \mu\text{M}$  and  $1 \mu\text{M}$ , respectively. The product inhibition was also dependent on the enzyme concentration as the negative impact appeared at lower product concentration for lower concentration of biocatalyst ( $0.5 \mu\text{M}$ ) compared to the concentration of  $1 \mu\text{M}$  of enzyme.



**Figure 9.** Influence of increased initial product concentrations [0-1.5 mM] on initial product formation rates performed *in vitro* and initiated by the addition of 1 mM of C<sub>12:0</sub> (10 % DMSO (v/v)). Boxes represent 1  $\mu$ M of enzyme and circles 0.5  $\mu$ M of enzyme tested.

## 2.2. *In vivo* reactions

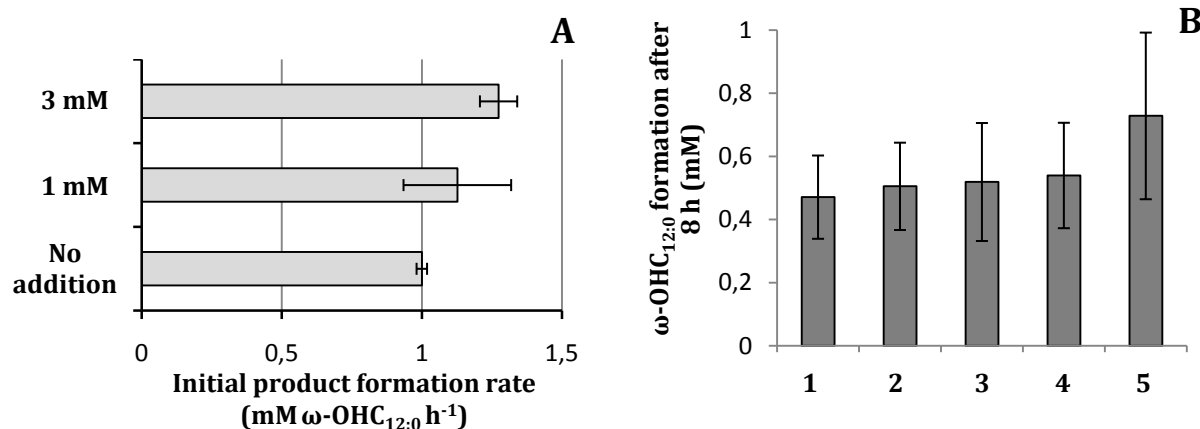
The product inhibition detected *in vitro* has been evaluated *in vivo* in shaking flasks with *E. coli* ( $\Delta$ *fadD*) resting cells, in which the beta-oxidation pathway of fatty acids consumption is blocked. 1.5 mM of C<sub>12:0</sub> was used as a constant substrate concentration provided at the start of the reaction and increased concentrations of  $\omega$ -OHC<sub>12:0</sub> have been tested (up to 5 mM, 10 % DMSO (v/v)) (Figure 10). The fastest rate of reaction was obtained when no product was added at the start of the reaction. Product inhibition occurring at the enzyme level *in vitro* was also demonstrated during whole-cell catalysis.



**Figure 10.** Product inhibition profile after 2 h biotransformations *in vivo* using 1.5 mM C<sub>12:0</sub> and increased initial product concentrations, performed in shake flasks with 50 g<sub>cww</sub> L<sup>-1</sup> resting *E. coli* cells. The substrate was dissolved in DMSO and applied using a final concentration of 5 % (v/v).

### 3. Cofactor regeneration limited

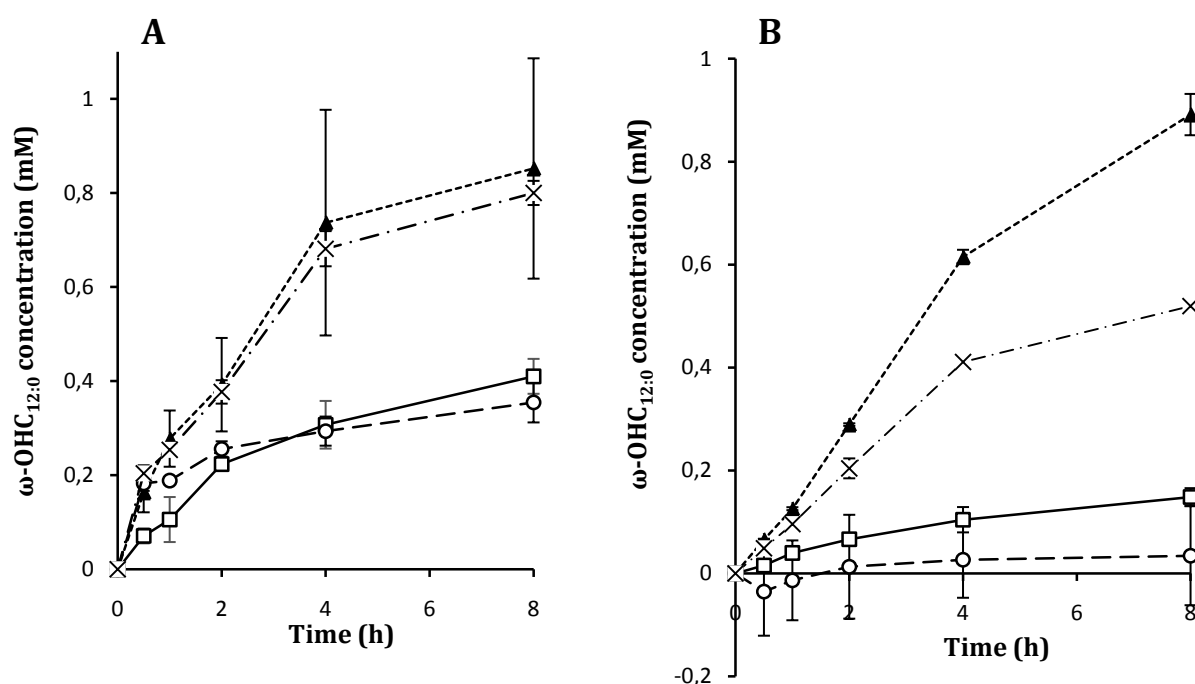
The regeneration of NADPH cofactor in whole-cells has been suggested to be limited.<sup>[116,146,156]</sup> Studies on the potential limited cofactor available along a biotransformation has been investigated. The addition of external cofactor was tested along 8 hours reaction: stoichiometric amount of NADPH (1.5 mM and twice the amount, 3 mM) have been supplemented at the beginning of the biotransformation started by 1.5 mM of C<sub>12:0</sub> compared to a control reaction (no NADPH addition) in shaking flasks with 50 g<sub>cww</sub> L<sup>-1</sup> *E. coli* resting cell (Figure 11A). According to the estimation of the initial rates, the efficiency of biotransformation was improved with the addition of increased amount of cofactor and the highest rate was obtained in the case of 3 mM of NADPH supplemented. The potential consumption of cofactor for other cellular metabolisms and/or the ineffective coupling efficiency, eventually led to a limited NADPH amount available for product formation. The large expenses attributed to the NAD(P)H cofactors have driven efforts in the generation of chemical alternatives. Numbers of cofactor analogues have been synthesized and have successfully replaced the natural one.<sup>[157,158]</sup> In our study, a similar experiment involving the use of an artificial cofactor, nicotinamide-3,4-chloromethoxy-phenyl-hydride (termed NACMH, Appendix A3 (Dr. Sebastian Löw PhD thesis, ITB, University of Stuttgart)) instead of NADPH confirmed the previous results of limited availability of NADPH to reach the maximum concentration of product (Figure 11B). Moreover the external addition of artificial cofactor demonstrates a practical and economically feasible approach to overcome the limited regeneration of the natural one.



**Figure 11.** Influence of cofactor addition on biotransformations with *E. coli* resting cells in shake flasks. 1.5 mM C<sub>12:0</sub> were used and supplementary NADPH (A) provided at the start of the reaction (1.5 mM or 3 mM). Addition of NACMH, synthetic cofactor (B), 1.5 mM provided at the start (2), addition of 1.5 mM at the beginning and after 2 h reaction (3), addition of cofactor 1.5 mM at the start, after 2 h and 4 h (4), addition of 3 mM at the beginning of the reaction (5) in comparison with a control reaction without providing extra cofactor analogue (1).

#### 4. Transport limitation

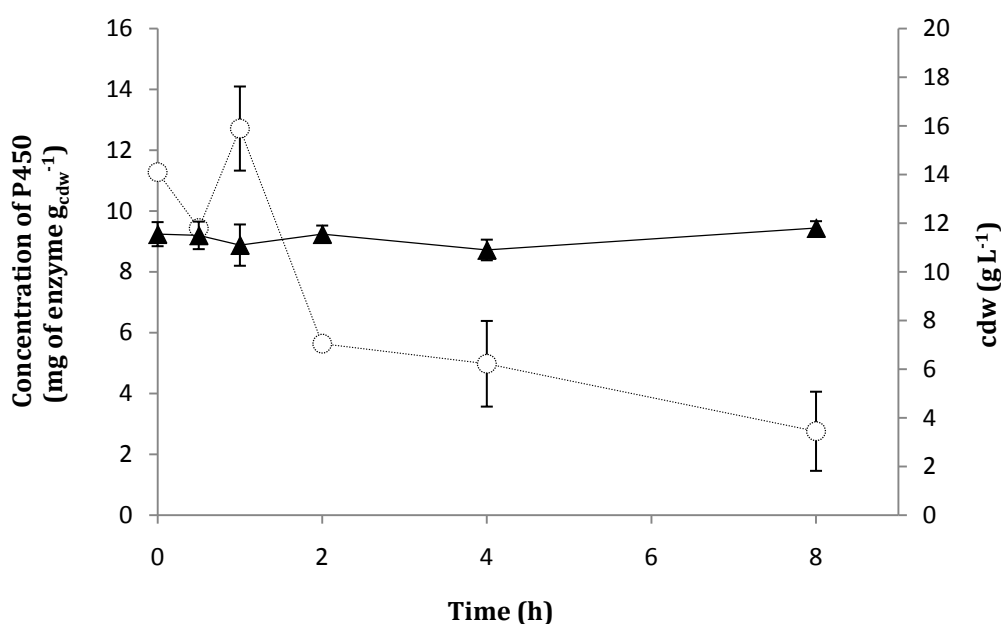
Previous studies have shown the benefit of using a permeabilization technique to enhance the substrate transport across the cell membrane.<sup>[159]</sup> Therefore the potential limitation of fatty acid substrate uptake/product release across the cells membrane was investigated (Figure 12). Permeabilization methods including sonication, acetone treatment and freezing/thawing on  $50 \text{ g}_{\text{cww}} \text{ L}^{-1}$  *E. coli* resting cells were tested during a biotransformation of 1.5 mM of substrate  $\text{C}_{12:0}$  compared to a control batch (no treatment applied to the cells). Two biotransformations in parallel were studied, comparing the stoichiometric addition of NADPH cofactor added at the beginning of the reaction and no external cofactor supplemented. By using  $\text{C}_{12:0}$  substrate, the maximum concentration, 0.9 mM of  $\omega\text{-OHC}_{12:0}$  within 8 hours, was achieved with the intact cells batch. No significant difference was observed between the flasks with addition of cofactor and the batch without, suggesting that the reaction was not limited by a potential lack of NADPH available. These results reveal no limitation of transport of substrate/product of medium chain-length fatty acids across the cell membrane of *E. coli*.



**Figure 12.** Progress curves of biotransformations of 1.5 mM  $\text{C}_{12:0}$  in shake flasks after different permeabilization methods of the cell membrane. The experiments were performed in the presence (A) as well as in the absence (B) of stoichiometric amounts of NADPH cofactor (1.5 mM) supplemented at the beginning of the reaction. Control without treatment (triangles), sonication (cross), acetone treatment (boxes) and freezing followed by thawing (circles).

## 5. P450 instability

A main limitation identified in a P450-based reaction is the decline of the biocatalyst concentration along a whole-cell biotransformation. The protein content decreasing will lead to a limiting rate of reaction and will prevent from the completion of conversion of  $C_{12:0}$  into  $\omega$ -OHC $_{12:0}$ . The decrease in the P450 concentration is significantly higher during a biotransformation in fermenter compared to *in vivo* reaction in shaking flasks. In a 1 L bioreactor a loss of 50 % of biocatalyst (from 12.7 mg of enzyme  $g_{cdw}^{-1}$  to 5 mg enzyme  $g_{cdw}^{-1}$ ) between 2 and 4 hours of reaction course can be observed for a constant amount of resting cells, cell dry weight evaluated (Figure 13).

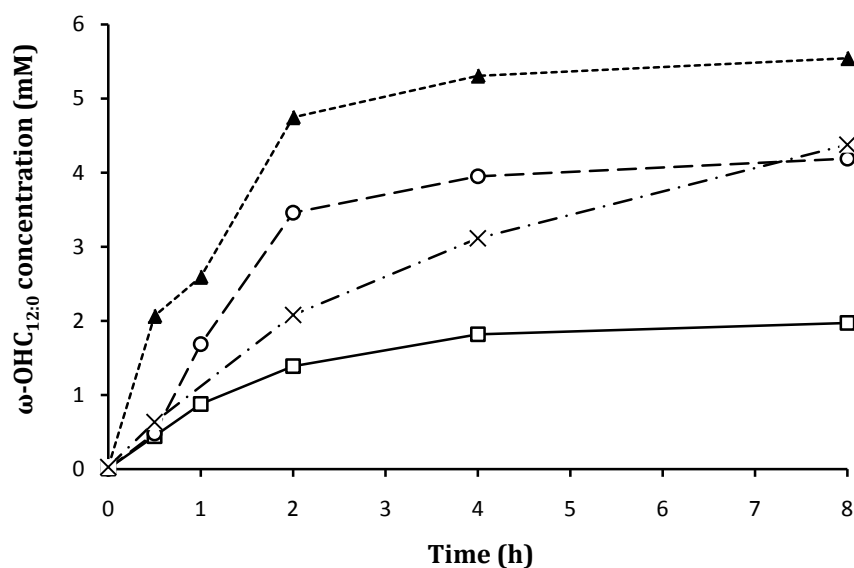


**Figure 13.** Constant cell dry weight (solid line, triangles) and decrease of the P450 concentration (dashed line, circles) along 8-hour whole-cell biotransformation of 10 mM of  $C_{12:0}$  in 1 L fermenter. Reaction conditions: 9  $g_{cdw} L^{-1}$ , 100 mM potassium phosphate buffer pH 7.4, 30 °C,  $pO_2$  30 %.

## 6. Feeding strategies

The substrate/product inhibition, the low P450 stability and the limited NADPH regeneration have been identified as limitations in the studied P450-catalysed reaction. To address present bottlenecks, substrate feeding strategies have been applied at larger scale (in 1 L stirred tank fermenter) (Figure 14). The starting concentration of substrate provided was 10 mM, dissolved in 5 % DMSO (v/v) at the start of the reaction as common batch system, loaded along the biotransformation via a pump system (fed-batch) or delivered as solid form with intact

cells and cells pre-treated with DMSO overnight. The fastest rate of reaction and the most efficient bioconversion of  $C_{12:0}$  into  $\omega$ -OHC $_{12:0}$  were reached by the application of the solid substrate feeding. Supplementing the solid substrate to cells pre-treated with 5 % DMSO reduced the efficiency of biotransformation signifying a negative impact of the addition of the co-solvent, potentially damaging the cells.<sup>[160]</sup> This approach of solid feeding allows a slow and continuous dissolving of the fatty acids along the biotransformation preventing from substrate and product inhibition limitations and enables the avoidance of harmful co-solvent.



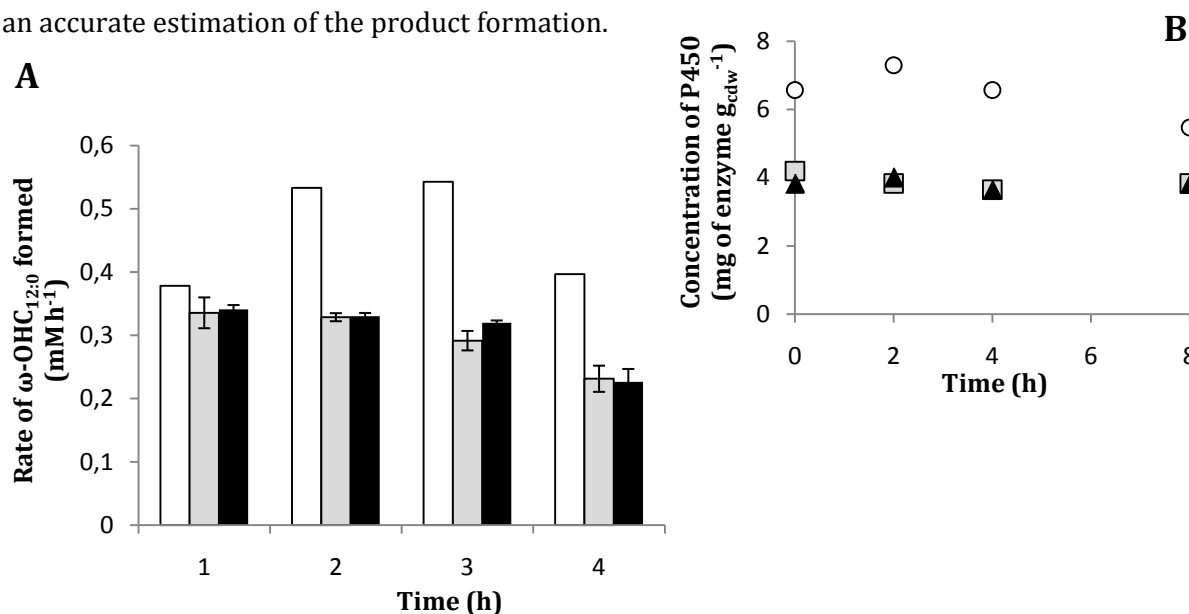
**Figure 14.** Progress curves of biotransformations performed in 1 L bioreactor. Batch reaction with 10 mM  $C_{12:0}$  dissolved in DMSO prior to addition (boxes, 5 % end concentration DMSO (v/v)), fed-batch reaction with  $C_{12:0}$  dissolved in DMSO (5 % end concentration (v/v)) and applied with a linear feed during the first 5 h of reaction (cross), solid application of 10 mM  $C_{12:0}$  to cells treated with 5 % DMSO overnight (v/v), centrifuged and resuspended in potassium phosphate buffer pH 7.4 (circles) and 10 mM solid  $C_{12:0}$  applied to non-DMSO treated cells (triangles). Reaction conditions: 9 g<sub>cdw</sub> L<sup>-1</sup>, 100 mM potassium phosphate buffer pH 7.4, 30 °C, pO<sub>2</sub> 30 %.

## 7. Standard conditions required to improve a process

To enhance the yield of  $\omega$ -OHC $_{12:0}$  and as a first step towards an industrial process, the model reaction was evaluated in a stirred tank bioreactor under essential conditions. Up-scaling is a key step, requiring the knowledge and the control of many parameters intended to improve the global process. The relevant parameters as agitation, oxygen supply, or carbon source have to be explored as well as an evaluation of the protein expression reproducibility, the choice of the bacterial strain and the amount of resting cells required for efficient biotransformation.

### 7.1. Protein expression rate

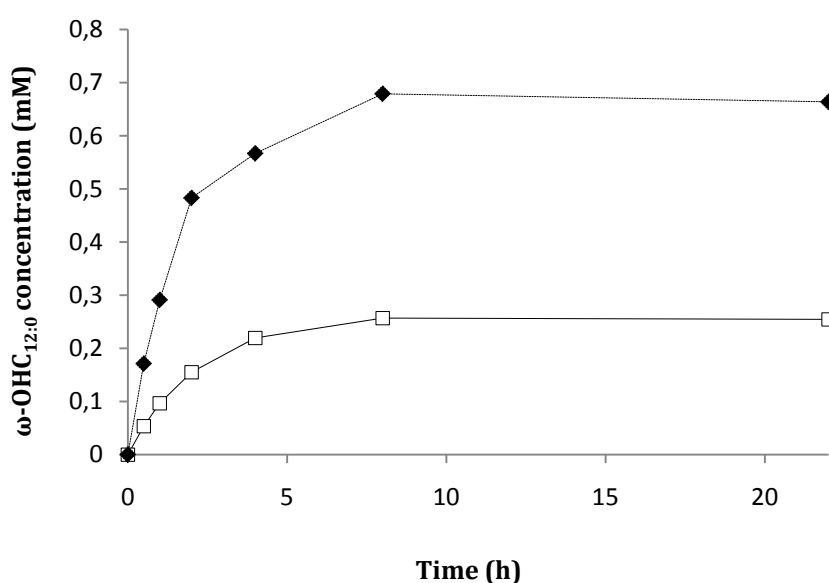
The efficiency of biotransformation was shown variable from one batch of fermentation to another one. Scale changing affects the biotransformation efficacy: under identical conditions, the rates and yields of product formed are enhanced in the stirred tank compared to shaking flasks reactions. The solubility limit of substrate provided was reached at 5 mM of  $C_{12:0}$  (5 % DMSO (v/v)) in shaking flasks, whereas up to 10 mM in fermenter can be examined (by excluding substrate inhibition possibility). Also, the benefit observed from the substrate solid feeding strategy in fermenter is unfortunately not applicable in shaking flasks as this requires high mixing conditions. Besides differences observed between various scales, the biotransformation variations depend on the efficiency of protein expression as the biocatalysis reactions in whole resting cells are performed based on the cell density ( $c_{ww}/cdw$ ). The figure 15 displayed the biotransformation of 1.5 mM  $C_{12:0}$  into  $\omega$ -OHC $_{12:0}$  using 50  $g_{cww} L^{-1}$  whole-cells carrying CPR2<sub>mut</sub> from three different fermentation batches (different overnight cultures) in shaking flasks. Only equivalent P450 expression batches led to similar biotransformation of substrate (black and grey data). Based on the cell density, the biotransformation efficacy is highly dependent on the rate of P450 expression. This observation made at small scale regarding the protein expression has to be taken in consideration prior any evaluation at larger scale for an accurate estimation of the product formation.



**Figure 15.** Variable efficiency of biotransformation of 1.5 mM of  $C_{12:0}$  between batches of cells expression (A) depending on the protein content (B). The two cell batches coloured in grey and black have similar protein concentration leading to comparable biotransformation rates in comparison to the third batch (in white) with a greater P450 content (white circles) and higher rates of product formation displayed. Error bars for the “black” and “grey” batches represent standard deviation of duplicates.

## 7.2. Bacterial strain

The expression rate was shown variable from one batch to another one with the bacterial strain being a crucial element. *E. coli* BL21 (DE3) expressing well pET28a(+) vector at small scale in shaking flasks is no longer the most suitable strain for protein expression in fermenter (Figure 16). The *E. coli* strain HMS174 (DE3) showed greater biotransformation efficiency of 10 mM  $C_{12:0}$  provided than BL21 (DE3). This can be related to the effective protein expression in *E. coli* HMS174 (DE3) strain due to of a higher plasmid stability in fermenter, as previously demonstrated elsewhere.<sup>[149]</sup>

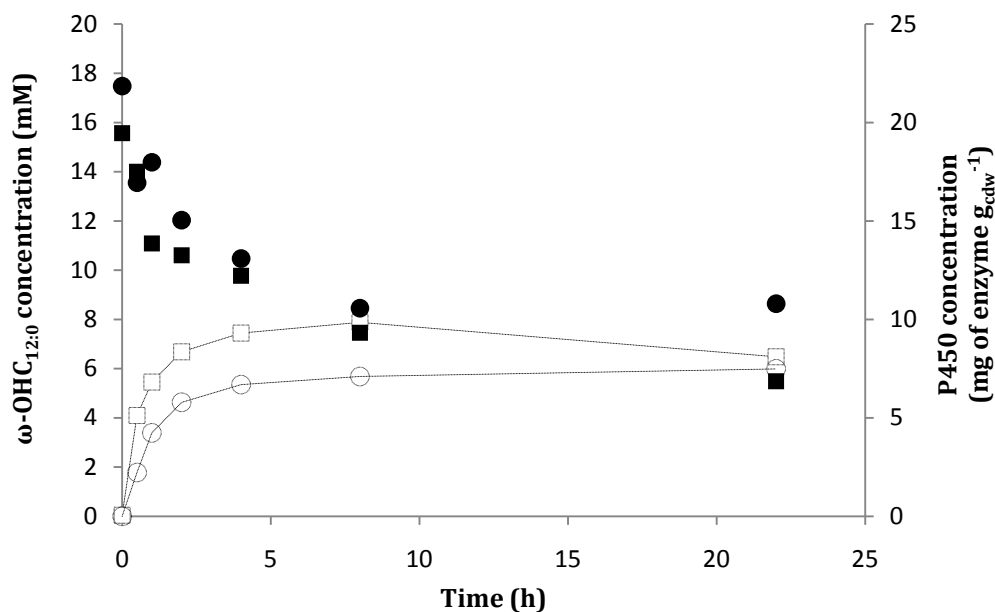


**Figure 16.** Evaluation of biotransformation efficiency of 10 mM  $C_{12:0}$  between *E. coli* strains HMS174 (DE3) (dashed line, black diamonds) vs. BL21 (DE3) (solid line, white boxes). HMS174 (DE3) possesses a greater plasmid stability leading to improved conversion of substrate. Reaction conditions: resting cells 9 g<sub>cdw</sub> L<sup>-1</sup>, 100 mM potassium phosphate buffer pH 7.4, 30 °C, pO<sub>2</sub> 30 %.

## 7.3. Biotransformation efficiency

Evaluation of the amount of resting cells required for an efficient biotransformation was also investigated in 1 L fermenter. 200 g<sub>cww</sub> L<sup>-1</sup> were set up during a biotransformation of 10 mM of  $C_{12:0}$  for 22 h and compared to the usual procedure of 50 g<sub>cww</sub> L<sup>-1</sup> also used in shaking flasks (Figure 17). Supplemental addition of whole resting cells did not increase the rate of reaction or the concentration of product formed. This can be explained by the similar P450 concentration detected leading to comparable biotransformation efficiency.





**Figure 17.** Estimation of the efficiency of biotransformation in the presence of a larger amount of resting cells, 200 g<sub>cdw</sub> L<sup>-1</sup> (white rounds) in comparison to the standard procedure 50 g<sub>cdw</sub> L<sup>-1</sup> (white boxes) and evaluation of the P450 concentration in both cases (black data). Reaction conditions: *E. coli* resting cells (50 g<sub>cdw</sub> L<sup>-1</sup> or 200 g<sub>cdw</sub> L<sup>-1</sup>), 100 mM potassium phosphate buffer pH 7.4, at 30 °C, pO<sub>2</sub> 30 %.



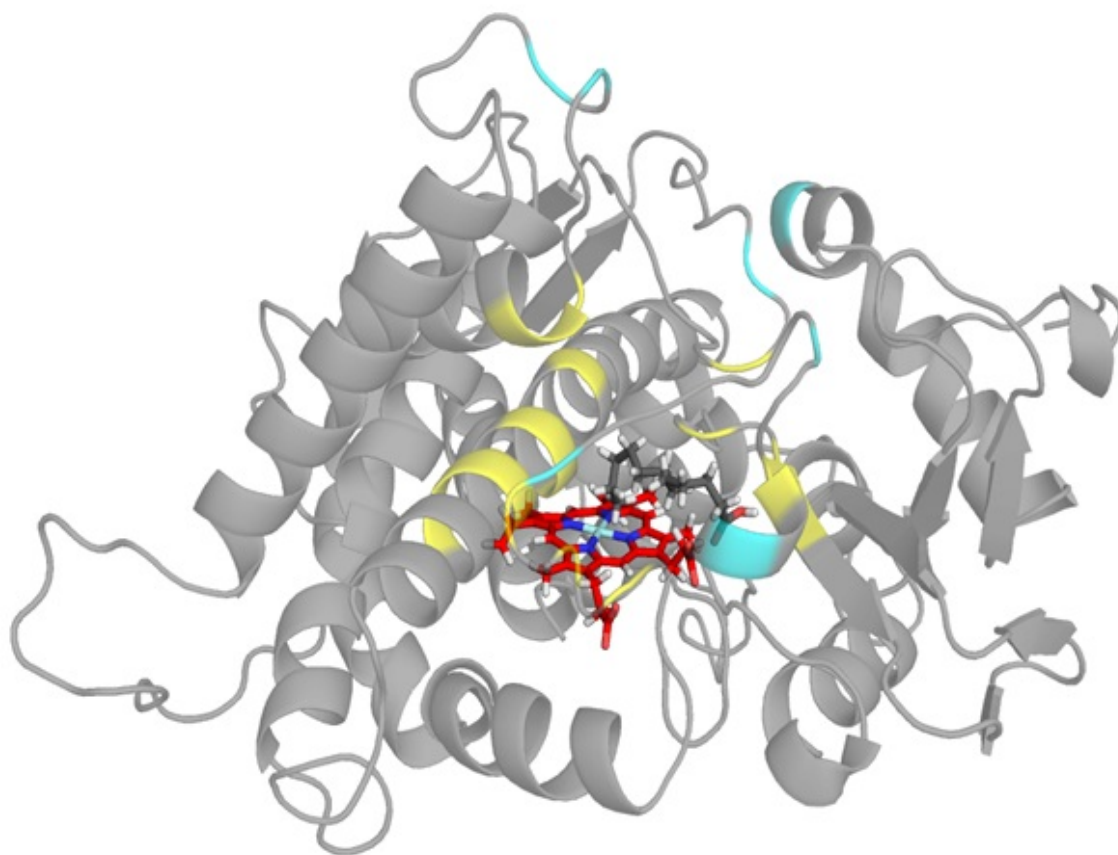
## Chapter 2: Generation of mutant libraries and screening for improved activity

The preparation of the publication **II** is based on the work entirely presented in this chapter.

### 1. Library generation

#### 1.1. Homology model

The choice of adequate positions targeted for mutagenesis was based on the creation of a homology model of CYP153A from *Marinobacter aquaeolei* (CYP153A<sub>M.aq.</sub>), using the crystal structure of CYP153A7 from *Sphingopyxis macrogoltabida* as template (Dr. Łukasz Gricman investigations, ITB, University of Stuttgart). The model substrate dodecanoic acid (C<sub>12:0</sub>) was docked into the model and the positions having most likely interactions with the substrate were selected for mutations (Figure 18). Thus, 6 positions located at the substrate entrance tunnel and 17 positions from the binding pocket were targeted for mutation, coloured in blue and in yellow, respectively.



**Figure 18.** Homology model of CYP153A<sub>M.aq.</sub> and C<sub>12:0</sub>, model substrate docked inside (in dark grey). Positions closed to the heme domain (red) in the binding pockets (yellow) and positions located at the substrate entrance channel (blue) potentially interacting with the substrate for terminal hydroxylation conformation were selected for mutations.

## 1.2. Positions and residues selected for mutations

The number and the nature of the residues chosen for substitution were based on amino acid conservation analysis after sequence alignments between CYP153A<sub>M.aq.</sub> and prokaryotic and eukaryotic P450s sequences (Dr. Łukasz Gricman, ITB, University of Stuttgart). The amino acids occurring the most frequently for each position selected, from the substrate entrance tunnel and from the binding pocket, were taken into consideration (Table 8, p65).<sup>[84,152]</sup> To increase the diversity, the strategy of using a degenerated codon to exchange between 2 and 5 amino acids was chosen considering the positions located at the substrate entrance channel (yellow squares).<sup>[153]</sup> Site-directed mutagenesis was a preferred method to generate mutants in the case of positions located in the active site and only amino acids with the greatest percentage of frequency were taken into account for substitution. The positions located within the active site with percentages marked in blue colour represent substitution already performed during the generation of a small focused library.<sup>[128]</sup> The most active mutant from this family identified and described so far to accomplish the terminal hydroxylation of fatty acid was the variant G307A, position from the binding pocket.

## 1.3. Substrate entrance tunnel

### 1.3.1. Degenerated codon

In order to expand the diversity, previously limited to a focused library targeting residues in the binding pocket only, positions located at the substrate entrance tunnel were selected and mutations were generated through the design and the application of a degenerated codon using CASTER 2.0 tool.<sup>[153]</sup> Due to the degeneracy of the genetic code, a large number of colonies were picked for the screening of each selected position by taking in consideration the theoretical number of clones to test for full coverage of amino acids substitutions (Table 9).

**Table 9.** Theoretical number of colonies required to be screened for a full coverage of amino acid substituted

Positions	Expected amino acids	Number of colonies screened	
		Theory for full coverage	Actual number screened
R77	L, P, V, A	10	> 20
D134	V, G	10	> 20
S140	F, L, C, W, R	22	> 30
S233	P, A, R, G	34	> 40
T235	L, V, R, G	22	> 30
S453	A, G	10	12

**Table 8.** Positions and residues selected for substitution based on amino acid conservation within prokaryotic and eukaryotic CYPs. Percentage of frequency for each amino acid at each selected positions. The positions coloured in yellow are located in the substrate entrance tunnel, any other positions are located into the binding pocket. The percentages coloured in blue correspond to substitutions previously tested and described somewhere else.<sup>[128]</sup>

Positions		Amino acids																			Substitutions	
		A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W		Y
77	R	8,18	2,12	10,20	5,16	6,36	6,46	2,17	3,60	1,51	7,57	2,01	2,46	11,23	3,52	5,85	4,45	6,51	3,44	1,30	4,82	L, P, V, A
129	Q	6,87	0,42	5,84	7,86	1,69	9,16	2,50	3,75	6,50	3,98	1,35	3,06	10,13	3,37	8,21	6,97	4,34	4,64	0,81	1,52	P
131	I	11,03	0,44	3,19	4,38	1,58	5,91	2,28	2,27	3,73	4,62	1,12	2,03	13,61	1,37	3,51	7,24	3,65	4,94	1,36	10,20	P
134	D	2,70	0,11	2,52	2,54	1,03	4,21	0,42	3,15	1,35	1,64	1,17	1,19	4,95	0,42	1,69	2,83	2,15	3,18	0,11	0,95	V, G
140	S	3,82	1,24	1,89	0,80	7,72	4,23	2,86	11,33	0,81	18,87	5,42	0,85	1,44	1,62	8,65	2,06	2,39	11,28	3,48	2,63	F, L, C, W, R
141	V	4,07	0,71	3,30	2,73	1,61	14,94	1,28	0,36	2,17	4,52	1,09	2,40	4,99	2,00	3,25	4,68	1,72	2,22	0,57	3,05	G
143	M	6,51	0,29	11,15	1,78	2,64	34,89	1,79	1,06	1,51	0,55	1,96	6,36	0,65	1,68	2,63	14,00	3,88	0,76	0,55	0,64	G, S
145	I	9,20	0,15	0,06	0,02	13,09	2,97	0,10	17,74	0,01	24,40	2,53	0,81	0,25	1,00	0,13	3,96	3,42	14,86	0,12	0,89	V, L
228	M	9,64	0,70	0,32	0,21	10,75	6,71	1,02	10,60	0,31	18,91	10,55	4,34	0,40	1,43	1,83	4,46	3,62	8,24	0,66	3,19	L
229	A	4,66	0,92	3,67	0,37	20,42	4,18	0,18	7,40	0,43	19,85	6,72	1,56	0,42	1,55	0,68	4,98	6,04	11,92	0,46	1,45	F, L
233	S	2,03	0,17	4,05	2,69	3,08	4,24	1,80	1,39	1,99	2,31	2,00	2,63	8,01	2,52	4,36	3,19	1,57	1,71	0,81	1,15	P, A, R, G
235	T	3,02	0,00	3,55	3,63	2,52	5,35	0,45	1,88	2,54	8,10	3,84	2,30	2,15	0,48	2,89	2,22	3,15	2,89	0,19	1,72	L, V, R, G
241	D	2,73	0,03	3,10	1,27	1,43	4,34	0,16	2,20	0,53	7,49	0,61	3,65	9,53	1,11	4,45	4,34	6,65	2,38	1,48	1,09	L, P, I, M, T
301	L	16,29	9,63	0,02	0,04	2,45	2,29	0,61	12,61	0,01	17,65	6,49	0,21	0,04	0,05	0,01	4,54	4,68	20,50	0,08	0,63	A, V
302	T	5,05	0,54	5,64	1,24	10,46	2,97	0,56	14,33	3,41	22,51	6,60	2,43	0,27	1,51	1,51	4,03	4,96	8,26	1,61	0,91	L, I
303	L	4,43	0,06	19,65	7,85	2,03	4,65	0,13	5,24	0,03	13,90	1,08	9,25	0,07	0,31	0,03	5,83	17,59	6,51	0,01	0,15	D, T
306	V	8,30	0,05	0,15	0,48	21,33	6,76	0,15	14,98	0,09	13,92	5,41	1,97	0,17	0,30	0,73	4,26	5,02	12,66	1,44	0,62	F, I
307	G	73,95	0,02	0,79	4,90	0,01	14,28	0,02	0,10	0,04	0,23	0,04	0,08	1,23	0,16	0,71	0,75	0,72	0,77	0,01	0,01	R
310	D	1,56	0,01	41,01	44,70	0,17	1,35	4,05	0,22	0,17	0,24	0,16	0,82	0,01	0,89	0,05	0,94	0,37	1,20	0,00	0,04	E
354	L	19,49	0,04	0,09	0,05	1,40	6,76	0,21	11,34	0,03	10,91	0,20	0,77	0,91	0,12	0,27	4,81	5,04	37,13	0,04	0,04	A, V
356	Y	5,45	0,28	0,14	0,94	15,70	6,43	2,66	5,45	1,20	21,56	6,71	1,46	1,08	5,05	6,42	6,09	3,37	6,29	0,37	2,76	F, L
357	M	4,24	0,05	2,38	0,47	7,89	8,22	0,37	10,16	0,18	24,99	5,01	8,51	1,00	2,11	1,01	4,87	8,16	8,04	0,81	0,59	L
358	R	8,12	0,30	2,76	2,47	7,16	6,38	0,79	2,63	1,75	4,46	5,46	2,66	29,42	2,07	4,63	7,89	5,75	2,53	0,36	2,03	P
453	S	4,33	0,16	3,19	3,01	0,36	21,01	1,01	0,90	3,81	1,16	1,48	2,29	0,28	1,82	2,62	9,49	5,58	2,45	0,09	0,31	A, G
455	F	3,21	0,19	0,09	0,08	15,70	4,80	0,51	7,69	0,13	20,79	3,85	1,67	2,82	0,30	2,31	2,98	3,16	6,73	0,21	0,54	V

### 1.3.2. Selection through *in vivo* tests

The screening of a large number of colonies issued from the degenerated codons was achieved at small scale (24 deep-well plates, 2 mL final volume for protein expression) with 50 g<sub>cww</sub> L<sup>-1</sup> whole resting cell biotransformation of 2 mM C<sub>12:0</sub> performed for 2 hours, and the samples were extracted and quantified by GC-FID. On one hand, this method enables to rapidly identify the well-expressed and active variants with the first identification based on the brown colour of the pellet suggesting a successful protein expression (due the iron-containing heme). On the other hand, as the biotransformation is normalized by the cell wet weight, a low protein expression rate will automatically lead to low biotransformation efficiency. As the reproducibility of protein expression can be a limiting factor, the identification of active variants after whole-cell biotransformation is biased by the identification of high protein expression not reflecting the activity of the mutants. Therefore, subsequently to the whole-cell screening, the best hits after GC-FID analysis were sent for sequencing before expression at larger scale and evaluation during *in vitro* reaction tests using a normalized concentration of enzyme (Table 10 and following section).

**Table 10.** Identification of the most active variants (positions located in the substrate access channel) based on whole-cell based-reactions, product quantification via GC-FID and comparison to the wild type enzyme and G307A mutant.

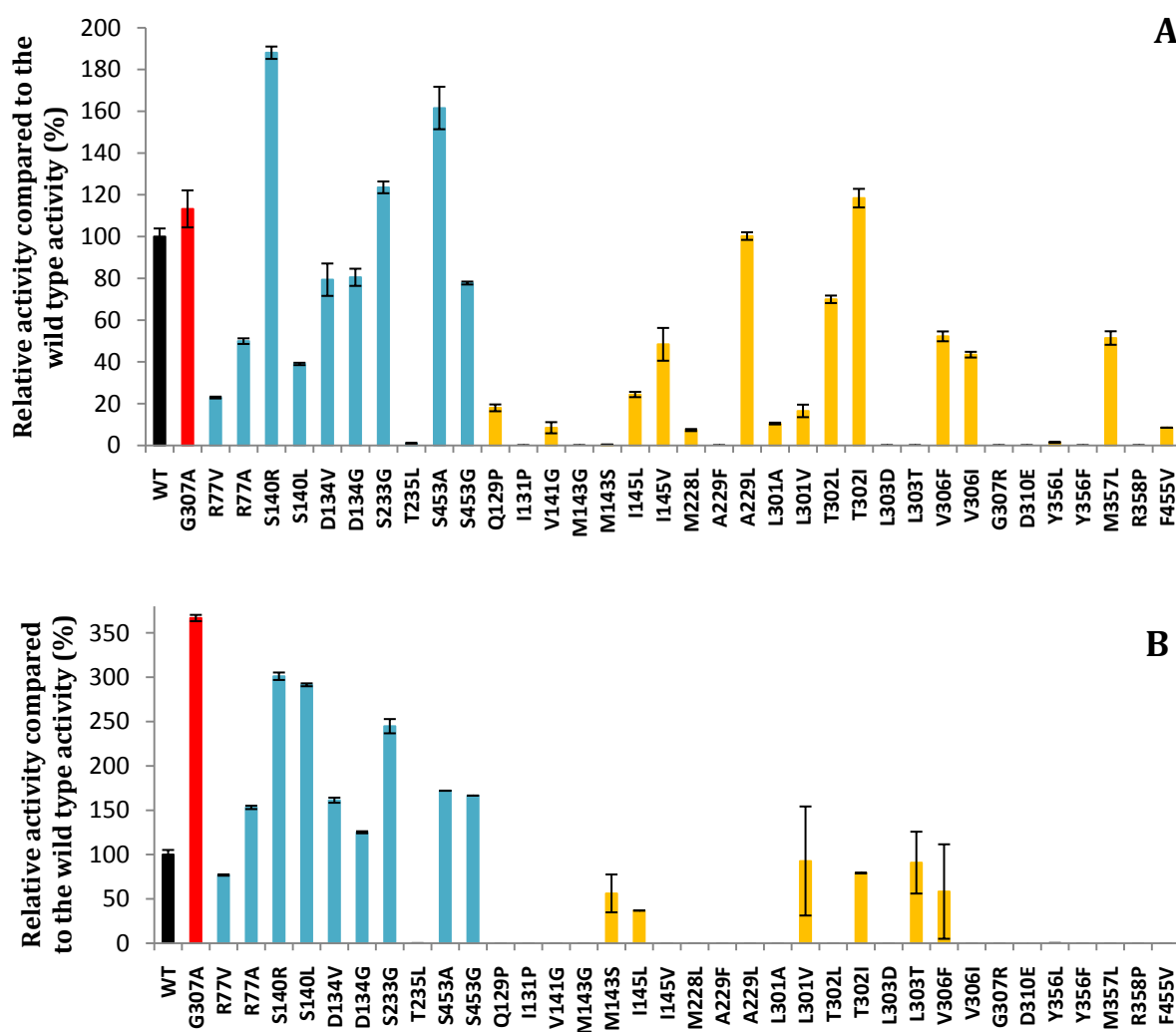
Library position	$\omega$ -OHC <sub>12:0</sub> (mM)
wild type	0.08-0.10
G307A	0.15
R77V	0.13
R77A	0.12
D134V	0.18
D134G	0.16
S140R	0.09
S140L	0.10
S233G	0.17
T235L	0.12
S453A	0.13
S453G	0.12

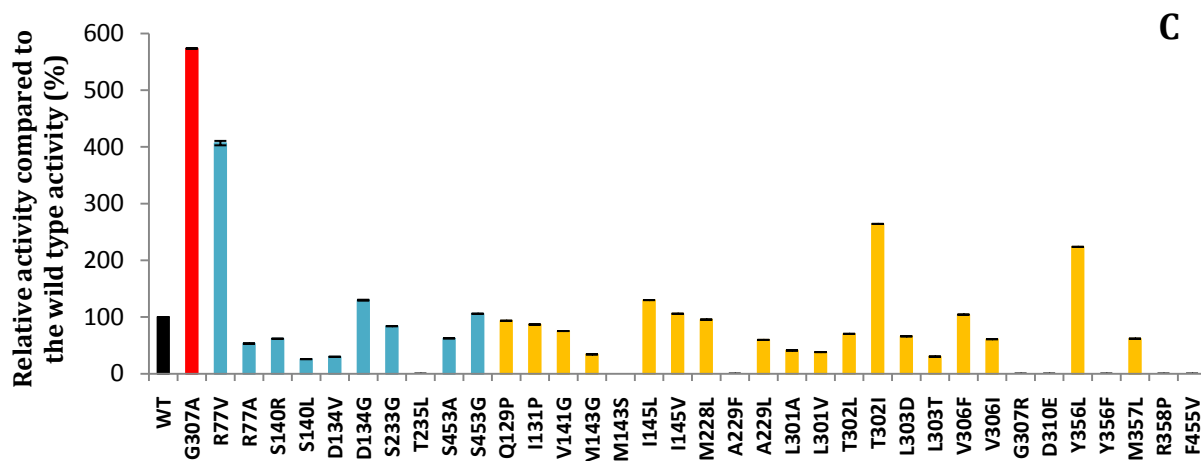
## 2. *In vitro* activity tests

Each new mutant generated from the positions within the active site and all the selected sequenced variants from the substrate entrance tunnel positions were expressed at larger scale (in 2 L flasks), the cells lysed and the protein lysate recovered to perform *in vitro* tests. Each reaction was performed for two hours, at 30 °C, 700 rpm using 0.5  $\mu$ M lysate, 1 mM of substrate, 1 mM NADPH and cofactor regeneration system. Three different substrates were tested, the



model medium chain-length substrate  $C_{12:0}$  (Figure 19A),  $C_{16:0}$  as long chain-length fatty acid (Figure 19B) and as short chain-length fatty acid,  $C_{8:0}$  (Figure 19C). The figures displayed the activity relative for each variant (positions in the substrate entrance in blue/ active site in yellow) compared to the wild type activity (black) set to 100 %. The best mutant described so far, G307A (in red), was also taking into consideration for comparison, also acting as positive control. The variants S140R, S453A, S233G and T302I displayed a higher activity (1.9-, 1.6-, 1.23- and 1.18-fold enhancement, respectively) compared to the wild type activity towards  $C_{12:0}$ . The activities of the mutants S140R, S233G, T302I and S453A were also shown higher than G307A in the case of  $C_{12:0}$ . Most of the mutations generated in the active site indicated a loss in activity to convert  $C_{16:0}$ , G307A remaining the most active variants (3.7-fold increase), followed by S140R, S140L, and S233G, positions located at the substrate entrance tunnel also beneficial to catalyse the conversion of  $C_{12:0}$  (3-, 2.9- and 2.4-fold increase, respectively). Short chain fatty acid  $C_{8:0}$  remains poorly converted by most of the new variants, G307A displaying the greatest activity for such substrate in comparison to the wild type but corresponding to 6 % of substrate converted only.





**Figure 19.** Relative activity of the new generated variants (substrate entrance tunnel (blue) and active site (yellow)) compared to the wild type activity set at 100 % (in black). G307A (red) was tested and used as positive control. *In vitro* tests were performed with 0.5  $\mu$ M lysate towards 1 mM  $C_{12:0}$  (A),  $C_{16:0}$  (B) and  $C_{8:0}$  as substrates (C). Error bars regarding the relative activity towards  $C_{12:0}$  and  $C_{16:0}$  represent standard deviation of triplicates or duplicates.

\* Protein content normalized based on the classic CO measure.<sup>[76]</sup>

### 3. Improved variants

Subsequently to the *in vitro* tests, the single mutants showing improved activity compared to the wild type, towards one or several of the selected substrates, were combined with G307A and further investigations were performed on the new created variants. This led to the generation of five double mutants G307A/S140R, G307A/S140L, G307A/S233G, G307A/T302I, S233G/T302I and one triple mutant G307A/S233G/T302I. Despite a good activity towards  $C_{12:0}$  for the mutant S453A, this variant was not taking into consideration for further analysis due to the difficulty to obtain reproducible expression.

#### 3.1. Characterization *in vitro*

##### 3.1.1. *In vitro* activity comparison

The activity of the five double mutants created and the triple mutant were compared during *in vitro* tests towards 1 mM of  $C_{12:0}$ , 1 mM  $C_{16:0}$  and 1 mM  $C_{8:0}$  as substrates. The percentages of conversion of the different substrates after 2 hours reaction are shown in the Table 11. The substrate  $C_{12:0}$  was entirely depleted after two hours by the double mutant G307A/S233G and the triple mutant G307A/S233G/T302I, forming the corresponding  $\omega$ -OHFA. The same double mutant G307A/S233G also displayed a higher activity than the single mutant G307A in the case of longer and shorter chain-length fatty acids (62 % and 14 % conversion,



respectively, against 45 % and 6 %, respectively for G307A); although the two tested substrates remain less converted by the biocatalyst in comparison to C<sub>12:0</sub>. The beneficial mutation T302I in the case of C<sub>12:0</sub>, led to a loss in activity towards C<sub>16:0</sub> and this can be observed for each combination with such position as with G307A/T302I, S233G/T302I and the triple mutant G307A/S233G/T302I. The single mutants S140R, shown highly active towards C<sub>16:0</sub> and C<sub>12:0</sub> (3- and 1.9-fold increase, respectively) and S140L with improved activity towards C<sub>16:0</sub> compared to the wild type (2.9-fold increase), were combined with G307A. However the creation of double mutants only led to a decreased activity in the case of G307A/S140L for C<sub>12:0</sub> and C<sub>16:0</sub> (28 % and 31 % conversion, respectively for the double mutant in comparison to 55 % and 45 % conversion, respectively for G307A) and a complete loss in activity towards the three substrates in the case of G307A/S140R. The two mutations highly valuable for improved activity towards short to long chain-length fatty acids are: S233G, position from the substrate entrance tunnel in combination with G307A, the position located within the binding pocket.

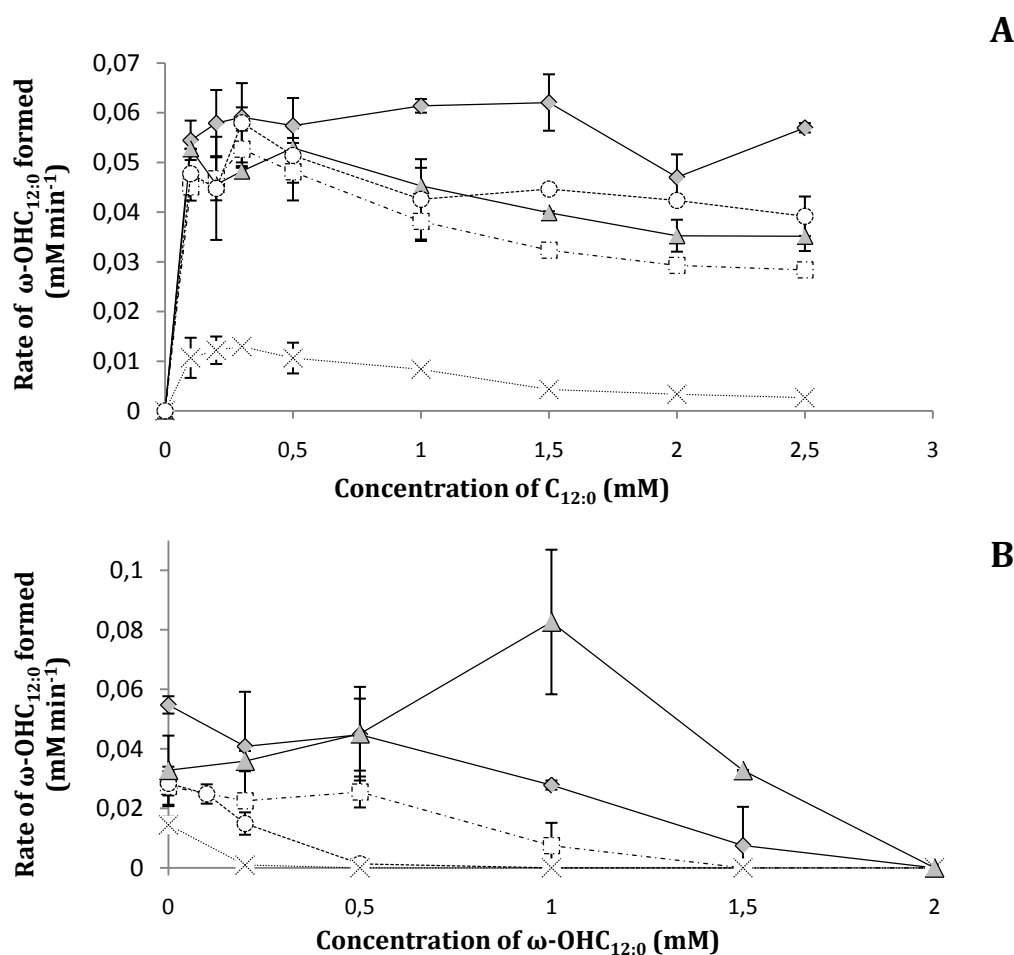
**Table 11.** Comparison of the percentage of conversion of the three substrates C<sub>12:0</sub>, C<sub>16:0</sub> and C<sub>8:0</sub> by the best selected variants and the created double and triple mutants. Conditions: 2 h *in vitro* tests, 0.5 lysate enzyme, 1 mM NADPH cofactor, cofactor regeneration and potassium phosphate buffer pH 7.4. \*Protein content normalized based on the reduction of the whole construct.<sup>[77]</sup>

Variants / Fatty acids	Conversions (%)		
	C <sub>12:0</sub>	C <sub>16:0</sub>	C <sub>8:0</sub>
wild type	50	13	1
G307A	55	45	6
G307A/S140R	0.9	1.2	0.5
G307A/S140L	28	31	0.9
G307A/S233G	99	62	14
G307A/T302I	95	2	7
S233G/T302I	60	0	0.9
G307A/S233G/T302I	100	0.1	5

### 3.1.2. Substrate/product inhibition

Further characterizations were performed with the double mutants G307A/S233G, G307A/T302I and the triple mutant G307A/S233G/T302I in comparison with the best single mutant described G307A to evaluate potential substrate and product inhibition (Figure 20A and B, respectively). Increased concentrations of C<sub>12:0</sub> in the range of [0.1-2.5 mM] were tested for

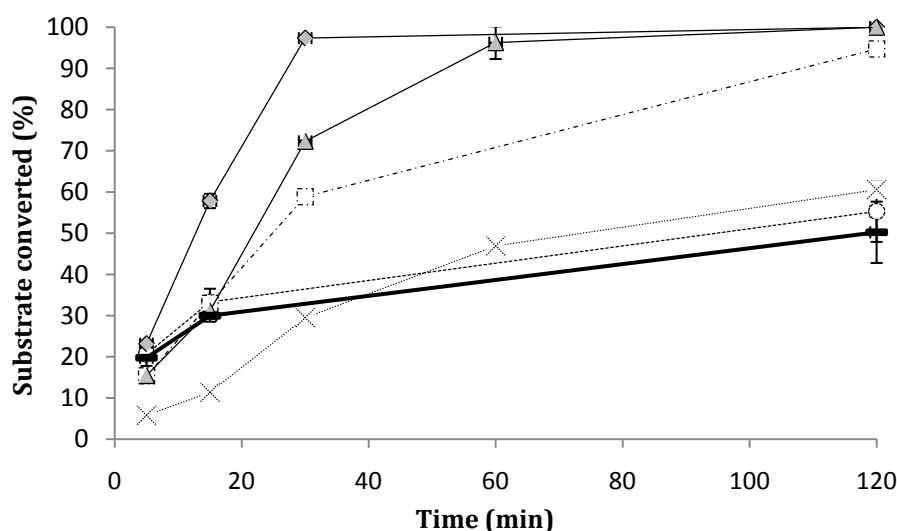
the substrate inhibition reactions and a constant concentration of 1 mM of  $C_{12:0}$  in the presence of increased concentration of  $\omega$ -OHC $_{12:0}$  [0-2 mM] were assayed. The tendency showed a decrease in the activity for each variant in the presence of greater amount of substrate up to 0.3 mM of substrate added for the majority of variants. The double mutant G307A/S233G showed a higher resistance than the other mutants for higher concentrations of substrate, up to 2 mM of  $C_{12:0}$ . The external addition of product had a strong negative impact on the efficiency of biotransformation of the substrate: a decrease in the rate of reaction can already be observed for the smallest quantity of 0.2 mM of  $\omega$ -OHC $_{12:0}$  supplemented. However the triple mutant is the most resistant one for greater quantities of product up to 1.5 mM of  $\omega$ -OHC $_{12:0}$  supplemented. The other variants were more robust than G307A regarding the presence of  $\omega$ -OHC $_{12:0}$ , excluding the double mutant S233G/T302I displaying a lower activity and rapidly susceptible to inhibition by substrate and product.



**Figure 20.** Substrate inhibition investigations (A) and product inhibition evaluation (B) for 2 minutes reaction with the different mutants (0.5  $\mu$ M): single variant G307A (circles), G307A/S233G (grey diamonds), G307A/T302I (boxes), S233G/T302I (cross) and the triple mutant G307A/S233G/T302I (grey triangles). \*Protein content normalized based on the reduction of the whole construct.<sup>[77]</sup>

### 3.1.3. Performance over time

The activities of the new created biocatalysts over time were estimated by performing *in vitro* reactions towards 1 mM of substrate  $C_{12:0}$  for 5 minutes, 15 minutes, 30 minutes, 1 h and 2 h with quantification of product (Figure 21). The resistance described previously for greater amount of substrate and product for the double mutant G307A/S233G, G307A/T302I and the triple mutant G307A/S233G/T302I can explain the rapid consumption of the substrate along the reaction. The fastest rate of conversion was achieved with the double mutant G307A/S233G with entire substrate depletion after 30 minutes time course leading to the formation of the corresponding  $\omega$ -OHC $_{12:0}$ . The triple mutant reached the total consumption of substrate after 1 h of reaction. The wild type, the single mutant G307A and the double mutant S233G/T302I, less resistant to increased concentrations of substrate and product (previously described) and with a lower activity, only reached the conversion of 50-60 % after 2 hours of reaction.

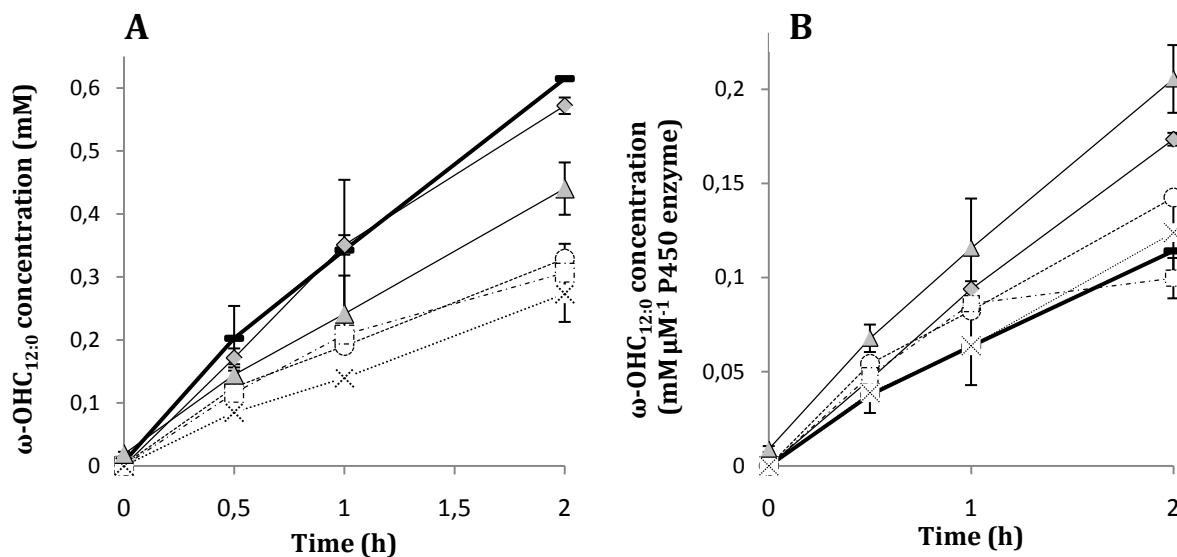


**Figure 21.** Evolution of the substrate conversion over time and comparison of activity among the different variants, using 0.5  $\mu$ M lysate, 1 mM substrate  $C_{12:0}$ . The fastest rate of reaction was obtained with the double mutant G307A/S233G (grey diamonds). Triple mutant G307A/S233G/T302I (grey triangles), G307A/T302I (boxes), S233G/T302I (cross), G307A (circles) and wild type (solid black line). \*Protein content normalized based on the reduction of the whole construct.<sup>[77]</sup>

### 3.2. *In vivo* validation

The improved activity towards  $C_{12:0}$  demonstrated *in vitro* with the generated double mutants and the triple mutant was evaluated during *in vivo* reactions in comparison with the wild type and the single variant G307A (Figure 22). The initial rates during a whole resting cell

biotransformation ( $50 \text{ g}_{\text{cdw}} \text{ L}^{-1}$ ) and the product concentration of each variant were evaluated, starting with  $2 \text{ mM}$  of  $\text{C}_{12:0}$  substrate. Prior to the estimation of protein expressed (by CO measurement), the wild type showed the greatest concentration of product ( $0.61 \text{ mM}$  of  $\omega\text{-OHC}_{12:0}$ ) followed by the double mutant G307A/S233G and the triple mutant G307A/S233G/T302I ( $0.57 \text{ mM}$  and  $0.44 \text{ mM}$  of  $\omega\text{-OHC}_{12:0}$ , respectively) after 2 hours reaction (Figure 22A). This suggests an efficient protein expression leading to a greater biotransformation efficacy. The single variant G307A and the double mutant G307A/T302I had similar product concentrations ( $0.33 \text{ mM}$  and  $0.31 \text{ mM}$  of  $\omega\text{-OHC}_{12:0}$  formed after 2 hours, respectively), whereas the lowest conversion of substrate was observed for the mutant S233G/T302I ( $0.27 \text{ mM}$  of  $\omega\text{-OHC}_{12:0}$ ). For an accurate comparison, the protein content was estimated for each variant at the beginning of the reaction in the cell lysate. The assessed product formation was no longer based on the cell density but based on the amount of P450 produced (Figure 22B). The tendency after CO measurement showed that the triple mutant, the double mutant G307A/S233G and G307A reached the highest concentration of product per micromolar of enzyme ( $0.21 \text{ mM}$ ,  $0.17 \text{ mM}$  and  $0.14 \text{ mM}$  of product per  $\mu\text{M}$  of P450, respectively, after 2 hours reaction). Lower product concentrations can be observed for the double variants S233G/T302I, G307A/T302I, similar to the wild type activity ( $0.12$ ,  $0.10$  and  $0.11 \text{ mM}$  of  $\omega\text{-OHC}_{12:0}$  formed, respectively).



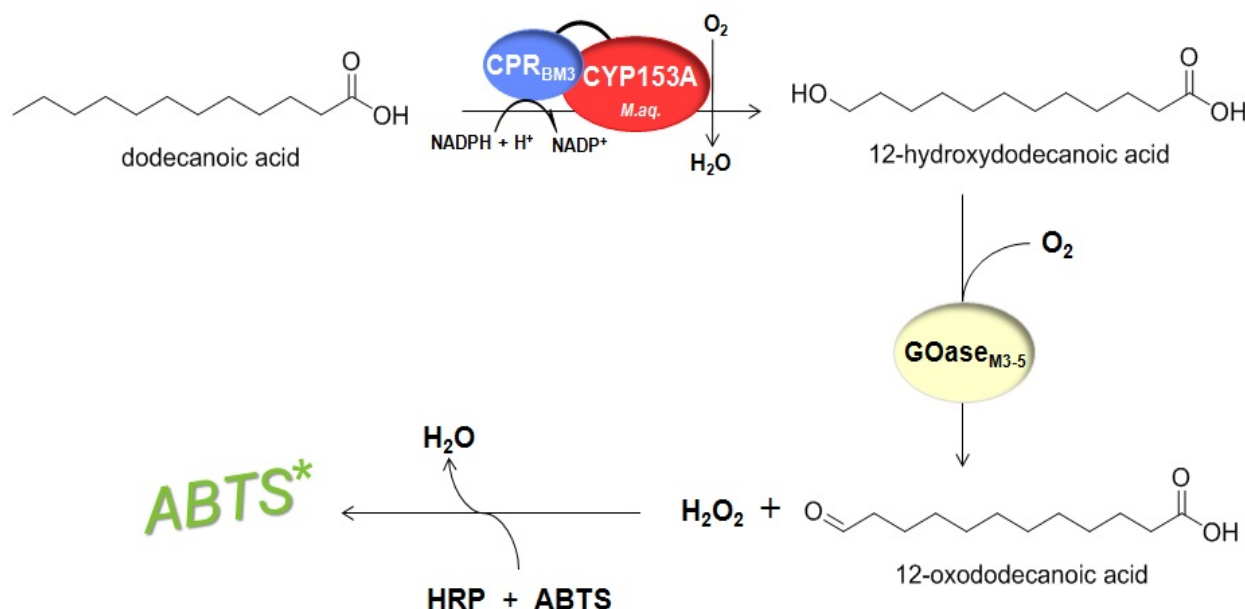
**Figure 22.** Product concentration along two hours biotransformation of  $2 \text{ mM}$   $\text{C}_{12:0}$  provided ( $5\% \text{ DMSO (v/v)}$ ) with whole resting cells ( $9 \text{ g}_{\text{cdw}} \text{ L}^{-1}$ ), prior to CO evaluation (A) and correlated to the protein content (B). Triple mutant G307A/S233G/T302I (grey triangles), double mutant G307A/S233G (grey diamonds), G307A/T302I (boxes), S233G/T302I (cross), G307A (circles) and wild type (solid black line). \* Protein content based on the classic CO measure.<sup>[76]</sup>

### Chapter 3: Establishment of a whole-cell screening assay for identification of improved P450-activity

The development of a screening assay is of high interest in this study to identify improved variant activity for the conversion of C<sub>12:0</sub> into ω-OHC<sub>12:0</sub>. The screening of large size mutant libraries usually relies on an efficient screening technique or selection method. The establishment of such assay was achieved in parallel with the generation of the mutant libraries described in *Chapter 2* (publication II) and several selected variants were tested to demonstrate the successful concept. Part of this work was included in a communication (publication III).

#### 1. Galactose oxidase assay

Based on previous work, a mutant galactose oxidase enzyme M3-5 (GOase<sub>M3-5</sub>) was shown to oxidize a panel of alcohols and terminally hydroxylated fatty acids.<sup>[147]</sup> (Bachelor Thesis Sarah-Luise Lang, ITB, University of Stuttgart). This reaction is associated with the production of hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) and by coupling a horseradish peroxidase (HRP) and the ABTS, the activity of the enzyme can be monitored at OD<sub>420</sub>, maximum of absorption of the oxidized dye. The enzyme GOase<sub>M3-5</sub> mutant was shown to oxidize the product formed by CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> and the activity of the enzyme and consequently of active P450 variants towards fatty acids can be evaluated through such assay (Figure 23).



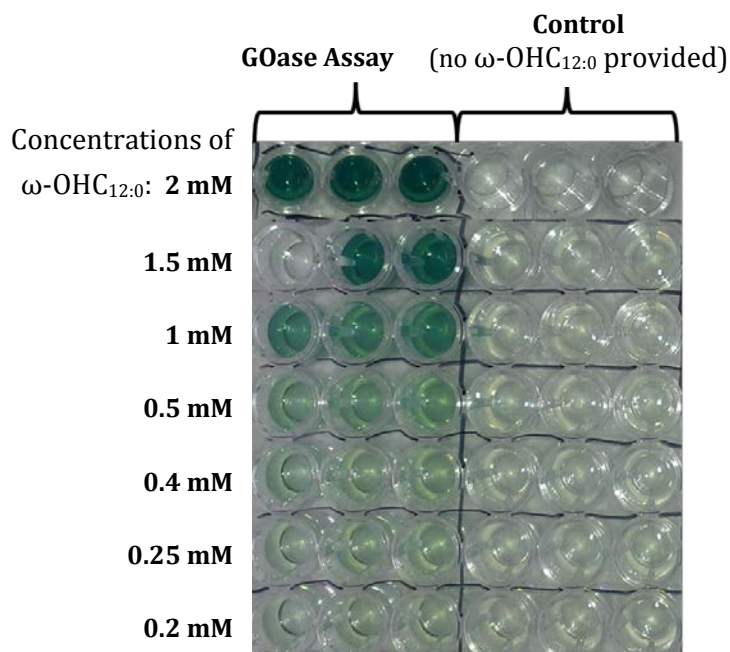
**Figure 23.** Subsequently to the P450-based reaction, the product ω-OHC<sub>12:0</sub> can be oxidized by GOase<sub>M3-5</sub>, generating H<sub>2</sub>O<sub>2</sub> and the activity of the enzyme and therefore of the P450 can be monitored at OD<sub>420</sub> after addition of the HRP and ABTS.

## 2. Establishment of the assay

The different steps required for establishing a suitable protocol were achieved with the use of CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> wild type as model for the P450-based reaction. C<sub>12:0</sub> and  $\omega$ -OHC<sub>12:0</sub> were used as model substrate and product, respectively and in the presence of the chosen catalyst for the assay, GOase<sub>M3-5</sub>.

### 2.1. Demonstration in a buffer system

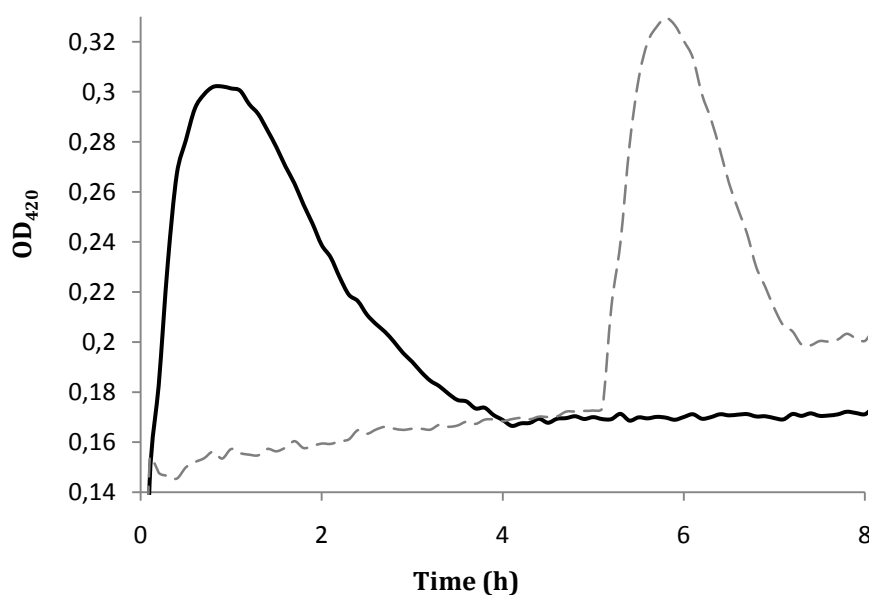
Preliminary experiments were performed in a buffer system with purified GOase<sub>M3-5</sub>, HRP, ABTS and directly addition of increased concentrations of  $\omega$ -OHC<sub>12:0</sub> in the range of [0.2-2 mM] to demonstrate the concept of reaction, feasibility of catalysis with an increase of absorption occurring and visible to the naked eye (oxidized ABTS turns into a green colour, Figure 24). The most suitable concentrations of enzyme, volume of buffer, pH, molarity, choice of dye or incubation time were established as well using the buffer system (Sarah-Luise Lang Bachelor thesis, ITB, University of Stuttgart). The appropriate conditions required the use of 1  $\mu$ L of HRP (900 units  $\mu$ L<sup>-1</sup> stock), 8  $\mu$ L of ABTS (10 mM stock), 10  $\mu$ L of GOase (3.5-8  $\mu$ M mL<sup>-1</sup> stock), 10  $\mu$ L of  $\omega$ -OHFA (5 % DMSO, end concentration (v/v)) and 100 mM potassium phosphate buffer pH 7.4 up to 200  $\mu$ L, final volume.



**Figure 24.** Purified GOase<sub>M3-5</sub>, HRP and ABTS in the presence of increased concentrations of  $\omega$ -OHC<sub>12:0</sub> in 100 mM potassium phosphate buffer pH 7.4. The colour of the solution turns into green due to the oxidation of ABTS dye upon GOase activity. The darkness of the green colour is correlated to higher amount of fatty acids present at the start of the assay.

## 2.2. Deactivation of the supernatant

The assay, successfully established in a buffer system with purified enzymes, was developed in a high-throughput screening manner. The enzymes GOase, HRP and ABTS were directly added to the supernatant issued from the whole-cell P450-based reaction. To ensure reproducible signals in terms of intensity and time of appearance, the deactivation at 90 °C for 30 minutes of the supernatant prior to the assay is required (Figure 25). The heating step must allow the suppression of external metabolic activity leading to a lag phase detected between 4 and 10 hours of incubation. The grey dashed curve corresponded to the same supernatant tested as the solid black line in which the deactivation has been applied for 30 minutes after the biotransformation of 2 mM of  $C_{12:0}$  by the CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> wild type. Subsequently to the heating step, the maximum of absorption was reached after 1 h incubation with ABTS, HRP and GOase<sub>M3-5</sub>.

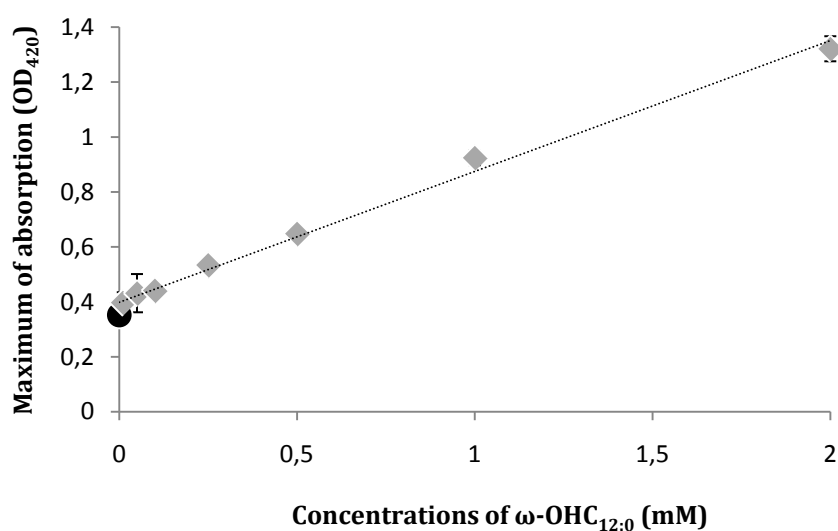


**Figure 25.** Signal of increase of absorption at OD<sub>420</sub> after P450-based reaction. The signal appeared after 1 h incubation in the case of a heating step at 90 °C for 30 min prior to the application of the assay (solid black line). In the absence of the deactivation procedure, the signal is not reproducible in terms of intensity and of time of appearance with a lag phase (grey dashed line).

### 2.3. Detection limit

#### 2.3.1. Increased concentrations of $\omega$ -OHC<sub>12:0</sub>

In order to evaluate the efficacy of the assay in the detection of small quantities of compounds of interest, increased concentrations of  $\omega$ -OHC<sub>12:0</sub> were supplemented to whole resting cells for two hours and the supernatant was analysed in the presence of HRP, ABTS and GOase<sub>M3-5</sub>. The negative control consisted in replacing the  $\omega$ -OHC<sub>12:0</sub> by the same amount of DMSO (black data). In the presence of the P450 product, the signal displayed a characteristic peak of absorption after 1 hour incubation and the negative control remained of lower intensity (Figure 26). The intensity of the signal peak depends on the product concentration: for a larger intensity of signal, higher amounts of fatty acid are present. The values below the negative control (black data), will correspond to background signal, therefore the limit of detection can be estimated between 0.01 mM and 0.02 mM of  $\omega$ -OHC<sub>12:0</sub>.



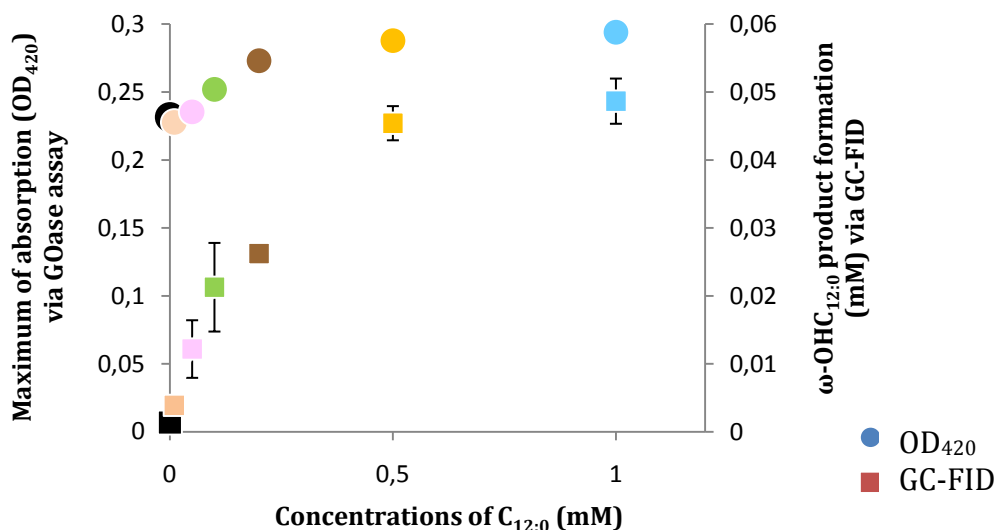
**Figure 26.** Maximum of absorption (OD<sub>420</sub>) displayed depending on increased concentrations of  $\omega$ -OHC<sub>12:0</sub>. Correlation observed between the amounts of  $\omega$ -OHC<sub>12:0</sub> present and the intensity of signal recorded (higher intensity of signal for greater quantities of fatty acids provided).

#### 2.3.2. Addition of increased concentrations of C<sub>12:0</sub>

A similar experiment was performed using increased amounts of C<sub>12:0</sub> substrate provided at the beginning of the P450-catalysed reaction in whole resting cells for 2 h (and a negative control consisting in the same amount of DMSO, black data points). The deactivated supernatants were afterwards analysed using the GOase assay (coloured circles, left axis Figure 27). In order to validate the assay and to determine the minimum of product detectable, the



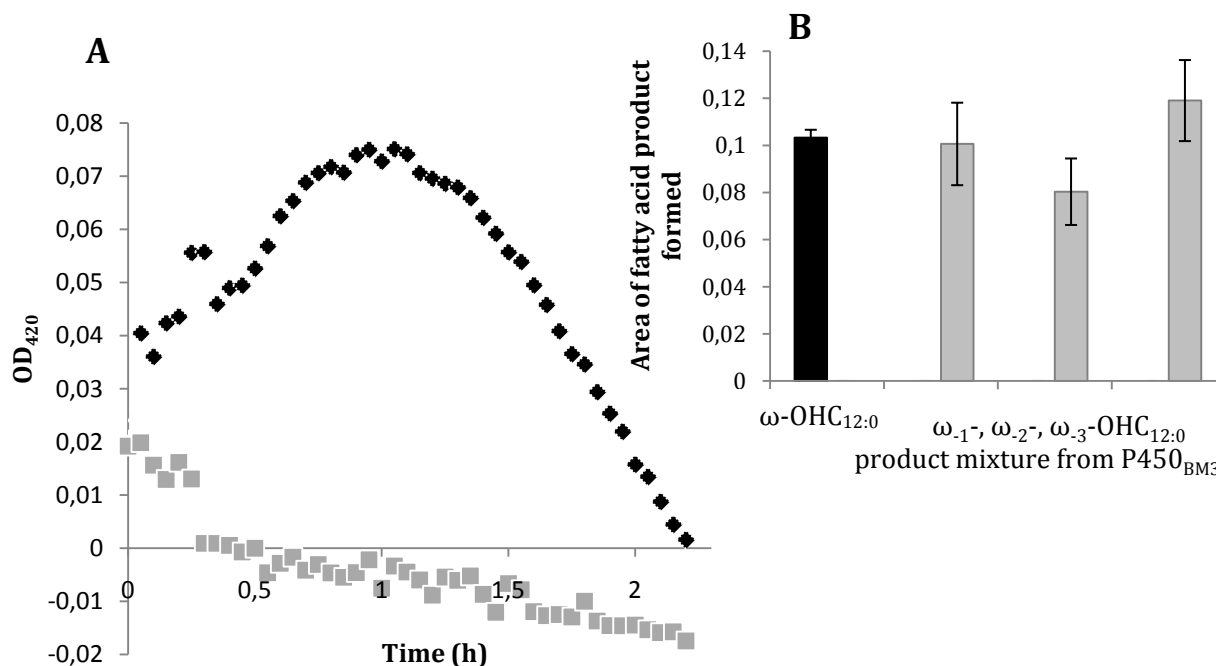
same samples were extracted and the product formed quantified after GC-FID measurement (coloured boxes, right axis Figure 27). The quantification of product by GC-FID confirmed a correlation between the intensity of signal and the amount of product formed after the P450-biocatalysis. The detection limit of the  $\omega$ -OHC<sub>12:0</sub> characterized via the GOase assay corresponds to > 0.012 mM of hydroxylated product.



**Figure 27.** Maximum of absorption (OD<sub>420</sub>), depending on increased concentrations of substrate C<sub>12:0</sub> provided for whole-cell biotransformation, reflecting the amount of product formed (coloured circles, left axis). The increase of absorption associated to the concentration of  $\omega$ -OHC<sub>12:0</sub> formed after P450-based reactions, can also be correlated to the GC-FID analysis, after product extraction (coloured boxes, right axis).

### 3. Selectivity of the galactose oxidase

The GOase assay was tested with the supernatant obtained after a biotransformation of 2 mM of C<sub>12:0</sub> with the P450<sub>BM3</sub> compared to CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> wild type. A peak of absorption can be observed in the case CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> wild type only (black curve, Figure 28A). P450<sub>BM3</sub> is well-known as a sub-terminal P450 hydroxylase and the GC results confirmed an activity of the enzyme with the production by P450<sub>BM3</sub> of  $\omega$ -1-,  $\omega$ -2- and  $\omega$ -3-OHC<sub>12</sub> product mixture, whereas CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> wild type produced the terminal  $\omega$ -OHC<sub>12:0</sub> (black and grey data, respectively Figure 28B).<sup>[101]</sup> The absence of the characteristic absorption peak in the case of P450<sub>BM3</sub> suggests a high selectivity of the GOase<sub>M3-5</sub> towards  $\omega$ -OHFA product only. The product mixture  $\omega$ -1-,  $\omega$ -2- and  $\omega$ -3-OHC<sub>12</sub> (resuspended in DMSO) were directly assayed in a buffer system with GOase, HRP and ABTS and confirmed the previous results (data not shown). This represents a major advantage to screen enzymes with high regioselectivity for the terminal hydroxylation of molecules.

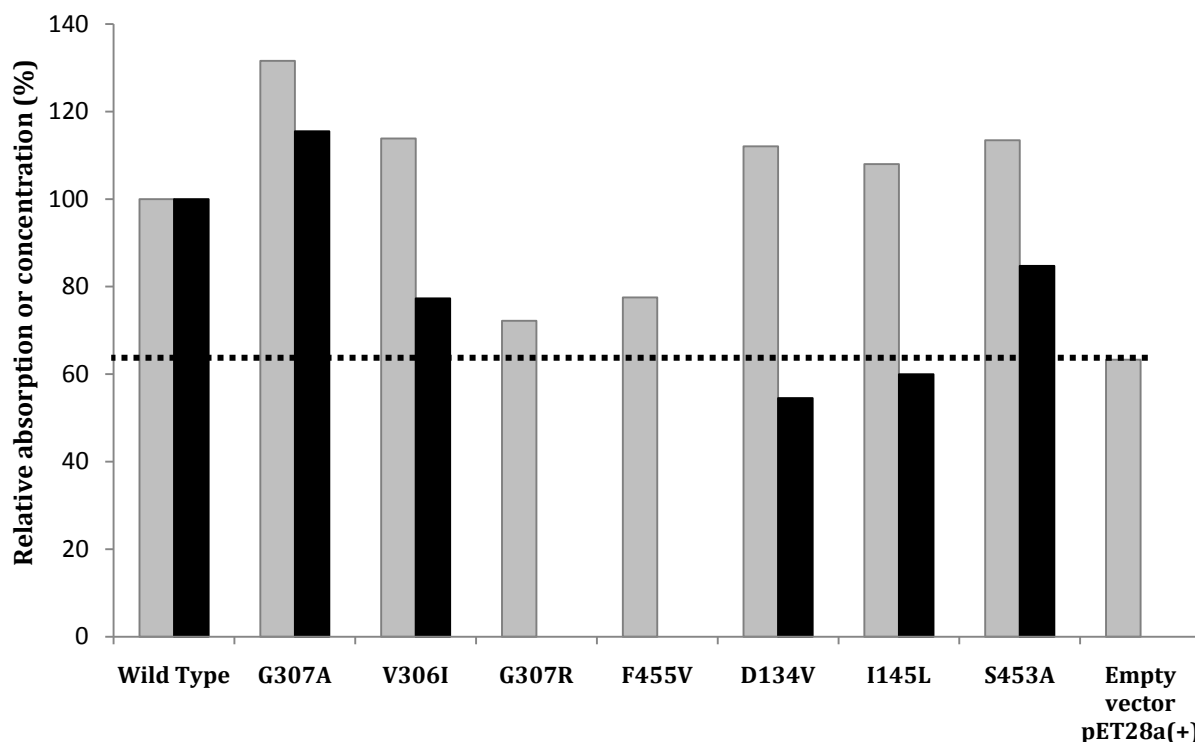


**Figure 28.** (A) Comparison of activity between CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> wild type and P450<sub>BM3</sub> (black diamonds and grey boxes respectively). The signals of absorption were subtracted to the negative control signal (no substrate provided). Signal of P450<sub>BM3</sub> was equivalent to the negative control reaction, with no peak of absorption, potentially due to the production of a mixture of fatty acids confirmed after extraction and GC-FID analysis (grey data) (B). Therefore GOase<sub>M3-5</sub> possesses a great selectivity for the terminal hydroxylated products and enables the identification of highly selective P450 variants for the terminal hydroxylation of fatty acids.

## 4. Screening of mutant libraries

### 4.1. GOase vs. GC-FID

To test the robustness of the assay and to validate its application for the screening of variants with improved activity, several mutants generated and characterized in *Chapter 2*, were tested with the GOase assay. Four mutations located in the active site and three mutations from the substrate entrance tunnel were selected. Subsequently to the biotransformation of 2 mM of C<sub>12:0</sub> for 2 h using 50 g<sub>cww</sub> L<sup>-1</sup> resting cells, the supernatants deactivated were tested via GOase assay (grey bars, Figure 29) and via GC-FID (black bars, Figure 29). The best mutant described so far (G307A) was tested as positive control and a negative control consisted in the empty vector pET28a(+). Each variant activity was compared to the activity of the wild type, set at 100 % (termed relative activity). The empty vector pET28a(+) enables the background limit detection (dashed line), therefore G307R and F455E showing negligible increased of absorption were confirmed inactive due to the absence of product detected after GC-FID analysis. According to the GOase assay and to the GC-FID quantification the most active mutant remaining is G307A.



**Figure 29.** Relative absorption compared to the wild type (set at 100 %) after the GOase assay (grey bars) and relative activity compared to the wild type after product quantification via GC-FID (black bars) for several selected mutants. G307A remains the most active variant after whole resting cell biotransformation of 2 mM  $C_{12:0}$ .

#### 4.2. Protein content correlation

The previous results obtained from the GOase assay or from the quantification by GC-FID can be biased by variable protein expression rates leading to variation in the biotransformation efficiency (absolute concentrations and absorptions, Table 12, p81). Therefore, the absorption (via GOase assay) and activity (via GC-FID) of each variant was compared to the activity of the wild type set at 100 % (relative conversion), prior to the correlation of protein content determined by CO spectra (relative specific conversion). By considering the absolute concentrations, G307A mutant appears as the most active variant, followed by the wild type (Table 12, p81 and Figure 29) but the protein expression rate was not taking into account. Based on the P450 content, the most active variant identified was S453A, followed by G307A, positions located at the substrate entrance tunnel and in the active site, respectively. The assay also enabled to characterize non-active variant as G307R and F455V with an intensity of signal similar to the background signal (empty vector pET28a(+)) and lacking the characteristic peak at 450 nm after CO spectra measurement. The results from whole-cell biotransformation

correlated to the protein content obtained via GOase assay and GC-FID were confirmed by the lysate activity (Table 12, p81). *In vitro* tests were performed with the normalized concentration of 0.5  $\mu\text{M}$  of P450 lysate towards 1 mM of  $\text{C}_{12:0}$  for 2 hours reaction and  $\omega\text{-OHC}_{12:0}$  were extracted and analysed by GC-FID. The mutant S453A was shown 1.6-fold more active than the wild type, G307A was found 1.1-fold more active than the wild type and the variants G307R and F455V were confirmed inactive. Combining the GOase assay with the determination of the P450 concentration enabled the accurate identification of active variants based on the protein content and not only based on the efficient protein expression rate.

**Table 12.** Screening for improved variant activity in comparison to the wild type activity set at 100 % (relative conversions). The comparison between the results obtained from the GOase assay, GC-FID analysis and lysate activity tests were determined prior to and after CO correlation (absolute and relative specific concentrations values, respectively). The CO correlation based on the protein content enabled the identification of the enhanced mutant S453A. (AS\*: active site; SE\*: substrate entrance; *Protein content based on the classic CO measure*<sup>[76]</sup>)

Mutants	Mutation location	Absolute product concentration GC-FID ( $\mu\text{M}$ )	Relative conversion GC-FID (%)	Relative Specific conversion GC-FID (%)	Absolute absorption GOase-Assay	Relative conversion GOase-Assay (%)	Relative Specific conversion GOase-Assay (%)	Relative specific conversion Lysate (%)
Wild Type	-	360 $\pm$ 0.04	100	100	0.181	100	100	100
G307A	AS*	416 $\pm$ 0.08	116	105	0.238	131	119	113
V306I	AS	279 $\pm$ 0.16	78	66	0.206	113	96	43
G307R	AS	0	0	0	0.131	72	0	0
F455V	AS	0	0	0	0.140	77	0	8
D134V	SE*	197 $\pm$ 0.05	55	47	0.203	112	96	79
I145L	SE	216 $\pm$ 0.12	60	54	0.196	107	96	48
S453A	SE	305 $\pm$ 0.02	85	<b>123</b>	0.206	113	<b>164</b>	<b>161</b>
Empty pET28a(+)	-	0	0	0	0.115	63	0	0



## Discussion

### *Chapter 1: Identification of bottlenecks in CPR2<sub>mut</sub>-catalysed reactions: how to address the limitations and improve whole-cell biocatalysis*

The first step towards P450-based process optimisation involves the identification of potential challenges preventing from industrial implementation. Despite the large number of reactions catalysed by these versatile enzymes, their potential for industrial application still suffer from reported limitations.<sup>[116]</sup> Numbers of bottlenecks associated to CYPs reactions and whole-cell systems have already been described, including substrate/product inhibition on the biocatalyst and/or the host, insufficient NADPH availability, catalyst instability or substrate/product transport limited.<sup>[107,116,146]</sup> Based on the previous literature, investigations have been executed using the model system fusion protein CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> (G307A, 3xGGS, termed CPR2<sub>mut</sub>) with identification of bottlenecks and alternatives to tackle the potential challenges.

In a whole-cell based reaction, depending on the substrate and/or product size, polarity or solubility, a limited transport across the cell envelope may occur leading to insufficient rates and unsatisfactory product concentrations.<sup>[146]</sup> To improve transport across the cell wall of hydrophobic substrates, as steroids, cyclodextrines and derivatized forms have demonstrated great permeabilization and solubilisation effect.<sup>[161]</sup> A synthetic testosterone derivative was successfully converted at 98 % in *Bacillus megaterium* after enhancement of the process with the substrate solubilised with 2-hydroxypropyl- $\beta$ -cyclodextrin.<sup>[162]</sup> Examples from the literature report the co-expression of the *alkL* gene from *Pseudomonas putida* GPo1 coding for a transporter and CPR2<sub>mut</sub> or CYP153A6, heterologous expressed in *E. coli* leading to an enhancement in the product yield of terminal hydroxylated fatty acid methyl esters and (*S*)-perillyl alcohol respectively, both in a biphasic system.<sup>[73,129]</sup> Furthermore, a 5.5-fold increase in the conversion of the long-chain hexadecanoic acid (C<sub>16:0</sub>) into terminally hydroxylated hexadecanoic acid ( $\omega$ -OHC<sub>16:0</sub>) was achieved through the deletion of the *fadD* gene, coding for FadD enzyme involved in the degradation of fatty acids and the overexpression of the *E. coli* FadL transporter.<sup>[137]</sup> In this study, in the case of dodecanoic acid (C<sub>12:0</sub>), permeabilization methods applied to enhance the substrate uptake had no beneficial consequences as the reaction rate was not shown limited by the substrate transport. This comes to the fact that the activity of FadL is not required for short and medium chain-length of fatty acids entering the cells via simple diffusion.<sup>[163]</sup>

Cofactor availability was revealed as a limiting parameter in a whole-cell biotransformations and the biocatalysis was improved by the addition of supplemental NADPH. The addition at the start of the reaction of twice the amount of NADPH compared to the substrate quantities (3 mM and 1.5 mM, respectively) led to the most efficient biotransformation using whole resting cells with 1.3-fold increase compared to the control reaction, lacking external addition of NADPH. However the addition of NAD(P)H for large-scale applications is hindered by the expense of such cofactor.<sup>[164]</sup> Evolution methods, chemical and electrochemical techniques or light-driven catalysis are potential alternatives to overcome the limited availability of NADPH. Enzyme engineering is one possible approach: P450<sub>cam</sub> mutant catalysing the hydroxylation of naphthalene through the peroxide shunt pathway using H<sub>2</sub>O<sub>2</sub> instead of NADPH cofactor was created.<sup>[165]</sup> As another option, the availability of cofactor can be improved by the enhancement of cofactor regeneration. A 22-fold increase in the epoxidation of 1-hexene by P450<sub>BM3</sub> (F87V) was demonstrated since the glucose dehydrogenase of *Bacillus amyloliquefaciens* was co-expressed in the *Bacillus subtilis* strain 3C5N.<sup>[156]</sup> Improvement in the production of  $\omega$ -OHC<sub>16:0</sub> was achieved by *E. coli* strain engineering, expressing CYP153A from *Marinobacter aquaeolei* (CYP153A<sub>M.aq.</sub>) for the catalysis and overproducing the essential NADH cofactor.<sup>[166]</sup> A switch in the cofactor specificity required for the enzyme from NADPH to NADH would reduce the cost of the process. NADH is a more stable compound and can be generated in larger amount in *E. coli*. This would also enable a greater choice of dehydrogenases providing the cofactor regeneration to co-express with P450s, as NAD<sup>+</sup>-dependent dehydrogenases are more abundant than NADP<sup>+</sup>-dependent ones.<sup>[167]</sup> P450<sub>BM3</sub> in combination with NADP<sup>+</sup> dehydrogenase have been successfully engineered to accept NAD<sup>+</sup> for the hydroxylation of cyclohexane, octane and C<sub>14:0</sub>, performed in a biphasic system.<sup>[168]</sup> Successful examples relate chemical approaches consisting in the replacement of NAD(P)H cofactor by artificial nicotinamide molecules. The combination of engineered P450<sub>BM3</sub> and utilization of cofactor-regeneration system showed catalysis through NADH biomimics.<sup>[158]</sup> More recently the synthetic cofactor analogue 1-benzyl-1,4-dihydronicotinamide was assayed to promote peroxygenases reactions by testing P450 peroxygenases CYP152A1, from *Bacillus subtilis*, and CYP152A2, from *Clostridium acetobutylicum*.<sup>[157]</sup> The use of the synthetic cofactor nicotinamide-3,4-chloromethoxy-phenyl-hydride (termed NACMH, Appendix A3) was shown in this study beneficial for the biotransformation of 1.5 mM of C<sub>12:0</sub> by the fusion construct CPR2<sub>mut</sub> expressed in *E. coli* cells. Indeed, 1.5-fold increase was observed in the case of addition of twice the stoichiometric amount of the NADPH analogue. This demonstrates the feasibility of external addition of synthetic cofactor to overcome the limited availability of natural NAD(P)H. Electrochemical methods are also developed to circumvent the need of expensive cofactor and to improve the catalytic response on electrodes; this implies engineering and immobilization of the P450 on



suitable electrode surfaces transferring the electrons.<sup>[169]</sup> The immobilization on glassy carbon electrodes of the heme domain of P450<sub>BM3</sub> and the human CYP3A4 fused to the flavodoxin protein from *Desulfovibrio vulgaris* for the electron transfer allowed the enhancement of the coupling efficiency and the catalytic activity towards *p*-nitrophenol and erythromycin, respectively.<sup>[169,170]</sup> The light-driven catalysis is another possibility, and the creation of hybrid P450<sub>BM3</sub> heme domain containing the photosensitizer Ru(II) bound was lately reported active for the hydroxylation of C<sub>12:0</sub> upon light and in the absence of cofactor.<sup>[171]</sup>

Substrate C<sub>12:0</sub> has been shown to have inhibitory effects at the enzyme level and during whole-cell reactions with the fusion construct CPR2<sub>mut</sub>. Substrate inhibition has already been reported at the enzyme level for concentration above 0.2 mM of saturated fatty acids, longer than 13 carbons in the case of CYP505 from *Fusarium oxysporum* during *in vitro* reactions.<sup>[172]</sup> In our fusion construct, the inhibitory concentration of substrate C<sub>12:0</sub> was estimated above 1 mM *in vitro* and 2 mM for whole-cell biotransformations. At larger scale, in 1 L fermenter, 15 mM of C<sub>12:0</sub> corresponded to the inhibitory concentration of substrate tested. It was also reported that monounsaturated fatty acids, such as oleic acid and polyunsaturated fatty acids like linoleic acid have an inhibitory impact on the eukaryote P450 CYP3A.<sup>[173]</sup> Regarding the influence on the host, higher concentrations of shorter chain-length fatty acids (20 mM for decanoic acid) have been described harmful for the cell membrane of *E. coli*, leading to an inhibition of the growth.<sup>[174]</sup> Inhibitory properties are mainly observed for short and medium chain-length of saturated fatty acids, explained by the fact that they cannot be metabolized by *E. coli* compared to longer chain-length fatty acids (> C12 carbons).<sup>[175]</sup> Furthermore unsaturated fatty acid as linolenic acid, palmitoleic acid and arachidonic acid display antibacterial activity on pathogenic bacteria as they inhibit the enoyl-acyl carrier protein reductase FabI, involved in the fatty acid synthesis.<sup>[176]</sup> Substrate inhibition of fatty acids observed at the biocatalyst level and related to hosts can be associated to the concentration, the chain-length and the number of double bonds of fatty acids.

Besides substrate inhibition, product inhibition from  $\omega$ -OHC<sub>12:0</sub> was also observed during CPR2<sub>mut</sub> reactions *in vivo* and *in vitro*. The rate of reaction was shown dramatically decreased according to increased concentrations of fatty acid product during whole-cell biotransformation. Small quantities (above 0.1 mM and 0.05 mM substrate for 1  $\mu$ M and 0.5  $\mu$ M of catalyst, respectively) were also shown as inhibitory at the CYP level. An alternative to address such limitation during whole cell biotransformation is the use of a two-phase system consisting in an aqueous phase and an organic phase. This has been demonstrated during whole-cell biotransformation expressing CPR2<sub>mut</sub> using fatty acid methyl ester as substrate, also providing the second liquid phase.<sup>[73]</sup> An aqueous/organic phase system with the carrier solvent bis(2-ethylhexyl) phthalate (BEHP) was successfully applied to extend the time of reaction from 48 h

to 120 h and enabled a longer production of the toxic compound octanol using resting *E. coli* cells harbouring CYP153A6.<sup>[159]</sup> The same biphasic system with BEHP solvent, employing CYP153A6 biocatalyst was used to enhance the hydroxylation of limonene.<sup>[129]</sup>

To address substrate inhibition observed for concentrations above 10 mM of C<sub>12:0</sub> in fermenter, as well as product inhibition, investigations in 1 L fermenters have demonstrated the importance of the feeding strategy in the biotransformation efficacy process. To enhance the efficiency of biocatalysis and to increase the rate of reaction, the use of a fed-batch strategy consisting of the continuous feeding of substrate along the reaction course via a pump system was shown beneficial (2-fold increase compared to the classic batch). Another successful approach involved the application of a solid form of the substrate provided (2.8-fold increase compared to the classic batch). The solid feeding alternative enabled a slow dissolving of the substrate along the biotransformation course. The rate of reaction was no longer limited by a large concentration of substrate and product in the aqueous phase when the co-solvent DMSO was used, which had led to inhibitions. The water solubility at 25 °C of C<sub>12:0</sub> is predicted to be 12.76 mg L<sup>-1</sup> and 278 mg L<sup>-1</sup> for  $\omega$ -OHC<sub>12:0</sub> corresponding to 64  $\mu$ M and 1.3 mM, respectively.<sup>[177,178]</sup> In the case of the solid substrate application, a higher solubility of the substrate, a higher rate of reaction and a higher product concentrations were achieved along the reaction course as demonstrated in a buffer system as well (Appendix A4.). The use of DMSO to dissolve the fatty acids was also avoided with the solid feeding approach, which is a major advantage to circumvent P450 instability caused by co-solvent.<sup>[107,160]</sup>

CPR2<sub>mut</sub> instability has been observed along a biotransformation course with the loss of 50 % of the biocatalyst after 2 h of reaction with whole-cells in fermenter. Such loss of catalyst observed at small and large scale, has naturally a negative impact on the biotransformation efficiency and this can explain that the reaction will stop between 4 and 8 hours and the product concentration remaining identical after 24 hours of reaction. Temperature, pressure, ionic strength, pH, addition of co-solvent, protein engineering and/or formation of reactive species are affecting the stability of secondary and tertiary structures of proteins.<sup>[170,179,180]</sup> The stability of the enzyme is a crucial parameter and protein engineering has been reported to be more effective through the direct selection of stable protein.<sup>[181]</sup> Directed evolution has been applied to P450<sub>BM3</sub> mutant possessing peroxygenase activity, leading to an improved thermostability and this technique has also provided P450<sub>BM3</sub> mutants with improved resistance to organic co-solvents such as acetone, acetonitrile, DMF, DMSO and ethanol.<sup>[182,183]</sup> Furthermore the linker region between the heme domain and the reductase domain of this specific P450 was described as a crucial element to increase such resistance. The linker region can increase the stability also because of a more efficient electron transfer and this prevents from the formation of reactive

oxygen species (ROs).<sup>[116]</sup> ROs can cause instability by accelerating the catalyst degradation, but the negative influence of this oxidative stress does not only affect the protein, in a whole-cell based reaction, it is also involved in the lipid peroxidation, in the cell toxicity and this leads to cell apoptosis.<sup>[184]</sup> Therefore, to enhance the coupling efficiency and to avoid the production of ROs, the generation of artificial fusion proteins has been widely performed. The chosen system CPR2<sub>mut</sub> is one example of the successful protein chimera displaying higher activity and greater coupling efficiency compared to the wild type enzyme.<sup>[73]</sup> Other examples describe the creation of diverse P450 heme domains fused to the reductase domain of P450<sub>BM3</sub>, or P450<sub>RhF</sub>, natural fusion proteins used as model catalyst.<sup>[108]</sup> As an example, the modification of the linker region of the protein chimera P450<sub>cam-RhFRed</sub> was shown indispensable for efficient electron transfer, increasing stability and improving catalysis.<sup>[110,185]</sup> The stability of an enzyme can be enhanced by immobilization, enabling recycling, long-term operation, storage and better separation catalyst/product for repeated re-use of the enzyme.<sup>[186]</sup> However only a few examples report such method applied to P450-based reactions, hampered by the complexity of structure requiring redox partner for catalysis. Nonetheless, sol-gel matrix immobilization was demonstrated to increase the stability and the activity of P450<sub>BM3</sub> towards  $\beta$ -ionone, octane and naphthalene.<sup>[187,188]</sup>

Higher yields of product and faster initial rates of reaction can be obtained under controlled environment, especially for poor soluble substrate. An accurate control of the pH, the stirring and the aeration in fermenter compared to shaking flasks has been shown to be essential to enhance a process.<sup>[162]</sup> Despite a better regulation of parameters, the protein expression influencing the biotransformation process is variable between different batches of fermentation in both scale conditions. The variation from clone to clone or culture to culture was shown to result from extrinsic noise corresponding to fluctuation in concentrations of biomolecules (e.g. ribosomes, polymerases) and intrinsic noise depending on the expression system (promoter, regulator, inducer).<sup>[189]</sup> In fermenter, the choice of the *E. coli* strain seems also critical to ensure plasmid stability and high protein expression as large differences were observed between *E. coli* BL21 (DE3) and HMS174 (DE3), the latest one leading to more efficient biotransformations.<sup>[149]</sup> Considering these possible fluctuations affecting the protein expression, the CO spectra measurements and correlation with biocatalyst concentration is one option for P450-based reactions to compare batches and efficiency of biotransformation.<sup>[190,191]</sup> Regarding the potential of activity of the enzyme, improvements can be achieved through the utilization of two main evolution techniques: rational protein design or directed evolution.<sup>[192]</sup>



## Chapter 2: Generation of mutant libraries and screening for improved activity

Evolutionary tools provide improved biocatalysts regarding activity, selectivity, stability or specificity.<sup>[192]</sup> This involves the generation of mutant libraries through gene shuffling, rational design or directed evolution approaches, followed by the screening of enhanced variants for a specific functionality.

Gene shuffling method consists in the *in vitro* recombination of homologous mutant genes to generate chimera sequences by random fragmentation and PCR reassembly. This method generally requires high level of sequence similarity between the genes to be shuffled in order to avoid the generation of parental genes in majority compared to the chimera sequences.<sup>[193]</sup> Pioneering work was performed with  $\beta$ -lactamase as enzyme model and nowadays array protocols for gene recombination have been established.<sup>[194]</sup> The principle of directed evolution is based on the generation of molecular diversity via random mutagenesis through generally the use of error prone PCR.<sup>[195]</sup> Arnold's lab has provided numbers of successful engineered P450<sub>BM3</sub> through directed evolution, performing non-conventional reactions.<sup>[196,197]</sup> The selection of enhanced variants for a purpose of interest is based on high-throughput screening assay or a method of selection.<sup>[198]</sup> On one hand the major advantage of such approach is the identification of mutations in the absence of structural data, and a gain in understanding the mechanism and structure functionalities, on the other hand, the experimental method suffers from the degeneracy of the genetic code.<sup>[199]</sup> Furthermore high-throughput screening assay or selection methods have to be developed to identify improved variants within large-size libraries of mutants created.<sup>[200]</sup> Prior to this study, considering the absence of a reliable high-throughput screening method to identify improved variants of CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> for the terminal hydroxylation of fatty acids, a rational design approach was the preferred method to create the mutant libraries.

Rational design strategy can be applied with the knowledge of the sequence, structure and function of the protein, with computational estimations as well.<sup>[198]</sup> Algorithms, docking, quantum-mechanical (QM) and molecular dynamics (MD) simulations allow the prediction of the impact of amino acid substitutions on protein structure and stability, leading to the selection of a limited number of residues to exchange. Bioinformatics studies provide directions to alter the substrate specificity, stereoselectivity and stability of the enzyme. As an example, P450<sub>BM3</sub> enzyme has been widely engineered after rational design and tested for new functionalities or improved activities.<sup>[201]</sup> In the absence of crystal structure, the first step in the rational design approach to identify suitable positions for mutagenesis is the creation of a homology model. The importance of homology modelling increases and provides reliable predictions of protein

structure, also contributing to understand unknown mechanisms and protein functions.<sup>[202]</sup> As an example, to gain insights into the structure and the function of CYP106A2, difficult to crystallise, mutant libraries based on rational design and docking studies were created.<sup>[203]</sup> The predictions led to the production of variants with a changed activity and selectivity of hydroxylation towards steroids and this was experimentally demonstrated. In our system, the substrate C<sub>12:0</sub> was docked into a homology model of the heme domain of CYP153A<sub>M.aq.</sub> and positions located within the active and at the substrate entrance tunnel and potentially interacting with the substrate, according to the model were selected for mutation.<sup>[151]</sup> The choice of residues selected for substitution was based on a visual assessment and on the conservation analysis after multiple CYPs sequences alignments.<sup>[84]</sup> Only residues with high frequencies at the selected positions and enabling the access/exit of the substrate/product to the binding pocket, were taken into consideration.<sup>[84,152]</sup>

A small focused library of CYP153A<sub>M.aq.</sub> has already been reported and targeted positions within the active site, therefore new substitutions were tested for such selected positions and new positions from the binding pocket were subjected to site-directed mutagenesis with the replacement of one or two amino acids (with high frequency).<sup>[128]</sup> A degenerated codon was used for positions located at the substrate entrance tunnel, enabling the substitution of two to five amino acids and increasing the diversity. Degenerated codon libraries attract attention as they target specific regions creating an enriched library with a limited size.<sup>[204]</sup> As an example, the CASTing technique introduced in 2005 helps to reshape the binding pocket and was used to alter the enantioselectivity employing a degenerated codon.<sup>[205]</sup> By this method, focused or reduced-size libraries are created, targeting specific sites in the protein.<sup>[153,206]</sup>

Consecutive mutagenesis steps, with the generation of double mutants and one triple mutant were performed successively to the identification of improved single variants. The screening *in vitro* of single variants for improved activity towards short to long chain-length fatty acids led to the identification of three promising positions located in the substrate entrance tunnel and one position from the binding pocket of high interest for further investigations. The mutants S140R, S233G and T302I displayed higher activity than the wild type towards C<sub>12:0</sub> (1.9-fold, 1.23-fold and 1.18-fold increase, respectively). Besides the variants S140R and S233G also with increased activity towards C<sub>16:0</sub> (3-fold and 2.4-fold improvement, respectively), the variant S140L presented 2.9-fold increase compared to the wild type. A loss of activity towards the three tested chain-length can be observed in the case of most of the substitutions addressed for positions in the active site. Substitutions in the binding pocket seem to have a negative impact on the activity and strongly affect the substrate specificity. As an example, most of the mutated positions in our fusion construct have beneficial influence for converting short and medium

chain-length but resulted in lack of activity for longer molecule as  $C_{16:0}$  tested. A remodelling of the active site by modification of the residues in close proximity to the prosthetic group has shown impacts on mediating the substrate specificity and regioselectivity.<sup>[207-209]</sup> Other studies on six mammalian CYP subfamilies suggests a common pattern involving residues at the entrance of the active site responsible for the substrate recognition and influencing the substrate specificity.<sup>[210]</sup>

The substrate access channel was also shown as a critical region in proteins involved in the substrate specificity enabling or preventing from the ligand access/exit.<sup>[211]</sup> Studies performed with CYP1A1 suggest substrate size discrimination from residues located at the substrate entrance tunnel. A two-step recognition process must occur with a first binding of the substrate at the surface of the entrance and a second binding of the ligand within the active site. Same observations were made comparing P450<sub>cam</sub>, P450<sub>BM3</sub> and CYP107A1, sharing the same overall fold and a similar substrate entrance tunnel, but displaying different mechanisms for the substrate access.<sup>[212]</sup> Simulations permitted the identification of several routes for access/release of the ligand but only one shared by the three proteins. The distinct and specific rearrangement of the substrate access pathways differing between the three proteins was shown highly dependent on the substrate properties. Most of the mutated positions located at the substrate access tunnel in CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub> improved the activity as demonstrated by the mutants S453A, S140R, S140L and S233G or remain identical to the wild type activity for the three different chains, suggesting no influence on the substrate specificity. This can be explained by a rather similar size and common physico-chemical properties shared between the three different substrates tested suitable for the access and release of the fatty acid ligand.

The next step consisting in combining the best single variants showing high activity towards  $C_{12:0}$  or  $C_{16:0}$  led to the creation of four double mutants combined with G307A: G307A/S140R, G307A/S140L, G307A/S233G and G307A/T302I. One double mutant deriving from two new selected mutations and one triple mutant were additionally generated and tested, S233G/T302I and G307A/S233G/T302I. The most active variant towards short to long chain-length fatty acids include the position located in the binding pocket G307A and S233G from the substrate access tunnel. This mutant displayed respectively towards  $C_{12:0}$ ,  $C_{16:0}$  and  $C_{8:0}$ , 2-fold, 4.8-fold and 14-fold improvement compared to the wild type activity and 1.8-fold, 1.4-fold and 2.3-fold increase compared to the single mutant G307A. The conversion of the different fatty acids tremendously improved *in vitro* can be explained by the fact that the exchange of a polar residue as serine by a glycine may result in either changes in opening/closing conformation of the binding pocket or reduce non-optimal hydrogen bonds with the substrate or product. The rate of reaction was also shown faster for this double mutant with a complete depletion of 1 mM

substrate reached after 30 minutes of reaction only. Further investigations *in vitro* regarding substrate and product inhibition suggest a greater resistance to large concentrations of substrate and product by this double mutant also leading to an increased stability of the biocatalyst. Indeed this double mutant showed resistance to concentrations up to 2 mM of  $C_{12:0}$  provided compared to 0.5 mM of substrate displaying inhibitory effects to G307A. In the presence of 2 mM of  $\omega$ -OHC $_{12:0}$ , the activity of the double mutant was abolished against 0.2 mM of product in the case of G307A. Comparing to G307A, similar observations can be made concerning the triple mutant G307A/S233G/T302I towards  $C_{12:0}$  regarding product resistance with an activity abolished for 2 mM of product present and the rate of reaction corresponded to 2-fold increase compared to the wild type with a complete depletion of the substrate reached after 1 hour of reaction. As previously discussed, mutations within the active site most likely have a high impact on the substrate specificity as G307A (and T302I) in this double (or triple) mutant but the catalytic activity was also shown enhanced in P450<sub>BM3</sub> by mutating substrate channel positions more distant from the active site, corresponding to S233G in our system.<sup>[213]</sup> Besides, studies have shown that mutations in the catalytic site that often increase the activity are poorly optimized for the stability, therefore the stability observed for the double and triple mutant can be attributed to the mutation part of the substrate entrance channel as well.<sup>[181]</sup>

Interestingly, the improved activity observed for the single variant S140R (1.9-fold increase compared to the wild type) had no longer beneficial impact after combining this position from the substrate entrance tunnel with G307A. The sudden difficulty of expression of this double mutant suggests an inappropriate folding of the protein and therefore a loss of activity towards any fatty acids tested. A leucine residue substituted at this same position instead of an arginine (S140L) does not have any consequences on the expression but the double mutant G307A/S140L does not have any greater effects on activity towards short and medium chain-length in comparison to the wild type enzyme (2.4-fold enhancement only towards  $C_{16:0}$  compared to the wild type). The dynamics of protein created by mutations and the mechanisms by which residues increase the expression or the stability of the protein still remains unclear.<sup>[214]</sup>

The mutation T302I combined with the mutation G307A, two positions located in the binding pocket of the enzyme, showed a beneficial effect on the activity towards medium chain-length  $C_{12:0}$  and short chain-length  $C_{8:0}$  (respectively, 1.9-fold and 7-fold increase compared to the wild type activity). Although the rate of reaction evaluated remained lower than the double mutant S233G/G307A or the triple mutant G307A/S233G/T302I with a substrate depletion of 95 % reached after 2 hours of reaction. Moreover, a dramatic loss in the activity for the conversion of  $C_{16:0}$  was observed: the single mutant T302I and any combination with the position T302I as G307A/T302I, S233G/T302I and the triple mutant G307A/S233G/T302I



displayed a reduced or a complete abolished activity towards long chain-length fatty acid. This can potentially be explained by the lack of space or a blocking conformation of the binding pocket after replacing the threonine by an isoleucine preventing from catalysis of such long fatty acid substrate and/or stopping the release of the product.

*In vivo* catalysis reactions correlate with the *in vitro* results at the exception that the fastest rate and the highest product concentration after 2 hours reaction was evaluated with the triple mutant, followed by the double mutant G307A/S233G and the single variant G307A. The improved rate of reaction and product concentration for the triple mutant (0.21 mM of  $\omega$ -OHC<sub>12:0</sub> per  $\mu$ M of P450 after 2 hours reaction compared to 0.11 mM of  $\omega$ -OHC<sub>12:0</sub> per  $\mu$ M of enzyme for the wild type) can be explained by several factors as great stability for high concentrations of substrate/product also observed *in vitro*. Further validation of the performance of the selected improved variants would require to upscale the process as described in *Chapter 1*. The fusion construct CPR2<sub>mut</sub> is well expressed in *E. coli* HMS 174 (DE3) and the process was optimized in 1 L fermenter. The next step would include the expression in the mentioned *E. coli* host of the best variants and a comparison in the concentration of product formed and efficiency of expression between the different candidates at larger scale.



### *Chapter 3: Establishment of a whole-cell screening assay for identification of improved P450-activity*

The identification of functional molecules and effective engineered P450s requires reliable, robust, sensitive, efficient and simple to implement screening or selection methods.<sup>[215]</sup> The performance of a screening assay or a selection technique is critical in the case of directed evolution techniques employed as it must permit large number of samples to be tested in a reduced time and in a cost-effective way.<sup>[216]</sup> Furthermore the screening method should be sensitive to allow the detection of quantities in the micromolar range.<sup>[200]</sup> Typical selection methods rely on the synthesis of essential amino acids, nucleotides and the degradation of toxic elements in solid or liquid cultures. Screening-based methods *in vitro* and *in vivo* have been established and are linked to the desired catalytic activity.<sup>[217]</sup>

In the literature, many screening assay have been already reported for the detection of improved P450 activity. A rapid method to screen P450 mutant libraries was developed based on the NAD(P)H depletion in partially lysed cells, which can also be used with purified enzyme.<sup>[218]</sup> During enzymatic activity the consumption of the cofactor is monitored at 340 nm and this helps to rapidly detect active variants and allows the determination of coupling efficiency as well. This technique can however yield misleading results due to uncoupling events; the cofactor is consumed forming reactive oxygen species as hydrogen peroxide instead of the product of interest, especially for unnatural substrates tested.<sup>[215]</sup> Furthermore this method requires the external addition of expensive cofactor NAD(P)H. As an alternative to the NAD(P)H depletion test, colorimetric assay have been developed to easily display the enzymatic reaction. The principle of a surrogate substrate-based colorimetric high-throughput assay is one option. The use of *p*-nitrophenyl (pNA) leads to the yellow form *p*-nitrophenolate monitored at 410 nm and the aldehyde generated after a hydroxylation reaction. P450<sub>BM3</sub> mutant activities were screened towards octane *para*-nitrophenoxy analogue and an improved P450 stability for organic co-solvents has been reported.<sup>[183,196,216]</sup> However this method can be biased by the use of an artificial substrate, only similar to the desired one.<sup>[215]</sup> Another assay described suitable for directed evolution was based on the epoxidation of styrene by P450<sub>BM3</sub>, in which the product can react with  $\gamma$ -(4-nitrobenzyl)pyridine forming a purple colour in the cell lysate.<sup>[200]</sup> A similar technique to screen for terminal hydroxylation of alkane activities was based on the terminal hydroxylation of a methylether derivative by P450s producing formaldehyde, which then can react with the Purpald<sup>®</sup> reagent leading to a dark purple colour.<sup>[219]</sup> Other methods involve indole which can be hydroxylated by mono- and dioxygenases such as styrene monooxygenase and naphthalene dioxygenase and upon air oxidation this subsequently leads to the formation of

indirubin and insoluble indigo colored-dye<sup>[220]</sup>. Solid screening assay on LB-agar plate or in liquid media have been developed; site-specific saturation mutagenesis was performed on P450<sub>BM3</sub>, and screened for enhanced variant activity towards indole substrate.<sup>[221]</sup> Another example report P450<sub>cam</sub> mutant libraries generated and the identification of variants catalysing the indigo formation.<sup>[220]</sup> High-performance liquid chromatography (HPLC) is often used in the pharmaceutical industry as standard method for P450 activity and inhibition estimation. In accordance to a high-throughput screening aspect, microtiter plate assays were developed using P450s oxidizing pro-fluorescent molecule to a fluorescent product.<sup>[222]</sup> An assay was developed using human CYP3A4 and 7-benzoyloxy-4-trifluoromethyl coumarin as probe substrate, enabling the determination of IC<sub>50</sub> of a competitive P450 inhibitor.<sup>[223]</sup> Furthermore, a genetically encoded fluorescent sensor was established to detect the production of hydrogen peroxide by an engineered P450<sub>BM3</sub> within whole cell.<sup>[223]</sup> Luminogenic P450-based assay can also be used as alternatives to the fluorescent probes in drug discovery process for the determination of inhibitory effect of molecules on P450s. The principle of such assay is based on the conversion of a luminogenic substrate by a CYPs into a luciferin product initiating a luminescent signal proportionally correlated to the luciferin produced in the first P450-catalysed step.<sup>[224]</sup>

Among various kind of screening methods established, assays using the enzyme galactose oxidase (GOase) have attracted large attention. Liquid and solid phase techniques have been developed to screen for GOase libraries yielding higher activity and specificity.<sup>[147]</sup> Such assay was of high interest to identify improved activity for the terminal hydroxylation of fatty acids as this protein can catalyse the oxidation of primary alcohols and hydroxylated fatty acids into aldehydes and generates hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>). By coupling a dye and the horseradish peroxidase (HRP), the activity of the GOase can be spectrophotometrically monitored. The use of ABTS dye, in the presence of H<sub>2</sub>O<sub>2</sub> and HRP, leads to the polymerization of the ABTS forming a soluble green-coloured product, for which the maximum of absorption can be monitored at 420 nm. The mutant GOase<sub>M3-5</sub> was the chosen catalyst due to its ability to convert primary alcohols and  $\omega$ -OHFA and each component is strictly required for a successful assay (Appendix A5). The overall process from the P450-based reaction until the final analysis and characterization via the GOase of active/improved variant would require 6.7 hours for 96 clones, thus this would enable the screening of 100-200 clones per day (Appendix A6).

In order to provide an efficient system of detection, the first key step in such process involves the choice of appropriate parameters as suitable additional enzymes for the assay, substrate form, buffers and monitoring technology.<sup>[225]</sup> Detection of substrate or cofactor depletion, monitoring the product formation or the direct binding ligand-enzyme are three

possible strategies to adopt in a cell-free system.<sup>[225]</sup> Commonly the development of an assay commences with a purified enzyme or cell lysate to verify the biocatalysis of interest, to establish the crucial parameters for each component as molarity, pH, salt concentration for the buffer, amount of enzyme and substrate, temperature of incubation or cofactor requirement. It is important to consider the degree of complexity of the setup and the availability of reagents or enzyme necessary for the assay reaction.<sup>[226]</sup> The GOase assay has the benefit of requiring only a plate reader as monitoring tool, and the ABTS dye, the HRP and the GOase enzymes are commercially available. Although auto-oxidation and lack of sensitivity were reported, ABTS has the advantage of being non-toxic, stable at high temperature, environmentally friendly and suitable for many reactions. It is also crucial to ensure the stability of each element over time during a reaction course. Such preliminary experiments were performed in the case of the GOase assay using a potassium phosphate buffer, purified GOase, HRP, ABTS and by testing different amounts of enzyme, volume of dye and different concentrations of the P450 product,  $\omega$ -OHC<sub>12:0</sub>. The assay was shown successful in a buffer system, the oxidized ABTS turning into a green colour, but no activity could be detected using cell lysates (data not shown) and this might be due to extra enzyme activities still present in the crude cell extract and potentially inhibiting the GOase activity.

A major advantage in the GOase screening assay in comparison to each above mentioned assays is the direct use of the substrate of interest ( $\omega$ -OHFA) instead of unnatural substrates to generate colorimetric or fluorescent signal. Another advantage from our approach is the low detection limit. The detection of improved variants in terms of activity, specificity or regioselectivity must be highly sensitive; the GOase assay has a detection limit of product of [10-20  $\mu$ M] range for  $\omega$ -OHC<sub>12:0</sub> and therefore successfully enables to distinguish improved active variants and/or non-active one. The mutant GOase<sub>M3-5</sub> was also shown highly selective towards the terminal hydroxylated compounds. Indeed, the product mixture ( $\omega$ -<sub>1,2,3</sub>-OHC<sub>12:0</sub>) generated by the biotransformation of C<sub>12:0</sub> by P450<sub>BM3</sub> could not be detected via the GOase assay. This reinforces the utilization of such screening assay to identify highly selective mutants for the terminal hydroxylation of great interest for industrial applications.

The use of purified enzymes in a buffer system for the screening is not a suitable high-throughput screening manner and the cell lysate technique was found unsuccessful in our approach, therefore we established a whole-cell screening assay. A whole cell system benefits from the avoidance of additional cells lysis and protein purification steps but it might be difficult to achieve accurate performance and results considering reproducibility issues and background noise.<sup>[191]</sup> In our case, the heating step at 90 °C for 30 minutes was found indispensable to prevent from external metabolic activity within the supernatant potentially leading to a lag

phase, with misleading results and non-reproducible signals.<sup>[227]</sup> The limited reproducibility of biotransformation from one batch of expression to another batch can be circumvented by the correlation of protein content based on the CO spectra method.<sup>[76,191]</sup> By combining the GOase assay and an evaluation of protein concentration for each variant, a greater accurate comparison between variants and wild type were demonstrated. Prior to the estimation of the protein content, the results of the GOase can be biased by the P450-biocatalysis only based on the cell density emphasizing the well-expressed variants. The single mutant S453A identified via *in vitro* activity in *Chapter 2* was confirmed as improved variant regarding activity towards C<sub>12:0</sub> after application of the GOase assay combined with the CO measurement (after whole-cell catalysis). The loss of activity observed for the variants G307R and F455E via the assay (absorption similar to the background reaction, empty vector) and after *in vitro* tests, confirmed the possibility of screening and discrimination between improved active variants and non-active mutants.

A non-negligible parameter in the development of a screening assay is the time. The GOase assay can be applied directly on the supernatant after the P450-based reaction performed in whole cells, the only required step prior to the assay is the heating procedure for 30 minutes. Then, the maximum of absorption displayed on the plate reader can be evaluated after 1 h incubation with the GOase, HRP and ABTS. The standard procedure to analyse fatty acids production via GC-FID would necessitate more than 46 hours to carry out the different steps for the preparation of 96 samples for the GC analysis and including the time of method measurement. The GOase assay would require supposedly 6 h only considering the samples preparation and the analysis method on plate reader for 96 clones to test (Appendix A6).

Thus, the GOase assay presents several advantages including the screening of 100-200 clones per day in microtiter plate and enabling the discrimination between improved active variants and non-active one towards C<sub>12:0</sub> with a low detection limit of product. This technique considerably shortens the time to identify active P450s in comparison to the classic methods of fatty acids extraction and GC analysis (the screening of 100 clones would require two days). Each component used for the assay can be commercially available and this colorimetric assay can be directly performed on the supernatant from whole-cell based-reactions with the characterization of the product of interest ( $\omega$ -OHFA, in our system). Furthermore this screening assay combined with the measurement of the P450 content via CO measurement allows an accurate identification of improved activity. The system would require further optimisations regarding the background signal and improvement on the limit of detection. Moreover, this assay was tested on a rather small ratio of samples as the system was demonstrated with seven different variants only. Therefore further steps would require the screening of a large number of clones (200-1000) and the identification of improved mutants through such methods.

## Conclusion and outlook

Oxygenated fatty acids are of high interest for industrial applications in the chemicals, cosmetics or perfumes industries. The production of terminal hydroxylated fatty acids and dicarboxylic acids can be achieved via the biological route with yeast.<sup>[59]</sup> Intensive strain engineering has improved the production of OHFA and DCA but the system still suffers from low space-time yield and the requirement of a specific technology for industrial implementation.<sup>[73]</sup> The heterologous expression of P450s in bacterial cells as *E. coli* is a promising alternative to the yeast. However the performance of recombinant CYPs in *E. coli* is rather poor in comparison to the eukaryotic systems.

Investigation of process limitations related to P450-catalysed reactions in *E. coli* is a first step towards an industrial application. Subsequently to the identification of bottlenecks, strategies can be established to circumvent the challenges and to improve the overall process. Evaluation of limitations reported in the literature and associated to CYPs and to whole-cells have constituted the basis for the studies of challenges associated to artificial fusion CPR2<sub>mut</sub>.<sup>[107,116]</sup> CPR2<sub>mut</sub> was selected for investigations regarding its high activity, regioselectivity and coupling efficiency to produce  $\omega$ -OHFA, medium to long chain-length.<sup>[128]</sup> The fusion construct was also shown well expressed in the *E. coli* host and enabled the production of 4 g L<sup>-1</sup> of hydroxylated fatty acid methyl ester in a two-phase system in 1 L fermenter.<sup>[73]</sup> For this study, the chosen model substrate was C<sub>12:0</sub>, well converted by the biocatalyst and deep characterizations of the protein chimera were made *in vitro* and *in vivo* at small to larger scale.

Reported limitations as substrate/product inhibition were also observed with our system and the application of the solid substrate was presented as a promising option to enhance the production of terminal hydroxylated product in stirred tank reactor. The choice of the feeding strategy of substrate in bioreactor under high speed stirring has been shown essential for an efficient biotransformation. The fed-batch approach, consisting in a slow supply of the substrate along the reaction, has shown improved concentrations of product formed in comparison to a classic batch with the substrate provided at the start. The best rate of reaction and the highest concentration of hydroxylated products were achieved with the application of the solid form of substrate. This can be explained by the combination of a slow dissolving of the fatty acids substrate along the reaction course and the avoidance of co-solvent to dissolve fatty acid, reported harmful to the biocatalyst and to the cells. Beyond 10 mM of substrate provided in the 1 L fermenter dissolved in DMSO, the system must be confronted to substrate inhibition preventing from any conversion of the substrate (observed for 15 mM substrate supplemented).

The application of the solid substrate was tested only for 10 mM concentration of substrate in this study and testing higher concentrations by this approach could tackle the limiting concentration abovementioned using co-solvent DMSO. Efforts in the improvement of the whole-cell process in fermenter using the feeding strategy of the solid substrate and tested with short to long chain-length of fatty acids, still need to be performed.

Regarding the low activity of the chosen biocatalyst, described as a limiting parameter for industrial process, evolutionary techniques can be employed to tailor enzymes and to improve activity towards fatty acids. A small focused library, targeting positions within the active site had led to the identification of the mutant G307A, more active towards fatty acids in the range of  $C_{12:0}$ - $C_{16:0}$ .<sup>[128]</sup> The creation of a homology model and docking studies of the substrate  $C_{12:0}$  into the model helped to initiate the second study regarding the selection of suitable positions interacting with the substrate in the binding pocket and located at the substrate entrance tunnel. As a shift in the substrate specificity was not targeted, enriched mutant libraries were created by the use of a degenerated codons on positions located at the substrate access channel. Besides playing an important role for the substrate specificity, the positions located in close proximity to the heme domain were addressed by site-directed mutagenesis and screened for improved catalytic activity. The best mutant described G307A/S233G resulted in the combination of a position located in the active site (G307A) and a mutation from a position in the substrate entrance tunnel (S233G). The improved activity could be explained by the access/exit of the ligand facilitated by the mutation from the substrate entrance tunnel and the improved activity by the mutation within the binding pocket. The enhanced activity can also be explained by an improved stability or resistance to increased concentrations of substrate/product fatty acids. The successful consecutive steps of mutagenesis support the idea of combinatorial approach as the key strategy to improve activity but also regioselectivity, enantioselectivity and stability. Our results indicate that mutating the more distal positions from the heme domain will facilitate the access of the substrate to the active site and will enable the exit of the product leading to improved activity. To enhance the overall catalytic activity or to create a shift in the substrate range, mutagenesis should be addressed towards positions located in the active site.

Directed evolution is an alternative approach to generate diversity in the absence of sequence or structural data. However this technique based on random mutations suffers from bias of the genetic code leading to the necessity of oversampling to maximize the library quality. Screening thousands of clones for full coverage can only be possible with a reliable assay or a robust selection method. In parallel to the mutant library created from the artificial fusion construct, a screening technique was established to enable the identification of improved



variants. A mutant of the enzyme GOase<sub>M3-5</sub> could generate H<sub>2</sub>O<sub>2</sub> and oxidize the hydroxylated product formed by P450 and by coupling a HRP and ABTS, signal of absorption at 420 nm can be recorded. The development of an effective whole-cell assay based on the identification of the P450 product was a challenging task including the establishment of each parameter as suitable concentrations of buffer, enzyme, dye and overcoming limiting factors as lag phase or background noise.<sup>[227]</sup> The concept was evaluated with selected mutants generated previously and the identification of enhanced active variants was successfully achieved by comparing the assay and the classic product extraction and GC-FID quantification. The development of a high-throughput screening assay is the first step towards the implementation of saturation mutagenesis or directed evolution to apply on CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub>. By such approach promiscuous reactions, enlarge substrate panel, stability or activity could be screened.

Based on this study using a specific P450 fusion construct, an integrative strategy can be applied to screen enhanced variants and to improve the overall process. Further research can start with the screening and the identification of improved variants for a specific functionality after mutant library generation. Through the whole-cell assay established, random strategies can be considered for mutagenesis and the screening of large-sized mutant libraries can be performed targeting enhanced CYPs towards short to long chain-length fatty acids and alkanes. The fusion construct is well-expressed in *E. coli* host and the investigations and identification of limiting parameters at the CYP level and related to whole-cell biotransformation can constitute the basis for further characterizations of the new generated mutants. Moreover, the model substrate of medium chain-length fatty acids provided as solid form was shown beneficial regarding the rate of reaction and the product concentration using the model fusion biocatalyst. The knowledge acquired to improve the process in bioreactor by applying adequate feeding strategy can also be applied to new mutants created and will improve the production of compounds at larger scale as the first step towards industrial implementation.



## References

- [1] G. H. Bruthland, *Our Common Future Report of the World Commission on Environment and Development*, **1987**.
- [2] P. H. Raven, *Science (New York, N.Y.)* **2002**, 297, 954–958.
- [3] W. C. Clark, N. M. Dickson, J. F. Kennedy, H. Universit, *PNAS* **2003**, 100, 8059–8061.
- [4] M. Gavrilescu, Y. Chisti, *Biotechnology advances* **2005**, 23, 471–99.
- [5] D. Elliott, *Renewables*, IOP Publishing, Bristol, UK, **2013**.
- [6] M. Dua, A. Singh, N. Sethunathan, A. K. Johri, *Applied microbiology and biotechnology* **2002**, 59, 143–52.
- [7] G. M. Cooper, *The Cell*, Sinauer Associates, **2000**.
- [8] A. Schmid, J. S. Dordick, B. Hauer, A. Kiener, M. Wubbolts, B. Witholt, *Nature* **2001**, 409, 258–268.
- [9] B. Zechendorf, *Trends in Biotechnology* **1999**, 17, 219–225.
- [10] T. Sutherland, I. Horne, K. Weir, C. Coppin, M. Williams, M. Selleck, R. Russell, J. Oakeshott, *Clinical and Experimental Pharmacology and Physiology* **2004**, 31, 817–821.
- [11] S. Kumar, *Expert opinion on drug metabolism & toxicology* **2010**, 6, 115–31.
- [12] E. L. Rylott, R. G. Jackson, J. Edwards, G. L. Womack, H. M. B. Seth-Smith, D. a Rathbone, S. E. Strand, N. C. Bruce, *Nature biotechnology* **2006**, 24, 216–9.
- [13] C. Scott, G. Pandey, C. J. Hartley, C. J. Jackson, M. J. Cheesman, M. C. Taylor, R. Pandey, J. L. Khurana, M. Teese, C. W. Coppin, et al., *Indian Journal of Microbiology* **2008**, 48, 65–79.
- [14] L. E. Erickson, *Choice: Current Reviews for Academic Libraries* **2005**, 42, 883.
- [15] Y. Gao, Y. Bach, P. Cacioli, P. Butler, I. Louis, *Enzyme and Microbial Technology* **2014**, 54, 38–44.
- [16] J.-M. Choi, S.-S. Han, H.-S. Kim, *Biotechnology advances* **2015**, 33, 1443–54.
- [17] U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, *Nature* **2012**, 485, 185–194.
- [18] D.-K. Ro, E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, J. M. Ndungu, K. A. Ho, R. A. Eachus, T. S. Ham, J. Kirby, et al., *Nature* **2006**, 440, 940–3.
- [19] C. J. Paddon, P. J. Westfall, D. J. Pitera, K. Benjamin, K. Fisher, D. McPhee, M. D. Leavell, A. Tai, A. Main, D. Eng, et al., *Nature* **2013**, 496, 528–32.
- [20] B. Erickson, J. E. Nelson, P. Winters, *Biotechnology Journal* **2011**, 7, 176–185.

- [21] S. N. Naik, V. V. Goud, P. K. Rout, A. K. Dalai, *Renewable and Sustainable Energy Reviews* **2010**, *14*, 578–597.
- [22] C. N. Ibeto, A. U. Ofoefule, K. E. Agbo, *Trends in Applied Sciences Research* **2011**, *6*, 410–425.
- [23] S. E. Sexton, L. A. Martin, D. Zilberman, *Agricultural and Resource Economics Update* **2006**, *9*, 1–4.
- [24] H. Liu, T. Cheng, M. Xian, Y. Cao, F. Fang, H. Zou, *Biotechnology advances* **2013**, *32*, 382–389.
- [25] Q. Kang, L. Appels, T. Tan, R. Dewil, *The Scientific World Journal* **2014**, *14*, 1–13.
- [26] H. Eggert, M. Greaker, E. Potter, *Policies for Second Generation Biofuels: Current Status and Future Challenges*, **2011**.
- [27] N. Sarkar, S. K. Ghosh, S. Bannerjee, K. Aikat, *Renewable Energy* **2012**, *37*, 19–27.
- [28] L. Brennan, P. Owende, *Renewable and Sustainable Energy Reviews* **2010**, *14*, 557–577.
- [29] L.-B. Yonghua, G. Peltier, *Oilseeds & fats Crops and Lipids* **2013**, *20*, D606.
- [30] J. C. Philp, R. J. Ritchie, J. E. M. Allan, *Trends in biotechnology* **2013**, *31*, 219–22.
- [31] B. . Hermann, K. Block, M. K. Patel, *Environ. Sci. Technol* **2007**, *41*, 7915–7921.
- [32] B. M. Nestl, S. C. Hammer, B. A. Nebel, B. Hauer, *Angewandte Chemie (International ed. in English)* **2014**, *53*, 3070–95.
- [33] R. A. Sheldon, *Royal Society of Chemistry* **2014**, *16*, 950–963.
- [34] C. Gao, C. Ma, P. Xu, *Biotechnology advances* **2011**, *29*, 930–9.
- [35] A. Cukalovic, C. V. Stevens, *Biofuels, Bioproducts and Biorefining* **2008**, *2*, 505–529.
- [36] M. Berovic, M. Legisa, *Biotechnology annual review* **2007**, *13*, 303–43.
- [37] J. Schrader, J. Bohlmann, *Advances in Biochemical Engineering/Biotechnology* **2015**, 483–487.
- [38] B. Cok, I. Tsiropoulos, A. L. Roes, M. K. Patel, *Biofuels, Bioproducts and Biorefining* **2014**, *8*, 16–29.
- [39] M. Villela Filho, C. Araujo, A. Bonfá, W. Porto, *Enzyme Research* **2011**, *2011*, 1–8.
- [40] G. Kaur, A. K. Srivastava, S. Chand, *Biochemical Engineering Journal* **2012**, *64*, 106–118.
- [41] T. Dishisha, L. P. Pereyra, S.-H. Pyo, R. A. Britton, R. Hatti-Kaul, *Microbial cell factories* **2014**, *13*, 76.

- [42] H. Yim, R. Haselbeck, W. Niu, C. Pujol-Baxley, A. Burgard, J. Boldt, J. Khandurina, J. D. Trawick, R. E. Osterhout, R. Stephen, et al., *Nature chemical biology* **2011**, 7, 445–52.
- [43] M. Pagliaro, M. Rossi, *Future of Glycerol*, Royal Society Of Chemistry, Cambridge, **2010**.
- [44] A. Schirmer, M. A. Rude, X. Li, E. Popova, S. B. del Cardayre, *Science (New York, N.Y.)* **2010**, 329, 559–62.
- [45] E. J. Steen, Y. Kang, G. Bokinsky, Z. Hu, A. Schirmer, A. McClure, S. B. Del Cardayre, J. D. Keasling, *Nature* **2010**, 463, 559–62.
- [46] R. J. Fox, G. W. Huisman, *Trends in biotechnology* **2008**, 26, 132–8.
- [47] W. Higashide, Y. Li, Y. Yang, J. C. Liao, *Applied and environmental microbiology* **2011**, 77, 2727–33.
- [48] J. V. Seppälä, H. Korhonen, J. Kylmä, J. Tuominen, in *Biopolymers Online* (Eds.: Y. Doi, A. Steinbüchel), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, **2005**, pp. 327–369.
- [49] A.-M. Clarinval, J. Halleux, in *Biodegradable Polymers for Industrial Applications*, Elsevier, **2005**, pp. 3–31.
- [50] M. L. Focarete, C. Gualandi, M. Scandola, M. Govoni, E. Giordano, L. Foroni, S. Valente, G. Pasquinelli, W. Gao, R. A. Gross, *Journal of biomaterials science. Polymer edition* **2010**, 21, 1283–96.
- [51] M. Desroches, M. Escouvois, R. Auvergne, S. Caillol, B. Boutevin, *Polymer Reviews* **2012**, 52, 38–79.
- [52] G. M. Whited, F. J. Feher, D. A. Benko, M. A. Cervin, G. K. Chotani, J. C. McAuliffe, R. J. LaDuca, E. A. Ben-Shoshan, K. J. Sanford, *Industrial Biotechnology* **2010**, 6, 152–163.
- [53] N. Farmer, *Trends in Packaging of Food, Beverages and Other Fast-Moving Consumer Goods (FMCG)*, Elsevier, **2013**.
- [54] B. N. M. van Leeuwen, A. M. van der Wulp, I. Duijnste, A. J. A. van Maris, A. J. J. Straathof, *Applied microbiology and biotechnology* **2012**, 93, 1377–87.
- [55] T. Beardslee, S. Picataggio, *Lipid Technology* **2012**, 24, 223–225.
- [56] C. Liu, F. Liu, J. Cai, W. Xie, T. E. Long, S. R. Turner, A. Lyons, R. a Gross, *Biomacromolecules* **2011**, 12, 3291–8.
- [57] S. Huf, S. Krügener, T. Hirth, S. Rupp, S. Zibek, *European Journal of Lipid Science and Technology* **2011**, 113, 548–561.
- [58] J. O. Metzger, U. Bornscheuer, *Applied microbiology and biotechnology* **2006**, 71, 13–22.
- [59] W. Lu, J. E. Ness, W. Xie, X. Zhang, J. Minshull, R. a Gross, *Journal of the American Chemical Society* **2010**, 132, 15451–5.

- [60] U. Biermann, I. Bornscheuer, M. A. R. Meier, J. O. Metzger, H. J. Schaefer, *Angewandte Chemie* **2011**, *50*, 3854–3871.
- [61] A. Abe, K. Sugiyama, *Anti-Cancer Drug* **2005**, *16*, 543–549.
- [62] J. A. Labinger, **2004**, *220*, 27–35.
- [63] M. Bordeaux, A. Galarneau, J. Drone, *Angewandte Chemie (International ed. in English)* **2012**, *51*, 10712–23.
- [64] M. Sono, M. P. Roach, E. Coulter, J. H. Dawson, **1996**, *96*, 2841–2888.
- [65] R. G. Lageveen, G. W. Huisman, H. Preusting, P. Ketelaar, G. Eggink, B. Witholt, *Applied and environmental microbiology* **1988**, *54*, 2924–2932.
- [66] D. J. Koch, M. M. Chen, J. B. van Beilen, F. H. Arnold, *Applied and Environmental Microbiology* **2009**, *75*, 337–344.
- [67] T. H. M. Smits, S. B. Balada, B. Witholt, J. B. van Beilen, *Journal of bacteriology* **2002**, *184*, 1733–1742.
- [68] J. H. Maeng, Y. Sakai, Y. Tani, N. Kato, *Journal of bacteriology* **1996**, *178*, 3695–3700.
- [69] A. Tani, T. Ishige, Y. Sakai, N. Kato, *Journal of bacteriology* **2001**, *183*, 1819–1823.
- [70] R. Bernhardt, *Journal of biotechnology* **2006**, *124*, 128–45.
- [71] A. Beopoulos, J. Verbeke, F. Bordes, M. Guicherd, M. Bressy, A. Marty, J.-M. Nicaud, *Applied microbiology and biotechnology* **2014**, *98*, 251–62.
- [72] Y. Waché, M. Aguedo, J.-M. Nicaud, J.-M. Belin, *Applied microbiology and biotechnology* **2003**, *61*, 393–404.
- [73] D. Scheps, S. Honda Malca, S. M. Richter, K. Marisch, B. M. Nestl, B. Hauer, *Microbial Biotechnology* **2013**, *6*, 694–707.
- [74] M. Klingenberg, *Archives of Biochemistry and Biophysics* **1958**, *75*, 376–386.
- [75] D. R. Nelson, *Philosophical Transactions of the Royal Society B: Biological Sciences* **2013**, *368*, 20120430.
- [76] T. Omura, R. Sato, *The Journal of biological chemistry* **1964**, *239*, 2370–8.
- [77] Y. Khatri, F. Hannemann, K. M. Ewen, D. Pistorius, O. Perlova, N. Kagawa, A. O. Brachmann, R. Müller, R. Bernhardt, *Chemistry & biology* **2010**, *17*, 1295–305.
- [78] D. C. Lamb, M. R. Waterman, *Philosophical Transactions of the Royal Society B: Biological Sciences* **2013**, *368*, 20120434.
- [79] M. Kubota, M. Nodate, M. Yasumoto-Hirose, T. Uchiyama, O. Kagami, Y. Shizuri, N. Misawa, *Bioscience, biotechnology, and biochemistry* **2005**, *69*, 2421–2430.

- [80] P. R. Ortiz de Montellano, *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 3rd Ed., New York, **2005**.
- [81] K. Auclair, P. Moënne-Loccoz, P. R. Ortiz de Montellano, *Journal of the American Chemical Society* **2001**, *123*, 4877–85.
- [82] S. Yoshioka, S. Takahashi, H. Hori, K. Ishimori, I. Morishima, *European journal of biochemistry / FEBS* **2001**, *268*, 252–9.
- [83] D. R. Nelson, T. Kamataki, D. J. Waxman, F. P. Guengerich, R. W. Estabrook, R. Feyereisen, F. J. Gonzalez, M. J. Coon, I. C. Gunsalus, O. Gotoh, *DNA and cell biology* **1993**, *12*, 1–51.
- [84] L. Gricman, C. Vogel, J. Pleiss, *Proteins: Structure, Function and Bioinformatics* **2014**, *82*, 491–504.
- [85] M. Fischer, M. Knoll, D. Sirim, F. Wagner, S. Funke, J. Pleiss, *Bioinformatics (Oxford, England)* **2007**, *23*, 2015–7.
- [86] J. Park, S. Lee, J. Choi, K. Ahn, B. Park, J. Park, S. Kang, Y.-H. Lee, *BMC genomics* **2008**, *9*, 402.
- [87] D. R. Nelson, *Human genomics* **2009**, *4*, 59–65.
- [88] D. Scheps, S. H. Malca, H. Hoffmann, B. M. Nestl, B. Hauer, *Organic & biomolecular chemistry* **2011**, *9*, 6727–33.
- [89] M. P. Mayhew, V. Reipa, M. J. Holden, V. L. Vilker, *Biotechnology Progress* **2000**, *16*, 610–616.
- [90] E. O'Reilly, M. Corbett, S. Hussain, P. P. Kelly, D. Richardson, S. L. Flitsch, N. J. Turner, *Catalysis Science & Technology* **2013**, *3*, 1490.
- [91] N. Murayama, N. Imai, T. Nakane, M. Shimizu, H. Yamazaki, *Biochemical pharmacology* **2007**, *73*, 2020–6.
- [92] B. Lalovic, B. Phillips, L. L. Risler, W. Howald, D. D. Shen, *Drug metabolism and disposition: the biological fate of chemicals* **2004**, *32*, 447–54.
- [93] J. Venhorst, R. O. B. C. a Onderwater, J. H. N. Meerman, J. a N. N. M. Commandeur, N. P. E. Vermeulen, *Pharmacology* **2000**, *28*, 1524–1532.
- [94] I. C. Gunsalus, T. C. Pederson, S. G. Sligar, *Annual review of biochemistry* **1975**, *44*, 377–407.
- [95] F. P. Guengerich, *The Journal of biological chemistry* **1991**, *266*, 10019–22.
- [96] M. Katagiri, B. N. Ganguli, I. C. Gunsalus, *J. Biol. Chem.* **1968**, *243*, 3543–3546.
- [97] K. He, L. M. Bornheim, A. M. Falick, D. Maltby, H. Yin, M. A. Correia, *Biochemistry* **1998**, *37*, 17448–57.
- [98] I. G. Denisov, T. M. Makris, S. G. Sligar, I. Schlichting, *Chemical reviews* **2005**, *105*, 2253–77.

- [99] T. L. Poulos, B. C. Finzel, A. J. Howard, *Journal of molecular biology* **1987**, *195*, 687–700.
- [100] K. G. Ravichandran, S. S. Boddupalli, C. A. Hasermann, J. A. Peterson, J. Deisenhofer, *Science (New York, N.Y.)* **1993**, *261*, 731–6.
- [101] L. Narhi, A. Fulco, *J. Biol. Chem.* **1986**, *261*, 7160–7169.
- [102] M. A. Noble, C. S. Miles, S. K. Chapman, D. A. Lysek, A. C. MacKay, G. A. Reid, R. P. Hanzlik, A. W. Munro, *The Biochemical journal* **1999**, *339 (Pt 2)*, 371–379.
- [103] 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, et al., *Trends in biochemical sciences* **2002**, *27*, 250–7.
- [104] M. Wang, D. L. Roberts, R. Paschke, T. M. Shea, B. S. Masters, J. J. Kim, *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 8411–6.
- [105] F. Hannemann, A. Bichet, K. M. Ewen, R. Bernhardt, *Biochimica et biophysica acta* **2007**, *1770*, 330–44.
- [106] G. A. Roberts, G. Grogan, A. Greter, S. L. Flitsch, N. J. Turner, *Journal of bacteriology* **2002**, *184*, 3898–908.
- [107] E. O'Reilly, V. Köhler, S. L. Flitsch, N. J. Turner, *Chemical communications (Cambridge, England)* **2011**, *47*, 2490–501.
- [108] P. R. Ortiz de Montellano, *Cytochrome P450 - Structure, Mechanism, and Biochemistry 4th*, **2015**.
- [109] V. R. Dodhia, A. Fantuzzi, G. Gilardi, *Journal of biological inorganic chemistry: JBIC: a publication of the Society of Biological Inorganic Chemistry* **2006**, *11*, 903–16.
- [110] A. Robin, V. Köhler, A. Jones, A. Ali, P. P. Kelly, E. O'Reilly, N. J. Turner, S. L. Flitsch, *Beilstein journal of organic chemistry* **2011**, *7*, 1494–8.
- [111] A. Robin, G. a Roberts, J. Kisch, F. Sabbadin, G. Grogan, N. Bruce, N. J. Turner, S. L. Flitsch, *Chemical communications (Cambridge, England)* **2009**, 2478–80.
- [112] M. Bordeaux, D. de Girval, R. Rullaud, M. Subileau, E. Dubreucq, J. Drone, *Applied microbiology and biotechnology* **2014**, *98*, 6275–83.
- [113] S. M. Hoffmann, M. J. Weissenborn, L. Gricman, S. Notonier, J. Pleiss, B. Hauer, *submitted to ACS Chemical Biology* **2015**.
- [114] C. R. Robinson, R. T. Sauer, *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 5929–34.
- [115] B. Goodwin, M. R. Redinbo, S. A. Kliewer, *Annual review of pharmacology and toxicology* **2002**, *42*, 1–23.
- [116] R. Bernhardt, V. B. Urlacher, *Applied microbiology and biotechnology* **2014**, *98*, 6185–203.



- [117] M. K. Trower, R. Lenstra, C. Omer, S. E. Buchholz, F. S. Sariaslani, *Molecular microbiology* **1992**, *6*, 2125–34.
- [118] H. Ichinose, H. Wariishi, H. Tanaka, *Applied microbiology and biotechnology* **2002**, *58*, 97–105.
- [119] J. S. Yadav, J. C. Loper, *Current genetics* **2000**, *37*, 65–73.
- [120] B. R. Baer, A. E. Rettie, *Drug metabolism reviews* **2006**, *38*, 451–476.
- [121] F. Xu, S. G. Bell, J. Lednik, A. Insley, Z. Rao, L.-L. Wong, *Angewandte Chemie (International ed. in English)* **2005**, *44*, 4029–32.
- [122] R. Fasan, M. M. Chen, N. C. Crook, F. H. Arnold, *Angewandte Chemie (International ed. in English)* **2007**, *46*, 8414–8.
- [123] C. F. Oliver, S. Modi, M. J. Sutcliffe, W. U. Primrose, L. Y. Lian, G. C. Roberts, *Biochemistry* **1997**, *36*, 1567–72.
- [124] M. Girhard, T. Klaus, Y. Khatri, R. Bernhardt, V. B. Urlacher, *Applied Microbiology and Biotechnology* **2010**, *87*, 595–607.
- [125] E. O'Reilly, S. J. Aitken, G. Grogan, P. P. Kelly, N. J. Turner, S. L. Flitsch, *Beilstein journal of organic chemistry* **2012**, *8*, 496–500.
- [126] P. Jurtshuk, G. E. Cardini, D. T. Gibson, *Critical Reviews in Microbiology* **1971**, *1*, 239–289.
- [127] J. B. Johnston, P. M. Kells, L. M. Podust, P. R. Ortiz de Montellano, *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 20687–92.
- [128] S. Honda Malca, D. Scheps, L. Kühnel, E. Venegas-Venegas, A. Seifert, B. Nestl, B. Hauer, *Chemical Communications* **2012**, *48*, 5115.
- [129] S. Cornelissen, M. K. Julsing, J. Volmer, O. Riechert, A. Schmid, B. Bühler, *Biotechnology and bioengineering* **2013**, *110*, 1282–92.
- [130] M. Bordeaux, A. Galarneau, F. Fajula, J. Drone, *Angewandte Chemie (International ed. in English)* **2011**, *50*, 2075–9.
- [131] T. H. M. Smits, S. B. Balada, B. Witholt, J. B. van Beilen, *Journal of bacteriology* **2002**, *184*, 1733–42.
- [132] T. Maier, H. H. Förster, O. Asperger, U. Hahn, *Biochemical and biophysical research communications* **2001**, *286*, 652–8.
- [133] T. Fujii, T. Narikawa, F. Sumisa, A. Arisawa, K. Takeda, J. Kato, *Bioscience, biotechnology, and biochemistry* **2006**, *70*, 1379–85.
- [134] N. Fujita, F. Sumisa, K. Shindo, H. Kabumoto, A. Arisawa, H. Ikenaga, N. Misawa, *Bioscience, biotechnology, and biochemistry* **2009**, *73*, 1825–30.

- [135] R. K. Gudimanchi, C. Randall, D. J. Opperman, O. A. Olaofe, S. T. L. Harrison, J. Albertyn, M. S. Smit, *Applied microbiology and biotechnology* **2012**, *96*, 1507–16.
- [136] T. Vallon, M. Glemser, S. H. Malca, D. Scheps, J. Schmid, M. Siemann-Herzberg, B. Hauer, R. Takors, *Chemie Ingenieur Technik* **2013**, *85*, 841–848.
- [137] J. H. Bae, B. G. Park, E. Jung, P.-G. Lee, B.-G. Kim, *Applied Microbiology and Biotechnology* **2014**, *98*, 8917–8925.
- [138] J. M. Weber, J. O. Leung, S. J. Swanson, K. B. Idler, J. B. McAlpine, *Science (New York, N.Y.)* **1991**, *252*, 114–7.
- [139] S. Jennewein, C. D. Rithner, R. M. Williams, R. B. Croteau, *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 13595–600.
- [140] K. Sonomoto, M. M. Hoq, A. Tanaka, S. Fukui, *Applied and environmental microbiology* **1983**, *45*, 436–43.
- [141] T. Fujii, Y. Fujii, K. Machida, A. Ochiai, M. Ito, *Bioscience, biotechnology, and biochemistry* **2009**, *73*, 805–10.
- [142] J. . Falck, Y. K. Reddy, D. C. Haines, K. M. Reddy, U. M. Krishna, S. Graham, B. Murry, J. A. Peterson, *Tetrahedron Letters* **2001**, *42*, 4131–4133.
- [143] E. M. Gillam, F. P. Guengerich, *IUBMB life* **2001**, *52*, 271–7.
- [144] A. Schmid, J. S. Dordick, B. Hauer, A. Kiener, M. Wubbolts, B. Witholt, *Nature* **2001**, *409*, 258–68.
- [145] D. J. Leak, R. A. Sheldon, J. M. Woodley, P. Adlercreutz, *Biocatalysis and Biotransformation* **2009**, *27*, 1–26.
- [146] M. T. Lundemo, J. M. Woodley, *Applied microbiology and biotechnology* **2015**, *99*, 2465–83.
- [147] F. Escalettes, N. J. Turner, *ChemBioChem* **2008**, *9*, 857–860.
- [148] S. E. Deacon, M. J. McPherson, *Chembiochem: a European journal of chemical biology* **2011**, *12*, 593–601.
- [149] K. Marisch, K. Bayer, M. Cserjan-Puschmann, M. Luchner, G. Striedner, *Microbial cell factories* **2013**, *12*, 58.
- [150] M. Jeske, J. Altenbuchner, *Applied microbiology and biotechnology* **2010**, *85*, 1923–33.
- [151] Y. Yang, J. Liu, Z. Li, *Angewandte Chemie (International ed. in English)* **2014**, *53*, 3120–4.
- [152] Ł. Gricman, C. Vogel, J. Pleiss, *Proteins* **2015**, *83*, 1593–1603.
- [153] M. T. Reetz, *Angewandte Chemie (International ed. in English)* **2011**, *50*, 138–74.
- [154] S. G. Sligar, D. L. Cinti, G. G. Gibson, J. B. Schenkman, *Biochemical and Biophysical Research Communications* **1979**, *90*, 925–932.

- [155] A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, **2012**.
- [156] A. Siriphongphaew, P. Pisnupong, J. Wongkongkatep, P. Inprakhon, A. S. Vangnai, K. Honda, H. Ohtake, J. Kato, J. Ogawa, S. Shimizu, et al., *Applied Microbiology and Biotechnology* **2012**, *95*, 357–367.
- [157] C. E. Paul, E. Churakova, E. Maurits, M. Girhard, V. B. Urlacher, F. Hollmann, *Bioorganic and Medicinal Chemistry* **2014**, *22*, 5692–5696.
- [158] J. D. Ryan, R. H. Fish, D. S. Clark, *Chembiochem : a European journal of chemical biology* **2008**, *9*, 2579–2582.
- [159] O. a Olaofe, C. J. Fenner, R. K. Gudimanchi, M. S. Smit, S. T. L. Harrison, *Microbial cell factories* **2013**, *12*, 8.
- [160] D. H. Rammner, *Annals of the New York Academy of Sciences* **1967**, *141*, 291–9.
- [161] M. V. Donova, O. V. Egorova, *Applied Microbiology and Biotechnology* **2012**, *94*, 1423–1447.
- [162] F. M. Kiss, M. T. Lundemo, J. Zapp, J. M. Woodley, R. Bernhardt, *Microbial Cell Factories* **2015**, *14*, 28.
- [163] S. R. Maloy, C. L. Ginsburgh, R. W. Simons, W. D. Nunn, *The Journal of biological chemistry* **1981**, *256*, 3735–42.
- [164] V. Urlacher, R. D. Schmid, *Current Opinion in Biotechnology* **2002**, *13*, 557–564.
- [165] H. Joo, Z. Lin, F. H. Arnold, *Nature* **1999**, *399*, 670–673.
- [166] C. Sung, E. Jung, K.-Y. Choi, J. Bae, M. Kim, J. Kim, E.-J. Kim, P. Il Kim, B.-G. Kim, *Applied Microbiology and Biotechnology* **2015**, *99*, 6667–6676.
- [167] J. T. Wu, L. H. Wu, J. A. Knight, *Clinical chemistry* **1986**, *32*, 314–9.
- [168] S. C. Maurer, K. Kühnel, L. a. Kaysser, S. Eiben, R. D. Schmid, V. B. Urlacher, *Advanced Synthesis and Catalysis* **2005**, *347*, 1090–1098.
- [169] A. Fantuzzi, Y. T. Meharena, P. B. Briscoe, C. Sassone, B. Borgia, G. Gilardi, *Chemical communications (Cambridge, England)* **2006**, 1289–1291.
- [170] S. J. Sadeghi, G. Gilardi, *Biotechnology and Applied Biochemistry* **2013**, *60*, 102–110.
- [171] N.-H. Tran, D. Nguyen, S. Dwaraknath, S. Mahadevan, G. Chavez, A. Nguyen, T. Dao, S. Mullen, T.-A. Nguyen, L. E. Cheruzel, *Journal of the American Chemical Society* **2013**, *135*, 14484–7.
- [172] T. Kitazume, A. Tanaka, N. Takaya, A. Nakamura, S. Matsuyama, T. Suzuki, H. Shoun, *European journal of biochemistry / FEBS* **2002**, *269*, 2075–82.
- [173] V. Hirunpanich, B. Sethabouppha, H. Sato, *Biological & pharmaceutical bulletin* **2007**, *30*, 1586–1588.

- [174] L. A. Royce, P. Liu, M. J. Stebbins, B. C. Hanson, L. R. Jarboe, *Applied Microbiology and Biotechnology* **2013**, *97*, 8317–8327.
- [175] J. P. Fay, R. N. Farias, *Journal of general microbiology* **1975**, *91*, 233–40.
- [176] C. J. Zheng, J.-S. Yoo, T.-G. Lee, H.-Y. Cho, Y.-H. Kim, W.-G. Kim, *FEBS letters* **2005**, *579*, 5157–5162.
- [177] “Lauric acid | C<sub>12</sub>H<sub>24</sub>O<sub>2</sub> | ChemSpider,” can be found under <http://www.chemspider.com/Chemical-Structure.3756.html>.
- [178] “12-Hydroxydodecanoic acid | C<sub>12</sub>H<sub>24</sub>O<sub>3</sub> | ChemSpider,” can be found under <http://www.chemspider.com/Chemical-Structure.71366.html>.
- [179] C. J. C. Whitehouse, S. G. Bell, L.-L. Wong, *Chemical Society reviews* **2012**, *41*, 1218–60.
- [180] Y. Li, D. A. Drummond, A. M. Sawayama, C. D. Snow, J. D. Bloom, F. H. Arnold, *Nature biotechnology* **2007**, *25*, 1051–1056.
- [181] J. D. Bloom, S. T. Labthavikul, C. R. Otey, F. H. Arnold, *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 5869–5874.
- [182] O. Salazar, P. C. Cirino, F. H. Arnold, *ChemBioChem* **2003**, *4*, 891–893.
- [183] T. S. Wong, F. H. Arnold, U. Schwaneberg, *Biotechnology and Bioengineering* **2004**, *85*, 351–358.
- [184] R. C. Zangar, D. R. Davydov, S. Verma, *Toxicology and applied pharmacology* **2004**, *199*, 316–31.
- [185] C. R. Robinson, R. T. Sauer, *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 5929–5934.
- [186] R. a Sheldon, S. van Pelt, *Chemical Society reviews* **2013**, *42*, 6223–35.
- [187] S. C. Maurer, H. Schulze, R. D. Schmid, V. Urlacher, *Advanced Synthesis & Catalysis* **2003**, *345*, 802–810.
- [188] E. Weber, D. Sirim, T. Schreiber, B. Thomas, J. Pleiss, M. Hunger, R. Gläser, V. B. Urlacher, *Journal of Molecular Catalysis B: Enzymatic* **2010**, *64*, 29–37.
- [189] M. Lindmeyer, D. Meyer, D. Kuhn, B. Bühler, A. Schmid, *Journal of Industrial Microbiology & Biotechnology* **2015**, *42*, 851–866.
- [190] S.-J. Choi, M. Kim, S.-I. Kim, J.-K. Jeon, *Journal of biochemistry and molecular biology* **2003**, *36*, 332–5.
- [191] W. A. Johnston, W. Huang, J. J. De Voss, M. A. Hayes, E. M. J. Gillam, *Journal of biomolecular screening* **2008**, *13*, 135–41.
- [192] U. T. Bornscheuer, M. Pohl, *Current Opinion in Chemical Biology* **2001**, *5*, 137–143.

- [193] P. L. Bergquist, M. D. Gibbs, *Methods in molecular biology (Clifton, N.J.)* **2007**, 352, 191–204.
- [194] W. P. Stemmer, *Nature* **1994**, 370, 389–391.
- [195] M. J. Dougherty, F. H. Arnold, *Current opinion in biotechnology* **2009**, 20, 486–91.
- [196] E. T. Farinas, U. Schwaneberg, A. Glieder, F. H. Arnold, *Advanced Synthesis & Catalysis* **2001**, 343, 601–606.
- [197] P. S. Coelho, E. M. Brustad, A. Kannan, F. H. Arnold, *Science (New York, N.Y.)* **2013**, 339, 307–10.
- [198] S. Lutz, *Current Opinion in Biotechnology* **2010**, 21, 734–743.
- [199] N. E. Labrou, *Current protein & peptide science* **2010**, 11, 91–100.
- [200] M. Alcalde, *Journal of Biomolecular Screening* **2004**, 9, 141–146.
- [201] A. Seifert, M. Antonovici, B. Hauer, J. Pleiss, *ChemBioChem* **2011**, 12, 1346–1351.
- [202] Z. Xiang, *Current protein & peptide science* **2006**, 7, 217–27.
- [203] M. Lisurek, B. Simgen, I. Antes, R. Bernhardt, *Chembiochem: a European journal of chemical biology* **2008**, 9, 1439–49.
- [204] T. M. Jacobs, H. Yumerefendi, B. Kuhlman, A. Leaver-Fay, *Nucleic Acids Research* **2015**, 43, e34.
- [205] M. T. Reetz, J. D. Carballeira, J. Peyralans, H. Höbenreich, A. Maichele, A. Vogel, *Chemistry - A European Journal* **2006**, 12, 6031–6038.
- [206] M. T. Reetz, L.-W. Wang, M. Bocola, *Angewandte Chemie (International ed. in English)* **2006**, 45, 1236–41.
- [207] A. Seifert, J. Pleiss, *Proteins* **2009**, 74, 1028–35.
- [208] B. Dueholm, C. Krieger, D. Drew, A. Olry, T. Kamo, O. Taboureau, C. Weitzel, F. Bourgaud, A. Hehn, H. T. Simonsen, *BMC evolutionary biology* **2015**, 15, 122.
- [209] T. W. . Ost, C. S. Miles, J. Murdoch, Y.-F. Cheung, G. A. Reid, S. K. Chapman, A. W. Munro, *FEBS Letters* **2000**, 486, 173–177.
- [210] M. S. Zharkova, B. N. Sobolev, N. Yu Oparina, A. V Veselovsky, A. I. Archakov, *Journal of molecular recognition : JMR* **2013**, 26, 86–91.
- [211] P. Urban, G. Truan, D. Pompon, *Biochimica et biophysica acta* **2015**, 1850, 696–707.
- [212] P. J. Winn, S. K. Lüdemann, R. Gauges, V. Lounnas, R. C. Wade, *Proceedings of the National Academy of Sciences of the United States of America* **2002**, 99, 5361–6.

- [213] C. F. Butler, C. Peet, A. E. Mason, M. W. Voice, D. Leys, A. W. Munro, *The Journal of biological chemistry* **2013**, *288*, 25387–99.
- [214] S. Kumar, Y. Zhao, L. Sun, S. S. Negi, J. R. Halpert, B. K. Muralidhara, *Molecular pharmacology* **2007**, *72*, 1191–9.
- [215] T. W. Johannes, R. D. Woodyer, H. Zhao, in *Enzyme Assays: High-Throughput Screening, Genetic Selection and Fingerprinting*, **2006**, pp. 77–93.
- [216] U. Schwaneberg, C. Otey, P. C. Cirino, E. Farinas, F. H. Arnold, *Journal of Biomolecular Screening* **2001**, *6*, 111–117.
- [217] D. P. Nannemann, W. R. Birmingham, R. A. Scism, B. O. Bachmann, *Future medicinal chemistry* **2011**, *3*, 809–19.
- [218] G. E. Tsotsou, a. E. G. Cass, G. Gilardi, *Biosensors and Bioelectronics* **2002**, *17*, 119–131.
- [219] P. Meinhold, M. W. Peters, A. Hartwick, A. R. Hernandez, F. H. Arnold, *Advanced Synthesis & Catalysis* **2006**, *348*, 763–772.
- [220] A. Celik, R. E. Speight, N. J. Turner, *Chemical communications (Cambridge, England)* **2005**, *7*, 3652–4.
- [221] Q. S. Li, U. Schwaneberg, P. Fischer, R. D. Schmid, *Chemistry (Weinheim an der Bergstrasse, Germany)* **2000**, *6*, 1531–1536.
- [222] M. T. Donato, *Drug Metabolism and Disposition* **2004**, *32*, 699–706.
- [223] Q. Cheng, C. D. Sohl, F. P. Guengerich, *Nature protocols* **2009**, *4*, 1258–1261.
- [224] B. Larson, P. Banks, J. J. Cali, M. Sobol, S. Shultz, *Journal of laboratory automation* **2011**, *16*, 47–55.
- [225] M. G. Acker, D. S. Auld, *Perspectives in Science* **2014**, *1*, 56–73.
- [226] J.-P. Goddard, J.-L. Reymond, *Current opinion in biotechnology* **2004**, *15*, 314–22.
- [227] T. S. Moon, D. R. Nielsen, K. L. J. Prather, *AIChE Journal* **2012**, *58*, 2303–2308.

## Appendix

### A1. Gene CYP153A<sub>M.aq.</sub>-CPR<sub>BM3</sub>

Sequences from the fusion construct CYP153A from *Marinobacter aquaeolei* (in red, NCBI accession number ABM17701) and the reductase domain of CYP102A1 from *Bacillus megaterium* (in purple, NCBI accession number WP\_034650526). Start and stop codons are underlined.

ATGCCAACACTGCCAGAACATTTGACGACATTCAGTCCCAGACTGATTAACGCCACCTCCAGGGTGGTGCCGATG  
CAGAGGCAAATTCAGGGACTGAAATTCCTAATGAGCGCCAAGAGGAAGACCTTCGGCCACGCCGACCGATGCC  
GAATTCGTTGAAACACCCATCCCGACGTTAACACGCTGGCCCTTGAGGACATCGATGTCAGCAATCCGTTTTTA  
TACCGGCAGGGTCAGTGGCGCGCCTATTTCAAACGGTTGCGTGATGAGGCGCCGGTCCATTACCAGAAGAACAGC  
CCTTTCGGCCCTTCTGGTCGGTAACTCGGTTTGAAGACATCCTGTTTCGTGGATAAGAGTACAGACCTGTTTTCC  
GCCGAGCCGCAAATCATTCTCGGTGACCCTCCGGAGGGGCTGTGCGTGAAATGTTCATAGCGATGGATCCGCCG  
AAACACGATGTGCAGCGCAGCTCGGTGCAGGGAGTAGTGGACCCGAAAACTGAAGGAGATGGAGGGGCTGATC  
CGATCACGCACCGGCGATGTGCTTGACAGCCTGCCTACAGACAAACCTTTAACTGGGTACCTGCTGTTTCCAAG  
GAACTCACAGGCCGATGCTGGCGACGCTTCTGGATTTTCTTACGAGGAACGCCACAAGCTGGTTGAGTGGTTCG  
GACAGAATGGCAGGTGCAGCATCGGCCACCGGCGGGGAGTTTGGCGATGAAAATGCCATGTTTGACGACGCGGCA  
GACATGGCCCGGTCTTTCTCCAGGCTTTGGCGGGACAAGGAGGCGCGCCGCGCAGCAGGCGAGGAGCCCGTTTT  
GATTTGATCAGCCTGTTGACAGCAACAAAGAAACGAAAAGACCTGATCAATCGGCCGATGGAGTTTATCGGTAAT  
TTGACGCTGCTCATAGTCGGCGGCAACGATACGACGCGCAACTCGATGAGTGGTGGCTGGTGGCCATGAACGAA  
TTCCCAGGGAATTTGAAAAATTGAAGGCAAAACCGGAGTTGATTCCGAACATGGTGTGCGAAATCATCCGCTGG  
CAAACGCCGCTGGCCTATATGCGCCGAATCGCCAAGCAGGATGTGAACTGGGCGGCCAGACCATCAAGAAGGGT  
GATCGAGTTGTATGTGGTACGCGTCGGGTAACCGGGACGAGCGCAAATTTGACAACCCCGATCAGTTCATCATT  
GATCGCAAGGACGCAGAACCCACATGTCTGTTCCGCTATGGGTTTACCCTTGCATGGGCAACCGTCTGGCTGAA  
CTGCAACTGCGCATCCTCTGGGAAGAAATACTCAAGCGTTTTGACAACATCGAAGTCGTGGAAGACCCGAGCGG  
GTGCACTCCAACCTTCGTGCGGGCTATTCCAGGTTGATGGTCAAACCTGACACCGAACAGTATTCTTCACCTAGC  
ACTGAACAGTCTGCTAAAAAAGTACGCAAAAAGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGT  
TCAAATATGGGAACAGCTGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCACCGCAG  
GTCGCAACGCTTGATTACACGCCGGAATCTTCCGCGCAAGGAGCTGTATTAATTGTAACGGCGCTTTATAAC  
GGTCATCCGCCTGATAACGCAAAGCAATTTGTGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTT  
CGTACTCCGATTTGGATGCGGCGATAAAAACCTGGGCTACTACGTATCAAAAAGTGCCTGCTTTTTATCGATGAA  
ACGCTTGCCGCTAAAGGGGACAGAAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACA  
TATGAAGAATGGCGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTTGAAAACAGTGAAGAT  
AATAAATCTACTCTTTCACTTCAATTTGTGACAGCGCCGCGGATATGCCGCTTGCGAAAATGCACGGTGCCTTT  
TCAACGAACGTCGTAGCAAGCAAAGAACTTCAACAGCCAGGCAGTGCACGAAGCACGCGACATCTTGAAATTGAA  
CTTCCAAAAGAAGCTTCTTATCAAGAAGGAGATCATTTAGGTGTTATTCTTCGCAACTATGAAGGAATAGTAAAC  
CGTGTAAACAGCAAGGTTCCGGCTAGATGCATCACAGCAAATCCGCTTGGAAGCAGAAAGAAAAAATAGCTCAT  
TTGCCACTCGCTAAAACAGTATCCGTAGAAGAGCTTCTGCAATACGTGGAGCTTCAAGATCCTGTTACGCGCACG  
CAGCTTCGCGCAATGGCTGCTAAAACGGTCTGCCCGCCGATAAAAGTAGAGCTTGAAGCCTTGCTTGAAAAGCAA  
GCCTACAAAGAACAAGTGTGGCAAAACGTTTAAACAATGCTTGAAGTCTTGAAAAATACCCGGCGTGTGAAATG  
AAATTCAGCGAATTTATCGCCCTTCTGCAAGCATAACGCGCGCTATTACTCGATTTCTTCATCACCTCGTGTG  
GATGAAAAACAAGCAAGCATCACGGTCAGCGTTGTCTCAGGAGAAGCGTGGAGCGGATATGGAGAATATAAAGGA  
ATTGCGTCAACTATCTTGCCGAGCTGCAAGAAGGAGATACGATTACGTGCTTTATTTCCACACCGCAGTCAGAA  
TTTACGCTGCCAAAAGACCCTGAAACGCCGCTTATCATGGTTCGGACCGGGAACAGGCGTCGCGCGTTTTAGAGGC  
TTTGTGAGGCGCGCAAACAGCTAAAAGAACAAGGACAGTCACTTGGAGAAAGCACATTTATACTTCGGCTGCCGT  
TCACCTCATGAAGACTATCTGTATCAAGAAGAGCTTGAACCGCCAAAAGCGAAGGCATCATTACGCTTCATACC  
GCTTTTTCTCGCATGCCAAATCAGCCGAAAACATACGTTTACGACGTAATGGAACAAGACGGCAAGAAATGATT  
GAACTTCTTGATCAAGGAGCGCACTTCTATATTTGCGGAGACGGAAGCCAAATGGCACCTGCCGTTGAAGCAACG  
CTTATGAAAAGCTATGCTGACGTTACCAAGTGAGTGAAGCAGACGCTCGCTTATGGCTGCAGCAGCTAGAAGAA  
AAAGGCCGATA

## A2. Mutant library

The following table classifies the different vector constructs generated and used in this study. Each variant created derives from the fusion construct previously described (Material and Method section 2. and Appendix A1.).

**Table A2.** Vector constructs and mutant libraries

Vector name	ITB number
pET30a-Goase <sub>M3-5</sub>	pITB1086
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> + (3xGGS)	pITB728
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (G307A)+(3xGGS) (CPR2 <sub>mut</sub> )	pITB431
pJOE4782.1- CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (G307A)+(3xGGS) (CPR2 <sub>mut</sub> )	pITB469
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub>	pITB1166
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (G307A)	pITB1167
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (R77V)	pITB1168
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (R77A)	pITB1169
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (D134V)	pITB1170
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (D134G)	pITB1171
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (S140R)	pITB1172
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (S140L)	pITB1173
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (S233G)	pITB1174
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (T235L)	pITB1175
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (S453A)	pITB1176
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (S453G)	pITB1177
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (Q129P)	pITB1178
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (I131P)	pITB1179
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (V141G)	pITB1180
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (M143G)	pITB1181
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (M143S)	pITB1182

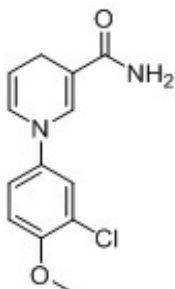


Table A2. (continued)

Vector name	ITB number
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (I145V)	pITB1183
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (I145L)	pITB1184
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (M228L)	pITB1185
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (A229F)	pITB1186
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (A229L)	pITB1187
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (L301A)	pITB1188
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (L301V)	pITB1189
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (T302L)	pITB1190
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (T302I)	pITB1191
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (L303D)	pITB1192
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (L303T)	pITB1193
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (V306F)	pITB1194
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (V306I)	pITB1195
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (G307R)	pITB1196
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (D310E)	pITB1197
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (Y356F)	pITB1198
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (Y356L)	pITB1199
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (M357L)	pITB1200
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (R358P)	pITB1201
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (F455V)	pITB1202
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (G307A/S140R)	pITB1203
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (G307A/S140L)	pITB1204
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (G307A/S233G)	pITB1205
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (G307A/T302I)	pITB1206
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (S233G/T302I)	pITB1207
pET28a(+)-CYP153A <sub>M.aq.</sub> -CPR <sub>BM3</sub> (G307A/S233G/T302I)	pITB1208

### A3. Cofactor analogue

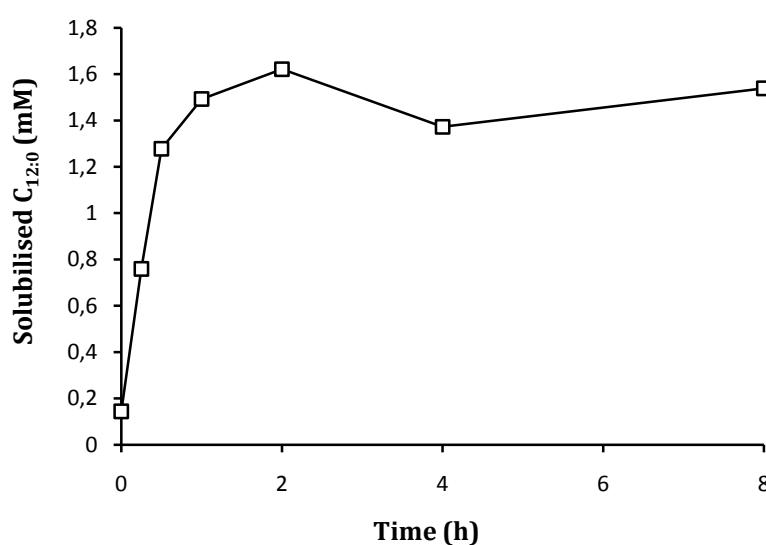
The NADPH analogue, nicotinamide-3,4-chloro-methoxy-phenyl-hydride (termed NACMH) was synthesized and described somewhere else (Dr. Sebastian Löw PhD thesis, ITB, University of Stuttgart).



**Figure A3.** Structure of NACMH, artificial NADPH cofactor.

### A4. Solid feeding strategy

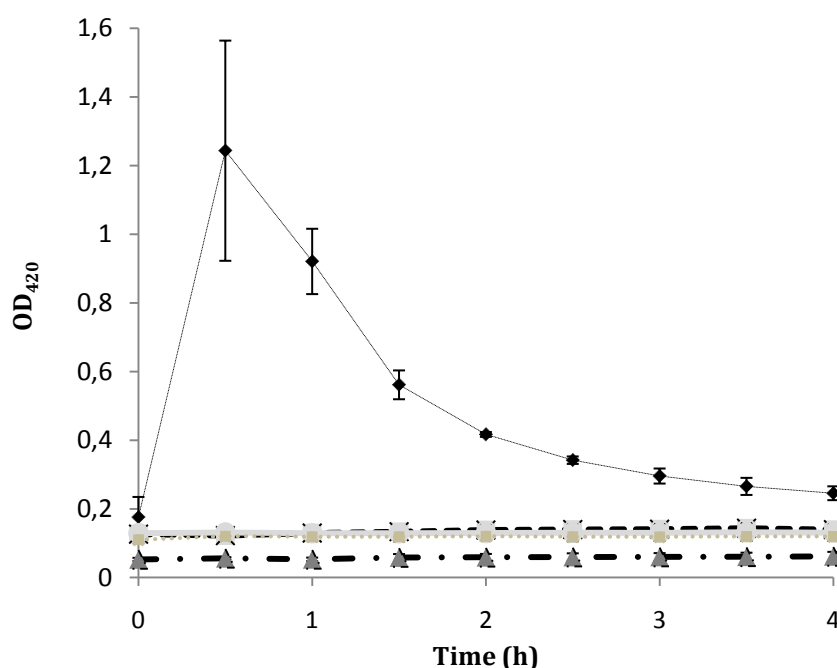
The solubility of the solid substrate  $C_{12:0}$  was evaluated in aqueous conditions along 8 hours stirring. Addition of 10 mM of solid substrate in 100 mM potassium phosphate buffer pH 7.4 (final volume of 300 ml in 1 L fermenter) and sampling over time to study the dissolving of substrate along a reaction course.



**Figure A4.** Solubility of 10 mM of  $C_{12:0}$  in 100 mM potassium phosphate buffer pH 7.4 over time in 1 L fermenter.

### A5. Control reactions (GOase assay)

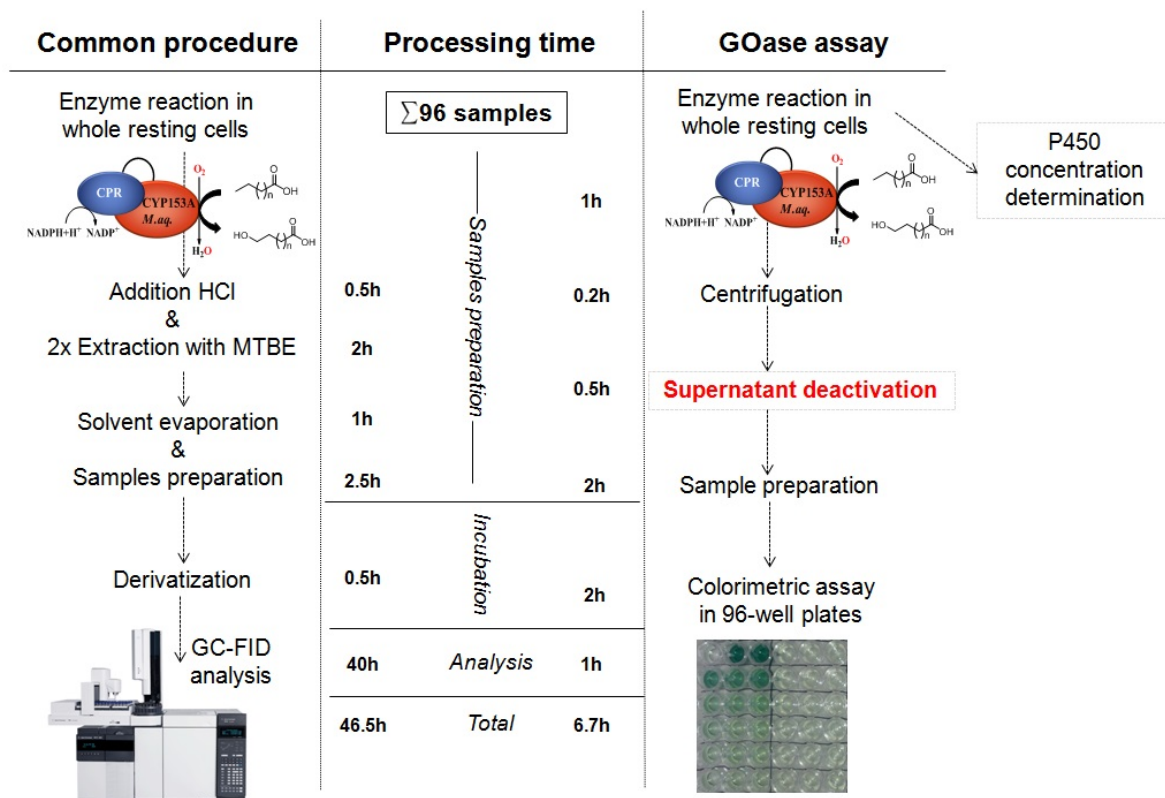
Several control reactions were performed after a biotransformation of 2 mM of  $C_{12:0}$  with  $50 \text{ g}_{\text{cww}} \text{ L}^{-1}$  *E. coli* resting cells or the same amount of solvent DMSO as negative control: tests of samples in the absence of HRP, of ABTS or GOase. The comparison was made with a positive control with all the reaction components added to resting cells supplemented with of 2 mM of  $\omega$ -OHC $_{12:0}$  and negative control with DMSO replacing the fatty acids, for 2 h. Only the presence of each reaction component and the presence of the  $\omega$ -OHFA product enable a displayed signal at OD $_{420}$ .



**Figure A5.** Different control reactions in the absence of HRP (triangles), of ABTS (circles), of GOase (crosses) and without substrate provided (boxes) were compared to a reaction analysing the supernatant after 2 h biotransformation of 2 mM  $\omega$ -OHC $_{12:0}$  and in the presence of each element to perform the assay (diamonds). These control experiments confirm the requirement of each component for a successful application of the GOase assay.

### A6. GOase assay vs. GC-FID

Comparison of the two procedures GC-FID and GOase assay regarding sample preparation and rapidity of analysis. Only 6.7 h are required to analyse the fatty acids from 96 well plates via the GOase assay, whereas over 46 h are needed by GC-FID method (including extraction, derivatization and method evaluation).



**Figure A6.** Comparison of the GOase assay and GC-FID regarding time for samples preparation and fatty acids analysis. Only 6.7 hours are required via the GOase method to characterize 96 variants whereas the common procedure including GC-FID analysis would necessitate about two days.





## Curriculum vitae

### Personal data

---

Name: Sandra Notonier  
 Nationality: French  
 Date of birth: 17.04.1988  
 Place of birth: Marseille, France  
 Email: sandra.notonier@gmail.com  
 Mobile: +49 176 618 900 83

### Education

---

07.2012 - 07.2015 **PhD Thesis supervised by Prof. Dr. Bernhard Hauer at the Institute of Technical Biochemistry, University of Stuttgart, Germany**  
 Dissertation: *"Development of highly efficient CYP153A-catalysed terminal hydroxylation of fatty acids"* P4FIFTY FP7 Funded Marie Curie

09.2008 - 09.2011 **Master Degree in Biotechnology at the School of Advanced Studies in Engineering, Biotechnology (Polytech-ESIL), Marseille, France**  
 Master Thesis: *"New strategies for the biosynthesis of bifunctional fatty acids derivatives"* at Renewable Energy Group (former LS9, Inc.), San Francisco, USA

09.2006 - 06.2008 **Two-year University Degree in Biology/Biochemistry at Aix-Marseille University, Marseille, France**

2006 **Baccalaureat, emphasis on Biology, with honours, Marseille, France**

### Work experiences

---

11.2013 - 12.2013 **Training in bioprocess engineering supervised by Prof. John Woodley at the Technical University of Denmark, Copenhagen, Denmark**  
 Bioprocess engineering for the application of P450s and upscaling strategies using bench-scale fermenters

02.2011 - 08.2011 **Internship in metabolic engineering at LS9, Inc. (acquired by Renewable Energy Group), San Francisco, USA**  
 Creation of P450 mutant, whole-cell biocatalysis development and research on metabolic pathways in *Escherichia coli* for the production of bifunctional fatty acids derivatives

03.2010 - 04.2010 **Internship in molecular biology, Biosensor Platform, CNRS, Marseille, France**  
 Development of a phagosensor as a tool for biologic detection

07.2009 - 08.2009 **Internship in the quality department at Catalyse company, Marseille, France**  
 Control of the documentation regarding chemicals and security procedures

