

'I am convinced that natural selection has been the most important, but not the exclusive, means of modification.'

(Charles Darwin; 'On the origin of species')

Published and patented items from this thesis:

As peer-reviewed publications:

Dennig, A.; Lültsdorf, N.; Liu, H.; Schwaneberg, U., Regioselective *o*-hydroxylation of monosubstituted benzenes by P450 BM3. *Angew. Chem., Int. Ed. Engl.* **2013**, 52, 8617-8620.

Ruff, A. J.; **Dennig, A.***; Schwaneberg, U., To get what we aim for - progress in diversity generation methods. *FEBS J.* **2013**, 280, (13), 2961-2978.

Dennig, A.; Marienhagen, J.; Ruff, A. J.; Guddat, L.; Schwaneberg, U., Directed Evolution of P450 BM3 into a *p*-Xylene Hydroxylase. *ChemCatChem* **2012**, 4, (6), 771-773.

Marienhagen, J.; **Dennig, A.;** Schwaneberg, U., Phosphorothioate-based DNA recombination: an enzyme-free method for the combinatorial assembly of multiple DNA fragments. *BioTechniques* **2012**, 0, (0), 1-6.

Dennig, A.; Shivange, A. V.; Marienhagen, J.; Schwaneberg, U., OmniChange: the sequence independent method for simultaneous site-saturation of five codons. *PLoS One* **2011**, 6, (10), e26222.

*equal contribution as shared first author

As poster presentations:

Dennig, A.; Shivange, A.; Marienhagen, J.; Schwaneberg U., OmniChange: the sequence independent method for simultaneous site-saturation of up to five codons (Biotrans; Sicily, Italy, 2011).

Marienhagen, J.; **Dennig, A.;** Bott, M.; Schwaneberg, U., Phosphorothioate-based DNA Recombination: an enzyme-free and sequence-independent method for the combinatorial assembly of multiple DNA fragments. (Biotrans; Sicily, Italy, 2011).

Ruff, A. J.; **Dennig, A.;** Schwaneberg U., Ultra-high throughput screening system for accelerated directed evolution of P450 monooxygenases. (11th International Symposium on Cytochrome P450 – Biodiversity & Biotechnology; Torino, Italy 2012).

Dennig, A.; Ruff, A. J.; Schwaneberg, U., Engineering P450 monooxygenases with OmniChange and flow cytometer-based screening. (Biocat; Hamburg, Germany, 2012).

As book chapter:

Dennig, A.; Marienhagen, J.; Ruff, A. J.; Schwaneberg, U., OmniChange: Simultaneous Site-Saturation of Five Codons. Directed Enzyme Evolution: Screening and Selection Methods (Methods in Molecular Biology) (in print).

As patents or patent application

Enzyme-free and sequence independent method for fusion of several DNA fragments. (The patent application was submitted in cooperation with BASF SE) [*Methods and Materials for Nucleic Acid Manipulation*] EP11 165 367.1] (US Patent filed in 2011).

Patent application on new P450 BM3 variants for aromatic hydroxylation of benzenes (The patent application was submitted in cooperation with DSM) [*Stereoselective hydroxylation of benzenes*] EP 12165386].

Abstract

The field of biocatalytic synthesis is of growing importance for the development of new and environmentally friendly synthetic routes to bulk and fine chemicals. P450 monooxygenases play an evident role in the functionalization of non-activated C-H atoms, a reaction that remains a major challenge in the field of organic chemistry. For instance, the synthesis of phenol with benzene as educt is inefficient and only profitable due to the high demand for the acetone byproduct. This work covers the engineering of the bacterial monooxygenase P450 BM3 from *Bacillus megaterium* towards the aromatic hydroxylation of benzenes. The main objective of this study was to develop an efficient P450 catalyst that is capable of producing *o*-phenols that are important building blocks for the synthesis of pharmaceuticals, plastics and vitamins. For this purpose, novel P450 BM3 variants with superior performance (e.g. >28-fold increased activity, more efficient use of the NADPH cofactor and a broader substrate spectra for benzene substrates) have been identified. Thirty putative substrates have been investigated for conversions with variant M2 (R47S, Y51W, I401M) and 12 benzene derivatives were additionally subjected to detailed characterizations with the WT enzyme and generated variants. In particular, the hydroxylation of anisole with variant M3 (R47S, Y51W, A330F, I401M) was very efficient and selective (>95 % *o*-position) reaching product concentrations higher than 1 g L⁻¹ and a TTN exceeding 10000. Besides the broadened substrate spectrum of P450 BM3 and the development of a protocol for preparative scale synthesis of phenols, the influence of commonly employed co-solvents such as DMSO was investigated. Unexpectedly, the best results for conversion of toluene to *o*-cresol were obtained without any co-solvent allowing a more sustainable route to *o*-phenols. A co-solvent free biocatalytic approach is cost-effective and allows a simplified purification of products. The catalytic characterization of P450 variants was complemented by docking studies to generate a hypothesis that explains the observed hydroxylation patterns. It was possible to show that the aromatic hydroxylation of benzene is governed through aromatic interactions in the active site, in particular between amino acid F87 and the aromatic substrates. Substituting phenylalanine by the 19 other canonical amino acids led to loss of efficient aromatic hydroxylation of *p*-xylene indicating that aromatic interactions are indispensable for this substrate class. Systematic investigations of a number of benzenes derivatives supported the hypothesis that regioselectivity and activity is governed strongly by π - π -interactions of substrate and protein. Introducing a phenylalanine at position 330, in close distance to the active site, further increased the binding and activity of P450 BM3 for substrates such as pseudocumene, mesitylene and nitrotoluene. Overall the measurements were in good agreement with the docking of substrates and proposed a T-shape orientation of the substrates in the binding pocket of P450 BM3. In the last part of this thesis the development of two novel methods for directed protein evolution, OmniChange and PTRec are described. OmniChange allows for the first time the efficient and reliable saturation of five independent codons in a DNA sequence, addressing current challenges and demands in focused mutagenesis. An in-depth statistical analysis of 48 clones showed that up to 86 % of all theoretical codons were present at five simultaneously targeted positions. The

robust and simple protocol allows a modular and combinatorial assembly of the targeted codons to study additive and cooperative effects in detail. Based on the successful OmniChange protocol a new method for DNA recombination, named 'PTRec', was established. PTRec enables recombination of proteins with less than 50 % sequence identity and requires only four conserved amino acids to design a crossover point. As proof of principle a model library of three phytase gene sequences was generated. The chimeras were sequenced and the distribution of specific domains revealed a close to ideal statistical distribution of DNA fragments. OmniChange and PTRec are based on phosphorothioate chemistry and allow for the first time a simple and robust fusion of multiple DNA fragments having low restrictions with regard to DNA sequence homology. In both cases the *E. coli* host was able to uptake and repair multiple DNA nicks to fuse up to five DNA fragments. If successfully accomplished, the protocol could be extended to develop further methods that would fuel progress in the engineering of proteins and in the implementation of biocatalyst into sustainable chemo-synthetic processes.

Table of Contents

1.	Introduction	1
1.1	Beyond Peak Oil – New Key Enabling Technologies for the 21 st Century	1
1.2	Milestones in Biotechnology – 6000 Years of History at a Glance.....	1
1.3	Sub-Divisions of Biotechnology and Economical Importance	2
1.4	Industrial Biotechnology: Success Stories	3
1.5	Industrial Biotechnology: Limitations and Challenges	5
1.6	Industrial Biotechnology: Perspectives.....	5
1.7	Key Performance Parameter for Enzymes in Industrial Biotechnology	7
1.8	Improving Operational Performance of Enzymes.....	7
1.9	<i>In Vitro</i> Evolution of Proteins	8
1.10	Strategies in Directed Protein Evolution.....	9
1.10.1	Engineering of Proteins by Directed Evolution and Random Mutagenesis.....	10
1.10.2	Structure Guided Rational Engineering by Focused Mutagenesis.....	12
1.11	Sampling Diversity in Mutant Libraries	13
1.12	Enzyme Classes and Catalyzed Reactions.....	14
1.13	Heme-Iron Oxidoreductases: Cytochrome P450 Monooxygenases	14
1.13.1	P450 Monooxygenase Electron Transfer Systems	15
1.13.2	The Catalytic Cycle of Cytochrome P450 Monooxygenases	16
1.13.3	The Uncoupling Reaction of Cytochrome P450 Monooxygenases	17
1.13.4	Cofactor Challenge in Biocatalytic Approaches with P450 Monooxygenases	17
1.14	The Third P450 monooxygenase from <i>Bacillus megaterium</i> (P450 BM3)	18
1.14.1	General Information on P450 BM3.....	18
1.14.2	Dimeric Structure and Electron Transfer	19
1.14.3	P450 BM3 – The Workhorse of Protein Engineers	20
1.14.4	Substrate Spectra and Reaction Conditions	21
1.15	Importance of Phenolic Building Blocks	22
1.15.1	Chemical Synthesis of Phenol.....	22
1.15.2	Biocatalytic Production of Phenol, Cresols and Dimethylphenols	23
1.15.3	Aromatic Hydroxylation of Benzenes by P450 BM3.....	23
1.16	Objectives	25
2.	Material and Methods	26
2.1	Chemicals	26
2.2	Enzymes and PCR Components.....	26
2.3	DNA Extraction and Purification with Commercial Kits	26
2.4	Machines and Equipment	26
2.5	Reaction Buffer.....	27

2.6	Bacterial Strains, Plasmids and Genes	28
2.7	Cultivation Media and Additives	31
2.7.1	Liquid Cultivation Media:	31
2.7.2	Solid Media:	31
2.7.3	Media Additives:	31
2.8	Microbiological Methods	32
2.8.1	Preparation of Chemical Competent <i>Escherichia coli</i> Cells	32
2.8.2	Transformation of Plasmid DNA into <i>Escherichia coli</i>	32
2.8.3	Preparation of <i>Escherichia coli</i> Cryo-Cultures.....	33
2.8.4	Shake Flask Expression of P450 BM3	33
2.8.5	Cell Lysis with High-Pressure Homogenizer and Sonicator	33
2.8.6	Mutant Library Preparation and Expression in Microtiter Plates	34
2.8.7	Preparation of Microtiter Plates for Screening.....	35
2.9	Molecular Biological Methods.....	36
2.9.1	DNA Extraction, Storage and Sequencing	36
2.9.2	Polymerase Chain Reaction (PCR)	36
2.9.3	Oligonucleotide Design for PCR Amplifications	36
2.9.4	Colony PCR	37
2.9.5	Phosphorothioate-Based Ligase-Independent Gene Cloning (PLICing).....	37
2.9.6	Agarose Gel Electrophoresis and DNA Quantification	38
2.9.7	Cloning of P450 BM3 Wild-Type into pALXtreme-1a Expression Vector.....	39
2.9.8	Site Directed and Site Saturation Mutagenesis.....	39
2.10	Biochemical Methods	40
2.10.1	Screening of Mutant Libraries in Microtiter Plates.....	40
2.10.2	The NADPH Depletion Assay.....	41
2.10.3	Phenol Detection by 4-Aminoantipyrine Assay (4-AAP Assay).....	41
2.10.4	Statistical Evaluation of Mutant Libraries Screening and Selection of Variants	42
2.10.5	Rescreening of Improved Variants	43
2.10.6	Purification and Lyophilization of P450 BM3	44
2.10.7	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis	45
2.10.8	Quantification of P450 Concentration by Carbon Monoxide Difference Spectra	45
2.10.9	Catalytic Characterization of P450 BM3 Wild-Type and Variants	46
2.10.9.1	Determination of NADPH Oxidation Rates and Coupling Efficiency	46
2.10.9.2	Determination of K_M and k_{cat}	47
2.10.9.3	Long Term Substrate Conversions with P450 BM3	48
2.10.9.4	HPLC Measurement of Phenols and Hydroquinones.....	48
2.11	Analytical and Chemical Methods	49
2.11.1	Two-Phase Solvent Extraction of Reaction Products.....	49
2.11.2	GC-Analysis.....	49

2.11.2.1	GC-FID Measurements.....	49
2.11.2.2	GC-MS Measurements.....	49
2.11.3	Visualization of Protein Structures and Docking of Substrates	50
3.	Results and Discussion	50
3.1	Directed Evolution of P450 BM3 into a <i>p</i> -Xylene Hydroxylase	50
3.1.1	Abstract.....	50
3.1.2	Introduction	50
3.1.3	Experimental.....	52
3.1.3.1	Focused Mutant Library Generation	52
3.1.3.2	Screening of Focused Mutant Libraries in Microtiter Plates	52
3.1.3.3	Kinetic Characterization of Improved P450 BM3 Variants	53
3.1.3.4	Product Detection via Gas-Chromatography.....	53
3.1.4	Results and Discussion	53
3.1.5	Summary and Conclusion	61
3.2	Substrate Screening with the Engineered Variant M2 and Investigation of the Influence of DMSO on the Activity of P450 BM3	62
3.2.1	Abstract.....	62
3.2.2	Introduction	62
3.2.3	Experimental.....	63
3.2.3.1	Selection and Conversion of Substrates	63
3.2.3.2	Investigating the Influence of DMSO and Isopropanol on Activity of P450 BM3	65
3.2.4	Results.....	65
3.2.4.1	Substrate Screening with Variant M2	65
3.2.4.2	Investigation of Co-solvent Influence on Activity of P450 BM3	68
3.2.5	Discussion	69
3.2.6	Summary and Conclusion	72
3.3	Regioselective <i>o</i> -Hydroxylation of Monosubstituted Benzenes by P450 BM3.....	73
3.3.1	Abstract.....	73
3.3.2	Introduction	74
3.3.3	Experimental.....	76
3.3.3.1	Determination of Catalytic Performance of P450 BM3 Wild-Type and Variant M2... ..	77
3.3.3.2	Gas Chromatography Analysis of Reaction Samples (GC-FID and GC-MS)	77
3.3.4	Results and Discussion	77
3.3.5	Summary and Conclusion	82
3.4	Hydroxylation of Trimethylbenzenes with P450 BM3 Wild-Type and Variants	82
3.4.1	Abstract.....	82
3.4.2	Introduction	83
3.4.3	Experimental.....	85
3.4.4	Results.....	85

3.4.4.1	Screening and Selection of Improved Variants for Conversion of Pseudocumene ..	85
3.4.4.2	Catalytic Characterization of P450 BM3 WT and Variants for Trimethylbenzene Hydroxylation	88
3.4.4.3	Docking of Mesitylene and Pseudocumene into the Active Site of P450 BM3	93
3.4.5	Discussion	96
3.4.6	Summary and Conclusion	99
3.5	One Pot Synthesis of Hydroquinones from Benzene with P450 BM3.....	100
3.5.1	Abstract.....	100
3.5.2	Introduction	100
3.5.3	Experimental.....	102
3.5.4	Results.....	102
3.5.4.1	Conversion of Toluene	103
3.5.4.2	Conversion of Anisole.....	104
3.5.4.3	Conversion of <i>p</i> -Xylene	105
3.5.4.4	Conversion of Chlorobenzene	106
3.5.4.5	Conversion of Bromobenzene	107
3.5.4.6	Estimation of Time Dependent Formation of Phenols and Hydroquinones	108
3.5.4.7	Conversion of Anisole, Toluene and <i>p</i> -Xylene with Variant M2	109
3.5.4.8	Conversion of Halogenated Benzenes with Variant M2	111
3.5.5	Discussion	112
3.5.6	Summary and Conclusion	116
3.6	Chemoselective α -Hydroxylation of Nitrotoluenes by P450 BM3.....	116
3.6.1	Abstract.....	116
3.6.2	Introduction	117
3.6.3	Experimental.....	118
3.6.4	Results.....	119
3.6.4.1	Hydroxylation of <i>o</i> -Nitrotoluene by P450 BM3 WT and Variants	119
3.6.4.2	Hydroxylation of <i>m</i> -Nitrotoluene by P450 BM3 WT and Variants	120
3.6.4.3	Hydroxylation of <i>p</i> -Nitrotoluene by P450 BM3 WT and Variants	121
3.6.4.4	Hydroxylation of <i>p</i> -Toluidine by P450 BM3 Variant M3.....	122
3.6.4.5	Docking of Nitrotoluenes into the Active Site of P450 BM3	122
3.6.5	Discussion	125
3.6.6	Summary and Conclusion	127
3.7	OmniChange: The Sequence Independent Method for Simultaneous Site-Saturation of Five Codons	128
3.7.1	Abstract.....	128
3.7.2	Introduction	128
3.7.3	Experimental.....	130
3.7.3.1	Employed Strains and Vectors	130

3.7.3.2	Mutagenic Oligonucleotide Design and Preparation of DNA Fragments	131
3.7.3.3	Mutant Library Generation	132
3.7.4	Results	133
3.7.4.1	PCR Amplification and Library Assembly	133
3.7.4.2	Quality Control of Generated Libraries	135
3.7.4.3	Statistical Evaluation of Mutant Libraries	136
3.7.4.4	Generation of Sub-Libraries with the OmniChange Protocol	138
3.7.5	Discussion	139
3.7.6	Summary and Conclusion	143
3.8	Phosphorothioate-based DNA Recombination: An Enzyme-free Method for the Combinatorial Assembly of Multiple DNA Fragments	144
3.8.1	Abstract	144
3.8.2	Introduction	144
3.8.3	Experimental	146
3.8.3.1	Sequence Alignment of Three Phytase Genes for Identification and Selection of Crossover Points	147
3.8.3.2	Identification of Mature Phytase Sequences by SignalP Algorithm	147
3.8.3.3	Oligonucleotide Design for Recombination of Mature Phytase Genes	148
3.8.3.4	Domain Amplifications by PCR, DNA Purification and Quantification	150
3.8.3.5	Library Assembly, Colony PCR and Sequencing of Chimeras	150
3.8.4	Results and Discussion	151
3.8.5	Summary and Conclusion	159
3.9	To Get What We Aim For – Progress in Diversity Generation Methods (Abstract from the original Review article)	160
4.	Final Summary and Conclusions	161
4.1	Engineering of P450 BM3 for Application in Phenol Synthesis	161
4.2	Development of New Methods for Directed Protein Evolution: OmniChange and PTRec..	165
5.	Appendix	168
5.1	List of Tables	168
5.2	List of Figures	169
5.3	List of Schemes	172
5.4	Abbreviations	172
5.5	Additional Experimental Information and Data	174
6.	References	188
7.	Further Contributions to Scientific Publications	208
8.	Acknowledgements	209
	Statement	211
	Lebenslauf	212

1. Introduction

1.1 Beyond Peak Oil – New Key Enabling Technologies for the 21st Century

One of the largest challenges in the 21st century is the steadily increasing demand for petrochemical products and transportation fuels. The so called “peak oil”, at which the maximum oil production is possible, was reached already.[1-5] Political instabilities and growing technical difficulties for facilitating natural resources such as gas[6] and oil[7] force to look for alternative routes to building blocks in organic chemistry or transportation fuels.[4, 8-11] Europe’s dependency on foreign oil and gas is high.[5, 8] Therefore, strategies have to be developed to ensure a constant and independent supply for energy and bulk chemicals. Another challenge is the establishment of an environmentally friendly or ‘green’ industrial production to reduce for instance the exhaust of CO₂ and the production of toxic waste, meanwhile saving energy expenses.[12-14] To face these challenges the European Union announced a new science and technology initiative (*‘Developing a common strategy for key enabling technologies in the EU’*)[15] to develop a more sustainable and economically attractive production in Europe. One of the key technologies to achieve these goals is biotechnology that should be implemented to a larger extent into the production of new bio-based materials, chemicals, pharmaceuticals and agricultural products.[8, 15] The final goal until 2020 is to build up a ‘bio-economy’ that generates leading edge technologies, new employment and the fundament for a sustainable industrial production.[8]

1.2 Milestones in Biotechnology – 6000 Years of History at a Glance

Although products from biotechnology processes are widely distributed in daily life (bread and cheese making, penicillin, insulin)[16-19], none of these products is affiliated by a label such as ‘Made by Biotechnology’. More common are negative affiliations for products generated in a biotechnological process at which genetically modified organisms in food products or cloning of humans became frequent topics in broadcasting giving only a small snapshot and resulting very often in misleading argumentations.[20] So, what is biotechnology and what does it stand for? The definition changed over decades, but a common agreement says:

The application of science and technology to living organisms, as well as parts, products and models thereof, to alter living or non-living materials for the production of knowledge, goods and services.[21]

The earliest examples of biotechnology processes are the fermentation of beer and wine (Sumerians), soya fermentation (Chinese) as well as production of cheese and bread.[19] All these processes were performed for more than 6000 years lacking the knowledge why, for instance, alcohol formation occurs in the prepared mash or coagulation of milk proteins takes place during cheese making.[22] With the first identification and characterization of microorganisms, especially of bacteria and yeast, the responsible organisms (e. g. *Saccharomyces cerevisiae*) could be cultivated in pure form and used

to enable today's processes of making beer, wine or bread with excellent quality as well as quantity.[23] Bringing light into the dark by the pioneer work of van Leuwenhoek (1683), Pasteur (1860) and Koch (1870), the reliable cultivation enabled employment of microorganisms for first biotechnological processes such as the alcoholic fermentation.[19] A large step in modern medicine was achieved by Alexander Fleming (1928), who discovered and described the antibiotic effect of *Penicillium chrysogenum*, which is a product from a biotechnological process, nowadays produced in ton scale.[24] The basis for today's 'modern' biotechnology was grounded in 1944 and 1953, when DNA was identified as genetic information carrier by Avery et al.[25] and its helical structure was deciphered by Watson and Crick.[26-28] Based on this knowledge several DNA modifying methods were developed that became standards tools in nearly every biology and biotechnology laboratory: DNA restriction and cloning (restriction enzymes in 1962 by Werner Arber) and amplification of DNA by the polymerase chain reaction (PCR; by Kary Mullis, 1983).[19] Having these tools in hand, it was possible to produce recombinant enzymes in bacterial strains, such as human insulin (1982).[29] The advances in molecular genetics enabled a more detailed view on regulation of different genetic and metabolic circles and allowed the modification by insertion or deletion of genes. The ability to amplify and decipher DNA sequences by Sanger sequencing[30, 31] initiated in the 1990s the human genome project which was finished in 2001 with 99.9 % completeness.[32] Sequencing of full genomes of bacteria, yeast, plants, animals and humans is today a simple and cheap analytical method that provides vast amount of useful data. Recent developments range from the re-design of proteins by protein engineering[33-37] to the re-creation of microbial factories from scratch that can be regarded as 'artificial' organisms.[38] The large and rapidly increasing amount of genetic data[39, 40] and growing knowledge on its function allows establishing new scientific fields such as the synthetic biology.[41, 42] This comparably young scientific field will allow the generation of new man made pathways for the production of important chemical products and biomaterials, meanwhile generating a deeper understanding how metabolic regulation in organisms is structured and functionally organized.[42] Today's biotechnology is a rapidly expanding and interdisciplinary science field merging the borders of different academic disciplines such as microbiology, genetics, biochemistry, chemistry, medical sciences and process engineering.[43, 44] The recent review from Buchholz and Collins summarizes small steps, major milestones and applications thereof in the field of industrial microbiology and biotechnology.[19]

1.3 Sub-Divisions of Biotechnology and Economical Importance

Biotechnology itself is a versatile scientific field which is not only focused on the production of value added compounds. Therefore, biotechnology was subdivided into six basic sections that are further divided into subsections. To simplify the categorizing a color coding was developed: green = plant biotechnology; red = medical biotechnology; white = industrial biotechnology; grey = bioremediation. Figure 1 gives an overview about economic importance of the five major sections in biotechnology in Germany in 2008.[45]

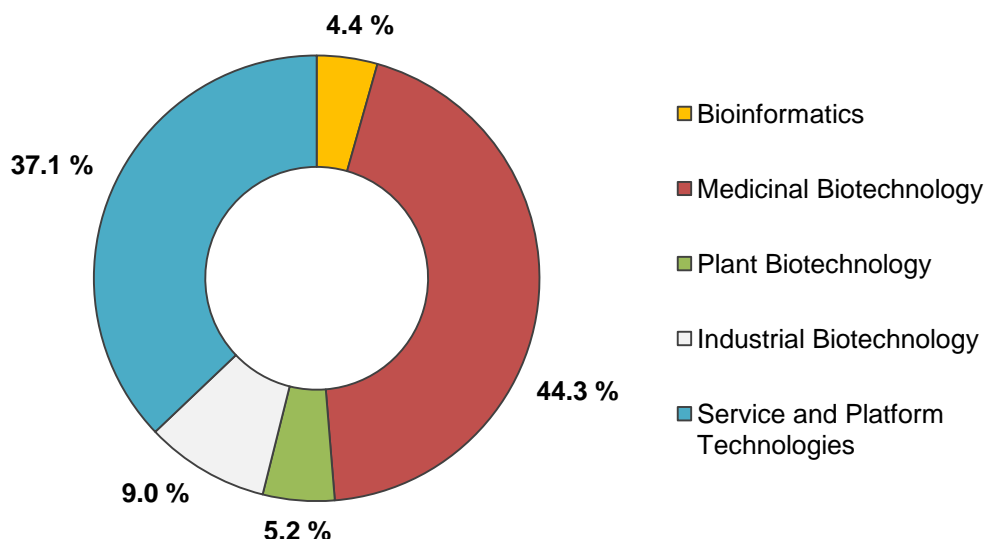
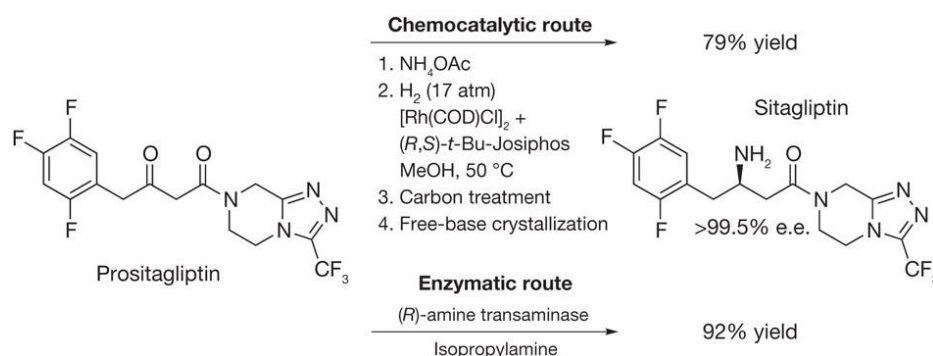


Figure 1. Business operation areas for biotechnology companies in Germany estimated for the year 2008. Most biotechnology companies are located at the moment in the areas of medicinal biotechnology (44.3 %) and service and platform technologies (37.1 %).[45]

Overall the economic importance of biotechnology, especially in Germany, is rapidly growing and reached in 2008 a total volume of more than 1.8 billion €.[45, 46] The largest economic volume is represented by the medicinal biotechnology with around 44.3 % turnover. With growth rates of nearly 10 % per year and already 30000 employees it is expected that the constant growth of the biotechnology related economy will contribute to a sustainable and growing economy in Germany and also in Europe.[45]

1.4 Industrial Biotechnology: Success Stories

White or industrial biotechnology is the third largest section according to total turnover in the biotechnology market.[45] Industrial biotechnology is commonly associated to the production of valuable chemical compounds using isolated enzymes or whole cells. A selection of biotechnologically produced chemicals and pharmaceuticals is summarized in a recent review from Bornscheuer et al.[47] A selected process is shown in Scheme 1 using an enzymatic route that employs an ω -transaminase to produce Sitagliptin that represents an important antidiabetic drug (treatment of diabetes mellitus type 2).[47] The transaminase reaction has several advantages compared to the original chemical process including a higher yield and an optically pure product. Furthermore, the enzymatic catalyst has higher productivity, produces less waste and omits transition-metal catalysts and high pressures.



Scheme 1. Biocatalytic transamination of Prositagliptin by a (*R*)-selective ω -transaminase to produce Sitagliptin. The original scheme and several more enzymatic driven industrial processes can be found in the report from Bornscheuer et al.[47]

Especially enzymes such as alcohol dehydrogenases (ADHs), lipases and transaminases are widely used due their potential of enantioselective catalysis for the production of optical pure chiral compounds (Figure 2) e.g. for synthesis of drugs such as (*S*)-Ibuprofen.[48-50] These enzymes are highly productive, exceeding a product concentration of $>100\text{ g L}^{-1}$ with good stability in organic solvents, high TTN and yields.[48, 51] Recent developments showed that alcohol dehydrogenases can even operate at very low water content or in neat substrates, which is convenient to most purification processes in chemical industry.[51, 52]

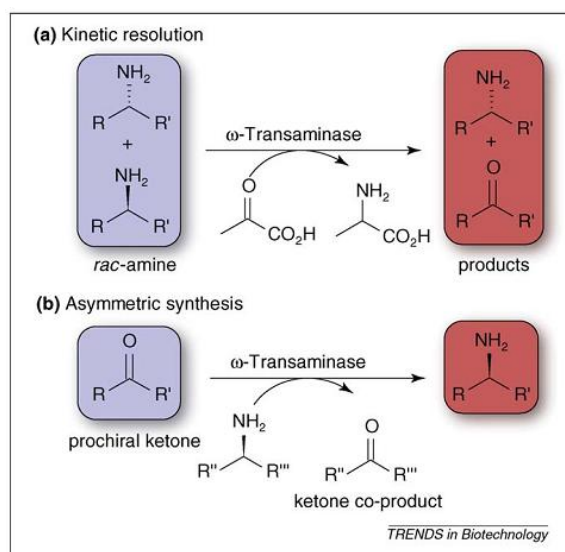


Figure 2. Enzymatic synthesis of chiral amines by (a) kinetic resolution and (b) asymmetric synthesis. (a) Starting from a racemic mixture of amines, an ω -transaminase with pyruvate as amine-acceptor depletes one of the enantiomers, meanwhile enriching the other enantiomer. (b) Asymmetric synthesis of a prochiral ketone, employing an ω -transaminase and an amine donor, leads to formation of a ketone co-product and a chiral amine product. The Figure was taken from Koszelewski et al.[49]

The application of whole cells in biocatalysis is highly attractive, especially when the organism can produce a later product by its own metabolism using glucose or other low priced and abundant resources.[53-55] A well-established process is the production of organic acids from glucose by *Corynebacterium glutamicum* with several million metric tons per year.[56, 57] Many more successful

examples for application of whole cells, cell free catalysts and enzyme cascade reactions were summarized in recent reviews.[11, 47, 53, 55, 56, 58-62]

1.5 Industrial Biotechnology: Limitations and Challenges

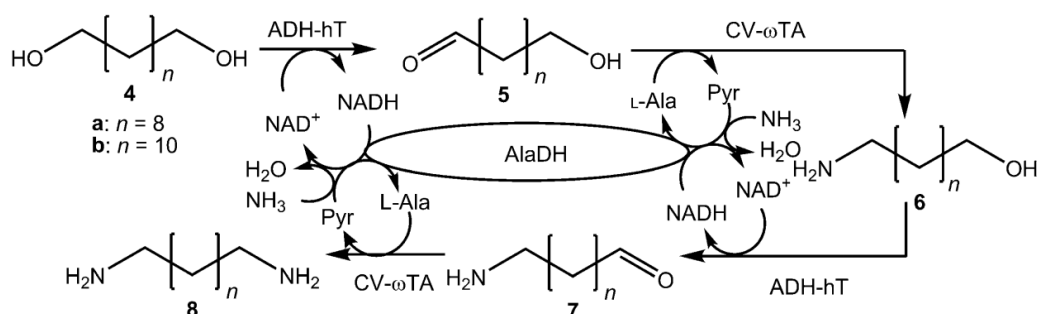
Although nature provides in theory an infinite amount of biocatalysts with outstanding reaction potential[33, 35, 37, 47, 63], the competition with chemosynthetic processes or refining from petrochemicals is high. Refining and purification of building blocks from oil or synthetic processes can be performed in large scale and at high purities sustaining petrochemicals, gas and coal as major source for building blocks.[64] Introducing a new synthetic route (biocatalytic or chemosynthetic) requires establishing of a reliable up-streaming and down-streaming process.[64] Products generated in a biocatalytic process require in many cases, other than in most chemosynthetic approaches, to be isolated from aqueous bulk environment. Purification from aqueous environments requires different strategies and equipment, while distillation of the bulk water is not as easy as compared to common organic solvents in synthesis. A further issue is the remediation of contaminated water that requires specific workup and treatment before disposal or re-use.[65] Limitations of biocatalyst are commonly summarized as a lack of operational stability meaning: low activity and productivity, poor stability in organic solvents or extreme pH, low conversion and yields in the reaction system.[65, 66] In chemical synthesis, these limitations can be overcome using a stoichiometric amount of reagents and the respective catalyst, which leads to higher costs and increased amount of used material. In biocatalytic processes whole cell approaches are commonly favored, especially when a valuable compound can be produced from a cheap carbon source using the cell-own metabolism.[53-56, 60, 67] However, some organisms are difficult to cultivate or require security level facilities. Further limitations could be the low solubility of a substrate, inhibitory effects (also by intermediates or the final product), and the destruction of the cell that could finally result in low product concentrations.

1.6 Industrial Biotechnology: Perspectives

Nonetheless, chemical synthesis has many well established processes, the dependency on petrochemical sources is very high and is expected to be intensified in the next decades.[5, 7, 13] In addition, some catalytic reactions are still very challenging in organic synthesis, for instance the stereoselective C-C coupling or the direct hydroxylation of non-activated carbon atoms with molecular oxygen and atmospheric pressure at RT.[63, 68-71] The potential to perform catalytic reactions under mild conditions with high chemo-, regio- and enantioselectivity is very attractive for organic synthesis.[71-73] Implementation of biocatalysts into synthetically processes can reduce purification efforts in multi-step synthesis, avoiding toxic or corrosive waste materials and therefore allowing a 'greener' route to chemical building blocks and products.[53, 56, 60, 63, 65, 72, 74, 75] The use of whole cell systems allows to use renewable and abundant (carbon-)sources for the production of chemical building blocks. Especially lignin, cellulose or organic acids from fermentation processes have a large potential since they are widely available and do not compete with food.[76] Metabolic

engineering of microorganisms and engineering of new catalytic pathways[10, 42, 54] would generate microbial cell factories to produce in theory any chemical structure from inexpensive carbon sources.[77, 78] However, the production of target compounds is always limited and regulated by the metabolisms of the cell. Therefore, the major challenge remains the control over fluxes for a target substrate, availability of cofactors and intermediates to obtain high product yields.[59]

New strategies in cell-free biocatalysis aim on multi-step reactions with two or more isolated enzymes in a one-pot reaction (Scheme 2).[62] However, the complexity of these cascades or tandem reactions is much lower compared to an entire cell metabolism and aims at reducing inhibitory effects, reversibility of the reaction as well as unwanted side product formation. Using isolated enzymes every step of the reaction can be monitored separately and if necessary be manipulated by the operator. [62, 79]



Scheme 2. Exemplary enzymatic cascade for the bioamidation of alcohols. An alcohol dehydrogenase (ADH-hT), an alanine dehydrogenase (AlaDH) and an ω -transaminase (CV- ω TA) are employed to drive the reaction towards the diamine products with 99 % conversion from 50 mM substrate. The scheme was taken from Sattler et al.[79]

Chemical synthesis very often does not allow multi-step synthesis as one pot reaction, for instance a reduction and an oxidation in the same reaction vessel, which is done with great ease in living cells.[62, 77, 79] To achieve a certain selectivity in chemical synthesis protection groups are employed that require additional purification and reaction steps.[80] Cell-free enzymatic cascade reactions could ease production of chemical compounds, omitting purification steps and therefore losses of educts or formation of unwanted side products.[62, 79]

By nature enzymes are designed to operate efficiently in cascades or in presence of hundreds of different catalysts, each performing a different reaction. The natural pathways and cascades are regulated by the metabolism using repressors, inducers or inhibitors that influence a target catalyst directly or on the genetic level.[81] A further regulatory mechanism is given by the affinity of an enzyme to a specific substrate which is described by the K_M value which plays an important role in cell free catalysis.[59] Since organism or enzymes are commonly not designed by nature to produce tailor made chemicals as bulk products, the application in an artificial cascade requires therefore a lot of fine tuning to operate all biocatalysts in an optimal way.[62, 79] The major challenge in cell free

enzyme cascades is still that all employed enzymes have to operate under identical conditions, for instance temperature, pH, ionic strength and changing educt, intermediate and product concentrations.[79, 82]

Immobilization of enzymes or whole cells can support stability of the respective catalysts and in some examples show that these constructs can serve for several years e.g. glucose isomerases for the production of fructose.[58, 83-85] On the other hand, not every catalyst retains its original activity after immobilization. However, immobilization is an important technique since the re-usability and long-term usage define economic potential of a biocatalyst.[85] Industrial biotechnology, however, will not replace chemical synthesis or vice versa. A combination of both would enable a powerful approach of high yields and conversions (chemistry) and excellent selectivities (biocatalysis), on the other hand jointly enabling sustainable production of bulk and fine chemicals.[86]

1.7 Key Performance Parameter for Enzymes in Industrial Biotechnology

Despite the attractiveness of selective catalysis using biocatalysts several parameters can help to figure out if an enzyme is suitable for a specific application (Table 1).[59] By nature proteins are designed to be stable and active in the defined and natural environment of the original host organism. A recent review from Schrewe et al. summarizes several parameters that are important to characterize biocatalysts in cell-free and whole cell industrial processes and their relation to parameters in chemical synthesis.[59]

Table 1. Parameters that are important for an application of biocatalysts in organic synthesis.[59]

Parameter	Denotation	Unit
Rate	Specific activity	$\text{U g}_{\text{enzyme}}^{-1}$
Rate	Turnover number	$\text{mol}_{\text{product}} \text{mol}_{\text{enzyme}}^{-1} \text{s}^{-1}$
Yield	Product yield on substrate	$\text{mol}_{\text{product}} \text{mol}_{\text{substrate}}^{-1}$
Yield	Total turnover number	$\text{mol}_{\text{product}} \text{mol}_{\text{enzyme}}^{-1}$
Productivity	Space time yield	$\text{g}_{\text{product}} \text{L}^{-1} \text{h}^{-1}$
Affinity	K_M	$\text{mol}_{\text{substrate}} \text{L}^{-1}$
Catalyst efficiency	k_{cat}/K_M	$\text{M}^{-1} \text{s}^{-1}$

In addition, parameters such as co-solvent tolerance, thermo and pH stability, stability at high substrate concentrations, cofactor dependency and selectivity of the catalytic reaction play important roles when designing biocatalytic processes.[65, 87]

1.8 Improving Operational Performance of Enzymes

The earliest methods to adopt an organism to the specific demands for an industrial process are the breeding of plants and feedstock to obtain higher yields in agricultural production.[88] The rather slow

process of breeding, was accelerated employing chemical mutagens as well as radiation to increase mutational rates in the target organism.[89] However, the application of mutagenic chemicals or radiation is rather inconvenient and commonly avoided due to low acceptance by the consumers.[89, 90] In addition, randomly and genome-wide introduced mutations are difficult to identify and explain. The lack for operational stability of many enzymes and whole cells in industrial processes[66] encouraged researchers to develop new strategies and methods to tailor biocatalysts to the non-natural conditions.[33, 91, 92] Various methods for immobilization of enzymes and whole cells were developed to protect the catalysts from harsh conditions or provide an environment which offers optimal or at least acceptable conditions for the specific catalytic activity.[58] Further developments were in direction to apply two-phase systems to separate e.g. a toxic substrate from the operating enzyme/cell, meanwhile, simplifying down streaming, cofactor regeneration and analysis of generated products.[93] Nonetheless, not every enzyme operates under an artificial or manmade environment.

As a consequence from evolution biocatalysts operate best in the environment where they were evolved for.[33] This includes also the fact that some enzymes were designed to operate under harsh conditions such as high temperatures (hot springs), at extreme pH levels (salt lakes, sulfur springs) or even in presence of high salt concentrations (halophilic bacteria).[94] As a conclusion from this: nature evolves catalyst with different operational stability and performance meanwhile catalyzing the same or similar reactions. Since most enzymes are also proteins, the 20 canonical amino acids are the building blocks for every cell to generate its specific toolbox of biocatalysts coded by DNA. Consequently, the exchange (mutation) of one or several nucleotide(s) within a DNA sequence can lead to an amino acid exchange that could have an impact on activity, stability or other protein properties. In most cases the gene and amino acid sequence of the target biocatalyst are known due to easy and fast access to DNA sequence information.[95, 96] This information allows modifying the DNA sequence by *in vitro* mutagenesis methods that are time-efficient, generate large diversity and do not require hazardous chemicals or radiation.[36, 37, 97-99] Furthermore the rather focused alteration of a defined DNA sequence commonly has no significant impact on the host metabolism and targets specifically the gene of interest (GOI).

1.9 *In Vitro* Evolution of Proteins

Only four building blocks (guanine, adenine, thymine and cytosine) code for the 20 chemically different canonical amino acids. This allows in a common size gene of around 900 bp a theoretical diversity of 20^{300} potential variants.[37, 98] However, this would generate a number that is beyond any imagination and exceeds the theoretically calculated amount of molecules in the universe.[98] Nonetheless, a very large number of variants can be generated by protein engineering, a process that is rather simple and generates highly diverse and large amount of variants in short time compared to design and synthesis of chemical catalysts.[32, 36, 37, 98] The generated variants can be sampled using high throughput screening systems to find the best or most suitable catalysts for a specific

approach upon selection pressure in the screening system.[37, 98] The correct folding of the peptide to a functional enzyme is therefore carried over by the host cell that is used for expression, ensuring reliable assembly of the catalyst. Mostly *Escherichia coli* strains are used for protein engineering campaigns due to ease in handling, fast growth and high transformation efficiency which is essential to generate mutant libraries with sufficient size of clones.[100]

The ability to modify DNA sequences by *in vitro* mutagenesis approaches allows modifying and tailoring of biocatalysts towards a specific demand.[36] A large number of methods and engineering strategies were developed in past and summarized in several reviews.[33-37, 74, 98, 101] Two basic strategies in protein engineering can be drawn: 1. Directed Evolution[33-37, 74, 98, 101]; employing iterative cycles of random mutagenesis or semi-rational engineering combined with a suitable screening system[36, 37, 102] and 2. Rational protein engineering; targeting pre-selected amino acid residues either by selective exchange or by randomization of a specific codon.[36, 98, 103-107]

1.10 Strategies in Directed Protein Evolution

At the beginning of each protein engineering campaign stands the question, what would be the right strategy to design the requested variant for a process?[36, 108] Is a random approach more likely to generate the desired variant or can the optimal mutant be generated by rational selection amino acids? Random mutagenesis is commonly performed if the gene sequence is the only available information and no structural model can be generated.[36, 37, 92, 98] Having a good homology model in hand or even a protein crystal structure model, a focused mutagenesis approach could be the strategy of choice. A strategy guide is given in Figure 4 that was taken from Chica et al.[108]

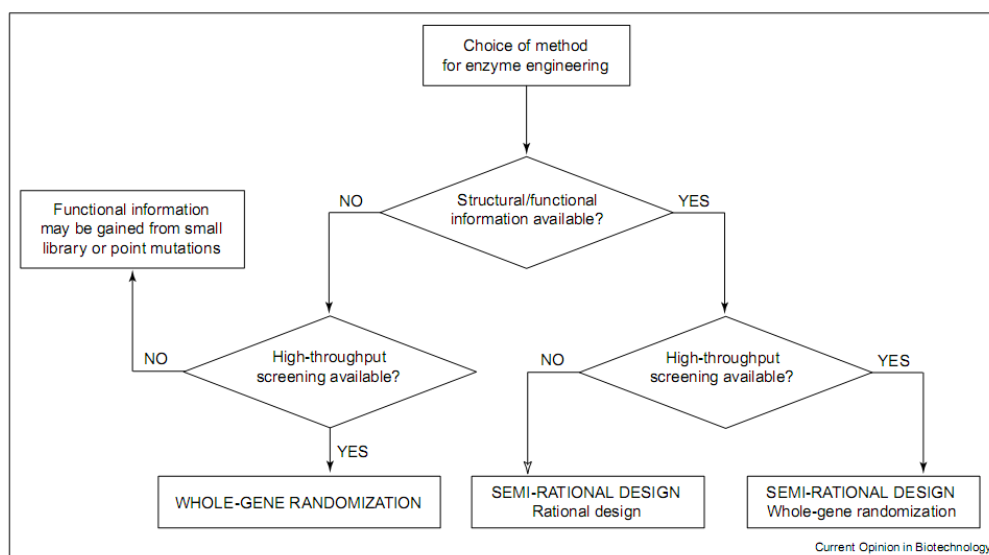


Figure 3. Strategy guide for engineering of enzymes based on structural information and throughput of the screening system. Depending on the availability of a good protein structural model and screening systems various mutagenesis strategies are available such as whole gene randomization, semi-rational design or rational design. The Figure was taken from Chica et al.[108]

However, not every property of an enzyme can be rationalized, for instance, stability in organic solvents or thermal inactivation. Therefore, selection of the right codons is difficult, though several computational programs are available for assistance.[108-112] Other properties such as activity or selectivity very often require a direct interaction of a substrate and target residues and therefore can be rationalized and marked down to amino acid residues within the substrate binding pocket.[103, 104, 107] A combined approach of rational, semi-rational and random mutagenesis can be helpful to explore a catalyst for the desired property. In this case, mutagenic hot spots can be identified by random mutagenesis and in a subsequent approach saturation mutagenesis can reveal the most beneficial substitution for the desired property.[37, 98] At the moment only guidelines that assist in decision are available and more statistical data is needed to formulate clear rules for engineering of proteins.[36]

1.10.1 Engineering of Proteins by Directed Evolution and Random Mutagenesis

In 20th century the concept of directed evolution was introduced which allows *in vitro* mutagenesis of a GOI using random mutagenesis, followed by selection of improved variants with desired properties by a suitable screening system.[91, 92] The cycle of mutagenesis and screening can be repeated in an infinite way to design a protein that has the desired catalytic property. Commonly, iterative cycles of mutagenesis and screening are performed where the best variant from previous directed evolution rounds serves as template for subsequent mutagenesis (Figure 4).[102]

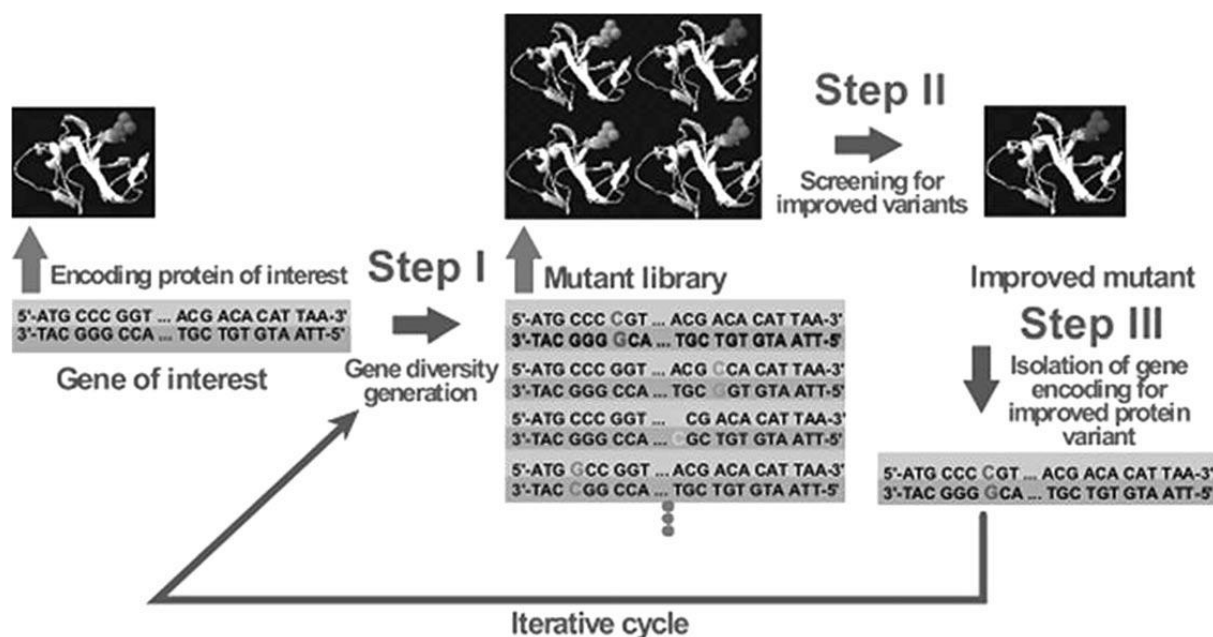


Figure 4. A directed evolution experiment includes three steps. Step 1: Generation of a mutant library. Step 2: Screening of mutant library with a target substrate or selection pressure. Step 3: Selection and isolation of variants with improved or desired properties. The whole process can be restarted and performed iteratively in infinite rounds until the protein of interest displays the desired performance. The Figure was taken from Güven et al.[102]

A large number of methods were developed and fine-tuned for *in vitro* engineering of proteins.[36, 37] Methods for random *in vitro* mutagenesis make use of the fact that DNA polymerases are not entirely error-free during DNA amplifications, especially when the enzyme lacks an exonuclease subunit for example the Taq DNA polymerase.[113] The error-rate of the DNA polymerase can be increased by varying cofactor concentrations (Mn^{2+})[114], employing unbalanced dNTP concentrations[114] or the application of non-natural nucleotides that pair during DNA amplification with two or more nucleotides equally with a specific bias as reported for SeSaM[115] and Mutazyme.[116] The probably most common method for random mutagenesis is error-prone PCR (epPCR)[117] which is a standard PCR protocol using either unbalanced dNTPs or varying concentrations of Mn^{2+} . Although epPCR is widely used, mainly single transition mutations (C vs. T or G vs. A) are introduced that lead to rather conservative amino acid exchanges (for instance alanine vs. valine).[115, 118-120] To overcome the bias of epPCR, the sequence saturation mutagenesis method (SeSaM) was developed that has a higher probability to introduce transversion mutations as well as subsequent mutations, generating a higher theoretical diversity that potentially increases the likeliness of finding improved variants.[115, 118]

Another powerful tool in directed evolution is the recombination of proteins which enables random exchange of larger parts of DNA sequences to generate protein chimeras, a method introduced as DNA or Family shuffling by Stemmer in 1994.[121, 122] Using DNA recombination methods it is possible to recombine parts of proteins from different environments such as extremophilic with mesophilic proteins, generating chimeras with high stability and activity.[36] Several reports underline

that these methods are powerful and can remove commonly present natural borders for horizontal exchange of genetic information between organisms.[123] A major advantage of recombination methods is for instance the low amount of inactive variants that are commonly generated and the possibility to oversample the generated diversity which is not possible when performing random mutagenesis.[124] However, methods for recombination of proteins are rarely applied compared to epPCR or focused mutagenesis, most likely due to their complexity of the protocols as well as high sequence dependency that is required to define a crossover point for recombination.[36]

1.10.2 Structure Guided Rational Engineering by Focused Mutagenesis

Methods in focused mutagenesis target one or more preselected amino acid residues within a known DNA sequence.[36, 37] A selected codon is exchanged to a specific amino acid e.g. by site directed mutagenesis (SDM)[125, 126] or saturated applying a defined codon degeneracy for instance by site saturation mutagenesis (SSM).[106] Commonly, the residues are selected based upon rational decision[36, 37, 103, 104], by browsing through the 3D structure of a homology model or crystal structure to identify promising residues.[112] Prediction of the “best” substitution remains difficult, therefore SDM only investigates a snapshot (5 % = 1 SDM) of the theoretical and potential diversity.[105, 127] Site saturation mutagenesis (SSM) allows identification of the best amino acid on a specific position by applying NNN (64 codons) or NNK degeneracy (32 codons), covering all 20 canonical amino acids.[105, 127] The saturation of a codon requires therefore a reliable medium throughput screening system for improved variants which is commonly performed in MTP format. Mostly whole plasmid amplification is used, also known as the ‘QuikChange’ method, which is based on a simple and robust protocol that generates a large amount of mutant colonies for efficient oversampling of the theoretical diversity.[127] Strategies for focused mutagenesis were further extended to approaches where multiple neighboring codons were saturated (CASTing)[103] or several triplets were iteratively saturated (ISM; Figure 5).[104, 107]

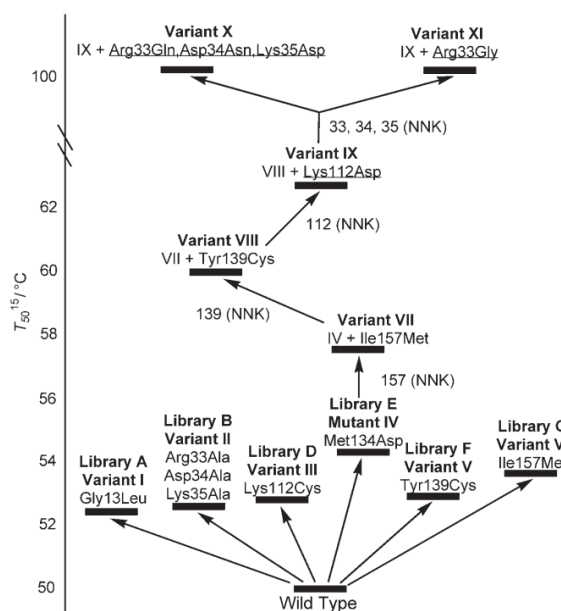


Figure 5. Evolutionary tree for improving thermostability of a lipase (Lip A) by iterative saturation mutagenesis (ISM). Selected positions are saturated with defined degeneracy, for example NNK.[105, 127] The best variants from each round are target for subsequent single site-saturation mutagenesis, allowing generation of variants with desired functionality (here increased thermostability). The Figure was taken from Reetz et al.[128]

A combination of both methods was reported as iterative CASTing, which served as an efficient strategy for the re-engineering of P450 BM3 towards hydroxylation of steroids.[107] Compared to random mutagenesis the overall generated diversity is limited and can be adjusted during every experiment to the availability and throughput of a screening system.[36] Nonetheless, the largest challenge remains unsolved: the selection of the correct residues for a focused mutagenesis experiment.

1.11 Sampling Diversity in Mutant Libraries

Basic requirements for a successful directed evolution experiment are: 1. High quality mutant libraries and 2. A screening platform with desired throughput and good precision to identify improved variants.[36, 37, 129] Screening of diversity in mutant libraries is performed using high- or ultra-high throughput screening platforms that enable reliable throughput of $>10^4$ clones per day.[36, 37, 129] Commonly, MTP based screening platforms are used providing high precision and easy handling.[98] Due to the high amount of inactive clones in random mutant libraries (epPCR and SeSaM)[36, 129] screening platforms that enable a rough pre-selection of active clones are favored, such as agar plate assays[130, 131], flow cytometer[132, 133] or micro fluidic devices[134-137] that allow selection of active clones with a throughput of theoretically 10^9 to 10^{10} clones per day. A further limitation in directed evolution is the generation of a high number of clones after transformation for instance of an epPCR library that contains $>10^{10}$ different variants.[36] Common transformation efficiencies hardly exceed 10^9 cfu μg^{-1} plasmid DNA and do not allow screening of the entire diversity generated during mutagenesis. Due to limitations in screening and transformation efficiency the full diversity generated within a random library is not 'screenable' with current methods.[36]

1.12 Enzyme Classes and Catalyzed Reactions

With more than 2000 fully sequenced genomes[96] and several thousand enzymes being isolated and characterized a distinct classification was necessary to handle the large and growing amount of available data. Enzymes are grouped into six classes (EC number = enzyme commission number) according to the specifically catalyzed reaction:

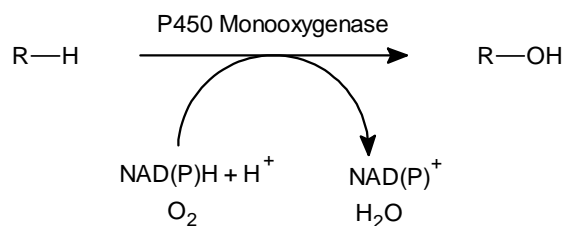
1. **Oxidoreductases:** All enzymes that catalyze oxidoreductions
2. **Transferases:** Enzymes that transfer a chemical group (e.g. methyl or glycosyl groups)
3. **Hydrolases:** Catalyzing the hydrolysis of various bond types
4. **Lyases:** Cleavage of C-C, C-O, C-N and other bonds by means other than by hydrolysis
5. **Isomerases:** Catalyze changes within a single molecule
6. **Ligases:** Catalyze the joining of two molecules with concomitant hydrolysis of the di-phosphate bond in ATP or a similar tri-phosphate

(definitions taken from BRENDA enzyme database)[138]

A large compilation of reported and characterized enzymes can be found in the 'BRENDA' database (<http://www.brenda-enzymes.org>).[138]

1.13 Heme-Iron Oxidoreductases: Cytochrome P450 Monooxygenases

Cytochrome P450 monooxygenases or shortly named CYPs or P450s are heme-iron containing oxidoreductases (EC 1.14.14.1) that catalyze the introduction of molecular oxygen into non-activated carbon atoms.[63, 139] With more than 5000 cloned members CYPs represent a large enzyme superfamily that can be found in all kingdoms of life.[139] The rapidly growing number of new P450s requires a clear classification which was done by Don Nelson in 1998, who reserved as follows: 1-49 for animals; 51-69 for lower eukaryotes; 71-99 for plants; from 101 ongoing for bacteria.[140] The name of this enzyme class is derived from the observation that the reduced form of porphyrin iron makes a shift in absorbance from 420 to 450 nm maximum upon binding of CO, which was first described by Omura and Sato in 1964.[141] P450s have a typical and conserved fold (20 % sequence identity) which suggested an ancestral form of this class of enzymes.[142] P450 monooxygenases catalyze various reactions such as reduction, desaturation, ester cleavage, ring expansion, ring formation, aldehyde scission, dehydration, *ipso* attack, one-electron oxidation, coupling reactions[143], molecular rearrangements[144] and cyclopropanation[145], to name a few examples. The hydroxylation reaction of a non-activated carbon atom is displayed in Scheme 3.



Scheme 3. Basic scheme for a P450 monooxygenase-catalyzed reaction.[146, 147] During the reaction one mol of substrate, one mol of oxygen and one mol of NAD(P)H are converted to one mol of hydroxylated product (R-OH), one mol of H₂O and one mol of NAD(P)⁺.

The natural function of P450s is highly diverse, ranging from detoxification of pollutants to the synthesis of steroids and flavor compounds.[63, 70, 147] Especially the ability to integrate molecular oxygen selectively into non-activated carbon atoms under mild reaction conditions (RT, atmospheric pressure) makes P450 monooxygenases highly attractive for applications in chemical syntheses.[70, 146]

1.13.1 P450 Monooxygenase Electron Transfer Systems

The supply with electrons in P450 monooxygenase systems is very sophisticated and highly diverse. Despite that P450s are present in all kingdoms of life and the structural fold of the monooxygenase domain is highly conserved, CYPs have an exceptional large diversity with regards to their electron transfer systems. Ten classes of electron transfer systems or electron transfer chains are known for P450 monooxygenases, which were summarized in a review by Hannemann et al.[148] Upon binding of a potential substrate (Reaction step II in Figure 7) the oxidation potential of the heme iron increases, allowing the oxidation of the NAD(P)H cofactor by a reductase followed by the transfer of the electrons to a mediator, a ferredoxin or flavin molecule (Figure 6).

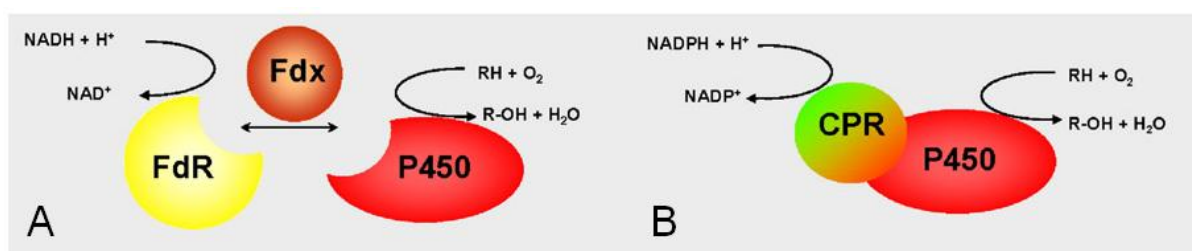


Figure 6. Two bacterial electron transfer systems of P450 monooxygenases. Parts of the Figure were taken from Hannemann et al.[148] **A:** Class I; Most prominent bacterial and mitochondrial eukaryotic electron transfer system. Electrons are transferred from NADH via a reductase (FdR; highlighted in yellow) to a ferredoxin mediator (Fdx). In the final step electrons are delivered to the P450 monooxygenase domain where the hydroxylation reaction is catalyzed. **B:** Class VIII. Self-sufficient fusion of a reductase (CPR) and a bacterial P450 monooxygenase, as known for P450 BM3.[149-151] Reductase and P450 monooxygenase are fused on a single peptide chain allowing a direct electron transfer without additional redox mediator.

Finally, the electrons are delivered to the monooxygenase domain where the oxygen cleavage is catalyzed. Different transfer systems are known which strongly depend on location of the respective P450 that can be membrane bound or soluble.[148] However, activity and coupling of the reaction are

not always optimal for a later synthetic application. Monooxygenases from class VIII[148] do not encounter this challenge, since these enzymes are self-sufficient and carry their natural reductase and electron transfer components fused with the respective monooxygenase domain as known for P450 BM3, the probably best investigated and widely applied self-sufficient P450 monooxygenase. Because only a small part of all P450s are self-sufficient (mainly CYP102A family)[148, 152] a clear trend is visible to design and screen for self-sufficient or artificial redox-partner systems for their application in a later process. [153]

1.13.2 The Catalytic Cycle of Cytochrome P450 Monooxygenases

For decades the catalytic cycle was widely studied to decipher the mechanism of oxygen activation and transfer to non-activated carbon atoms by P450 monooxygenases.[70, 147, 154] More than 50 years after Omura and Sato first introduced the term cytochrome P450 ('pigment 450') [141, 155], the catalytic cycle and its details are accepted as drawn in Figure 7 that was taken from Whitehouse et al.[70]

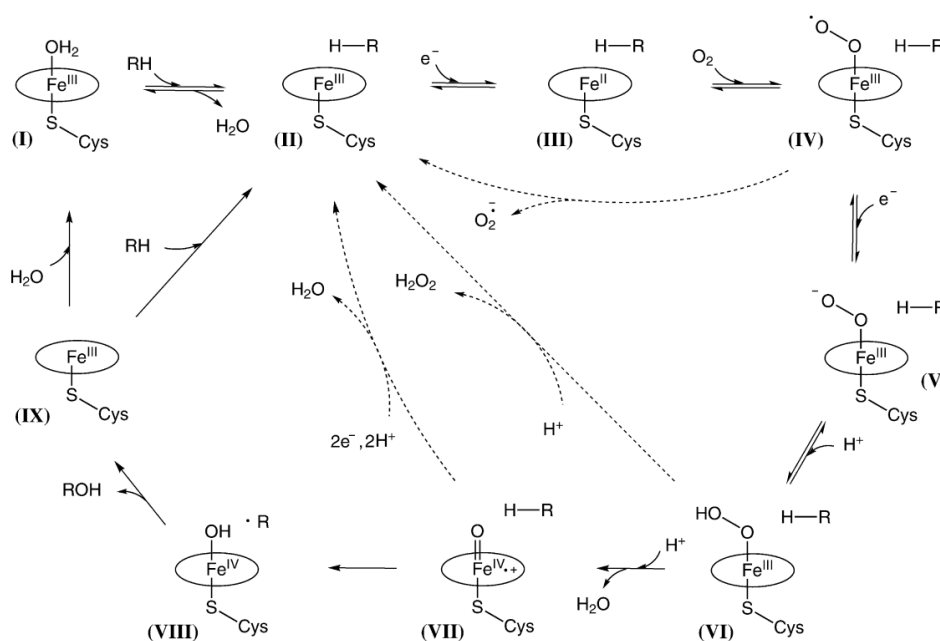


Fig. 1 The P450 catalytic cycle.

Figure 7. The catalytic cycle of cytochrome P450 monooxygenases. Also indicated with dotted arrows are the three uncoupling reactions that can occur during the monooxygenase reaction. The Figure was taken from Whitehouse et al.[70]

The catalytic cycle involves ten steps that have to be performed as concerted action to ensure efficient hydroxylation of the target substrate.[151] In resting state (I) a water ligand is bound to the heme iron, which is upon binding of a substrate removed (II). One electron from a NAD(P)H cofactor is transferred via the adjacent reductase and/or a redox mediator molecule (III) to the heme iron that, upon removal of the water ligand, becomes more oxidizing. The reduced heme iron (ferrous state) allows binding of a molecule of oxygen (IV; oxy-complex), which is reduced by the second electron

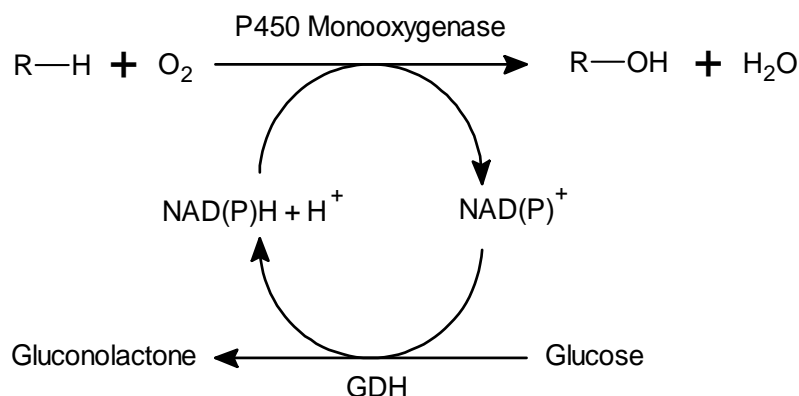
from the cofactor (NAD(P)H) to the peroxy-complex (V). After protonation of the peroxy-complex (VI) compound 0 is formed, which leads after second protonation to release of a molecule of water and formation of compound I (VII), the active species for abstraction of the H-atom from the substrate molecule (VIII). The abstracted proton is later rebounded as OH radical to the substrate and the catalytic cycle restarts with the hydroxylated substrate leaving the active site. Finally, the heme iron returns to the pentaferrous state (IX) that allows binding of a new water molecule (I). Besides hydroxylation via rebounding also epoxidation of substrates can occur that does not include H-abstraction from the substrate molecule.[154] P450-catalyzed epoxidation involves a direct attack by compound I (VII) at the unsaturated carbon atoms to form the respective epoxide.[154, 156]

1.13.3 The Uncoupling Reaction of Cytochrome P450 Monooxygenases

The catalytic cycle of P450 monooxygenases requires a concerted action of substrate and O₂ binding as well as electron transfer from the reductase to the heme iron to activate the oxygen for hydroxylation of a target C-atom (Figure 7). In many experimental investigations it was observed that the amount of oxidized NAD(P)H cofactor does not generate a stoichiometric amount of product. [70, 157] The ratio between transferred electrons and generated product concentration is known as ‘coupling’ and describes the efficiency of a P450-catalyzed reaction. However, the coupling depends strongly on the target substrate and O₂ transfer or the second electron transfer step.[157, 158] Three types of uncoupling are described for P450 monooxygenases: 1. *superoxide uncoupling* (O₂^{•-}-formation): Due to a slow transfer of the second electron. 2. *peroxide uncoupling* (H₂O₂-formation): Due to a loosely fitting substrate and 3. *oxidase uncoupling* (H₂O-formation): the H-bond cannot be abstracted from the substrate due to a large distance to compound I (VII). A perfect coupling would produce a stoichiometric amount of hydroxylated product, which commonly is the case when employing the ‘natural substrate’ of a P450, for example the conversion of camphor with P450_{cam}. [159] High coupling is not only desired for achieving high productivities, but also to avoid reactive oxygen species that could lead to inactivation of the monooxygenase catalyst.[139]

1.13.4 Cofactor Challenge in Biocatalytic Approaches with P450 Monooxygenases

Oxidoreductases require an equimolar amount of a co-factor such as NADH or NADPH to catalyze a reaction. Nicotine-amide cofactors are commonly regarded as expensive[160-162] and sometimes as the limiting factor for catalytic application of enzymes in industry. Furthermore, NADH and NADPH have comparably low stability with regards to temperature, pH as well as ionic strength.[163] Due to the uncoupling of P450s with most non-natural substrates, equimolar application of cofactor and substrate is not sufficient to ensure full conversion.[63, 70] Therefore, additional enzymes can be supplemented in a reaction system for regeneration of the respective cofactor and permanent supply of the monooxygenase domain with electrons (Scheme 4).



Scheme 4. NADPH cofactor regeneration with a glucose dehydrogenase (GDH) and glucose. The system can be employed together with a P450 monooxygenase, for instance P450 BM3, for efficient cofactor regeneration.[164, 165]

Alternatively, NADH and NADPH can be regenerated using whole cells and their respective cell own cofactor regeneration systems. This strategy is commonly cheaper since no additional enzyme catalyst has to be produced and purified.[59, 77] Employing whole cells, NADH would be the favored cofactor since it is more abundant in cells and theoretical yields from one mole of glucose are 10 fold higher than for NADPH under aerobic conditions. To overcome challenges with low yields of NADPH in whole cells, one strategy allows to invert selectivity towards NADH by introducing mutations in the cofactor binding domain of reductases as shown for P450 BM3 (R966D, W1046S).[164, 166] Although regeneration systems based on glucose are highly efficient, recent developments go either in direction to use inorganic sources such as phosphite for NADPH regeneration.[167] Alternatively, it is also possible to replace NAD(P)H by biomimetic cofactors such as N-Benzyl-1,4-dihyronicotinamide that can be regenerated with a rhodium catalyst and sodium formate (HCO_2^-) as electron source.[63, 168] Very attractive is also the application of abundant and cheap electron sources such as zinc dust or an electric current that requires significant more development to reach productivities comparable to natural reductase-mediator systems.[160, 161]

1.14 The Third P450 monooxygenase from *Bacillus megaterium* (P450 BM3)

1.14.1 General Information on P450 BM3

The cytochrome P450 monooxygenase BM3 (EC 1.14.14.1) was described for the first time by Fulco et al. as third monooxygenase from *Bacillus megaterium* (P450 BM3)[149, 150] or CYP102A1 which is the systematical name. P450 BM3 is a soluble, self-sufficient monooxygenase with a length of 1049 amino acids. It is known to hydroxylate a broad variety of substrates, whereas highest coupling efficiencies of the WT enzyme are obtained for oxidation of fatty acids (medium to long chain).[70] Though not fully coupled, the natural function in *Bacillus megaterium*, a soil bacterium, is attributed to detoxification of xenobiotic lipids (unsaturated fatty acids) produced by plants or in regulation of membrane fluidity of the host organism.[169, 170] Due to its self-sufficient character and extraordinary high catalytic activity towards hydroxylation of fatty-acids, which are nearly exclusively

hydroxylated in sub-terminal positions, P450 BM3 became one of the best investigated enzymes with around 20 to 30 publications each year.[70] With its exceptional high activity and broad substrate spectra P450 BM3 grows interest as industrial catalysts for synthetic applications. Together with P450_{cam}[171] it is probably the best investigated bacterial P450 monooxygenase.[151]

1.14.2 Dimeric Structure and Electron Transfer

In its natural form P450 BM3 is a dimeric protein formed by two identical peptide chains (Figure 8). The reductase and monooxygenase domains are connected via a short linker (six amino acids from 468 to 473)[172, 173] that also enables efficient electron transfer between both monomers.[70]

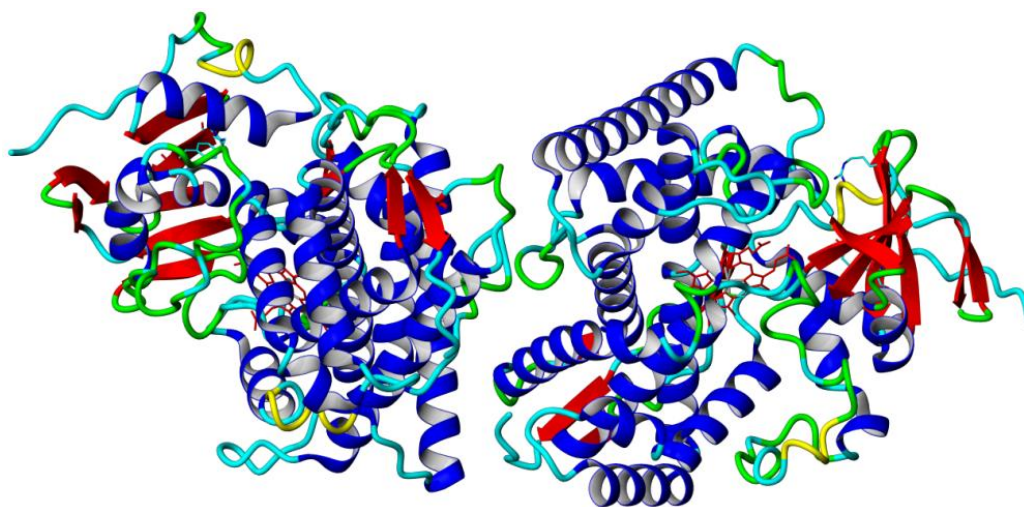


Figure 8. Visualization of the dimeric structure of the two P450 BM3 monooxygenase subunits. The peptide sequence is shown as ribbon model with α -helices (dark blue), β -sheets (red) and random coil structures. The heme iron buried in the peptide structure is highlighted in red and stick representation.

Both monomers interact during catalysis leading to the exceptional high activity for a P450 monooxygenase. Girvan et al. could show experimentally that electrons from one subunit are transferred to the monooxygenase domain of the second monomer.[166] Further indications that the dimeric protein form is essential to maintain high activity is the observation that dilution of P450 BM3 below a concentration of 10 nM leads to a non-linear decrease in activity, which was reported by Neeli et al.[174] A number of 49 crystal structures have been resolved, with and without a substrate bound (see: <http://www.rcsb.org/pdb>). Most structures are available for the monooxygenase domain and very recently a structure model of the crystallized reductase domain was published[175], allowing insight into cofactor binding and electron transfer. However, crystallization of the full protein is not accomplished till today. The large number of crystal structures and homology models thereof allowed a deeper insight into the catalytic mechanism and substrate-protein interaction in P450 BM3 and other monooxygenases.

1.14.3 P450 BM3 – The Workhorse of Protein Engineers

The large number of available crystal structures, the soluble expression in *E. coli* and its self-sufficient character make P450 BM3 the ideal monooxygenase catalyst for investigation of structure function relationships. An outstanding number of variants have been generated, targeting a vast number of important catalytic residues.[70] In addition, a large library of substrates was investigated allowing a detailed insight into the catalytic mechanism, substrate-protein interactions as well as structure function relationships. Various evolving strategies were pursued to fine tune this enzyme accepting substrates ranging from gaseous methane[176] to large steroid molecules.[177] The detailed review of Whitehouse et al. summarizes efforts being made on generating variants with subsequent screening for an inconceivable number of substrates.[177] The comparably huge substrate binding pocket as well as large molecular movements of the peptide structure during catalysis provides a highly promiscuous monooxygenase in contrary, for instance, to P450_{cin} that (so far) hydroxylates only one known substrate.[178] Several key-residues are known that influence the activity and selectivity of P450 BM3 such as F87, R47, Y51, T268 and A330.[151] The location of these residues in the monooxygenase domain of P450 BM3 are highlighted in Figure 9.

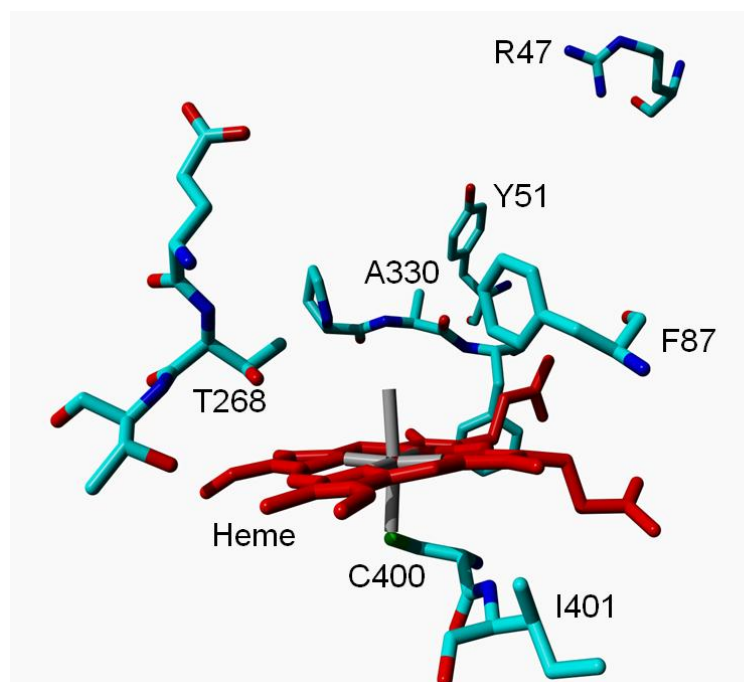


Figure 9. Selected key residues in the substrate binding domain of P450 BM3. C400 and I401 are involved in heme binding and coordination.[151] T268 is responsible for substrate recognition as well as oxygen activation and orientation.[179] F87 and A330 directly interact with bound substrates and influence activity and regioselectivity.[69, 180] R47 and Y51 are the so called ‘gaters’[151] that control substrate inlet to the active site. R47 is also involved into fatty acid substrate binding through formation of a salt bridge between carboxyl-group of the substrate and the guanidine-group of R47.[181]

Besides engineering P450 BM3 towards improved activity, new substrates or selectivity, many efforts go in direction to optimize catalytic performance with regards to increase tolerance against co-solvents[182] or employing an electric current or inorganic electron sources such as zinc dust.

[160, 161] Mutational strategies involve very often site-directed mutagenesis of known key residues allowing fast catalytic improvements and rational explanations.[151] However, not always the best amino acid for a specific target can be identified by this strategy.

Recent developments also employ decoy molecules in biocatalytic approaches with P450 BM3. [176, 183] Decoy molecules are commonly perfluorinated fatty acids that cannot be hydroxylated by P450 BM3 but assist in narrowing the active site and therefore providing additional sterical interaction for substrate orientation.[176, 183] The key-residue R47 serves for anchoring the carboxyl group within the active site.[183, 184] Employing these decoy molecules it was possible to hydroxylate small and industrially important molecules such as methane[176] and benzene[183] with the P450 BM3 WT enzyme, leading to increased coupling efficiency and activity.

1.14.4 Substrate Spectra and Reaction Conditions

The P450 BM3 WT accepts various substrates, but preferentially medium chain fatty acids are hydroxylated with highest efficiency and activity.[70] The high preference for these substrate class was further elucidated by crystallization and mutagenesis studies where the guanidium group of R47 was proposed to interact with the carboxyl-group of the fatty acid substrates forming a salt bridge.[185, 186] P450 BM3 has a broad substrate scope ranging from propane, cycloalkanes to large aromatics such as pyrene or sterically demanding testosterone (Figure 10).[151] In many cases the activity could be increased towards the specific substrates by introducing specific mutations in the active site or by residues close to the substrate channel. A detailed summary of investigated substrates and related mutations is given in the review by Whitehouse et al., compiling information on all substrates and their conversion by P450 BM3 reported until 2012.[70, 151]

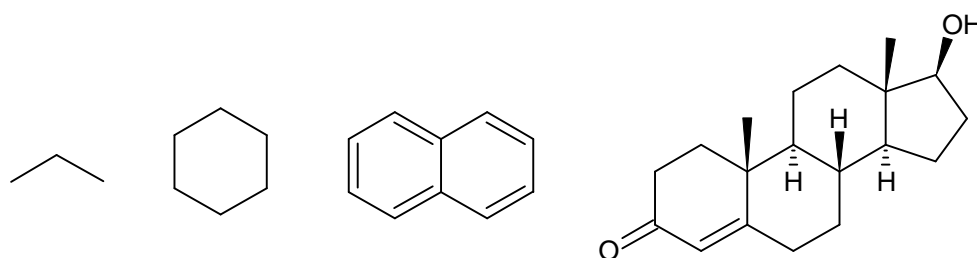


Figure 10. P450 BM3 accepts various substrates for hydroxylation such as propane, cyclohexane, naphthalene and testosterone.[70, 176, 177, 187-189]

Although enzymes are known to catalyze reactions with high regioselectivity and in several cases also with excellent enantioselectivity, P450 BM3 not always provides good selectivity for a biocatalyst. This can be attributed to the comparably large active site allowing substrates to bind in varying orientations. Compared to P450cam, which has enantio- and regioselectivity of >99 % towards camphor, P450 BM3 is regarded as a rather unselective enzyme. Nonetheless, engineering of the active site allowed tailoring of P450 BM3 towards pre-dominant products or inverting enantioselectivity as shown successfully for styrene epoxidation to produce styrene epoxide.[190, 191]

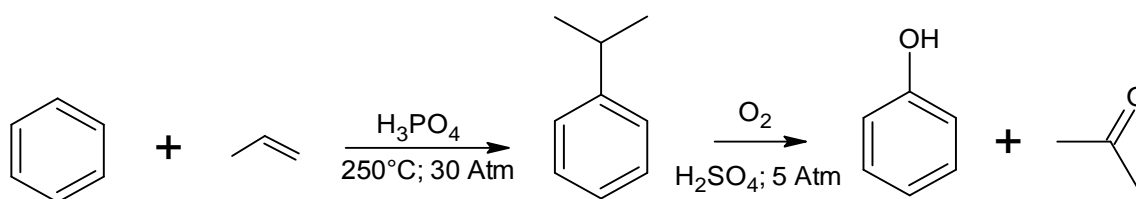
The pH-optimum of P450 BM3 is around 7.4 to 8.0, therefore commonly phosphate, Tris- or MOPS buffers are used with conductivities between 50 and 100 mM.[151] Also the applied temperatures for conversion reactions vary strongly as summarized in a recent review ranging from 15 to 37°C.[151]

1.15 Importance of Phenolic Building Blocks

Regioselective aromatic hydroxylation is one of the most desired reactions in organic chemistry for the generation of precursors for pharmaceuticals and flavors.[192-196] Especially the demand for production of phenols and alkylphenols (cresols and xylenols) was steadily increasing within the last century, as feedstock for resins, plastics or bisphenol-A.[197] With around nine megatons in 2008, phenols are one of the most synthetically produced chemicals worldwide underlining the economic importance of these compounds.[197] Phenols are mainly produced in a three-step cumene process (“Hock-Process”)[198] or can alternatively be extracted from natural sources such as tar coal, biomass or gasification processes requiring intensive extraction efforts and energy, especially for the separation of the isomeric alkylphenols.[197]

1.15.1 Chemical Synthesis of Phenol

Phenol is mainly produced by the three step cumene process also known as “Hock synthesis” (Scheme 5)[198], which is only profitable because of the valuable acetone byproduct.[197] In step 1 benzene is alkylated (Friedel-Crafts alkylation) with propylene employing AlCl_3 or H_3PO_4 , 30 atm pressure and high temperatures ($> 200^\circ\text{C}$) under anhydrous conditions. The second step, the acidic oxidation of cumene with O_2 , is performed at 5 atm and using H_2SO_4 to yield phenol and acetone. Starting from benzene as the educt, the yields are comparably low since only 5 % of the initially used benzene are converted to phenol.



Scheme 5. The cumene process or ‘Hock Process’ for the industrial production of phenol from benzene. In a first step a cumene intermediate is formed under high pressure, temperature and application of a strong acid. Oxygen binds to cumene in α -position and forms a peroxide intermediate that initiates a chain reaction. After several rearrangement steps phenol and acetone are formed. A detailed description of the phenol synthesis mechanism can be found in the original publication from Hock et al.[198]

Due to low yield of the cumene process new synthetic routes were developed circumventing the “Hock-Process” by employing for instance inorganic catalysts[193, 199], but still these strategies require large amounts of energy and harmful or toxic chemicals for synthesis.[197, 200] A direct chemical hydroxylation of benzene with molecular oxygen is till today not possible, even under high pressure and temperature employing Pd/Au catalysts.[201]

1.15.2 Biocatalytic Production of Phenol, Cresols and Dimethylphenols

Aromatic substrates have always been targets in biotechnological processes either for bioremediation or in functionalization reactions to produce valuable building blocks under mild conditions. Several organisms and enzymes were identified that are involved in aerobic as well as anaerobic pathways for degradation of these widely distributed compounds. Table 2 summarizes reported enzymes and whole cell catalysts that perform aromatic hydroxylation of benzenes.

Table 2. Selected organisms and enzymes reported for the aromatic hydroxylation of benzenes and biocatalytic production of phenol.

Substrate	Product	Enzyme	Ref.
L-tyrosine	Phenol	Tyrosine phenol lyase	[202]
Benzene	Phenol	P450 BM3	[203]
Toluene/ <i>o</i> - and <i>m</i> -Xylene	<i>o</i> -Cresol/Dimethylphenols	P450 BM3	[156, 180]
Benzene	Phenol	ToMo	[204]
Benzene/Toluene	Phenol/Cresols	Xamo	[205]

1.15.3 Aromatic Hydroxylation of Benzenes by P450 BM3

P450 BM3 could be an ideal catalyst to perform aromatic hydroxylation, for instance of *p*-xylene and pseudocumene, since several aromatic substrates are known to be hydroxylated by P450 BM3 such as toluene, *m*- and *o*-xylene, ethylbenzene, styrene, phenols, naphthalene and pyrene.[70, 151, 156, 180, 189, 191, 206-208] However, activities are comparably low for most of these substrates. In 2011, Whitehouse et al. reported the rational engineering of P450 BM3 towards improved aromatic hydroxylation of *m*- and *o*-xylene as well as toluene.[156] The *p*-xylene isomer was excluded by the authors as substrate and a previous NMR study revealed exclusively α -hydroxylation.[209] Another promising approach to enable aromatic hydroxylation of benzene substrates is the application of decoy molecules that were also used to improve activity for small and gaseous substrates such as ethane and propane by increasing sterical interactions within the active site of P450 BM3.[176, 183] By employing a perfluorinated fatty acid it was possible to hydroxylate benzene with P450 BM3 to produce phenol.[183] However, due to their persistent environmental impact[210], toxicity and cancerogenicity[211] perfluorinated fatty acids are currently under investigation and discussion to be banned in Europe.[212] The mechanism for aromatic hydroxylation of benzenes was elucidated by de Visser and Shaik and involves the epoxidation of the benzene by Compound I (Figure 7) without abstraction of a proton from the substrate.[154] After epoxidation, a proton rearrangement and a NIH shift takes place, so that the aromatic ring re-aromatizes and forms the respective phenol (Figure 11). Whitehouse et al. confirmed this mechanism experimentally by investigating the hydroxylation of *o*- and *m*-xylene with P450 BM3.[156]

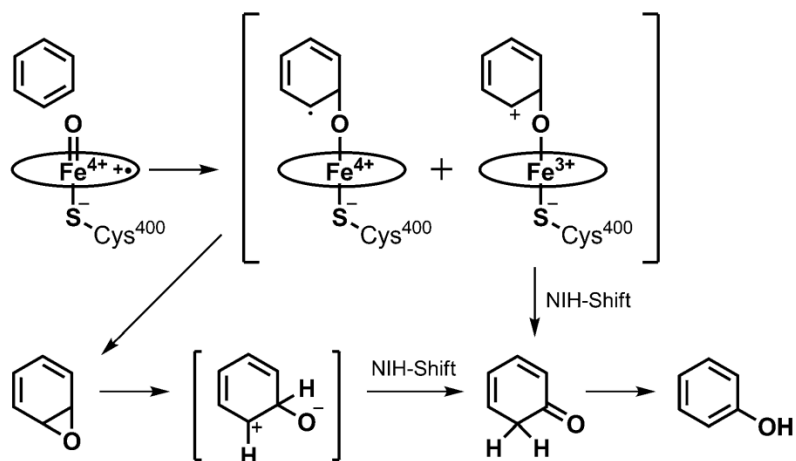


Figure 11. The mechanism for P450-catalyzed epoxidation of benzene for the biocatalytic production of phenol (taken from Shoji et al.).[183]

1.16 Objectives

Within the framework of the EU funded project ‘OXYGREEN’ (FP7-KBBE; Project Reference: 212281) the design and application of oxygenases into chemo-synthetic processes should be promoted. Designer monooxygenases can perform reactions that are till today difficult to achieve in organic synthesis. In this thesis the industrially important P450 monooxygenase from *Bacillus megaterium* BM3 should be re-engineered towards aromatic hydroxylation of *p*-xylene and pseudocumene to produce 2,5-dimethylphenol and trimethylphenols. Both compounds represent important building blocks for vitamins and pharmaceuticals. A semi-rational approach should be applied, together with a regioselective screening assay for phenols, to design a P450 BM3 variant that is capable of performing efficient and selective aromatic hydroxylation of the two target substrates. Furthermore, the semi-rational approach should also enable a deeper understanding on how P450 BM3 performs aromatic hydroxylation reactions employing benzene substrates. In combination with the catalytic data, the aromatic character of the employed substrates and their accommodation in P450 BM3’s active site should be investigated by docking studies to enable a general understanding of benzene-protein interactions. To benchmark P450 BM3 as catalyst for phenol production, the obtained P450 BM3 variants should be employed in further hydroxylation reactions with industrially important benzene educts. Key catalytic parameters such as turnover frequencies, product yields, regio- and chemoselectivity, total turnover numbers and product concentrations should be measured to evaluate the potential of P450 BM3 as catalyst for synthesis of phenols.

An important factor in every protein engineering campaign is the selection of the most promising mutagenesis strategy. Although site saturation mutagenesis is a well-established and reliable method, several limitations remain, such as the simultaneous focused mutagenesis of distant codons in a DNA sequence. As a consequence the full diversity in a protein can only be investigated sparsely, therefore eliminating potentially best amino acid combinations and to a larger extent also cooperative interactions that could lead to new and exceptional proteins. Consequently, the need for a new reliable method for saturation of multiple codons independent from their location on gene sequence is highly desirable for protein engineers. To overcome limitations in focused mutagenesis a new method for saturation of multiple codons should be developed based on the PLICing technology. If successfully accomplished the protocol could be extended to develop further methods that would fuel progress in engineering of proteins and for the implementation of biocatalyst into sustainable chemo-synthetic processes.

2. Material and Methods

2.1 Chemicals

All chemicals used were of analytical grade and purchased from Sigma Aldrich (Steinheim, Germany), ABCR (Karlsruhe, Germany), Fluka (Neu-Ulm, Germany), TCI Europe (Eschborn, Germany) and AppliChem (Darmstadt, Germany). Only a few chemicals were commercially available with less than analytical grade (e.g. 1,2,3-trimethylbenzene) and therefore not used for detailed kinetic characterization or quantification.

2.2 Enzymes and PCR Components

Restriction enzymes were purchased from New England Biolabs (Frankfurt, Germany). Phusion and Taq DNA Polymerases as well as dNTPs were purchased from Fermentas (St.Leon-Rot, Germany). Glucose dehydrogenase from *Pseudomonas sp.*, catalase from bovine liver and lysozyme from chicken egg white was provided by Sigma Aldrich (Steinheim, Germany). Oligonucleotides used for focused mutagenesis and cloning were obtained from Eurofins MWG Operon (Ebersberg, Germany) and produced at HPLC purity.

2.3 DNA Extraction and Purification with Commercial Kits

Plasmid extraction and purification of PCR products was achieved using the NucleoSpin Plasmid kit as well as NucleoSpin Gel and PCR Clean-up kit (Macherey Nagel, Düren, Germany). The kits were used according to producer's manual instructions for optimal yields and purities of DNA. All extracted and purified DNA samples were analyzed for quantity and purity on a NanoDrop spectrophotometer and by agarose gel electrophoresis.

2.4 Machines and Equipment

Äktaprime plus with UV detector	GE Healthcare (München, Germany)
BioPhotometer Plus	Eppendorf (Hamburg, Germany)
Centrifuges 5810 R and 5424	Eppendorf (Hamburg, Germany)
CO-lecture bottle station	Sigma Aldrich (Steinheim, Germany)
Drying cabinet FP 53	Binder GmbH (Tuttlingen, Germany)
Evaporator IKA RV 10 basic	IKA-Werke (Staufen, Germany)
Mini-Sub Cell GT System electrophoresis	Bio-Rad (München, Germany)
EV222 Power Supply Consort	Anachem Ltd (Luton, UK)
Freeze dryer Alpha 1-2 LD plus	Christ (Osterode am Harz, Germany)
GC-FID 2010 and 2010-Plus	Shimadzu GmbH (Duisburg, Germany)

GC-MS QP2010	Shimadzu GmbH (Duisburg, Germany)
Gel documentation U:Genius	Syngene (Cambridge, UK)
Glass column TAC15/125PE5-AB-2	YMC Europe GmbH (Dinslaken, Germany)
High Pressure Homogenizer EmulsiFlex C3	Avestin (Ottawa, ON, Canada)
HPLC	HPLC System Gold (Beckmann Coulter, München, Germany)
HPLC column Nucleosil 100-5 C18	Macherey Nagel (Düren, Germany)
Microtron microtiter plate shaker	INFORS (Heinsbach, Germany)
Microtiter plate reader Tecan Sunrise	Tecan Group AG (Männendorf, Switzerland)
Mini-Protean Tetra Cell system	Bio-Rad (München, Germany)
Multipoint magnet stirrer Variomag 15	Thermo Fischer Scientific (Wilmington, DE, USA)
NanoDrop Spectrophotometer 1000	Thermo Fischer Scientific (Wilmington, DE, USA)
Quartz glass cuvette (10.00 mm)	Hellma (Muellheim, Germany)
Quartz glass microtiter plate (96 wells)	Hellma (Muellheim, Germany)
Rotary shaker Certomat S2	Sartorius (Göttingen, Germany)
PCR thermocycler Mastercycler pro S	Eppendorf (Hamburg, Germany)
Sonicator Vibra-Cell VCX130	Sonics & Materials (Newton, CT, USA)
Spectrophotometer Varian Cary 50 UV	Agilent Technologies (Darmstadt, Germany)
Thermo mixer Comfort	Eppendorf (Hamburg, Germany)
Ultra centrifuge Sorvall RC-6	Thermo Fischer Scientific (Wilmington, DE, USA)
Vortex Genie2	Scientific Industries, Inc. (Bohemia, NY, USA)

2.5 Reaction Buffer

If not stated otherwise during all P450 BM3 reactions a phosphate buffer (KPi; pH 7.5; 50 mM) was used as frequently reported in literature.[70] Buffers for commercial enzymes (DNA polymerases, restriction enzymes) were used as recommended by producer's manual instructions. Reactions with commercial glucose dehydrogenase (GDH) were performed also with the respective KPi buffer.

2.6 Bacterial Strains, Plasmids and Genes

Employed bacterial strains:

E. coli DH5 α : [(supE44 Δ lacU169 Φ 80 lacZ Δ M15) hsdR17 recA1 gyrA96 thi-1 relA1)]

E. coli BL21 (DE3): [F- ompT gal dcm lon hsdSB (rB- mB-) λ (DE3 [lacI lacUV5-T7 gege 1 ind1 sam7 nin5])]

E. coli BL21 (DE3) lacI^{Q1}: [F- ompT gal dcm lon hsdSB (rB- mB-) λ (DE3 [lacI lacUV5-T7 gege 1 ind1 sam7 nin5])] (the lacI gene was inserted into the genome of *E. coli* BL21-Gold (DE3).[213]

(*E. coli* BL21 (DE3) lac^{Q1} cells were kindly provided by Dr. Alexander Schenk (Jacobs University Bremen, Germany)).

Employed bacterial plasmids are shown as vector maps in Figures 12 and 13.

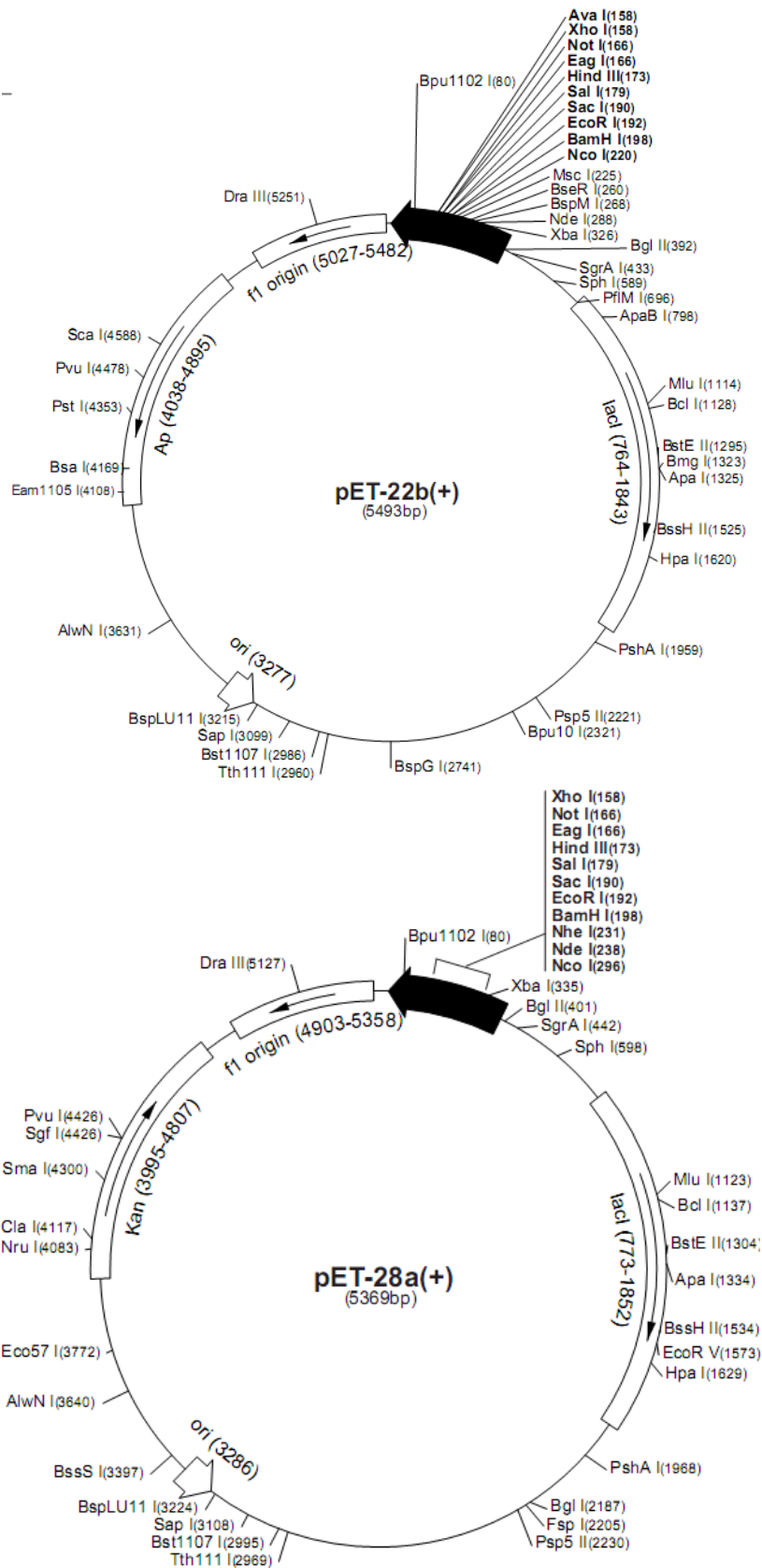


Figure 12. Vector maps of employed pET-22b(+) and pET-28a(+) high-copy expression plasmids from Novagen. The P450 BM3 gene was cloned into pET28a(+) using restriction ligation cloning.

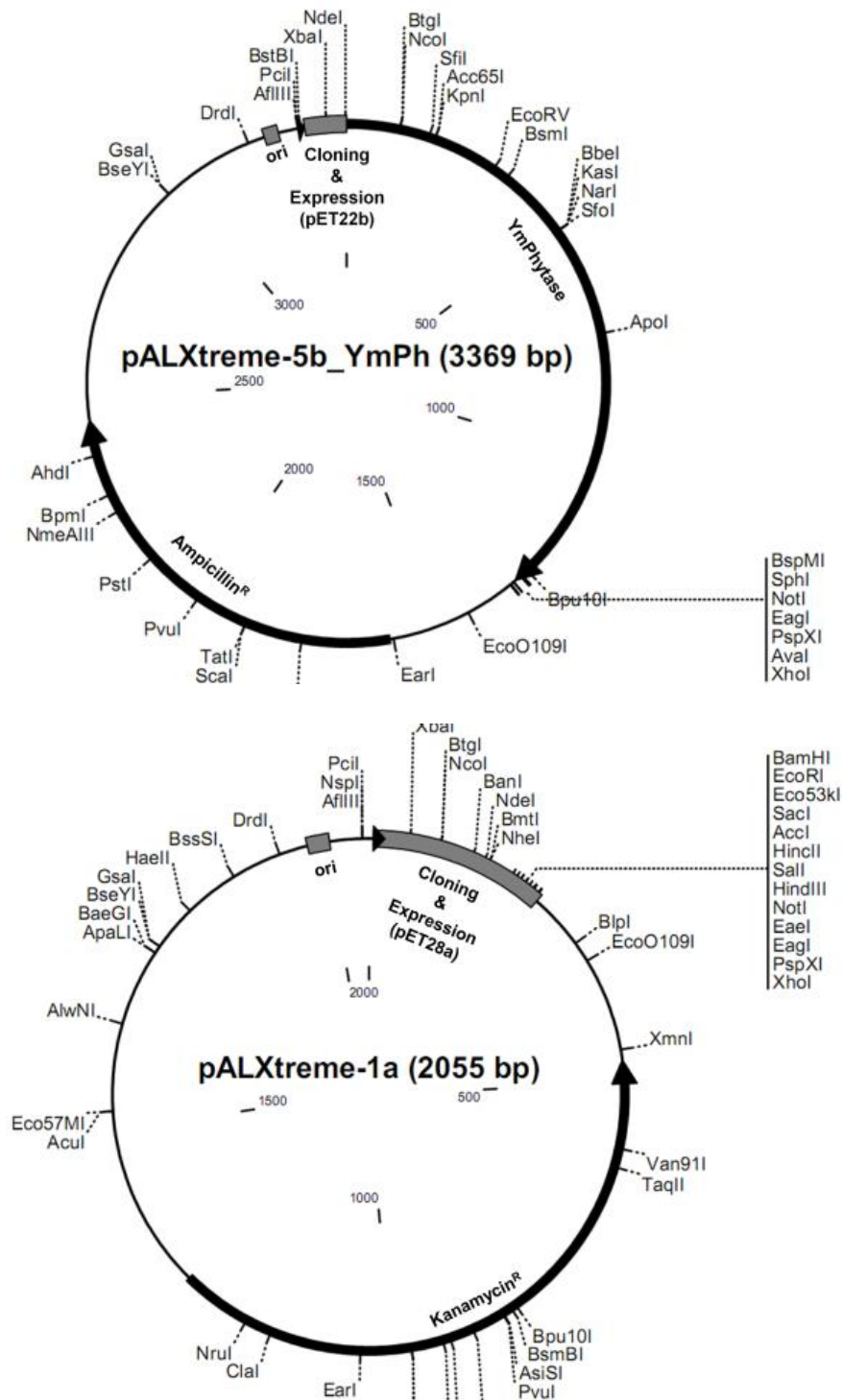


Figure 13. Vector maps of pET-derived pALXtreme-5b (pET-22b(+)) and pALXtreme-1a (pET-28a(+)), kindly provided by Dr. Alexander Schenk (Jacobs University Bremen, Germany). Both plasmids were constructed by removing >60% of sequence from the template pET-vectors.[213] The plasmid pALXtreme-5b_YmPh containing the phytase gene from *Yersinia mollaretii* (YmPhytase) was kindly provided by Dr. Amol Shivange (Jacobs University Bremen, Germany).[214]

For controlled and reliable expression the pALXtreme vectors were always used in combination with the genetically engineered *E. coli* BL21 (DE3) $lacI^{Q1}$ strain which contains the *lacI* repressor gene

integrated into the genome of *E. coli* BL21 (DE3).[213] Table 3 summarizes the genes used in this work and their origin.

Table 3. Employed genes in this work and their origin.

Gene Name	Origin	Accession number	Reference
BM3	<i>Bacillus megaterium</i>	J04832	[150]
YmPhytase	<i>Yersinia mollaretii</i>	JF911533	[214]
appA	<i>Escherichia coli</i>	AF537219	[215]
phyX	<i>Hafnia alvei</i>	---	[216]

A full list of all generated constructs and libraries can be found in the appendix section of this thesis (Table 28)

2.7 Cultivation Media and Additives

2.7.1 Liquid Cultivation Media:

LB media: 10 g L⁻¹ tryptone, 5 g L⁻¹ yeast extract, 10 g L⁻¹ NaCl, pH 7.0

TB media: 10 g L⁻¹ tryptone, 5 g L⁻¹ yeast extract, 4 g L⁻¹ glycerol, 0.1 M KPi pH 7.0

SOC media: 20 g L⁻¹ tryptone, 5 g L⁻¹ yeast extract, 0.5 g L⁻¹ NaCl, 0.1 M KCl,

All media components were solubilized in ddH₂O and autoclaved for 20 min. Sterilized culture media were stored at RT until use or as frozen 1 mL stocks at -20°C (SOC media).

2.7.2 Solid Media:

LB agar: 10 g L⁻¹ tryptone, 5 g L⁻¹ yeast extract, 10 g L⁻¹ NaCl, pH 7.0, 15 g L⁻¹ agar agar.

All media components were solubilized in ddH₂O and autoclaved for 20 min. Corresponding antibiotics were added to hand warm (~ 50°C) LB agar. Around 15 mL of liquid LB agar were filled into plastic Petri dishes for hardening. Prepared LB agar plates were stored at 4°C until use.

2.7.3 Media Additives:

All media additives were prepared as 1000x solutions. The additives were solubilized in ddH₂O and sterilized by pressing solutions through a 0.2 µM filter.

ALA (Aminolevulinic acid): 0.5 M solubilized in ddH₂O.

Thiamine: 100 g L⁻¹ solubilized in ddH₂O.

Kanamycin: 50 g L⁻¹ in ddH₂O.

Ampicillin: 50 g L⁻¹ in ddH₂O.

IPTG (Isopropyl- β -D thiogalactopyranoside): 0.1 M solubilized in ddH₂O.

Trace Element Solution: 0.5 g L⁻¹ CaCl₂ · 2H₂O, 0.18 g L⁻¹ ZnSO₄ · 7H₂O, 0.1 g L⁻¹ MnSO₄ · H₂O, 20.1 g L⁻¹ Na₂-EDTA, 16.7 g L⁻¹ FeCl₃ · 6H₂O, 0.16 g L⁻¹ CuSO₄ · 5H₂O dissolved in ddH₂O.

The trace element solution was stored at 4°C until use. All other solutions were aliquoted as 1 mL stocks and stored at -20°C until use.

2.8 Microbiological Methods

2.8.1 Preparation of Chemical Competent *Escherichia coli* Cells

Many bacterial strains have the natural ability for the admission of external DNA by natural competence mechanisms.[217] In this work, the chemical transformation was used to introduce plasmid DNA into competent *E. coli* cells.[217, 218] To achieve a high amount of transformants the natural competence was increased by using a standard preparation protocol for *E. coli*. [217, 218] An *E. coli* pre-culture was grown in 4 mL LB media (250 rpm, 16 h, 37°C). One milliliter of pre-culture was used to inoculate 200 mL LB media in a 1000 mL shake flask. The main culture was grown until OD₆₀₀ 0.4 (250 rpm, 37°C). Subsequent steps were performed on ice or 4°C. Residual LB media was removed by centrifugation (4500 x g, 10 min, 4°C). The obtained cell pellet was resuspended in 15 mL TFB1 solution (30 mM K-Acetate, 50 mM MnCl₂, 100 mM RuCl₂, 10 mM CaCl₂, 15 % glycerol; pH 6.8; sterile filtrated; stored at 4°C) and incubated for 10 min on ice. Resuspended cells were centrifuged and resuspended in 2 mL TFB2 solution (10 mM MOPS, 75 mM CaCl₂, 10 mM RuCl₂, 15% glycerol; pH 6.8; sterile filtrated; stored at 4°C). TFB1 and TFB2 were sterilized using a sterile 0.2 μ M filter. Cells resuspended in TFB2 solution were aliquoted into sterile and pre-cooled 2 mL plastic tubes. Aliquots of 100 μ L competent cells were directly frozen in liquid N₂ and stored until use at -80°C.

2.8.2 Transformation of Plasmid DNA into *Escherichia coli*

For transformation of plasmid DNA 100 μ L of frozen competent *E. coli* cells and plasmid DNA were thawed on ice (15 minutes). Commonly, 1 μ L containing total 5 ng of isolated and purified plasmid DNA were mixed with 100 μ L of chemical competent cells. Cells and plasmid DNA were incubated (15 min; on ice) prior to transformation by heat shock.[217] The transformation was achieved by placing tube with cells and plasmid DNA into a water bath (42°C) for 45 sec. After the heat shock, tubes were placed on ice for five minutes. Subsequently 900 mL of pre-heated (37°C) SOC media (section 2.7.1) were supplemented to the transformation mixture. Cells were recovered for 45 min (250 rpm, 37°C) before plating on LB agar plates with corresponding antibiotics (section 2.7.3).

The competence of new generated batches of competent cells was always confirmed by transforming cells with 1 ng pUC19 plasmid DNA.[219] The transformed and recovered cells were plated subsequently in dilutions of 1:10, 1:25, 1:50 and 1:100 (in triplicate) on LB agar plates containing the

corresponding antibiotics. Cells were grown overnight at 37°C and counted on the next day for statistical evaluation.[217, 218] The amount of cells and the corresponding dilution factor were used to calculate the competence of the prepared cells with following equation:

$$\text{Transformation efficiency [cfu } \mu\text{g}^{-1}] = \text{Number of colonies on an agar plate} \times \text{dilution factor} \times \mu\text{g}^{-1} \text{ DNA}$$

2.8.3 Preparation of *Escherichia coli* Cryo-Cultures

Glycerol stocks of bacterial strains were prepared by striking one colony from a plasmid transformation (section 2.8.2) with a sterile inoculation loop on a LB agar plate containing appropriate antibiotics. After incubation (16 h, 37°C), the grown cells were removed from the agar plates using a sterile inoculation loop. Cell mass was resuspended in an autoclaved plastic test tube filled with 1 mL 50 % glycerol (w/w). The cells were resuspended by vortexing the tube thoroughly prior to long time storage at -80°C.

2.8.4 Shake Flask Expression of P450 BM3

Pre-cultures for expression in shake flasks were prepared by inoculating a sterile glass tube containing 3 mL LB media with cells from the glycerol stock using a sterile inoculation loop (section 2.8.3). Cells were grown until reaching stationary phase (250 rpm, 16 h, 37°C). From the pre-culture 1 mL of culture broth was transferred into a sterile Erlenmeyer flask (1000 mL) containing 100 mL TB media, trace elements as well as appropriate antibiotics. P450 BM3 expression was induced by addition of IPTG, ALA and thiamine at optical density of 0.8 at 600 nm wavelength (OD_{600}).[160] Cells were harvested after 16 h of expression by centrifugation (4500 x g, 15 min, 4°C). After discarding remaining media, the pellets containing expressed P450 BM3 were stored until use at -20°C. As a control the expression protocol was performed in the same way with the empty vector plasmid containing no gene for heterologous expression (section 2.6).

2.8.5 Cell Lysis with High-Pressure Homogenizer and Sonicator

Frozen cell pellets from shake flask expressions were thawed on ice and resuspended thoroughly in 10 to 20 mL KPi buffer (50 mM; pH 7.5) until no solid particles were visible. Efficient cell disruption was achieved by a high pressure homogenizer (“French press”) performing three iterative cycles at 1500 bar pressure. Cell debris was removed by centrifugation in an ultra centrifuge (24000 x g, 15 min, 4°C). The supernatant was stored at 4°C until and used for P450 quantification with the CO binding assay (section 2.10.8). Alternatively, smaller cell pellets were lysed by ultrasonication. Therefore the cell pellet was thawed on ice and resuspended in KPi buffer (50 mM, pH 7.5) thoroughly by vortexing. Resuspended cells were placed on ice and disrupted by sonication. Following sonication program was employed for three cycles: 40 % amplitude, 3 cycles à 30 s, 30 s break (cooling). Finally lysed cells were centrifuged (20000 x g, 10 min, 4°C) for entire removal of cell debris and the clear supernatant was used for further experiments.

2.8.6 Mutant Library Preparation and Expression in Microtiter Plates

Mutant colonies were picked with sterile toothpicks from agar plates into flat bottom 96well MTPs filled with 100 μ L of LB media (section 2.7.1) and appropriate antibiotics. Screening for improved variants requires reliable references (positive and negative controls). As negative control on each MTP serves the plasmid without any cloned gene within the multiple cloning site (MCS), the empty vector (EV) (Figure 12 and 13). For a positive control in a screening experiment and for selection of improved variants serves the wild-type (WT) or the best enzyme variant from a previous round of screening. To ensure a high confidence level for selection of improved variants, both controls were placed in triplicate on each MTP (Figure 14).

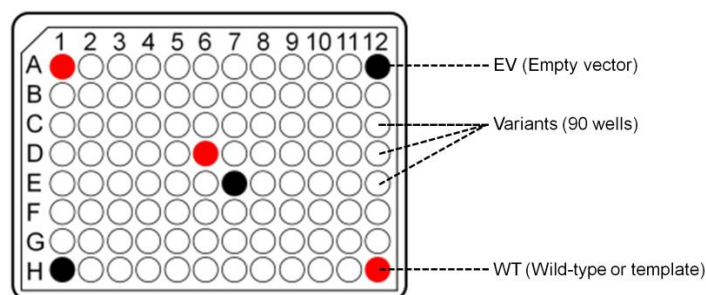


Figure 14. Scheme for MTP arrangement of generated variants (white colored wells) together with a positive control (e.g. WT; red colored wells: A1; D6 and H12) and negative control (EV; black colored wells; A12, E7 and H1) for reliable screening.

After each well was inoculated with the respective colony from an agar plate the plate was closed with a sterile lid and sealed with tape to prevent evaporation. Cells were grown overnight in a MTP shaker (900 rpm, 37°C, humidity control). To grown cultures 100 μ L of 50 % (W/V) sterile glycerol solution were supplemented and the plates stored at -80°C prior to expression and screening for improved variants (Figure 15).

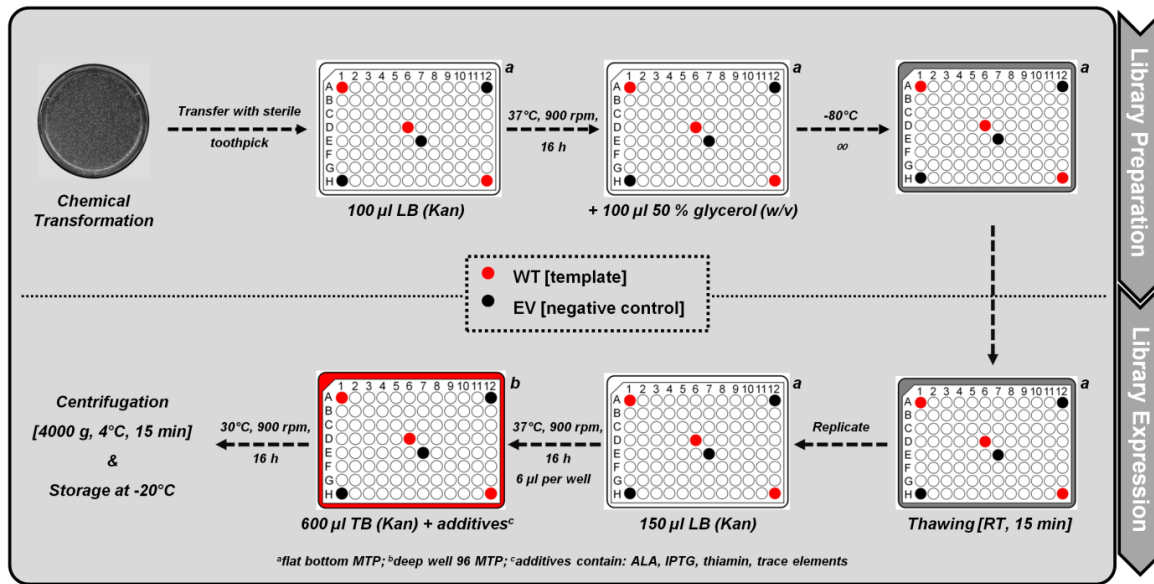


Figure 15. Preparation and expression process for mutant libraries in MTP format. **Library preparation:** After chemical transformation clones are picked into a MTP pre-filled with LB media. The plate is incubated for 16 h to grow cells to high density. Finally glycerol is supplemented prior to long time storage at -80°C. **Library Expression:** Frozen plates were thawed at RT for 15 min and cell mass is replicated into a new MTP filled with LB. After 16 h cultivation at 900 rpm and 37°C, 6 µL from each well were transferred into a deep well MTP filled with TB media and all additives necessary for expression of P450 BM3.[160, 161] Protein expression was performed for 16 h at 900 rpm and 30°C. Residual media was removed by centrifugation and the plates were stored at -20°C until use.

Glycerol stocks of mutant libraries in MTPs were thawed at RT for 15 min. After the plates thawed completely, the cells were replicated with a sterile metal-pin replicator into flat-bottom 96 well MTPs filled with 150 µL sterile LB media and the corresponding antibiotics. Inoculated plates were closed with a sterile MTP lid and sealed with tape to prevent evaporation. Pre-culturing of libraries was achieved in a MTP shaker (900 rpm, 37°C, humidity control) for 16 h. Six µL from each well of the pre-culture were pipetted into 600 µL TB media pre-filled into a 96 deep-well MTP. The TB-media was supplemented with all additives necessary for reliable expression of P450 BM3 (section 2.7.1).[160] Deep well MTPs were closed with a sterile lid and sealed with tape. Expression in deep well MTPs was performed for 20 h in a MTP shaker (900 rpm, 30°C). After expression of mutant libraries the MTPs were centrifuged to remove residual media (3400 x g, 15 min, 4°C). Cell pellets were stored in deep well MTPs for at least 12 h at -20°C until use in screening assays (section 2.10).

2.8.7 Preparation of Microtiter Plates for Screening

Deep well MTPs containing pellets of expressed P450 BM3 monooxygenase were thawed at RT for 15 min. Subsequently 150 µL KPi buffer (50 mM, pH 7.5) were supplemented and the MTP incubated for 5 min at RT. Pellets were resuspended thoroughly by vortexing deep well MTPs until all pellets were entirely resolved. For full lysing of *E. coli* BL21 lacI^{Q1} cells additional 150 µL KPi buffer (50 mM, pH 7.5) containing 5 mg mL⁻¹ lysozyme were supplemented to the resuspended pellets.[160] The plates were closed with a lid and incubated for 1 h in a MTP shaker (900 rpm, 37°C). Cell debris, insoluble or solid compounds were removed by centrifugation (3400 x g, 15 min, 4°C). Cell lysates in

deep well MTPs were stored at 4°C until final application in activity screening with the respective assays (section 2.10.2 and 2.10.3).

2.9 Molecular Biological Methods

2.9.1 DNA Extraction, Storage and Sequencing

All plasmid extractions from *E. coli* strains were performed using the Macherey Nagel Plasmid extraction kit and recommend steps according to producer's manual instructions. For extraction *E. coli* BL21 (DE3) lacI^{Q1} cells containing corresponding plasmids were grown in 3 mL LB media supplemented with corresponding antibiotics as overnight culture (250 rpm, 16 h, 37°C). Extracted and purified plasmids were quantified using a NanoDrop spectrophotometer prior to long time storage at -20°C. For sequencing 1 µg of plasmid DNA was sent for sequencing. Primers for sequencing can be found in appendix section of this thesis (Table 27). Obtained sequencing files were analyzed using CloneManager 9 software for instance by annealing the sequence data of variants with the template DNA sequence used for mutagenesis (see appendix Figure 89).

2.9.2 Polymerase Chain Reaction (PCR)

In vitro amplification of DNA was achieved by polymerase chain reaction (PCR) using Taq DNA or high fidelity Phusion DNA polymerase.[113, 220] Reaction mixtures (if not stated otherwise) contained in a final volume of 50 µL: 1x DNA Polymerase Puffer; 5 U DNA Polymerase (Taq or Phusion); 0.2 mM dNTP mix; 20 ng plasmid DNA; 0.2 µM forward und reverse primer. Amplification of DNA was achieved in a thermocycler employing the program shown in Table 4.

Table 4. Thermocycler program for PCR amplification with Taq DNA polymerase.

Step	Temperature [°C]	Duration [sec]	
Denaturation	94	20	} 25 cycles
Annealing	55	30	
Elongation	72	60 per kb DNA [#]	
Final Elongation	72	300	
Storage	4	∞	

[#]Elongation with Phusion DNA Polymerase was performed with 2 kb min⁻¹

2.9.3 Oligonucleotide Design for PCR Amplifications

Primers were designed as recommended for efficient amplification of DNA fragments (GC content: 40-60 %; Tm: > 55°C; avoiding primer self- or hetero-dimers and hairpins). Designed oligonucleotides were re-checked with OligoAnalyzer Software from IDT (<http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/>).[221]

2.9.4 Colony PCR

Colony PCR was performed to control correct assembly of generated DNA constructs. Therefore, a cell material from a colony grown on solid LB agar or 5 μL liquid cell culture were re-suspended in 50 μL ddH₂O. Plasmid DNA was released by heating the sample to 95°C for 15 min. Cell debris was removed by centrifugation (20000 x g, 5 min, 20°C) and 5 μL from the supernatant was used in total 50 μL PCR mixture (section 2.9.2). All colony PCRs were performed using Taq DNA polymerase and the respective PCR amplification protocol (Table 4).

2.9.5 Phosphorothioate-Based Ligase-Independent Gene Cloning (PLICing)

In this thesis all DNA cloning steps were performed employing the phosphorothioate-based ligase-independent DNA cloning method (PLICing).[213] PLICing allows a nearly sequence independent cloning of DNA fragments, with high efficiency and very low background of false positive colonies. Making use of a chemical cleavage step to generate specific DNA overhangs for later hybridization, limitations of the widely used restriction-ligation cloning are circumvented. Figure 16 illustrates the principle of the PLICing cloning method (B) and the modification in the phosphorothiolated nucleotides applied in the protocol (A).

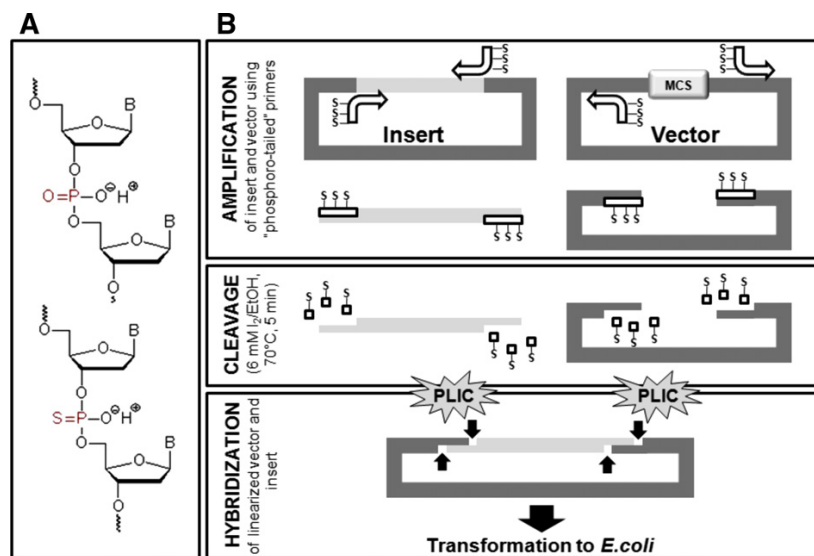


Figure 16. A schematic representation of the PLICing method (taken from Blanus et al.)[213] **A:** Difference between natural phosphodiester bonds (upper) and the PTO modified nucleotide bonds (lower) employed in the PLICing method. **B:** Three step PLICing protocol for cloning of DNA fragments. The Figure was taken from Blanus et al.[213]

In the first step, the amplification of insert and vector was achieved with modified (phosphorothiolated) primers employing a standard PCR protocol (section 2.9.2). The products were analyzed by agarose gel electrophoresis for quality of the amplified DNA (section 2.9.6). After DpnI digest and purification with a DNA purification kit, the PCR products were diluted to 0.03 pmol μL^{-1} insert and 0.01 pmol μL^{-1} vector, respectively. Cleavage of phosphorothiolated nucleotides was achieved by mixing 4 μL of insert DNA and 4 μL vector DNA, each with 2 μL of iodine cleavage mix

(6 mM I₂ in EtOH; 250 mM Tris-HCl pH 9.0). The samples were heated in a thermocycler (5 min, 70°C) for cleavage of PTO nucleotides from 5' endings. The mechanism for cleavage of PTO containing DNA was elucidated by Eckstein and Gish[222] and is shown exemplary in Figure 17.

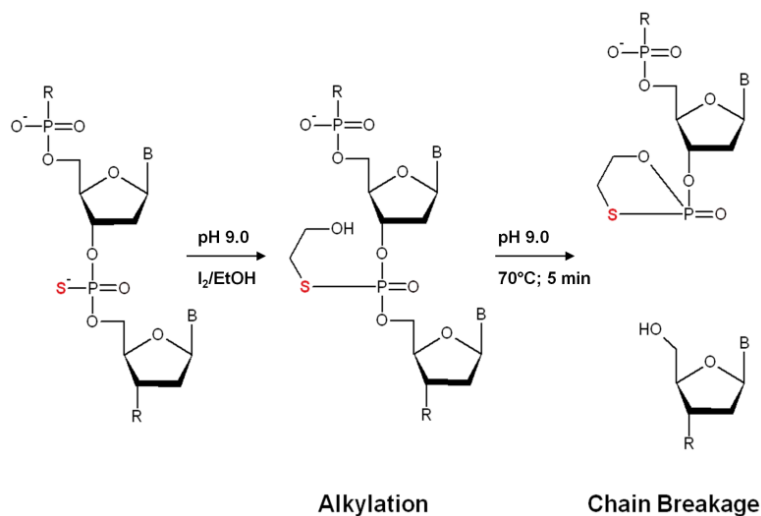


Figure 17. Chemical cleavage of phosphothiolated nucleotides in the presence of iodine and ethanol (I₂/EtOH) under alkaline conditions. In step one the sulfur atom is alkylated by iodoethanol leading to an instable intermediate which releases the phosphorothiolated nucleotide from the DNA. The detailed mechanism was introduced by Eckstein and Gish in 1989.[222]

Cleaved DNA fragments were mixed and transformed by heat shock into chemical competent *E. coli* cells (section 2.8.2) without further purification step.[124, 213, 223] The obtained clones were analyzed by colony PCR and agarose gel electrophoresis (section 2.9.4 and 2.9.6). One colony containing the construct in expected size was used for plasmid extraction and preparation of glycerol stocks. Correct assembly of the plasmid was verified by sequencing of the generated DNA constructs.

2.9.6 Agarose Gel Electrophoresis and DNA Quantification

Quality control of all DNA samples (PCR products, enzymatic restriction, plasmid DNA) was achieved by gel electrophoresis on 1 % agarose gels. Smaller DNA fragments (<500 bps) that were generated during development of OmniChange[223] (3.7) and PTRec[124] (3.8) were analyzed on 2 % agarose gels for higher resolution. Solid agarose powder was solubilized by heating in Tris EDTA acetate buffer (TAE-buffer; 40mM Tris; 2 mM EDTA; pH 8.0) in a kitchen microwave. Five μ L of DNA samples were mixed with 6x loading dye. Samples were loaded on an agarose gel and separated for 30 min at 110 mV (2 h for 2 % agarose gels). Staining of DNA was achieved employing ethidium bromide (40 μ L ethidium bromide solubilized in 1000 mL ddH₂O). Visualization of DNA samples on agarose gels was achieved on a UV illuminator connected with a gel scanning system. Quantification of DNA (plasmid DNA or PCR products) was done by using a NanoDrop spectrophotometer.

2.9.7 Cloning of P450 BM3 Wild-Type into pALXtreme-1a Expression Vector

Smaller plasmids have significant advantages in directed evolution campaigns due to a higher efficiency during amplification of DNA as well as increased transformation efficiencies.[224] Therefore the P450 BM3 WT gene was cloned from pET-28a(+) into pALXtreme-1a using the PLICing protocol[213] which leads to a decrease in plasmid size of more than 30 % (section 2.6). The primer sequences employed during cloning can be found in the appendix section of this thesis (Table 27). The generated construct was sequenced and used for all engineering work and characterizations of P450 BM3 variants (see full DNA sequence of P450 BM3 WT in appendix section).

2.9.8 Site Directed and Site Saturation Mutagenesis

Focused mutagenesis is a standard approach in protein engineering for the substitution of single amino acids in protein sequences in a defined manner.[32, 36, 98, 103] The most common protocol for site directed mutagenesis (SDM) and site saturation mutagenesis (SSM) is the whole plasmid amplification using complementary synthetic oligonucleotides also known as the “QuikChange protocol”. [126, 225] Main challenge of this protocol are the complementary primers which prevent efficient DNA amplification due to formation of strong heterodimeric structures which is exemplary shown for saturation of positions R47 and Y51 (P450 BM3) in this thesis (Figure 18).[223]

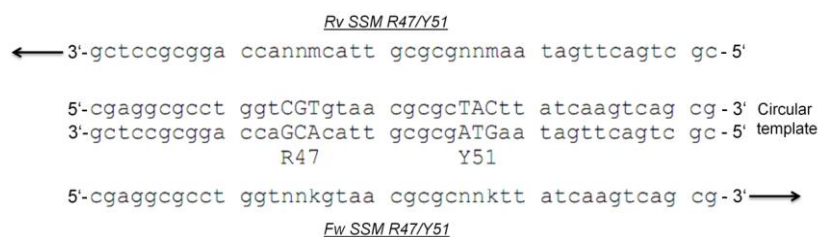


Figure 18. Simultaneous site saturation mutagenesis of rationally selected amino acid residues R47 and Y51 of P450 BM3. The principle for hybridization of complementary primers containing NNK-degenerated nucleotides to a circular plasmid is shown. PCR amplification is in 3'-direction, catalyzed by Phusion DNA Polymerase.

In this thesis a modified two-step PCR protocol is used to obtain higher yields of PCR products during generation of focused mutant libraries by reducing formation of heterodimers of oligonucleotides.[221] All primers were designed according to recommendations provided in the QuikChange SDM manual and using OligoAnalyzer Software from IDT (www.eu.idtdna.com/analyzer/applications/oligoanalyzer). [125, 226] In step 1 of the modified protocol the PCR is performed with single primers for three cycles (section 2.9.2) to generate linear templates. Step 2 comprises mixing of both reaction mixtures from step 1 and performing 15 additional PCR cycles under identical conditions. After 15 cycles of exponential DNA amplification, the samples were analyzed for PCR product quantity and quality by agarose gel electrophoresis (section 2.9.6). Figure 79 in appendix section displays PCR products from SSM at position F87 of P450 BM3. Correctly amplified products were supplemented with 10 U of DpnI

restriction enzyme to remove methylated template DNA (16 h at 37°C). Digested PCR products (5 µL) were directly transformed into chemical competent *E. coli* lacI^{Q1} cells (section 2.8.2). Obtained colonies were used to prepare mutant libraries for screening in MTP format (section 2.8.6).

2.10 Biochemical Methods

2.10.1 Screening of Mutant Libraries in Microtiter Plates

A directed evolution experiment is based on diversity generation combined with efficient selection, mimicking Darwinian evolution in a test tube format as *in vitro* approach.[33, 92] A reliable selection of variants is only achieved together with an efficient screening process that represents a major bottleneck for most protein engineering campaigns.[36, 37, 129, 227] Common thresholds for efficient medium or high-throughput systems should not exceed 15 % of deviation in activity for the same variant under identical conditions.[191, 227, 228] Before starting a screening process the standard deviation has to be determined using references such as a positive (active) as well as a negative control (inactive), where for instance the WT enzyme serves as positive and the EV as negative control. In case of screening for improved P450 enzymes *E. coli* is the ideal host organism since it does not contain monooxygenase enzymes (Figure 19 A).[70]

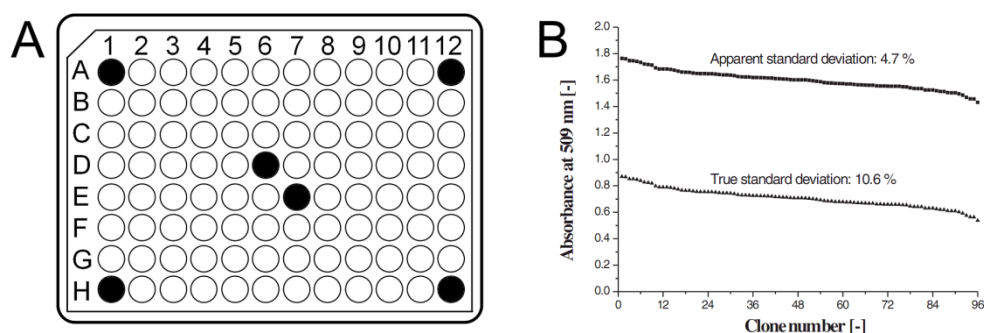


Figure 19. A: Distribution of clones into a MTP for estimation of standard deviation of hydroxylation reaction of P450 BM3. Black labeled wells contain *E. coli* BL21 (DE3) lacI^{Q1} cells harboring the empty vector plasmid (no P450 monooxygenase; negative control). White filled wells contain *E. coli* BL21 (DE3) lacI^{Q1} cells expressing P450 BM3 WT (90 times). **B** Reported standard deviation estimation for the hydroxylation of 11-phenoxyundecanoic by P450 BM3 and detection by 4-AAP assay.[228] Apparent standard deviation displays absolute values measured in each of the wells in a MTP. The 'true standard deviation' displays activity without measured background for instance absorbance values for EV controls. Figure 19 B was taken from Wong et al.[228]

In this thesis the widely applied NADPH depletion assay is used for identification of improved P450 BM3 variants.[229] Due to the 'uncoupling challenge' in P450 catalysis, product based detection systems are favored such as the 4-AAP assay for detection of phenolic products[228] or the continuous pNCA assay that can be applied as dummy substrate for fatty acid hydroxylation.[184] For every substrate, co-solvent or modified parameter in the screening procedure the standard deviation has to be recorded and calculated as shown in Figure 19 B.

2.10.2 The NADPH Depletion Assay

Recording NAD(P)H depletion is a widely used assay for oxidoreductases such as P450 monooxygenases.[229] In this thesis, the assay was used during all screening procedures to identify P450 M3 variants with higher activity towards target substrates. Therefore lysate from variant libraries was prepared as described before (section 2.8.7). Two separate MTPs are prepared to identify variants with improved oxidation activity for the NADPH cofactor, one MTP without substrate (control MTP) and one containing the target substrate (sample MTP) for hydroxylation. The two plates were prepared according to Table 5.

Table 5. Preparation of MTP plates for screening with the NADPH depletion assay. The sample MTP contains the target substrate whereas the control MTP does not contain a substrate for hydroxylation. Five minutes incubation are required to adjust substrate concentration within the active site of P450 BM3.

	Compound	Sample MTP [μL]	Control MTP [μL]
	DMSO	10 (+ substrate)	10 (- substrate)
	Buffer (KPi; pH 7.5; 50 mM)	140	140
	Crude cell extract	50	50
5 min incubation of P450 with the enzyme; in the meantime: measurement of blank-value of both plates with at a wavelength of 340 nm.			
<i>supplement</i>	NADPH (0.8 mM)	50	50

Immediately after supplementing NADPH the absorption at 340 nm is recorded for 10 min with intervals of 10 seconds (60 cycles). Depending on P450 activity in the variant library and employed amount of protein the length for intervals or total measurement time can be adjusted in a way that linear depletion of NADPH is monitored reliably.

2.10.3 Phenol Detection by 4-Aminoantipyrine Assay (4-AAP Assay)

Sensitive detection of phenols is of great importance for assaying toxic compounds in environmental samples and health control. Therefore the 4-aminoantipyrine assay was introduced in the 1940s that serves efficiently for reliable and sensitive detection of phenols ($\mu\text{g L}^{-1}$) in aqueous solutions.[230] The assay is based on the formation of a phenol-4-aminoantipyrine (4-AAP) dye complex by oxidative coupling, leading to an extended conjugated electron system with strong absorbance at 509 nm.[228, 230] In 2005, the assay was transferred to a MTP-based screening format for protein engineering campaigns with P450 BM3 and aromatic substrates such as 11-Phenoxyundecanoic acid (Figure 20).[228]

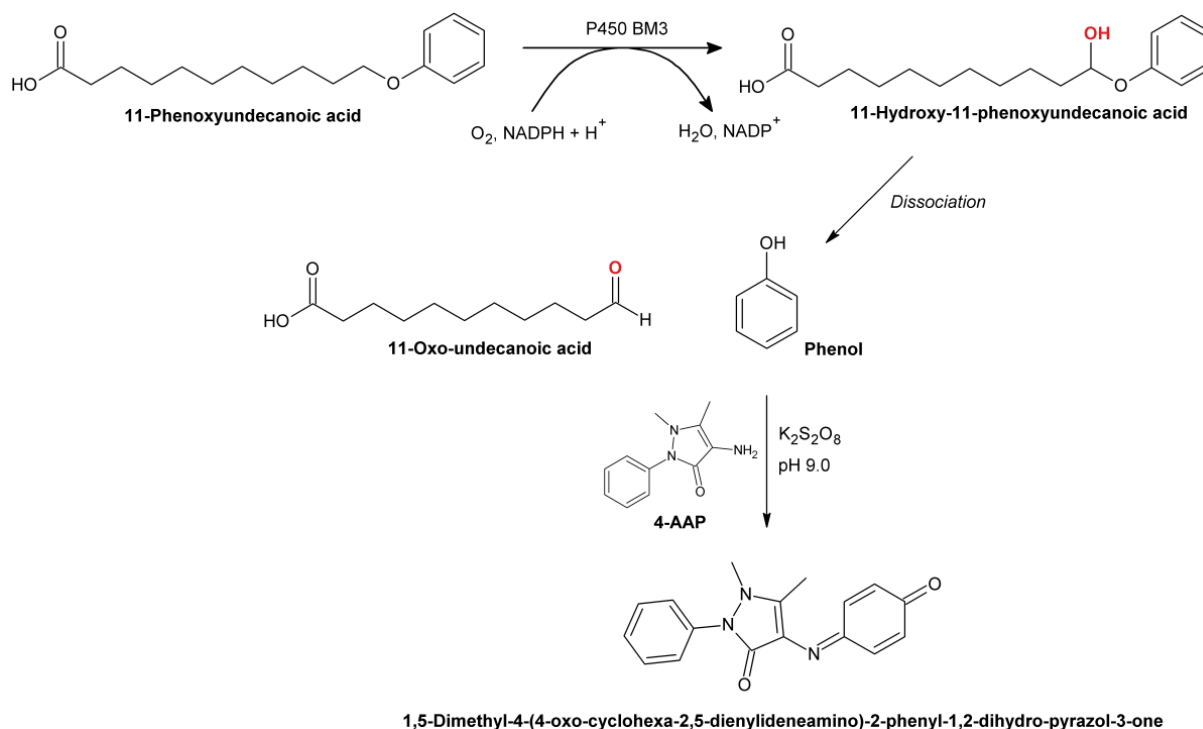


Figure 20. The 4-AAP assay for phenolic product detection in MTP format.[228] After hydroxylation of 11-phenoxyundecanoic acid one molecule of phenol is dissociated due to formation of an instable hemiacetal compound. The phenol reacts in *p*-position under alkaline condition and catalyzed by potassium peroxodisulfate with 4-AAP forming a product that displays strong absorbance maxima at 509 and 600 nm. The Figure was redrawn from Wong et al.[228]

In this work, the 4-AAP assay was slightly modified compared to the published protocol and directly applied for screening and selection of P450 BM3 variants that are capable to perform aromatic hydroxylation of *p*-xylene (3.1) and pseudocumene (3.2).[69] During all experiments 96-well quartz glass MTPs were used since *p*-xylene and pseudocumene solubilize poly-propylene and poly-styrene based materials. The NADPH depletion assay was performed as described before (section 2.10.9.1). After full depletion of NADPH in the first wells the reaction was stopped by pipetting 25 μl quenching buffer (4 M urea in 0.1 M NaOH) into the reaction mixtures. Finally 20 μl of 4-aminoantipyrene (4-AAP) (5 mg mL⁻¹) and 20 μl potassium peroxodisulfate (5 mg mL⁻¹) are supplemented for phenolic product detection (Figure 20).[69] After 30 min incubation at 750 rpm on a MTP shaker the amount of phenol was measured at 509 nm and 600 nm wavelengths in a MTP reader. The combination of NADPH depletion and 4-AAP assay allows detection of the most active and the most productive variants in a single MTP.

2.10.4 Statistical Evaluation of Mutant Libraries Screening and Selection of Variants

For fast and reliable evaluation of large data sets from the NADPH depletion and 4-AAP assay, a Microsoft Excel evaluation spread sheet was developed to identify the improved P450 BM3 variants. Major criteria for selection of variants are the improved NADPH depletion activity and highest relative amount of phenolic product formed during the catalytic reaction. The screening procedure and obtained data formats are exemplarily shown in Figure 21.

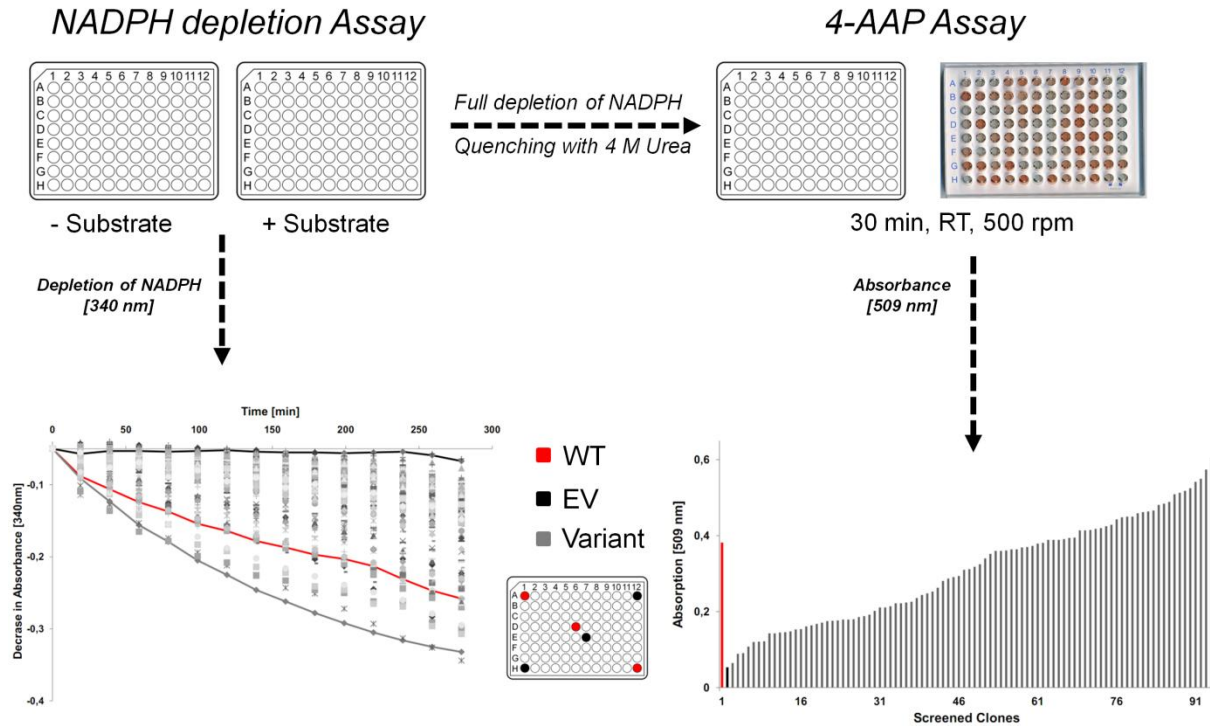


Figure 21. Combined screening with the NADPH depletion assay and the 4-AAP detection system for phenols.

The selection of P450 BM3 variants was done refining recorded assay data with following equations:

For the NADPH depletion assay:

$$\text{Improved variant}_{\text{NADPH}} = \text{Variant}_{\text{NADPH}} [\Delta\text{Abs}_{+\text{Sub}} - \Delta\text{Abs}_{-\text{Sub}}] > \text{WT}_{\text{NADPH}} [\Delta\text{Abs}_{+\text{Sub}} - \Delta\text{Abs}_{-\text{Sub}}] \times \text{TQ}$$

For the 4-AAP assay:

$$\text{Improved variant}_{\text{4-AAP}} = \text{Variant}_{\text{4-AAP}} [\text{Abs}_{\text{Variant}} - \text{Abs}_{\text{EV}}] > \text{WT}_{\text{4-AAP}} [\text{Abs}_{\text{WT}} - \text{Abs}_{\text{EV}}] \times \text{TQ}$$

ΔAbs : NADPH depletion velocity (slope) at 340 nm; +/-Sub: reaction sample with (+) or without (-) supplemented substrate; Abs: Absorbance at 509 nm; WT: wild-type; EV: empty vector; TQ: threshold quantity.

The threshold quantity (TQ) for selection of improved variants was adjusted depending on relative activity of WT and variants throughout the library. Variants with increased NADPH depletion velocity and/or phenolic product formation were selected for a statistical more evident rescreening.

2.10.5 Rescreening of Improved Variants

An important factor in statistical evaluation of large data sets is the reduction of false positive and false negative events. Especially deviations in the expression of variants[164] or temperature effects in MTP preparation, expression and screening can lead to wrong data interpretation.[231] To identify improved variants a deep statistical analysis is necessary. Therefore selected variants, commonly 5 best hits, from pre- screenings are picked multiple times (>3) into a new MTP together with

reference controls (section 2.8.6). The screening was repeated again employing the NADPH depletion assay and 4-AAP assay (Figure 22).

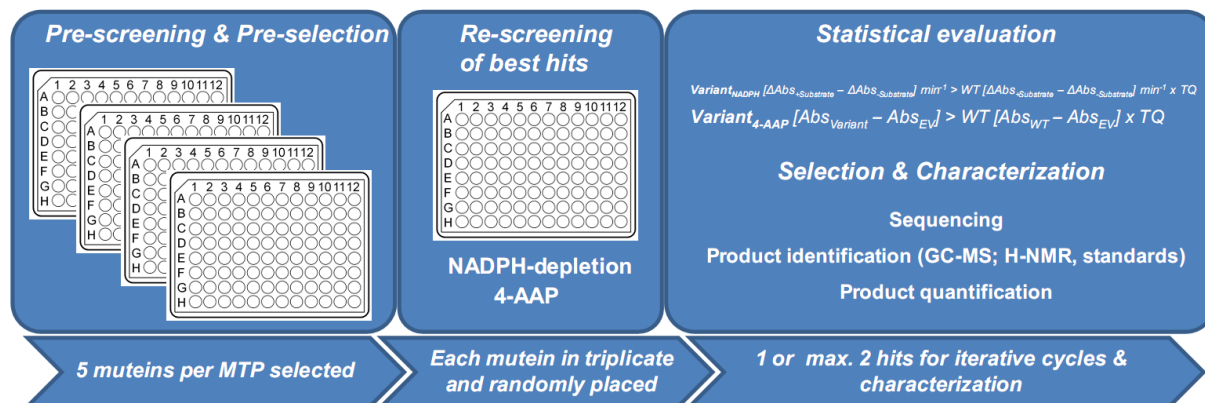


Figure 22. Entire process of screening, selection of variants for rescreening and final characterization of confirmed improved variants.

After evaluation of mean values and relative deviations for each prospective improved variant, one or two variants are selected for sequencing, product identification and catalytic characterization. This strategy was applied in this thesis to identify improved variants, and can also be employed for iterative saturation mutagenesis, allowing a reliable selection and engineering towards a desired property.

2.10.6 Purification and Lyophilization of P450 BM3

Purification of P450 BM3 was performed according to a published protocol employing anion exchange chromatography.[232] Following solutions were prepared, filtered and degassed to perform the purification protocol: 500 mL ddH₂O; 500 mL 20 % (V/V) EtOH; 1000 mL Buffer A (0.1 M Tris-HCl, pH 7.8); 1000 mL Buffer B (1 M NaCl, 0.1 M Tris-HCl, pH 7.8). For purification a glass chromatography column filled with anion exchange matrix (Toyopearl DEAE 650S, Tosoh Bioscience, Stuttgart, Germany) was packed and connected to an Äkta pumping system with UV-detection. The column was packed at 20 mL min⁻¹ flow rate using 20 % EtOH as liquid phase as well as storage solution. During all purification steps a flow rate of 5 mL min⁻¹ was used and all parameters were monitored and controlled by the Äkta pumping system (pressure, UV absorbance at 280 nm, flow rate, conductivity; see Figure 80 in appendix section). Following program was used to prepare the column prior to loading of protein samples: 1. Removal of 20 % EtOH storage solution by flushing column with 10 column volumes (CV) ddH₂O. 2. Equilibration of column with 10 CV Buffer A. A pellet from 100 mL expression culture of *E. coli* BL21 (DE3) lacI^{Q1} cells was entirely resuspended in pre-cooled Buffer A. The resuspended cells were disrupted in a high pressure homogenizer and cell debris as well as solid particles were removed by ultra centrifugation (24500 x g, 30 min, 4°C). Smaller particles were removed by pressing centrifuged cell lysate through a 0.45 µm sterile filter. Filtered cell lysate was loaded onto the equilibrated column using Buffer A inlet. Proteins not binding to the anion exchange matrix were removed by washing the column with 10 CV of Buffer A. The

published 2-step-protocol was performed (14 % Buffer B for 10 CV; 24 % Buffer B for 10 CV) to elute in the first step *E. coli* proteins (14 % Buffer B) and in the second step P450 BM3 (24 % buffer B). Eluted proteins were collected as fractions of 5 mL volume by the Äkta auto sampler. All fractions were stored on ice and analyzed by SDS-PAGE (see Figure 80 in appendix section) as well as CO difference spectra for active P450 content (section 2.10.8). Fractions containing active P450 BM3 and highest purity (>80 %) were combined and desalted with a gel filtration column equilibrated with KPi buffer (50 mM; pH 7.5). Purified and desalted P450 BM3 protein was shock frozen in liquid N₂ and lyophilized under vacuum. Freeze-dried P450 BM3 protein was stored until use at -20°C.

2.10.7 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) for protein analysis was performed according to the protocol of Laemmli et al.[233] A 10 % SDS-PAGE gel was used for analysis of samples containing P450 BM3. A resolving gel (1.9 mL ddH₂O, 1.7 mL 40% acrylamide mix, 1.3 mL 1.5 M Tris (pH 8.8), 0.05 mL 10% SDS, 0.05 mL 10% APS and 0.002 mL TEMED) and stacking gel (1 mL) 0.68 mL ddH₂O, 0.17 mL 40% acrylamide mix, 0.13 mL 1.0 M Tris (pH 6.8), 0.01 mL 10% SDS, 0.01 mL 10% APS and 0.001 mL TEMED) were prepared as described in the original protocol. Sample preparation was done incubating 15 µL of protein samples (lysate or purified protein) in 5 µL SDS-PAGE loading buffer (1 x Roti-Load) in a heat block (15 min, 98°C) for entire denaturation of peptides. Four µL of each samples and a pre-stained reference protein ladder were loaded on the SDS-PAGE gel. Peptides were separated by gel electrophoresis (16 mA; 45 min) employing Tris-glycine buffer (0.3% Tris, 1.5% glycine, 0.1% SDS in ddH₂O, no pH adjustment). Visualization of protein bands was achieved by coomassie-brilliant blue (CBB) staining (0.25 % CBB, 45.5 % methanol, 9.2 % acetic acid, in ddH₂O).[234]

2.10.8 Quantification of P450 Concentration by Carbon Monoxide Difference Spectra

Quantification of active P450 BM3 monooxygenase content was performed according to the published protocol from Omura and Sato employing CO-differential spectra evaluation.[155] Therefore 1.5 mL KPi buffer (50 mM, pH 7.5) containing P450 protein (lysate, purified or lyophilized protein) were incubated with sodium dithionite (a few milligrams) to reduce the heme-iron of P450 BM3.[141, 155] After 1 min incubation the absorbance spectra was measured from 600 to 400 nm in a cuvette spectrophotometer. The sample was saturated with CO for 20 sec and the absorbance spectrum was recorded again from 600 to 400 nm. The gassed sample containing P450 BM3 displays the typical absorbance maxima at 450 nm (Figure 23).[141]

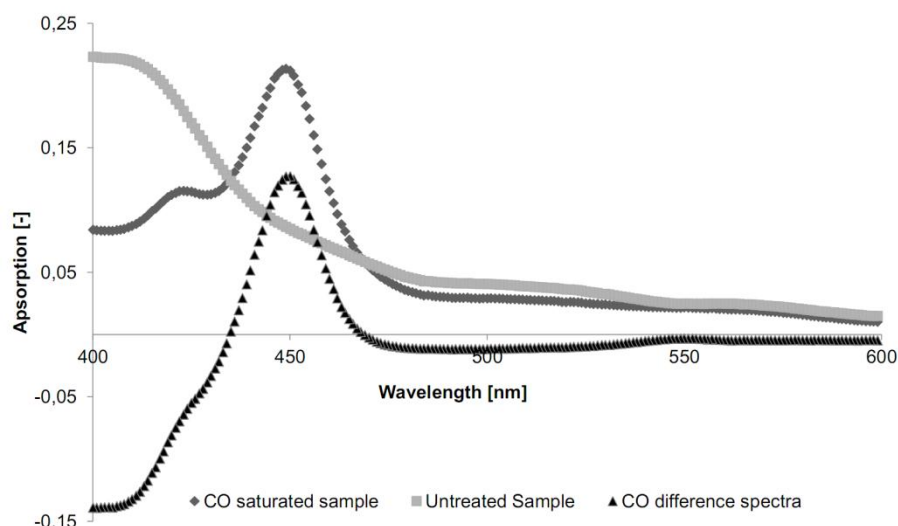


Figure 23. Absorbance spectra recorded before (grey squares) and after gassing (grey diamonds) P450 BM3 containing samples with carbon monoxide (CO). The characteristic P450 peak appears at 450 nm in the CO-gassed sample. Active protein concentration was calculated using the CO-difference spectra (black triangles).

Active P450 monooxygenase content was calculated with following equation:

$$\text{P450 concentration [mM]} = \frac{A_{450\text{nm}} - A_{500\text{nm}}}{\epsilon} \times \text{DF}$$

$A_{450\text{nm}}$ and $A_{500\text{nm}}$ = Absorbance estimated from CO-difference spectra at 450 and 500 nm; ϵ = specific extinction coefficient for P450 BM3 = $91 \text{ mM}^{-1}\text{cm}^{-1}$ [235]; DF = dilution factor

2.10.9 Catalytic Characterization of P450 BM3 Wild-Type and Variants

Key parameters to assess catalytic performance of P450 monooxygenases are the measurement of initial oxidation rates for NAD(P)H and estimation of coupling efficiency for a target substrate. [70, 180, 229] In addition parameters such as substrate affinity, chemoselectivity, regioselectivity, and enantioselectivity play an important role in biocatalytic conversions with P450 monooxygenases. [70, 146]

2.10.9.1 Determination of NADPH Oxidation Rates and Coupling Efficiency

The P450 monooxygenase domain requires constant supply with electrons to perform the hydroxylation reactions that are delivered in P450 BM3 by the associated reductase domain.[166] Consequently the depletion of NADPH can be measured continuously as oxidation of the cofactor at 340 nm wavelength and related to the presence and conversion of a substrate in the active site of the monooxygenase domain.[229] Measurement of NADPH depletion by the P450 BM3 WT and variants was done in quartz glass cuvettes (path length 1 cm) placed in a spectrophotometer. Reactions were assembled in the cuvette with a final volume of 5 mL. The reaction mixture contains following components that were adjusted accordingly depending on the substrate and variant employed for the

measurements: KPi buffer (pH 7.5, 50 mM); DMSO; purified P450 BM3 enzyme; substrate; NADPH (0.2 mM). For every substrate class investigated in this thesis a small experimental section is provided to underline specific alterations in the reaction mixtures. After 5 min incubation the reaction is started by addition of NADPH. Depletion of NADPH is monitored at 340 nm using intervals of 2 sec. A control reaction was performed without substrate to monitor the background depletion/oxidation of NADPH in the reaction system which was subtracted from the activity determined with substrate.

Following equation was used to determine NADPH oxidation activity of P450 BM3:

$$\text{NADPH oxidation rate [min}^{-1}] = \frac{\Delta A_{340} \text{ min}^{-1}}{\epsilon_{\text{NADPH}} \times C_{\text{P450}} \times d}$$

$\Delta A_{340} \text{ min}^{-1}$ = decrease in absorbance per minute at 340 nm wavelength; ϵ_{NADPH} = specific extinction coefficient of NADPH at 340 nm [$6.22 \times 10^{-3} \mu\text{M}^{-1} \text{ cm}^{-1}$]; C_{P450} = concentration of P450 monooxygenase in the reaction [μM]; d = path length [cm].

After full depletion of NADPH the products were extracted by partitioning using two-phase extraction (section 2.11.1). Products were quantified on GC-FID applying calibration curves of the commercially available standards. The coupling efficiency was calculated from the quantity of depleted NADPH and yielded product amount. Following equation was used for calculation of the coupling efficiency:

$$\text{Coupling efficiency [\%]} = \frac{\text{Product } [\mu\text{M}]}{\text{Consumed NADPH } [\mu\text{M}]} \times 100$$

2.10.9.2 Determination of K_M and k_{cat}

In general, the catalytic activity of enzymes can be described as substrate concentration dependent reaction velocity which was determined as k_{cat} and K_M that are key parameters that describe activity and affinity of a catalyst for a specific substrate.[236] In this thesis the K_M and k_{cat} values for WT and two engineered variants were estimated employing *p*-xylene as substrate. Therefore a mixture of 3 mL KPi (50 mM ; pH 7.5) containing *p*-xylene (1 to 30 mM), 1.5 % DMSO (V/V), catalase (1200 U mL⁻¹) and P450 BM3 monooxygenase (WT: 0.5 μM ; M1: 0.133 μM ; M2: 0.036 μM) was pre-incubated for 5 min in an air-closed 25 mL glass-flask shaking on an Eppendorf MixMate at 750 rpm. Substrate conversion by P450 BM3 was induced by supplementing 330 μL NADPH (5 mM). Reaction mixture samples of 180 μL were taken every 10 s and pipetted onto 25 μL quenching buffer (section 2.10.3) pre-filled into a quartz glass MTP. The 4-AAP assay was performed as described before.[69, 228] The products were quantified using calibration curves ranging from 0 to 200 μM for the commercial standard of 2,5-dimethylphenol (2,5-DMP) (see Figure 81 in appendix section). Due to the dimeric structure of P450 BM3 and known co-operative interaction of both monomers during catalysis

measured activities were fitted with the Hill-equation to estimate k_{cat} and K_M for the three monooxygenase variants.[69, 166] Data analysis and fitting of catalytic parameters was done using Origin 7.0 software (OriginLab Corporation, Northampton, MA, USA) and the Hill equation:

$$k_{\text{cat}} = V_{\text{max}} \times \frac{S^n}{k^n + S^n}$$

k_{cat} [sec^{-1}]; V_{max} [U mg^{-1}]; n = Hill coefficient [-]; S = Substrate concentration [mM]; k = affinity constant [mM]

2.10.9.3 Long Term Substrate Conversions with P450 BM3

To access productivity (substrate conversion, product yield and concentrations, TTN) and regioselectivity of the WT and engineered variants long term reactions with a glucose dehydrogenase (GDH) regeneration system were performed using either cell free lysate (section 2.8.5) or purified P450 BM3 protein (section 2.10.6). All reactions were performed in glass tubes sealed with a septum and lid (Figure 34). For efficient catalysis a magnetic stirring bar was used and solutions were stirred at 300 rpm and RT. Reactions were assembled as 2 or 4 mL volumes leaving a minimal head space of 50 % gaseous volume for oxygen supply. Reaction mixtures for long term conversions contained following components: P450 BM3 monooxygenase lysate or purified protein (0.05 to 1 μM); GDH (1.25 U mL^{-1}); 60 mM glucose; substrate; DMSO (0 to 1.5 %); catalase (1200 U mL^{-1}); NADPH (500 μM); KPi buffer (50 mM, pH 7.5). Modifications of the set up in particular for the employed substrate concentrations are mentioned as short experimental part in each section of this thesis. Before initiating the reaction by addition of NADPH, the solution was stirred for 5 min.[160, 161] The reaction was performed for 24 h (300 rpm, RT) prior to product extraction (section 2.11.1) and analysis by GC and GC-MS (section 2.11.2).[68] Product and substrate quantifications were done by applying commercially available product standards and internal standards for calibration. A full list of employed standards, their specific retention times, employed GC-column and GC heating programs can be found as full compilation in the appendix section of this thesis (Table 29).

2.10.9.4 HPLC Measurement of Phenols and Hydroquinones

High performance liquid chromatography (HPLC) was used to access the time and concentration dependent formation of products from P450 BM3 hydroxylation reactions. The advantage of HPLC lies especially in the fast and simple preparation of samples for analysis. Therefore samples of 200 μL were taken at defined time intervals from the reaction mixtures which were assembled as described before (section 2.10.9.3). The samples were quenched in 200 μL acetonitrile (ACN), mixed thoroughly for 2 min and centrifuged for 5 min at 20000 $\times g$ to remove insoluble particles. Clear supernatant containing products were separated on a reverse phase column (flow rate: 1 mL min^{-1} , 100 bar pressure, 40°C) with a mobile phase consisting of 50 % (V/V) KPi (20 mM, pH 3.0) and 50 % ACN (V/V). Absorbance spectra of the target products were identified by analyzing commercial

standards of expected products. An exemplary HPLC spectrum of a conversion with chlorobenzene can be found in the appendix section of this thesis (Figure 82).

2.11 Analytical and Chemical Methods

2.11.1 Two-Phase Solvent Extraction of Reaction Products

Products from P450 BM3 conversions were identified by gas-chromatographic mass analysis (GC-MS) and quantified on GC-FID using calibration curves for identical standards. Therefore the generated products were extracted by partitioning (two-phase extraction) employing methyl *tert*-butyl ether (MTBE), CHCl₃ (chloroform) or ethyl acetate (EtOAc) as extraction solvent. Commonly MTBE was used due to the low toxicity and easy handling during extractions (0.74 g/cm³ density). For quantification of products, always a defined amount of internal standard was supplemented to the extraction solvent (commonly 20 mM) for reliable quantification of generated products. A compilation of all substrates, internal standards and extraction solvents can be found in the appendix section (Table 29). Reaction mixtures and extraction solvent were mixed in a ratio of 2:1 thoroughly on a vortex (5 min, RT) to achieve a high level of reproducibility. Both phases were separated by centrifugation (20000 x g, 5 min, RT) and the organic phase was removed and dried over anhydrous Na₂SO₄ to remove residual water. As last step, the dried and water-free organic phase was centrifuged for entire removal of salts. Liquid supernatants were filled into glass vials containing 200 µL glass inlets.

2.11.2 GC-Analysis

2.11.2.1 GC-FID Measurements

Substrate depletion and product formation was measured using a GC-FID detection system with H₂ as mobile phase. Calibration curves for expected products were recorded using commercially available standards. The standards were supplemented in varying concentrations to 4 mL buffer (KPi, 50 mM, pH 7.5). Solvent extraction via partitioning was performed as described before using an internal standard (section 2.11.1). Extracted calibration samples were injected (1 µL; injector temperature 300°C) and analyzed on a GC-FID (detector temperature 250°C) with optimized heat program for separation of products, substrates and internal standards (Table 29 in appendix section). Calibration curves were calculated by taking the ratio of the area of a product or substrate divided by the area of the internal standard. Data was processed as linear regression for product quantification. All extractions were performed in triplicate to achieve a high confidence level of accuracy and a coefficient of determination of 0.99.

2.11.2.2 GC-MS Measurements

Initial identification of product masses was done on a gas chromatograph connected to a mass spectrometer using helium as carrier gas. Extraction samples were injected (1 µL; injector temperature

300°C) and separated on a Optima-17 MS column isothermally at 100°C and with a solvent cut off of 2 min. Obtained fragmentation patterns were analyzed with the provided compound library.

2.11.3 Visualization of Protein Structures and Docking of Substrates

For visualization of protein structures YASARA[237] software package was used. PDB files of the P450 BM3 WT crystal structure 1BU7 (heme-domain)[238] and 4DQL of the recently crystallized and resolved reductase domain were used.[175] All structure models are available at: <http://www.rcsb.org/pdb/home/home.do>. Data from homology model structures of P450 BM3 and docking data with substrates was kindly provided by Tsvetan Kardashliev (RWTH Aachen University).

3. Results and Discussion

3.1 Directed Evolution of P450 BM3 into a *p*-Xylene Hydroxylase

3.1.1 Abstract

Regioselective hydroxylation of *p*-xylene to 2,5-dimethylphenol (2,5-DMP) is of high interest for the production of thermostable polymers (>500 K) and to generate building blocks for “next-generation” plastics. Direct aromatic hydroxylation with monooxygenases has significant advantages compared to chemical synthesis routes, especially in terms of sustainability and selectivity by performing biocatalysis at RT, in water and by use of molecular oxygen. The adjacent chapter describes the reengineering of P450 BM3 into an efficient monooxygenase for aromatic hydroxylation of *p*-xylene. An iterative approach was used to saturate five key amino acid residues (R47, Y51, F87, A330 and I401), to enable the first aromatic hydroxylation of *p*-xylene by P450 BM3. Mutating three out of five amino acid residues (R47, Y51 and I401) improved activity, coupling efficiency and led to an excellent selectivity (>98 %) during the hydroxylation reaction. Finally, the P450 BM3 variant M2 (R47I, Y51W, I401M) was obtained with the highest reported activity (k_{cat} 1950 min⁻¹; 30-fold increase compared to P450 BM3 WT), showing excellent coupling efficiency (66 %) and selectivity (>98 %) for generation of 2,5-DMP. Residue F87 proofed to be essential for efficient coupling and aromatic hydroxylation of *p*-xylene, indicating aromatic-aromatic interactions of substrate and protein.

3.1.2 Introduction

Regioselective hydroxylation of aromatic rings in aqueous environment using molecular oxygen as sole oxidizing agent is a “dream reaction” in organic chemistry.[69] The demand for phenol and alkylphenols as a feedstock for resins, plastics, and bisphenol A has increased constantly throughout the last century and is of high interest for the production of precursors for pharmaceuticals and flavors.[30, 68, 197] Direct hydroxylation with enzymes is an attractive alternative, reducing reaction and purification steps, as well as waste generation and process energy demands.[63, 75] The engineering of biocatalysts by directed evolution and focused mutagenesis further allows to tailor enzymes to be more cost-effective as well as having higher operational stability which is very often

the limiting step in synthetic applications.[36, 91, 92] Cytochrome P450 monooxygenases (CYPs), a protein super family with more than 5000 cloned members,[139] catalyze oxygenation reactions with molecular oxygen at RT.[63] The direct hydroxylation of aromatic rings is a synthetically attractive reaction that has been reported for P450 monooxygenases. Whitehouse et al. reported the rational driven reengineering of the P450 monooxygenase BM3 (CYP 102A1) towards aromatic hydroxylation of toluene, *m*- and *o*-xylene.[156, 180] In further studies a reaction mechanism was proposed to explain the diverse product pattern after aromatic hydroxylation of *o*- and *m*-xylene.[154, 156] However, *p*-xylene was not investigated as a substrate. The *p*-xylene isomer was reported in a previous NMR study, employing the P450 BM3 WT as catalyst, to be exclusively hydroxylated at the α -position.[156, 209]

Only a few enzymes (toluene 4-monooxygenases and microsomal P450 monooxygenases) were reported to catalyze the aromatic hydroxylation of *p*-xylene to the corresponding phenol.[239-241] The toluene 4-monooxygenase suffers from low activity ($k_{\text{cat}} = 36 \text{ min}^{-1}$), stability (four-protein complex), and selectivity (20% aromatic and 80% benzylic hydroxylation) when *p*-xylene is applied as substrate.[241] Nonetheless, the direct aromatic hydroxylation of *p*-xylene to 2,5-dimethylphenol (2,5-DMP) is of high interest for the production of temperature stable polymers (>500 K), as is 2,5-DMP itself, which can serve as a building block for ‘next-generation’ plastics.[242, 243] Further applications for phenols and in particular 2,5-DMP are also synthetic processes for vitamin E (β -tocotrienol)[244] and pharmaceuticals such as Gemfibrozil[245] that was patented by Pfizer in 1969 as lipid lowering agent (Figure 24).

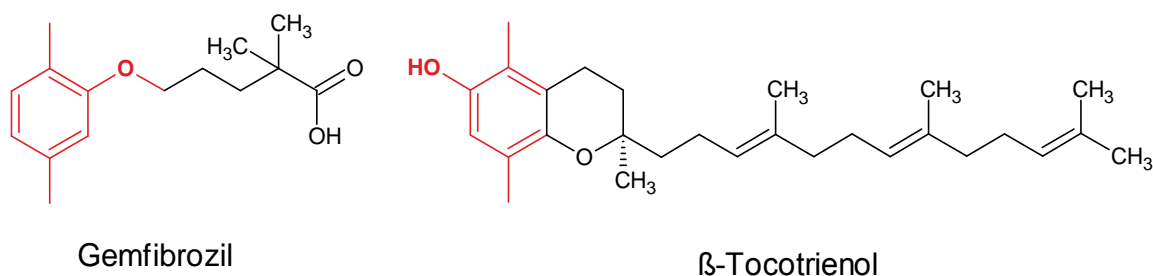
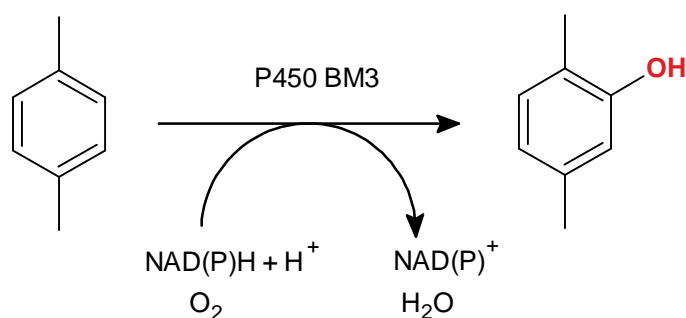


Figure 24. Structure of Gemfibrozil[245] and β -tocotrienol (vitamin E)[244], both containing substructures of 2,5-DMP. The substructure of 2,5-DMP is highlighted in red. Further information with regards to related drugs containing a 2,5-DMP substructure can be found at www.drugbank.ca.

A major objective in this chapter is the development of an efficient P450 BM3 catalyst for aromatic hydroxylation of *p*-xylene. At first, a platform had to be established for screening in MTP format with the target substrate.[228] Subsequently, amino acid residues should be selected for an iterative saturation mutagenesis approach[104] to engineer a catalyst that is capable of performing the aromatic hydroxylation of *p*-xylene. Since several key residues are already known influencing selectivity and activity in P450 BM3 for catalysis of benzene substrates[180, 206, 246], an iterative saturation mutagenesis approach was chosen to make aromatic hydroxylation of *p*-xylene possible (Scheme 6).



Scheme 6. Conversion of *p*-xylene by P450 BM3 to produce 2,5-DMP.

Obtained variants should be purified, kinetically characterized and discussed with respect to their ability to hydroxylate *p*-xylene.

3.1.3 Experimental

Large parts of this chapter were published in *ChemCatChem* (2012).[69] Lukas Guddat assisted during catalytic characterization of generated P450 BM3 variants within the frame of his master thesis project at the Institute of Biotechnology (RWTH Aachen University).

3.1.3.1 Focused Mutant Library Generation

Altogether five amino acid residues were chosen for iterative NNK saturation: R47, Y51, F87, A330 and I401.[68-70, 156, 180, 206] Residues R47 and Y51 were simultaneously saturated due to known cooperativity between the two amino acids at the entrance of the substrate channel ('gaters').[70] The other three residues (F87, A330 and I401) were saturated as single positions. Design of mutagenic oligonucleotides (section 2.9.3), the 2-step-PCR protocol (section 2.9.8) and preparation of MTPs for screening (section 2.8.7) was done as described in the material and methods section of this thesis. The list of used oligonucleotides for saturation mutagenesis of P450 BM3 can be found in Table 27 in the appendix section of this thesis. *E. coli* Gold (DE3) IaI^{Q1} cells and the vector pALXtreme-1a(+)[213] were used during all experiments in this chapter (section 2.6). The required oversampling of variants was achieved by generating a library size of 180 clones for single site-saturated positions (F87, A330 and I401) and 2000 clones for double site-saturation (R47, Y51), achieving a confidence level of 99 % for all possible codons.[127]

3.1.3.2 Screening of Focused Mutant Libraries in Microtiter Plates

The clear supernatant was used for assaying monooxygenase activity by measuring NADPH depletion at 340 nm in 96-well quartz glass MTPs (section 2.10.2). A reaction mixture in MTP format contained 1.5 % DMSO, 200 mM *p*-xylene, 10 to 50 μl cell lysate and 0.16 mM NADPH in a final volume of 250 μl KPi buffer (50 mM, pH 7.5). The amount of lysate was adjusted for each round of screening in a way that linear depletion of NADPH could be monitored reliable for at least 2 min. For quantification of phenolic products the selective 4-AAP assay was used (section 2.10.3).[228] Mutants

for subsequent rounds of saturation were selected based upon best performance for NADPH depletion and highest formation of 2,5-DMP (section 2.10.4). Plasmids of improved variants of P450 BM3 were isolated, sequenced and analyzed for amino acid substitutions as described in material and methods section.

3.1.3.3 Kinetic Characterization of Improved P450 BM3 Variants

Sequenced variants of P450 BM3 were expressed, purified and stored until kinetic characterization as lyophilized powder (section 2.10.6). Catalytic characterization was done as described in the material and methods section of this thesis (section 2.10.9.2). Fitting and evaluation of catalytic data (K_M , k_{cat}) was done using Origin 7.0 software and the implemented Hill-equation (section 2.10.9.2) with Hill-coefficients of 2.96 (M2), 2.81 (M1) and 2.37 (WT) to obtain a root mean square value of 99 %. Coupling efficiency of the generated variants was determined by applying the same set up assay mixture concentrations that were used for kinetic characterization of P450 BM3 variants (section 2.10.9.1).[156] A final substrate concentration of 20 mM *p*-xylene was used and after 5 min pre-incubation in 5 mL glass cuvettes (1 cm width) the reaction was induced by addition of 166 μ L NADPH (5 mM). Depletion of NADPH was measured in a Varian Cary 50 spectrophotometer at 340 nm. Product quantification was achieved via 4-AAP assay (section 2.10.3). In addition, the formation of H_2O_2 was determined as reported by a standard protocol using the horse-reddish peroxidase assay (HRP).[247] All measurements for kinetic characterization as well as determination of coupling efficiency were done in triplicate.

3.1.3.4 Product Detection via Gas-Chromatography

Qualitative product detection using toluene and all three xylene isomers as substrates was performed via GC-FID measurement. Therefore, 3 mL reaction mixtures were prepared as described above in a 25 mL shake flask. After 30 min incubation at 750 rpm on an Eppendorf MixMate additional 330 μ L of 10 mM NADPH were supplemented followed by 30 min shaking at 750 rpm. Two-phase extraction of products was realized using chloroform as solvent (section 2.11.1). Extracted products were injected into a GC-2010 Plus gas chromatograph and separated isothermally at 120°C on an Optima-17 MS column. As standards all expected phenols, dimethylphenols (DMP) as well as benzylalcohols were injected applying the identical GC program.

3.1.4 Results and Discussion

Five key amino acid residues of P450 BM3 were selected and iteratively saturated: R47 Y51, F87, A 330, and I401. All residues were reported before to influence either substrate selectivity, activity or both using benzene substrates for conversions (Figure 25).[68-70, 156, 180, 206]

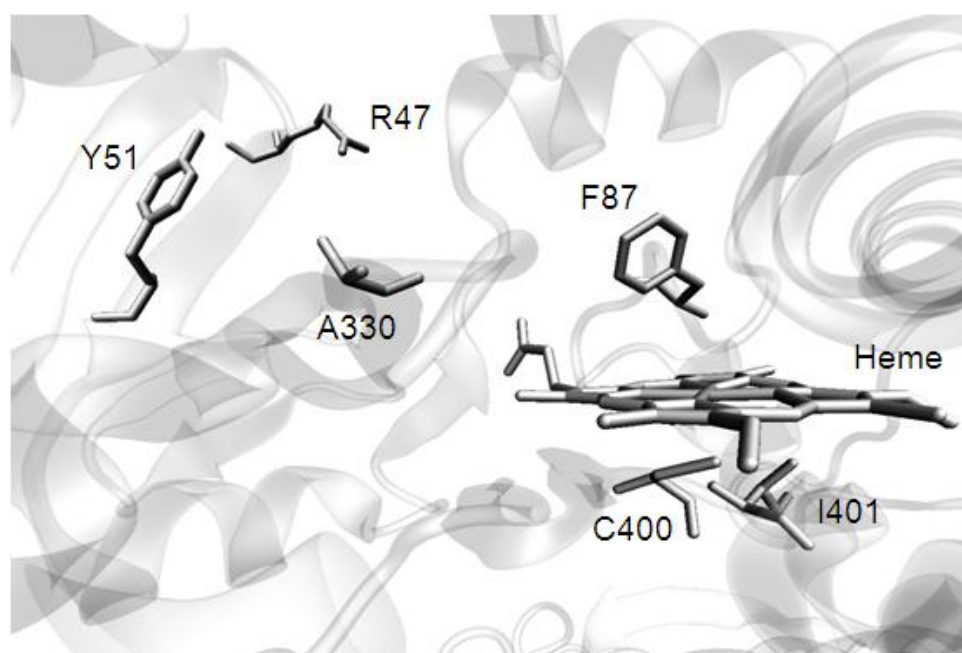


Figure 25. P450 BM3 active site snapshot. Selected amino acid residues targeted for mutagenesis (R47, Y51, F87, A330, and I401) and the heme iron cluster, which is bound through C400, are depicted in a stick representation. Amino acid residues R47 and Y51 lie at the entrance to the substrate channel and are key residues in controlling substrate access to the binding pocket ('gaters').^[70] I401 is in close proximity to C400 and was reported to have a significant effect on the conformation of the iron cluster within the catalytic center. Positions F87 and A330 were the only residues targeted for saturation mutagenesis within the binding pocket. Both amino acid residues lie directly above the heme and play an important role in substrate orientation and binding as well as controlling coupling efficiency. The Figure was taken from Dennig et al.^[69]

Basis for successful engineering of biocatalysts are high quality mutant libraries as well as a reliable screening assays for identification of improved protein variants.^[36, 37, 98] The 4-AAP assay^[228] in combination with *p*-xylene as substrate offers the possibility for direct identification of variants with improved product formation, circumventing limitations of the broadly applied NADPH depletion assay which is commonly used for P450 monooxygenases.^[70, 229] The NADPH-depletion assay (18 % SD) and 4-AAP assay (509 nm: 11.6 % SD; 600 nm: 21.1 % SD) were optimized and fulfill all the requirements for reliable detection of P450 BM3 activity in 96 well MTP format (Figure 26).^[129, 228]

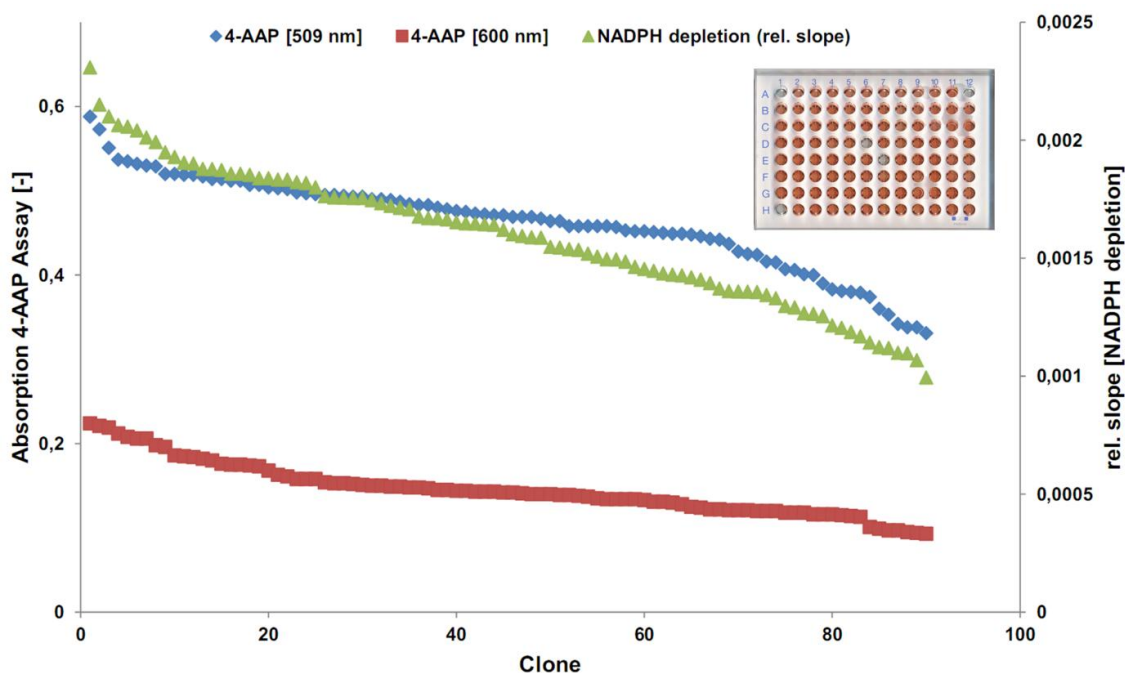


Figure 26. Validation of the MTP screening system using NADPH depletion and 4-AAP assay in 96-well format with P450 BM3 and *p*-xylene (200 mM) as substrate. Standard deviation of NADPH depletion assay was estimated to be 18 % (green triangles), whereas standard deviation for 4-AAP was calculated to be 11.6 % at 509 nm (blue diamonds) and 21.1 % at 600 nm (red squares) respectively. 2,5-DMP and 4-AAP form a red colored complex (section 2.10.3) which is shown in the exemplary MTP. Wells with red color contain active P450 BM3, whereas non-colored wells represent the EV control.

A slightly higher deviation in the NADPH depletion assay is acceptable, since the *E. coli* host has several NADPH consuming enzymes which can be measured as background activity.[191, 228, 229] The possibility to combine the two assays in the same screening MTP enables selection of more active (increased NADPH oxidation) as well as more productive variants (increased absorbance with 4-AAP) within the same MTP. Although two absorption wavelengths (509 and 600 nm) can be used for linear detection of 2,5-DMP with the 4-AAP assay[228], only absorption at 509 nm was used for data evaluation due to the lower standard deviation (Figure 26). Since P450 BM3 was reported to hydroxylate *p*-xylene only in α -position[209, 246] the first target in this project was to improve substrate inlet into the active site which can be monitored by an increased NADPH depletion rate. [69, 229] Therefore, in the first round of P450 BM3 mutagenesis, positions R47 and Y51 were saturated simultaneously because both amino acids jointly control substrate access to the active site (Figures 9 and 25).[70] After screening 2000 clones (94 % coverage)[127] and re-screening the best 30 variants in triplicate with the NADPH depletion assay and the 4-AAP assay, the variant M1 (R47S, Y51W) was obtained with a more than seven-fold increased product formation ($k_{\text{cat}} = 500 \text{ min}^{-1}$) compared to the P 450 BM 3 WT (Table 6).

Table 6. Catalytic parameters for hydroxylation of *p*-xylene by P450 BM3 (WT, M1, and M2). [a] Units [$\text{nmol}_{2,5\text{-dimethylphenol}} \text{nmol}^{-1}_{\text{P450}} \text{min}^{-1}$]. [b] Units [$\text{mM}^{-1} \text{min}^{-1}$]. [c] Coupling efficiency. Data from this table was published in Dennig et al.[69]

Variant	$k_{\text{cat}}^{\text{[a]}}$	$K_{\text{M}} \text{ [mM]}$	$k_{\text{eff}}^{\text{[b]}}$	$C^{\text{[c]}} \text{ [%]}$	2,5-DMP [%]	4-MBA [%]
WT	68.0 ± 3	7.9	8.6	45	>98	1
M1	500.6 ± 35	7.1	70.4	54	>98	1
M2	1953.1 ± 37	6.1	320.2	66	>98	1

On average 5 to 10 variants per screened plate showed improved NADPH depletion and formation of phenolic products, which is shown exemplary for MTP 19 (Figure 27). The high amount of improved clones and only a few inactive clones (< 50 % per plate) underlines that simultaneous saturation of both residues for engineering campaigns of P450 BM3 allows generation of several improved variants without modifying the active site of the enzyme.

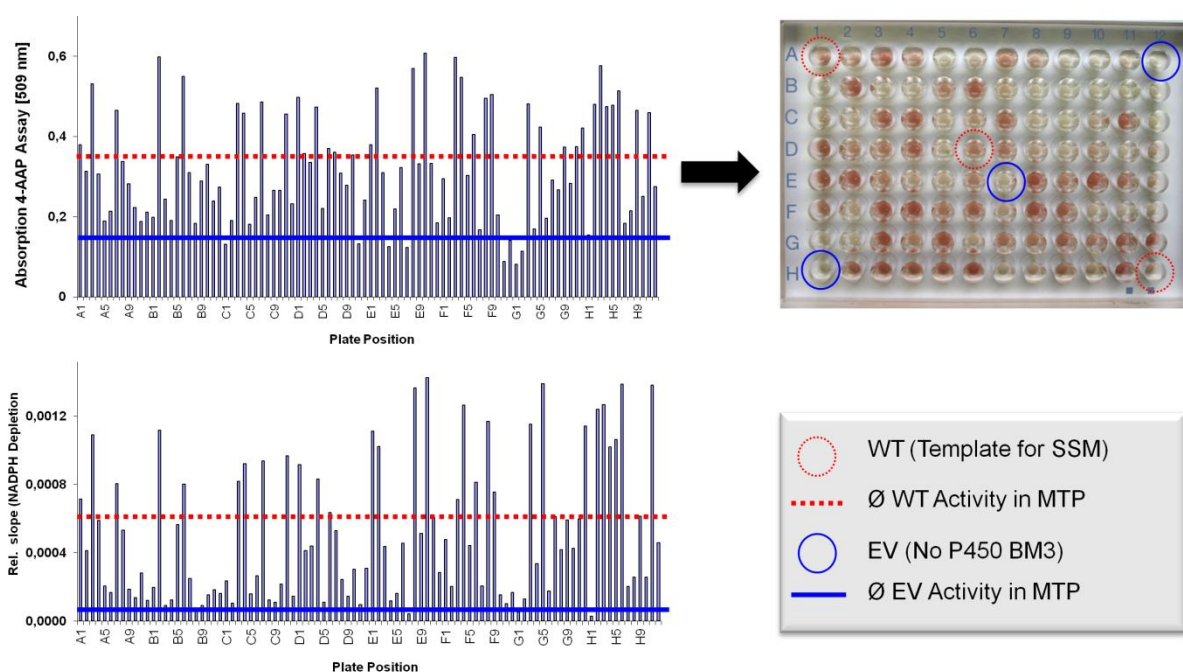


Figure 27. Exemplary screening of plate 19 from first round of saturation mutagenesis (SSM on R47 and Y51) is shown using the 4-AAP (upper left diagram; upper right picture of MTP incubated for 30 min) and NADPH-depletion assay (left bottom diagram). The red dotted line indicates average activity and productivity of the WT protein. The blue line represents average activity of the empty vector control (no P450). Blue and red dotted circles in the MTP plate picture (upper right) indicate the position of the respective controls. Dark red wells indicate formation of phenolic product from 4-AAP assay.

Unexpectedly, the P450 BM3 WT displayed low activity for aromatic hydroxylation of *p*-xylene ($k_{\text{cat}} = 68 \text{ min}^{-1}$; 45 % coupling efficiency), producing 2,5-DMP with high selectivity (>98 %) (Figure 28), which contradicts a previous report.[209]

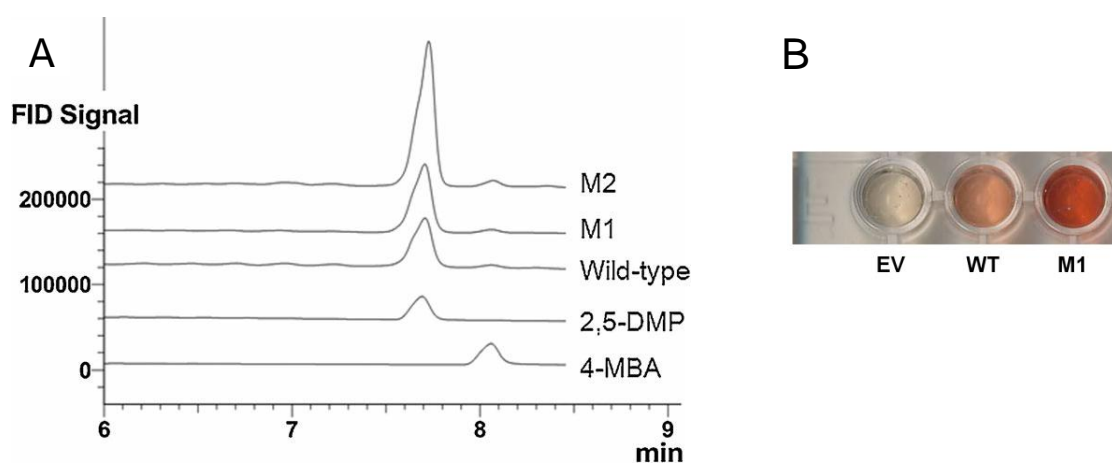


Figure 28. A: The gas chromatograms show the separation of conversion products of *p*-xylene by P450 BM3 WT, M1 (R47S, Y51W) and M2 (R47S, Y51W, I401M). A flame ionization detector (FID) was used to detect the compounds after two-phase extraction. 2,5-Dimethylphenol (2,5-DMP) and 4-methylbenzylalcohol (4-MBA) were employed as commercial standards with an identical GC program. Figure 28A was taken from Dennig et al.[69] **B:** 4-AAP assay with CO normalized proteins. No P450 BM3 (EV; only cell lysate), WT P450 BM3 (WT) and variant M1 (M1).[228]

This observation can be attributed to the significant higher substrate concentrations (1 to 30 mM *p*-xylene; Figure 29) employed in the reaction setup and screening (200 mM *p*-xylene) (3.1.3.2).[209] The recently proposed mechanism for *o*- and *m*-xylene hydroxylation, however, does not exclude aromatic hydroxylation of *p*-xylene.[154, 246] Besides of an increase in productivity (7 fold), variant M1 showed a remarkable increase in the coupling efficiency by 20 % (total coupling 54 %). This was accompanied by a slightly decreased K_M (from 7.9 to 7.1 mM) (Table 6). The bulky tryptophan residue at position 51 is thought to promote aromatic–aromatic interactions with *p*-xylene, which could accelerate substrate uptake.[69] Furthermore, the serine at position 47 widens access to the catalytic center and the exchange of arginine to serine reduces polarity at the entrance to the substrate channel, hindering the entrance of the hydrophobic *p*-xylene (Figure 25).[70, 180, 248] After characterization of variant M1, the position I401 was saturated and screened for improved product formation using the so far best variant M1 as the template. Position I401 is located below the heme iron cluster (Figure 25) and was reported to have a significant influence on the conformation of the heme iron within the catalytic center (I401P).[156, 180] However, I401 was not saturated before and is located in a rather sensitive area since C400 is responsible for correct binding and coordination of the heme iron.[151, 180] After screening 180 clones, variant M2 (R47S, Y51W, I401M)[69] was extracted from the mutant library, which showed after purification and normalization an activity that was 5 times higher than that of M1 (29 fold improved compared to WT; $k_{cat} = 1953 \text{ min}^{-1}$) (Table 6 and Figure 29).[69]

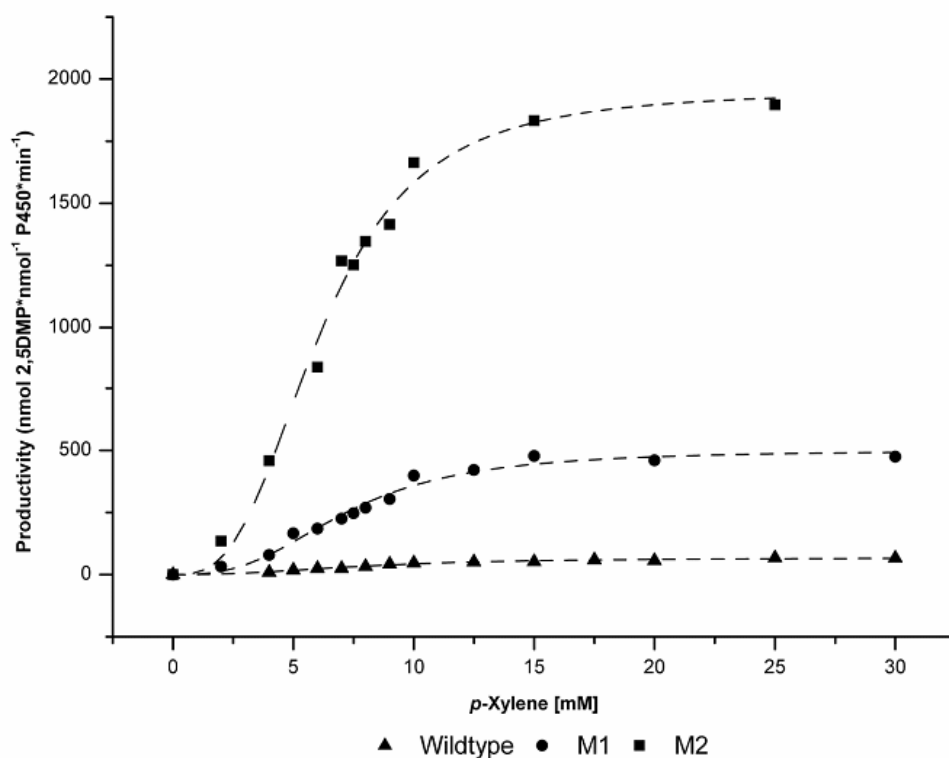


Figure 29. Kinetic characterization of WT P450 BM3 (▲), variant M1 (R47S, Y51W) (●) and variant M2 (R47S, Y51W, I401M) (■) for determination of k_{cat} and K_{M} with *p*-xylene as substrate. All measurements were performed in triplicate with less than 10 % standard deviation. Fitting of kinetic parameters was achieved using the Hill equation (see section 2.10.9.2) with Hill coefficients of 2.96 (M2), 2.81 (M1) and 2.37 (WT) to obtain a root mean square value of 99 %. The Figure was taken from Dennig et al.[69]

The mode of interaction and the substitution to methionine suggested that the heme conformation might be altered, resulting in changes in electron density and oxidizing potential of the iron, which could lead to a more active P450.[180] In addition, a 46 % increase in coupling efficiency (66 % in total) was observed, whereas the K_{M} was reduced again by 15 % to 6.1 mM. The high k_{cat} value makes the M2 variant a very efficient monooxygenase catalyst ($k_{\text{eff}} = 320.2$; Table 6) although no active site residues had to be mutated. Positions F87 and A330 were also saturated with NNK degeneracy using M2 as the template, but no further improvements in catalytic performance towards aromatic hydroxylation of *p*-xylene could be obtained. Figure 30 summarizes the evolution pathway of P450 BM3 towards aromatic hydroxylation of *p*-xylene and the potential influence of amino acids and mutations on protein activity.

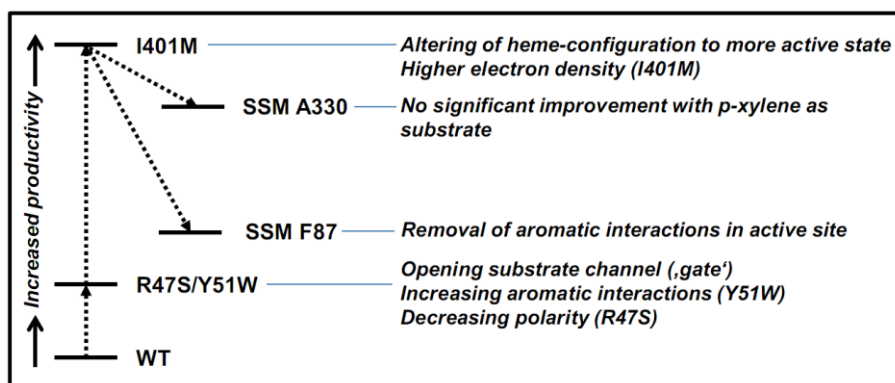


Figure 30. ISM strategy for generation of variants for *p*-xylene hydroxylation. Effects of amino acid exchanges on catalytic performance and their deduced reason within the protein structure are given. Saturation of R47, Y51 and I401 lead to significant improved variants, whereas saturation of A330 and F87 lead to no improvement in hydroxylation of *p*-xylene.

Out of 180 screened variants at position F87, only four clones ($\approx 2\%$) retained activity of M2 (Figure 31) and all four clones contained a codon for phenylalanine indicating the importance of this residue for aromatic hydroxylation.[151] F87 is located directly above the heme center (Figure 25) and controls substrate orientation and binding through hydrophobic or aromatic interactions.[68-70] Results indicate that phenylalanine is indispensable for aromatic hydroxylation of *p*-xylene and a key residue in substrate orientation most likely through π - π binding.[69] Site-directed mutagenesis of position A330 to proline was reported before and is known to improve coupling efficiency by up to 35 % for aromatic hydroxylation of toluene.[180, 206, 246] SSM on this position using M2 as template for mutagenesis did not reveal any further improvement in the coupling efficiency during aromatic hydroxylation of *p*-xylene (Figure 31).

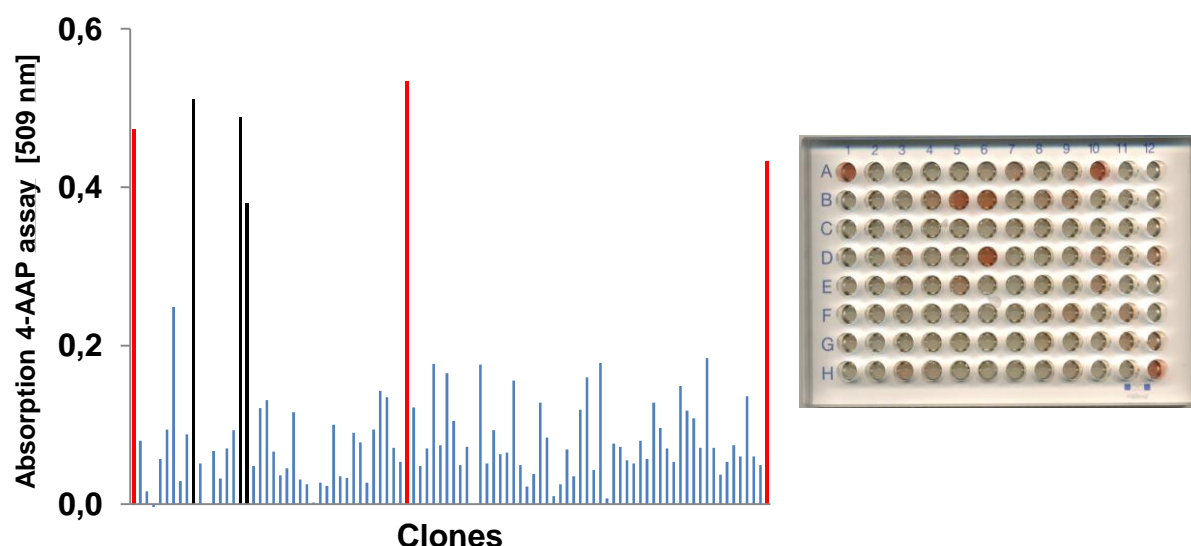


Figure 31. Screening of SSM library at position F87 with the 4-AAP assay and *p*-xylene as substrate. Variant M2 was used as template for mutagenesis. The diagram shows absorption values recorded for 90 variants, variant M2 (wells A1, D6 and H12) and EV (wells, A12, E7 and H1). Variant M2 is also highlighted as red bars in the diagram. Black bars highlight variants that were sequenced and identified having phenylalanine as amino acid at position 87 (B5, B6 and A10). Sequenced variants containing similar productivity for 2,5-DMP as variant M2 displayed distinct color formation in the 4-AAP assay as shown in the MTP picture on the right side. All other amino acid substitutions displayed no comparable formation of 2,5-DMP.

Still the question of why *p*-xylene was reported to be hydroxylated mainly at the α -position[209] and not at the aromatic ring remains. The aromatic hydroxylation of *p*-xylene is feasible because methyl groups direct hydroxylation by P450 BM3 to the aromatic rings of *m*- and *o*-xylene.[156] Unlike Rock et al.[209] significantly higher substrate concentrations were applied ranging from 1 to 30 mM (vs. a maximum of 0.5 mM) to kinetically characterize the obtained P450 BM3 variants and the WT (Figure 29).[69] An accumulation of hydrophobic *p*-xylene in the substrate binding site of P450 BM3 might reduce the number of water molecules present and therefore promote aromatic hydroxylation.[249] As there are no residues exchanged within the binding site of any of the characterized variants, the high coupling (66%) and high activity (1950 min^{-1}) are most likely attributable to the chemical structure of *p*-xylene, which has four identical carbon atoms in the aromatic ring and two methyl groups that assist in orientation for the hydroxylation reaction. Consequently, the probability of aromatic hydroxylation for *p*-xylene might be higher than that of its two isomers. Lower coupling efficiencies for P450 BM3 with the substrates *o*-xylene and *m*-xylene (12 and 29%, respectively)[246] might also be attributed to fewer carbon atoms available in the *o*-position. In addition to *p*-xylene, also toluene, *m*-xylene and *o*-xylene were employed as substrates with P450 BM3 variant M2. GC analysis (Figure 32) revealed similar hydroxylation patterns to those reported by Whitehouse et al [246] indicating that the effect on selectivity of the introduced mutations is minor.

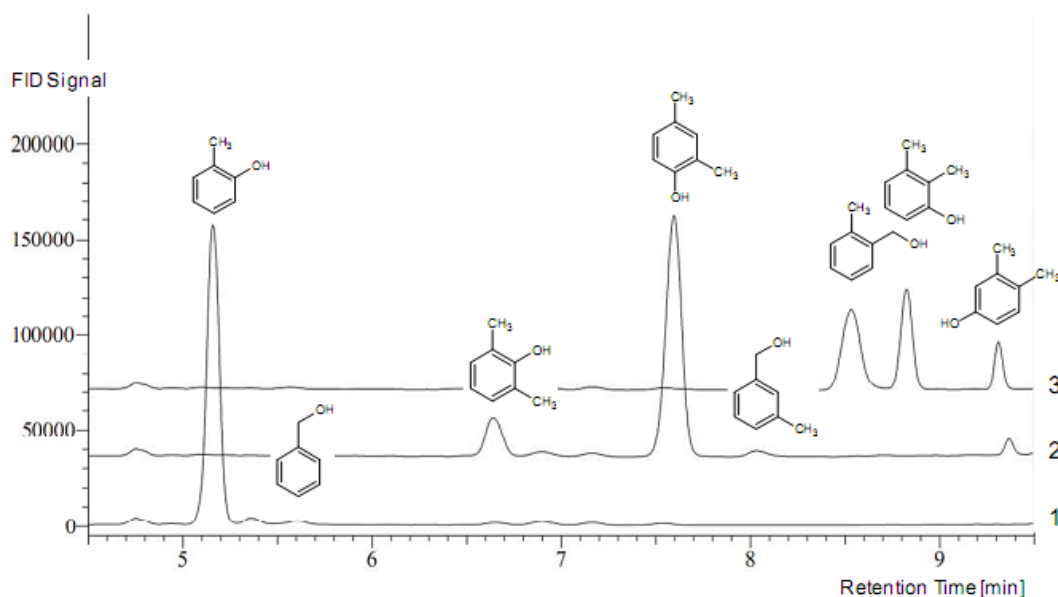


Figure 32. Gas chromatograms show the hydroxylation of toluene (1), *m*-xylene (2) and *o*-xylene (3) by P450 BM3 variant M2 (R47S, Y51W, I401M). All potential products (from aromatic and α -hydroxylation) were injected and separated as standards with the identical GC-FID program. The Figure was taken from supplementary information from Dennig et al.[69]

The results support the hypothesis that aromatic hydroxylation of xylenes by P450 BM3 takes preferentially place in *o*-position to a CH_3 substituent (Figure 32). It is likely that the methyl groups in xylenes and the catalytic key residue F87 govern jointly the regioselective hydroxylation of the aromatic ring system in xylenes and toluene.[69, 156]

3.1.5 Summary and Conclusion

In summary, this is the first report for the efficient hydroxylation of *p*-xylene to 2,5-DMP by a P450 BM3 monooxygenase. Five selected amino acid residues were iteratively saturated and investigated for improving aromatic hydroxylation of *p*-xylene. Three out of five amino acid residues (R47, Y51 and I401) significantly improved the activity of P450 BM3, whereas saturation of F87 leads to a loss in activity. Residue F87 proved to be essential for efficient coupling and aromatic hydroxylation of *p*-xylene whereas no other amino acid could retain the catalytic performance. The saturation of residue A330 did not reveal an improvement in catalytic activity. Finally, a reengineered P450 BM3 variant (M2: R47S, Y51W, I401M) was obtained with the highest reported activity ($k_{\text{cat}} = 1953 \text{ min}^{-1}$; a 30-fold increase compared P450 BM3 WT), excellent coupling efficiency (66 %) and selectivity (>98 %) for 2,5-DMP production. The 4-AAP assay was successfully used to identify and characterize new P450 BM3 variants. Based on the obtained catalytic data a new hypothesis on aromatic-aromatic interactions of *p*-xylene and P450 BM3 could be generated that would help to investigate protein substrate interactions to a further extend.

3.2 Substrate Screening with the Engineered Variant M2 and Investigation of the Influence of DMSO on the Activity of P450 BM3

3.2.1 Abstract

In this chapter the engineered variant M2 (R47S, Y51W, I401M) is applied for conversion of 28 selected benzene substrates. Screening for new potential substrates was performed using the regioselective 4-AAP assay to detect phenols prior to identification of products by GC-FID and GC-MS. Thirteen substrates were hydroxylated by variant M2 either on the aromatic ring or aliphatic substituents (terminal or sub-terminal positions). Aromatic hydroxylation was possible on 9 out of 13 converted substrates. In addition variant M2 performs epoxidation of styrene and *p*-methylstyrene as well as α -hydroxylation of ethylbenzene, enabling synthesis of valuable building blocks. Besides the substrate screening the influence of the co-solvents DMSO and isopropanol on catalytic performance of P450 BM3 was investigated. Increasing the concentration of co-solvents from 0 to 30 % revealed that productivity for *o*-cresol decreases significantly with increasing the concentrations of co-solvents. The best catalytic performance was achieved in a co-solvent-free reaction that can save cost for chemicals; simplify purification and resulting in a 'greener' synthetic approach for the production of phenols.

3.2.2 Introduction

Engineering of biocatalysts by means of directed evolution became a standard approach to tailor enzymes to a specific demand.[33-37] After finalizing the characterization process of an improved variant the question appears in which direction to proceed with the new catalyst? If the catalysts still has low performance towards the desired property, engineering with an altered strategy would be suitable, e.g. using directed evolution instead of semi-rational engineering, to optimize the enzyme to a further extend.[36, 108] On the other hand, if the performance is satisfactory, widening of the substrate spectra can lead to new insights on functionality of the protein as well as exploring new routes in biocatalytic synthesis. The engineered P450 BM3 variant M2, that was evolved and characterized in chapter 3.1, performs aromatic hydroxylation of *p*-xylene with appealing productivity ($k_{\text{cat}} = 1953 \text{ min}^{-1}$) and good coupling efficiency (66 %).[69] In addition the new P450 variant could hydroxylate also the *m*- and *o*-xylene isomers as well as toluene (Figure 32), indicating that the new variant is a promising candidate for conversion of more substrates. Since the high selectivity and activity of the monooxygenase catalyst was attributed to π - π interactions of phenylalanine and the substrate[68, 69], the screening for additional substrates could lead to new products as well as to provide a more detailed insight into substrate-P450 BM3 interactions.

Another important consideration applying water insoluble substrates such as xylenes or trimethylbenzenes in biocatalysis is the addition of co-solvents to the reaction system.[182] Only a few enzymes are reported that can operate efficiently in neat organic substrates.[51, 52] In a few cases water even impairs or directs reactions as known for esterification by lipases.[250-252] Common co-

solvents used in biocatalysis are for instance DMF, DMSO, THF or ACN, which should increase the solubility of hydrophobic substances in the polar aqueous bulk. Especially when the K_M of an enzyme is higher than the solubility of the substrate in the aqueous solution the application of co-solvents is helpful.[253] However, co-solvents can strongly interfere with the later purification of products or activity and stability of a catalyst.[182] Alternative co-solvents can be ionic liquids or deep-eutectic solvents that are able to increase solubility of substrates or replace bulk water[254], but applications in industrial processes with enzymes are not reported so far and interactions of enzymes and co-solvents are not well understood. Nonetheless, the best strategy in a biocatalytic approach would be to avoid any co-solvents in the reaction and to operate in a pure bulk environment, either neat substrates or water.[51] P450 BM3 is a comparably sensitive catalyst with regards to activity in widely used co-solvents (20 % DMSO, 8 % ACN, 12.5 % EtOH or 3 % THF).[182] Although the relative activity of P450 BM3 towards co-solvents could be improved by directed evolution,[182] the generated variants have significantly decreased activity beyond DMSO concentrations of 25 %. Elevated DMSO concentrations were reported to prevent efficient substrate binding in the active site and effects in the peptide linker region which is responsible for dimerization and efficient electron transfer.[182, 255, 256]

During the characterization of P450 BM3 towards aromatic hydroxylation of *p*-xylene a comparably low concentration of DMSO (1.5 %) was sufficient to achieve good hydroxylation activities (3.1.3.2).[69] One aim in this part of the thesis is therefore the investigation of activity and productivity of P450 BM3 variant M2 under varying co-solvent concentrations (DMSO and isopropanol). Furthermore the engineered variant M2 is investigated for the potential to hydroxylate a selection of new benzene substrates to explore the synthetic potential of the monooxygenase catalyst to a further extend. In addition to the alkyl substituents on the benzene ring also larger substituents such as sulfur-, amine- and halogen-substituents should be tested, since these groups are important for further synthetical approaches.

3.2.3 Experimental

3.2.3.1 Selection and Conversion of Substrates

Twenty-eight benzene substrates were selected, having one to five substituents with varying functionalities and sterical demands. Selected substrates are shown in Figure 33 whereas the full list of all substrates is summarized in Table 7.

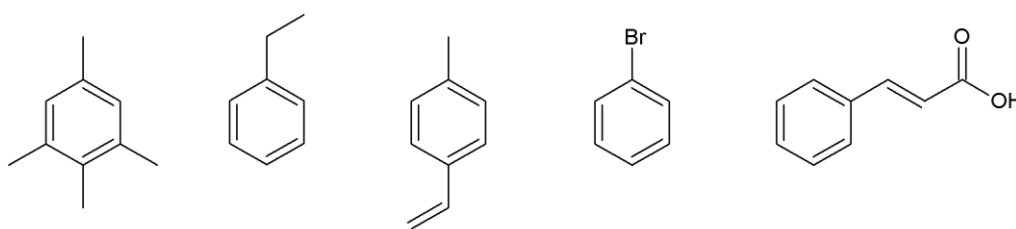


Figure 33. Five selected substrates used for hydroxylation by variant M2.[69] From left to right: 1,2,3,5-tetramethylbenzene; ethylbenzene; 4-methylstyrene; bromobenzene; *trans*-cinnamic acid.

All substrates were obtained in purities of 98 % or higher. An exception are 1,2,3,5-tetramethylbenzene (70%) and 1,2,3,4-tetramethylbenzene (90 %) which were not available in higher purities. Nonetheless, both substrates were applied in the biocatalytic reactions.

Table 7. Substrate library used for conversions with P450 BM3 variant M2

Substrate No.	Name	Substrate No.	Name
1	1,2,4,5-Tetramethylbenzene	15	4-Methylanisole
2	Mesitylene	16	Hydrocinnamic acid
3	1,2,3-Trimethylbenzene	17	<i>trans</i> -Cinnamic acid
4	Pentamethylbenzene	18	1,2,3,5-Tetramethylbenzene
5	Chlorobenzene	19	1,2,3,4-Tetramethylbenzene
6	4-Chlorotoluene	20	Cumene
7	4-Methylbenzenethiol	21	1,4-Diethylbenzene
8	Thiophenol	22	Styrene
9	Bromobenzene	23	4-Methylstyrene
10	4-Bromotoluene	24	Pyrazine
11	<i>p</i> -Toluidine	25	2-Methylpyrazine
12	Aniline	26	2,3-Dimethylpyrazine
13	Ethylbenzene	27	4-Fluorotoluene
14	Isobutylbenzene	28	<i>p</i> -Nitrotoluene

All conversions were performed in closed glass vessels, at RT and 300 rpm stirring (Figure 34). P450 BM3 variant M2 was used as cell free lysate (section 2.8.5). In a total volume of 3 mL KPi (50 mM, pH 7.5) following components were supplemented: P450 BM3 variant M2 (2 μ M), DMSO (1.5 % V/V), substrate (50 mM), catalase (1200 U mL⁻¹) and NADPH (0.5 mM). The reaction mixture was stirred for 2 h at RT prior to product detection with the 4-AAP assay (section 2.10.3)[228] and

extraction with CHCl_3 (section 2.11.1). As a positive control the conversion of *p*-xylene was performed. Furthermore, two negative controls were prepared: 1. Cell free lysate from *E. coli* without P450 monooxygenase (EV) and *p*-xylene as substrate; 2. Reactions containing variant M2 and all components including substrate but leaving NADPH out of the reaction.



Figure 34. Reaction vessels standing on a magnetic stirring board and filled with the reaction mixture for hydroxylation of selected benzene substrates. The photo was taken after 2 h reaction prior to extraction of products for GC-FID and GC-MS analysis.

Extracted compounds were separated isothermally (100°C for 30 min; FS-Supreme-5ms column) for detection on a GC-FID chromatograph. Chromatograms of conversions were compared with the chromatogram of an EV extraction control. New product peaks were analyzed on GC-MS to identify product masses and confirm the formation of a P450-catalyzed product.

3.2.3.2 Investigating the Influence of DMSO and Isopropanol on Activity of P450 BM3

Toluene was selected as substrate for all reactions. A calibration curve for *o*-cresol was established in MTP format using the 4-AAP assay for detection of phenolic products.[69, 228] P450 BM3 activity was measured in quartz glass MTP as described in the material and methods part of this thesis employing the NADPH depletion assay (section 2.10.2). 50 μL of cell free lysate were therefore incubated for 5 min in 150 μL KPi containing 0 to 30 % of DMSO or isopropanol (IPA) respectively, and 50 mM substrate. The hydroxylation reaction was initiated by supplementing 50 μL NADPH (final concentration 0.33 mM). Depletion of NADPH was monitored in a MTP reader until NADPH was fully depleted in the first well and the reaction was quenched immediately in the whole plate. Quantification of the reaction products was achieved by applying a calibration curve for *o*-cresol (similar to Figure 81 in appendix section).

3.2.4 Results

Parts of this chapter were published online as supplementary information in *Angewandte Chemie* (2013).[68] Moritz Gauer assisted during catalytic characterization of P450 BM3 variant M2 at varying cosolvent concentrations within the frame of his research internship project at the Institute of Biotechnology (RWTH Aachen University).

3.2.4.1 Substrate Screening with Variant M2

Initially, all reactions in the substrate screening were investigated with the 4-AAP assay to identify produced phenols. Figure 35 shows the EV (negative control) (A), the positive control with P450 BM3 variant M2 and *p*-xylene as substrate (B) as well as the conversion of *p*-chlorotoluene (C).

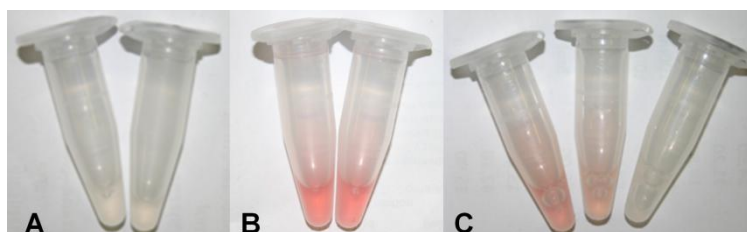


Figure 35. Screening of new substrates for conversion by P450 BM3 variant M2 with the 4-AAP assay. A: Empty vector (negative control) B: *p*-xylene (positive control for 4-AAP assay) C: Substrate *p*-chlorotoluene (Substrate 6); left tube: 1 mM NADPH added; middle tube: 0.5 mM NADPH added; right tube: no NADPH added.

Table 8 summarizes the results from the screening of the substrate library with variant M2. Nine out of 28 substrates showed positive response in the 4-AAP assay which was monitored by red or pink color formation (Figure 35), whereas the control reaction without NADPH displayed no color formation (Figure 35 C; right tube). However, out of the ten positive hits in the 4-AAP assay, four false positive hits were identified in the later GC-FID and GC-MS analysis, either no product at all (hydrocinnamic acid and *p*-toluidine) nor a phenol (cumene and styrene) could be identified. During GC-MS analysis six hits from the 4-AAP assay screening were confirmed for formation of a phenolic product (chloro- and bromobenzene; 4-chloro- and 4-bromobenzene; ethylbenzene; 4-methylanisole) (Table 8).

Table 8. Results from conversion screening of 28 selected benzenes with variant M2. A qualitative evaluation was done to explore substrate scope of variant M2. Parts of the content in this table have been published as supplementary information.[68]

Substrate	4-AAP	GC-FID	GC-MS	Aromatic*	Non-aromatic*
1	-	-	-	-	-
2	-	+	+	+	-
3	-	-	-	-	-
4	-	-	-	-	-
5	+	+	+	+	-
6	+	+	+	+	+
7	-	-	-	-	-
8	-	-	-	-	-
9	+	+	+	+	-
10	+	+	+	+	+
11	+	+	-	-	-
12	-	-	-	-	-
13	+	+	+	+	+
14	-	-	-	-	-
15	+	+	+	+	-
16	+	-	-	-	-
17	-	-	-	-	-
18	-	-	-	-	-
19	-	-	-	-	-
20	+	+	-	-	-

Substrate	4-AAP	GC-FID	GC-MS	Aromatic*	Non-aromatic*
21	-	+	+	+	+
22	+	+	+	-	Epoxidation
23	-	+	+	-	Epoxidation
24	-	-	-	-	-
25	-	-	-	-	-
26	-	-	-	-	-
27	n.t.	+	+	+	+
28	n.t.	+	+	-	+

(n.t. = not tested), '+' (product detected); '-' (no product detected); *chemoselectivity

All other substrates did not show any color formation in the 4-AAP assay. Altogether, 14 out of 28 substrates displayed one or more distinct product peak on GC-FID chromatograms. The products for 11 converted substrates could be identified by GC-MS and are summarized in Table 9.

Table 9. Products identified by GC-MS analysis. Parts of the content in this table have been published as supplementary information.[68]

Substrate	Products	Product 1	Product 2	Product 3
2	1	2,4,6-Trimethylphenol	-	-
5	1	2-Chlorophenol	-	-
6	3	2-Chloro-5-methylphenol	5-Chloro-2-methylphenol	4-Chlorobenzylalcohol
9	1	2-Bromophenol	-	-
10	3	2-Bromo-5-methylphenol	5-Bromo-2-methylphenol	4-Chlorobenzylalcohol
13	3	Phenylethanol (<i>R/S</i>) [#]	Acetophenone	2-Ethylphenol
15	1	2-Methoxy-5-methylphenol	-	-
20	1	2-Methyl-2-phenylpropan-1-ol	-	-
21	1	n.d.	-	-
22	1	Styrene epoxide (<i>R/S</i>)	-	-
23	1	4-Methylstyrene oxide (<i>R/S</i>)	-	-
27	3	2-Fluoro-5-methylphenol	5-Fluoro-2-methylphenol	4-Fluorobenzylalcohol
28	1	<i>p</i> -Nitrobenzylalcohol	-	-

[#]main product (GC area); n.d. (not determined)

Four of the converted substrates displayed more than one product peak on GC-FID and GC-MS (Table 9; Substrates: 6, 10, 13 and 27). For instance employing 4-chloro- or 4-bromotoluene in the reaction, two phenols and the respective benzyl alcohol were formed. Chloro- and bromobenzene revealed formation of only one distinct product that was identified as the respective *o*-phenol (5 and 9). 4-Methylanisole was also hydroxylated exclusively in *o*-position to 2-methoxy-5-methylphenol (15) with no further side product formation. Both styrene substrates (22 and 23) were converted to the respective styrene epoxides. Ethylbenzene conversion displayed three product peaks on GC-FID (13).

The substrate was converted mainly to phenylethanol (>90 %) which was in small parts further oxidized to acetophenone. The third peak was identified as 2-ethylphenol, however, intensity of the product peak was very low on GC-FID. The only compound with a methylbenzene substructure that was exclusively hydroxylated in α -position was *p*-nitrotoluene (28).

3.2.4.2 Investigation of Co-solvent Influence on Activity of P450 BM3

The second part of this chapter focuses on the influence of co-solvents on catalytic performance during hydroxylation of toluene by P450 BM3 variant M2. DMSO and IPA were selected as co-solvents since they are widely applied in biocatalytic approaches either to solubilize substrates in aqueous bulk[182, 253] or as a cheap source for cofactor regeneration and racemic resolution for instance in ADH catalysis.[51, 62] In the performed experiments co-solvent concentrations from 0 to 30 % were used. Figure 36 (A and B) shows the co-solvent dependent catalytic performance of the engineered variant M2 in the presence of increasing concentrations of DMSO.

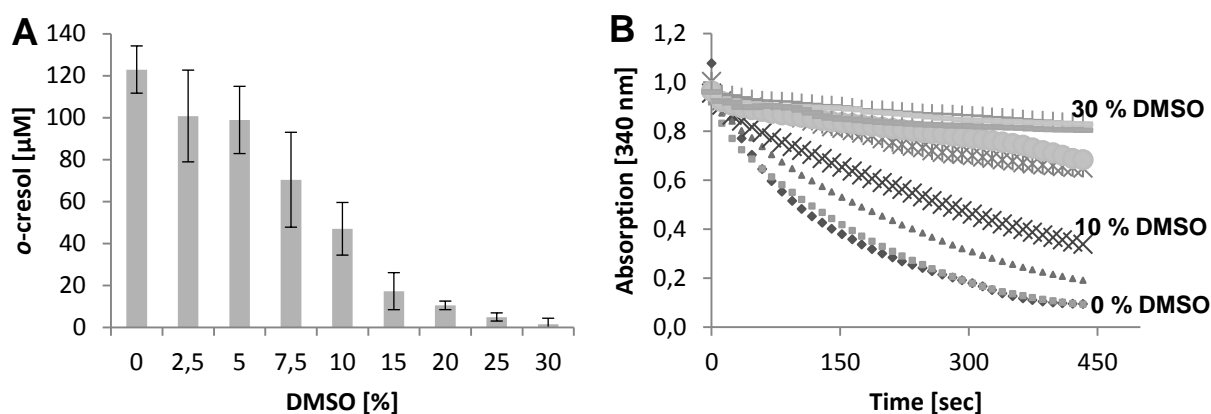


Figure 36. A: Formation of *o*-cresol by variant M2 over time at varying DMSO concentrations (0 to 30 %). The reaction was quenched after NADPH was fully depleted in the first well (here at 0 % DMSO). **B:** NADPH depletion assay using P450 BM3 variant M2 and increasing amount of DMSO (0 to 30 %) in the reaction system.

The engineered variant M2 displayed highest activity and product formation (\varnothing 120 μM *o*-cresol) when no DMSO was present in the reaction system (Figure 36 A). Increasing the amount to 5 % DMSO the productivity was decreased by 10 % which is in the deviation range of the applied 4-AAP assay (3.1.4).[228] Half of the activity and productivity was obtained when 7.5 to 10 % of co-solvent was applied in the reaction system. Beyond 20 % DMSO, only spare hydroxylation activity was retained ($< 5 \mu\text{M}$ *o*-cresol) within the same reaction period. The oxidation of NADPH followed the same trend; therefore, the highest NADPH oxidation rate was obtained without any additional co-solvent (Figure 36 B). Increasing the DMSO concentration to 25 or 30 % respectively, led to an inactivation of the P450 monooxygenase. Applying IPA as co-solvent, under the same reaction and detection conditions, a similar trend was visible. Without co-solvent in the reaction the highest product formation rate was achieved (on average 95 μM *o*-cresol). Up to 5 % IPA no change or loss in productivity or activity was visible (Figure 37 A and B). At 10 to 15 % IPA the half maximal

productivity was reached, whereas beyond 15 % IPA the P450 catalyst did not display product formation or detectable NADPH depletion activity.

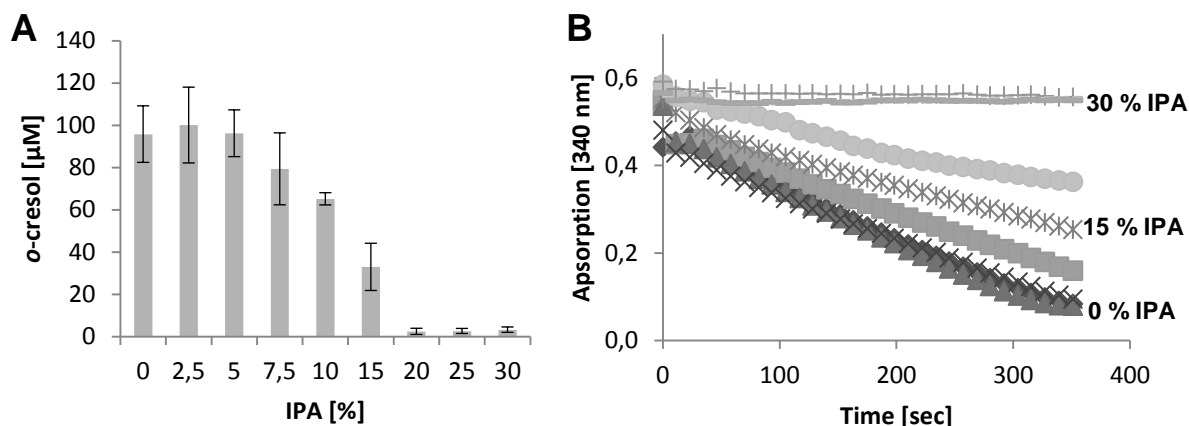


Figure 37. A: Time dependent formation of *o*-cresol by variant M2 at increasing IPA concentrations (0 to 30 %). The reaction was quenched after NADPH was fully depleted in the first well (here at 0 % IPA). **B:** NADPH depletion assay using P450 BM3 variant M2 and increasing amount of IPA (0 to 30 %) in the reaction system.

Compared to DMSO, the increasing concentrations of IPA had a similar detrimental effect on productivity and activity of the P450 BM3 variant M2 (Figures 36 and 37). However, with IPA half of the productivity was retained at a concentration between 10 and 15 % IPA (for DMSO between 7.5 and 10 %).

3.2.5 Discussion

The screening with 28 different substrates revealed that the engineered P450 BM3 variant is a versatile catalyst for the hydroxylation of benzenes, although no additional mutations in the substrate binding site had to be introduced.[69] Applying the regioselective 4-AAP assay it was possible to identify with good precision which substrate is hydroxylated on the aromatic ring to produce phenol (Table 8 and 9). Although false positive hits were obtained, 6 out of 10 positive hits in the initial screening were confirmed later as phenols by GC-MS analysis (Table 9). In addition to aromatic hydroxylation, six substrates were hydroxylated at α - or sub-terminal positions for instance *p*-nitrotoluene, ethylbenzene or cumene (Table 9). The observation that ethylbenzene ($-\text{CH}_2-\text{CH}_3$ substituent) is preferentially hydroxylated in sub-terminal position (>90 % GC area) was reported before.[257] However, toluene (one $-\text{CH}_3$ substituent) and *p*-xylene (two $-\text{CH}_3$ substituents) are preferentially hydroxylated on the aromatic ring by P450 BM3 WT and variant M2 (3.1).[68, 69] This would suggest that BM3 can efficiently distinguish between associated benzene ring substituents.[68-70, 156, 180, 206, 257] The ethyl-substituent provides more sterical interaction in the active site of P450 BM3, leading most likely to a different substrate orientation in the binding pocket. In addition, the preferential hydroxylation of mesitylene on the aromatic ring (Table 9; section 3.4) also supports this observation, which means that through space-binding interactions a chemoselective hydroxylation of benzenes by P450 BM3 becomes feasible.[70, 179] Consequently, a sterically more demanding substituent such as an ethyl-

group promotes a different orientation in the active site compared to toluene, *p*-xylene or mesitylene. The hydroxylated sub-terminal C-atom should be in closer orientation than the C2-atom in the aromatic ring; however, traces of a second product were detected and identified as ethylphenol by GC-MS analysis (Table 9). The fact that variant M2 catalyzes also benzylic hydroxylation of e.g. ethylbenzene generates access to other valuable building blocks in organic synthesis for pharmaceuticals, flavors, pigments and dyes.[192-196] All *p*-substituted toluenes (fluor-, chloro-, and bromotoluene) were converted by variant M2 leading to formation of three products, the two isomeric phenols as well as the respective benzylalcohol (Table 9). The observation that three products are obtained suggests that also for this class of substrates more than one orientation mode within the active site of P450 BM3 is possible. However, from the recorded data it is not possible to disseminate which substrate orientation is the preferred one, but BM3 seems to recognize also other functional groups for orientation of the substrate in the active site. This finding is also supported by the data recorded for monosubstituted benzenes such as toluene, chloro- and bromobenzene where only one major product peak was identified, indicating selective substrate recognition and preferred binding mode. A rather unusual result was obtained for conversion of *p*-nitrotoluene and 4-methylanisole. Compared to the *p*-halogenated-toluenes, *p*-nitrotoluene contains a large and strong electronegative NO₂ substituent (4.2 Pauling Units)[258] that seems to direct hydroxylation with excellent selectivity towards α -hydroxylation (section 3.6). The fact that no phenol was detected in the reaction is rather unusual when compared to the substrates hydroxylated preferentially on the aromatic ring such as the '*p*-halogenated-toluenes' and especially 4-fluorotoluene, which also contains a strong electronegative group (3.9 Pauling Units)[258] or bromine, which has a larger atom radius than most of the investigated substituents. The investigation of *p*-nitrotoluene hydroxylation by P450 BM3 was further preceded in chapter 3.6. In addition, the results obtained during the conversion of 4-methylanisole were unexpected since hydroxylation led to one distinct product identified as 2-methoxy-5-methylphenol (confirmed by GC-MS). The observation that 4-methylanisole is solely hydroxylated on the aromatic ring in *o*-position to the OCH₃ substituent is in contradiction to the results obtained for *p*-halogenated-toluenes. In this respect, an isomeric product mixture of two phenols and the respective benzyl alcohol was obtained employing for example 4-chlorotoluene as substrate (Table 9). The question therefore remains more current than ever: How does P450 BM3 differentiate between methoxy-, nitro- as well as halogen-substituents to perform reactions with varying chemoselectivity? Sterical differences play a key role in substrate recognition, orientation and selectivity[70, 179], whereas aromatic interactions between arene substrates and aromatic residues in proteins can strongly influence catalytic performance.[68, 259] The latter was shown for the substitution of residue F87 to all 19 canonical amino acids (Figure 31).[69] The associated substrate substituents very likely influence by electronegativity[258], inductive and/or mesomeric effects the π -electron density leading to 'substituent specific' orientations, contributing to the broad diversity in chemoselectivity.[68, 69, 154, 156] Figure 38 summarizes the products that were produced by P450 BM3 variant M2 as major products (GC area >90 %) and highlights the requirement for different chemo-selectivity.

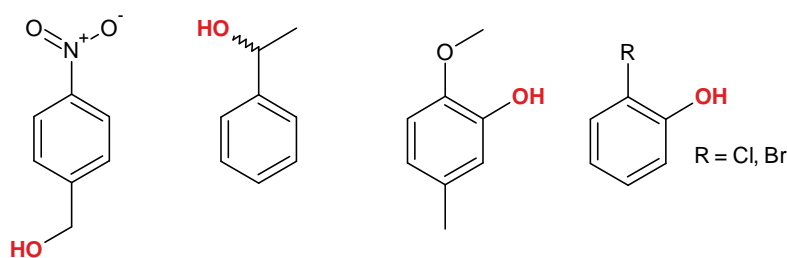


Figure 38. Selected products obtained after hydroxylation with P450 BM3 variant M2 and their abundance as relative GC-area. From left to right: *p*-nitrobenzylalcohol (>99 %); 2-phenylethanol (>90 %); 2-methoxy-5-methylphenol (>99 %), 2-bromo and 2-chlorophenol (>99 %).

No product formation was obtained under the described conditions employing benzene substrates with NH_2 and sulfur substituents or hetero-aromatic compounds (Table 8). One reason might be that the substrates do not enter the active site or bind in a non-productive orientation within the binding pocket or substrate channel. Substrates containing larger substituents such as hydrocinnamic and *trans*-cinnamic acid were also not converted to a detectable amount. Carboxylic and charged substituents tend to interact with other charged or polar residues (e.g. fatty acids and arginine 47 in P450 BM3).[185, 186] Consequently, the binding of the substrate close to the heme iron might be prevented by ionic interactions with other residues of P450 BM3. Styrene and methylstyrene were hydroxylated as reported before, solely to the respective epoxides[145, 191] which are valuable building blocks in polymer chemistry. In case of diethylbenzene, one product peak was identified on GC-FID, however GC-MS proposed either 2,5-diethylphenol or 1-(4-ethylphenyl)ethan-1-ol as products. In this case, a purification of products followed by NMR analysis or application of commercial standards would be necessary to verify the obtained structure. Although products from P450 BM3 hydroxylation could be identified, the mass analysis gives only a certain probability for a potential product. Some of the investigated substrates were applied and investigated in more detail in subsequent chapters such as mesitylene, *p*-nitrotoluene and the halogenated benzenes.

An important consideration for poorly soluble substrates in biocatalysis is the application of co-solvents to increase the accessibility to a target enzyme.[52, 182, 253] In chapter 3.1, the characterization of variant M2 for hydroxylation of *p*-xylene, which is insoluble in aqueous solutions, revealed that only 1.5 % DMSO is sufficient to perform the hydroxylation reaction with high productivity ($k_{\text{cat}} \sim 2000 \text{ min}^{-1}$) and excellent coupling efficiencies (66 %) (3.1.4).[69] Therefore, the influence of co-solvents on catalytic activity and productivity of P450 BM3 variant M2 was further investigated to examine the effect of co-solvents during aromatic hydroxylation of a benzene substrate (here toluene was employed as model substrate). Results show that the engineered P450 catalyst performs best in solely aqueous solution without addition of extra co-solvent (Figures 36 A and 37 A). Since commonly used co-solvents are miscible in hydrophobic bulk the influence on P450 BM3 could be diverse.[182, 255] The highly hydrophobic substrate channel of P450 BM3 promotes inlet of hydrophobic molecules as reported for the natural substrates, medium chain fatty acids.[70] Applying

high concentrations of DMSO and IPA could lead to accumulation of the co-solvents within the active site, enabling inlet of water or salt molecules, which could generate a hydrophilic environment that reduces access for hydrophobic substrates. A hydrophilic ‘plugging’ would also slow down initiation of the catalytic cycle (Step 1, Figure 7)[70] resulting in a slower transfer of electrons from NADPH to the heme iron (Step 2, Figure 7), which can be measured by depletion of NADPH.[227, 229] The fact that NADPH depletion is significantly decreased in the reaction mixture by increasing co-solvent concentrations (15 to 30 %) (Figure 36 B and 37 B) could be an indicator that the substrate is not entering the active site. This was reported and discussed before for a P450 BM3 crystal structure with DMSO as ligand.[255, 256] DMSO was reported to bind (in close distance) to the heme iron which could prevent binding of the substrate to initiate the catalytic mechanism of P450s. However, during the measurements with DMSO and IPA similar effects were observed for both co-solvents, indicating that other effects could also play a role leading to a decrease in activity. A previous protein engineering study with P450 BM3 towards higher activity in presence of various co-solvents showed that the mutation F87A decreases relative activity significantly when catalysis with P450 BM3 is performed in organic co-solvents[182] indicating on the one hand a strong effect of the respective co-solvent in close distance to the heme iron. On the other hand, since P450 BM3 is a dimeric protein (Figure 8)[174] it requires the interaction of both subunits for its exceptional high activity.[166] The addition of co-solvents could lead to a dissociation or loss of interaction between both monomers leading to the observed decrease in catalytic activity and productivity. Generally, most proteins (soluble and not membrane bound) are covered with a shell of water[260] that ensures solubility and a defined fold in aqueous environments, jointly ensuring the catalytic function. DMSO and IPA can move between hydrophilic and hydrophobic compartments and therefore could also lead to structurally unfolding of secondary protein structures which could result in decreased activity.[260, 261] The fact that both co-solvents display similar effects on the catalytic activity could be explained by a structural change in a catalytically important area of P450 BM3. In the mentioned engineering study, variants with higher tolerance against several commonly applied co-solvents were found.[182] All mutations improving catalytic performance at increasing co-solvent concentrations were accumulated in the reductase domain and peptide linker[182], which is important for efficient electron transfer between the two monomers of BM3 enabling the exceptional high catalytic activity.[166, 174] The presence of high concentrations of co-solvent could lead to the dissociation of P450 BM3 monomers which would explain the significant loss in activity that was observed in a similar way for both co-solvents (Figure 36 and 37).

3.2.6 Summary and Conclusion

One objective in this chapter was the exploration of the engineered P450 BM3 variant M2 for a larger substrate scope and its potential to generate new phenolic building blocks. Additionally, the hydroxylation of further substrates should enable a more detailed understanding on how P450 BM3 converts benzene substrates. The screening of 28 benzene substrates with variant M2 revealed that

several substrates could be hydroxylated efficiently on the aromatic ring, on α - or sub-terminal positions. Halogenated monosubstituted substrates were hydroxylated with excellent selectivity, whereas *p*-halogenated toluenes were hydroxylated at all three available positions, indicating that several binding modes are possible for this class of substrates. *p*-Nitrotoluene was hydroxylated exclusively in α -position indicating that the nitro-substituent strongly influences orientation and electron distribution in the aromatic ring enabling a highly specific orientation within the P450 BM3 active site. Furthermore, 4-methylanisole was exclusively hydroxylated in *o*-position to the OCH₃-substituent. The screening enabled a pre-selection of new substrates that were further investigated in detail in subsequent chapters of this thesis.

The investigation of increasing concentrations of DMSO and IPA during P450 BM3 catalysis revealed that the application of a co-solvent can be omitted for efficient hydroxylation of toluene. Although the results display only a snapshot for a specific substrate, the application of co-solvents is not always optimal, even though the substrate retains low solubility in the aqueous bulk. Major effects can be attributed to a blocking of the substrate channel or the heme, as well as detrimental effects on the dimeric structure of P450 BM3. However, from the recorded data it remains unclear which is the strongest influencing effect. Concentrations beyond 5 % (DMSO and IPA) should be avoided. The use of an entirely co-solvent free biocatalytic approach would be optimal, with regards to sustainability as well as economic criteria. However, a general transfer to all substrate classes is not possible; especially for solid substrates the application of a co-solvent might be inevitable due to technical challenges during application, for balancing and adjustment of concentrations in the reaction vessel. The fact that the engineered P450 BM3 variant catalyzes hydroxylation of new and unexplored substrates and the surplus of a co-solvent free catalysis in water was transferred to further investigation of substrates in this thesis.

3.3 Regioselective *o*-Hydroxylation of Monosubstituted Benzenes by P450 BM3

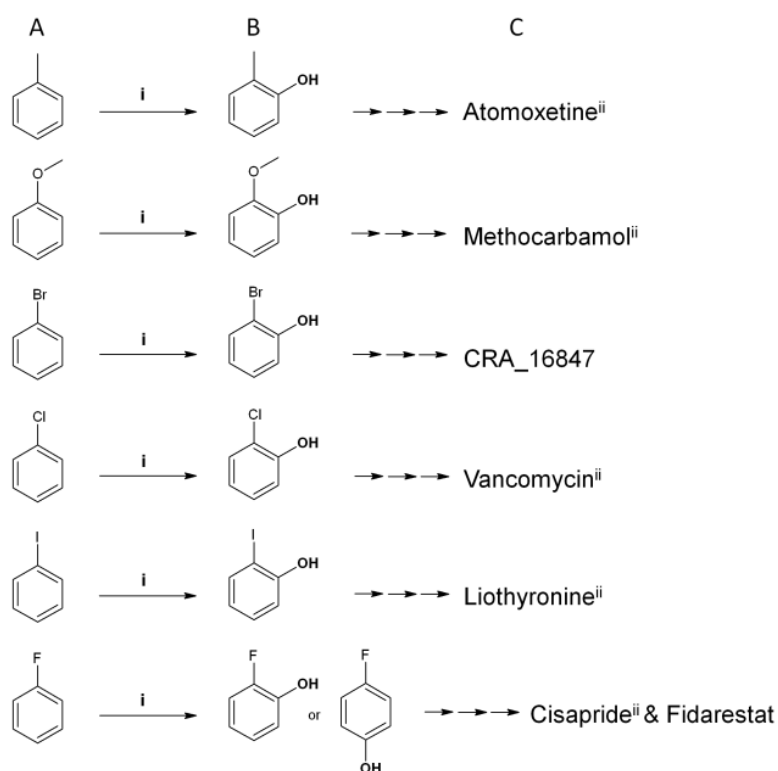
3.3.1 Abstract

The previously engineered P450 BM3 variant M2 (R47S, Y51W, I401M) was investigated for the conversion of six selected monosubstituted benzenes with varying functional groups (F, Cl, Br, I, CH₃ and OCH₃). The catalyst displayed excellent initial turnover frequencies (iTOF of 38.5 s⁻¹; anisole) for the production of important building blocks for pharmaceutical intermediates (substituted *o*-phenols). All substrates were hydroxylated in water without co-solvent and using molecular oxygen as sole oxidant. GC-FID and GC-MS analysis revealed for variant M2 excellent regioselectivity (>99 %) and appealing coupling efficiency (48 %) outperforming the WT. In addition, this is the first direct hydroxylation reported for fluoro- and iodobenzene by a P450 monooxygenase. The results are discussed with respect to influence of substituents on the benzene ring and aromatic interactions

within the active site of P450 BM3. Docking of the substrate into the enzyme's active site supported the theory that a T-shape orientation of benzene substrates is favored over face-to-face or dimer interactions with residue F87.

3.3.2 Introduction

Phenols are important building blocks for plastics or resins and key intermediates for the synthesis of pharmaceuticals and pro-drugs (Scheme 7).[193, 197, 262, 263] In particular, halogenated phenols are widely used in vitamin synthesis, lipid lowering agents or as drugs in treatments of the respiratory system; halogenated phenols are therefore regarded as “active pharmaceutical ingredient” (API) (Scheme 7).[192-196] The halogenated phenols are barely found in nature and mainly produced in chemosynthetic processes with moderate selectivity.[193, 197, 263, 264] Synthesis and purification of isomeric phenols requires energy-intensive down-streaming processing and hazardous chemicals such as mercury/thallium acetate derivatives for the synthesis of iodophenol.[193, 197, 200] Regioselectivity of the electrophilic substitution reactions is governed through inductive and mesomeric effects of substituents at the benzene ring.[194] Chlorination of phenol yields in equal amounts 2- and 4-chlorophenol; furthermore, di-chlorophenols are formed after elongated reaction times.[194] Consequently, careful purification of isomeric product mixtures is required to obtain pure regiomers for drug synthesis.[73] However, products with high purity (>99 %) are required in applications as synthons for drugs and to prevent unwanted side effects.[73, 196]



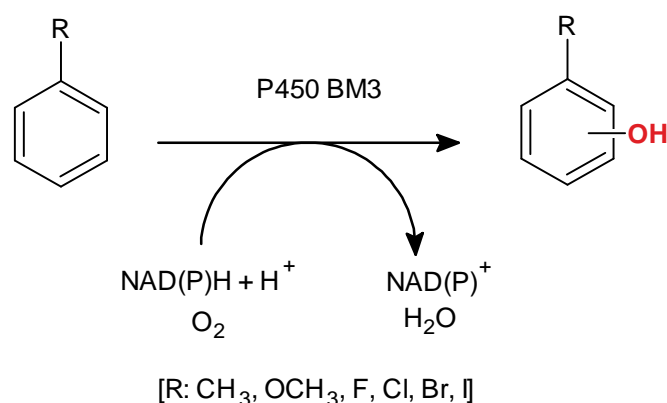
Scheme 7. A: Aromatic substrates employed for hydroxylation with P450 BM3 variant M2 (R47S, Y51W, I401M).[69] **B:** Products obtained after aromatic hydroxylation with P450 BM3 variant M2. **C:** Selected FDA approved drugs that contain in the product structure from the “(B) column” (<http://www.drugbank.ca/on>). The Scheme was taken from Dennig et al.[68]

Guaiacol as well as its derivatives, such as vanillin and eugenol, are common antioxidants in treatment of cancer, cardiovascular disorders, as well as Parkinson’s and Alzheimer’s diseases.[265] In addition, guaiacol is an important intermediate for the production of flavors and pharmaceuticals[266] and can be isolated from plants[267] or synthesized using catechol as educt in presence of corrosive reagents and stoichiometric amounts of NaOH.[266] In 2011, Chen et al. reported an environmental friendly synthesis employing rhenium-promoted mesoporous zirconia as catalysts for halogenation of phenols which represents the state-of-the-art technology.[193] The reported catalysts brominates with moderate selectivity (84 % 4-bromophenol, 9 % 2-bromophenol and 7 % 2,4-di-bromophenol) and modest productivity (71 % conversion of 0.5 mmol phenol with 100 mg catalyst in 24 h). The direct hydroxylation of substituted benzenes with chemical catalysts is unattractive since reactive oxygen species target non-activated C-atoms unselectively[71] and preferentially not at the aromatic ring.[201]

Biocatalysis offers alternative and attractive routes to direct aromatic hydroxylation of benzenes to the corresponding phenols.[72, 73, 146] The biocatalysis in water can reduce significantly purification/reaction steps, energy demands and expenses for waste disposal of hazardous organic solvents or unwanted byproducts, such as corrosive HBr and HCl. Hydrogen halides are produced in large amounts during electrophilic halogenation reactions.[71, 75, 146, 192, 193] Especially the

potential of regio- and enantioselective catalysis in aqueous environment makes direct enzymatic hydroxylations attractive for the production of pharmaceuticals.[71-73] A major limitation for application of biocatalysts in technical processes is the often insufficient operational stability.[66] For example, only a few peroxidases are known to halogenate phenols with moderate selectivity.[146]

P450 BM3 (CYP102A1), a heme-dependent and industrially important monooxygenase[70], is capable of performing direct aromatic hydroxylation of toluene and *p*-xylene with high selectivity and activity.[68-70, 180, 206, 246] To date, P450 monooxygenases are mainly used to hydroxylate aromatic compounds with alkyl-substituents[70, 146] whereas hydroxylations of halogenated benzenes are nearly neglected despite of their synthetic potential.[193, 265] Earlier in this thesis (section 3.1), the P450 BM3 variant M2 (R47S, Y51W, I401M) was engineered towards efficient and regioselective aromatic hydroxylation of *p*-xylene, to produce 2,5-dimethylphenol.[69] During the substrate screening with the same variant (section 3.2.4.1), bromo- and chlorobenzene showed selective formation of a product that was identified in both cases as the respective *o*-phenol (Table 8 and 9). Based on this observation four additional substrates (anisole, toluene, fluoro- and iodobenzene) were selected to investigate the hydroxylation of monosubstituted benzene substrates by P450 BM3 WT and variant M2 (Scheme 8).



Scheme 8. Aromatic hydroxylation of selected monosubstituted benzenes by P450 BM3

The observation that P450 BM3 displays highest productivity when no additional DMSO co-solvent is present was included in this study (Figure 36 and 37, section 3.2.4.2). No additional co-solvent was therefore employed in the reaction system, which is also beneficial for a later synthetic application where co-solvents such as DMSO, DMF or THF interfere with most purification methods.

3.3.3 Experimental

Large parts of this chapter were published in *Angewandte Chemie* (2013).[68] Nina Lültsdorf assisted in catalytic characterization of P450 BM3 variants during her master thesis project at the Institute of Biotechnology (RWTH Aachen University).

3.3.3.1 Determination of Catalytic Performance of P450 BM3 Wild-Type and Variant M2

Every reaction set-up contained substrate (30 mM), NADPH (550 μ M), purified P450 BM3 WT (0.135 μ M) and P450 BM3 variant M2 (R47S, Y51W, I401M; 0.075 μ M) in a final volume of 5 mL KPi buffer (50 mM; pH 7.5). The reaction mixture was incubated in a glass flask (25 mL; 5 min; 700 rpm; RT) on an Eppendorf MixMate. Coupling efficiencies were performed as described before in the material and methods section (section 2.10.9.1). P450 BM3 activity was induced by addition of 550 μ M NADPH. For variant M2, samples (200 μ L) were taken every 20, 40, 60, 90, 120, 360 and 600 s and quenched in 25 μ L urea (4 M) solubilized in NaOH (0.1 M). The detection of phenolic products was performed as described before employing the 4-AAP detection system (section 2.10.3).[68] Calibration curves with commercial standards of the respective *o*-phenols were recorded (from 0 to 200 μ M) and used for quantification of reaction products.

3.3.3.2 Gas Chromatography Analysis of Reaction Samples (GC-FID and GC-MS)

All products from conversions were extracted as described before using MTBE as extraction solvent with/and without phenol as internal standard (section 2.11.1). Identification of products from P450 BM3 conversions was achieved via GC-MS analysis and GC-FID employing commercially available standards of all three phenol isomers. A detailed list of all substrates and products separated on GC (retention times, temperature programs and employed columns can be found in appendix section; Table 29).

3.3.4 Results and Discussion

In this part of the thesis, the aromatic hydroxylation of six substituted benzenes (Scheme 8) with the engineered variant M2[69] was investigated for the production of *o*-substituted phenols by direct hydroxylation with molecular O₂ at room temperature and in water (50 mM KPi; pH 7.5). Despite of the comparably poor solubility of substrates in water (<2 g L⁻¹)[268], no additional co-solvents were employed during catalysis. In a systematical approach the influence of ring-substituents (F, Cl, Br, I, CH₃ and OCH₃) on regioselective hydroxylation was investigated. The highly sensitive, colorimetric and regioselective 4-AAP assay[228] allowed reliable quantification of *o*-hydroxylated phenols. GC-FID and GC-MS were employed with commercial standards to determine regioselectivity of variant M2. Table 10 and Figure 39 show that the P450 BM3 WT catalyzes product formations with moderate activities which do not exceed 0.6 U mg_{P450}⁻¹ (e.g. guaiacol).

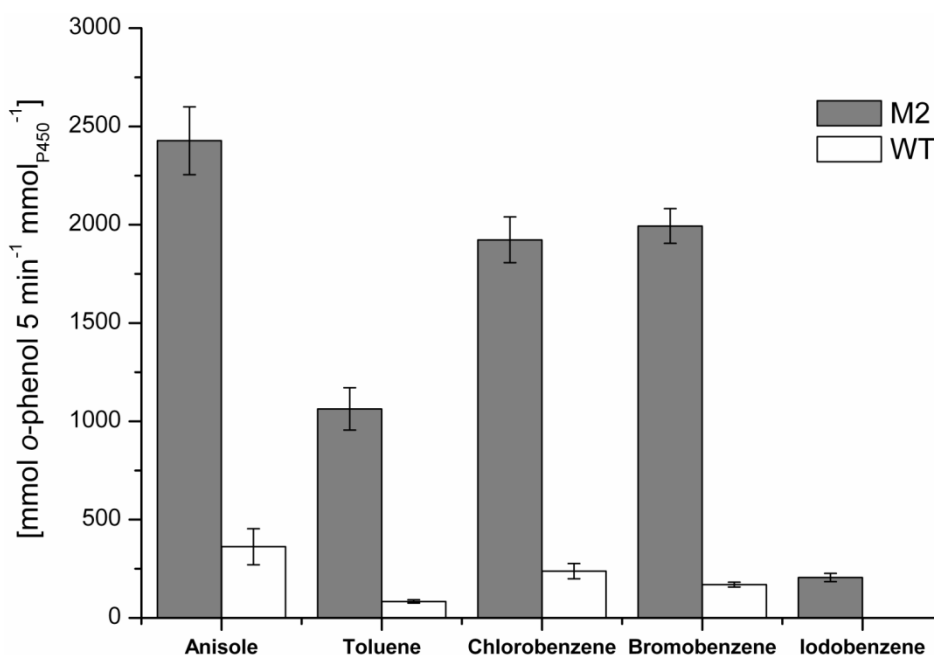
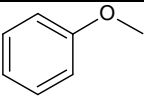
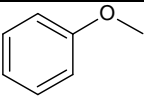
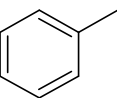
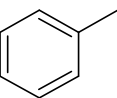
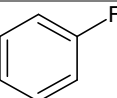
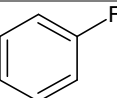
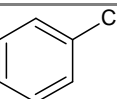
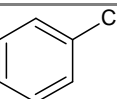
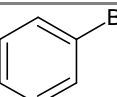
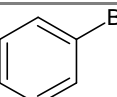
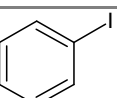
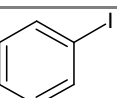


Figure 39. Product formation of *o*-hydroxylated phenols from anisole, toluene, chloro-, bromo- and iodobenzene by P450 BM3 WT as well as variant M2 (R47S, Y51W, I401M).[69] The reaction was quenched after 5 min and phenolic products were quantified using the regioselective 4-AAP assay.[228] All experiments were performed in triplicate with 30 mM of substrate. Parts of the content in this table have been published as supplementary information.[68]

Fluoro- and iodobenzene are not converted to a detectable amount by the WT enzyme. Variant M2 could efficiently hydroxylate all six substrates including fluoro- and iodobenzene. Based on initial turnover frequency (iTOF) values[68], particularly the aromatic hydroxylation of anisole was performed very efficiently with $19.5 \text{ U mg}_{\text{P450}}^{-1}$. An iTOF of 38.6 s^{-1} and a specific activity of $19.5 \text{ U mg}_{\text{P450}}^{-1}$ is, so far, the highest reported product formation rate for a direct aromatic hydroxylation of a benzene compound by a monooxygenase.[68-70, 180, 206, 246] Interestingly, the productivity of guaiacol is about 3-fold higher than for *o*-cresol ($7.4 \text{ U mg}_{\text{P450}}^{-1}$), most likely due to differences in sterical demands (Table 10).[69, 70] Toluene, bromo- and chlorobenzene are also converted much faster by P450 BM3 M2 than by the WT (12.7-fold in case of toluene; 8-fold and 11.7-fold) even under NADPH limited conditions (TOF values; Table 10).

Table 10. Catalytic performance of P450 BM3 WT and variant M2 for the aromatic hydroxylation of six substrates. ^[a]Initial turnover frequency (iTOF) [$\mu\text{mol product } \mu\text{mol}_{\text{P450}}^{-1} \text{ min}^{-1}$ or $\text{U } \mu\text{mol}_{\text{P450}}^{-1}$] (maximal initial productivities were measured after 20 s for variant M2 and after 300 s for WT); ^[b]turnover frequency ($\mu\text{mol}_{\text{Product}} \mu\text{mol}_{\text{P450}}^{-1} 300 \text{ s}^{-1}$) ^[c]specific enzyme activity ($\mu\text{mol}_{\text{Phenol}} \text{ min}^{-1} \text{ mg}_{\text{P450}}^{-1}$) ^[d]coupling efficiency; n.d. (not detected); #~3 % benzylalcohol is produced; *51.1 % (GC area) of measured products is phenol; [§]Determined after 30 min reaction time. Ph = phenol. The Table was taken from Dennig et al.[68]

Variant	Substrate	iTOF [s^{-1}] ^[a]	TOF [300 s^{-1}] ^[b]	U $\text{mg}_{\text{P450}}^{-1}$ ^[c]	C [%] ^[d]	<i>o</i> -Ph	<i>m</i> -Ph/ <i>p</i> -Ph
WT		n.d.	362 ± 92	n.d.	11 ± 4	>90	n.d./<10
M2		39 ± 2	2427 ± 173	19.5	44 ± 4	>95	n.d./<5
WT		n.d.	84 ± 8	n.d.	10 ± 3	>95	<1 [#]
M2		15 ± 1	1063 ± 108	7.4	48 ± 4	>99	n.d.
WT		n.d.	n.d.	n.d.	0	n.d.	n.d.
M2		n.d.	67 ± 31 [§]	0.02	6 ± 3	49	n.d./51
WT		n.d.	238 ± 39	n.d.	6 ± 3	>96	<4
M2		12 ± 1	1923 ± 88	6.2	28 ± 4	>99	<1
WT		n.d.	170 ± 12	n.d.	5 ± 3	>97	<3
M2		8 ± 1	1994 ± 88	3.8	29 ± 6	>99	n.d./<1
WT		n.d.	n.d.	n.d.	0	n.d.	n.d.
M2		2 ± 0.7	205 ± 21	0.97	23 ± 8	>99	n.d.*

The WT protein and variant M2 displayed excellent regioselectivity (>90 % *o*-hydroxylation); M2 reaches >99 % in *o*-hydroxylation of toluene and halogenated substrates. An exception is fluorobenzene (Table 10) that was converted with a ratio of 49:51 (*o*- to *p*-hydroxylation), and lacks *m*-hydroxylation that was reported for eukaryotic P450s from rat.[269] The fluor substituent is highly electronegative and can potentially reduce or alter π - π -interactions with the phenyl ring of F87.[258] Changed π - π -interactions could result in a loss of a precise sterical control yielding a racemic *o*-/*p*-product mixture.[270] The importance of amino acid residue F87 for aromatic hydroxylation of *p*-xylene was investigated in chapter 3.1 and attributed to strong π - π -interactions between substrate and phenylalanine 87.[69, 259] Regioselective hydroxylation of substituted benzenes in BM3 WT and M2 is controlled through space by binding and orienting benzene substrates in a defined manner towards T268, a key residue in O₂ activation and substrate recognition (Figure 40).[70, 179]

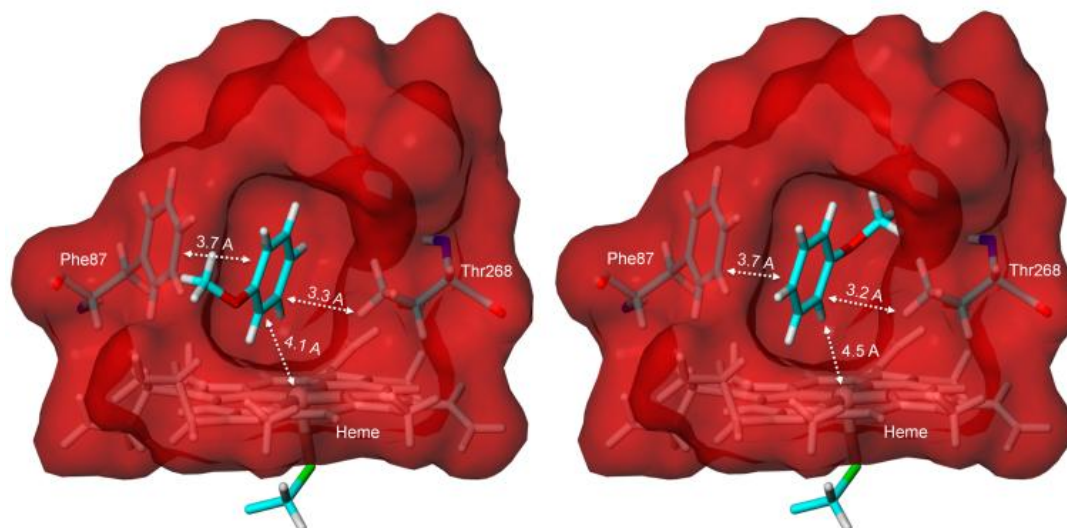


Figure 40. Visualization of P450 BM3's (PDB 1BU7)[238] active site cavity with the anisole substrate docked in two T-shape binding orientations. The red surface indicates space filling of residues in P450 BM3 active site within <4 Å distance to the substrate. Distances are given in Å between anisole and the key residues F87, T268 as well as the heme iron. Docking was achieved using the VINA docking plug-in for YASARA.[109, 237] The Figure was taken from Dennig et al.[68]

Small compounds such as benzenes have to be positioned towards the active site for regioselective hydroxylation since the large substrate binding pocket of P450 BM3 can harbor for instance polycyclic aromatic hydrocarbons such as naphthalene or pyrene.[70, 189, 208] P450 BM3 controls reactions through space, and the intrinsic reactivity of substrates plays not a dominant role during catalysis in contrast to most chemical catalysts.[271] Figure 40 shows the substrate docked into an active site cavity located between the key catalytic residues F87, T268 and the heme iron. The substrate fits excellently in T-shape orientation[259] with strong binding energies of -4.98 and -5.46 kcal mol⁻¹. A key-performance parameter of P450 monooxygenases is the coupling efficiency, as a measure for an efficient use of the reduction equivalent NADPH that is regarded as major cost factor in biocatalytic approaches with P450 monooxygenases.[70, 162] High coupling efficiencies require that a substrate is positioned in a specific orientation so that an efficient transfer of the activated oxygen can be achieved from BM3 to the targeted C-atom (Step VII and VIII in the catalytic cycle of P450 BM3; Figure 7).[157] The WT displayed coupling efficiencies ranging from 5 to 11 % for converted substrates (Table 10). No NADPH consumption was observed for iodobenzene suggesting that iodobenzene is not a substrate for P450 BM3 WT. One reason could be, as reported for DMSO, an interaction of the substrate with the amino acid residue R47 which restricts the entrance for iodobenzene to the active site.[255] In case of fluorobenzene, NADPH was consumed without hydroxylation of the substrate, indicating that fluorobenzene is bound within the active site and induces the electron transfer to the heme iron (low spin to high spin)[272] that enables oxygen binding. [273] Oxygen is finally reduced to hydrogen peroxide or other reactive oxygen species if a bound substrate is 'loosely fitting'[70] and no hydrogen atom is positioned for abstraction (Step IV and step VI; Figure 7).[157] Variant M2 displayed, compared to BM3 WT, significantly improved coupling

efficiencies ranging from 23 to 48 % for all investigated substrates (Table 10). Especially toluene and anisole were converted with very high coupling efficiencies for a P450 monooxygenase (~48 %).[70] Coupling efficiencies of M2 for bromo- and chlorobenzene were improved by 6- and 5-fold making P450 BM3 M2 a very efficient catalyst for the conversion of substituted benzenes. Table 10 reveals that product formation rates and coupling efficiencies differ significantly in contrast to the determined regioselectivity values. Main differences between the substituents (OCH₃, CH₃, Cl, Br, I, F) are steric demands as well as varying electronegativity.[258] The so called π - π -interactions between residue F87 and monosubstituted benzenes can occur in four orientations: (a) π - π stacked, (b) dimer, (c) T-shaped, and (d) inverse T-shaped.[259] Depending on the binding state the distance between two aromatic rings can vary from 3.5 (b) to 6.0 Å (d) leading to interaction energies ranging from -5.38 (a) to -0.88 kcal mol⁻¹ as reported for benzene and hexafluorobenzene dimers.[259] The high regioselectivity of BM3 and variant M2 is a first indication that the selected benzene substrates have a strongly preferred orientation for *o*-hydroxylation. Docking of anisole into the active site of P450 BM3 yielded only the T-shape orientation (Figure 41 A and B) as preferential binding mode with binding energies (Gibbs free energy of binding; ΔG_{bind}) of -4.98 kcal mol⁻¹ (A) and -5.46 kcal mol⁻¹ (B).

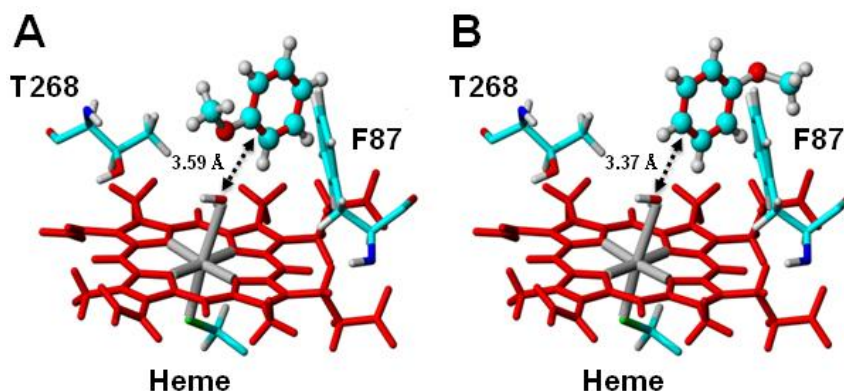


Figure 41. Visualization of P450 BM3's active site (PDB: 1BU7)[238] with the docked anisole substrate (ball and stick representation) in T-shape orientation and highest interaction energies (**A**: -4.98 kcal mol⁻¹; **B** -5.46 kcal mol⁻¹).[109, 237] Conformation **A** suggests hydroxylation at *o*-position; conformation **B** suggests hydroxylation preferentially at *p*-position. The heme center, the residues T268 (O₂ activation)[179] and F87 (indispensable for aromatic hydroxylation)[69] are depicted in stick representation. The distance (in Å) between the closest C-atoms of anisole to the water ligand bound to the heme-iron are indicated with arrows. The Figure was taken from Dennig et al.[68]

The two orientations (inversion of the OCH₃-group) with highest binding energies, suggest an interaction of anisole with F87 in T-shape orientation.[68, 259] Conformation (A) would expose the *o*-position directly towards the reactive oxygen species which is located between T268 and the heme center yielding the *o*-phenol, in this case guaiacol.[154, 246] Conformation (B) in Figure 42 would preferentially yield *p*- and *m*-hydroxylated phenols that were produced in minor amounts (*p*-methoxyphenol) (Table 10). The catalytic results from hydroxylation of selected substrates support the hypothesis that conformation A (Figure 41) is the preferred orientation mode for P450 BM3 WT and variant M2. Despite the well agreement of modeling and experimental data, the docking studies

neglect dynamics of the binding pocket and heme during hydroxylation reactions.[70] The following correlation between atom size, electronegativity and reactivity of the halogenated substituents was experimentally observed: M2 activity: Cl > Br > I (Table 10). Fast chlorobenzene hydroxylation could be attributed to an electron withdrawing effect which could promote faster formation of the epoxide intermediate.[154] Electron density donating substituents such as iodine tend to stabilize the epoxide intermediate, therefore re-aromatization could be slower according to the proposed mechanism.[274] In addition, hydroxylation by P450 BM3 M2 leads to dehalogenation of iodobenzene to phenol (51.1 % GC area), likely due to the length and weakness of the C-I bond.[275, 276]

3.3.5 Summary and Conclusion

In summary, this is the first report on direct hydroxylation of halogenated benzenes with nearly perfect regioselectivity for chloro-, bromo- and iodobenzene by an engineered P450 BM3 variant (M2) and the first P450 hydroxylation of iodobenzene. All phenols were produced at RT, in water with molecular oxygen and without co-solvent saving material and costs in an entirely green process for *o*-phenol synthesis. The reported phenols are important synthons and direct hydroxylation offers novel options in the retro synthesis, for instance, of antibiotics such as vancomycin. Notably, the engineered M2 variant has excellent selectivity (>99 %) and for a P450 enzyme excellent activity (e.g. anisole hydroxylation: 19.5 U mg_{P450}⁻¹). The results allow a deeper insight into P450 catalyzed aromatic hydroxylation, but further structural studies are necessary to decipher binding, orientation and potential π - π interactions in the active site of P450 BM3.

3.4 Hydroxylation of Trimethylbenzenes with P450 BM3 Wild-Type and Variants

3.4.1 Abstract

In this part of the thesis, hydroxylation of mesitylene and pseudocumene by P450 BM3 is investigated. Mesitylene is preferentially hydroxylated at the aromatic ring (up to 90 %), whereas hydroxylation of pseudocumene is less selective with formation of benzylalcohols in excess (65 %) employing the WT enzyme. Both substrates are hydroxylated at very low rates (PFRs of maximum 17 min⁻¹) by variant M2 (R47S, Y51W, I401M). In order to improve the catalytic performance of variant M2, a SSM library at position A330 (template M2) was screened with pseudocumene as substrate. Screening resulted in a new variant M3 (R47S, Y51W, A330F, I401M) that displayed improved hydroxylation activity (PFR 226 min⁻¹; pseudocumene as substrate) as well as higher coupling efficiency (6 fold; mesitylene as substrate) for both substrates. The chemoselectivity of the identified variant is inverted towards aromatic hydroxylation of pseudocumene (60.3 %) compared to the WT. Docking of both substrates into a new homology model structure of variant M3 revealed that the exchange of alanine to phenylalanine generates additional aromatic interactions in the active site, providing a stronger

binding of both investigated substrates and resulting in a 231-fold higher PFR of 2,4,6-trimethylphenol relative to the P450 BM3 WT.

3.4.2 Introduction

Aromatic hydroxylation of trimethylbenzenes (TMB) is of high interest for the production of vitamins and educts for C-C-coupling reactions.[277, 278] The respective benzylalcohols and benzaldehydes find broad application in food industry (almond flavor) or as synthons for C-C coupling reactions.[279] In particular 2,3,6- and 2,3,5-trimethylphenol (TMP) are of high interest in synthetic applications for the production of vitamin E (α -tocopherols) (Figure 42 A and B).[244]

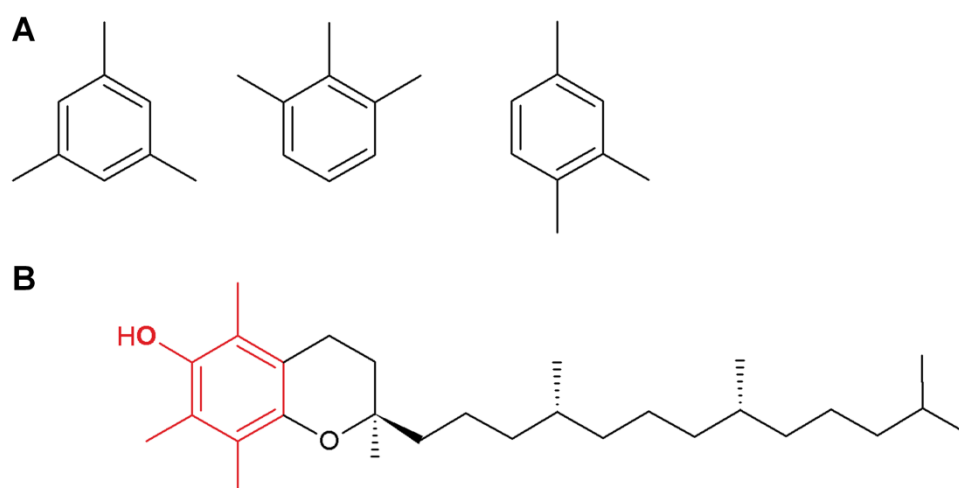


Figure 42. A: The three isomers of trimethylbenzene (TMB). From left to right: mesitylene (1,3,5-trimethylbenzene), hemellitol (1,2,3-trimethylbenzene; not used in this study) and pseudocumene (1,2,4-trimethylbenzene). **B:** Chemical structure of vitamin E (α -tocopherol).[244] The substructure of 2,3,6-TMP is highlighted in red.

Chemical synthesis of TMPs is commonly performed using phenol as educt.[280] Chemo- and regioselective hydroxylation of TMBs bonds is challenging due to difficulties in sterical control of the benzene substrate. Therefore chemical hydroxylation of e.g. pseudocumene is rather unattractive. [71, 201] Besides the known challenges in chemical O_2 -activation and hydroxylation,[201] a non-selective hydroxylation of pseudocumene would require increased purification efforts (Figure 43).

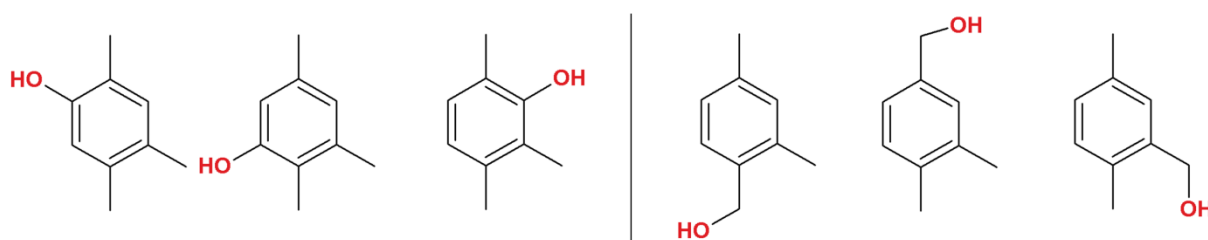


Figure 43. A non-selective hydroxylation of pseudocumene with P450 BM3 could generate six isomeric products: three phenols (left) and three benzylalcohols (right). Products names from left to right: 2,4,5-TMP; 2,3,5-TMP; 2,3,6-TMP; 2,4-DMBA; 3,4-DMBA; 2,5-DMBA. TMP = Trimethylphenol; DMBA = Dimethylbenzylalcohol.

Biocatalysis with P450 BM3 offers attractive possibilities as shown for regioselective hydroxylation of toluene and xylenes[68-70, 146, 156, 180, 206]. TMBs comprise an additional CH_3 -group that increases sterical demands and requires a catalyst with higher selectivity especially for pseudocumene hydroxylation. In chapters 3.1 and 3.3 of this thesis, P450 BM3 was applied to rather ‘symmetrical’ molecules such as *p*-xylene[69] and monosubstituted benzenes that were hydroxylated with high selectivity on the aromatic ring.[68] The studies for the hydroxylation of *o*- and *m*-xylene with P450 BM3 indicated decreased regio- and chemoselectivity compared to the *p*-xylene isomer[69, 156] (relative GC peak area: 55.9 % 2-methylbenzylalcohol; 36.7 % 2,3-di-methylphenol; 7.4 % 3,4-dimethylphenol; conversions of *o*-xylene done with GDH regeneration system for 24 h; Figure 32; section 3.1). In addition to the respective phenols also a significant amount of benzylalcohol was produced through α -hydroxylation.

Considering the industrial importance of the phenolic hydroxylation products, this work should investigate how P450 directs regio- and chemoselectivity during TMB hydroxylation. The most active variant for aromatic hydroxylation of benzenes, variant M2 (section 3.1)[68, 69, 156] was used for characterization and conversion of TMBs, as well as the P450 BM3 WT as reference. Recent reports show that mutating position A330 leads to increased activity and coupling efficiency for hydroxylation of benzene substrates.[156, 180, 206] Whitehouse et al. performed one SDM to substitute alanine by proline (A330P), a substitution that impact the active site by relocation of residue 329 generating a catalyst with improved catalytic performance for aromatic hydroxylation of toluene (4.8-fold increased coupling efficiency vs. WT).[180] In this chapter a SSM at position 330 was performed for the first time, followed by screening for improved catalytic activity and production of the two target products 2,3,6- and 2,3,5-DMP (Figure 43). Application of the colorimetric and selective 4-AAP assay allows detection of the two target phenols (section 2.10.3).[228] SSM at position A330 provides a deeper insight into what might be the optimal amino acid exchange for the selected substrates.[70] In addition to the major objective, the production of 2,3,5- and 2,3,6-TMP, the obtained experimental data will be used for analysis of variants at molecular level with focus on sterical demand and symmetry of substrates.

3.4.3 Experimental

The kinetic characterization was done using purified P450 protein at concentrations of 0.3 μM (WT) and 0.1 μM (variants). The activity of the P450 monooxygenases was measured as initial NADPH oxidation rates at 340 nm in 5 mL quartz glass cuvettes, containing 10 mM of substrate (pseudocumene or mesitylene) and 1.5 % DMSO. The 4-AAP assay was performed as described in material and methods (section 2.10.3). Regioselectivity, product yields and total turnover number (TTN) were determined using a GDH regeneration system for continuous supply of the NADPH cofactor (section 2.10.9.3).[164] The 4 mL reaction mixture was assembled containing 0.3 μM P450 BM3 WT or 0.1 μM variant, 1.5 % DMSO (V/V), 10 mM substrate, 30 mM glucose, 200 μM NADPH, 1200 U mL^{-1} catalase and 12 U GDH. Products were extracted after 24 h reaction time by two phase extraction prior to analysis on GC-FID (section 2.12.2). Detailed protocols for extraction and GC-FID analysis of substrates and products can be found in the appendix section (Table 29). Calibration curves for all theoretical products (phenols and benzylalcohols) were prepared using commercially available standards.

3.4.4 Results

Tsvetan Kardashliev (Institute of Biotechnology, RWTH Aachen University) performed the docking of substrates and contributed in parts to the collection of catalytic data. Marco Grull assisted in catalytic characterization of P450 BM3 variants during his bachelor thesis project at the Institute of Biotechnology (RWTH Aachen University).

3.4.4.1 Screening and Selection of Improved Variants for Conversion of Pseudocumene

The SSM library at position A330 using the best variant M2 (R47S, Y51W, I401M)[68, 69] as template was screened to identify a P450 BM3 variant that is capable of hydroxylating pseudocumene preferentially on the aromatic ring. Two MTPs (180 clones; 99.9 % coverage)[127] containing pseudocumene as substrate were screened with the NADPH depletion assay and the 4-AAP detection system (section 2.10) for production of phenols. Several variants with improved NADPH oxidation rates were identified as indicated in Figure 44.

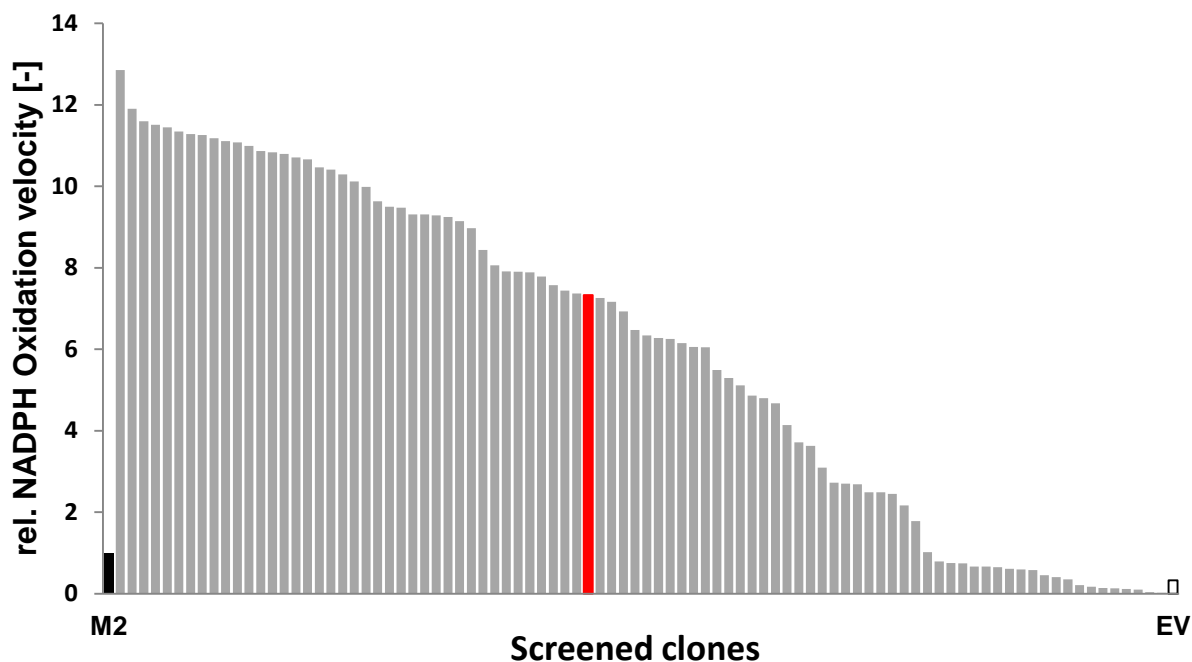


Figure 44. Screening of SSM library A330 (NNK degeneration) with the NADPH depletion assay employing pseudocumene as substrate. Variant M2 was used as template for mutagenesis.[68, 69] Activity of variant M2 (black bar) and EV (white bar with black frame) is shown as average activity of three wells (section 2.8.6). The later characterized variant M3 is highlighted as red bar.

Variant M2, used as template in this work, showed very low activity with pseudocumene as substrate (relative NADPH oxidation velocity; black bar; Figure 44). Around 20 % of the clones in the library were counted as inactive, whereas the largest fraction of clones (>70 %) displayed more than 2-fold increased NADPH depletion rates compared to variant M2. The variants with highest activity indicated 10 to 12 fold higher NADPH oxidation rates compared to variant M2 (Figure 44). In order to identify clones that produce specifically more 2,3,5- or 2,3,6-TMP the 4-AAP assay was performed after full depletion of NADPH (section 2.10.3).[69] Figure 45 shows the result of the 4-AAP assay in MTP measured after 30 min incubation time. Compared to the screening with *p*-xylene in the previous chapter (Figure 27, section 3.1.4) the color intensity was much lower and only a few wells scored positive by visual inspection (Figure 45; MTP at upper right corner).

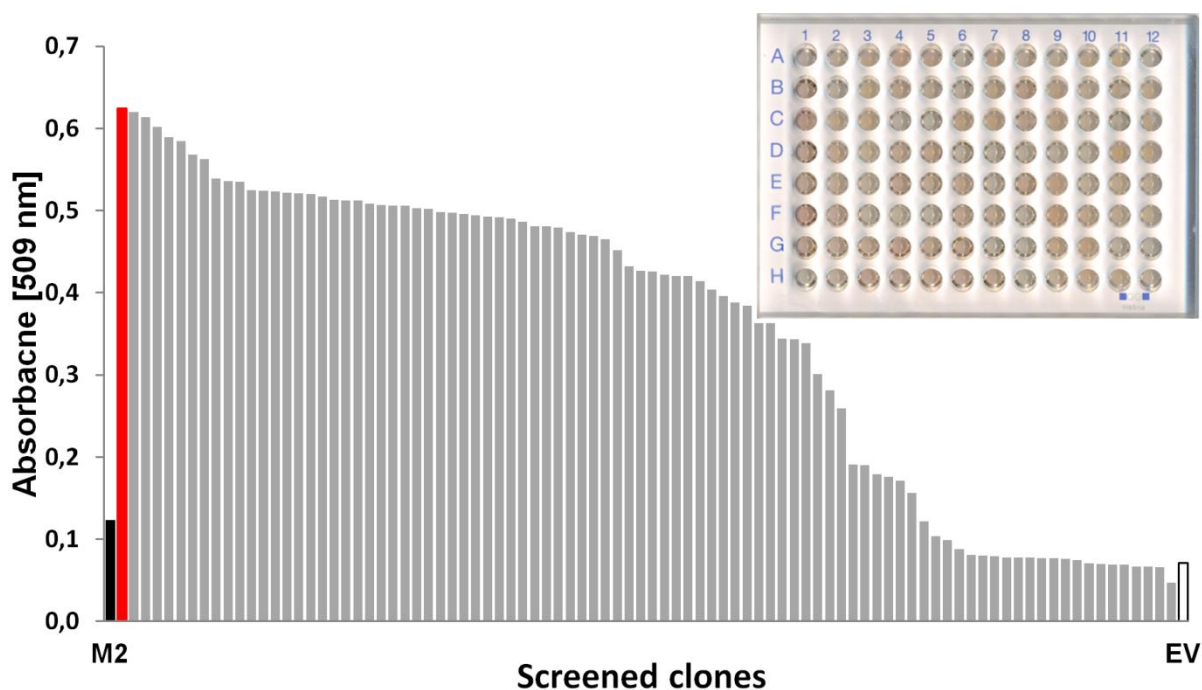


Figure 45. Screening of SSM library A330 (NNK degeneracy) with the 4-AAP assay and pseudocumene as substrate.[69, 228] Variant M2 was used as template for mutagenesis. Activity of variant M2 (black bar) and EV (white bar with black frame) is shown as average activity of three wells (section 2.8.6). The later characterized variant M3 is highlighted as red bar. The MTP after 30 min color development of the 4-AAP assay is shown in the upper right corner. Variant M2 does not show significant color formation (wells A1, D6 and H12).

Several variants (Figure 45; grey bars) displayed up to 5-fold higher absorbance than the M2 variant (black bar), which has a slightly higher absorbance than the EV control (white bar with black frame). Although the measured absorbance values are comparably high for the best variants (> 0.5), the visible color formation is low when comparing to the hydroxylation of *p*-xylene where an absorbance of 0.5 leads to a deep red color formation in the wells (see Figure 26 in section 3.1.4 for comparison). During measurement with pseudocumene the formation of emulsions was visible. The data from 180 clones was evaluated according to the description in material and methods section of this thesis (Figure 21 and 22, section 2.10.4). The best hits (selection criteria: a) visible color formation; b) improved NADPH oxidation; c) highest absorbance in the 4-AAP assay) were used for a statistically significant rescreening in triplicate (section 2.10.5). Especially, clones with distinct red color formation were selected since they indicate formation of the target phenolic products. Three clones (P12G7, P12F10 and P11E9; P11E9 is highlighted as red bar in Figure 44 and 45) were investigated in more detail using CO normalized cell free lysate.[141, 155] All three variants displayed 2- to 3-fold improved formation of phenolic products in the 4-AAP assay compared to the variant M2 (Figure 46).

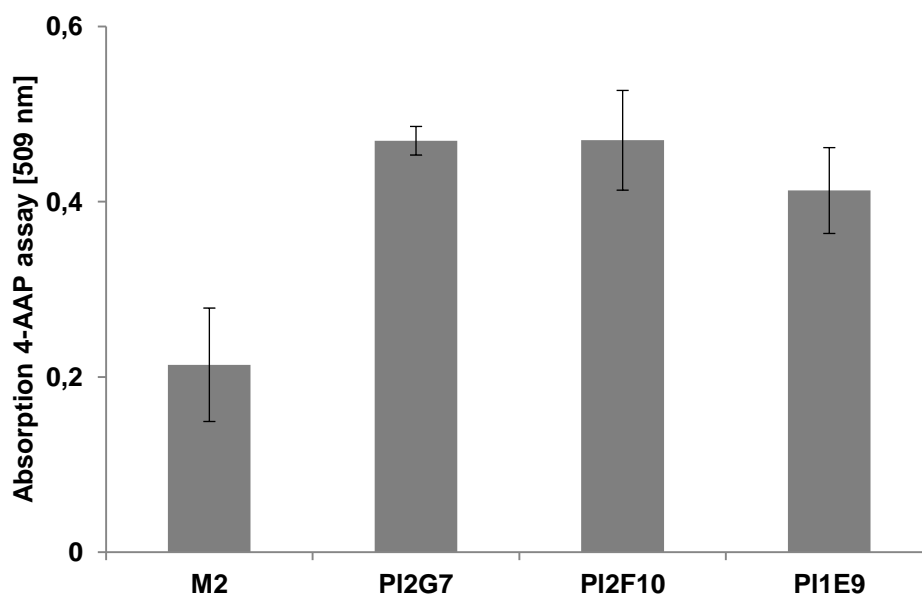


Figure 46. Rescreening of three selected clones from SSM library A330 (template M2) with the 4-AAP assay (section 2.10.3).[69, 228] The P450 BM3 concentration was normalized using CO-titration (section 2.10.8). [141, 155] Activity measurements were done in triplicate to ensure statistical confidence of the obtained results.

The plasmids from PI2G7, PI2F10 and PI1E9 were extracted and sent for sequencing. All clones (PI2G7, PI2G10 and PI1E9) had alanine substituted by phenylalanine (A330F). Therefore, the clone ‘PI1E9’ was selected, named to ‘M3’ (R47S, Y51W, A330F, I401M) and used for further detailed kinetic characterizations.

3.4.4.2 Catalytic Characterization of P450 BM3 WT and Variants for Trimethylbenzene Hydroxylation

The activity of the P450 BM3 WT for conversion of pseudocumene was lowest with 22.2 min^{-1} and moderate coupling efficiency of 14.6 %, leading to a final PFR of 3.2 min^{-1} (Table 11). Variant M2 displayed a 4-fold improved catalytic activity and slightly improved coupling efficiency providing a 5-fold higher PFR than the WT. The new variant M3 is around 22-fold more active than the WT and 5.7-fold than the variant M2, which is in agreement with the results from screening in MTP format (Figure 44 and 45).

Table 11. Catalytic performance of WT, M2 and M3 for the hydroxylation of pseudocumene. ^[a]NADPH depletion rate [$\text{mol}_{\text{NADPH}} \text{ mol}_{\text{P450}}^{-1} \text{ min}^{-1}$]; ^[b]coupling efficiency [$\text{mol}_{\text{Product}} \text{ mol}_{\text{NADPH}}^{-1}$]; ^[c]product formation rate [$\text{mol}_{\text{Product}} \text{ mol}_{\text{P450}}^{-1} \text{ min}^{-1}$].

Variant	N ^[a]	C ^[b] [%]	PFR ^[c]
WT	22.2 ± 4.2	14.6 ± 2.1	3.2
M2	87.4 ± 3.3	19.4 ± 2.4	17.0
M3	499.3 ± 90.9	45.2 ± 6.0	225.7

In addition to a significantly higher activity, variant M3 has a 3-fold improved coupling efficiency compared to the WT (2.3-fold vs. M2). Overall, variant M3 represents a 70-fold more productive catalyst than the WT enzyme with a PFR of 225.7 min⁻¹. An important factor in synthetic approaches is the regioselectivity that should be, in optimal case, >99.9 % for the target product. Hydroxylation of pseudocumene revealed that P450 BM3 hydroxylates all six carbon atoms with varying ratios to produce phenols and benzylalcohols (Table 12; Figure 43).

Table 12. Regioselectivity of WT, M2 and M3 for the hydroxylation of pseudocumene. Values were recorded after 24 h reaction time with GDH regeneration system.[164]

Variant	2,4-DMBA [%]	2,5-DMBA [%]	3,4-DMBA [%]	2,3,5-TMP [%]	2,3,6-TMP [%]	2,4,5-TMP [%]
WT	40.0	20.5	2.9	9.4	4.1	23.0
M2	25.3	11.2	1.8	5.8	16.9	39.0
M3	29.1	9.4	1.1	5.5	14.0	40.9

The WT produces preferentially benzylalcohols (63.5 %), whereas 2,4-DMBA is the major product with a fraction of 40 % (Table 12). In addition, 2,5-DMBA and 2,4,5-TMP were produced at 20.5 % and 23 % respectively. The main targets, 2,3,5- and 2,3,6-TMP, were produced in minor quantities with 9.4 and 4.1 %. In contrary to the WT, the chemoselectivity of the two variants (M2 and M3) is inverted towards formation of phenolic products that in total represent 61.7 % for M2 and 60.3 % for M3 respectively (Figure 47).

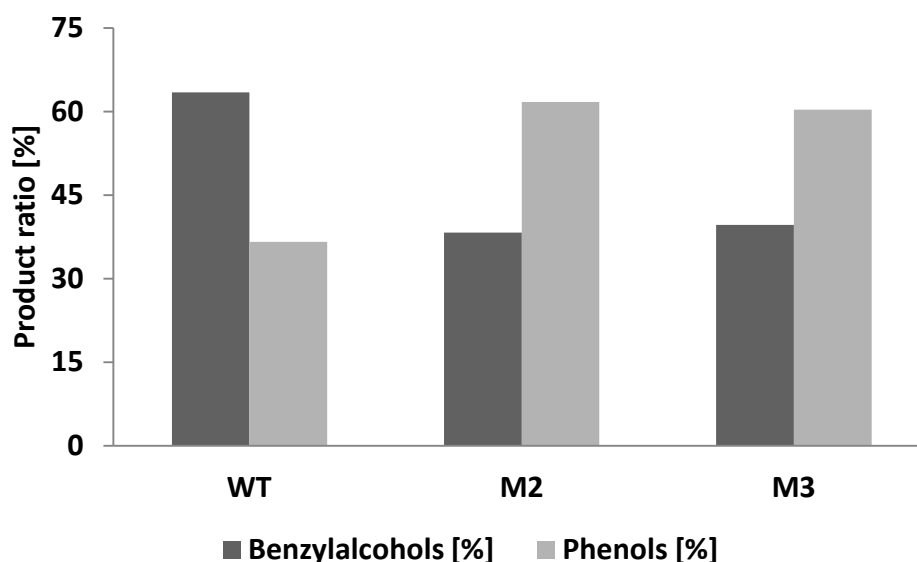


Figure 47. Chemoselectivity of P450 BM3 WT, variant M2 (R47S, Y51W, I401M) and M3 (R47S, Y51W, A330F, I401M) after 24 h conversion of pseudocumene employing a GDH cofactor regeneration system[164] and 30 mM glucose (section 2.10.9.3).[164] P450 BM3 WT predominantly produces benzylalcohols, whereas the two variants perform mainly aromatic hydroxylation.

Regioselectivity of variants M2 and M3 are comparable since both display nearly equal product patterns for the respective benzylalcohols and phenols (Figure 47 and 48). Table 12 summarizes products ratios and Figure 48 displays absolute changes in regioselectivity of the two variants compared to the WT. Variant M2 and M3 produce significantly more phenolic products, in particular 12.8 % more 2,3,6-TMP for M2 and 17.9 % more 2,4,5-TMP for M3. Altogether, 2,4,5-TMP reaches up to 40.9 % of the total product during conversions, which is a nearly 2-fold increase compared to the WT (Table 12).

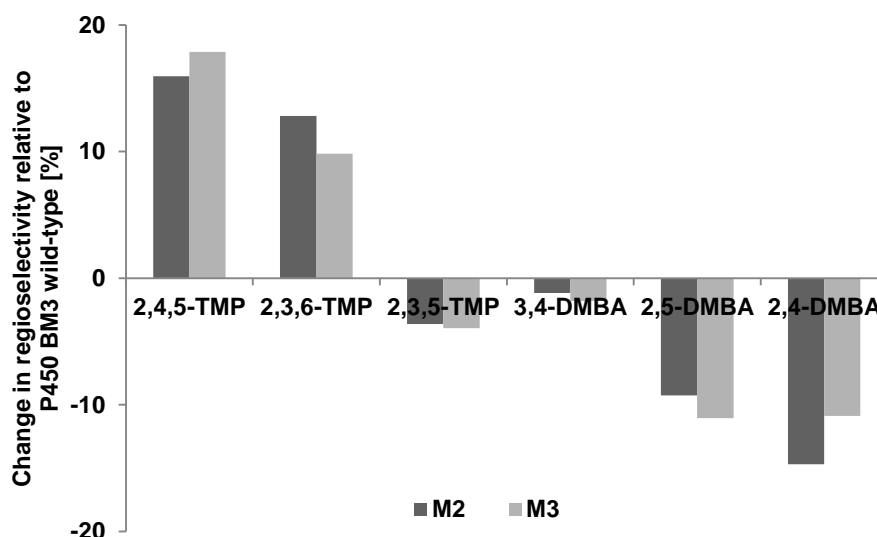


Figure 48. Determined alterations in product distributions (ratios) of variants M2 and M3 compared the P450 BM3 WT. The data was obtained from 24 h reactions employing a GDH for efficient NADPH cofactor regeneration (section 2.10.9.3).[164] The values represent absolute changes in percentage between WT and variants calculated from Table 12.

Especially the relative amount of 2,5- and 2,4-DMBA was decreased by -11 to -14 % when applying M2 and M3 in conversion reactions. Minor changes were observed for 3,4-DMBA and the target product 2,3,5-TMP that was decreased by 1.8 % for M3. Applying M2 and M3 the relative amount for the target product 2,3,6-TMP was increased by 12.8 % (M2) and 9.9 % (M3), whereas the relative quantity of the second target compound 2,3,5-TMP was decreased up to 3.9 %.

In addition to pseudocumene the two other TMB isomers, mesitylene and hemellitol, were also investigated in conversions with all three P450 variants. Hemellitol was only available at low purity (<95 %) and several product peaks were obtained during conversion reactions that could not be attributed to the target substrate after GC-MS analysis. Consequently, hemellitol was excluded from further investigations. Conversions with mesitylene as substrate revealed two distinct products peaks that were identified as 2,4,6-TMP and 3,5-DMBA, while the phenol was produced in excess (around 90 % GC area). The activity was in general lower compared to pseudocumene reaching a maximum NADPH oxidation rate of 285 min^{-1} for M3 (Table 13), which corresponds to 42 % of the activity when employing pseudocumene as substrate. Compared to the P450 BM3 WT, the variant M3 was 22-fold more active for hydroxylation of mesitylene. Coupling efficiencies for mesitylene were very low for WT and variant M2, with 2.6 % and 4.1 % respectively. The variant M3 had a 9-fold higher coupling efficiency than the WT which is also in agreement with the improvements observed during conversion of pseudocumene with this new variant (Table 12). Overall the coupling efficiencies for mesitylene were significant lower compared to those for pseudocumene.

Table 13. Kinetic characterization of P450 BM3 and variants M2 and M3 for conversion of mesitylene. ^[a]NADPH depletion rate [$\text{mol}_{\text{NADPH}} \text{mol}_{\text{P450}}^{-1} \text{min}^{-1}$]; ^[b]coupling efficiency [$\text{mol}_{\text{Product}} \text{mol}_{\text{NADPH}}^{-1}$]; ^[c]product formation rate [$\text{mol}_{\text{Product}} \text{mol}_{\text{P450}}^{-1} \text{min}^{-1}$].

Variant	N ^[a]	C ^[b] [%]	PFR ^[c]	2,4,6-TMP [%]	3,5-DMBA [%]
WT	13 ± 2.4	2.6 ± 1.4	0.3	90	10
M2	34.6 ± 14.1	4.1 ± 2.2	1.4	87	13
M3	285.4 ± 20.4	24.3 ± 2.1	69.3	84	16

The improvements in activity and coupling efficiency make M3 a 231-fold better catalyst for mesitylene hydroxylation compared to the P450 BM3 WT and 50-fold compared to variant M2 (Table 13). The differences in regioselectivity are however minor between the WT and the two investigated variants. All three P450s preferentially hydroxylate mesitylene on the aromatic ring, whereas the WT displays highest selectivity towards 2,4,6-TMP with 90 % and around 10 % hydroxylation in α -position leading to formation of 3,5-DMBA (Table 13). Variant M2 and M3 have slightly decreased selectivity towards the aromatic ring, producing 87 % (M2) and 84 % (M3) of 2,4,6-TMP (Table 13 and Figure 49). After 24 h reaction time with a GDH cofactor regeneration system (section 2.10.9.3), the total amount of product was quantified by GC-FID. Conversion of 10 mM mesitylene with WT generated 118 μM total products corresponding to a yield of 1.2 % and a TTN of 393. Employing the engineered variants more than 3-fold increased product concentrations were obtained (M2: 373 μM ; M3: 458 μM), corresponding to a product yield of 3.7 % and a TTN 1243 for M2 and 4.6 % yield and TTN of 1526 for M3 respectively (Figure 49).

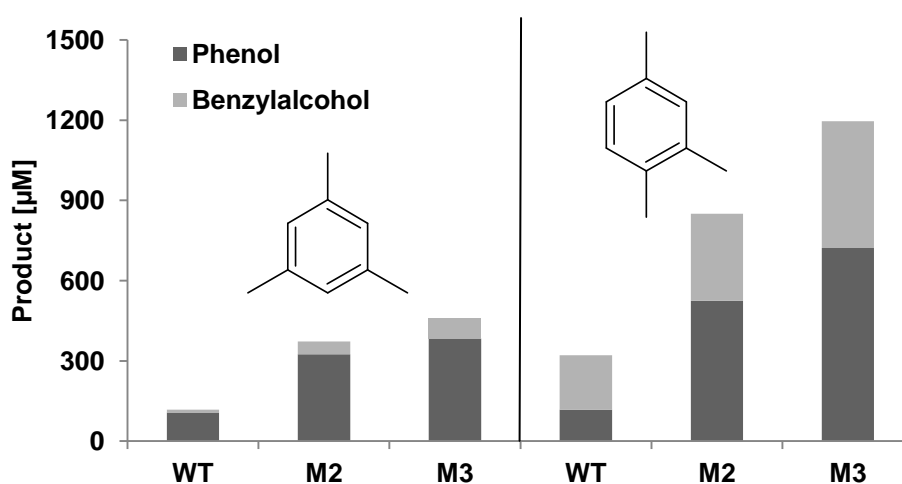


Figure 49. Absolute product concentration (sum of all benzylalcohols and phenols) after conversion of mesitylene (left) and pseudocumene (right) for 24 h applying a GDH regeneration system.[164] Ratios of benzylalcohols and phenols are shown as separation of displayed bars: upper part (benzylalcohols); lower part (phenols).

Under identical reaction conditions the three investigated P450s displayed a better performance employing pseudocumene as substrate. The WT produced the lowest amount of product (321 μM ;

3.2 % yield; 1070 TTN), whereas the variant M2 generated 2.6-fold more product (850 μM ; 8.5 % yield; 2833 TTN). Highest product concentrations were obtained with variant M3 reaching 1196 μM product (12 % yield; 3986 TTN) which corresponds to a final product concentration of 0.16 g L^{-1} .

3.4.4.3 Docking of Mesitylene and Pseudocumene into the Active Site of P450 BM3

Docking of substrates into the binding pocket of enzymes allows rationalizing the measured catalytic data of variants and to investigate the influence of mutations on altered performance parameters. [109, 177] The likeliness of strong binding (more free Gibbs energy of binding (ΔG_{bind}) in kcal mol^{-1}), [281] avoiding loose fitting of substrates while placing an H-atom in position for abstraction, should allow a higher coupling efficiency and activity in P450 catalyzed reactions. [70, 157] In chapter 3.3 the docking of anisole was in well agreement with the catalytic data, enabling further insights on how regioselectivity in P450 BM3 is governed through aromatic interactions. [68] The preference for aromatic hydroxylation of mesitylene suggest a defined binding mode for the substrate producing mainly 2,4,6-TMP (Table 13). Herein, only binding modes are shown that are in agreement with catalytic data (productive orientations) and suggested by the docking software [109, 237] which does not exclude other binding orientations of the two substrates. Two substrate orientations close to the heme iron and residue F87 are shown in Figure 50 (A and B) after docking mesitylene into the P450 BM3 WT crystal structure (1BU7). [238]

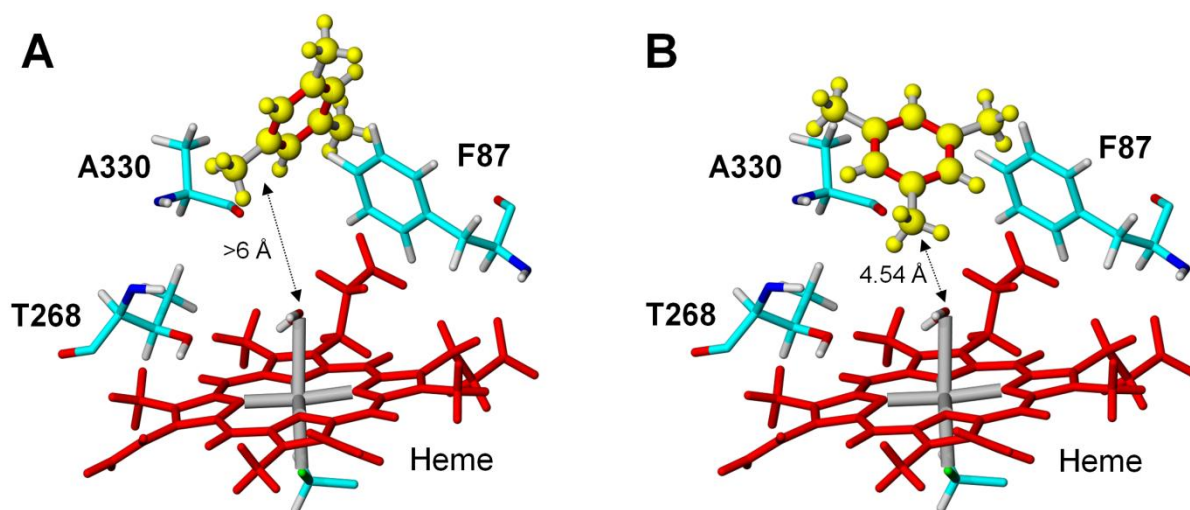


Figure 50. Docking of mesitylene into the active site of P450 BM3 WT (1BU7) [238]. **A:** Mesitylene docked in orientation allowing hydroxylation on the aromatic ring. ΔG_{bind} : $-4.93 \text{ kcal mol}^{-1}$. **B:** Mesitylene docked in orientation allowing hydroxylation preferentially in α -position; ΔG_{bind} : $-4.44 \text{ kcal mol}^{-1}$. Arrows indicate the closest distance between heme-bound water ligand and a C-atom of the substrate.

The strongest binding mode (Figure 50 A; $-4.93 \text{ kcal mol}^{-1}$) leads to a conformation where the α -C-atom and the aromatic ring atoms are in equal distance to the heme-bound water ligand ($>6 \text{ \AA}$). As second binding mode the substrate is oriented with the α -C-atom towards the heme iron with binding energies of $-4.44 \text{ kcal mol}^{-1}$ and a distance of 4.54 \AA to the heme-bound water ligand (Figure 50 B). Since no crystal structure of variant M3 is available a homology structure model was

generated by FoldX software package.[282] Despite that the generated structure has a more stable fold than the P450 BM3 WT, the binding energies for mesitylene are lower suggesting a stronger interaction of substrate and M3 than for the WT (Figure 51 A and B). The strongest binding mode (ΔG_{bind} : $-5.74 \text{ kcal mol}^{-1}$) represents a substrate orientation that allows hydroxylation at α -position, as well as hydroxylation of the aromatic ring (Figure 51 A) with the water ligand in 4.60 \AA distance to the closest C-atom. Another binding mode with a similar distance suggested also a preferential α -hydroxylation of mesitylene (Figure 51 B), however ΔG_{bind} is $0.92 \text{ kcal mol}^{-1}$ higher.[109]

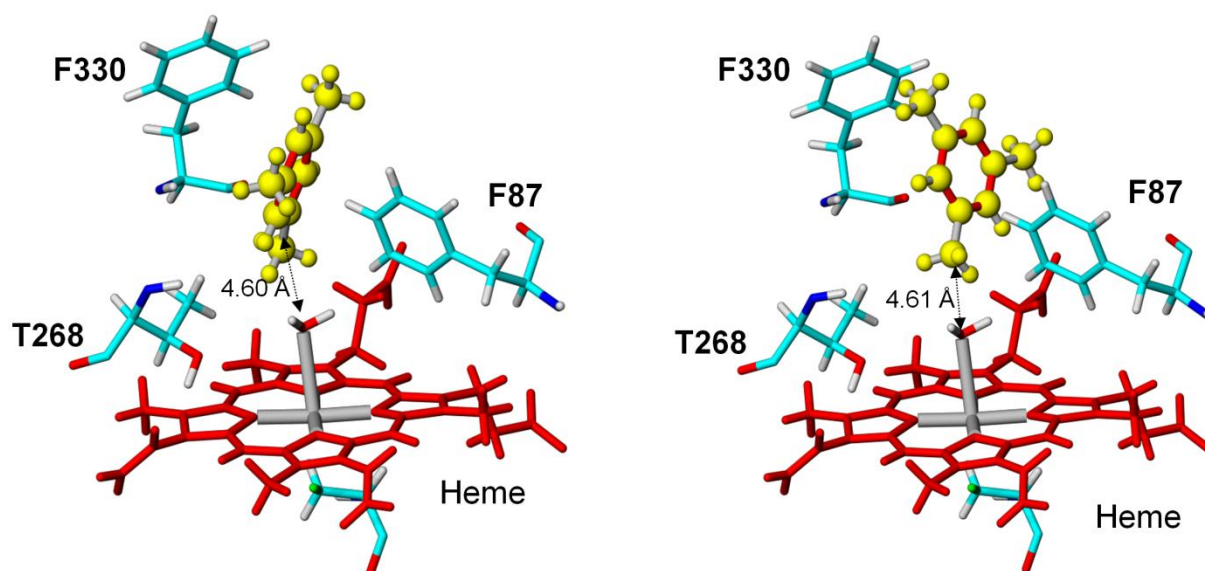


Figure 51. Docking of mesitylene into the active site of the homology model (generated with FoldX and YASARA)[109, 237] of P450 BM3 variant M3 (R47S, Y51W, A330F, I401M). **A:** Mesitylene docked in orientation allowing hydroxylation on the aromatic ring. ΔG_{bind} : $-5.74 \text{ kcal mol}^{-1}$; **B:** Mesitylene docked in orientation allowing hydroxylation preferentially in α -position. ΔG_{bind} : $-4.82 \text{ kcal mol}^{-1}$. Arrows indicate the closest distance between heme-bound water ligand and a C-atom of the substrate.

Orientation of phenylalanine residue (F330) towards the substrate binding pocket would allow additional aromatic interactions with the mesitylene substrate. The WT amino acid residue (A330) seems not to interact directly with the substrate since the functional group (CH_3) is not pointing towards the heme iron or the substrate (Figure 50).

Pseudocumene is converted to all six potential products by the P450 BM3 WT, variant M2 and M3, all with altered regio- or chemoselectivity (Figure 48; Table 12). The substrate was also docked into the WT crystal structure (PDB: 1BU7)[238] as well as into the homology model structure of variant M3. Docking of pseudocumene into the active site revealed several binding orientations. Herein, only the substrate orientations with lowest binding energies and productive binding mode are shown. For pseudocumene docked into the WT three binding modes were obtained with strong binding energies (A: $-5.60 \text{ kcal mol}^{-1}$; B: $-5.53 \text{ kcal mol}^{-1}$; C: $5.37 \text{ kcal mol}^{-1}$; Figure 52) and close distance of the target C-atom to the water bound ligand (A: 3.37 \AA ; B: 3.34 \AA ; C: 3.01 \AA ; Figure 52). The three binding orientations would lead to formation of 2,5-DMBA, 2,4-DMBA and 2,4,5-TMP that represent 83.5 % of the total product after conversion with the WT enzyme (Table 12).

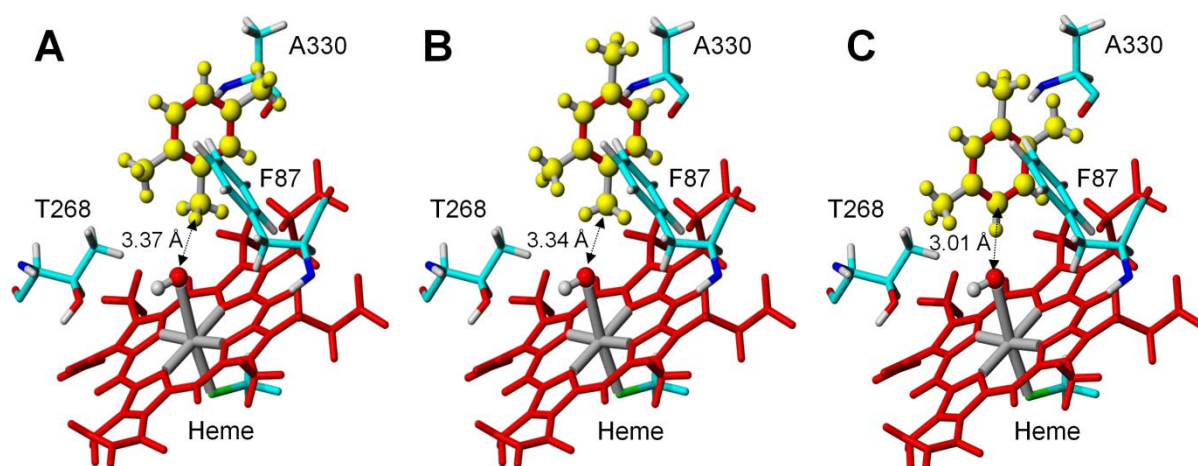


Figure 52. Docking of pseudocumene into active site of P450 BM3 WT (1BU7)[238]. **A:** Orientation allowing hydroxylation preferentially in α -position (potential product: 2,5-DMBA); ΔG_{bind} : $-5.60 \text{ kcal mol}^{-1}$. **B:** Orientation allowing hydroxylation preferentially in α -position (potential product: 2,4-DMBA); ΔG_{bind} : $-5.53 \text{ kcal mol}^{-1}$. **C:** Orientation allowing preferentially aromatic hydroxylation (potential product: 2,4,5-TMP); ΔG_{bind} : $-5.37 \text{ kcal mol}^{-1}$. Arrows indicate the closest distance between heme-bound water ligand and a C-atom of the substrate.

Docking of pseudocumene into the model structure of variant M3 proposed even stronger binding modes reaching up to $-6.33 \text{ kcal mol}^{-1}$ for binding mode A (Figure 53). This substrate orientation leads to a larger distance between the closest C-atom and the heme iron (5.04 \AA). Binding mode B ($-5.53 \text{ kcal mol}^{-1}$) is in the closest orientation of a C-atom to the heme iron (3.21 \AA) that would lead to an increased formation of 2,3,6-TMP which was confirmed for M3 compared to the WT (Table 12). Binding modes C and D (Figure 53) would also bind the substrate with larger distance to the heme-bound water ligand compared to the WT dockings (Figure 52).

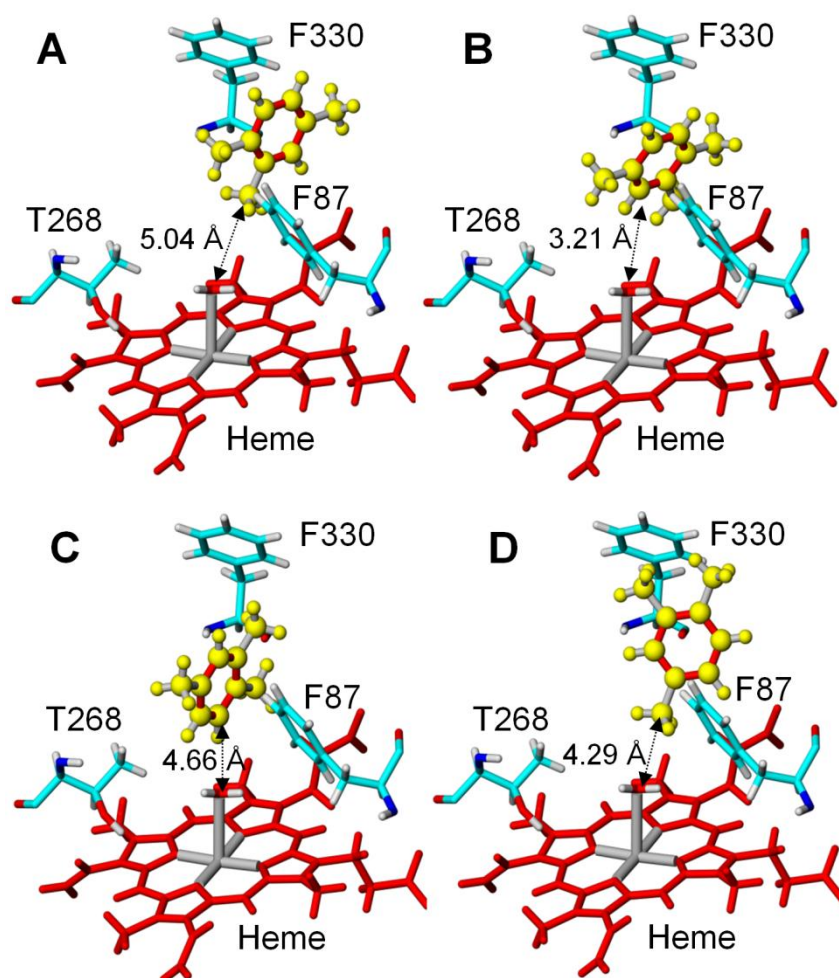


Figure 53. Following binding energies were obtained for the four different binding modes of pseudocumene docked into the active site of a structure model of variant M3. Suggested products are given together with the respective ΔG_{bind} values. **A:** ΔG_{bind} : $-6.33 \text{ kcal mol}^{-1}$; 2,5-DMBA; **B:** ΔG_{bind} : $-5.01 \text{ kcal mol}^{-1}$; 2,4,5-TMP. **C:** ΔG_{bind} : $-5.51 \text{ kcal mol}^{-1}$; 2,3,6-TMP. **D:** ΔG_{bind} : $-4.86 \text{ kcal mol}^{-1}$; 2,4-DMBA. The arrow indicates the distance between the closest C-atom of the substrate and the heme-bound water ligand.

In three out of four cases the docking of pseudocumene into variant M3 suggested a substrate orientation in close distance towards residue F330 (Figure 53 A, B and D), which leads to a more distant binding orientation of the substrate to the heme iron. Binding modes A, B and D (Figure 53) proposed that pseudocumene could be arranged in a π -X- π stacking (X = benzene substrate) between residues F330 and F87, which would allow a stronger substrate binding.

3.4.5 Discussion

In this part of the thesis the major objective was the selective aromatic hydroxylation of pseudocumene to produce 2,3,5- and 2,3,6-TMP, both important building blocks in vitamin E synthesis.[244] Hydroxylation of pseudocumene by the P450 BM3 WT protein leads to formation of all six theoretical products (Figure 43; Table 12), among which benzylalcohols represent the predominant fraction (63.4 %; Figure 47). Considering that P450 BM3 can perform hydroxylation of benzene substrates with excellent selectivity (Section 3.1 and 3.3), the generation of all six theoretical products from pseudocumene was rather unexpected, although P450 BM3 is in many cases known to

hydroxylate unselective.[151, 184, 191] To enable a first insight into selectivity of P450 BM3 towards pseudocumene hydroxylation the substrate was docked into the active site of P450 BM3 WT (Figure 52) revealing three major binding modes. As shown previously for the docking of anisole, the strongest binding is in close T-shape orientation to F87. This indicates strong π - π interactions (Figure 52 A, B and C), which are in well agreement with the reported interaction energies of benzene molecules.[259] Furthermore, the binding modes with lowest energies (highest probability) correlate very well with the obtained product patterns for P450 BM3 WT and pseudocumene as substrate (Table 12). In absolute numbers this means that 83.5 % of the total products obtained during conversion employing WT protein reactions correspond to the three major binding modes obtained from the docking experiments (Figure 52). The distance between heme-bound water ligand is close to the pseudocumene substrate (3.01 to 3.37 Å) and would allow an efficient activation of the catalytic cycle of the monooxygenase upon substrate binding (Step I, Figure 7).[70] However, the activity and coupling efficiency of the WT protein is low indicating that the substrate is hindered to induce the catalytic mechanism, rendering the P450 BM3 WT's catalytic performance low. Simulation studies and docking of substrates do not reflect neither access to the active site nor solubility limitations of substrates,[109] but these two factors play an evident role in (bio)catalytic approaches.[52] In case of P450 BM3 WT low activity is most likely caused by polar amino acid residues R47 and Y51 at the substrate channel entrance together with the low solubility of the used substrates (0.057 g L^{-1}) that is one magnitude lower compared to the monosubstituted benzenes (0.5 to 1.6 g L^{-1}).[268] A higher co-solvent concentration could increase solubility of the hydrophobic substrate in the aqueous bulk phase leading to an increased activity; nonetheless the application of co-solvents is not always beneficial as shown in section 3.2 (Figure 36 and 37). Variant M2 (R47S, Y51W, I401M) does not restrict substrate entrance and provides additional aromatic interaction at the entrance of the substrate channel (Y51W)[68, 69], altogether increasing activity and coupling efficiency for pseudocumene hydroxylation. Nevertheless, the performance of variant M2 is still very low for P450 BM3[70] that is known to hydroxylate for instance *p*-xylene with a k_{cat} of 1953 min^{-1} or an iTOF of 2427 min^{-1} for anisole (section 3.1 and 3.3).[68, 69]

Despite the significant improvements in catalytic performance during pseudocumene hydroxylation, variant M2 displays a distinct shift in chemoselectivity towards phenolic products (61.7 %) (Figure 47 and 48). No active site residue was mutated in variant M2 making rational explanation difficult. Commonly, altered sterical interactions between enzyme and substrate are responsible for changing regioselectivities and can be rationalized by structure-function relationships.[36, 283, 284] An explanation for the current results would be highly speculative since no crystal structure with a bound benzene substrate based on the WT active site configuration is available.[151] Explanations for the altered selectivity (without active site modifications) could be the increased inlet of H_2O molecules (altered hydrophobic-hydrophilic interactions),[285] increased amount of DMSO (heme blocking)[255, 256] or a joint interaction of substrate molecules within the active site (decoying).[176,

183] Since P450 BM3 is capable of hydroxylating even pyrene[189] the possibility of multiple benzene substrates in active site is not unlikely[286, 287], potentially influencing activity and chemoselectivity as shown during pseudocumene hydroxylations (Table 11). However, as final proof a crystal structure with benzene as ligand would be essential to confirm such a hypothesis. The only known crystal structure of P450 BM3 with a benzene ligand (styrene)[190] has a mutation in the active site (F87A) therefore missing important interactions for the investigated substrates in this thesis.[191]

Screening of a SSM library at position A330 for improved production of two target phenols (2,3,5- and 2,3,6-TMP) generated a variant (M3) with significantly increased catalytic activity as well as improved product formation (Table 11 and 13; Figure 49). For the first time a red color formation was visible in the 4-AAP assay employing pseudocumene as substrate, most likely due to the 2.3-fold increased coupling of the variant M3 (Table 11); the product ratio for 2,3,5- and 2,3,6-TMP was not increased compared to variant M2 (Table 12 and Figure 48). Though high absorbance values (~ 0.5 ; 509 nm) were obtained, these do not correlate with the results from hydroxylation of *p*-xylene with variant M2 (Figure 31; section 3.1). One reason might be the high substrate concentration (200 mM) that leads to formation of emulsions in the aqueous bulk phase in the MTP wells. The new and unreported variant M3 displayed a mutation of the unpolar alanine to the sterically more demanding and aromatic phenylalanine (A330F), that leads to additional aromatic interactions in the substrate binding pocket of P450 BM3 in close distance to residue F87 (Figure 53). Despite the fact that an active site residue was mutated, the influence on regioselectivity was minor for all products with the largest variation for 2,4-DMBA (+ 3.8 %) and 2,3,6-TMP (- 2.9 %) compared to variant M2 (Figure 48). On the other hand the formation of 2,4-DMBA and 2,5-DMBA are decreased by 11.1 and 10.9 % (variant M3 vs. WT) which would require an inversion of the substrate orientation in the binding pocket compared to the WT (Figure 52).

Although regioselectivity was improved for the target product 2,3,6-TMP the overall selectivity of the variant M2 and M3 is far away from the demands in synthetic applications since purification of isomeric products would be tedious. A PFR of 225.7 min^{-1} is already a remarkable improvement compared to the WT, generating around 1.2 mM (0.12 g L^{-1}) product that corresponds to a TTN of 3986 and can be regarded as a good productivity.[63, 70] Especially the coupling was increased remarkably to 45.2 % which is more than 2-fold higher than the coupling efficiency of variant M2, indicating that binding of pseudocumene is improved. A higher coupling efficiency can be attributed to the additional aromatic interaction within the active site provided by residue F330 (Figure 53). However, distances between substrate and heme iron are increased according to docking simulation results. A second substrate molecule does not fit in close distance to the heme iron. Interaction of two or more substrates in the substrate channel might be responsible for the remarkable improvements, for example in case of the 231-fold increased PFR employing mesitylene as substrate (Table 13). It is worth mentioning that docking studies do not reflect structural movements in P450 catalysis that are

well known and can reach up to 6 Å movements as shown for laurate.[288] Nonetheless, the substrate requires to be bound in close distance to the heme-bound water ligand to initiate the catalytic cycle. The mutation F330 increases activity and coupling efficiency by several fold and might be attributed to a strong stacking of the substrate between amino acid residues F87 and F330 (π -X- π ; X = substrate) (Figure 53).[289] The binding or stacking of a substrate in larger distance to the heme iron should theoretically decrease activity and coupling efficiency of the monooxygenase.[151, 157, 158] However, in this study it was the inverted case (Figure 53 and Table 11). Most likely, the generated homology model or docking of the substrate do not reflect exactly enough the interactions between substrate and protein. Undoubtedly, the introduction of phenylalanine at position 330 provides additional aromatic interactions in the active site[259] resulting in up to 70-fold higher PFR (Table 11), which comes closer to demands in synthetic applications (final PFR of 225.7 min⁻¹; variant M3 and pseudocumene). Despite the low selectivity, the hydroxylation of pseudocumene with variant M3 is more efficient, greener and faster compared to chemical hydroxylations of similar substrates.[201]

In addition to pseudocumene also the isomeric mesitylene was investigated to allow a deeper insight into the selectivity and activity of P450 BM3 for hydroxylation of this class of substrates. Mesitylene was hydroxylated with lower activities and poor coupling efficiencies employing the WT enzyme or variant M2 (<5 % for P450 BM3 WT and variant M2; Table 13). The absolute product yields were also significant lower with a TTN of 393 for WT and 1243 for M2 (Figure 49). In contrast to pseudocumene, P450 BM3 is able to perform selective aromatic hydroxylation of mesitylene whereas the WT produces 2,4,6-TMP with 90 % selectivity. The two variants have slightly decreased selectivity (Table 13). Although selectivity comes close to what is demanded in a synthetic application, the PFR values (67 min⁻¹; M3) and coupling efficiency (24.3 %) are very low indicating that the substrate is not properly inducing the P450 catalytic mechanism, most likely due to differences in orientation of mesitylene within the active site. Docking revealed that mesitylene, other than pseudocumene, is not binding in close distance to the heme iron (Figure 50 and 51) which explains significant lower activity and coupling efficiencies (Table 13). The introduction of an additional phenylalanine (F330) could also not improve distance restrictions to the heme iron. Binding modes of mesitylene in the WT structure and the structure model of M3 do not allow rational explanation for high selectivity towards the aromatic ring.

3.4.6 Summary and Conclusion

Regioselective aromatic hydroxylation of pseudocumene enables a new route to important building blocks for vitamin E synthesis, e.g. 2,3,5- and 2,3,6-TMP. The WT and engineered variant M2 have poor activity and selectivity employing pseudocumene as substrate, producing mainly benzylalcohols (P450 BM3 WT). Chemoselectivity of variant M2 could be inverted towards the production of phenol but rationalization of the obtained data remains unresolved since no active site residue was mutated. To overcome low activity, productivity and selectivity a SSM library at position A330 was screened

for increased formation of 2,3,5- and 2,3,6-TMP, which can be detected by the regioselective 4-AAP assay. Screening revealed that the substitution A330F led to a significantly increased activity (22-fold) and coupling efficiency (3-fold) resulting in a 70-fold increased PFR for pseudocumene compared to the WT. For mesitylene even a 231-fold improvement of PFR compared to the WT could be achieved. The introduced mutation, though located in the substrate binding pocket, did not alter regioselectivity of the catalyzed reaction. Based on docking results with pseudocumene and mesitylene a hypothesis was generated that additional aromatic interactions (π -X- π) improve binding and increase catalytic performance of P450 BM3. Although the engineered variants are far away from a synthetic application, the improvements in catalytic performance are remarkable employing the variant M3. In this respect, employing pseudocumene as substrate, a TTN of 3986 with a final product concentration of 0.12 g L⁻¹ was reached. Further active site engineering is required to shift the selectivity towards a preferred product, which could not be achieved so far for pseudocumene. Invaluable would be a crystal structure of WT and M3 with a bound benzene substrate to decipher and rationalize the inverted chemoselectivity and increased activities.

3.5 One Pot Synthesis of Hydroquinones from Benzene with P450 BM3

3.5.1 Abstract

P450 BM3 hydroxylates mono-substituted benzenes to *o*-phenols with excellent regioselectivity and activity. This section focuses on the estimation of key catalytic parameters in production scale synthesis: total turnover numbers (TTNs), yields and product titers. The WT and two designer monooxygenases (M2 and M3) were employed for the conversion of seven selected benzenes in presence of a GDH cofactor regeneration system. The highest product concentrations were obtained for variant M3 reaching up to 6 mM guaiacol (0.74 g L⁻¹), a TTN of 6000 and a yield of 30 % *o*-phenol from 20 mM anisole. Besides the respective *o*-phenols, additional products were detected that correspond to hydroquinones. Time- and *o*-phenol concentration-dependent formation of the hydroquinones was measured by HPLC and could be attributed to the altered substrate concentrations of benzene substrates and the pK_a values of phenolic intermediates. The formation of hydroquinones was also performed with excellent regioselectivity (>99 %) enabling overall the first enzymatic one pot synthesis of substituted hydroquinones from benzene educts.

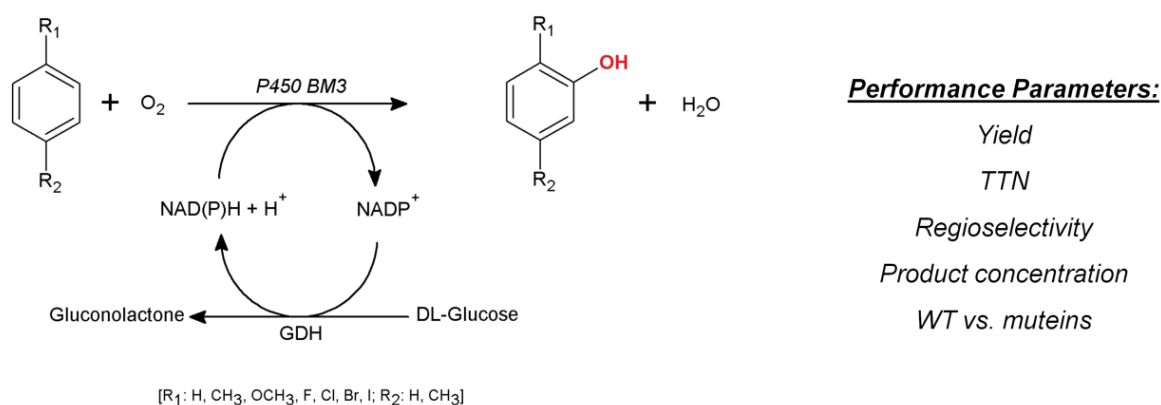
3.5.2 Introduction

In sections 3.1 and 3.3 of this thesis, the P450 BM3 WT and variants were characterized for the conversion of *p*-xylene and mono substituted benzene substrates. It was shown that the new variants have higher coupling efficiency, activity and are capable to hydroxylate substrates such as fluoro- and iodobenzene which could not be hydroxylated by the WT under the described conditions (Table 10; section 3.3.4).[68] In previous experiments the NADPH cofactor was limited within the reaction system[68, 69]; therefore only a snapshot of the catalytic performance could be recorded such as initial

activity, productivity, coupling efficiency and regioselectivity. However, enzymes and specifically P450 monooxygenases are rarely used for synthesis of bulk chemicals firstly due to the lack of operational stability[66] and mainly because of a strong competition with petrochemical sources.[148] A chemical route from substituted benzenes to phenols employing molecular oxygen, moderate temperature and atmospheric pressure is not reported.[201] Therefore, the application of an enzyme catalyst can be a highly competitive and appealing approach to fill an existing gap in synthetic chemistry.

Two major challenges are present when employing P450 monooxygenases in a biosynthetic approach: 1. Sufficient NAD(P)H cofactor supply and 2. Control over chemoselectivity. Limitation of the NAD(P)H cofactor can be overcome using efficient regeneration systems such as whole cells [77, 78, 290] or *in vitro* regeneration with supplemented dehydrogenase enzymes (section 2.10.9.3).[164, 191] The target substrates in this thesis render the application of whole cells challenging (e.g. *E. coli*) due to high toxicity at elevated concentrations.[291] *In vitro* production systems can therefore be advantageous, overcoming limitations of substrates and products, allowing easy access of substrates to the catalyst, avoiding undesired side reactions and for some processes and products a simplified purification. The formation of successive products is commonly encountered by applying a secondary water immiscible organic phase such as octanol or *n*-hexane, which could also serve as reservoir for substrates and storage for products.[51, 292] Since P450 BM3 accepts a broad spectrum of substrates[70], the ideal strategy would be to apply the substrate itself as secondary organic phase. Further issues in synthetic approaches are stability of the catalyst and varying conditions for ionic strength, pH or thermo- and (co-)solvent stability, which were not covered herein and would have to be considered in larger scale processes.

To access important catalytic parameters for a preparative synthesis, the *in vitro* GDH system[164] was selected for regeneration of the NADPH cofactor during long term reactions for 24 h (section 2.10.9.3). Scheme 9 illustrates the target reaction and a list of parameters that should be evaluated and discussed with focus on a future synthetic potential of the engineered variants in large scale synthesis.



Scheme 9. Aromatic hydroxylation of selected benzenes using a P450 BM3 monooxygenase and a glucose dehydrogenase (GDH) regeneration system[164] for production of *o*-phenols. Key parameters that should be monitored after 24 h reaction time are the product yields, TTN, regioselectivity, product concentrations and a further comparison of WT and the new designer monooxygenases that were characterized in earlier sections of this thesis.

3.5.3 Experimental

Conversion of selected substrates with the GDH cofactor regeneration system[164] was performed as described in material and methods section of this thesis (section 2.10.9.3). Conversions for estimation of TTNs, yields and product concentrations over 24 h reaction time were done employing 1 μ M P450 protein (WT, M2 and M3) from cell free lysate and three different substrate concentrations (1, 2 and 20 mM). DMSO was only used in case of 1 and 2 mM substrate concentrations to ensure accurate supplementation of the substrates; all other reactions were performed without additional co-solvent (section 3.3).[69] After 24 h stirring at RT the products were extracted (section 2.11.1) using MTBE as solvent and 20 mM cyclododecanol (CDDC) as internal standard for quantification. Calibrations curves for substrates and all reaction products were prepared using commercial standards. A full list of all products and substrates can be found in the appendix section, including retention times and applied columns for separation (Table 29). Unknown products were identified by GC-MS and commercial standards were acquired (if available) to confirm and quantify the obtained compounds. Time-dependent product formation was measured by HPLC analysis employing 1 μ M purified P450 (WT and M2) protein (section 2.10.6) and 20 mM benzene substrate (Scheme 9). Analysis of phenolic products was performed by HPLC as described in section 2.10.9.4.

3.5.4 Results

The result section is separated in two parts. The first part covers batch reactions over 24 h with WT and variants at different substrate concentrations to describe product yields, regioselectivity, TTN and product concentrations. Only results are shown that were reproducible at minimum in triplicate with less than 15 % deviation. The second part focuses on the time dependent product formation by P450 BM3 variant M2. Nina Lülendorf and Kübra Kömürçü assisted in catalytic characterization of P450 BM3 variants within the frame of their master thesis project at the Institute of Biotechnology (RWTH Aachen University).

3.5.4.1 Conversion of Toluene

Key performance parameters (product yields, selectivity, product concentrations and TTN) for preparative synthesis of *o*-phenols were recorded from small scale batches of 4 mL after 24 h reaction time. Toluene was hydroxylated by all three enzymes (WT, M2 and M3) with high regioselectivity and activity to *o*-cresol as described before (Table 10; section 3.3.4).[68, 69, 180] Extraction and analysis of the products by GC-FID revealed that 1 and 2 mM toluene are fully depleted after 24 h corresponding to a conversion of >99 %. The yields however, did not match with the conversions. The WT displayed lowest yields with 5 % and 23 % *o*-cresol from 1 and 2 mM toluene, respectively (Table 14). Product yields for the variants M2 and M3 are 9 and 13-fold higher at the same substrate concentrations (1 and 2 mM), leading to final product concentrations of 50 mg L⁻¹ and 70 mg L⁻¹, respectively. In addition to *o*-cresol, one additional product peak was identified as methylhydroquinone that represents the major product with 73 % of total product when applying M3 and 2 mM toluene as substrate (Table 14). The appearance of methylhydroquinone proves that *o*-cresol is preferentially hydroxylated for a second time in *p*-position to the firstly introduced hydroxyl group. In a previous characterization with limited amount of cofactor and 20 mM substrate the methylhydroquinone was not detected (section 3.1 and 3.3).[68, 180] TTN values increase for all three investigated P450 when employing 2 mM substrate compared to 1 mM indicating a substrate limitation. The P450 BM3 WT displayed lowest TTN (487) while the variant M2 reached a TTN of 1954. Variant M3 displayed the highest TTN (2182) and methylhydroquinone represented the major product for this variant (27 % *o*-cresol; 73 % methylhydroquinone). Compared to M3 the variant M2 generates more *o*-cresol from 2 mM toluene (41 % *o*-cresol; 59 % methylhydroquinone).

Table 14. Catalytic performance of P450 BM3 WT, variant M2 and M3 at three different concentrations of toluene (1, 2 and 20 mM) and excess of NADPH cofactor provided by a GDH cofactor regeneration system (section 2.10.9.3). [164] Products were extracted after 24 h and analyzed using commercial standards.

Variant (1 mM Substrate)	Yield [%]	Ph : HQu [%]	Product concentration [g L ⁻¹]	TTN
WT	5	100 : 0	0.01	48
M2	44	58 : 42	0.05	623
M3	62	33 : 67	0.07	1031
Variant (2 mM Substrate)	Yield [%]	Ph : HQu [%]	Product concentration [g L ⁻¹]	TTN
WT	23	96 : 4	0.05	487
M2	62	41 : 59	0.14	1954
M3	63	27 : 73	0.15	2182

Variant (20 mM Substrate)	Yield [%]	Ph : HQu [%]	Product concentration [g L ⁻¹]	TTN
WT	16	80 : 20	0.35	3789
M2	21	69 : 31	0.47	5421
M3	34	74 : 26	0.75	8462

Ph = *o*-phenol; HQu = hydroquinone; TTN = total turnover number

When employing 20 mM toluene, large amounts of substrate remain unconverted (30 to 50 %) resulting in lower yields as compared to lower substrate concentrations. However, the total amount of formed product is increased; the WT reaches a final product concentration of 0.35 g L⁻¹ corresponding to a TTN of 3789 (Table 14). Variant M2 displays improved product concentrations as well as higher TTN values exceeding 5000, whereas the highest amount of product was obtained employing variant M3 which could reach 0.75 g L⁻¹ product with a TTN of 8462. The product ratio for *o*-cresol and methylhydroquinone is shifted towards the phenolic product compared to lower substrate concentrations (WT 80:20; M2 69:31; M3 74:26). It is worth mentioning that the formation of the second product occurs with excellent regioselectivity (>99 %) and no catechol or resorcinol were detected. In addition to the methylhydroquinone standard also resorcinol and catechol were employed as analytical standards under identical conditions.

3.5.4.2 Conversion of Anisole

Anisole conversions were performed with 2, 4 and 20 mM substrate which is a variation to the other reaction setups in this chapter. Also for anisole a second product was identified corresponding to the respective methoxyhydroquinone. Product ratios of guaiacol and methoxyhydroquinone in the reactions with WT and M3 were comparable to those measured for toluene, whereas M2 produces more of the respective *o*-phenol (Table 14). Conversion of 20 mM anisole employing variant M2 yielded the highest product concentrations of 1.11 g L⁻¹ (Table 15), corresponding to a TTN of 10162 and a good yield of 44 %. Also the WT and M2 display very high TTN values (both beyond 7000) and product concentrations of 0.79 g L⁻¹ and 0.83 g L⁻¹, respectively.

Table 15. Catalytic performance of P450 BM3 WT, variant M2 and M3 at three different concentrations of anisole (2, 4 and 20 mM) and excess NADPH cofactor provided by a GDH cofactor regeneration system (section 2.10.9.3).[164] Products were extracted after 24 h and analyzed using commercial standards.

Variant (2 mM Substrate)	Yield [%]	Ph : HQu [%]	Product concentration [g L ⁻¹]	TTN
WT	1	100 : 0	0.02	174
M2	46	71 : 29	0.11	1172
M3	96	32 : 68	0.26	3244

Variant (4 mM Substrate)	Yield [%]	Ph : HQu [%]	Product concentration [g L ⁻¹]	TTN
WT	3	100 : 0	0.02	135
M2	45	50 : 50	0.35	4026
M3	65	35 : 65	0.35	4255

Variant (20 mM Substrate)	Yield [%]	Ph : HQu [%]	Product concentration [g L ⁻¹]	TTN
WT	31	80 : 20	0.79	7430
M2	33	81 : 19	0.83	7733
M3	44	84 : 16	1.11	10162

Ph = *o*-phenol; HQu = hydroquinone; TTN = total turnover number

The main difference to the conversion of toluene is that M3 can yield beyond 95 % product from 2 mM anisole and the investigated P450s tend to produce preferentially guaiacol over methoxyhydroquinone. Moreover, the WT enzyme displays very high product formation at 20 mM reaching nearly the performance of the M2 variant which was not the case for toluene.

3.5.4.3 Conversion of *p*-Xylene

p-Xylene was the first substrate investigated in this thesis and variant M2 was specifically engineered towards the conversion of this compound (section 3.1).[69] The WT conversions are not shown. Due to the fact that the commercial standard for 2,5-dimethylhydroquinone (2,5-DMHQ) was not available the relative GC-areas compared to 2,5-DMP are given in Table 16. During all measurements with the variants (M2 and M3) employing 1 and 2 mM substrate the dominant product was always 2,5-DMHQ with a GC-area of 70 to 88 % (Table 16). Product yields for 2,5-DMP are lower due to the fact that M2 and M3 perform also secondary hydroxylation of *p*-xylene which is not included in the yield calculations. The GC-MS fragmentation of 2,5-DMHQ can be found in the appendix section of this thesis. In addition, a third minor peak (< 5 % GC area) was identified that corresponds to the oxidized form of 2,5-DMHQ, the respective 2,5-dimethylquinone (see appendix Figure 85 for GC-MS fragmentation). The 2,5-DMP product did not exceed 0.046 g L⁻¹ for substrate concentrations of 1 and 2 mM (M2 and M3).

Table 16. Catalytic performance of P450 BM3 WT, variant M2 and M3 at three different concentrations of *p*-xylene (1, 2 and 20 mM) and excess NADPH cofactor provided by a GDH cofactor regeneration system (section 2.10.9.3).[164] Products were extracted after 24 h and analyzed using commercial standards.

Variant (1 mM Substrate)	Yield* [%]	2,5-DMP [mM]	Hydroquinone [% GC-area]	<i>o</i> -Phenol [g L ⁻¹]
WT	n.r.	n.r.	n.r.	n.r.
M2	6	0.06	88	0.007
M3	10	0.10	84	0.013

Variant (2 mM Substrate)	Yield* [%]	2,5-DMP [mM]	Hydroquinone [% GC-area]	<i>o</i> -Phenol [g L ⁻¹]
WT	n.r.	n.r.	n.r.	n.r.
M2	12	0.24	80	0.03
M3	19	0.37	71	0.046

Variant (20 mM Substrate)	Yield* [%]	2,5-DMP [mM]	Hydroquinone [% GC-area]	<i>o</i> -Phenol [g L ⁻¹]
WT	n.r.	n.r.	n.r.	n.r.
M2	21	5.6	18	0.68
M3	34	4.6	19	0.56

*Yields were calculated for the 2,5-DMP product since no commercial standard of 2,5-DMHQ was available.

Employing a substrate concentration of 20 mM significantly increases product concentrations for M2 and M3, whereas M2 produces 17.5 % (5.6 vs. 4.6 mM 2,5-DMP) more 2,5-DMP. As shown before during conversions of toluene and anisole the ratio of the phenolic product increases at the highest applied substrate concentration of 20 mM. Finally, a product concentration of 0.68 g L⁻¹ 2,5-DMP was obtained for variant M2, corresponding to a TTN of 5600 not considering the second hydroxylation to the hydroquinone product.

3.5.4.4 Conversion of Chlorobenzene

Chlorobenzene was converted with excellent regioselectivity to 2-chlorophenol (>99 %; Table 17)[68] and in this chapter in a second hydroxylation step to the respective hydroquinone (chlorohydroquinone), which was performed with >99 % regioselectivity. The total product concentration did not exceed 0.26 g L⁻¹ (WT and M3; both 20 mM substrate) (Table 17). Compared to anisole and toluene the formation of chlorohydroquinone was dominant over the 2-chlorophenol (except for WT at 1 mM substrate concentration) but slightly decreased at higher substrate concentrations. Employing 1 mM substrate together with the engineered variants >90 % of the product was chlorohydroquinone, corresponding to a TTN of around 1500.

Table 17. Catalytic performance of P450 BM3 WT, variant M2 and M3 at three different concentrations of chlorobenzene (1, 2 and 20 mM) and excess of NADPH cofactor provided by a GDH cofactor regeneration system[164] (section 2.10.9.3). Products were extracted after 24 h and analyzed using commercial standards.

Variant (1 mM Substrate)	Yield [%]	Ph : HQu [%]	Product concentration [g L ⁻¹]	TTN
WT	23	74 : 26	0.03	282
M2	78	7 : 93	0.11	1515
M3	78	7 : 93	0.11	1508

Variant (2 mM Substrate)	Yield [%]	Ph : HQu [%]	Product concentration [g L ⁻¹]	TTN
WT	49	43 : 57	0.13	1525
M2	65	18 : 82	0.15	2393
M3	66	19 : 81	0.19	2390

Variant (20 mM Substrate)	Yield [%]	Ph : HQu [%]	Product concentration [g L ⁻¹]	TTN
WT	9	36 : 64	0.26	3045
M2	9	24 : 76	0.24	2977
M3	9	24 : 76	0.26	3261

Ph = o-phenol; HQu = hydroquinone; TTN = total turnover number

Highest TTN values are obtained for 20 mM substrate whereas all three investigated P450s displayed nearly equal product concentrations. However, the amount of 2-chlorophenol was slightly higher for the WT (ratio 36:64) compared to variants M2 and M3 that displayed equal product ratios (24:76). Product yields for chlorobenzene conversions were lower compared to toluene and anisole when employing 20 mM substrate (see WT; 20 mM toluene; Table 14).

3.5.4.5 Conversion of Bromobenzene

Bromobenzene conversion was performed in a previous chapter (section 3.3) with excellent regioselectivity (>99 %) towards the 2-bromophenol. Once more the secondary hydroxylation to bromohydroquinone was performed with excellent selectivity comparable to the conversion of chlorobenzene (Table 17). One difference between chloro- and bromobenzene is that the ratio of phenol and hydroquinone is more on the side of the second product (for bromobenzene) reaching nearly complete selectivity with M2 and M3 (1 mM both; 2 mM only M3; Table 18). The measured TTN for bromobenzene (TTN 4877; M3; 20 mM substrate) is higher than for chlorobenzene (TTN 3261; M3; 20 mM substrate), yielding a nearly doubled product concentration of 0.54 g L⁻¹ for variant M3 with 20 mM substrate.

Table 18. Catalytic performance of P450 BM3 WT, variant M2 and M3 at three different concentrations of bromobenzene (1, 2 and 20 mM) and excess of NADPH cofactor provided by a GDH cofactor regeneration system[164] (section 2.10.9.3). Products were extracted after 24 h and analyzed using commercial standards.

Variant (1 mM Substrate)	Yield [%]	Ph : HQu [%]	Product concentration [g L ⁻¹]	TTN
WT	29	48 : 52	0.1	838
M2	81	1 : 99	0.15	1628
M3	70	2 : 98	0.13	1406

Variant (2 mM Substrate)	Yield [%]	Ph : HQu [%]	Product concentration [g L ⁻¹]	TTN
WT	40	35 : 65	0.22	2009
M2	54	24 : 76	0.26	2478
M3	75	1 : 99	0.28	3003

Variant (20 mM Substrate)	Yield [%]	Ph : HQu [%]	Product concentration [g L ⁻¹]	TTN
WT	9	36 : 64	0.34	3019
M2	9	27 : 73	0.32	2984
M3	16	36 : 64	0.54	4877

Ph = *o*-phenol; HQu = hydroquinone; TTN = total turnover number

The ratio of 2-bromophenol and bromohydroquinone is clearly towards the secondary product under all investigated concentrations and P450 variants. Altogether, the halogenated substrates display a preferential formation of hydroquinone products whereas alkyl-substituted benzenes display at 20 mM substrate concentrations a dominant formation of *o*-phenolic products. In all cases, despite the formation of a second non-isomeric reaction product, the regioselectivity for hydroxylations was higher than 99 %.

3.5.4.6 Estimation of Time Dependent Formation of Phenols and Hydroquinones

High-performance liquid chromatography (HPLC) is a widely applied method to analyze metabolites and reaction products in chemical or biological samples. Other than measurements with GC-FID or GC-MS, sample preparation for HPLC does not require an additional extraction step for product quantification. HPLC allows therefore the sampling of a large number of substrates and variants in parallel in a simple and straight forward approach. Analysis by GC-FID and GC-MS of the P450 BM3 catalyzed hydroxylation of benzene substrates under unlimited cofactor supply revealed formation of secondary products that could be identified as hydroquinones (Tables 14 to 18). This indicates that the first product becomes substrate for a second reaction step. Resulting from this observation, it is essential to measure the time-dependent product formation in order to identify the time point at which

the selectivity of the monooxygenases catalyst switches. Therefore, over 48 h samples were collected from 4 mL reaction batches (section 2.10.9.4) at defined time intervals to monitor the switch in substrate selectivity from benzene to phenol for the variant M2.

3.5.4.7 Conversion of Anisole, Toluene and *p*-Xylene with Variant M2

Time dependent product formation employing anisole as substrate revealed that in the first 8 h of the reaction solely guaiacol is produced reaching around 4 mM product (Figure 54). After 8 h reaction time and 4 mM guaiacol variant M2 starts to produce methoxyhydroquinone. The highest product concentration was reached after 24 h with 5.4 mM guaiacol and 1.4 mM methoxyhydroquinone that corresponds to a product titer of 0.67 g L⁻¹ and 0.2 g L⁻¹, respectively (Figure 54).

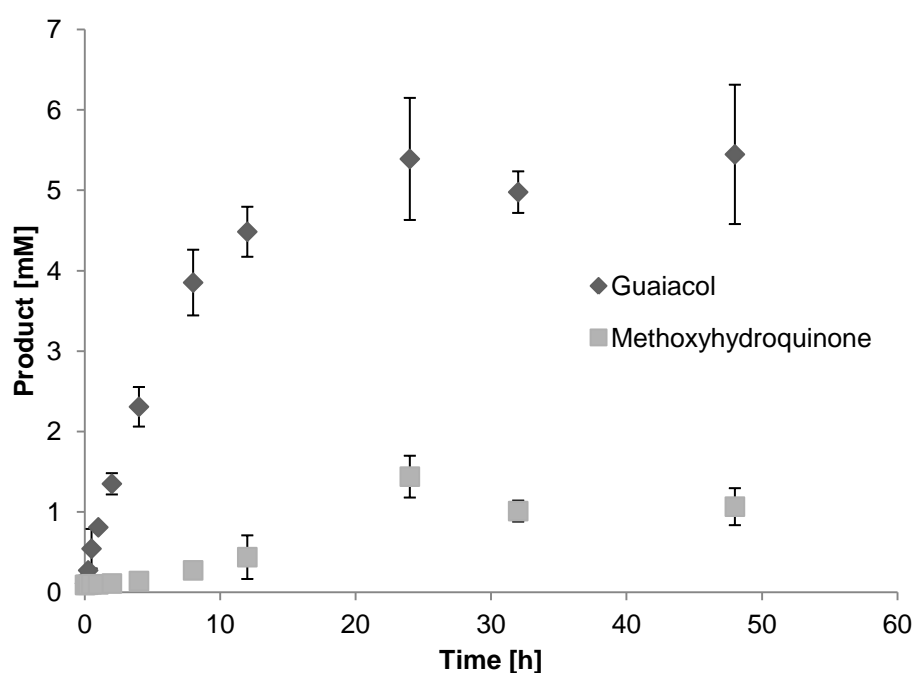


Figure 54. Conversion of 20 mM anisole by variant M2 (1 μ M purified enzyme) employing a GDH cofactor regeneration system.[164] Time dependent formation of guaiacol and methoxyhydroquinone was recorded for 48 h and measured by HPLC.

The final product concentrations after 24 h reaction are in good agreement with the amount of product measured on GC-FID during 24 h reactions with 20 mM substrate (0.83 g L⁻¹; Table 15).

Aromatic hydroxylation of *p*-xylene led to the formation of two products, 2,5-DMP and 2,5-DMHQ (see appendix section for GC-MS fragmentation analysis of 2,5-DMHQ; Figure 83). Despite the detection of a second product peak only 2,5-DMP was quantified during HPLC measurements since no standard for the respective hydroquinone (2,5-DMHQ) was available at that time. After 4 h reaction time already half of the final product was produced corresponding to 2.16 mM 2,5-DMP (Figure 55). The conversion stopped after 12 h whereas anisole displayed increased product formation until 24 h (Figure 54). The finally obtained concentration of 2,5-DMP was 3.65 mM (0.45 g L⁻¹) after 48 h.

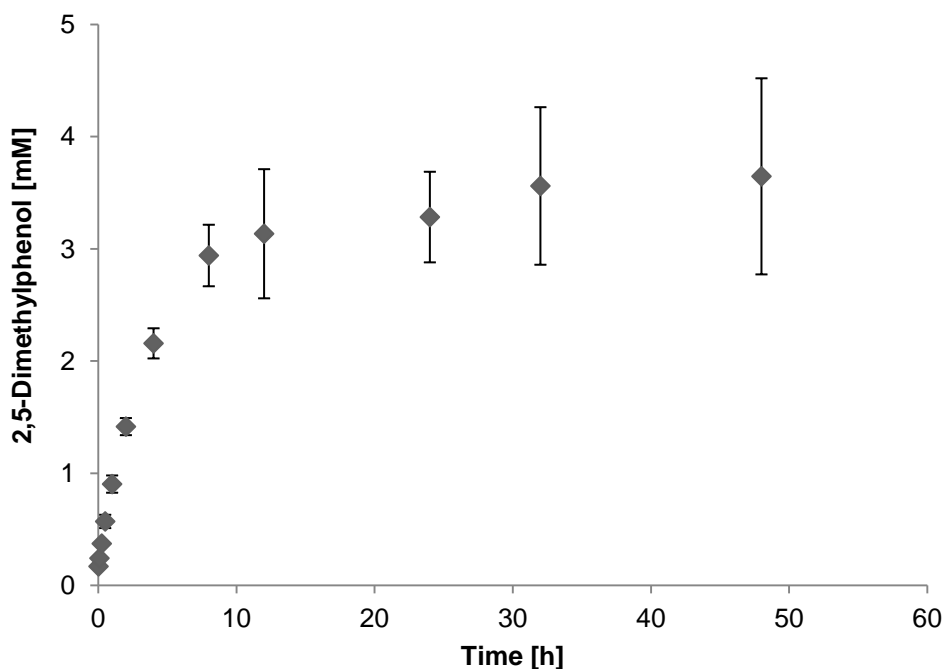


Figure 55. Conversion of 20 mM *p*-xylene by variant M2 (1 μ M purified enzyme) employing a GDH cofactor regeneration system.[164] Time dependent formation of 2,5-DMP was recorded over 48 h and measured by HPLC.

The final product concentration of 2,5-DMP reaches 65 % of the product concentration in the 24 h reactions samples (5.6 mM 2,5-DMP; Table 16).

The conversion of toluene displayed formation of two distinct products, *o*-cresol and methylhydroquinone (Table 14). *o*-Cresol is produced within the first 4 hours reaching the maximum of 1.3 mM (Figure 56). After this time point the concentration decreases and remains at 0.65 mM until end of the reaction. The formation of methylhydroquinone, however, increases steadily reaching 1.5 mM after 32 h, which is the maximum concentration under the described reaction and sampling conditions. In contrast to anisole and *p*-xylene, the formation of the respective hydroquinone starts very early after 5 min reaction time (Figure 56); for anisole and *p*-xylene between 4 and 8 h.

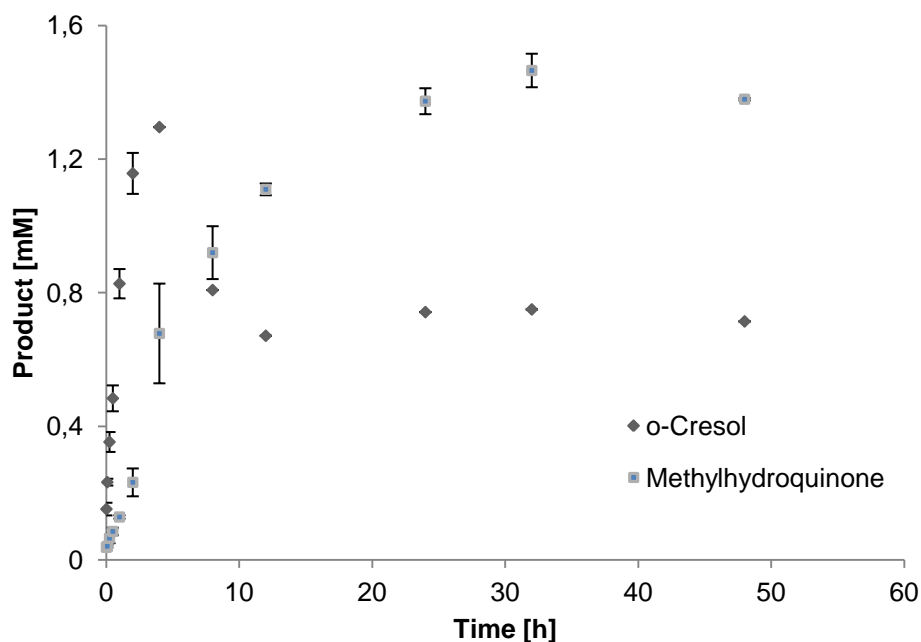


Figure 56. Conversion of 20 mM toluene by variant M2 (1 μ M purified enzyme) employing a GDH cofactor regeneration system.[164] Time dependent formation of *o*-cresol and methylhydroquinone was recorded over 48 h and measured by HPLC.

Compared to measurements on GC-FID (Table 14) the ratio for phenol to hydroquinone is inverted during HPLC measurements producing 65 % methylhydroquinone and leading to 0.25 g L⁻¹ product (1.38 mM methylhydroquinone and 0.71 mM *o*-cresol) corresponding to a TTN of 3470 which is 36 % lower than measured before in the 24 h conversion (TTN 5421; Table 14).

3.5.4.8 Conversion of Halogenated Benzenes with Variant M2

Although fluorobenzene was employed as substrate, no significant amount of product was generated during the reaction that could be detected by HPLC-UV. Further HPLC analysis with fluorobenzene was therefore not pursued. Conversion of chlorobenzene leads mainly to the formation of the chlorohydroquinone as determined before (Table 17) with a product ratio of 2:1 (hydroquinone to phenol) after 48 h (Figure 57). The final product concentration after 24 h reached 0.55 mM 2-chlorophenol and 1 mM chlorohydroquinone.

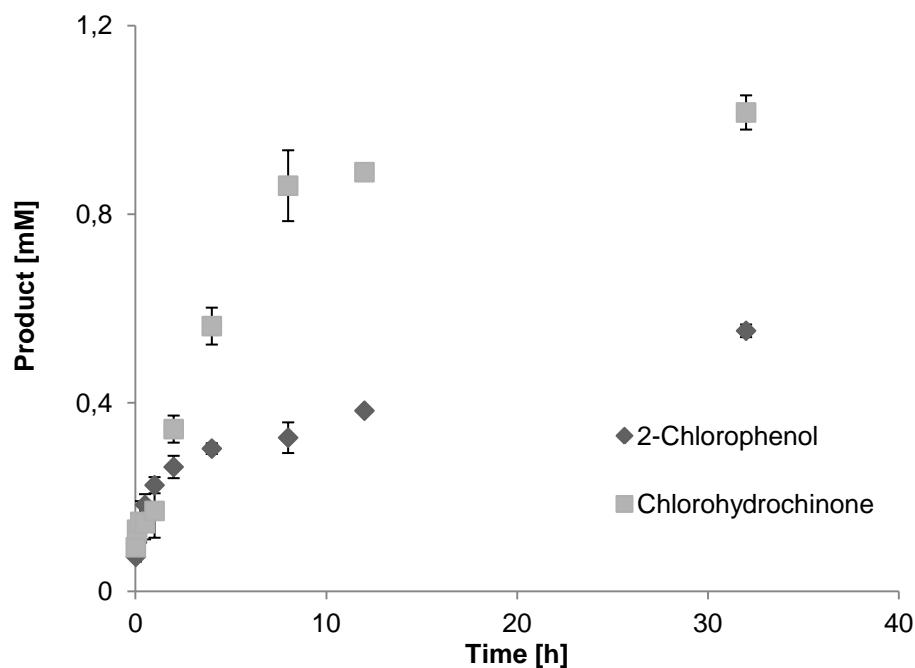


Figure 57. Conversion of 20 mM chlorobenzene by variant M2 (1 μ M purified enzyme) employing a GDH cofactor regeneration system.[164] Time dependent formation of 2-chlorophenol and chlorohydroquinone was recorded over 32 h and measured by HPLC.

Other than for anisole and toluene (Figure 54 and 56) the concentration of chlorohydroquinone exceeded after two hours reaction time the amount of 2-chlorophenol (Figure 57). The formation of the respective chlorohydroquinone started after 1 h at a 2-chlorophenol concentration of 0.23 mM. After 48 h conversion a product concentration of 0.22 g L⁻¹ was obtained, corresponding to a TTN of 2573. Bromobenzene displayed a similar product formation profile as chlorobenzene with similar product ratios for bromohydroquinone and 2-bromophenol (data not shown). The total product concentrations obtained for variant M2 and 20 mM of the respective substrates showed the following order: anisole (6.8 mM) > *p*-xylene (3.3 mM; excluding hydroquinone) > toluene (2.1 mM) > chlorobenzene (1.5 mM) > bromobenzene (1.6 mM) > iodobenzene (0.3 mM). The highest concentration for *o*-phenols was obtained using anisole (5.4 mM) and *p*-xylene (3.3 mM) as substrate. All other substrates did not exceed concentrations of 0.7 mM for the *o*-phenolic product. The *o*-phenol-hydroquinone ratio for toluene, chloro- and bromobenzene was similar. Compared to anisole the conversion of toluene, chloro- and bromobenzene led to formation of the respective hydroquinone in excess. The lowest product concentration was monitored for iodobenzene, which did not exceed 0.33 mM 2-iodophenol.

3.5.5 Discussion

In previous sections of this thesis the engineered P450 variants M2 and M3 displayed high activity and regioselectivity towards hydroxylation of *p*-xylene and monosubstituted benzenes (section 3.1 and 3.3). To generate a deeper insight into the operational performance and synthetic potential of P450 BM3 WT and variants M2 and M3 a GDH cofactor regeneration system was used to access

parameters such as regioselectivity, TTN, product concentrations and yields for selected substrates. The engineered P450 BM3 variants displayed for all substrates highest product yields when compared to the WT. This can be attributed to the increased coupling efficiency and activity (section 3.1 and 3.3).[68, 69] Interestingly, the differences in product yields between WT and both variants (M2 and M3) are higher at low substrate concentrations (1 and 2 mM) indicating that the substrate influx into the active site is restricted by R47 and Y51 for efficient catalysis as it was shown before (section 3.1).[69, 70] In addition, the improved K_M value (*p*-xylene; variant M2; Table 6; Figure 29) and higher binding affinity of benzene substrates together with increased coupling efficiencies (sections 3.1 and 3.6) make variants M2 and M3 better catalyst, especially at low substrate concentrations.[69]. Since evaporation of the substrates cannot be fully excluded the low productivity of the WT can also be attributed to the low activity[68, 69] and non-enzymatic substrate depletion in the reaction vessel over time. Increasing the substrate concentration to 20 mM allows also efficient hydroxylation of most substrates with the WT catalyst reaching up to 93 % TTN of the best variant M3 for chlorobenzene hydroxylation (Table 17). A reason for the unexpected high TTN of the P450 BM3 WT at 20 mM substrate concentration could be that no destabilizing mutations were introduced allowing a long catalyst ‘lifetime’ [293] as well as a decreased effect of evaporation for halogenated substrates. Higher TTN values of variant M2 and M3 can be attributed to the increased coupling efficiencies that reduce the amount of reactive oxygen species that could lead to inactivation of the monooxygenase catalyst.[151, 294, 295] For a further optimized process, space time yields might be more significant as an economic parameter and therefore the P450 variants M2 and M3 with higher activities, coupling efficiencies and yields are better candidates compared to the WT enzyme.

After 24 h reaction time an additional product was detected for all substrates except iodobenzene. The products were identified as the respective hydroquinones. Under limited cofactor supply the hydroquinones were not detected (section 3.1 and 3.3.), but a previous report showed that hydroxylation of various phenols by P450 BM3 (WT and variant F87V) results in formation of hydroquinones.[296] At comparably low substrate concentrations of 1 and 2 mM benzene the amount of produced hydroquinone is predominant as shown for variant M2 and 2 mM bromobenzene (99 % bromohydroquinone; Table 18).By increasing the substrate concentration to 20 mM the ratio of phenol to hydroquinone is shifted towards the phenolic products. One reason for the substrate-concentration directed selectivity might be that the first reaction product is transported into the organic substrate phase that serves as a bilayer in the reaction system.[297] The high pK_a value of the reaction products *o*-cresol ($pK_a >10$) and guaiacol (pK_a of ~ 10) leads to accumulation of the products in the organic phase.[298] Consequently, the products formed in the first reaction step are less accessible for a second hydroxylation by P450 BM3, especially at high substrate concentrations (20 mM) where a permanent organic bilayer (emulsion) is present. Over time the substrate phase is depleted and evaporation effects can occur especially at 1 and 2 mM substrate concentration, leading to a reduction of the organic bilayer. The increasing amount of phenolic product and depletion of the benzene in the

reaction vessel could lead to a 'switch' in substrate selectivity from benzene to phenol and to the preferential formation of hydroquinone at lower substrate concentrations. This observation was predominant in reaction setups with 1 and 2 mM benzene at which the substrates were fully depleted after 24 h reaction time. Unfortunately the measured conversions were not matching the product yields most likely due to evaporation of the substrates. Depletion of substrate could therefore be a misleading parameter when describing the catalytic performance.[299] Another important parameter that influences selectivity with two substrates present in a reaction system would be the K_M and k_{cat} value of P450 BM3 for the benzene and phenol substrates. More experimental investigations are necessary to draw a clear conclusion.

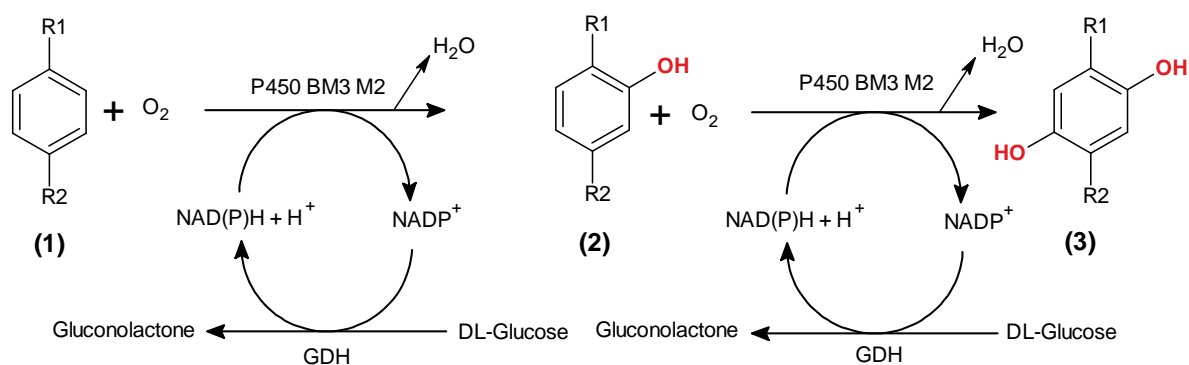
To access information on the time dependent product formation and altered substrate selectivity, the catalytic performance of variant M2 was analyzed via HPLC. Therefore, the formation of phenol and hydroquinone was measured for 48 h reaction time (Figures 54 to 57). The finally obtained concentrations for phenol and hydroquinone were in good agreement with the previous end-point measurements after 24 h reaction time (Tables 14 to 17). Toluene was the only exception displaying a higher concentration of methylhydroquinone (1.38 mM) than *o*-cresol (0.71 mM) after 4 h reaction time. One reason for the altered product ratios could be the higher volatility of toluene compared to the other substrates (boiling points: 111°C for toluene; 154°C for anisole) leading to a faster evaporation of the substrate. Sampling requires that at every time point the reaction vessel is opened leading to a loss of substrate which results (as discussed before) to an earlier and increased formation of methylhydroquinone in the reaction system. Other than for anisole and toluene, chloro- and bromobenzene display formation of the respective hydroquinones already within the first hour of the reaction (Figure 57). Compared to guaiacol and *o*-cresol the reaction products 2-chloro and 2-bromophenol have a lower pK_a value (< 9) that allows partitioning of the respective phenols into the aqueous and organic phase. The competition between benzene and phenol as substrates becomes therefore more evident for the halogenated substrates, which is supported by the measured data (Tables 14, 15, 17 and 18). Lower TTN values compared to anisole, *p*-xylene and toluene can be explained by the lower coupling efficiency and an earlier inactivation upon formation of reactive oxygen species.[295]

The possibility to produce hydroquinones from benzene educts with complete selectivity (no catechol or resorcinol was detected by GC-MS analysis) allows a new and attractive route to building blocks for vitamins (tocopherols), antioxidants, pharmaceuticals and polymerization inhibitors.[244, 300] Commonly hydroquinones are synthesized from benzene via the Hock process (60 % global hydroquinone production)[301] that has low product yields and leads to formation of side products (catechol).[296] Direct synthesis of hydroquinone from benzene was reported for *n*-alkane assimilating bacteria by Yoshikawa et al.[302], however, no detailed performance parameters are given.[301] Alternatively, hydroquinone can be obtained from quinic acid that is produced at 49 g L⁻¹ by a modified *E. coli* strain using glucose as starting material. Despite that most of the glucose is used

for biomass and metabolite production (79 to 85 %), the final step is performed in presence of stoichiometric amount of strong oxidizing reactants ($K_2S_2O_8$ and Ag_3PO_4 ; $(NH_4)_2Ce(SO_4)_3$) with hydroquinone yields of 51 to 91 %.[301]

Employing P450 BM3, together with a GDH regeneration system and no additional co-solvents, allows a simple and highly selective one pot synthesis of substituted hydroquinones with good yields (2 mM bromobenzene; variant M3; 75 % yield; 99 % bromohydroquinone; Table 18). Despite that a product mixture is obtained, the purification and separation of phenol and hydroquinone can be performed by a simple re-crystallization of the hydroquinone. The high selectivity for both reaction steps towards the hydroquinone (Scheme 10) circumvents energy intensive distillation of catechol or resorcinol making the reaction highly attractive for synthetic approaches. Overall the productivities of the generated variants allow a preparative scale synthesis of phenols and hydroquinones in milligram to gram scale for anisole with variant M3. With a TTN higher than 10000 (anisole hydroxylation) variant M3 reaches a catalytic performance that is rarely reported in P450 BM3 conversions.[151] For all other substrates product concentrations higher than 0.26 g L^{-1} were reached (from 20 mM substrate) which opens the possibility for further improvements of the catalyst and also for reaction engineering. Despite that a buffered reaction system was employed, the influence of the gluconolactone formation has to be investigated as well as the effect of the pH on P450 BM3 (activity and selectivity) and its effect on partitioning of phenolic products in the reaction system.

The initially drawn Scheme 9 has therefore to be revised to address also the formation of secondary products which is rarely reported in literature despite that compound I allows hydroxylation or epoxidation of nearly every C-H atom of a substrate (Figure 7).[70] Scheme 10 shows the re-drawn reaction scheme for conversion of substituted benzene substrates with P450 BM3.



[R₁: H, CH₃, OCH₃, Cl, Br; R₂: H, CH₃]

Scheme 10. Revised reaction scheme employing P450 BM3, benzene as substrate and a GDH cofactor regeneration system. After the first reaction step the benzene substrate (1) is hydroxylated to a phenolic product (2). Depending on the substrate and phenol concentration in the reaction system the engineered variants produce also the respective hydroquinones (3) in considerable amount.

3.5.6 Summary and Conclusion

In this chapter the operational performance of P450 BM3 WT and two engineered variants was investigated for the synthesis of *o*-phenols. Despite that the catalytic reaction to *o*-phenols was performed with high regioselectivity towards *o*-hydroxylation (>95 %) all substrates (except iodobenzene) were hydroxylated in a second reaction step to the corresponding hydroquinones. The regioselectivity towards production of hydroquinones was even higher (>99 %), allowing a new one-pot synthesis route to hydroquinones starting from benzene as educt which was not reported before employing a P450 monooxygenase. Responsible for the switch in substrate selectivity are varying substrate concentrations, the pK_a value of the phenolic products and the affinity of the monooxygenase for the *o*-phenolic products. Despite the challenges in product down streaming from aqueous solutions the high selectivity towards the hydroquinone allows a simple purification from the respective phenol intermediates. Further protein engineering, optimization of process parameters such as pH control and catalyst concentrations could yield higher product concentrations and TTN values. In addition, it is further necessary to investigate enzymatic parameters such as the K_M value, allowing a deeper understanding in respect to catalyst selectivity for benzene and phenol as substrates. Nonetheless, the one pot reaction from benzene to produce hydroquinone with P450 BM3 as catalyst was not reported before and the biocatalytic process has several advantages over chemical synthesis, avoiding several reaction and purification steps (Hock synthesis) as well as low product yields. The additionally produced hydroquinones can directly be used in synthesis of vitamins, pharmaceuticals and antioxidants.

3.6 Chemoselective α -Hydroxylation of Nitrotoluenes by P450 BM3

3.6.1 Abstract

In this section P450 BM3 WT and three derived variants are characterized with regards to activity, coupling efficiency, productivity and chemoselectivity for the hydroxylation of *o*-, *m*- and *p*-nitrotoluene. In contrast to the previously reported benzene substrates, all BM3 variants hydroxylate the nitrotoluene isomers preferentially at the α -position. All substrates were hydroxylated with regioselectivities higher than 95 %, providing new options for biocatalytic synthesis of nitrobenzylalcohols that have large potential as building blocks for pharmaceuticals and pigments. The variant M3 displayed 84-fold higher productivity (>200 U min⁻¹; *p*-isomer) and 2.5-fold higher coupling efficiencies (26 %; *o*-isomer) in comparison to the WT enzyme. With a TTN of more than 1500 and a final product concentration of 0.23 g L⁻¹ (*p*-NBOH) this is the first report of a biosynthetic route for this class of compounds using P450 BM3. Differences in catalytic performance are outlined especially with regards to influence of the NO₂-substituent on chemoselectivity as well as protein-substrate interactions, leading to a rather unusual chemoselectivity in P450 BM3-hydroxylation.

3.6.2 Introduction

Nitroarenes such as nitrotoluene (NT) (Figure 58), di-nitrotoluene (DNT) and tri-nitrotoluene (TNT) are produced in million ton-scale mainly for the use as explosives.[303] Despite their persistent incidence in nature and their environmental impact[303, 304] these compounds can also be used, after reduction of the NO₂-group, as synthons for pharmaceuticals and pigments.[305-307]

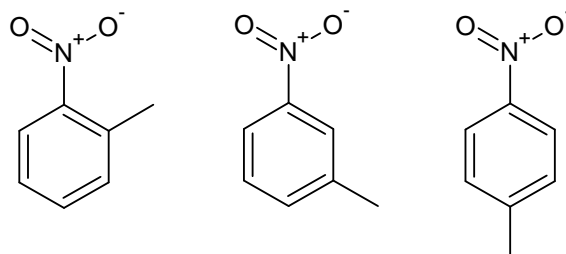


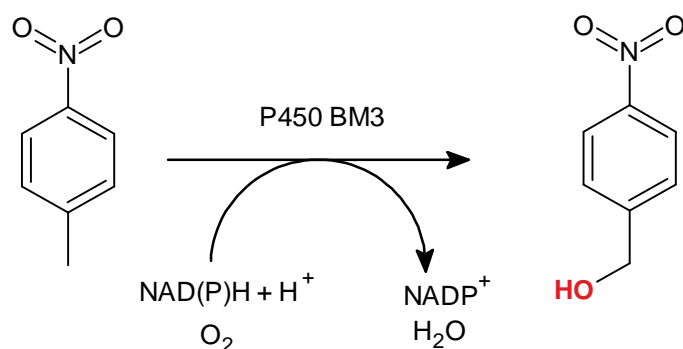
Figure 58. The three isomers of nitrotoluene; *o*-, *m*- and *p*-nitrotoluene (from left to right).

Hydroxylation of NT at α -position followed by reduction of the NO₂-group leads to formation of aminobenzylalcohols that represent substructures in several FDA-approved pharmaceuticals such as Ibutilide, Sotalol or Clenbuterol.[305] Nitrobenzylalcohols (NBOH) cannot be isolated from petrochemical resources such as oil or coal, thus requiring synthetic processes. Selective chemical hydroxylation of the α -C-atom was reported and patented in 1920 by Diefenbach (patent number DE214949). The reported synthesis strategy requires strong acids (H₂SO₄), a lead catalyst and low temperatures (<10°C) for direct oxidation of the α -position. Product yields range from 30 to 42 % with additional side products such as acids and aldehydes being formed. Few modifications for the synthesis of nitrobenzylalcohols were reported since 1920 (altogether 9 reports on Reaxys database; <https://www.reaxys.com>) leading to minor advancements of the synthetic route. Alternatively, *p*-nitrobenzylalcohol can also be generated in a two-step synthesis starting from *p*-nitrobenzylchloride. This route requires stoichiometric amounts of NaOH and high temperatures.[308, 309] The application of corrosive reactants and production of halogenated waste (HCl) is undesired, especially when considering increasing regulations for waste disposal and the pursuit of ‘greener’ synthetic routes.[12, 13, 146]

Biocatalysis can be very attractive to replace synthetic processes with problematic waste generation and requirements for extraction and remediation of toxic metals.[71, 75, 146] Several bacterial strains and enzymes such as peroxygenases, the non-heme iron dependent xylene monooxygenase (XMO) or P450 monooxygenases were reported that are capable to hydroxylate NT.[307, 310, 311] The biocatalytic route to NBOH is comparably inefficient with regards to activities and selectivity of the reaction as well as product yields employing biocatalysts. Another challenge using the reported catalyst is the second hydroxylation step where NBOH serves as substrate leading to the formation of 4-nitrobenzoic acid[307, 310], whereas the benzylalcohol and nitrobenzaldehyde represent only reaction intermediates. The very high toxicity of nitroarenes even at low concentrations limits

applications with whole cells; consequently, an *in vitro* biocatalytic approach could be the method of choice to produce NBOH. This could also circumvent the reduction of the NO₂-group by whole cells which leads to formation of toluidine that might not be accepted as substrate for the hydroxylation reaction.[307, 311]

P450 BM3 is the ideal monooxygenase catalyst for *in vitro* applications, due to its self-sufficient character, easy purification in large quantity, solubility in water as well as its tolerance to high substrate concentrations.[70] Hydroxylation reactions with NT and P450 BM3 have not been reported.[70] However, the engineered P450 BM3 variant M2 (R47S, Y51W, I401)[69] was able to hydroxylate *p*-NT at α -position (section 3.2; Table 9; Substrate 28). The observation that P450 BM3 M2 hydroxylates the α -position of *p*-NT with high selectivity was unexpected, especially when other investigated substrates in this thesis are taken into account (section 3.2; Table 8). To investigate this rather unusual chemoselectivity, the BM3 WT and three variants thereof (M1, M2 and M3; engineered and characterized in section 3.1 and 3.4 of this thesis) were selected for conversions of all three isomers of NT to generate an activity and selectivity profile for P450 BM3 (Scheme 11).



Scheme 11. Hydroxylation of *p*-NT by P450 BM3 in α -position leads to the formation of *p*-NBOH.

3.6.3 Experimental

P450 BM3 was expressed and purified as described before in the material and methods section (section 2.10.6). NADPH oxidation rate and coupling efficiency (section 2.10.9.1) were determined using 2 mM substrate and a final concentration of 2 % DMSO as co-solvent. P450 BM3 concentrations were adjusted (WT: 0.8 μ M; M1: 0.4 μ M; M2: 0.2 μ M; M3: 0.1 μ M) for reliable measurement of NADPH depletion over time in 5 mL quartz cuvettes. Product formation rates (PFR) were calculated from initial NADPH oxidation rates and coupling efficiency values as reported before.[156] Regioselectivity of all variants and WT was estimated using a GDH cofactor regeneration system[164] for semi-preparative product generation as described in material and methods section in a final volume of 2 mL (section 2.10.9.3). A final substrate concentration of 4 mM was employed together with 2 % DMSO (V/V) and 1 μ M purified P450 BM3 enzyme. Products from P450 BM3 conversions with a GDH cofactor regeneration system were extracted after 24 h reaction time by two-phase extraction with 1 mL EtOAc containing 5 mM 2,5-DMP as internal standard. Extracted

compounds were analyzed by GC-FID. For identification of products commercial standards of substrates and potential products (NBOHs and nitrobenzaldehydes) were separated and calibration curves were prepared for quantification of NBOHs. A detailed list of retention times for all analyzed compounds, together with heating profiles and used columns for separation, can be found in the appendix section of this thesis (Table 29).

3.6.4 Results

Tsvetan Kardashliev performed the docking of substrates and contributed in parts to the collection of catalytic data. Bastian Zirpel assisted in catalytic characterization of P450 BM3 variants within the frame of his master thesis project at the Institute of Biotechnology (RWTH Aachen University).

3.6.4.1 Hydroxylation of *o*-Nitrotoluene by P450 BM3 WT and Variants

The hydroxylation of *o*-NT by P450 BM3 WT was catalyzed with a NADPH depletion rate of 35 min⁻¹ and coupling efficiency of 18.7 %, leading to a final PFR of 7 min⁻¹ (Table 19). Compared to WT, the engineered variants displayed increased catalytic activity which follows the iterative development of the three variants (WT > M1 > M2 > M3) in this thesis (section 3.1 and 3.4).[68, 69] Opening the substrate channel (variant M1) increased the activity by 3-fold, while an increased activity of up to 6-fold compared to the P450 BM3 WT was measured for variant M2. The variant M3 displayed the highest activity which represents an 18-fold increased activity compared to the WT. In addition, the coupling efficiency of M3 was 4 % higher than for the P450 BM3 WT, leading to a maximum PFR of 146 min⁻¹ for *o*-NBOH and making variant M3 a 20-fold more productive catalyst.

Table 19. Catalytic characterization of P450 BM3 WT and variants towards hydroxylation of *o*-NT. Further identified products on GC-MS: **M3**: 33 % *o*-nitrobenzaldehyde and 5.8 % phenolic product; **M2**: 0.6 % *o*-NBA and 2.7 % phenolic product; **WT**: 3.5 % *o*-nitrobenzaldehyde and 1.2 % phenolic product.

Variant	N ^[a]	C ^[b]	PFR ^[c]	<i>o</i> -NBOH [%]	<i>o</i> -NBOH [mM]	Yield [%]	TTN
WT	35	18.7	7	95 (99)*	0.62	16	620
M1	97	22.4	22	n.d.	n.d.	n.d.	n.d.
M2	220	20.6	45	97 (98)*	1.48	37	1480
M3	643	22.7	146	61 (94)*	0.95	24	950

^[a]NADPH depletion rate [mol_{NADPH} mol_{P450}⁻¹ min⁻¹]; ^[b]coupling efficiency [mol_{Product} mol_{NADPH}⁻¹]; ^[c]product formation rate [mol_{Product} mol_{P450}⁻¹ min⁻¹]; *sum of NBOH and NBA (as % GC-area).

The regioselectivity of WT and variant M2 was similar with 95 % and 97 % of *o*-NBOH formed respectively. A minor fraction of *o*-NBA (WT: 3.5 %; M2: 0.6 %) was also produced when a GDH cofactor regeneration system employed. The amount of side-products in the reaction with M3 was significant higher, with 33 % *o*-NBA and 5.8 % of a phenolic product that could not be fully resolved by GC-MS analysis. The significantly increased amount of *o*-NBA formed during the hydroxylation reaction of *o*-NT with variant M3 is shown in Figure 59 (black line, peak at 2.25 min retention time).

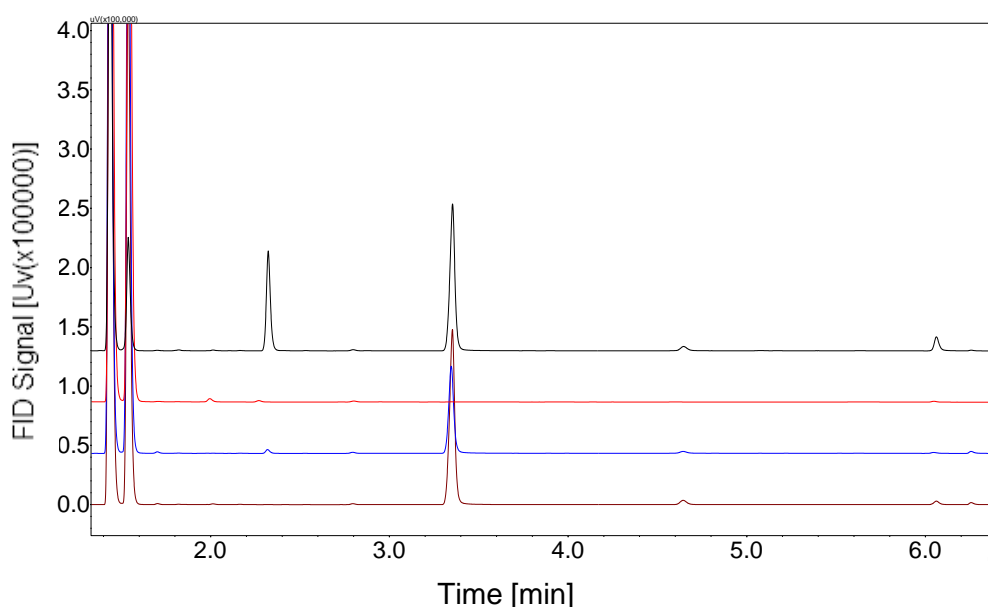


Figure 59. Chromatograms show separation and detection of products from the 24 h conversion of *o*-NT with selected P450 variants. M3 (black lane), EV (red lane), P450 BM3 WT (blue lane) and variant M2 (brown lane). Identified compounds elute in following order: *o*-NT (1.52 min), *o*-NBA (2.25) and *o*-NBOH (3.24 min). After a retention time of 6 min phenolic products elute from the column (e.g. M3; black line; 6.1 min).

Long term reactions with P450 BM3 WT and excess of NADPH cofactor[164] generated in total 0.62 mM *o*-NBOH, corresponding to a TTN of 620 and a product yield of 16 %. Variant M2 displayed improved catalytic performance under the described conditions, having a more than doubled product yield (37 %) and a final product concentration of 0.23 g L⁻¹ of *o*-NBOH. M2 is a much better catalyst than P450 BM3 WT and is the best variant with regards to production of *o*-NBOH. Variant M3 has a 3-fold higher PFR but generates a larger amount of side products. The 33 % of *o*-NBA produced by the variant M3 are not considered in the calculated values for product yield, TTN and final product titer.

3.6.4.2 Hydroxylation of *m*-Nitrotoluene by P450 BM3 WT and Variants

The hydroxylation of *m*-NT followed the same trend as described before employing the *o*-NT isomer as substrate (Table 19). Activities were increased in following order: WT > M1 > M2 > M3, reaching a maximum NADPH oxidation rate of 737 min⁻¹ (Table 20). Except for the WT, the NADPH depletion rate is higher for *m*-NT than for *o*-NT, whereas the coupling efficiencies are slightly decreased for the variants and WT.

Table 20. Catalytic characterization of P450 BM3 WT and variants towards hydroxylation of *m*-NT. Further identified products on GC-MS: **M3**: 3.8 % *m*-NBA and 15.5 % phenolic product; **M2**: 2.6 % *m*-NBA and 7.7 % phenolic product; **WT**: 5.4 % *m*-NBA and 5.2 % phenolic product.

Variant	N ^[a]	C ^[b]	PFR ^[c]	<i>m</i> -NBOH [%]	<i>m</i> -NBOH [mM]	Yield [%]	TTN
WT	31	11.0	4	89 (94)*	0.33	8	330
M1	180	16.5	30	n.d.	n.d.	n.d.	n.d.
M2	384	17.9	69	90 (93)*	0.45	11	450
M3	737	26.3	194	81 (85)*	0.52	13	520

^[a]NADPH depletion rate [mol_{NADPH} mol_{P450}⁻¹ min⁻¹]; ^[b]coupling efficiency [mol_{Product} mol_{NADPH}⁻¹]; ^[c]product formation rate [mol_{Product} mol_{P450}⁻¹ min⁻¹]; *sum of NBOH and NBA (as % GC-area).

Nevertheless all variants displayed improved coupling efficiencies during hydroxylation of *m*-NT, 5.5 % for M1 and 6.9 % for M2, respectively. The highest coupling efficiency was obtained for variant M3, with a more than doubled coupling efficiency (26.3 %) compared to the WT. The improved coupling consequently resulted in higher PFR values, in the case of variant M3 (194 min⁻¹) outperforming the WT enzyme (4 min⁻¹) by 48-fold. Long time conversions with an excess of NADPH cofactor revealed minor differences between catalytic performance of WT and variant M3 (Table 20). Although P450 BM3 WT has a significant lower PFR than the variant M3 (4 min⁻¹ vs. 194 min⁻¹; Table 20), the values for TTN and final product concentration (for *o*-NBOH) reached 63 % of that measured for the best variant (M3). Overall the amount of product after 24 h reaction time reached at maximum 79 mg L⁻¹ with M3, which is three times lower than the product titer from conversion of *o*-NT (Table 19). The amount of phenolic by-products is highest for the *m*-isomer, reaching 15.5 % of total GC-area for variant M3, 7.7 % for variant M2 and 5.2 % for P450 BM3 WT. Minor side products were identified as *m*-NBA reaching up to 5.4 % for P450 BM3 WT.

3.6.4.3 Hydroxylation of *p*-Nitrotoluene by P450 BM3 WT and Variants

In all reactions with *p*-NT, the regioselectivity for the α -position was very high (>93 %) independent of the employed P450 BM3 variant. M3 accomplishes a selectivity of 99 % for α -hydroxylation if the 10 % *p*-NBA formed as side-product is also considered (Table 21). The coupling efficiency of WT, M1 and M2 with *p*-NT as substrate was the lowest among all investigated NT isomers. Only variant M3 exceeds 10 % coupling efficiency (23.8 %). The P450 BM3 WT has the lowest activity and productivity with a PFR of 2.4 min⁻¹. The PFRs of variants M1 and M2 are significantly improved compared to the WT (3.5-fold and 13.5-fold). Due to a 3-fold higher coupling efficiency for *p*-NT (23.8 %) and the highest NADPH oxidation rate for all NT isomers (847 min⁻¹; Table 21), the variant M3 outperforms the WT by 84-fold. Besides of the highest regioselectivity (99 %), TTN (1560) and product yield (39 %), the highest product titer was obtained with variant M3 and *p*-NT as substrate (1.56 mM; 0.24 g L⁻¹).

Table 21. Catalytic characterization of P450 BM3 WT and variants towards hydroxylation of *p*-NT. Further identified products on GC-MS: **M3**: 10 % *p*-NBA and 1 % phenol; **M2**: 0 % *p*-NBA and 7 % phenol; **WT**: 5 % *p*-NBA.

Variant	N ^[a]	C ^[b]	PFR ^[c]	<i>p</i> -NBOH [%]	<i>p</i> -NBOH [mM]	Yield [%]	TTN
WT	31	7.7	2	94 (99)	0.28	7	280
M1	97	7.3	7	n.d.	n.d.	n.d.	n.d.
M2	309	8.7	27	93	1.05	26	1050
M3	847	23.8	202	89 (99)	1.56	39	1560

^[a]NADPH depletion rate [$\text{mol}_{\text{NADPH}} \text{mol}_{\text{P450}}^{-1} \text{min}^{-1}$]; ^[b]coupling efficiency [$\text{mol}_{\text{Product}} \text{mol}_{\text{NADPH}}^{-1}$]; ^[c]product formation rate [$\text{mol}_{\text{Product}} \text{mol}_{\text{P450}}^{-1} \text{min}^{-1}$]; *sum of NBOH and NBA (as % GC-area).

3.6.4.4 Hydroxylation of *p*-Toluidine by P450 BM3 Variant M3

The importance of aminobenzylalcohols as valuable building blocks for pharmaceuticals and pigments was briefly explained in the introduction of this section of the thesis. The direct α -hydroxylation of *p*-toluidine could open an additional route to aminobenzylalcohols as alternative to the hydroxylation of NTs and subsequent reduction of the NO₂ substituent. However, during the substrate screening, conversion of *p*-toluidine and aniline did not lead to a detectable product formation (section 3.2.4.1; Table 8). To evaluate the possibility for direct hydroxylation of *p*-toluidine a 24 h conversion of 30 mM *p*-toluidine (2 % DMSO co-solvent) was performed under the conditions described in material and methods section employing 1 μM P450 protein from variant M3 (section 2.10.9.3). After extraction with MTBE (section 2.11.1) a single product peak was detected on GC-FID and further analyzed by GC-MS. Mass analysis revealed that *p*-toluidine is preferentially hydroxylated on the aromatic ring to 6-amino-*m*-cresol (122.95 *m/z*) (see appendix for GC-MS fragmentation pattern; Figure 85). The hydroxylation of aminobenzenes was not further investigated. It is worth mentioning that 24 h conversions of *p*-nitrotoluene with cell free lysate produced a significant amount of *p*-toluidine which was attributed to an enzymatic reduction of the NO₂-group.[307, 311]

3.6.4.5 Docking of Nitrotoluenes into the Active Site of P450 BM3

In section 3.1 and 3.3 of this thesis the excellent regioselectivity of P450 BM3 towards aromatic *o*-hydroxylation of various benzenes was determined and attributed to strong interactions with the active site residue F87 (section 3.3; Figure 41 A and B; Table 10).[68, 69] A docking study with anisole bound in the P450 BM3 substrate binding pocket was in well agreement with the obtained activities and regioselectivities and supported the hypothesis of a T-shape benzene orientation with residue F87.[68, 259] The chemoselectivity for the α -position can be regarded as rather unusual, also when considering that various 1,4-substituted benzenes (4-fluoro, 4-chloro and 4-bromtoluene as well as for 4-methylanisole) were hydroxylated preferentially on the aromatic ring (section 3.2.4.1; Table 9). To investigate chemoselectivity for this class of substrates the isomers were docked into the

active site of P450 BM3 WT (PDB: 1BU7).[238] Figure 60 shows the active site of P450 BM3 WT with all three isomers bound in a productive binding mode; A: *o*-NT, B: *m*-NT; C: *p*-NT.

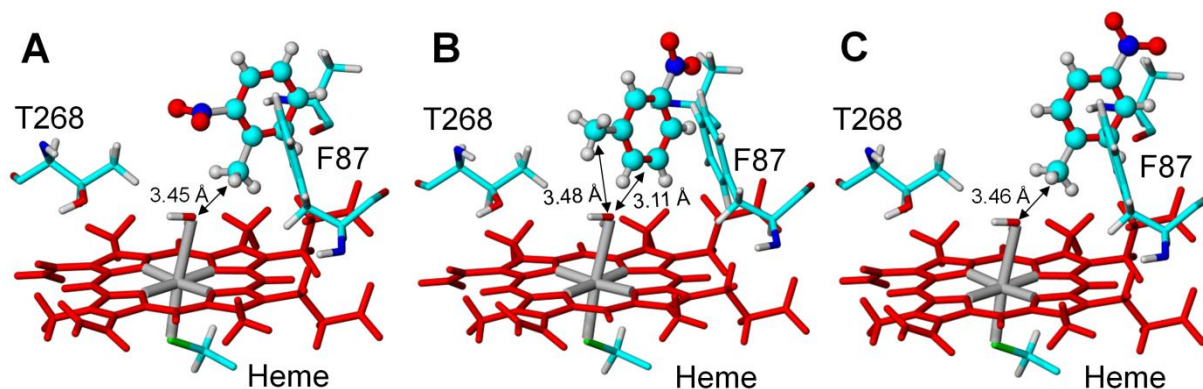


Figure 60. Docking of NT isomers into the active site of P450 BM3 WT (PDB: 1BU7).[238] The heme is depicted in stick representation (red) together with key residues F87, T268 and A330 (behind substrate and F87). Substrates are shown in ball and stick representation. **A:** *o*-NT in T-shape orientation with ΔG_{bind} of $-5.12 \text{ kcal mol}^{-1}$. **B:** *m*-NT in T-shape orientation with ΔG_{bind} of $-5.52 \text{ kcal mol}^{-1}$. **C:** *p*-NT in T-shape orientation with ΔG_{bind} of $-4.84 \text{ kcal mol}^{-1}$. Nitrogen atoms are highlighted in dark blue. Distance (in Å) between closest and hydroxylated C-atoms is indicated with arrows.

Docking of the isomeric substrates into the active site of P450 BM3 showed that the hydroxylated α -position is in all cases oriented towards the heme-bound water ligand and residue T268. The closest binding orientation was 3.11 \AA for *m*-NT which represents also the strongest binding (ΔG_{bind} : $-5.52 \text{ kcal mol}^{-1}$) obtained during docking of the NT isomers (Figure 60, B). Docking of *o*- and *p*-NT occurred with lower Gibbs free binding energy of $-5.12 \text{ kcal mol}^{-1}$ and $-4.84 \text{ kcal mol}^{-1}$ for the *o*- and *p*-isomers, respectively. The *m*-NT isomer was bound in an orientation that would also allow epoxidation of the aromatic ring and formation of a phenolic product.[68, 154, 156] In all cases further binding modes were suggested, but only the productive orientations with highest ΔG_{bind} that are in well agreement with the catalytic data are shown here (Table 19, 20 and 21). The binding mode of all NT isomers was in close proximity to F87 and in distinct T-shape orientation (section 3.3).[68, 69, 259] The NO_2 -substituent was however proposed in two different modes, pointing either towards residue T268 (Figure 61 A; *o*-NT) or towards residue A330 (Figure 61 B and C; *m*- and *p*-NT).

Variant M3 displayed the highest activity, PFR and coupling efficiency for all three isomeric substrates and in particular for *o*-NT an increased formation of *o*-NBA (Table 19). The major difference between M3 and all other investigated P450s is that this variant has an amino acid exchange directly in the substrate binding pocket close to the heme iron. A homology model was generated based on the crystal structure of P450 BM3 WT (PDB: 1BU7).[238] Figure 61 shows *o*-NT in 3 binding modes in the active site of the generated homology model of M3 (section 3.4). Using the generated model structure for docking, the *o*-NT is in close distance to residue F87 and in a distinct T-shape orientation, as it was shown before for the WT enzyme (section 3.3; Figure 41 A and B). Compared to the WT structure the variant M3 provides three binding orientations that could lead to product formation (Figure 61). Conformation A retains the original binding mode obtained also for the

WT enzyme (Figure 60 A), but with increased distance (0.81 Å) between the CH₃-substituent and the heme iron and a slightly lower binding energy (+ 0.3 kcal mol⁻¹).

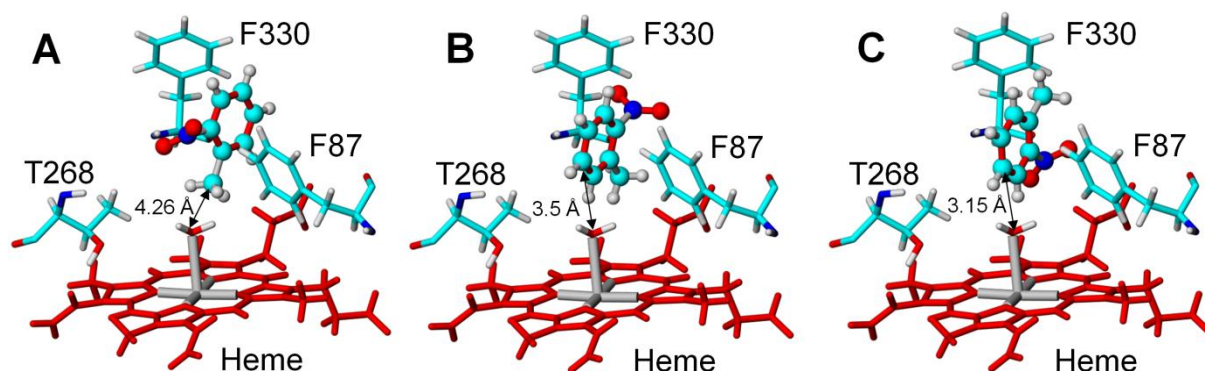


Figure 61. Docking of *o*-NT into the active site of the generated homology model of P450 BM3 variant M3 (PDP: 1BU7)[238]. The heme is depicted in stick representation (red) together with key residues F87, T268 and F330. Substrates are depicted as ball and stick representation. **A:** *o*-NT in T-shape orientation with ΔG_{bind} of -4.82 kcal mol⁻¹. **B:** *o*-NT in T-shape orientation with ΔG_{bind} of -4.69 kcal mol⁻¹. **C:** *o*-NT in T-shape orientation with ΔG_{bind} of -4.22 kcal mol⁻¹. Nitrogen atoms are highlighted in dark blue. Distance (in Å) between closest and hydroxylated C-atoms is indicated with arrows.

Binding energies for the two other alternative orientations are slightly lower. However, both orientations allow a closer distance of substrate and heme iron (Figure 61 B and C). The additional binding modes could explain the increased formation of phenolic products compared to the WT (5.8 % M3 vs. 1.2 % WT) (Table 19). Docking of *p*-NT into the model structure of M3 suggested two binding modes, both pointing the CH₃-substituent towards the heme iron. Orientations of the substrate leading to phenolic products were not suggested which is in good agreement with the determined chemoselectivity for *p*-NT (99 %; Table 21).

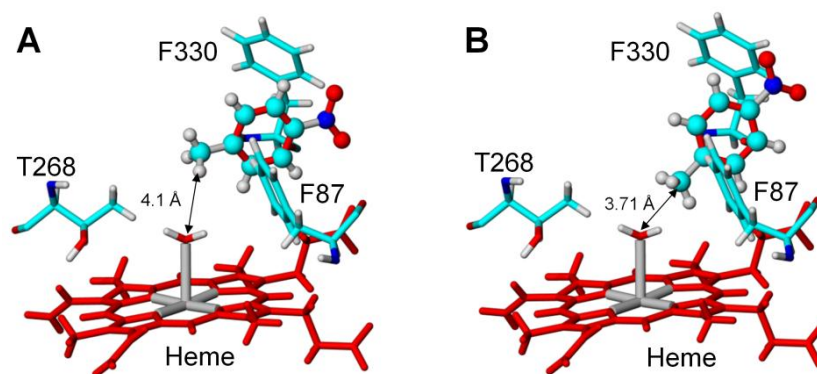


Figure 62. Docking of *p*-NT isomers into the active site of the generated homology model of P450 BM3 variant M3 (PDP: 1BU7).[238] The heme is depicted in stick representation (red) together with key residues F87, T268 and F330. Substrates are depicted as ball and stick representation. **A:** *p*-NT in T-shape orientation with ΔG_{bind} of -6.00 kcal mol⁻¹. **B:** *p*-NT in T-shape orientation with ΔG_{bind} of -5.44 kcal mol⁻¹. Nitrogen atoms are highlighted in dark blue. Distance (in Å) between closest and hydroxylated C-atoms is indicated with arrows.

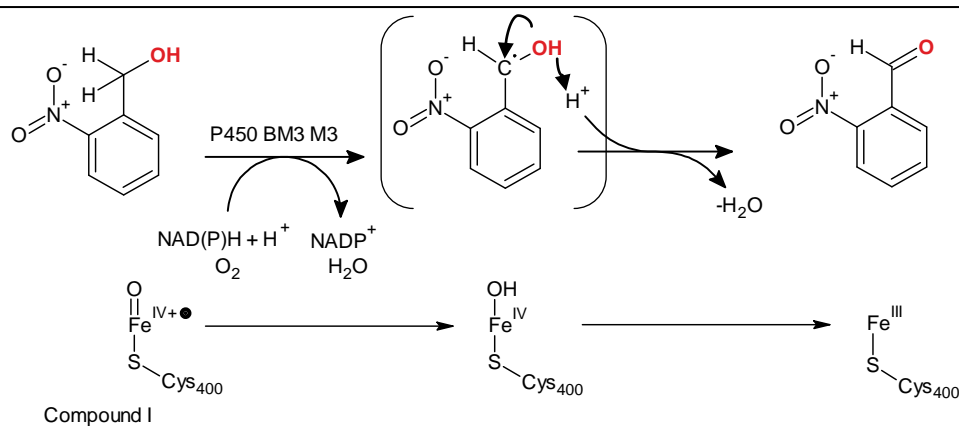
Compared to the WT the binding of *p*-NT is significantly stronger with ΔG_{bind} of -6.00 kcal mol⁻¹ and 5.44 kcal mol⁻¹ for the two suggested orientations in the active site of M3 (Figure 62 A and B).

Employing *m*-NT as ligand for docking into the active site of M3 did not provide any productive binding mode in close distance to the heme iron, therefore the results of this isomer are not shown.

3.6.5 Discussion

Characterization of P450 BM3 towards NT hydroxylation ratified the results from section 3.2 of this thesis, where *p*-NT is hydroxylated with almost perfect chemoselectivity in α -position (99 %; M3; Table 21). Surprisingly, all NT isomers are hydroxylated with good to excellent selectivity in α -positions; for *o*- and *p*-NT selectivities of 94 % and higher were obtained for the α -position (Table 19 and 21). The docking of NT isomers into the WT crystal structure model is in good agreement with the measured chemoselectivity. The generated homology model of M3 could, in parts, reconstitute the product pattern obtained for *p*-NT (Table 21; Figure 62 A and B) and in parts also for *o*-NT (Table 19; Figure 61). Despite that measured and simulated data correlate well, previous results on conversion of monosubstituted benzenes and 1,4-substituted benzenes by P450 BM3 displayed very high chemoselectivity towards the aromatic ring (section 3.1, 3.2 and 3.3).[68, 69] The observation that P450 BM3 performs terminal α -hydroxylation on benzene substrates was reported before, however, with significantly less selectivity, loss in activity and requiring a mutation at position F87 (such as F87A).[156] Commonly, P450 BM3 performs selective aromatic hydroxylation on symmetrical benzenes substrates with one and two substituents (1,4-position) or three methyl groups as shown for mesitylene (section 3.4; Table 13).[68-70] Furthermore, also 4-methylanisole, 4-bromo- and 4-chlorotoluene and fluorobenzene, all with varying atom sizes, different sterical demands and different electron densities were hydroxylated preferentially on the aromatic ring (section 3.2.4.1; Table 9).[258, 275] Though NO₂ has a strong electronegative influence (NO₂ = 4.19 Pauling Units; fluor: 3.98 Pauling Units),[258] the hydroxylation is directed towards α -position for all three isomers, a rather unexpected result since the size and polar character of the NO₂-group would provide a strong interaction potential by holding the substrate in a distinct binding mode. This, however, would result also in a different chemo-selectivity for each substrate. Since no polar residue is available within the binding pocket, that could directly interact with NO₂ (e.g. salt-bridge formation), the directing forces are most likely π - π -interactions of substrate and F87.[68, 69, 259] High binding energies suggest that the NT isomers are interacting with phenylalanine 87 and also F330 (variant M3) reaching a ΔG_{bind} of -6.0 kcal mol⁻¹ for *p*-NT, suggesting a close aromatic interaction and strong binding (Figure 62 A).[259] From the catalytic data and docking experiments it remains unclear, how the substrate binding orientation is influenced by the NO₂-substituent. One reason for the high selectivity towards CH₃-groups could be that electron densities are too low for epoxidation and re-aromatization to take place.[154] Instead, P450 BM3 performs the rebound mechanism to produce the respective NBOH (section 1.14.1; Step VII and VIII; Figure 6).[70] The coupling efficiency for a specific substrate allows further insights into binding and conversion of a substrate during the catalytic reaction.[68, 151] A loosely fitting substrate will lead to low coupling efficiencies as well as an unfavorable positioning where the target H-atom is not in abstraction distance.[151, 157, 158] The coupling

efficiencies differ significantly between investigated variants and employed substrates (Tables 19, 20 and 21). The lowest coupling efficiencies were obtained for *p*-NT (WT, M1 and M2) whereas variant M3 has a 3-fold improved coupling efficiency that results in an 84-fold improved PFR compared to the WT (Table 21). The increase in coupling efficiency can be attributed to a tighter binding of the substrate in the active site, where the phenylalanine at position 330 provides additional aromatic interactions (Figure 62).[259] In all conversions M3 displays highest coupling efficiencies, NADPH depletion rates and PFRs. As shown before in section 3.4, the two phenylalanine residues in the active site provide a better binding and orientation of substrates leading to a faster initiation of the catalytic mechanism and making variant M3 a more efficient catalyst.[70] The improved sterical control governed by F330 ensures coupling efficiencies beyond 20 % for the investigated substrates. A stacking of the arene substrates in the proposed and hypothetical ' π -X- π '[289] structure becomes therefore more likely. As a final proof a crystal structure of variant M3 with benzene ligand would be necessary. Improvements in activity for the respective variants follow the same trend as outlined for other substrates in previous sections of this thesis: WT<M1<M2<M3. The reason for the low productivity of the WT (PFR < 10 min⁻¹) is probably caused by residue R47 since substitution abolishes ionic interactions with the NO₂-group of NT substrates (salt bridge formation) and generates access to the active site.[312] Variant M2 has, as shown before (section 3.1 and 3.3), significant higher activity than WT and M1. However, for all NTs the PFR is comparably low, in all cases below 69 min⁻¹ (*m*-NT; Table 20) and more than 20-fold lower than for anisole.[68] The highest product titers, TTNs and products yields were obtained for variant M3 (*o*- and *p*-NT; Table 19 and 21) indicating that the additional aromatic interaction also improves the operational performance of the monooxygenase. Increased product yields and TTN values of M3 can also be explained by the increased coupling efficiencies indicating a more efficient hydroxylation of the substrates that reduces formation of reactive and catalyst inactivating oxygen species such as H₂O₂ (section 1.14.1; IV and VI; Figure 7). The major by-products during conversion of the NTs are NBAs and phenolic products (Tables 19, 20 and 21; Figure 59). Aromatic hydroxylation is the highest for *m*-NT (15.5 % GC area; M3; Table 20) which can be explained by the substrate binding mode obtained from the docking study that allows 2,3-epoxidation and re-aromatization to a phenolic product (Figure 60 B). The high amount of *o*-NBA produced by M3 (~33 %; Table 19; Figure 59) cannot be attributed to a dehydrogenase activity, since purified protein was applied during all reactions.[232] One explanation could be that the product re-enters the active site of BM3 and is hydroxylated a second time leading to formation of nitrobenzaldehyde and a molecule of water as proposed in Scheme 12.[313]



Scheme 12. Proposed mechanism for oxidation of *o*-NBOH by variant M3 (Table 19; Figure 59). In the first step P450 BM3 abstracts via compound I a proton from *o*-NBOH, forming an unstable α -C-atom radical (hypothetical intermediate highlighted in brackets). Instead of rebounding the OH-group to the substrate, the C-bound hydroxyl-group performs an electron rearrangement leading to the formation of *o*-NBA under oxidation of the OH-group. The proton, reacts with the heme-bound OH and is released as a molecule of water to restart the P450 cycle.

In addition, the hydroxylation of *p*-toluidine was investigated to develop a new biocatalytic route to aminobenzylalcohols. During substrate screening neither aniline nor *p*-toluidine could be hydroxylated by variant M2 (section 3.2; Table 9). In this part of the thesis, a GDH regeneration system was used together with variant M3 (section 2.10.9.3). GC-FID analysis revealed one distinct product peak that was identified by GC-MS as 6-amino-*m*-cresol indicating that amino substituents are directing hydroxylation to the aromatic ring. The amino group is less electronegative ($\text{NH}_2 = 2.7$ Pauling Units; $\text{NO}_2 = 4.2$ Pauling Units)[258] and has a smaller atomic radius, therefore, π - π -interaction between substrate and F87 could be different, leading to formation of a phenolic product instead of 4-aminobenzylalcohol. A direct route to aminobenzylalcohols employing *p*-toluidine and the presented P450 BM3 catalysts is therefore not feasible and would require more active site engineering.

3.6.6 Summary and Conclusion

Direct *in vitro* hydroxylation of NTs to NBOHs employing a biocatalytic approach can be an attractive alternative to chemical routes omitting multi-step synthesis, intensive use of energy and application of strong acids and toxic metal catalysts. This is the first report of a direct hydroxylation of all three NT isomers to produce NBOHs in an *in vitro* approach using P450 BM3. With a yield of 39 %, a TTN of 1560 and PFR of 202 min^{-1} for *p*-NT conversion, the variant M3 is a promising catalyst for chemical synthesis of *p*-NBOH. Although the catalyst has a good to excellent regioselectivity (>95 %) additional rounds of mutagenesis and screening would potentially lead to a variant with even higher productivity and coupling efficiency. The phenylalanine (F330) in the binding pocket provides additional aromatic interaction which increases activity, productivity, coupling efficiency and TTN of P450 BM3 for all three isomeric substrates. In addition to wet lab investigations, docking of all three isomers was attempted. The docking results were in good agreement with the catalytic data and indicate a peculiar T-shape π - π -interaction of substrate and phenylalanine 330. The possibility of using NO_2 -substituents to direct selectivity in aromatic substrates can further be explored and possibly

transferred to other enzymes and substrates. A successful application of P450 BM3 as *in vitro* catalyst could therefore overcome the limitations of known chemo-synthetic processes as well as drawbacks in biocatalytic approaches, meanwhile offering a direct and green route to NBOHs.

3.7 OmniChange: The Sequence Independent Method for Simultaneous Site-Saturation of Five Codons

3.7.1 Abstract

Focused mutagenesis is one of the key drivers in protein evolution to design biocatalysts towards a specific demand. Mainly localizable and already well investigated and understood properties such as activity and selectivity are target for this strategy. However, available methods are limited by simplicity of the protocols or by freedom to mutate distant codons simultaneously. To overcome these limitations the herein described multiple site saturation mutagenesis method, OmniChange, was developed that efficiently saturates five independent codons. As proof of principle, five chemically cleaved DNA fragments, each carrying one NNK-degenerated codon, were generated and assembled to full length genes in a one-pot-reaction without additional PCR-amplification or use of restriction enzymes or ligases. Sequencing of 48 clones revealed the presence of up to 27 different codons at individual positions, corresponding to 84.4% of the theoretical diversity offered by NNK-degeneration. OmniChange allows for the first time to study systematically and in a modular manner cooperative effects through multiple site substitution of amino acids which are located closely to each other in the 3D protein structure. With the generated protocol it is possible to fuse five DNA fragments in a predefined order without frameshifts, opening new methodological strategies for multi DNA fragment cloning in synthetic biology or metabolic engineering.

3.7.2 Introduction

Iterative rounds of focused mutagenesis and screening represent powerful tools to engineer a biocatalyst towards higher activity or selectivity (section 3.1).[36, 37, 98, 108] Plenty of successful examples for engineering biocatalysts by focused mutagenesis are reported, improving activity, selectivity or stability.[37, 69, 103-107, 110] However, using this strategy, cooperative effects of amino acid residues cannot be investigated to full extent by excluding potentially promising combinations.[36, 223] Cooperativity between closely located amino acid residues is expected to influence properties such as stability by formation of e.g. salt-bridges or aromatic interactions that are reported to stabilize secondary elements in the protein fold. Multi site-saturation mutagenesis will allow identifying cooperative amino acid substitutions which would not have been discovered by saturating single codons individually[106, 314] or iteratively.[36, 107, 223, 315] The success of finding such cooperative substitutions depends strongly on the quality of the generated mutant library as well as complementing screening systems.[36, 37, 98] The semi-rational CASTing approach can efficiently saturate neighboring codons[103, 106, 177] and makes use of the very robust whole

plasmid amplification method[126, 226] which was commercialized by Stratagene as QuikChange Site-Directed Mutagenesis Kit. However, the CASTing method is limited to amino acid residues that are located close to each other on the DNA sequence.[36, 103, 223] As further extension of the method, an iterative CASTing approach was successfully employed to improve enantioselectivity of an epoxide hydrolase[107] and a P450 monooxygenase[177] through iterative rounds of multi site-saturation mutagenesis of closely located residues combined with high throughput screening.

Although ISM[104], CASTing[103] and iterative CASTing[177] lead to impressive improvements in re-design of proteins, the high DNA sequence dependency encouraged scientist to develop new methods for multi-site saturation mutagenesis to explore the full sequence space offered by nature. [36, 37, 98] Several multi codon mutagenesis methods were developed over the last years differing in the number of simultaneously mutated positions, PCR steps and application of DNA modifying enzymes.[97, 104, 125, 316-320] A compilation of the most prominent methods for single and multi-site saturation mutagenesis methods and recent developments are summarized in two recent reviews.[36, 37] Major limitations of the reported methods are dependency on the DNA sequence, multiple PCR steps, multiplex PCRs, additional enzymatic modifications and final cloning steps.[223] Although chemical synthesis could be an alternative to generate *the* perfect mutant library[321, 322], the quality, time requirements and costs are not competitive to polymerase based methods at the moment.[36, 322] Consequently, a robust and reliable method for multi-site saturation mutagenesis is highly desirable enabling to explore the full potential of a biocatalyst.

Based on the known limitations, a new and robust method for multi site-saturation independent from DNA sequence restrictions should be developed in this part of the thesis. The concept should be based on a simple and robust protocol which enables introduction of up to five NNK degenerated codons[105] into a GOI, independent from positioning of the targeted positions in the DNA and protein sequence.[36] Therefore, a concept for a simple and robust 4-step protocol based on the PLICing cloning technology was developed.[213] The concept includes a standard PCR step for introduction of NNK codons into fragments of the target gene (Figure 63). Reassembly of the full plasmid should be achieved by cleavage of phosphorothiolated nucleotides from amplified DNA fragments as described by Eckstein et al.[222] and applied successfully in the PLICing gene cloning protocol.[213] After re-hybridization to full length plasmids the construct should contain NNK codons on desired positions.

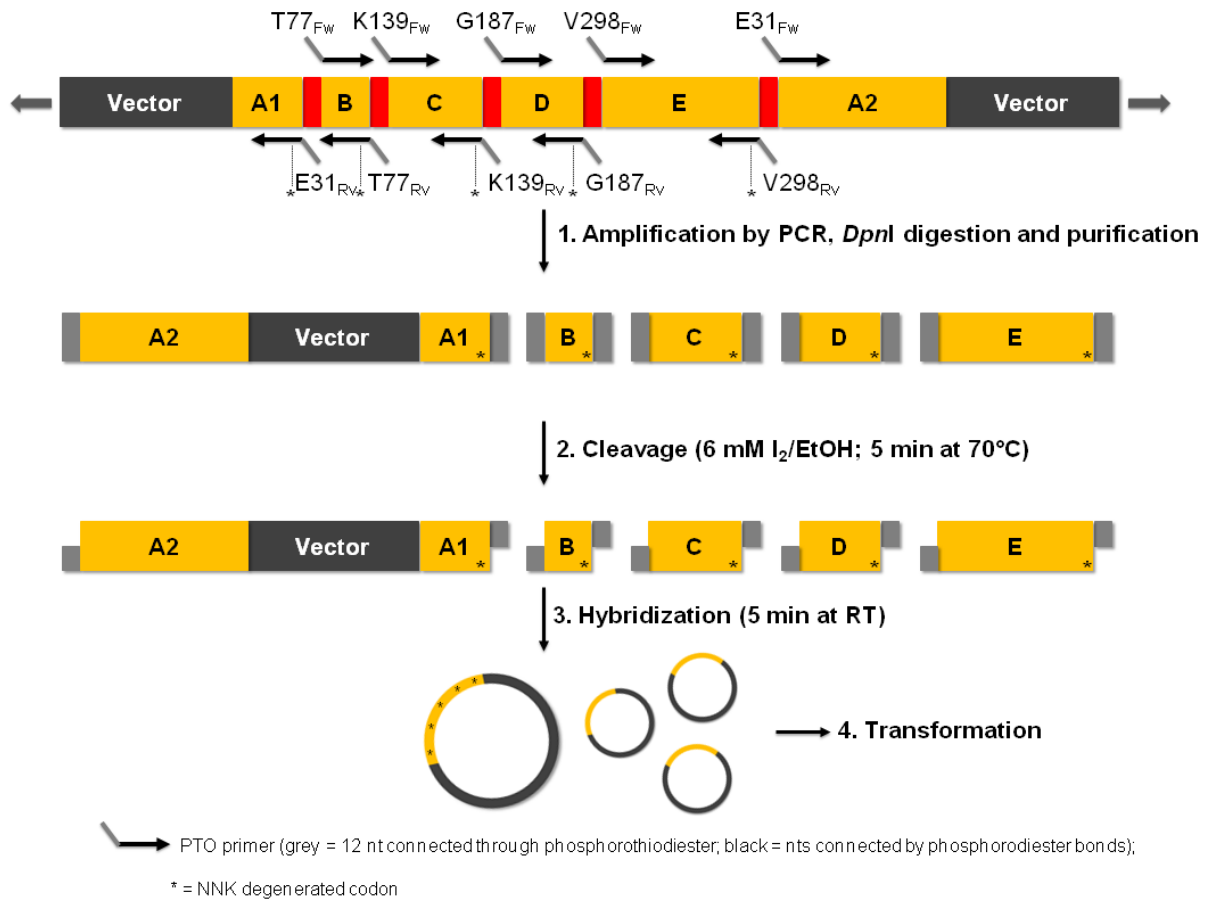


Figure 63. The OmniChange concept comprises four steps[223]: Step 1. Amplification of five DNA fragments introducing a NNK-saturated codon (indicated with *). Step 2. Chemical cleavage to generate complementary single-stranded 5'-overhangs. Step 3. Hybridization of cleaved fragments to a full circular plasmid containing ten DNA nicks. Step 4. Transformation and nick-repair in *E. coli* BL21-Gold (DE3) lacI^{Q1}. The Figure was redrawn and adapted from Dennig et al.[223]

Five codons were selected in the phytase from *Yersinia mollaretii*[214] to serve as a model gene during development of a protocol. The theoretically generated diversity (32^5 codon combinations[127]) lies in the range of state-of-the-art screening platforms for oversampling (10^8 to 10^9 clones).[129, 223] Generated libraries should be benchmarked to reported methods with regards to efficiency, robustness and quality of the generated diversity.

3.7.3 Experimental

3.7.3.1 Employed Strains and Vectors

The 1.3 kb phytase gene from *Yersinia mollaretii*[214] was cloned into the pALXtreme-5b plasmid by using the PLICing cloning method (section 2.9.5).[213] Transformation of plasmids was achieved by a heat shock protocol and using chemically competent *E. coli* BL21-Gold (DE3) lacI^{Q1} cells with a determined transformation efficiency of 1.4×10^6 cfu μg^{-1} pUC19 (section 2.8.2).

3.7.3.2 Mutagenic Oligonucleotide Design and Preparation of DNA Fragments

Target positions for simultaneous saturation mutagenesis were selected based on a phytase variant obtained from a SeSaM directed evolution experiment[115] that displayed improved thermostability.[214] Ten oligonucleotides were designed (Table 22) for simultaneous saturation of the five selected codons (E31, T77, K139, G187, V298), as indicated in the 4-step concept (Figure 63). All oligonucleotides were designed to maximize product yields in PCRs (section 2.9.3). The reverse primers harbored the degenerated NNK codon that was introduced by PCR into the respective DNA fragments (Step 1; Figure 63). For efficient and selective hybridization of amplified DNA fragments in defined order the ten oligonucleotides have at least 12 or 13 phosphorothioate nucleotide bonds at the 5'-terminus (Table 22).[213] The principle for cleavage and hybridization of amplified DNA fragments is described in more detail in material and methods (section 2.9.5; Figure 17) and in the work of Eckstein et al. and Blanusa et al.[213, 222]

Table 22. List of used primers during method development. Purpose, primer name and the respective sequence are shown. The degenerated codon is underlined, whereas the phosphorothiolated nucleotides are indicated as lowercase letters on each 5'end. Capital letters indicate the standard (natural) nucleotides. Parts of the content in this table have been published by Dennig et al.[223]

Fragment Amplification	Primer Name	Primer Sequence (5'-3')
Vector backbone	E31 _{FW}	ctagtgcctcagCGTAAGGGGCAAG
	E31 _{RV}	gataaccactcgMNNCAAAGTGTAACCCGTC
B	T77 _{FW}	cgagtggttatcTTGAGCCGCCATG
	T77 _{RV}	catagaagccgccCATCAGMNNCACTAATTGCGC
B₂	D52 _{FW}	ccggataagtggCCTCAATGGCCGGTAC
	D52 _{RV}	ccacttatccggTGTGACMNNATTCATTA ACTCTG
C	K139 _{FW}	ggcggcttctatgGTGATTATTTCCG
	K139 _{RV}	gaaacagtgatcAACCTTMNNCAAATCAGCCTG
D	G187 _{FW}	gatccactgtttcACCCCGTCGAAG
	G187 _{RV}	cagtgaaattgagAATCTCMNNCATCTGGGCAAATG
E	V298 _{FW}	ctcaatttcactgCTTCCCCCTATTGC
	V298 _{RV}	ctgaagcactagCGCCGTMNNAATCTGTTGCAAC

Since the primer design is a crucial step during the method development, a detailed guidance is shown exemplary for positions E31 and T77 in the model phytase gene (Figure 64). At first two or more positions are selected within the DNA sequence together with the respective degeneracy (Figure 64 A). Subsequently 12 nts downstream of the selected codon are selected which represent the

later PTO part and 5' end in the primer sequence (Figure 64 B). In the last step, 3' ends of the respective oligonucleotides are designed with regards to optimal PCR product yields (Figure 64 C).

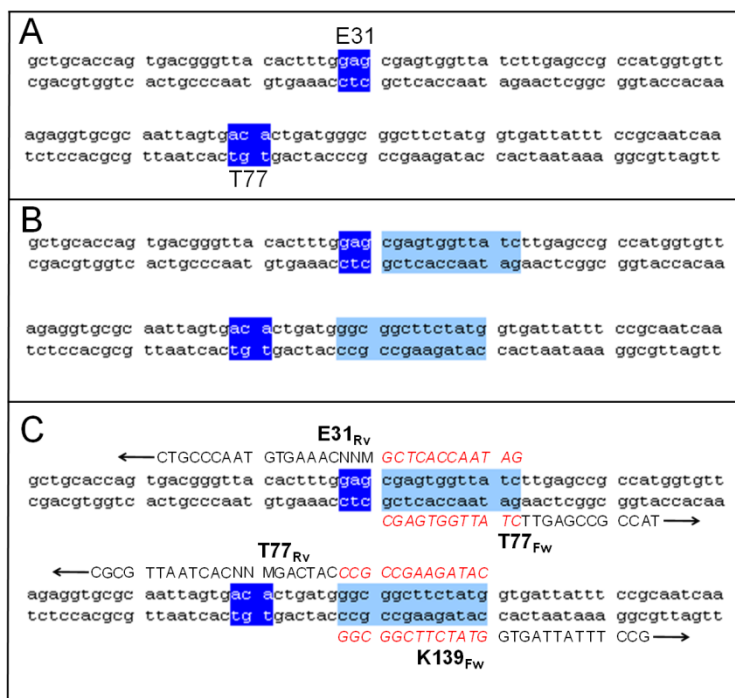


Figure 64. Three steps in oligonucleotide primer design explained exemplary for positions E31 and T77.[223] **A:** Selection of targeted codons for NNK saturation (dark-blue highlighted letters). **B:** Twelve nucleotides downstream of targeted codon are selected as phosphorothiolated nucleotides that can be cleaved for overhang generation, enabling hybridization of generated DNA fragments to full length plasmids (light-blue highlighted letters). **C:** Design of oligonucleotide 3' end for successful PCR amplification of DNA fragments. Arrows indicate Phusion DNA Polymerase amplification direction. Red labeled italicized letters represent phosphorothiolated nucleotides within primers used for method development.[213, 222] The Figure was taken from Dennig et al.[223]

3.7.3.3 Mutant Library Generation

All DNA fragments were amplified using Phusion DNA polymerase and standard PCR conditions as described in the material and methods section (section 2.9.2). PCR products were analyzed on a 2 % agarose gel to estimate quality of amplified DNA (section 2.9.6). Residual template DNA was removed by DpnI digest. After purification the PCR products were eluted in 50 μL ddH₂O water to avoid interference with the chemical cleavage reaction in step 2 (Figure 63). The DNA concentration was estimated and calculated to $\text{pmol } \mu\text{L}^{-1}$ to ensure equimolar amount of fragments in the reaction mixtures. The final protocol[223] employs purified DNA fragments diluted to 0.02 $\text{pmol } \mu\text{L}^{-1}$ (A₂-Vector-A₁) and 0.11 $\text{pmol } \mu\text{L}^{-1}$ (fragments B, C, D, and E). The PCR-products were cleaved according to the PLICing cloning technology[213] using 2 instead of 1 μL cleavage mixture. Hybridization of DNA fragments was performed stepwise by first adding 6 μL of cleaved fragment B to 6 μL of cleaved fragment A₂-Vector-A₁ (20°C with 5 min incubation time). Subsequently, remaining fragments C, D and E were supplemented gradually with 5 min incubation between each fragment. The assembly mixture was kept on ice for additional 5 min to allow hybridization of DNA overhangs. As self-hybridization control of the A₂-Vector-A₁ fragment the inserts (B, C, D, E) were replaced each

by 4 μL Milli-Q water (replacing the four inserts). Four μL from both hybridization mixtures were transformed via heat shock of 100 μL chemical competent *E. coli* BL21 (DE3) lacI^{Q1} cells (section 2.8.1). After recovery on LB agar plates the obtained colonies were counted and evaluated for statistical purpose (section 2.8.2). Correct assembly of full length plasmids was verified by colony PCR (section 2.9.4) and agarose gel electrophoresis on 1 % (W/V) agarose gels. Plasmids were sequenced and analyzed for quality of the mutant library. To proof functionality of the re-assembled genes the phosphatase activity of 20 randomly picked colonies was measured in a plate reader, according to a published protocol with the 4-methylumbelliferylphosphate assay.[214] As reference, colonies with the template gene (phytase WT) and colonies only carrying the empty pALXtreme-5b vector were used.

3.7.4 Results

Large parts of this chapter were published in *PLoS One* (2011)[223] including a statement for each author's contributions.

3.7.4.1 PCR Amplification and Library Assembly

Successful PCR-amplification of all DNA-fragments (Step 1) was verified on a 2 % agarose gel (Figure 65). After purification the DNA fragments were quantified spectrophotometrically and diluted to the respective concentration.

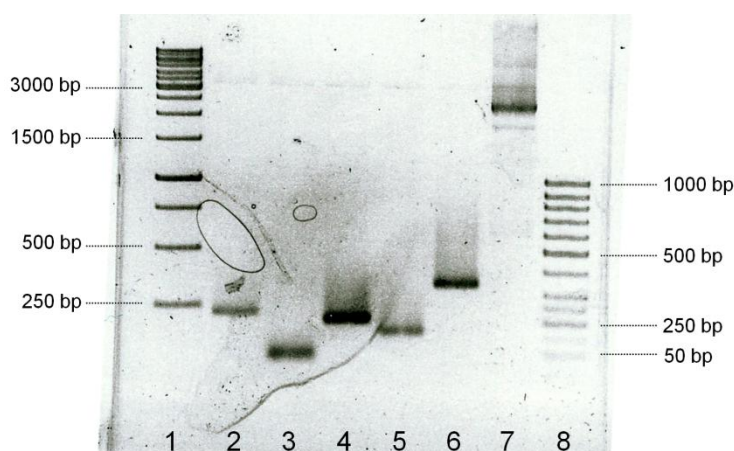


Figure 65. Exemplary separation of DNA on a 2 % agarose gel from step 1 in the OmniChange protocol: PCR amplification of DNA fragments for an OmniChange library (target positions D52/T77/K139/G187/V298). Lane 1: 1 kb DNA ladder. Lane 2: Fragment B; Lane 3: Fragment B₂; Lane 4: Fragment C; Lane 5: Fragment D; Lane 6: Fragment E; Lane 7: Vector backbone; Lane 8: 50 bp DNA ladder. In the final protocol the fragment in lane 2 was integrated into vector backbone amplification to omit an additional fragment and to decrease likelihood of wrong hybridization.

The smallest fragment that was amplified by PCR was fragment B₂ with 88 bp length (Figure 65; lane 3). DNA amplification using the PTO modified primers was performed as with standard PCR primers containing phosphodiester bonds without further optimization of PCR conditions. Product yields after purification of PCR products (from 50 μL reaction volume) commonly range from 20 to 60 $\text{ng } \mu\text{L}^{-1}$ for the inserts (corresponding to 0.2 to 0.5 $\text{pmol } \mu\text{L}^{-1}$) and 80 to 140 ng

μL^{-1} for the vector backbone (~ 0.05 to $0.08 \mu\text{mol } \mu\text{L}^{-1}$). Assembly of the mutant library (Step 2 and 3) was performed according to the developed concept (Figure 63) and as described in the material and methods section of this chapter. In the first experiments, using identical DNA insert and vector ratios (1:3) as reported for the PLICing cloning method,[213] no significant changes between the library plate and vector self-hybridization control were observed. Therefore, the amount of each insert was increased linearly by $0.02 \mu\text{mol } \mu\text{L}^{-1}$ DNA up to a final concentration of $0.11 \mu\text{mol } \mu\text{L}^{-1}$ (Figure 66). All transformations were done in triplicate for each assembly.

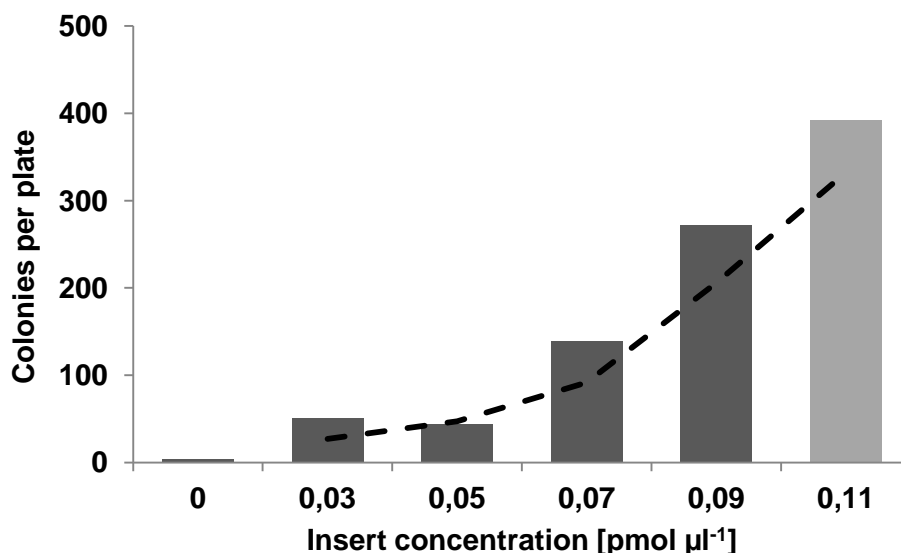


Figure 66. Dependency of insert concentration and obtained colonies after transformation of an OmniChange library and regeneration on agar plates. The amount of used vector backbone was kept constant with $0.02 \mu\text{mol } \mu\text{L}^{-1}$ while insert concentration was linearly increased by $0.02 \mu\text{mol } \mu\text{L}^{-1}$. Most colonies were obtained using a ratio of 2:11 (vector to insert).

The amount of vector was kept constant to ensure a low background of EV in the later library. After transformation, most colonies (≈ 400 per plate) were obtained applying a vector to insert ratio of 2:11 corresponding to a transformation efficiency of $22000 \text{ cfu } \mu\text{g}^{-1}$ for A_2 -Vector- A_1 (Figure 66; grey column) (section 2.8.2). The amount of colonies on all other plates was significantly lower showing a nearly linear dependency between insert concentration and obtained colonies. The vector re-hybridization control displayed 4 to 10 colonies per plate corresponding to 1 % of the total library with vector to insert ratio 2:11. Since libraries of $>10^7$ clones are necessary to oversample the theoretically generated diversity[127] the application of T4 DNA ligase within the protocol was considered to increase the efficiency by ligation of the ten DNA nicks in the final assembled plasmid. However, application of T4 DNA ligase did not increase transformation efficiency. As next step during library optimization a touchdown protocol was tested for gently and specific hybridization of cleaved fragments by cooling down the assembled mixture gradually from 70 to 4°C with $1^\circ\text{C } \text{min}^{-1}$. Applying this additional heating protocol the amount of colonies was decreased to 75 % (119 colonies obtained from vector to insert ratio 2: 11) compared to the samples without touchdown PCR.

3.7.4.2 Quality Control of Generated Libraries

Twenty randomly picked colonies from the generated library (Figure 66; ratio 2:11) were analyzed by colony PCR (section 2.9.4). All clones showed that the complete phytase gene information (~1.3 kb)[214] could be reassembled by the developed protocol, leading to full size plasmids (Figure 67). A restriction analysis was performed (data not shown) and confirmed in all cases full length plasmids.

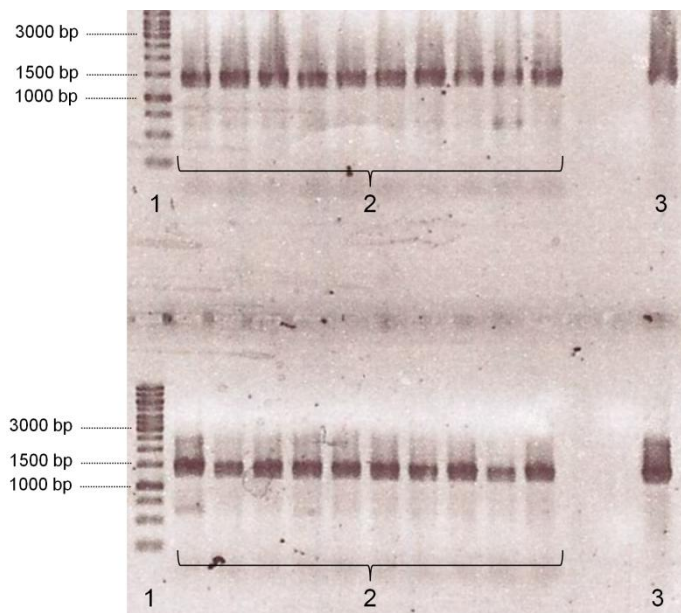


Figure 67. Colony PCR of 20 clones randomly picked from agar plates after transformation of *E. coli* BL21 (DE3) $lacI^{Q1}$. 1: 1 kb DNA ladder; 2: 20 random clones from a phytase OmniChange library. 3: Positive control of colony PCR applying the template plasmid (pALXtreme-5b) containing the WT phytase gene.

Five clones from the library were sequenced to verify the correct assembly of plasmids and the introduction of mutated codons at the selected positions. The sequencing confirmed that all five randomized positions were unique and different from the corresponding WT codons (Figure 68).

		D52	T77	K139	G187	V298
pALXtreme-5b	97	gaatgatgtca	tagtgacactgat	atttgaaaaagg	gatggcgaga	gattgtgacg
MSSMC5-T7.ab	146	gaatgtggtca	tagtgcttctgat	atttgctgaagg	gatgactgaga	gatttagacg
MSSMC1-T7.ab	151	gaattcgggtca	tagtgaactctgat	atttggtgaagg	gatgatggaga	gattaatacgt
MSSMC2-T7.ab	147	gaatatggtca	tagtgccggtgat	atttgagtaagg	gatgtatgaga	gatttcgacg
MSSMC3-T7.ab	148	gaatagggtca	tagtgacactgat	atttgaaaaagg	gatggcgaga	gattaatacgt
MSSMC4-T7.ab	147	gaattcgggtca	tagtgtttctgat	atttgcgtaagg	gatgatggaga	gattgacgacg

Figure 68. Sequence alignment of five clones from an initially prepared library. Positions D52, T77, K139, G187 and V298 are highlighted with white background. The phytase WT gene[214] was used as template for the alignment.

A detailed statistical analysis was performed by full gene sequencing and alignment of 48 randomly picked clones to the WT template. The alignment with the WT sequence displayed that all genes were in frame and all five codons fully saturated. None of the sequenced and analyzed clones retained the original WT sequence. In a new library that was generated by the identical protocol, from here on

called the OmniChange method[223], the result was confirmed by simultaneously saturating the positions E31, T77, K139, G187 and V298 (Figure 69).

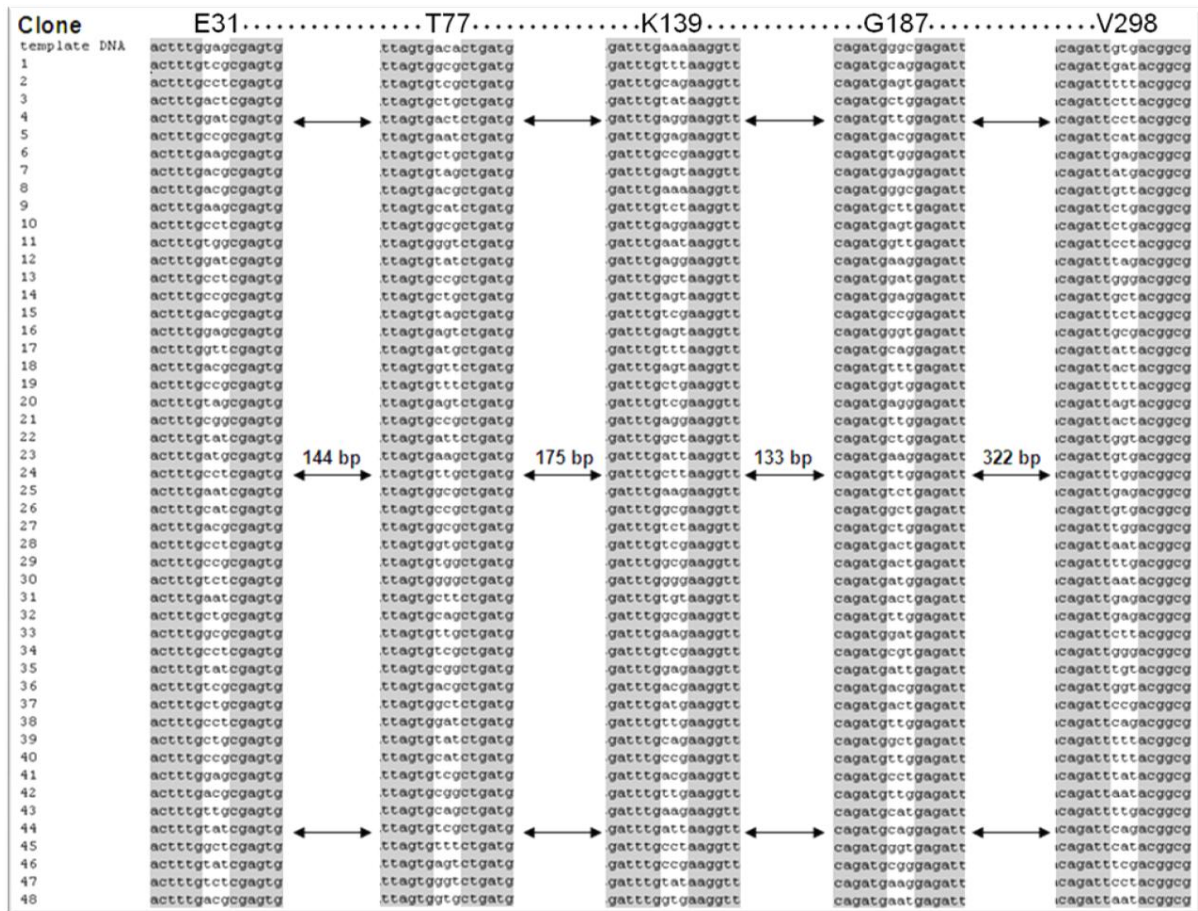


Figure 69. Obtained diversity shown as partial sequence alignment of 48 randomly selected clones from a generated OmniChange library (E31/T77/K139/G187/V298). The phytase WT gene (template DNA) was used for alignment of DNA from sequenced clones. The five sites targeted for saturation are highlighted in white whereas the original WT sequence is highlighted in grey. Arrows indicate the distance between the targeted positions in base pairs. The Figure was taken from Dennig et al.[223]

Detailed comparison of the sequenced clones revealed that one WT codon could be found at three amino acid positions (E31, K139 and G187) respectively two WT codons were detected at position V298 (Table 23). At amino acid position T77 not a single WT codon was identified since the NNK degeneracy does not contain the respective triplet (ACA).[127]

3.7.4.3 Statistical Evaluation of Mutant Libraries

To compare the generated diversity with state-of-the-art methods in multi-site saturation mutagenesis the aligned sequences were statistically evaluated. Although NNK does not contain AAA as codon the respective triplet was introduced at position K139 (Table 23; Clone 8).[127] Twenty-one different codons at position E31 could be observed, representing already 65.6% of the possible diversity after sequencing of 48 clones.

Table 23. Detailed statistical overview on codon diversity generated by OmniChange in a model library. Grey highlighted boxes indicate that the specific triplet occurs two or more times in the sequenced library. Red highlighted and bold type codons represent the original codon in the phytase WT sequence (template DNA). The Table was taken from Dennig et al.[223]

Position	E31	T77	K139	G187	V298
1	aaq	aaq	aaa	aat	aat
2	aag	aat	aag	aag	aat
3	aat	acg	aag	aag	aat
4	aat	acg	aag	aag	aat
5	acg	act	aat	acg	act
6	acg	agt	acg	acg	act
7	acg	agt	acg	act	agt
8	acg	agt	agg	act	atg
9	acg	atg	agg	act	att
10	acg	att	agg	act	cag
11	acg	cag	agg	agg	cag
12	act	cag	agt	agt	cat
13	atg	cat	agt	agt	cat
14	cat	cat	agt	atg	ccg
15	ccg	ccg	agt	att	cct
16	ccg	ccg	atg	cag	cct
17	ccg	ccg	att	cag	cct
18	ccg	cgg	att	cag	ctg
19	ccg	cgg	cag	cat	ctg
20	cct	ctg	cag	ccg	ctt
21	cct	ctg	ccg	cct	ctt
22	cct	ctg	ccg	cgg	gag
23	cct	ctt	ccg	cgt	gag
24	cct	gat	cct	ctg	gag
25	cct	gcg	ctg	ctg	gag
26	cct	gcg	ctt	ctg	gat
27	cgg	gcg	gag	ctt	gcg
28	ctg	gcg	gag	gag	gct
29	ctg	gct	gtg	gag	ggg
30	ctg	ggg	gcg	gat	ggg
31	gag	ggt	gcg	gat	ggt
32	gag	ggt	gcg	gct	ggt
33	gat	gtg	gct	gct	gtg
34	gat	gtg	gct	ggc	gtg
35	gcg	gtt	ggg	ggt	gtt
36	gct	tag	tat	ggt	tag
37	gtt	tag	tat	gtg	tat
38	tag	tat	tcg	gtt	tcg
39	tat	tat	tcg	tct	tct
40	tat	tcg	tcg	tgg	tgg
41	tat	tcg	tcg	tgg	tgg
42	tat	tcg	tct	ttg	tgt
43	tcg	tcg	tct	ttg	ttg
44	tcg	tgg	tgt	ttg	ttg
45	tct	ttg	ttg	ttg	ttt
46	tct	ttg	ttg	ttg	ttt
47	tgg	ttt	ttt	ttg	ttt
48	ttg	ttt	ttt	ttt	ttt
Different codons	21	26	24	27	27
Maximal diversity	32	32	32	32	32
Obtained diversity [%]	65.6	81.3	75.0	84.4	84.4

For all other positions, even a higher coverage could be obtained (Table 23; Figure 70): T77 (26 unique codons; 81.3%), K139 (24 unique codons; 75.0%), G187 (27 unique codons; 84.4%), V298 (27 unique codons; 84.4%).

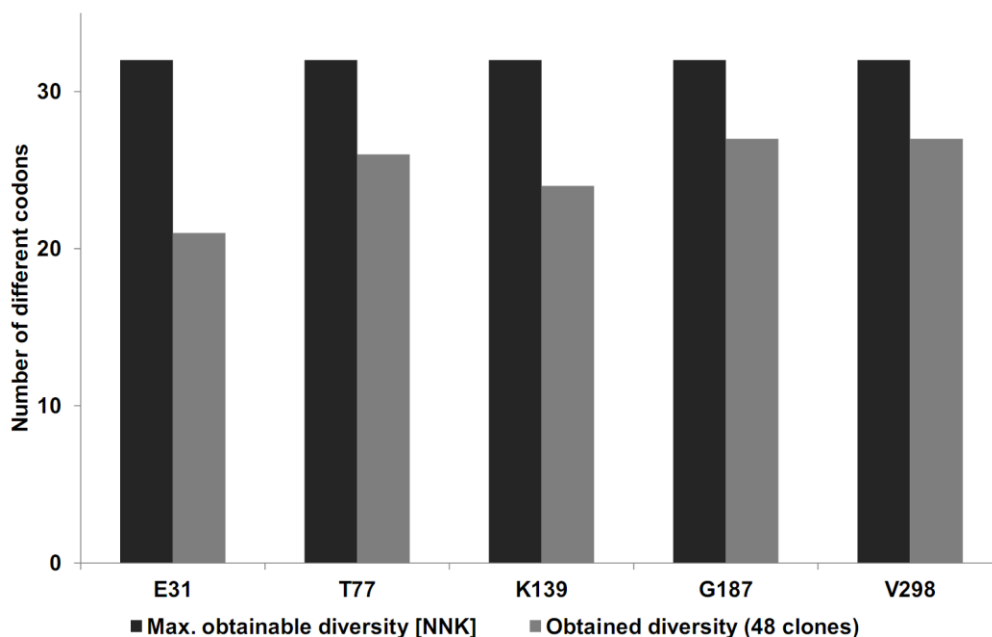


Figure 70. Obtained diversity from NNK saturation of five target codons with the OmniChange protocol (grey bars) compared to the theoretically possible diversity offered by NNK degeneracy (black bars). The lowest diversity is found at position E31 (65.6 %), the most diverse positions are G187 and V298 (each 84.4 %).[223]

As reported for the PLICing cloning method,[213] no additional mutations were introduced during the cleavage reaction with I_2 /EtOH and assembly of the mutated plasmids.[213, 222] Looking at the full sequence a mutation frequency of 0.17 per kb (11 mutations in 62.3 kb) was found within the 1.3 kb phytase model gene, at minimum 21 bp away from the targeted codons.

3.7.4.4 Generation of Sub-Libraries with the OmniChange Protocol

In total five different OmniChange libraries were generated: Library 1: D52/T77/K139/G187/V298; Library 2: G11/T77/K139/G187/V298; Library 3: E31/T77/K139/G187/V298; Library 4: T77/K139/G187/V298; Library 5: D52/T77) varying the combination of oligonucleotides used in PCR. Two additional primers had to be ordered to generate library 1 and 2 (appendix section; Table 30). Especially the generation of library 5 (primer combinations $K139_{Fw}/D52_{Rv}$ for fragment C-D-E-A₂-Vector-A₁; primer combinations $T77_{Fw}'/T77_{Rv}$ for fragment B'; Table 30) was important, to show that codons with less than 100 bp can be saturated efficiently. Sequencing of clones from both libraries reconfirmed that solely selected codons are saturated with the OmniChange method. Employing the subsets of available primers a modular and combinatorial assembly of DNA fragments is possible with the OmniChange protocol. Although the primers were not further checked for compatibility during the PCR reaction, the respective fragments were amplified and assembled according to the developed OmniChange protocol.

3.7.5 Discussion

Rational and semi-rational protein engineering are key drivers for tailoring proteins towards their specific applications.[32, 36] Especially activity and selectivity are well understood properties that can be improved through iterative rounds of focused mutagenesis and screening for improved variants as shown exemplary in this thesis in chapters 3.1 and 3.4.[69, 103, 104, 106, 107, 177] Commonly, focused mutagenesis studies are performed by SDM or SSM on single amino acid residues with pre-defined codon degeneracy.[37, 97, 98, 126] Using this strategy cooperative effects of amino acid exchanges can only be investigated to a limited extent.[36, 106, 223, 317] Methods for multi site-directed mutagenesis offer therefore the possibility to discover synergistic amino acid substitutions, opening novel opportunities for rational and semi-rational protein engineering and to study cooperative effects of amino acid residues located in close proximity in the secondary structure.[106] Exemplary this was shown in this thesis when saturating R47 and Y51 simultaneously leading to an unreported P450 BM3 variant for *p*-xylene hydroxylation (section 3.1).[69] Integrating also residues such as F87 or A330 into a simultaneous saturation strategy would be challenging to achieve with current methods.[36]

OmniChange was therefore developed for the simultaneous saturation of five independent codons with the possibility to extend the protocol to a modular and combinatorial assembly of independent codon combinations (see appendix; Figure 88 A and B).[223] The opportunity to study additive and/or cooperative effects in proteins opens completely new strategies for engineering biocatalyst to specific demands.[36, 106, 314] This part of the thesis illustrated the challenges in DNA-fragment preparation (optimization of DNA concentration; cleavage mix optimization) and DNA-hybridization (sequential addition of fragments) during the development of OmniChange. Since multi DNA fragment assembly is commonly regarded as a challenging task[323], it was unpredictable whether the five assembled DNA fragments (Figure 63, Step 3) would be hybridized to full size plasmids and if so, whether the reading frame of the model phytase gene would be fully retained. In addition, the hybridization is based on a thermodynamic hybridization of complementary DNA overhangs.[213] Therefore, it was unknown if the hybridized fragments are bound strong enough especially with regards to the used heat shock transformation protocol (Step 4; Figure 63). To benchmark OmniChange to state-of-the-art focused mutagenesis methods amino acid positions in very close distance (<100 bp; residues D52 and T77) were selected to underline that nearly every position can be targeted without sequence constraints, which is the case for CASTing[103] or the QuikChange Multi Site Saturation Kit.[125]

OmniChange solved all remaining technical challenges for saturation of five codons simultaneously in a four step method (Figure 63).[36, 223] Sequence analysis of the five generated libraries and in depth statistical analysis of 48 randomly picked clones of a model library showed that up to 27 of 32 possible codons could be obtained at individual positions (Table 23; Figure 69). The overall distribution of different codons was unbiased in the final library, since no codon was preferentially

integrated on targeted amino acid positions (Table 23). One key to success in the development was the introduction of the first saturated codon into the vector fragment (Figure 63). In this way, the amplification and hybridization of an additional DNA fragment was circumvented which might increase the efficiency by decreasing the likeliness of wrong fragment hybridization and also demand for reliable DNA ligation in the *E. coli* host (2 additional DNA nicks). Although NNK does not offer AAA as codon[127] the respective triplet was introduced at position K139 (Codon 1; Table 23). One reason could be the exonuclease (3'-5') activity of the Phusion DNA polymerase, but it is more likely a result of errors during synthesis of the respective oligonucleotides.

The possibility to amplify and hybridize a fragment of only 88 bp allows targeting of the whole gene sequence with OmniChange, which is in case of the QuikChange Multi Site Directed Mutagenesis protocol very challenging[125] and with CASTing impossible (Table 24).[103] All analyzed clones were unique and contained full-length and functional plasmids, which is essential in a subsequent screening approach where high quality mutant libraries are requested.[37, 98] A transformation efficiency of 2.2×10^4 cfu per μg A₂-Vector-A₁ using chemically competent *E. coli* BL21-Gold (DE3) lacI^{Q1} cells (1.4×10^6 cfu per μg pUC19) can be regarded as good for multi-fragment DNA cloning.[323] However, even if highly competent cells with transformation efficiencies of up to 10^9 cfu μg^{-1} would be used, the total amount of colonies would reach 10^7 cfu μg^{-1} , which is still 10-fold lower than the 10^8 variants required to oversample the full diversity offered by NNK degeneracy.[127]

The transformation of nicked DNA is commonly regarded as a crucial and limiting factor for fusion of two or more DNA fragments.[323] Enzymatic ligation of nicked DNA employing for instance the T4 DNA ligase is a standard approach used for cloning experiments to increase transformation efficiencies. OmniChange, however, enables transformation of circular DNA fragments with up to ten DNA nicks without further enzymatic treatment.[213, 223] The high efficiency underlines that 12 nt long single-stranded DNA overhangs are sufficient for reliable, selective and strong hybridization of the DNA fragments (Figure 63).[36, 124, 213, 223] The development of OmniChange underlines once more the exceptional potential of ligase-free or enzyme-free methods in genetic engineering[213], as well as the potential of the *E. coli* host to repair DNA nicks with its own natural ligase(s).

Another important issue during the development of OmniChange was the potential of I₂/EtOH in respect to secondary mutations within the generated library. Eckstein and Gish reported in detail the application and cleavage mechanism of phosphorothiolated nucleotides in DNA and RNA as well as their beneficial use for DNA sequencing applications.[222] The proposed cleavage mechanism for phosphorothiolated nucleotides through alkylation (Figure 17) is highly regioselective and does not lead to a traceable cleavage or modification of natural phosphodiester bonds.[124, 213, 223] In addition, the recently reported PLICing cloning method makes also use of this specific cleavage reaction using phosphorothiolated nucleotides reporting similar transition mutation frequency (1 transition in 9 kbp sequenced DNA)[213] as shown for the OmniChange protocol (11 transitions in

62 kb sequenced DNA).[223] The 11 transition mutations in the evaluated OmniChange library can be attributed to the employed DNA polymerase.[117]

In future, chemical gene synthesis could develop into an attractive alternative for multi site-saturation mutagenesis libraries reducing oversampling requirements in highly diverse libraries (>5 positions saturated).[321, 324] However, current state-of-the-art chemical DNA synthesis methods are at least one order of magnitude more expensive, employ expensive DNA synthesizers and are still restricted to the size of DNA fragments for an error-free synthesis.[322] Furthermore, saturation of more than five positions at the same time (with e.g. NNK degeneracy) is however not desired in directed evolution experiments due to limitations in throughput of screening technologies that do not imply selection systems.[36, 129, 223] State-of-the-art ultra high throughput screening technologies (micro fluidics or flow cytometer based technologies) allow efficient oversampling of mutant libraries with up to five simultaneously NNK-saturated codons (3.2×10^6 different protein variants; 3.4×10^7 different codon variants).[132, 135, 137, 325] For instance, sampling of up to 10^8 variants would completely cover the theoretical sequence space of five simultaneously NNK-saturated codons with a confidence level of 99.6%.[127] By introduction of a 6th NNK-saturated codon only 52% of the generated diversity can be explored when screening 10^8 variants.[127] Nevertheless, the degeneracy of the genetic code would allow varying degeneracy of codons to limit diversity and therefore targeting more sites that fulfill requirements in oversampling as shown successfully for the engineering of a glucose oxidase with OmniChange.[36, 105, 326]

Benchmarking OmniChange with other methods in multi-site saturation/directed mutagenesis revealed that major limitations such as enzymatic modifications, sequence restriction, targeting of multiple independent codons and long and tedious protocols were circumvented.[36] A compilation and comparison of performance parameters and new developments in focused mutagenesis methods is enclosed in recent reviews.[36, 37] Table 24 summarizes and compares the most prominent, advanced or “best” characterized methods for multi-site mutagenesis according to key performance parameters and experimental requirements.

Table 24. Overview of methods and strategies for focused mutagenesis on multiple positions. The Table was published in Dennig et al.[223]

Performance	POEP	OD SPM	ISOR	Iterative CASTing	QuikChange Multi Site-Directed Kit	OmniChange
SDM*/SSM**	+/-	+/-	-/+	-/+	+/+	-/+
Simultaneously mutated sites	8	11	5.6***	3	3 (SDM* 5)	5
Mutation efficiency in %	100	100	12	n.r.	55 (SDM* 32)	100
Max. SSM coverage % (NNK)	n.r.	n.r.	n.r.	n.r.	n.r.	66-84 (48 clones)
Unique clones (%)	n.r.	n.r.	n.r.	n.r.	50 (40 clones)	100 (48 clones)
Frameshifts, deletions, insertions	n.r.	n.r.	n.r.	n.r.	n.r.	0
Restricted in distance between targeted positions ^x	No	Yes	Yes	Yes	Yes	No
Experimental procedure						
No. of DNA amplification steps	2 (b, c)	2 (a, d)	3 (a, b, c)	1 (e)	1 (a)	1 (d)
Primer modification	No	5' phosphorylation	5' biotinylation	No	5' phosphorylation	5' PTO [#]
Agarose gel extraction	Yes	Yes	Yes	No	No	No
Endonuclease/Ligase	No	Yes	Yes	No	Yes	No
Final cloning step	Yes	Yes	Yes	No	No	No
Transformed DNA	dsDNA	dsDNA	dsDNA	dsDNA	ssDNA	nicked dsDNA

*Site Directed Mutagenesis; **Site Saturation Mutagenesis; ***45 positions targeted; n.r. (not reported);^xLimitations or restrictions for distance between targeted codons; [#]12 phosphorothiolated nucleotides on 5' end; a) Linear DNA amplification b) OEP-PCR c) Assembly/nested PCR d) Standard PCR e) PCR with complementary primers.

Especially the low technical requirements and simplicity of the protocols, which is a key to become a successful method[36], where fulfilled by OmniChange (Table 24). For instance OmniChange outperforms ISOR[317], POEP[318], OD-SPM[319] and the QuikChange Multi Site-Directed Kit[125] by avoiding agarose gel extraction and final cloning steps as well as use of restriction enzymes. CASTing is based on the very robust whole plasmid amplification[126], widely used and possibly the method of choice for saturation of multiple codons in close distance.[36] It is, however, not possible to saturate independent codons (in a single PCR step) with the same protocol due to technical limitations. OmniChange, however, will allow saturation of neighboring codons or independent codons with minimal sequence restrictions.[36, 223] To underline this wide flexibility in saturating multiple independent codons a modular assembly of five different libraries was shown in this chapter. The prediction and selection of the optimal residues and codon degeneracy remains difficult to achieve[36], therefore, the possibility to saturate only two or four codons with the same set of oligonucleotides makes OmniChange even more versatile (appendix section; Figure 88 A and B). By applying the same OmniChange protocol it was possible to generate “sub-libraries” which makes OmniChange a modular method matching throughputs of any employed screening system.[132, 135, 137, 325]

3.7.6 Summary and Conclusion

Rational protein engineering is a key driver in directed evolution of enzymes in all fields of biotechnology. OmniChange is a conceptually novel, robust, time-efficient and sequence-independent method for focused multi site-saturation mutagenesis that addresses the current demands in protein engineering. The developed 4-step protocol is based on a simple standard PCR amplification, combined with a robust chemical cleavage step circumventing enzymatic modifications and additional purification steps. The minimum requirements are twelve complementary nucleotide overhangs that are generated by selective chemical cleavage of PTO modified DNA. Sequence analysis of the five generated libraries and in depth statistical analysis of 48 randomly picked clones of a model library displayed that up to 27 of 32 possible codons (84 % diversity) could be obtained at individual positions. In addition, OmniChange offers for the first time to study systematically and in a modular manner cooperative effects through multiple site substitution of amino acids which are located closely to each other in the 3D protein structure. OmniChange will very likely allow saturation of more than five codons simultaneously or implementing primers with multiple saturated codons or different codon degeneracy. Deciphering structure-function relationships in proteins will allow not only improving activity but also stability in organic solvents and thermal inactivation. With the present protocol it is possible to fuse five DNA fragments in a predefined order without frameshifts, although multi nicked plasmid DNA is transformed directly into *E. coli*, opening new methodological strategies for multi DNA fragment cloning for instance in the fields of synthetic biology or metabolic engineering.

3.8 Phosphorothioate-based DNA Recombination: An Enzyme-free Method for the Combinatorial Assembly of Multiple DNA Fragments

3.8.1 Abstract

In this part of the thesis the development of a new method for efficient and nearly sequence independent recombination of genes with low similarity is described, the so called Phosphorothioate-based DNA recombination (PTRec). This conceptually new recombination method offers the possibility of rational crossover adjustment and to recombine genes on the amino acid level, meaning only a short stretch of four consecutive and conserved amino acids (12 nt) is necessary to design a crossover point. PTRec is based on the use of the phosphorothioate chemistry that allows generation of complementary DNA overhangs to fuse multiple DNA fragments in a selective and defined order as shown before during development of OmniChange. The distribution of the five different domains displayed a nearly ideal statistical distribution and after analysis of a model library with three phytase genes (45 to 53 % identity) 88 % of the 42 sequenced clones was unique. With this new method, major drawbacks of known methods were overcome by providing a simple protocol, easy rational adjustment of crossover points and the potential to recombine sequences with less than 50 % sequence dependency.

3.8.2 Introduction

In vitro recombination of DNA sequences was first introduced by Stemmer in 1994 as Family or DNA Shuffling.[122] The peak of development for new recombination methods was reached between 1998 and 2003[36, 124] where a broad array of recombination methods were reported such as StEP[327], RACHIT[328], ITCHY[329], SCRATCHY[330], SHIPREC[331], OE-PCR[121] and SISDC[332]. The main difference between DNA recombination and random mutagenesis methods is the overall generated diversity.[36, 37, 98] DNA recombination methods generate chimeric proteins by recombining large DNA molecules.[333] This strategy allows circumventing stop codons while larger functional elements are exchanged, generating more “meaningful” modifications in a protein structure.[36] Oversampling of recombined protein libraries commonly does not require ultra-high throughput screening platforms since the generated diversity is comparably lower than for random mutagenesis. The most diverse methods recombine up to three genes and are capable to rationally introduce between 1 and 4 cross-over points per gene (Table 26).[124] Although several methods are available for generation of protein chimeras the main challenges could not be solved such as high sequence dependency, low amount of crossover points, technical difficulties during multiplex PCR amplifications and additional enzymatic steps.[36, 124] As a consequence the recombination of distant related genes is difficult to achieve, for instance for genes from extremophiles with mesophilic enzymes.[36]

Main conceptual advances in the last three years were achieved by the methods TMGS-PCR,[334] USERec,[335] Golden Gate Shuffling,[336] and Integron[337]. These have in common the capability to recombine sequences with very low identity with predefined and increased number of crossover points, which overall increases diversity in chimeric libraries.[36] An increased number of crossover points is of great importance to generate large and diverse chimeric libraries, therefore more significantly improved variants can be expected from such a library.[36] Reliable and robust protocols are important to become a widely used method as shown for instance for epPCR[117] or whole plasmid amplification methods (SDM and SSM)[126, 225] which are widely applied in many successful protein engineering campaigns.[103, 104, 107, 128] Due to its simplicity, overlap extension PCR (OE-PCR)[121, 338] is a widely used method for recombination, but limited at least to 15 identical nucleotides to generate a crossover point which is not given especially at low DNA sequence identities. Therefore, important criteria during development of a new method for *in vitro* recombination are, besides the quality of a later library, the simplicity, time-requirements and robustness of the protocol.[36] However, compared to focused and random mutagenesis methods, the applications of methods for recombination of genes is rare, most likely due to the technical challenges of many protocols.

Inspired by the successful development of OmniChange (section 3.7)[223] and the expansion of the PLICing protocol[213] from a gene cloning technology into a multi-fragment assembly method, a new concept for generation of protein chimeras should be developed in this chapter (Figure 71). The fact that PLICing and OmniChange need only a minimum of sequence identity (12 nt) on the gene level to fuse DNA fragments should be the outline for a successful method development. The new concept is based on a simple and robust 4-step protocol where conserved regions of three model phytase genes are identified and used to generate crossovers for recombination, which have at minimum four amino acids identity.

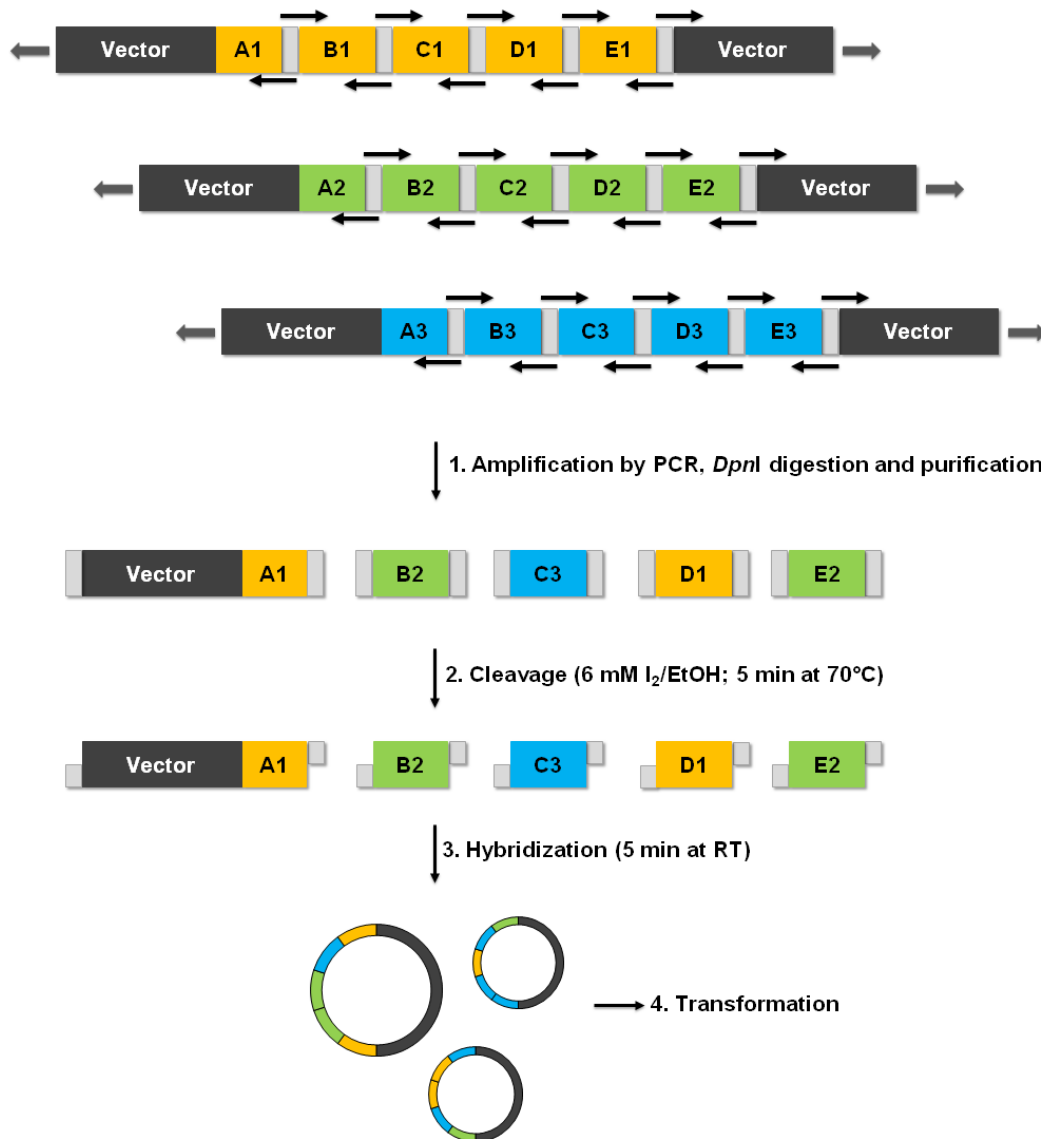


Figure 71. Concept of a 4-step protocol for recombination of three model genes during the development of PTRec. PTRec comprises four steps: (1) amplification of gene specific domains, (2) chemical cleavage using I₂/EtOH (pH 9.0) to generate single-stranded overhangs at pre-defined crossover sites, (3) recombination and hybridization of domains and (4) transformation into chemical competent *E. coli* cells.[124] The Figure was redrawn from Marienhagen et al.[124]

The hybridization areas are generated by the cleavage of the introduced PTO nucleotides on each 5' end of each fragment (Figure 71; Step 1) which represent the rationally generated crossover sections. After library generation a large set of clones should be sequenced and the method benchmarked towards reported methods for *in vitro* recombination of proteins. As proof of concept three bacterial phytase genes should be recombined that have a DNA sequence identity of 45 to 53 %.[124]

3.8.3 Experimental

The described PTRec libraries were generated by Dr. Jan Marienhagen, who also drafted the manuscript published in *BioTechniques*.[124] Alexander Dennig contributed in designing the PTRec

concept and experiments (selection of crossover points, primer design and preparation of original constructs for recombination), analyzing data and in parts during manuscript preparation.

3.8.3.1 Sequence Alignment of Three Phytase Genes for Identification and Selection of Crossover Points

For sequence alignment Clone Manager Professional 9 software was used. The three phytase genes were aligned by amino acid as well as DNA sequence (Figure 72; full DNA sequence alignment is shown in Figure 89 in the appendix section). Four crossover points were selected based on amino acid alignments by selecting four consecutive and conserved residues in all three phytases.

```

appA -----mkailipflsllipltpqsaafaqsepelk-----lesvvivsrhgvraptkatqlmqdvtpdawp
phyX mtislfnrnkpaiaqrilcplivalfsglpayasdtapagfq-----lekvvilsrhgvraptkmtqtmrnvtpqhwp
ympH -----mvkiamnsmrltalglmlgsfaisaapvaapvtgytlervvilsrhgvraptkqtelmndvtpdkwp

appA twpvklgwltprggeliaylghyqqrqlvadgllakkgcpcqsgqvaiiadvderttrktgeafaaglapdcaityhtqadt
phyX ewpvklgyitprgehlislmggfyrerfqqqgllpkdncptpdavyvwdvdqrtrktgeaflaglapqcdlaihhqqni
ympH qwpvpagyiltprgaqlvtlmggfygdyfrnqgll-pagcpadgtlyaqadidqrtrltgqafldgiapgcgklkvyhqadl

appA sspdpflfhpktgvcqldnanvtdailsraggssiadftghrgtafrelervlnfpqsnlclckrekqdescsltqalpsel
phyX qqadpflfhpvkagicsmdksqahaavekqagtpietlnqrygaslalmssvldfpkspycq-qhnigklcdfsgampsrll
ympH kkvdplfhpvpeagvcqldstqtthraieaqlgaplselsqryakpfaqmgeilnftaspycksllqqgkscdfatfaanev

appA kvsadn--vsltgavslasmlteifllqgaqgmpepgwgritdshqwnllslhnaqfyllqrtpevarsratp1llklik
phyX ainddgnkvallegavglstlaeifllhaqgmekvawgnihteqqwdsllklhnaqfdlmsrtppyiaknqtp1lgtia
ympH kvnqqgtkvs1sgplalsstlgeifllqnsqgmvdvahr1sgaenwvslslhnaqfdlmaktpyiarhkqtp1lqqiv

appA taltphppgkqaygvltlptsv----lfiaghdtnlalnlggalelnwtlpgqpdntppggelvferrr1sdnsgwiqvsl
phyX halgsniasrplpdispdnki----lfiaghdtnianisgm1gmtwtlpgqpdntppggalvferrw---dnagkpyvsv
ympH tal---vlqrkgqgqtlplseqtkllflgghdtnianigcm1ganwqlpgqpdntppggglvfelwqnpdnhqyvavkm

appA --vfqt1cqmrdktp1slntpp-gevkl1lagceernagcm1slagftqivnearipacs1*-
phyX nmvyqt1aq1hdqtp1tlqhp1-gsvr1n1pgcsdqtpdgycplstf1r1vnhsvepacq1p*
ympH --fyqtm1dqlrnsek1dkshpagivpieieg1cenigt1dk1c1q1dt1fq1krvaqv1epachi*-

```

Figure 72. Amino acid sequence alignment of three model phytases (appA, phyX and ympH) using Clone Manager Professional 9 software. Blue highlighted letters indicate highly conserved areas between all three phytases. Black dotted frames highlight crossover sections that were selected for recombination of the three phytase genes. The Figure was taken from supplementary information of Marienhagen et al.[124]

Selection of crossover points was not done rationally to functional properties of the respective domains; therefore, no functional screening was pursued and the quality of the library was judged upon sequencing data.

3.8.3.2 Identification of Mature Phytase Sequences by SignalP Algorithm

A pre-cloning step of the phytases was necessary since all three genes were obtained in different bacterial vectors and contained the host specific phytase signal sequence.[214] The signal sequence is commonly removed by the host (*E. coli*) after expression of the protein to ensure for instance transport into the periplasma. Therefore, the signal sequences of each phytase were identified by SignalP algorithm.[339] The highest ‘C score’ indicates the most likely cleavage site in gram negative bacteria

to produce the mature phytase protein. SignalP algorithm calculations were done for protein sequences of all three phytases. Figure 73 shows exemplary the signal sequence prediction for the phytase appA from *E. coli* (section 2.6; Table 3), whereas calculations for the two other phytases can be found in the appendix section (Figure 90 and 91). The mature phytase sequences were cloned by PLICing (section 2.9.5)[213] downstream of the *pelB* leader sequence of the pET22b plasmid (section 2.6; Figure 12) that was used during all experiments in this part of the thesis.

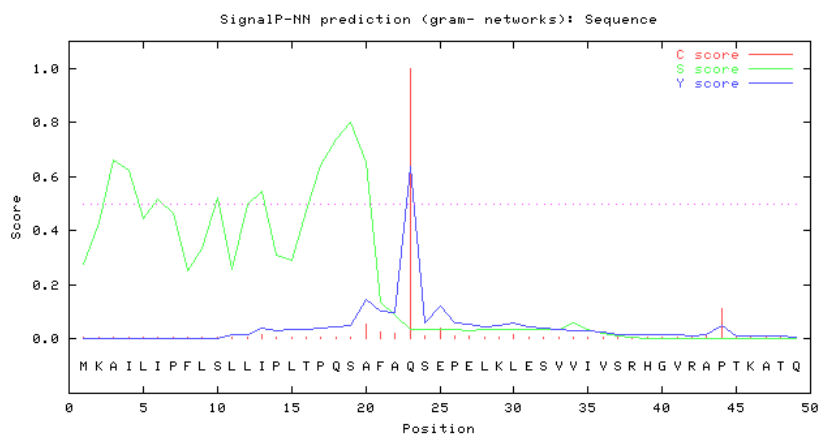


Figure 73. Prediction of cleavage site of appA phytase signal sequence in *E. coli* (gram negative bacteria) using SignalP algorithm.[339] The red vertical line indicates the most likely cleavage position due to highest calculated C score. With 99.9 % probability the cleavage takes place between residue 22 and 23 (Ala and Gln). Diagrams for calculation of the other two phytases (*ympH* and *phyX*) can be found in appendix section (Figures 90 and 91).

A uniform vector for all three phytase allows an unbiased evaluation, independent from varying transformation efficiencies of the generated chimeras. Primers for the initial cloning step of all three phytases (*appA* of *E. coli* (GenBank, AF537219), *ympH* of *Y. mollaretii* (GenBank, JF911533) and *phyX* of *Hafnia alvei* (Patent, U.S. no.2008/0263688 A1) were designed as described before in material and methods section (section 2.9.3). The mature phytase genes were amplified by PCR using Phusion DNA polymerase (section 2.9.2) and phytase specific PTO primers (*appA*, P1 and P2; *ympH*, P3 and P4; *phyX*, P5 and P6; Table 25). The recipient pET22b(+)vector was amplified accordingly. Cloning of the two genes into the pET22b(+) was proceeded as described in the standard PLICing protocol applying the recommended amount of DNA for insert and vector. Colony PCR (primers P33 and P34) was used to confirm correctly cloned constructs before sending isolated plasmids for sequencing.

3.8.3.3 Oligonucleotide Design for Recombination of Mature Phytase Genes

Although the selected crossover points are conserved on the amino acid level, the DNA sequence differs, which is attributed to the degeneracy of the genetic code.[127] For specific and efficient hybridization of DNA fragments, 35 silent mutations were introduced into the 5'- PTO-labeled endings of 19 out of 30 primers (Table 25). Exemplary the primer design at crossover point 3 is shown in Figure 74.

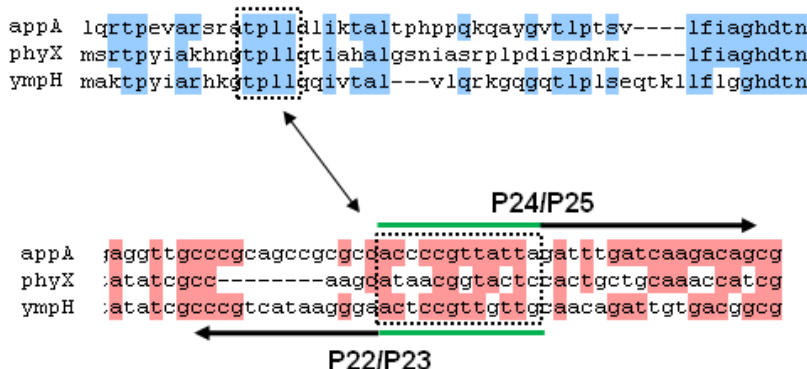


Figure 74. Snapshot of amino acid sequence (upper) and DNA sequence alignment (lower) at crossover point 3 (between domain 3 and 4) for the three model phytases appA, phyX and ympH. Blue highlighted letters (upper alignment) represent identical amino acid residues within the protein sequence. Black-dotted frames highlight the crossover point within the amino acid sequence and the corresponding DNA sequence. Red highlighted nucleotides (lower alignment) indicate identical nucleotides within the aligned gene sequences. The arrows display the binding regions of oligonucleotides P22, P23, P24, and P25 for amplification of appA and ympH. The green line represents the 12 PTO nucleotides on 5'-endings of each primer sequence. The Figure was taken from supplementary information of Marienhagen et al.[124]

The full list of all designed and used primers during the development of PTRec are summarized in Table 25.

Table 25. Full list of oligonucleotides used during the development of PTRec and their specific purpose, including oligos for sequencing and initial sub-cloning of mature phytase sequences into pET22b vector. Lowercase letters indicate PTO nucleotides. Capital letters represent nucleotides connected through natural phosphodiester bonds; Underlined nucleotides are mutated silently to generate complementary 5'-overhangs. The table was taken from supplementary information of Marienhagen et al.[124]

Primer Name	Purpose	Sequence (5'-3')
P1	ampl. of pET22b(+) for subcloning / universal fw-primer for ampl. of domain 1	gcttgcggccgcACTC
P2	amplification of pET22b(+) for subcloning	gtgatgatgatgATGATGCGGATCCGAATTAATTCCGATATCCATG GCCATCGCCGGC
P3	subcloning of appA into pET22b(+)	catcatcatcacCAGAGTGAGCCGGAGCTGAAGCTGGAAAG
P4	subcloning of appA into pET22b(+) / reverse ampl. of domain 5, appA	gcggccgcaagcCCTTACAACTGCACGCCGGTATGCG
P5	subcloning of ympH into pET22b(+)	catcatcatcacGCACCAGTTGCTGCACCAGTGACG
P6	subcloning of ympH into pET22b(+) / reverse ampl. of domain 5, ympH	gcggccgcaagcTTTTAAATATGGCATGC
P7	subcloning of phyX into pET22b(+)	catcatcatcacAGTGATACCGCCCTGCTGGGTTC
P8	subcloning of phyX into pET22b(+) / reverse ampl. of domain 5, phyX	gcggccgcaagcTTTTAAGGAAGC
P9	reverse ampl. of domain 1, appA	acgtacaccatgACGACTGACAATCACCAC
P10	reverse ampl. of domain 1, ympH	acgtacaccatgGCGGCTCAAGATAACCACTC
P11	reverse ampl. of domain 1, phyX	acgtacaccatgTCTGCTTAGGATAACAACCTTTTCCAACCTG
P12	forward ampl. of domain 2, appA	catggtg g acgtGCTCCAACCAAGTTTACGC

Primer Name	Purpose	Sequence (5'-3')
P13	forward ampl. of domain 2, ympH	catggtgtacgtTCGCCGACGAAACAAACAGAG
P14	forward ampl. of domain 2, phyX	catggtgtacgtGCGCCAACCAAAATGACAC
P15	reverse ampl. of domain 2, appA	aaacagcgggatcGGGACTGGACGTATCTGCCTG
P16	reverse ampl. of domain 2, ympH	aaacagcgggatcAACCTTTTTCAAATCAGCCTG
P17	reverse ampl. of domain 2, phyX	aaacagcgggatcGGCCTGCTGAATGTTTTGCTG
P18	forward ampl. of domain 3, appA	gatccgctgttAATCCTCTAAAACTGGCG
P19	forward ampl. of domain 3, ympH	gatccgctgttCACCCCGTCGAAGCCGGTG
P20	forward ampl. of domain 3, phyX	gatccgctgttCATCCTGTGAAAGCCG
P21	reverse ampl. of domain 3, appA	taataacggggtGGCGCGGCTGC
P22	reverse ampl. of domain 3, ympH	taataacggggtTCCCTTATGACGGGCG
P23	reverse ampl. of domain 3, phyX	taataacggggtACCGTTATGCTTGCGGATATAGGGCGTGC
P24	forward ampl. of domain 4, appA	accccgttattaGATTTGATCAAGACAGCG
P25	forward ampl. of domain 4, ympH	accccgttattaCAACAGATTGTGACGGCGCTAG
P26	forward ampl. of domain 4, phyX	accccgttattaCAAACCATCGCACACGCACTGGGTTC
P27	reverse ampl. of domain 4, appA	gttatccggctgACCGGGAAGCG
P28	reverse ampl. of domain 4, ympH	gttatcggctgTTGCGGTAGCTGC
P29	reverse ampl. of domain 4, phyX	gttatcggctgTCCCGGAAGTGCCATGTC
P30	forward ampl. of domain 5, appA	cagccagataacACGCCGCCAGGTGGTG
P31	forward ampl. of domain 5, ympH	cagccagataacACCCCGCCG
P32	forward ampl. of domain 5, phyX	cagccagataacACGCCTCCG

3.8.3.4 Domain Amplifications by PCR, DNA Purification and Quantification

Based on the alignments and selection of crossover points (Figure 72), five protein domains were designed varying in size (domain A, ~50 amino acids; domain B, ~105 amino acids; domain C, ~155 amino acids; domain D, ~50 amino acids; and domain E, ~55 amino acids). Amplifications of selected domains was achieved using Taq DNA polymerase (section 2.9.2) and domain specific oligonucleotides (Table 25). Correct amplification of DNA fragments was verified by agarose gel electrophoresis (section 2.9.6). Residual template DNA was removed by DpnI digestion and fragments were purified using a commercial DNA purification kit. The obtained DNA concentrations were measured spectrophotometrically.

3.8.3.5 Library Assembly, Colony PCR and Sequencing of Chimeras

The assembly of full length plasmids was performed similar to the OmniChange protocol (section 3.7)[223] with a few alterations and optimization steps. A mixture of 4 μ L DNA (0.03 pmol vector and 0.6 pmol of each insert) and 1 μ L cleavage mixture[213] proved to be best for cleavage of the PTO nucleotides and assembly of the chimeric plasmids. After iodine cleavage, fragments

(3 vector backbones and 12 inserts) were mixed in one PCR tube and the DNA was heated to 70°C (3 min; PCR-Cycler) followed by cooling to 20°C (5 min; PCR thermocycler) and storage on ice until transformation. This step ensures that no preferential hybridization occurs during assembly of full length plasmids. The assembled library was transformed without further purification or ligation into 50 µL chemically competent *E. coli* BL21 Gold (DE3) cells with a transformation efficiency of 1.4×10^6 cfu µg⁻¹ pUC19 (section 2.8.2). Cells were recovered on agar plates containing 50 µg mL⁻¹ ampicillin (section 2.7.3). Cells were recovered overnight at 37°C and correct assembly of chimeric phytase genes was verified by colony PCR (section 2.9.4) and agarose gel electrophoresis (section 2.9.6). Forty-eight colonies were sent for sequencing as agar plate cultures and using oligonucleotides PTO F1 Fw and PTO R3 Rv for sequencing (Table 27 in appendix section). Sequencing data was analyzed by alignment of phytase chimeras with the WT mature phytase sequences.

3.8.4 Results and Discussion

Three bacterial phytases (*appA*, *ympH* and *phyX*) were selected to serve as model genes for the development Phosphorothioate-based DNA Recombination (PTRec) method.[124] After protein sequence alignment of all three genes, four crossover points were designed, since most recombination methods have been performed with one to four crossover points.[36, 37] The generated protocol for PTRec includes four steps: 1) DNA fragment amplification by PCR using primers with complementary phosphorothioate nucleotides at the 5'-end; 2) Chemical cleavage of the PCR products using an alkaline mixture of I₂ and EtOH[222] to produce complementary single-stranded overhangs (each 12 nt long)[213, 223]; 3) DNA fragment assembly to full-length chimeric genes, including an additional heating step after all fragments were combined in one tube (70°C for 5 min) to prevent preferential hybridization, and 4) Direct transformation into *E. coli* BL21 (DE3) lacI^{Q1} (section 3.7) without further DNA purification steps (Figure 71).[223] As initial starting point, the three phytases were cloned into the pET22b(+) vector by PLICing[213] to ensure a comparability between all three genes and to reduce influences of e.g. varying transformation efficiencies influencing later statistical evaluations. Cloning of all phytases was done with the PLICing protocol and after sequencing all phytases were in frame. The 4-step protocol was preceded as outlined in Figure 71 including the additional heating step and a final incubation on ice until transformation that was included to prevent a preferential hybridization of domains. Transformation of 4 µL hybridization mixture, yielded in 500 to 1000 clones on a single agar plate corresponding to a transformation efficiency of 5.1×10^3 cfu µg⁻¹ DNA. On the control plate (cleaved vector fragment with ddH₂O instead of inserts) around 100 colonies were obtained (10 to 20 % background), indicating a successful assembly of the plasmids, as reported for OmniChange and PLICing.[213, 223] Despite the comparably high background (EV) [213, 223] the quality of the library was checked first by colony PCR. Therefore 24 random clones from the PTRec library were picked and colony PCRs was performed. All 23 clones

revealed full length genes in the expected size range from 1400 to 1500 bp (pelB leader sequence + His₆-Tag + phytase) (Figure 75).

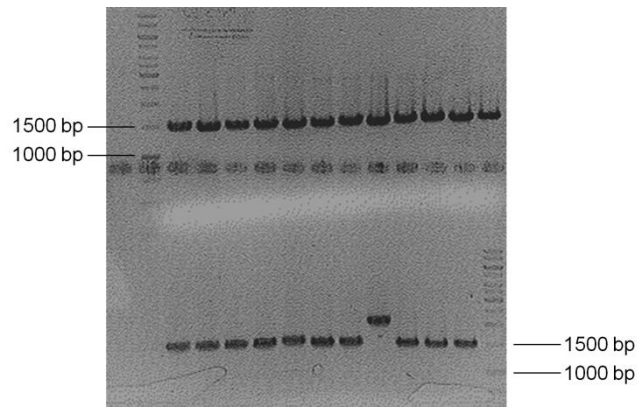


Figure 75. Colony PCR of 24 clones from a model PTRec library. The expected size of the chimeric phytase sequences is between 1400 and 1500 bp. A 1 kb DNA ladder (left upper/right lower) is used for estimation of obtained DNA fragment sizes. (This figure was kindly provided by Dr. Jan Marienhagen).

Only one clone displayed a significantly larger product of around 2.1 kb after colony PCR (Figure 75; lower lanes; lane 8), which is most likely to wrong hybridization of DNA fragments during step 3 or introduction of an additional fragment for instance an additional domain C which has a size of around 455 bp. Although it was already shown for PLICing (4 DNA nicks) and OmniChange (10 DNA nicks) the *E. coli* host is capable of ligating the respective nicks with an unexpectedly high efficiency. Consequently, it was unforeseen if the PTRec protocol, using a larger plasmid (two fold in size compared two OmniChange), still ensures a reliable repair of ten DNA nicks. The chemically generated nucleotide overhangs were stable and sufficient enough to withstand transformation conditions and nick repair in the *E. coli* host without causing frame shifts or mutations. Supported by the statistical evaluation (based on colony counting statistics; 10 to 20 % EV background) and the result from colony PCR (Figure 75) a sequence analysis of the generated library should reveal the quality of the library and the distribution of specific domains. Therefore 48 were picked and sequenced covering the whole gene sequence of all phytase chimeras. The number of clones is large enough to determine the efficiency of PTRec for DNA recombination and for benchmarking against state-of the art methods for DNA recombination.[124]

```

B12Shuffle-T7 atcaccagagtgagccggagctga: appA
C7Shuffle-T7  atcacgcaccagttgctgcaccag: ympH
C8Shuffle-T7  atcacgcaccagttgctgcaccag: ympH
C9Shuffle-T7  atcacgcaccagttgctgcaccag: ympH
C10Shuffle-T7 atcacagtgataccgcccctgctg: phyX
C11Shuffle-T7 atcaccagagtgagccggagctga: appA
C12Shuffle-T7 atcacgcaccagttgctgcaccag: ympH
D7Shuffle-T7  atcacagtgataccgcccctgctg: phyX
D8Shuffle-T7  atcacagtgataccgcccctgctg: phyX
D9Shuffle-T7  atcaccagagtgagccggagctga: appA
D10Shuffle-T7 atcacgcaccagttgctgcaccag: ympH
D11Shuffle-T7 atcaccagagtgagccggagctga: appA

```

Figure 76. Snapshot of the sequence alignment of 12 random clones from a PTRec library. Green highlighted nucleotides represent identical sequence parts (PTO area). White highlighted nucleotides represent domain specific sequences of phyX, ympH and appA.

Six out of 48 clones failed during sequencing due to bad sequence reads. However, 42 fully sequenced clones provided enough data to ensure a reliable statistical evaluation. Figure 76 shows a snapshot of the sequence alignment for domain 1 of the chimeric phytase genes. In total, all 42 genes contained the five expected domains in the correct order, demonstrating an efficient assembly through oriented hybridization. For statistical evaluation the respective domains were analyzed for their genetic origin and summarized in Figure 77 with a color coding to highlight the generated diversity of chimeras.

Clone	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					

Clone	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					

phyX

ympH

appA

Figure 77. Results from analysis of 42 clones generated by the PTRec protocol. All clones were sequenced at full gene length to investigate diversity and completeness of the recombined phytase genes. All clones are arranged in the order of sequencing, and the origin of each domain is highlighted: black: phyX; gray: ympH; white: appA. The Figure was redrawn from Marienhagen et al.[124]

The generated library was not screened for activity, since the quality of the library should be judged by the obtained diversity. Moreover, the crossover points were not selected based on a functional criterion which is still the major challenge to design recombination libraries.[36, 124] The distribution of individual fragments in the 42 clones was analyzed for each domain separately as well comparing the whole library with regards to uniqueness of the clones (Figure 77 and 78). The analyzed clones contained all five domains in correct order, without additional or redundant insertion of domains. The statistical distribution showed that, none of the domains occurred preferentially for the domains 1, 2, 4 and 5 (Domain 1: appA, 18; ympH, 12; phyX, 12/ Domain 2: appA, 19; ympH, 13; phyX, 10/ Domain 4: appA, 14; ympH, 14; phyX, 14/ Domain 5: appA, 14; ympH, 15; phyX, 13). Only phyX domain 3 can be regarded as being underrepresented (Domain 3: appA, 21; ympH, 14; phyX, 7) in the library since this domain was identified only seven times (obtained probability: 15.3 %; ideal probability: 33.3 %) (Figure 78).

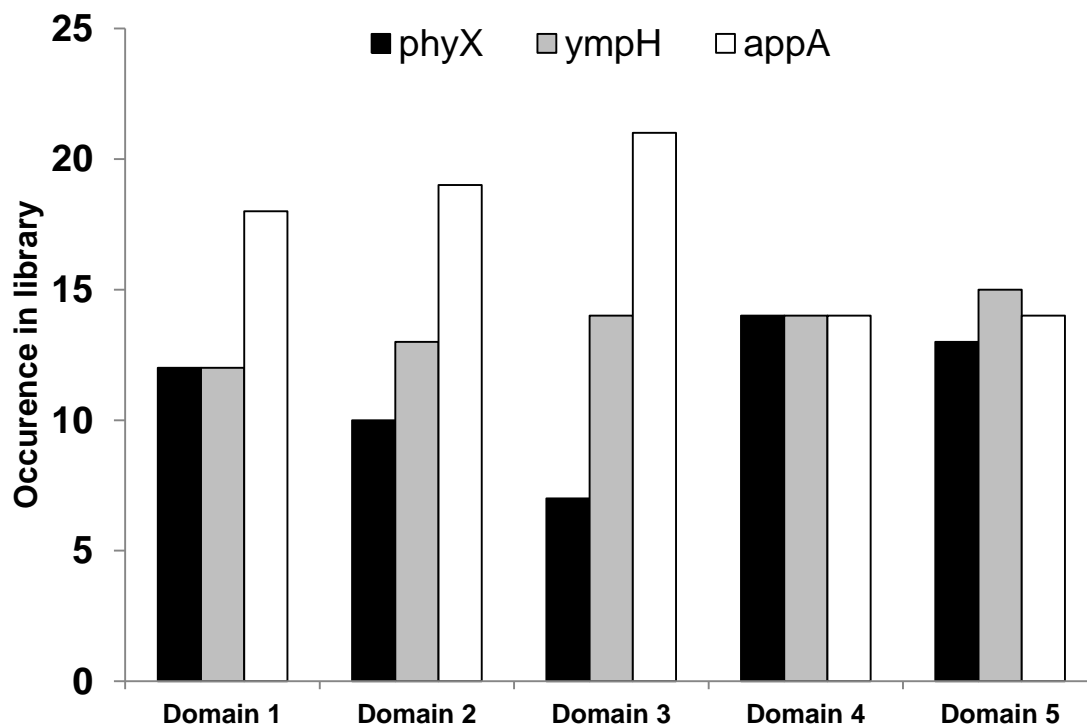


Figure 78. Statistical distribution of domains in a PTRec library of 42 sequenced clones. The occurrence of each phytase specific fragment is shown for the respective domain. Origin of the respective domains is indicated by colors: black: phyX; gray: ympH; white: appA. Domain 4 displays a statistical ideal distribution of domains (14:14:14), whereas in domain 3 phyX is under- and appA overrepresented.[124]

Although the library displayed a nearly ideal statistical distribution of phytase domains (14:14:14; domain 4; Figure 78) five sequenced clones were not recombined (3x *appA* and 2x *ympH*) whereas all the other 36 clones were unique in sequence (Figure 77). This can likely be attributed to incomplete *DpnI* digestion of template DNA after the PCR amplification and prior to chemical cleavage. In addition, two chimeric phytase genes occurred twice (Figure 77). No additional mutations were detected around the five crossover sites, which is well in agreement with the reported PLICing method and the results from sequencing a large amount of clones during development of OmniChange (section 3.7).[213, 223] Therefore this result underlines once more that the chemical cleavage with $I_2/EtOH$ under alkaline conditions (pH 9.0) specifically cleaves PTO nucleotides (section 2.9.5; Figure 17). In total three transition mutations were identified in all sequenced chimeric genes, corresponding to an error-rate of $2.2 \times 10^{-5} \text{ bp}^{-1}$ that can be attributed to the employed Taq DNA polymerase. The theoretical diversity of the chimeric phytase library includes 243 combinations (5 domains of 3 genes: 3^5). A library of this diversity would require an oversampling of 729 clones to explore 95 % of all possible combinations.[127] The required amount of colonies was obtained on a single agar plate after transformation of *E. coli* cells with comparably low competence ($1.4 \times 10^6 \text{ cfu } \mu\text{g}^{-1} \text{ pUC19}$), therefore generation of more colonies will be possible employing cells with higher competence. It is worth mentioning that an earlier library of 48 clones generated under identical conditions, and identical primers displayed a strong bias for one domain that was underrepresented (2 domains; probability: 4.2 %). Analyzing of sequencing data revealed a single point mutation within

the PTO area of a chemically synthesized primer that was identified in the hybridization area of underrepresented domain (2 clones). However after reordering the same primer the distribution was ideal as shown in Figure 78. A single point mutation in the PTO area therefore could influence the quality of a whole library and underlines the highly specific and selective hybridization employing phosphorothioate nucleotides in molecular cloning approaches.[124, 223]

DNA recombination is a powerful tool for *in vitro* evolution of enzymes, and broad range of methods are available for generating such genetically diverse libraries.[36, 37] Table 26 summarizes the most advanced methods for non-homologous gene recombination for direct comparison of PTRec at a glance for key performance parameters.

Table 26. Benchmarking methods for recombination with PTRec. The Table was taken from Marienhagen et al.[124]

Method	ITCHY[329]	SCRATCHY[330]	SHIPREC[331]	SISDC[332]	OE-PCR[121, 338]	USERec[335]	PTRec[124]
<i>Number of genes recombined as a proof of concept</i>	2	2	2	2	2	2	3
<i>Maximum number of crossover points as proof of concept</i>	1	3	1	4	3	9	4
<i>Maximum number of recombined genes reported</i>	2	2	2	3	Multiple	2	3
<i>Rational adjustment of crossover points is possible</i>	No	No	No	Yes	Yes	Yes	Yes
<i>Sequences included/removed in recombined genes</i>	No	No	No	Yes	No	No	No
<i>Endo-/Exonucleases employed during fragment generation</i>	Yes	Yes	Yes	Yes	No	Yes	No
<i>PCR amplification step before cloning is necessary</i>	No	Yes	Yes	Yes	Yes	Yes	No
<i>Ligase-activity is required</i>	Yes	Yes	Yes	Yes	Yes	Yes	No

The methods ITCHY, SCRATCHY and SHIPREC allow recombination of two genes with one to three crossover points generating only a few chimeras in a library. However, limitations due to low amount of crossover points forced to develop methods with rationally guided selection of crossover points. Application of SISDC, OE-PCR or USERec allows recombination of genes with up to 9 crossover points leading to a significant higher amount of protein chimeras (Table 26). Nonetheless, with increasing number of crossover points also more enzymatic steps are necessary accompanied with additional ligation and final PCR assembly steps making protocols tedious and less efficient.[213] Compared to the above mentioned non-homologous recombination methods, PTRec is unique, since any enzymatic steps during fragment generation is circumvented as well as final cloning or PCR steps are not required (Table 26). Key-to-success for recombination of the three model phytase genes by PTRec is the efficient and specific chemical cleavage of PTO bonds (section 2.9.5; Figure 17)[222] generating single-stranded DNA overhangs for efficient hybridization (section 2.9.5; Figure 16). As shown for OmniChange and PLICing[213, 223], any DNA purification step can be omitted prior to transformation, since iodine and EtOH do not interfere with hybridization of complementary single-stranded DNA. In addition, the time requirements for generation of a PTRec library are comparably low to all other non-homologous recombination methods (Table 26). One reason is the circumventing of enzymatic steps during fragment generation and eliminating ligation and PCR assembly steps meanwhile reducing technical challenges in library preparation to a minimum.[36] Commonly 500 to 1000 clones were obtained on a single plate when transforming the assembly mixture containing vector and all inserts. The vector re-hybridization control showed around 150 to 200 colonies as background. Compared to the background in OmniChange libraries, this is a very high value that might be attributed to different hybridization areas used for multi-fragment assembly. Despite the high empty vector background all sequenced clones contained full length plasmid, which is an excellent value. Furthermore, the high amount of clones from one transformation experiment is already sufficient to oversample the generated diversity in a later screening approach.[127] Another advantage of PTRec is the minimal sequence requirement at each crossover point: a stretch of only four amino acids that is identical among the protein sequence is sufficient to rationally define a crossover point (Figure 72). Even in gene sequences with very low similarity (<50 %) this can be compensated by incorporation of silent mutations in the primers employed in Step 1. The possibility to overcome the degeneracy of the genetic code decreases the sequence dependency to a further extent meaning that less nucleotide identity for each crossover point is necessary, which is a large advantage for instance over DNA Shuffling.[122] As a proof of concept, 19 silent mutations were integrated into the PCR primers on PTO 5'endings to ensure complementary single-stranded 5' overhangs for efficient hybridization (Table 25). The very low sequence identity requirements of PTRec should allow researchers to recombine more distantly related or even unrelated proteins to generate biocatalysts with improved or unknown functionalities. As shown already during the development of

OmniChange (section 3.7) a modular assembly of fragments will be possible to allow generation of chimeric sub-libraries.

3.8.5 Summary and Conclusion

Methods for molecular recombination are generally less frequently applied in protein engineering campaigns due to the technical challenges in preparation of the libraries and the high sequence restrictions to generate crossover points. In this chapter the development of the phosphorothioate-based DNA recombination method (PTRec) is described, that allows efficient recombination of gene sequences with similarities below 50 %. PTRec allows flexible generation of crossover points; each crossover point requires only a short stretch of four conserved amino acids to be generated. The additional heating step after mixing all fragments allows a nearly statistical ideal distribution of recombined domains in the later library. With this, a nearly sequence independent generation of protein chimeras is possible. PTRec is the first method that allows recombination on AA level, which also tolerates differences on the DNA level in particular within the crossover sections. PTRec is based, as PLICing and OmniChange, on the use of the phosphorothioate chemistry that enables reliable generation of complementary DNA overhangs. Applying this technology many technical limitations of other well established recombination methods have been solved, by avoiding multiple PCR steps, restriction enzymes or final cloning steps. The largest challenge in recombination of proteins remains unsolved: the rational selection of crossover points.

3.9 To Get What We Aim For – Progress in Diversity Generation Methods (Abstract from the original Review article)

Anna J. Ruff*, Alexander Dennig* and Ulrich Schwaneberg[36]

*These authors contributed equally to this work

Protein re-engineering by directed evolution has become a standard approach for tailoring enzymes in several fields of science and industry. Advances in screening formats and screening systems are fueling progress and enabling novel directed evolution strategies, despite the fact that the quality of mutant libraries can still be improved significantly. Diversity generation strategies in directed enzyme evolution comprise three options: (a) focused mutagenesis (selected residues are randomized); (b) random mutagenesis (mutations are randomly introduced over the whole gene); and (c) gene recombination (stretches of genes are mixed to chimeras in a random or rational manner). Either format has both advantages and limitations depending on the targeted enzyme and property. The quality of diverse mutant libraries plays a key role in finding improved mutants. In this review, we summarize methodological advancements and novel concepts (since 2009) in diversity generation for all three formats. Advancements are discussed with respect to the state-of-the-art in diversity generation and high-throughput screening capabilities, as well as robustness and simplicity in use. Furthermore, limitations and remaining challenges are emphasized ‘to get what we aim for’ through ‘optimal diversity’ generation.

4. Final Summary and Conclusions

4.1 Engineering of P450 BM3 for Application in Phenol Synthesis

Phenols represent key building blocks for plastics, resins, pharmaceuticals, pro-drugs and many more daily life products. The chemical synthesis of phenol is commonly performed in million ton scale via the Hock process that, despite of the low product yields (5 % with benzene as educt) and harsh reaction conditions, is only profitable due to the acetone byproduct. In order to circumvent harsh reaction conditions and energy intensive product downstreaming a more efficient and “greener” route to phenols is desired, preferentially involving molecular oxygen as sole oxidant. Enzymes, such as P450 monooxygenases, allow selective hydroxylation reactions under mild process conditions, using molecular oxygen as sole oxidant, including aromatic hydroxylation of benzene.

This work covers the development and characterization of unreported P450 BM3 variants that can be applied for biocatalytic hydroxylation of benzenes to produce aromatic building blocks such as phenols, hydroquinones and nitrobenzylalcohols. In total, 16 new biocatalytic routes (chemoselectivity >90 %) to *o*-substituted phenols (8), hydroquinones (5) and benzylalcohols (3) were described and investigated allowing a g L⁻¹ scale production for instance of guaiacol and methoxyhydroquinone without additional co-solvents. The new variants of the industrially important P450 monooxygenase BM3 were initially designed for the hydroxylation of *p*-xylene to produce 2,5-DMP. Therefore, five amino acid residues (R47, Y51, F87, A330, and I401) of P450 BM3 were rationally selected, iteratively mutated and the libraries screened for improved activity towards aromatic hydroxylation of *p*-xylene. Finally, a previously unreported P450 BM3 mutant named M2 (R47S, Y51W, I401M) was obtained. Variant M2 has an excellent catalytic performance towards aromatic hydroxylation of *p*-xylene ($k_{\text{cat}} = 1950 \text{ min}^{-1}$), high coupling efficiency (66 %) and selectivity (>98 %) to produce 2,5-DMP. Surprisingly, P450 BM3 WT was also able to hydroxylate *p*-xylene with nearly identical chemoselectivity, which is in contradiction to a previous report and has to be attributed to the significantly higher substrate concentrations employed in this thesis. Altogether, no residue in the immediate vicinity of the heme prosthetic group had to be mutated to obtain a variant with 37-fold improved catalytic efficiency compared to the WT enzyme, enabling for the first time the direct and selective aromatic hydroxylation of *p*-xylene to 2,5-DMP by P450 BM3.

The second target substrate, pseudocumene, was hydroxylated by WT and variant M2 at very low rate and with poor selectivity. All six theoretical products from single hydroxylations (three benzylalcohols and three phenols) were obtained with WT and variant M2. The WT produces preferentially benzylalcohols from pseudocumene (63.4 %), whereas the engineered variant M2 generates phenols in excess (61.7 %). From a SSM library, which additionally targeted position A330, a new P450 BM3 variant M3 (R47S, Y51W, A330F, I401M) was isolated. Variant M3 shows with pseudocumene as substrate a 70-fold higher productivity than the WT and 13-fold compared to the variant M2. For the isomeric mesitylene even more significant improvements were obtained, resulting in a 231-fold

higher PFR (M3 vs. WT) and indicating a better binding of both trimethylbenzene substrates in the active site. In order to explore the catalytic potential of the engineered P450 variants, a library of 28 substrates, predominantly monosubstituted benzenes, was screened whereby hydroxylation with high selectivity in *o*-position was mainly observed. For some substrates such as styrene (epoxidation), ethylbenzene and nitrotoluene (both α -hydroxylation) the aromatic hydroxylation was not predominant. A detailed characterization of the P450 BM3 WT and variant M2 for six selected monosubstituted benzenes revealed selectivities above 95 % towards *o*-hydroxylation of toluene, chloro-, bromo- and iodobenzene and high activity, especially for anisole (iTOF of 39 s⁻¹; 19.5 U mg⁻¹_{P450}). Compared to the WT protein, variant M2 displays improved activities ranging from 8 to 13-fold for the investigated monosubstituted benzenes. Accompanied with the improvements in catalytic activity, the coupling efficiency to non-natural substrates impressively reached in some cases nearly 50 %. On the other hand, the conversion of fluorobenzene is catalyzed with comparable low efficiency and selectivity, possibly due to the strong electron withdrawing effect of the fluor atom. Despite the low activity with fluorobenzene and iodobenzene, this is the first report of a P450 BM3 catalyzed hydroxylation of these substrates; in the case of iodobenzene no direct biocatalytic hydroxylation was previously reported. The hydroxylation of iodobenzene by variant M2 (TOF of 120 min⁻¹) also enables a new route to thyroid medications such as Liothyronine. Varying turnover frequencies for chloro- and bromobenzene can be attributed to the size and electronegativity of the substituents, restricting entrance to the active site and interactions close to the heme iron and residue F87.

After engineering and detailed catalytic characterization of P450 mutants the two best variants (M2 and M3) were benchmarked in long term biocatalytic conversions of monosubstituted benzenes and *p*-xylene relative to the WT. The following key performance parameters for implementation of catalysts into a synthetic process were determined: chemo- and enantioselectivity, TTN, product yields and final product concentrations. Therefore, a GDH-based cofactor regeneration protocol that is compatible with long-term hydroxylation reactions with P450 BM3 under unlimited cofactor supply was developed. It was further shown that co-solvents (e.g. DMSO or THF) are dispensable in the described hydroxylation of monosubstituted benzenes and *p*-xylene, despite their poor solubility in the aqueous solution. The janiform effect of DMSO on P450 BM3 was reported before; however, the investigation on conversion of toluene and variant M2 under varying co-solvent concentrations (0 to 30 % DMSO and IPA) revealed that P450 BM3 has highest productivity for *o*-cresol when no co-solvent is present in the KPi-buffered reaction system. Product pattern analysis revealed that in addition to the *o*-phenols unexpected side-products were formed in the GDH supported reactions. The latter were identified by GC-MS analysis and commercial standards as hydroquinones. The altered substrate preference can be attributed to varying K_M values for the respective benzenes and phenols, changes in substrate concentrations over time as well as varying pK_a values for the respective *o*-phenols influencing the partitioning of reaction products. All hydroquinones were formed with nearly

perfect selectivity (>99 %). The ratio of phenol and hydroquinone in the reaction can be controlled by adjusting substrate concentrations, e.g. a higher benzene concentration results in a higher phenol to hydroquinone ratio. Despite the significantly lower activity and TOF for all investigated substrates, the P450 BM3 WT alone is a very good catalyst for hydroxylation of benzenes generating a product titer of up to 0.79 g L⁻¹ from anisole. The new P450 BM3 variants provide higher coupling efficiency, resulting in higher product concentrations, increased product yields and TTNs. The highest product titer was obtained employing M3 and 20 mM of anisole, yielding 1.1 g L⁻¹ guaiacol and methylhydroquinone with a TTN of more than 10000 which is a rarely obtained value for a P450 monooxygenase and can be used as a starting point for further up-scaling. Several important parameters were investigated, bringing P450 BM3 and the variants closer to an application in organic synthesis. Yet, further work has to be done with regards to process engineering, co-factor supplementation/regeneration and catalyst concentrations.

Altogether, the new P450 BM3 variants enable not only a more efficient but also a 'greener' route to important phenolic building blocks by using an inexpensive catalyst, mild reaction conditions and molecular oxygen as sole oxidant – process parameters that are, till today, unreported for chemical catalysts. By providing a protocol for semi-preparative scale synthesis and bypassing the use of co-solvents or toxic and hazardous chemicals, the application of the described P450 variants can be a good alternative to traditional chemosynthetic processes. The P450-driven one-pot synthesis of hydroquinones from benzene educts was not reported so far and allows a new route to building blocks for vitamins, antioxidants, pharmaceuticals and polymerization inhibitors.

In addition to aromatic hydroxylation by P450 BM3, the benzylic hydroxylation of all three nitrotoluene isomers was investigated in detail. Of the three engineered variants (M1, M2 and M3) and the WT protein, M3 displayed the highest activity. The coupling efficiency and NADPH oxidation of variant M3 resulted in an 84-fold improved PFR compared to the WT enzyme for conversions of *p*-nitrotoluene. With 0.24 g L⁻¹ and a TTN beyond 1500 the engineered catalyst could be further improved, especially for the relatively low coupling efficiency (23.8 %; *p*-nitrotoluene). Nonetheless, the α -hydroxylation of nitrotoluenes allows a new biocatalytic route to nitrobenzylalcohols that after reduction of the NO₂-substituent are valuable building blocks for pharmaceuticals, pigments and dyes.

The importance of aromatic interactions in substrate recognition and regioselectivity was experimentally proven and largely verified by docking studies of the respective substrates allowing new insights into protein-substrate interactions. The obtained binding energies and orientation modes of benzene substrates are in a very good agreement with the catalytic data and show that aromatic interactions close to the heme iron and residue F87 play the dominant role in activity, selectivity and coupling.

A promising way to start re-engineering of P450 BM3 proved to be the simultaneous SSM of R47 and Y51, enabling entry of substrates into a wider active site pocket that provides enough space for binding and conversion of a broad range of substrates. Generally, the simultaneous removal of the polar R47 and Y51 ('gaters') enables a straight forward approach to increase activity of P450 BM3. The strategy to screen a library containing simultaneous saturated codons offers new possibilities and potentially leads to the best possible amino acid combination for a target substrate. This approach can easily be transferred to other substrate classes and very likely to other P450 enzymes that display similar structural folds. The catalytic data on these synthetically important compounds allows further studies of protein-substrate interactions especially with a focus on π - π -interactions.

In case of monosubstituted benzenes and *p*-xylene, it could be shown that P450 BM3 can perform direct hydroxylation with nearly complete *o*-selectivity due to orientation and binding of molecules in a precise orientation. Chemical synthesis very often does not allow selective reactivity and the intrinsic reactivity of the substrate plays the dominant role. An important observation in the mutagenesis study was also that residue F87 is essential for efficient aromatic hydroxylation of *p*-xylene, indicating the importance of strong interactions between the aromatic substrate and the active site of P450 BM3. The F87 residue is very often substituted by smaller residues such as alanine or valine, which can lead to remarkable improvements in catalytic activity of P450 BM3. The results shown in this thesis indicate that the exchange of phenylalanine is an inadvisable strategy if one aims at aromatic hydroxylation of benzene substrates. High coupling efficiencies and regioselectivity as well as a good catalytic activity for the relatively small benzene substrates, together with the indispensable presence of F87, suggests a strong interaction and selective positioning of benzene substrates within the comparably large active site of P450 BM3. In particular, anisole was converted very efficiently which prompted a docking study to investigate the hypothesis of a π - π interaction of benzene substrates and F87. Docking of anisole into the crystal structure model of the P450 BM3 WT revealed that the substrate most likely binds in T-shape orientation to residue F87 and leads to a strong atomic interaction. This orientation could explain the close orientation towards the heme iron ($< 4.0 \text{ \AA}$) and the adjacent water ligand, resulting in a chemo- and regioselective catalytic reaction. Introduction of phenylalanine (A330F) in the substrate binding pocket provides additional aromatic interaction, which was also suggested by analyzing the structural model of variant M3 and subsequent substrate docking. Although it remains unclear from the docking data results, the mutation A330F leads to a stronger binding of pseudocumene and mesitylene with both substrates stacked between residues F330 and F87. The hypothesis of a " π -X- π " stacking was generated which as future prospect requires more and reliable structural data. Chemo- and regioselective hydroxylation of non-symmetrical benzene substrates such as pseudocumene remains a major challenge in P450 BM3 catalyzed reactions as shown in previous reports with *o*- and *m*-xylene. Although the product pattern of the P450 BM3 WT could be rationally explained by docking of pseudocumene into the active site, the inverted chemoselectivity of M2 could not rationally be explained. To solve this question a crystal

structure with bound substrate would be required. Nonetheless, the new variant M3 has superior catalytic performance compared to variant M2 which was also proven during long term conversions with selected substrates. Unexpectedly, all three nitrotoluene isomers were preferentially hydroxylated at α -positions with regioselectivities beyond 90 %. This rather unusual selectivity for P450 BM3 can be attributed to the presence of the NO₂-substituent which has a strong directing influence during catalysis. A docking study revealed substrate binding modes in the WT enzyme that are in good agreement with the catalytic data. A close interaction of F87 and the substrate was also observed, however, formation of phenolic products was only in minor excess in nitrotoluene conversions. The reason why P450 BM3 provides excellent selectivity towards the α -positions of nitrotoluenes remains unsolved. In addition to the nitrobenzylalcohols, the respective nitrobenzaldehydes could be detected in a significant amount with up to 33 % for variant M3 and *o*-nitrotoluene as substrate. Oxidation of the alcohol was attributed to the abstraction of a proton from the α -carbon atom of nitrobenzylalcohol employing purified P450 BM3 protein. A mechanism for the reaction was proposed that requires binding of the first reaction product in a similar orientation as the initial substrate for efficient proton abstraction. In this way, a new synthetic route to nitrobenzylalcohols was provided and the possibility to use NO₂-groups as substituents to direct regioselectivities in P450 catalysis was postulated. As conclusion from the structure-functional investigations, only a crystal structure of P450 BM3 WT or variant M2 with a bound benzene substrate (which is so far not available) could serve as a final proof.

Transfer of this information from P450 BM3 to different enzymes could be helpful to engineer biocatalysts towards conversion of arenes or other aromatic substrates. The new P450 variants and their catalytic reactivity contribute to the growing knowledge on P450 BM3 in particular and monooxygenases in general. With respect to the demands in a steadily growing bioeconomy, the engineered P450 variants fulfill the current requirements for sustainable production of key building blocks in the chemical and pharmaceutical industry. The possibility to perform catalysis under mild reaction conditions, using molecular oxygen as sole oxidant, in conjunction with high selectivity and good turnover frequencies emphasizes the attractiveness of biocatalytic processes in the field of white biotechnology and contributes to the concept of green chemistry.

4.2 Development of New Methods for Directed Protein Evolution: OmniChange and PTRec

The successful outcome in directed protein evolution depends on two key factors: A reliable screening system and high quality mutant libraries. In this thesis two new methods for directed protein evolution were established. The OmniChange protocol solved all technical challenges in focused mutagenesis for saturating multiple independent codons and nucleotides in gene sequences while allowing assembly of multiple DNA fragments of various sizes in a simple four step protocol. Based on OmniChange, a new method for DNA recombination (PTRec) was developed that enables the generation of protein chimeras from gene sequences sharing less than 50 % homology.

A major limitation in focused mutagenesis is the saturation of multiple codons independent from their position within a target DNA sequence. In this thesis, the simultaneous saturation of two cooperatively interacting key residues in P450 BM3 (R47 and Y51) led to a new and unreported variant perfectly tailored for *p*-xylene hydroxylation. However, these variants would most likely remain undetected by saturating the positions iteratively or individually. In this regard, including additional positions into a simultaneous saturation experiment would be very attractive, allowing a holistic approach for optimization of proteins covering substrate entrance regions as well as deeply buried catalytic center residues. Mutagenesis methods that could generate such a focused mutant library are currently available but these are technically challenging and are often restricted to highly homologous regions. To solve this challenge a new method for homology independent multiple SSM, OmniChange, was developed as part of this protein engineering endeavor. Addressing the current challenges and demands for multiple SSM, OmniChange represents a robust method for saturation of independent codons throughout a gene sequence employing a simple four step protocol. The key to success is the employment of phosphorothiolated nucleotides at the 5'-ending of primers. After incorporation into a double stranded DNA they can be cleaved to generate up to 12 base pairs-long overhangs that selectively hybridize with high accuracy.

To benchmark OmniChange against state-of-the-art methods in multi-site saturation mutagenesis, 48 random clones from a model library were sequenced. All clones were unique, with all five codons being fully saturated. Up to 27 different triplets out of 32 possible (NNK degeneracy) were identified for a single codon corresponding to a diversity of 84.4 %, which is in conclusion an excellent value. Unexpectedly, the transformation of multi nicked plasmid DNA (10 nicks) into *E. coli* performs with high robustness omitting application of ligases or additional PCR amplification steps. The twelve complementary nucleotide overhangs are sufficient to hold the five DNA fragments together in a default order and withstand the heat shock transformation protocol, resulting in full length plasmids with the desired size. In addition, it was possible to employ the same set of oligonucleotides for saturation of two, three and four positions making OmniChange a versatile and modular method for focused mutagenesis. In future it will be possible to mutate more than five codons or introduce multiple saturated codons in one primer. Full saturation of more than five codons exceeds the throughput of state-of-the-art screening technologies[36, 223], thus degeneracies other than NNK can be employed to allow oversampling, if required. OmniChange is, in essence, a conceptually novel and powerful method for multi site-saturation mutagenesis which allows studying the influence of synergistic amino acid substitutions. Using OmniChange it will be possible to fully saturate substrate binding pockets of proteins to design better catalysts, especially for localizable properties such as activity and selectivity. The generation of libraries with high diversity will allow further exploration of the natural protein sequence space, particularly for properties that are difficult to rationalize, such as thermal stability or resistance against organic solvents that very often require cooperative amino acid substitutions to form salt or disulfide bridges.

Based on the successful development of OmniChange and the ability to fuse several DNA fragments in a rather simple methodological way a further method for efficient and nearly sequence independent recombination of genes was developed and named PTRec. This conceptually new method for DNA recombination offers the possibility of rational crossover adjustment and to recombine genes with low sequence identity (>50 %). Only a short stretch of four consecutive and conserved amino acids (12 nt) is necessary to design a crossover point. Generation of the chimeric libraries was similar to OmniChange, but an additional heating step was implemented as well as modifications for the employed DNA concentrations. This led to a nearly ideal statistical distribution of the respective domains of three model phytase genes (45 to 53 % identity). From 42 clones, 88 % of the sequenced genes were unique. PTRec provides an efficient way to overcome the major drawbacks of known methods for DNA recombination, providing a simple protocol, flexibility in terms of crossover points and the potential to recombine sequences with low sequence identity. It is worth mentioning that PTRec is to date the only method allowing recombination on the amino acid level and enables rational crossover adjustment by introducing silent mutations. With this large degree of freedom it will be possible to generate chimeras of unrelated protein families. Developments in both aspects, focused mutagenesis and DNA recombination, drive the progress in biocatalyst engineering for future applications in chemistry, medicine, material sciences and biology. The construction of genetic circuits and metabolic pathways very often requires assembly of multiple genetic elements. In this regard, the application of phosphorothioate-based cloning technologies such as PLICing, OmniChange and PTRec will fuel the process in the blooming field of synthetic biology.

5. Appendix

5.1 List of Tables

Table 1. Parameters that are important for an application of biocatalysts in organic synthesis.....	7
Table 2. Selected organisms and enzymes reported for the aromatic hydroxylation of benzenes and biocatalytic production of phenol.....	23
Table 3. Employed genes in this work and their origin.....	31
Table 4. Thermocycler program for PCR amplification with Taq DNA polymerase.....	36
Table 5. Preparation of MTP plates for screening with the NADPH depletion assay.....	41
Table 6. Catalytic parameters for hydroxylation of <i>p</i> -xylene by P450 BM3 (WT, M1, and M2).....	56
Table 7. Substrate library used for conversions with P450 BM3 variant M2.....	64
Table 8. Results from conversion screening of 28 selected benzenes with variant M2.....	66
Table 9. Products identified by GC-MS analysis.....	67
Table 10. Catalytic performance of P450 BM3 WT and variant M2 for the aromatic hydroxylation of six monosubstituted benzenes.....	79
Table 11. Catalytic performance of WT, M2 and M3 for the hydroxylation of pseudocumene.....	88
Table 12. Regioselectivity of WT, M2 and M3 for the hydroxylation of pseudocumene.....	89
Table 13. Kinetic characterization of P450 BM3 and variants M2 and M3 for conversion of mesitylene. ...	92
Table 14. Catalytic performance of P450 BM3 WT, variant M2 and M3 at three different concentrations of toluene.....	103
Table 15. Catalytic performance of P450 BM3 WT, variant M2 and M3 at three different concentrations of anisole.....	104
Table 16. Catalytic performance of P450 BM3 WT, variant M2 and M3 at three different concentrations of <i>p</i> -xylene.....	106
Table 17. Catalytic performance of P450 BM3 WT, variant M2 and M3 at three different concentrations of chlorobenzene.....	107
Table 18. Catalytic performance of P450 BM3 WT, variant M2 and M3 at three different concentrations of bromobenzene.....	108
Table 19. Catalytic characterization of P450 BM3 WT and variants towards hydroxylation of <i>o</i> -NT .	119
Table 20. Catalytic characterization of P450 BM3 WT and variants towards hydroxylation of <i>m</i> -NT	121
Table 21. Catalytic characterization of P450 BM3 WT and variants towards hydroxylation of <i>p</i> -NT .	122
Table 22. List of used primers for development of OmniChange.....	131
Table 23. Detailed statistical overview on codon diversity generated by OmniChange.....	137
Table 24. Overview of methods and strategies for focused mutagenesis on multiple positions.....	142
Table 25. Full list of oligonucleotides used during the development of PTRec and their specific purpose.....	149
Table 26. Benchmarking methods for recombination with PTRec.....	157
Table 27. List of oligonucleotides used for cloning, sequencing and mutagenesis of P450 BM3.....	174
Table 28. Strains and plasmids generated in this thesis.....	178
Table 29. List of all analyzed compounds including retention times, respective GC-programs, extraction solvents and applied internal standard for quantification.....	180

Table 30. Primer combinations employed for saturation of codons D52 and T77 in the YmPhytase sequence during development of the OmniChange method.....	186
---	-----

5.2 List of Figures

Figure 1. Business operation areas for biotechnology companies in Germany	3
Figure 2. Enzymatic synthesis of chiral amines by kinetic resolution and asymmetric synthesis	4
Figure 3. Strategy guide for engineering of enzymes based on structural information and throughput of the screening system.....	10
Figure 4. Three steps in a directed evolution experiment	11
Figure 5. Evolutionary tree for improving thermostability of a lipase (Lip A) by iterative saturation mutagenesis (ISM)	13
Figure 6. Two bacterial electron transfer systems of P450 monooxygenases.....	15
Figure 7. The catalytic cycle of cytochrome P450 monooxygenases	16
Figure 8. Visualization of the dimeric structure of the two P450 BM3 monooxygenase subunits	19
Figure 9. Selected key residues in the substrate binding domain of P450 BM3.....	20
Figure 10. P450 BM3 accepts various substrates for hydroxylation such as propane, cyclohexane, naphthalene and testosterone	21
Figure 11. The mechanism for P450-catalyzed epoxidation of benzene for the biocatalytic production of phenol.....	24
Figure 12. Vector maps of employed pET-22b(+) and pET-28a(+) high-copy expression plasmids ...	29
Figure 13. Vector maps of pET-derived pALXtreme-5b (pET-22b(+)) and pALXtreme-1a (pET-28a(+))	30
Figure 14. Scheme for MTP arrangement for reliable screening of P450 BM3 variants.....	34
Figure 15. Preparation and expression process for mutant libraries in MTP format.....	35
Figure 16. A schematic representation of the PLICing method.....	37
Figure 17. Chemical cleavage of phosphothiolated nucleotides in the presence of iodine and ethanol under alkaline conditions	38
Figure 18. Simultaneous site saturation mutagenesis of rationally selected amino acid residues R47 and Y51 of P450 BM3	39
Figure 19. A: Distribution of clones into a MTP for estimation of standard deviation of hydroxylation reaction of P450 BM3	40
Figure 20. The 4-AAP assay for phenolic product detection in MTP format	42
Figure 21. Combined screening with the NADPH depletion assay and the 4-AAP detection system for phenols	43
Figure 22. Entire process of screening, selection of variants for rescreening and final characterization of confirmed improved variants.	44
Figure 23. Absorbance spectra recorded before and after gassing P450 BM3 containing samples with carbon monoxide (CO)	46
Figure 24. Structure of Gemfibrozil and β -tocotrienol	51
Figure 25. P450 BM3 active site snapshot	54

Figure 26. Validation of the MTP screening system using NADPH depletion and 4-AAP assay in 96-well format with P450 BM3 and <i>p</i> -xylene as substrate	55
Figure 27. Exemplary screening of plate 19 from first round of saturation mutagenesis	56
Figure 28. A: Gas chromatograms of conversion products of <i>p</i> -xylene by P450 BM3 WT, variant M1 and M2.....	57
Figure 29. Kinetic characterization of WT P450 BM3 with <i>p</i> -xylene as substrate.....	58
Figure 30. ISM strategy for generation of variants for <i>p</i> -xylene hydroxylation.....	59
Figure 31. Screening of SSM library at position F87 with the 4-AAP assay and <i>p</i> -xylene as substrate	60
Figure 32. Gas chromatograms for the hydroxylation of toluene, <i>m</i> -xylene and <i>o</i> -xylene by P450 BM3	61
Figure 33. Five selected substrates used for hydroxylation by variant M2	64
Figure 34. Reaction vessels standing on a magnetic stirring board and filled with the reaction mixture for hydroxylation of selected benzene substrates	65
Figure 35. Screening of new substrates for conversion by P450 BM3 variant M2.....	66
Figure 36. A: Formation of <i>o</i> -cresol by variant M2 over time at varying DMSO concentrations.....	68
Figure 37. A: Time dependent formation of <i>o</i> -cresol by variant M2 at increasing IPA concentrations	69
Figure 38. Selected products obtained after hydroxylation with P450 BM3 variant M2.	71
Figure 39. Product formation of <i>o</i> -hydroxylated phenols from anisole, toluene, chloro-, bromo- and iodobenzene by P450 BM3 WT.....	78
Figure 40. Visualization of P450 BM3's (PDB 1BU7) active site cavity with the anisole substrate docked in two T-shape binding orientations	80
Figure 41. Visualization of P450 BM3's active site with the docked anisole substrate in T-shape orientation and highest interaction energies.....	81
Figure 42. A: The three isomers of trimethylbenzene (TMB).....	83
Figure 43. Six isomeric products from hydroxylation of pseudocumene with P450 BM3	84
Figure 44. Screening of SSM library A330 with the NADPH depletion assay employing pseudocumene as substrate	86
Figure 45. Screening of SSM library A330 with the 4-AAP assay and pseudocumene as substrate. .	87
Figure 46. Rescreening of three selected clones from SSM library A330.....	88
Figure 47. Chemoselectivity of P450 BM3 WT, variant M2 and M3 for hydroxylation of pseudocumene	90
Figure 48. Determined alterations in product distributions of variants M2 and M3 compared the P450 BM3 WT.....	91
Figure 49. Absolute product concentration after conversion of mesitylen and pseudocumene.....	92
Figure 50. Docking of mesitylene into the active site of P450 BM3 WT.....	93
Figure 51. Docking of mesitylene into the active site of the homology model of variant M3.....	94
Figure 52. Docking of pseudocumene into active site of P450 BM3 WT	95
Figure 53. Four different binding modes of pseudocumene docked into the active site of variant M3.	96
Figure 54. Conversion of 20 mM anisole by variant M2	109
Figure 55. Conversion of 20 mM <i>p</i> -xylene by variant M2.....	110

Figure 56. Conversion of 20 mM toluene by variant M2.....	111
Figure 57. Conversion of 20 mM chlorobenzene by variant M2.....	112
Figure 58. The three isomers of nitrotoluene.....	117
Figure 59. Chromatograms for separation and detection of products from the conversion of <i>o</i> -NT with selected P450 variants	120
Figure 60. Docking of NT isomers into the active site of P450 BM3 WT	123
Figure 61. Docking of <i>o</i> -NT into the active site of P450 BM3 variant M3.....	124
Figure 62. Docking of <i>p</i> -NT isomers into the active site of P450 BM3 variant M3	124
Figure 63. The OmniChange concept	130
Figure 64. Three steps in oligonucleotide primer design for OmniChange	132
Figure 65. Exemplary separation of DNA on a 2 % agarose gel from step 1 in the OmniChange protocol.....	133
Figure 66. Dependency of insert concentration and obtained colonies after transformation of an OmniChange library.....	134
Figure 67. Colony PCR of 20 random clones from a phytase OmniChange library	135
Figure 68. Sequence alignment of five clones from an initially prepared OmniChange library.....	135
Figure 69. Obtained diversity shown as partial sequence alignment of 48 randomly selected clones from a generated OmniChange library	136
Figure 70. Obtained diversity from NNK saturation of five target codons with the OmniChange protocol.....	138
Figure 71. Concept of a 4-step protocol for recombination of three model genes during the development of PTRec.....	146
Figure 72. Amino acid sequence alignment of three model phytases.....	147
Figure 73. Prediction of cleavage site of appA phytase signal sequence in <i>E. coli</i>	148
Figure 74. Snapshot of amino acid and DNA sequence alignment for the three model phytases appA, phyX and ympH.....	149
Figure 75. Colony PCR of 24 clones from a model PTRec library	152
Figure 76. Snapshot of the sequence alignment of 12 random clones from a PTRec library	153
Figure 77. Results from analysis of 42 clones generated by the PTRec protocol.....	154
Figure 78. Statistical distribution of domains in a PTRec library of 42 sequenced clones	155
Figure 79. Separation of DNA on a 1% agarose gel	179
Figure 80. Purification of P450 BM3.....	179
Figure 81. Calibration curve for 4-AAP assay and the commercial standard of 2,5 DMP.	182
Figure 82. UV-detection profile from a conversion of chlorobenzene by P450 BM3 variant M2.....	182
Figure 83. GC-MS fragmentation profile of 2,5-dimethylhydroquinone	183
Figure 84. GC-MS fragmentation profile of 2,5-dimethylquinone	183
Figure 85. GC-MS fragmentation of the product obtained from conversion of <i>p</i> -toluidin with variant M2	184
Figure 86. GC-MS fragmentation of phenolic products obtained from conversion of <i>o</i> -nitrotoluene with variant M3.....	184

Figure 87. Analysis of PCR products by agarose gel electrophoresis during development of the OmniChange method	185
Figure 88. Sequence alignment of OmniChange sub-libraries.....	185
Figure 89. Full DNA sequence alignment of three model phytases employed during the development of PTRec.....	186
Figure 90. Prediction of cleavage site of ympH phytase signal sequence in <i>E. coli</i>	187
Figure 91. Prediction of cleavage site of phyX phytase signal sequence in <i>E. coli</i>	187

5.3 List of Schemes

Scheme 1. Biocatalytic transamination of Prositagliptin by a (<i>R</i>)-selective ω -transaminase	4
Scheme 2. Exemplary enzymatic cascade for the bioamidation of alcohols	6
Scheme 3. Basic scheme for a P450 monooxygenase-catalyzed reaction.....	15
Scheme 4. NADPH cofactor regeneration with a glucose dehydrogenase	18
Scheme 5. The cumene process for the industrial production of phenol from benzene	22
Scheme 6. Conversion of <i>p</i> -xylene by P450 BM3 to produce 2,5-DMP.....	52
Scheme 7. Aromatic substrates employed for hydroxylation with P450 BM3 variant M2	75
Scheme 8. Aromatic hydroxylation of selected monosubstituted benzenes by P450 BM3.....	76
Scheme 9. Aromatic hydroxylation of selected benzenes using a P450 BM3 monooxygenase and a glucose dehydrogenaseregeneration system for production of <i>o</i> -phenols	102
Scheme 10. Revised reaction scheme employing P450 BM3, benzene as substrate and a GDH cofactor regeneration system	115
Scheme 11. Hydroxylation of <i>p</i> -NT by P450 BM3 in α -position to produce <i>p</i> -NBOH.....	118
Scheme 12. Proposed mechanism for oxidation of <i>o</i> -NBOH by variant M3.....	127

5.4 Abbreviations

°C	Degree Celsius
4-AAP	4-aminoantipyrine
ACN	Acetonitrile
ADH	Alcohol dehydrogenase
ALA	Aminolevulinic acid
ampl.	Amplification
APS	Ammonium persulfate
atm	Atmospheric pressure
bar	Pressure
bp	Base pair(s)
BM	<i>Bacillus megaterium</i>
BSA	Bovine Serum Albumin
CBB	Coomassie-brilliant blue
CDDC	Cyclododecanol
cfu	Colony forming units
CYP	Cytochrome P450 monooxygenase
CV	Column volume
ddH ₂ O	Double-distilled water

DEAE	Diethylethanolamine
DMBA	Dimethylbenzylalcohol
DMHQ	Dimethylhydroquinone
DMSO	Dimethyl sulfoxide
DMP	Dimethylphenol
DNA	Deoxyribonucleic acid
dsDNA	Double stranded Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
<i>E. coli</i>	<i>Escherichia coli</i>
e.g.	<i>exempli gratia</i> (for example)
epPCR	Error-prone polymerase chain reaction
et al.	<i>et alli</i>
EtOAc	Ethyl acetate
EtOH	Ethanol
EV	Empty vector
FAD	Flavin adenine dinucleotide
FDH	Formate dehydrogenase
FID	Flame ionization detector
FMN	Flavin mononucleotide
Fw	Forward
g	Gravitational force
ΔG_{bind}	Gibbs free energy of binding
GC	Gas chromatography
GDH	Glucose dehydrogenase
GOI	Gene of interest
h	Hour(s)
HTS	High Throughput Screening
HPLC	High performance liquid chromatography
HQu	Hydroquinone
IPA	Isopropanol
IPTG	Isopropyl β -D-thiogalactopyranoside
ISM	Iterative saturation mutagenesis
iTOF	Initial turnover frequency
K	Kelvin
Kan	Kanamycin
kb	Kilobase(s)
k_{cat}	Catalytic activity
K_{M}	Michaelis Menten constant/substrate affinity constant
KPi	Potassium phosphate buffer with inorganic phosphate
LB	Lysogeny broth
MBA	Methylbenzylalcohol
MCS	Multiple cloning site
min	Minute(s)
mol	Molecule
MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	Messenger ribonucleic acid
mM	Milli molar
MS	Mass spectrometry
MTBE	Methyl <i>tert</i> -butyl ether
MTP	Microtiter plate
<i>m/z</i>	Mass-to-charge ratio
NADP ⁺	Nicotine amide adenine dinucleotide phosphate (oxidized form)
NADPH	Nicotine amide adenine dinucleotide phosphate (reduced form)
NBA	Nitrobenzaldehyde
NBOH	Nitrobenzylalcohol
NT	Nitrotoluene

nt	Nucleotide
OD ₆₀₀	Optical density at a wavelength of 600 nm
P450	Cytochrome P450 monooxygenase
PCR	Polymerase chain reaction
PFR	Product formation rate
pH	Decimal logarithm of the reciprocal of the hydrogen ion activity
Ph	Phenol
pK _a	Logarithmic measure of the acid dissociation constant
PLICing	Phosphorothioate-based ligase-independent gene cloning
PTRec	Phosphorothioate-based DNA recombination
Ref.	Reference
Rv	Reverse
rpm	Revolutions per minute
RT	Room temperature
SDM	Site Directed Mutagenesis
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
s	Second(s)
ssDNA	Single stranded Deoxyribonucleic acid
SOC	Super optimal broth with catabolite repression
SSM	Site Saturation Mutagenesis
TB	Terrific broth
TEMED	N,N,N',N'-tetramethylethylenediamine
TFB	Transformation buffer
THF	Tetrahydrofuran
TMP	Trimethylphenol
TOF	Turnover frequency
Tris	Tris(hydroxymethyl)aminomethane
TTN	Total turnover number
U	Units (enzyme activity)
UV	Ultraviolet
vs.	versus
V/V	Volume per volume
WT	Wild-type
W/V	Weight per volume
W/W	Weight per weight

5.5 Additional Experimental Information and Data

Table 27. List of oligonucleotides used for cloning, sequencing and mutagenesis of P450 BM3. Information in this Table was published in parts by Dennig et al.[69]

Use	Target	Primer Name	Primer Sequence (5'-3')
SSM	R47/Y51	R47/Y51 Fw	CGAGGCGCCTGGT <u>NNK</u> GTAACGCGC <u>NNK</u> TTATCAAGTCAGCG
		R47/Y51 Rv	CGCTGACTTGATA <u>MNN</u> GC GCGTTAC <u>MNN</u> ACCAGGCGCCTCG
SSM	F87	F87 Fw	CAGGAGACGGGTT <u>ANNK</u> ACAAGCTGGACGCATG
		F87 Rv	CATGCGTCCAGCTTGT <u>MNN</u> TAACCCGTCTCCTG
SSM	A330	A330 Fw	CAACTGCTCCT <u>NNK</u> TTTTCCCTATATGC
		A330 Rv	CAACTGCTCCT <u>NNK</u> TTTTCCCTATATGC
SSM	I401	I401 Fw	CAGCGTGCGTGT <u>NNK</u> GGTCAGCAGTTC
		I401 Rv	GAACTGCTGAC <u>MNN</u> ACACGCACGCTG

Cloning of P450 BM3	Vector	PTO F1 Rv	ctatagtgagtcGTATTAATTTCGAACATGTGAGC
	Vector	PTO R3 Fw	cgggcttgtagCAGCCGGATCTCAG
	Gene	PTO F1 Fw	gactcactatagGGGAATTGTGAGCGG
	Gene	PTO R3 Rv	ctaacaaagcccgAAAGGAAGCTGAGTTG
Sequencing P450 BM3	---	T7 Promoter	TAATACGACTCACTATAGGG
	---	1	CTTTAACAGCTTTTACCGAGATCAGC
	---	2	CAGCTTAAATATGTCGGCATGGTC
	---	3	GTGCTATACGGTTCAAATATGG
	---	4	GTTTTCAACGAACGTCGTAGC
	---	5	GTCGATGAAAAACAAGCAAGC

Capital letters = Standard nucleotides connected through phosphodiester bonds;

Lower case letters = Nucleotides connected through a phosphorothioatediester bond

Full DNA and protein sequence of P450 BM3 wild-type

atgacaattaaagaaatgcctcagccaaaaacgtttgagagccttaaaaatttacggtta
M T I K E M P Q P K T F G E L K N L P L
 ttaaacacagataaacgggttcaagctttgatgaaaattgcggatgaattaggagaaatc
L N T D K P V Q A L M K I A D E L G E I
 tttaaattcgaggcgcctggctgtaaacgcgctacttatcaagtcagcgtctaattaaa
F K F E A P G R V T R Y L S S Q R L I K
 gaagcatgcgatgaatcacgctttgataaaaaacttaagtcaagcgccttaaatttgtagct
E A C D E S R F D K N L S Q A L K F V R
 gattttgcaggagacgggttatTTACAAGCTGGACGCATGAAAAAATTGGAAAAAGCG
D F A G D G L F T S W T H E K N W K K A
 cataatcttacttccaagcttcagtcagcaggcaatgaaaggctatcatgcatgatg
H N I L L P S F S Q Q A M K G Y H A M M
 gtcgatatcgccgtgcagcttggttcaaaagtgggagcgtctaataatgcagatgagcatatt
V D I A V Q L V Q K W E R L N A D E H I
 gaagtaccggaagacatgacacgtttaacgcttgatacaattggcttttgcggtttaac
E V P E D M T R L T L D T I G L C G F N
 tatcgctttaacagcttttacggagatcagcctcatccatttattacaagtatggctcgt
Y R F N S F Y R D Q P H P F I T S M V R
 gcactggatgaagcaatgaacaagctgcagcagcagcaaatccagacgaccagcttatgat
A L D E A M N K L Q R A N P D D P A Y D
 gaaaacaagcgcagtttcaagaagatatcaaggtgatgaacgacctagtagataaaatt
E N K R Q F Q E D I K V M N D L V D K I
 attgcagatcgcaaagcaagcgggtgaacaaagcagatgatttattaacgcatatgctaaac
I A D R K A S G E Q S D D L L T H M L N
 ggaaaagatccagaaacgggtgagccgcttgatgacgagaacattcgctatcaaattatt
G K D P E T G E P L D D E N I R Y Q I I
 acattcttaattgcgggacacgaaacaacaagtggtcttttatcatttgcgctgatttc
T F L I A G H E T T S G L L S F A L Y F
 ttagtgaaaaatccacatgtattacaaaaaagcagcagaagaagcagcagcaggttctagta
L V K N P H V L Q K A A E E A A R V L V
 gatcctgtttccaagctacaacaagtcacaacagcttaaatatgctcggcatggctttaaac
D P V P S Y K Q V K Q L K Y V G M V L N
 gaagcgtgcgcttatggccaactgctcctgcttttccctatatgcaaaagaagatagc
E A L R L W P T A P A F S L Y A K E D T
 gtgcttggaggagaatatccttttagaaaaaggcagcgaactaatggttctgattcctcag
V L G G E Y P L E K G D E L M V L I P Q
 cttcaccgtgataaaacaatttggggagacgatgtggaagagttccgtccagagcgtttt
L H R D K T I W G D D V E E F R P E R F
 gaaaatccaagtgcgattccgcagcatgcgctttaaacggtttgaaacggtcagcgtgcg

E N P S A I P Q H A F K P F G N G Q R A
tgtatcgggtcagcagttcgcctcttcatgaagcaacgctgggtacttggtatgatgctaaaa
C I G Q Q F A L H E A T L V L G M M L K
cactttgactttgaagatcatacaaaactacgagctggatattaagaaactttaacgтта
H F D F E D H T N Y E L D I K E T L T L
aaacctgaaggctttgtggtaaaagcaaaatcgaaaaaaattccgcttggcggtattcct
K P E G F V V K A K S K K I P L G G I P
tcacctagcactgaacagctctgctaaaaagtagcaaaaaggcagaaaacgctcataat
S P S T E Q S A K K V R K K A E N A H N
acgcccgtgcttgtgctatacggttcaaatatgggaacagctgaaggaaacggcgctgat
T P L L V L Y G S N M G T A E G T A R D
ttagcagatattgcaatgagcaaaggatttgcaccgcaggtcgcaacgcttgattcacac
L A D I A M S K G F A P Q V A T L D S H
gccggaaatcttccgcgcgaaggagctgtattaattgtaacggcgctttataacggctcat
A G N L P R E G A V L I V T A S Y N G H
ccgcctgataacgcaaagcaatttgcgactggtttagaccaagcgtctgctgatgaagta
P P D N A K Q F V D W L D Q A S A D E V
aaaggcgcttccgctactccgtatttggatgcccgcgataaaaactgggctactacgtatcaa
K G V R Y S V F G C G D K N W A T T Y Q
aaagtgcctgcttttatcgatgaaacgcttgcgcgctaaaggggcagaaaacatcgctgac
K V P A F I D E T L A A K G A E N I A D
cgcggtgaagcagatgcaagcgacgactttgaaggcacatatgaagaatggcggtgaacat
R G E A D A S D D F E G T Y E E W R E H
atgtggagtgacgtagcagcctactttaacctcgacattgaaaacagtgaaagataataaa
M W S D V A A Y F N L D I E N S E D N K
tctacttttcaacttcaatttgcgacagcgcgcgggatatgccgcttgcgaaaatgcac
S T L S L Q F V D S A A D M P L A K M H
ggctgcgttttcaacgaacgctcgtagcaagcaaaagaacttcaacagccaggcagtgcaacga
G A F S T N V V A S K E L Q Q P G S A R
agcacgcgacatcttgaaattgaaacttccaaaagaagcttcttatcaagaaggagatcat
S T R H L E I E L P K E A S Y Q E G D H
ttaggtgttattcctcgcgaactatgaaggaatagtaaacgctgtaacagcaaggttcggc
L G V I P R N Y E G I V N R V T A R F G
ctagatgcatcacagcaaatccgctctggaagcagaagaagaaaaattagctcatttgcca
L D A S Q Q I R L E A E E E K L A H L P
ctcgcataaaacagtatccgtagaagagcttctgcaatacgtggagcttcaagatcctggt
L A K T V S V E E L L Q Y V E L Q D P V
acgcgcacgcagcttccgcgcaatggctgctaaaacggctctgcccgcgcataaaagtagag
T R T Q L R A M A A K T V C P P H K V E
cttgaagccttgcttgaagcaagcctacaaaagaacaagtgctggcaaacgtttaaca
L E A L L E K Q A Y K E Q V L A K R L T
atgcttgaactgcttgaaaaaataccggcgctgtgaaatgaaattcagcgaatttatcgcc
M L E L L E K Y P A C E M K F S E F I A
cttctgccaagcatacgcgcgctattactcgatttcttcatcacctcgtgctgatgaa
L L P S I R P R Y Y S I S S P R V D E
aaacaagcaagcatcacggctcagcgttgtctcaggagaagcgtggagcggatattggagaa
K Q A S I T V S V V S G E A W S G Y G E
tataaaggaattgctcgaactatcttgcgcgagctgcaagaaggagatacgttacgtgc
Y K G I A S N Y L A E L Q E G D T I T C
tttatttccacaccgcagtcagaattttagctgccaagaccctgaaacgcgccttatc
F I S T P Q S E F T L P K D P E T P L I
atggctggaccgggaacagggcgtcgcgcgctttagaggctttgtgcaggcgcgcaaacag
M V G P G T G V A P F R G F V Q A R K Q
ctaaaagaacaaggacagtcacttggagaagcacatttatacttccggtgcccgttcacct
L K E Q G Q S L G E A H L Y F G C R S P
catgaagactatctgtatcaagaagagcttgaaaaacgcccagcaaggcatcattacg
H E D Y L Y Q E E L E N A Q S E G I I T
cttcataaccgcttttctcgcgatgccaaatcagccgaaaacatacgttcagcacgtaatg
L H T A F S R M P N Q P K T Y V Q H V M
gaacaagacgggaagaaattgattgaacttcttgatcaaggagcgcacttctatatttgc
E Q D G K K L I E L L D Q G A H F Y I C
ggagacgggaagccaaatggcacctgcccgttgaagcaacgcttatgaaaagctatgctgac

G D G S Q M A P A V E A T L M K S Y A D
gttcaccaagtgagtgaaagcagacgctcgcttatggctgcagcagctagaagaaaaaggc
V H Q V S E A D A R L W L Q Q L E E K G
cgatacgcaaaagacgtgtgggctgggtaa
R Y A K D V W A G -

Grey highlighted codon (ATG, M) = Start codon for translation in *E. coli*; Red highlighted codon (TAA)
= Stop codon for translation in *E. coli*.

Table 28. Strains and plasmids generated in this thesis.

Name	Host	Plasmid	Resistance	Protein	Mutation	Origin
EV	<i>E. coli</i> BL21Gold (DE3) lacI ^{Q1}	pALXtreme-1a	Kan	---	---	---
P450 BM3 WT	<i>E. coli</i> BL21Gold (DE3) lacI ^{Q1}	pALXtreme-1a	Kan	P450 BM3	---	<i>B. megaterium</i>
P450 BM3 M1	<i>E. coli</i> BL21Gold (DE3) lacI ^{Q1}	pALXtreme-1a	Kan	P450 BM3	R47S, Y51W	<i>B. megaterium</i>
P450 BM3 M2	<i>E. coli</i> BL21Gold (DE3) lacI ^{Q1}	pALXtreme-1a	Kan	P450 BM3	R47S, Y51W, I401M	<i>B. megaterium</i>
P450 BM3 M3	<i>E. coli</i> BL21Gold (DE3) lacI ^{Q1}	pALXtreme-1a	Kan	P450 BM3	R47S, Y51W, A330F, I401M	<i>B. megaterium</i>
P450 BM3 WT	<i>E. coli</i> BL21 DH5alpha	pALXtreme-1a	Kan	P450 BM3	---	<i>B. megaterium</i>
P450 BM3 M1	<i>E. coli</i> BL21 DH5alpha	pALXtreme-1a	Kan	P450 BM3	R47S, Y51W	<i>B. megaterium</i>
P450 BM3 M2	<i>E. coli</i> BL21 DH5alpha	pALXtreme-1a	Kan	P450 BM3	R47S, Y51W, I401M	<i>B. megaterium</i>
P450 BM3 M3	<i>E. coli</i> BL21 DH5alpha	pALXtreme-1a	Kan	P450 BM3	R47S, Y51W, A330F, I401M	<i>B. megaterium</i>
YmPhytase	<i>E. coli</i> BL21 DH5alpha	pALXtreme-5b	Amp	Phytase	Mature Sequence	<i>Y. mollaretii</i>
Phytase 1	<i>E. coli</i> BL21 DH5alpha	pET22b	Amp	Phytase	Mature Sequence	<i>Y. mollaretii</i>
Phytase 2	<i>E. coli</i> BL21 DH5alpha	pET22b	Amp	Phytase	Mature Sequence	<i>E. coli</i>
Phytase 3	<i>E. coli</i> BL21 DH5alpha	pET22b	Amp	Phytase	Mature Sequence	<i>H. alvei</i>

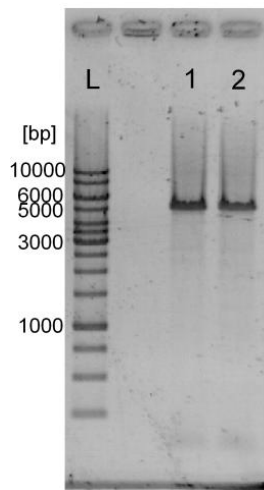


Figure 79. Separation of DNA on a 1% agarose gel. L = DNA ladder; Lane 1 and 2 PCR products from SSM at position F87.

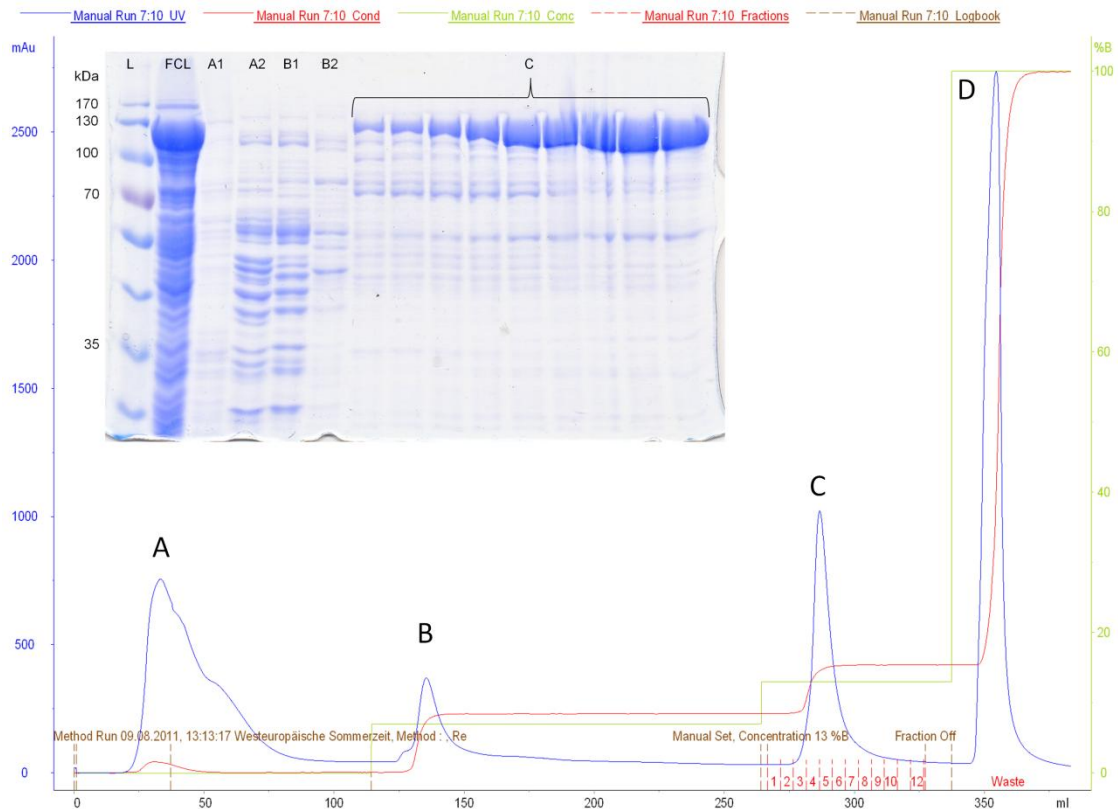


Figure 80. Purification of P450 BM3. Blue line: UV detector signal. Red line = solvent conductivity; Green line = percentage of elution buffer B (1 M Tris buffer pH 7.8). Peak A = Proteins not binding to the resin matrix. Peak B = Wash fraction. Peak C = Fraction containing P450 BM3. Peak D = Column cleaning and elution of residual proteins. Upper left shows the corresponding SDS-PAGE with proteins from all fractions. L = Molecular weight reference ladder. FCL = filtered cell lysate before loading on a column; A1 and A2 = fractions of proteins that do not bind to the resin matrix; B1 and B2 = proteins eluted by washing the column with low salt concentration. C = collected fractions containing P450 BM3.

Table 29. List of all analyzed compounds including retention times, respective GC-programs, extraction solvents and applied internal standard for quantification.

Compound	Column	RT [min]	GC-program	Extraction solvent	Co-solvent	Int. Standard
Toluene	Supreme 5	0.93	110°C isothermal	MTBE	DMSO (1 %)	CDDC
<i>o</i> -Cresol	Supreme 5	1.83	110°C isothermal	MTBE	DMSO (1 %)	CDDC
<i>m</i> -Cresol	Supreme 5	1.99	110°C isothermal	MTBE	DMSO (1 %)	CDDC
<i>p</i> -Cresol	Supreme 5	1.97	110°C isothermal	MTBE	DMSO (1 %)	CDDC
Benzylalcohol	Supreme 5	1.75	110°C isothermal	MTBE	DMSO (1 %)	CDDC
Phenol	Hydrodex- β -TBDAC	3.81	140°C isothermal	MTBE	DMSO (1,5 %)	---
Phenol	Supreme 5	1.41	110°C isothermal	MTBE	DMSO (1 %)	CDDC
Hydroquinone	Supreme 5	4.71	110°C isothermal	MTBE	DMSO (1 %)	CDCC
<i>p</i> -Xylene	Supreme 5	1.11	110°C isothermal	MTBE	DMSO (1 %)	CDDC
2,5-DMP	Supreme 5	2.77	110°C isothermal	MTBE	DMSO (1 %)	CDDC
2,5-DMHC	Supreme 5	5.70	110°C isothermal	MTBE	DMSO (1 %)	CDDC
4-Methylbenzylalcohol	Supreme 5	2.65	110°C isothermal	MTBE	DMSO (1 %)	CDDC/Phenol
<i>o</i> -Xylene	Optima-17 MS	3.01	120°C isothermal	CHCl ₃	DMSO (1,5 %)	Phenol
3,4-DMP	Optima-17 MS	9.38	120°C isothermal	CHCl ₃	DMSO (1,5 %)	Phenol
2,3-DMP	Optima-17 MS	8.77	120°C isothermal	CHCl ₃	DMSO (1,5 %)	Phenol
2-Methylbenzylalcohol	Optima-17 MS	8.48	120°C isothermal	CHCl ₃	DMSO (1,5 %)	Phenol
<i>m</i> -Xylene	Optima-17 MS	2.84	120°C isothermal	CHCl ₃	DMSO (1,5 %)	Phenol
2,4-DMP	Optima-17 MS	8.51	120°C isothermal	CHCl ₃	DMSO (1,5 %)	Phenol
3,5-DMP	Optima-17 MS	7.64	120°C isothermal	CHCl ₃	DMSO (1,5 %)	Phenol
2,6-DMP	Optima-17 MS	6.55	120°C isothermal	CHCl ₃	DMSO (1,5 %)	Phenol
3-Methylbenzylalcohol	Optima-17 MS	8.20	120°C isothermal	CHCl ₃	DMSO (1,5 %)	Phenol
Mesitylene	Supreme 5	1.41	110°C isothermal	MTBE	DMSO (1,5 %)	2,5-DMP
2,4,6-TMP	Supreme 5	3.56	110°C isothermal	MTBE	DMSO (1,5 %)	2,5-DMP
3,5-DMBA	Supreme 5	4.05	110°C isothermal	MTBE	DMSO (1,5 %)	2,5-DMP
Pseudocumene	Hydrodex- β -TBDAC	1.48	140°C isothermal	MTBE	DMSO (1,5 %)	Phenol
2,5-DMBA	Hydrodex- β -TBDAC	5.11	140°C isothermal	MTBE	DMSO (1,5 %)	Phenol
2,4-DMBA	Hydrodex- β -TBDAC	5.28	140°C isothermal	MTBE	DMSO (1,5 %)	Phenol
2,3,6-TMP	Hydrodex- β -TBDAC	5.63	140°C isothermal	MTBE	DMSO (1,5 %)	Phenol
2,3,5-TMP	Hydrodex- β -TBDAC	6.89	140°C isothermal	MTBE	DMSO (1,5 %)	Phenol
3,4-DMBA	Hydrodex- β -TBDAC	7.12	140°C isothermal	MTBE	DMSO (1,5 %)	Phenol
2,4,5-TMP	Hydrodex- β -TBDAC	7.70	140°C isothermal	MTBE	DMSO (1,5 %)	Phenol
Fluorobenzene	Hydrodex- β -TBDAC	1.12	140°C isothermal	MTBE	DMSO (1 %)	CDDC
2-Fluorophenol	Hydrodex- β -TBDAC	5.89	140°C isothermal	MTBE	DMSO (1 %)	CDDC
4-Fluorophenol	Hydrodex- β -TBDAC	8.65	140°C isothermal	MTBE	DMSO (1 %)	CDDC
3-Fluorophenol	Hydrodex- β -TBDAC	8.84	140°C isothermal	MTBE	DMSO (1 %)	CDDC
Fluorohydroquinone	Hydrodex- β -TBDAC	19.45	140°C isothermal	MTBE	DMSO (1 %)	CDDC
Iodobenzene	Supreme 5	1.84	110°C isothermal	MTBE	DMSO (1 %)	CDDC
2-Iodophenol	Supreme 5	3.41	110°C isothermal	MTBE	DMSO (1 %)	CDDC
3-Iodophenol	Supreme 5	5.593	110°C isothermal	MTBE	DMSO (1 %)	CDDC
4-Iodophenol	Supreme 5	5.591	110°C isothermal	MTBE	DMSO (1 %)	CDDC

Compound	Column	RT [min]	GC-program	Extraction solvent	Co-solvent	Int. Standard
Bromobenzene	Supreme 5	1.3	110°C isothermal	MTBE	DMSO (1 %)	CDDC
2-Bromophenol	Supreme 5	2.04	110°C isothermal	MTBE	DMSO (1 %)	CDDC
3-Bromophenol	Supreme 5	4.78	110°C isothermal	MTBE	DMSO (1 %)	CDDC
4-Bromophenol	Supreme 5	4.80	110°C isothermal	MTBE	DMSO (1 %)	CDDC
Bromohydroquinone	Supreme 5	5.73	110°C isothermal	MTBE	DMSO (1 %)	CDDC
Chlorobenzene	Supreme 5	1.06	110°C isothermal	MTBE	DMSO (1 %)	CDDC
2-Chlorophenol	Supreme 5	1.52	110°C isothermal	MTBE	DMSO (1 %)	CDDC
3-Chlorophenol	Supreme 5	3.45	110°C isothermal	MTBE	DMSO (1 %)	CDDC
4-Chlorophenol	Supreme 5	3.46	110°C isothermal	MTBE	DMSO (1 %)	CDDC
Chlorohydroquinone	Supreme 5	5.19	110°C isothermal	MTBE	DMSO (1 %)	CDDC
Anisole	Supreme 5	1.22	110°C isothermal	MTBE	DMSO (1 %)	CDDC
Guaiacol	Supreme 5	2.14	110°C isothermal	MTBE	DMSO (1 %)	CDDC
3-Methoxyphenol	Supreme 5	3.97	110°C isothermal	MTBE	DMSO (1 %)	CDDC
4-Methoxyphenol	Supreme 5	3.81	110°C isothermal	MTBE	DMSO (1 %)	CDDC
Methoxyhydroquinone	Supreme 5	5.33	110°C isothermal	MTBE	DMSO (1 %)	CDDC
<i>o</i> -Nitrotoluene	Supreme 5	1.52	135°C isothermal	EtOAc	DMSO (2 %)	2,5-DMP
<i>o</i> -Nitrobenzylalcohol	Supreme 5	3.24	135°C isothermal	EtOAc	DMSO (2 %)	2,5-DMP
<i>o</i> -Nitrobenzaldehyde	Supreme 5	2.25	135°C isothermal	EtOAc	DMSO (2 %)	2,5-DMP
<i>m</i> -Nitrotoluene	Supreme 5	1.56	150°C isothermal	EtOAc	DMSO (2 %)	2,5-DMP
<i>m</i> -Nitrobenzylalcohol	Supreme 5	4.23	150°C isothermal	EtOAc	DMSO (2 %)	2,5-DMP
<i>o</i> -Nitrobenzaldehyde	Supreme 5	2.40	150°C isothermal	EtOAc	DMSO (2 %)	2,5-DMP
<i>p</i> -Nitrotoluene	Supreme 5	1.68	150°C isothermal	EtOAc	DMSO (2 %)	2,5-DMP
<i>p</i> -Nitrobenzylalcohol	Supreme 5	4.67	150°C isothermal	EtOAc	DMSO (2 %)	2,5-DMP
<i>p</i> -Nitrobenzaldehyde	Supreme 5	2.32	150°C isothermal	EtOAc	DMSO (2 %)	2,5-DMP

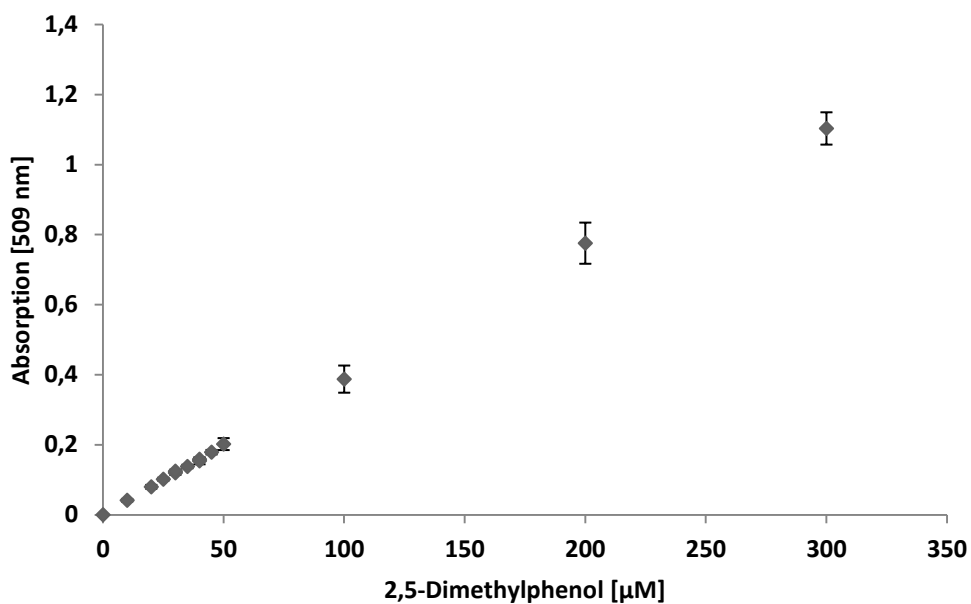


Figure 81. Calibration curve for 4-AAP assay and the commercial standard of 2,5 DMP in MTP format. The coefficient of determination was calculated to be $R^2 = 0.999$ from 0 to 300 μM 2,5-DMP.

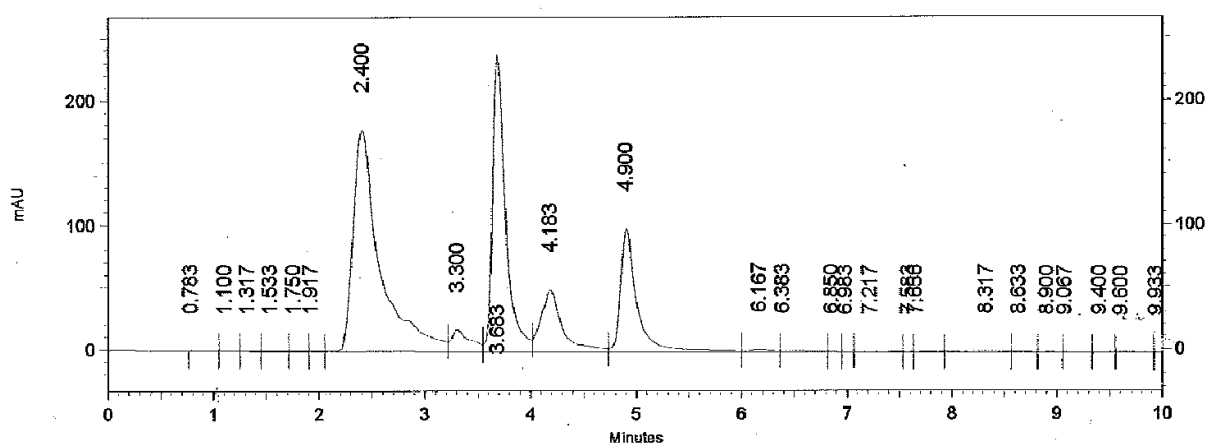


Figure 82. UV-detection profile from a conversion of chlorobenzene by P450 BM3 variant M2 after 8 h reaction time analyzed by HPLC chromatography. Chlorobenzene (4.9 min); 2-chlorophenol (4.183 min); chlorohydroquinone (3.683 min); NADPH and glucose (2.4 min).

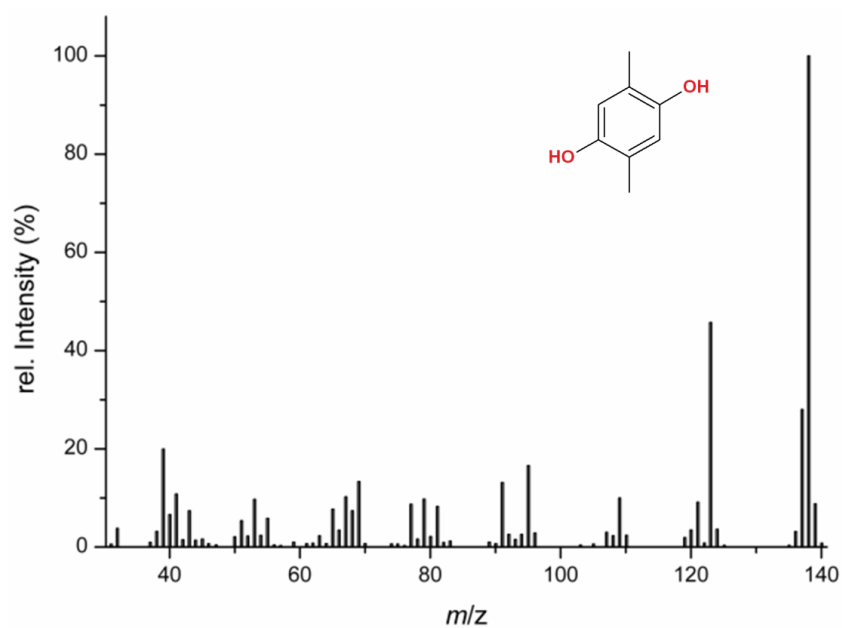


Figure 83. GC-MS fragmentation profile of 2,5-dimethylhydroquinone ($m/z = 138.05$). The product was detected during conversions of *p*-xylene with P450 BM3 WT, variant M2 and M3.

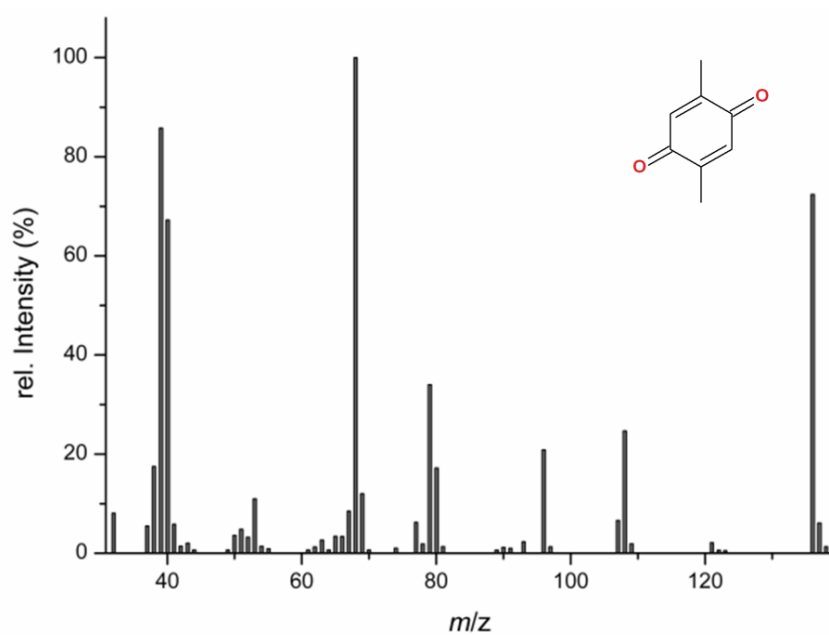


Figure 84. 2,5-dimethylquinone (oxidized product of 2,5-dimethylhydroquinone) was detected as minor fraction during long term conversion with *p*-xylene. Commonly this product represents less than 5 % (GC-peak area) in the reaction.

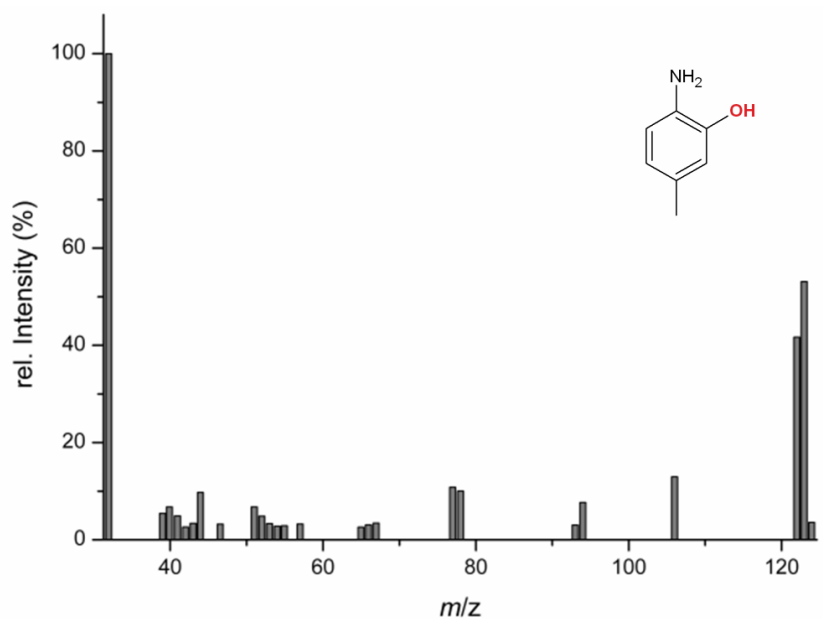


Figure 85. GC-MS fragmentation of the product obtained from conversion of *p*-toluidin with variant M2 in presence of a GDH regeneration system. The product was identified as 2-amino-5-methylphenol.

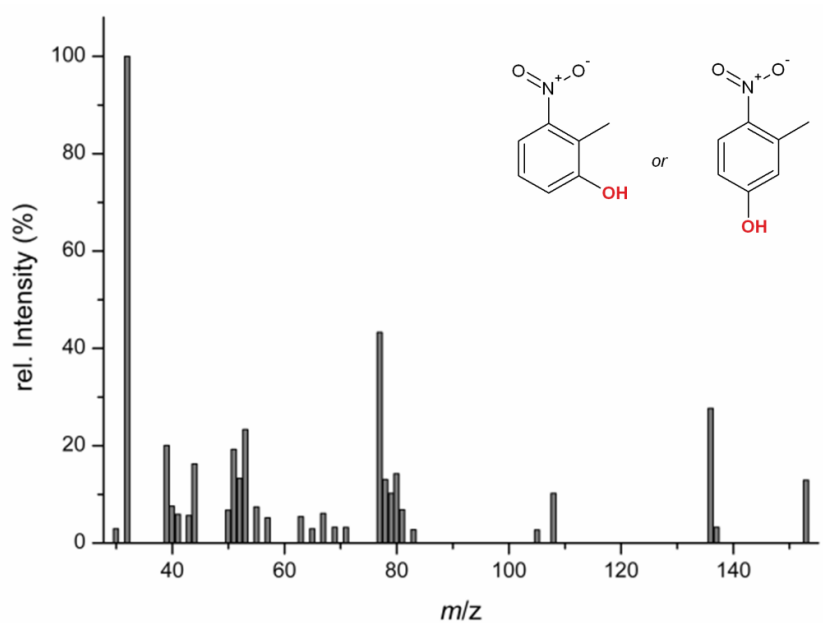


Figure 86. GC-MS fragmentation of phenolic products obtained from conversion of *o*-nitrotoluene with variant M3. Two isomeric product structures were suggested: 2-methyl-3-nitrophenol and 3-methyl-4-nitrophenol.

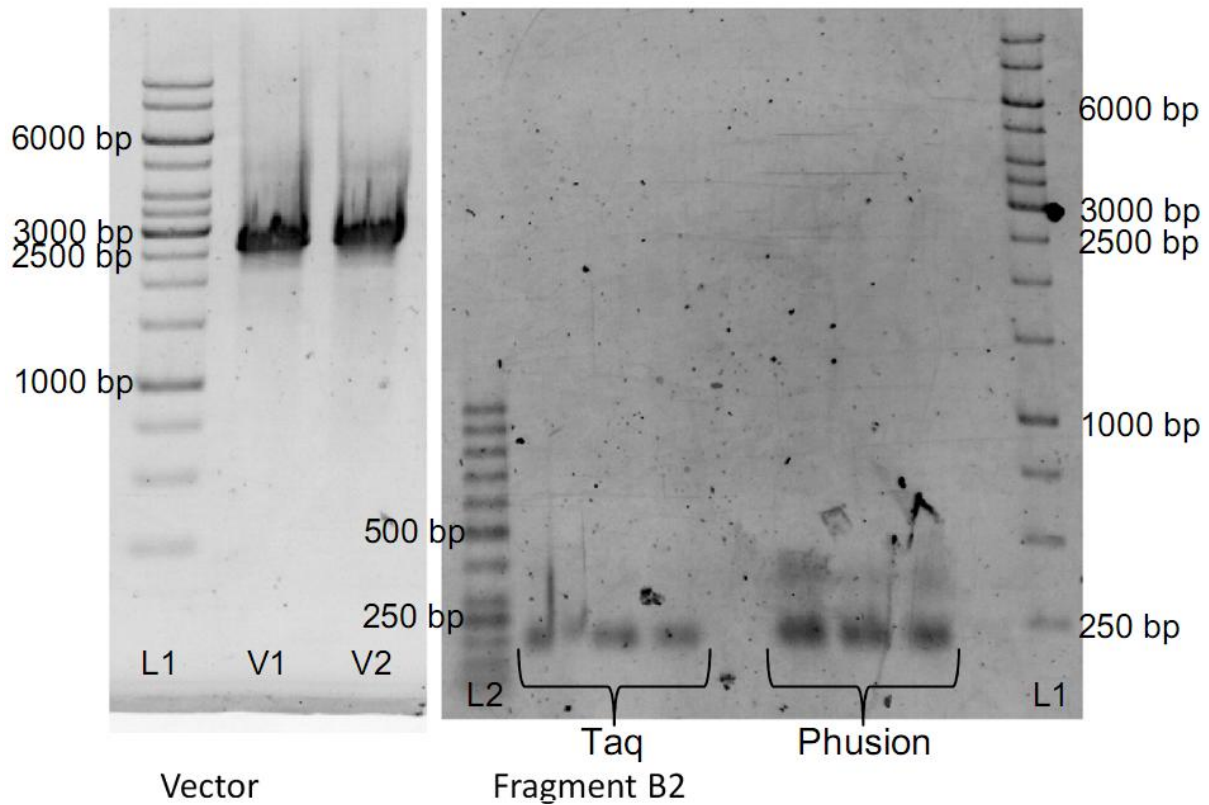


Figure 87. Analysis of PCR products by agarose gel electrophoresis during development of the OmniChange method. Exemplary the amplified vector and fragment B2 are shown. L1 = 1 kb DNA ladder; L2 = 50 bp DNA ladder; V1 & V2 = Amplified vector fragments. Lanes “Taq” display amplified fragment B using Taq DNA polymerase. Lanes “Phusion” display amplified fragment B using Phusion DNA polymerase.

		T77	K139	G187	V298
A	pALXtreme-5b	igtgacactg	atttgaaaaagg	gatgggcgag	gattgtgacg
	6103302.seq	igtgacactg	atttgaaaaagg	gatgggcgag	gattgtgacg
	6103296.seq	igtgaagctg	atttgacgaagg	gatgtgggag	gattacgacg
	6103298.seq	igtggcgctg	atttgttgaagg	gatgcttgag	gattgagacg
	6103300.seq	igtggagctg	atttgaaaaagg	gatgggcgag	gattgtgacg
B		D52	T77		
	pALXtreme-5b	gaatgatgtc	tagtgacactgat		
	14Fw-T7.ab1	gaatthtctg	tagtggttctgat		
	11Fw-T7.ab1	gaatcctgtc	tagtgaggctgat		
	12Fw-T7.ab1	gaatthtctg	tagtggttctgat		
13Fw-T7.ab1	gaataatgtc	tagtggttctgat			

Figure 88. Sequence alignment of OmniChange sub-libraries. **A:** Saturation of positions T77, K139, G187 and V298. **B:** Saturation of positions D52 and T77. All other codons remained unchanged. Obtained nucleotide exchanges are highlighted in white. WT nucleotides of the YmPhytase are highlighted in yellow.

Table 30. Primer combinations employed for saturation of codons D52 and T77 in the YmPhytase sequence during development of the OmniChange method (see result in Figure 89).

Fragment Amplification	Primer Name	Primer Sequence (5'-3')
C/D/E/A ₂ -Vector-A ₁	K139 _{FW}	ggcggcttctatgGTGATTATTTCCG
	D52 _{RV}	ccacttatccggTGTGACMNNATTACTTAAGTCTG
A ₂ -Vector-A ₁ (including position G11)	G11 _{FW}	cgatcagtgctgCACCAGTTGC
	G11 _{RV}	cagcactgatcgCAAACCTMNNTAGCATTAAAGCCCAGTGC
B'	T77 _{FW}	ccggataagtggtCCTCAATGGCCGGTAC
	T77 _{RV}	catagaagccgccCATCAGMNNCACTAATTGCGC

appA	1	atg----a-aagcagctc-----taateccatt-----t--ttatctctctg-----attcc--g--ttaaccgcccaat--
phyX	1	atg----acaatctctctgtttaaccgtaataaacccgctattgcacagcgtatt--ttatgtctctg-----atcgt--ggctttattctcaggttta
ympH	1	atggtgaa-aatagcaat-----gaatagcatg-----cgattaaactgcactgggcttaatggtag--gtagttttgcat--
appA	56	-ctg-----catcgcctcagagtg-----gccggagctga-----agctggaaagtgtggtgatgtcagctcgtcatggtgtggtgctccaaccaa
phyX	88	ccgg-----catacgc-----cagtgatcacgcccctgctgggtccagttggaaaaggttgttatcctaaagcagacatggctacgcccggcaaccaa
ympH	69	-cagtgctgcaccagttgctgcaccagt-----gaegggttaca-----ctttggagcagtggttatcttgagccgccatggttctgctgcgcgagaa
appA	138	ggccacgcaactgatgcagatgtcaccccagacgcatggccaacctggccggtaaaactgggttggctgacaccgcccgggtggtgagctaatcgctat
phyX	177	aatgacacaaacgatgcgcaactgcacacctcacagtgccctgaatggccggtaaaactcggctatatacagcccggcggtaaacatctgatagcctg
ympH	159	acaacacagagttaatgaatgatgtcacaccggtaagtggcctcaatggccggtaccagcggggtatttaacccccagaggtgcgaatagtgacactg
appA	238	ctcggacattacaaacgccagcgtctggtagcgc--acggatgctggcgaaaagggtgcccgcagctcgtgctcagctcgcgattattgctgatgtcgc
phyX	277	atgggcggtttttatcgagagcg-ctttcagcaacaaggcttatacctaaaggataaactgtcctacaccagatcccgtgatgtttgggcagacgctcgat
ympH	259	atggcggcttctctatggtgatattcccgcaatc--aaggatgctcccggca--ggatgcccggcagatggaacctctatgcccgaagccgatctcgat
appA	337	gagcgtaccgctaaaacagcgaagccttgcgcggcggctggcactgactgtgcaataaacgctacataccaggcagatagctccagct--ccgatcc
phyX	376	caaccacacagtaaaaacggcgaggcctcttagcgggtctgctcccagtgatgattagcagatccaccatcag-ca--aaacattcagcagggcagctcc
ympH	355	caacgaacccggttaactgggcaagcattccttggatagctagctcccgggtcgggtcctaaagtgcatatcaggtcatttgaaaaagg--ttgatcc
appA	435	gttatttaactcctcaaaaactggcgtttgccaaactggata--acggcaac-----gtgactgacgcgactcctcagcagggcaggaggtcaatgct
phyX	474	gctgttccatcctgtgaaagccggtatctgctcagtgata--aatcacag-----gcacacgcccggcttgaaaagcagggcagccacaccgatg-a
ympH	453	actgtttccaccgctgaaagccggtgctgtcagctagatcgacacaaacccataggggccatgagggcgaact-----ggggggccatbaagt
appA	526	gactttaccgggcatcggcaaacggcgtttcggcaactggaacg--ggtgcttaattttccgcaatcaaacctgtgctttaaactgtgagaacacaggag
phyX	564	gagcct--caatcaacgctatcaagcacttttagcctgatgagttcggtaactcgatttccaaaatccccctatg--ccagcagcacaacatggga
ympH	544	gaacttagccagcgttatgctaaagcatttgccagatggcgca--gattctcaattcactgcttccccctattgcagtcactacagcaacaagaa
appA	623	aaagctgttcattaacgcagcattaccatcg--gaactcaaggtgagc-----gccgac-----aatg-tctcataaacgggtgcccgaagcctcgca
phyX	659	aaactctcgatttttcacagggcagctcagc--agcctggcagataat-----gacagcggtaataaag-tggctctcgaaggtgcccgggactttca
ympH	641	aaactctgtgatttt--gccacctttgctgccaatgaagtaaggtgaactcagcagggggac-----aaaggtatccctcagctggcactggcactca
appA	709	tcaatgctgacggagatatttctcctgcaaacagcagcaggaatgccggagccgggtggggaaggatcacagctcacaccagtggaaca-----c
phyX	751	tcgacgttggctgaaatttctcctgctggaacacgctcaggaatgctaaagtggttgggggaatattcacactgagcagcaatgggact-----c
ympH	733	tcacatctgggtgaaatcttcttctgctacaaaattgcgaagggatgccggatgctgcttgg-----catcgattaaagtgagcgggaaattgggtctc
appA	801	cttgtaagtttgcataaacgcgcaattttatttgctacaacgcagccagaggttgcccgcagccgcccaccccgcttattagatttgatcaagacagcg
phyX	843	ctcgttasaattgcataaatgcccgcttgacttgacttgatgctcggcagccctatatcgcc-----aagcataaacggtaactccactgctgcaaacatcg
ympH	825	attatatacgcctgcataaatgcgcaatttgatttgatggctaaaacaccttatatcgccgctcataaagggaactccgttgttgcaacagatgtgacggcg
appA	901	ttgacgcccattccaccgcaaaaacagggct----atgg-----tgtgacattacc-----acttcagtg----ctgtttatcgccgacacgat
phyX	935	cacacgcac--tgggttccaatatacggagctgcccactg-----ccggatatttctg-----cagacaataagatcctgttattgcccgtcacgac
ympH	925	ctagtcg----ttcagcgtta----agggcg-----aagcccaaaccttgccattatctgagcagaccaaactc-----ctttcttggcgggtcagat
appA	979	actaatctggcaaatctcggcggcgaactggagctcaactggagcctcccggtcagccggataaacacgcccaggtggtgaactggtgtttgaagcct
phyX	1021	accaatattgccaatatttctggcatgctagggatgacatggacactccgggacagccagataaacacgctccggggcgggctttagtgtttgaagcgtt
ympH	1006	accaatattgccaatattgcccgtatgctaggagcaactggcagctaccgcaacagcccagataaacaccccggcgggtgggggctggtgtttgaactat
appA	1079	ggcgtcggctaaagcagataaacagccag--tggattcaggtttcgcctgcttccagactttacagcagatcgcgtgata----aaacggcctgtcattaa
phyX	1121	gg-gtagataaacgggggaaa--ccg--tatggttagcgtgaataggtgtatcaaacactggcagcttgcacagc--agacggcgcctaacgcttgc
ympH	1106	ggcagatacca--gataacatcagcaatattgctcagcttaagatgtctatcaaacatggtcagctacgaaatagtgaaaagttagaccfgaaaag
appA	1172	ataccgcccggagaggtgaaactgacctggcagatgtgaagagcgaatggcagggcagctgttctgttggcaggttt-tacg--caaatcgtgaa
phyX	1211	agcatcctgcccagcgtacgactaaacatccgggttgcagcagatcaaacgcccagtggtctatgcccgcctcaccctt-cagc--cgtttagtcaa
ympH	1203	tcatacagccggtattgttccattga-gatcgaaggttgtga----gaacatcg--gtacagacaacctttgcccagcttgatccttccaaaagagag
appA	1269	tgaagcagcagc-----accggcgtgc-agtttg--taa
phyX	1308	ccacagcgttga----gcctcgtgcccagcttccctaa
ympH	1295	tggctcaggtgatgaacctgcatgc-catat--taa

Figure 89. Full DNA sequence alignment of three model phytases employed during the development of PTRec. Areas with high sequence homology are highlighted in blue whereas areas with low sequence identity are highlighted in white. CloneManager 9 software was used for DNA sequence alignments.

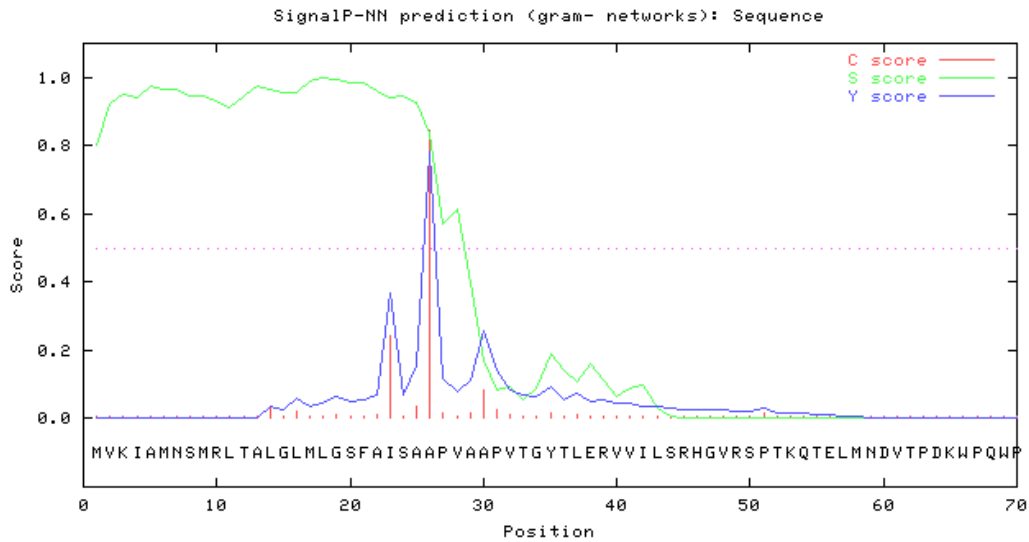


Figure 90. Prediction of cleavage site of ymPH phytase signal sequence in *E. coli* using SignalP algorithm for gram negative bacteria. The red vertical line indicates the most likely cleavage by *E. coli* according to highest calculated C score. With 99.9 % probability the cleavage takes place between residue 25 and 26 (Ala/Ala).

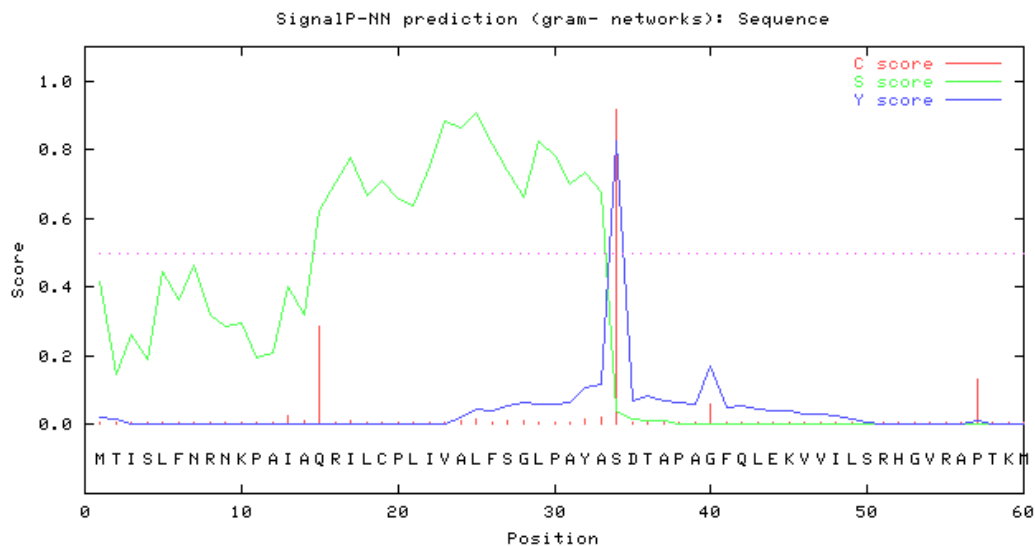


Figure 91. Prediction of cleavage site of phyX phytase signal sequence in *E. coli* using SignalP algorithm for gram negative bacteria. The red vertical line indicates the most likely cleavage position due to highest calculated C score. With 99.9 % probability the cleavage takes place between residue 34 and 35 (Ser/Ala).

6. References

- 1 Kerr, R. A. (2011) *Energy supplies. Peak oil production may already be here*. Science. 331, 1510-1511
- 2 Murphy, D. J. (2012) *Fossil fuels: Peak oil is affecting the economy already*. Nature. 483, 541
- 3 Witze, A. (2007) *Energy: that's oil, folks*. Nature. 445, 14-17
- 4 Brandt, A. R., Millard-Ball, A., Ganser, M. and Gorelick, S. M. (2013) *Peak oil demand: The role of fuel efficiency and alternative fuels in a global oil production decline*. Environ. Sci. Technol. 47, 8031-8041
- 5 Murray, J. and King, D. (2012) *Climate policy: Oil's tipping point has passed*. Nature. 481, 433-435
- 6 Howarth, R. W., Ingraffea, A. and Engelder, T. (2011) *Natural gas: Should fracking stop?* Nature. 477, 271-275
- 7 Giles, J. (2004) *Oil exploration: every last drop*. Nature. 429, 694-695
- 8 European, Commission and (2012) *Innovating for Sustainable Growth: A Bioeconomy for Europe*. http://ec.europa.eu/research/bioeconomy/pdf/201202_innovating_sustainable_growth_en.pdf
- 9 Antoni, D., Zverlov, V. V. and Schwarz, W. H. (2007) *Biofuels from microbes*. Appl. Microbiol. Biotechnol. 77, 23-35
- 10 Peralta-Yahya, P. P., Zhang, F., del Cardayre, S. B. and Keasling, J. D. (2012) *Microbial engineering for the production of advanced biofuels*. Nature. 488, 320-328
- 11 de Jong, B., Siewers, V. and Nielsen, J. (2012) *Systems biology of yeast: enabling technology for development of cell factories for production of advanced biofuels*. Curr. Opin. Biotechnol. 23, 624-630
- 12 Philp, J. C., Ritchie, R. J. and Allan, J. E. (2013) *Biobased chemicals: the convergence of green chemistry with industrial biotechnology*. Trends Biotechnol. 31, 219-222
- 13 Christensen, C. H., Rass-Hansen, J., Marsden, C. C., Taarning, E. and Egeblad, K. (2008) *The renewable chemicals industry*. ChemSusChem. 1, 283-289
- 14 Vennestrom, P. N., Osmundsen, C. M., Christensen, C. H. and Taarning, E. (2011) *Beyond petrochemicals: the renewable chemicals industry*. Angew. Chem., Int. Ed. Engl. 50, 10502-10509
- 15 European, Commission and (2009) *"Preparing for our future: Developing a common strategy for key enabling technologies in the EU"*. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0512:FIN:EN:PDF>
- 16 Feldbaum, C. (2002) *Biotechnology. Some history should be repeated*. Science. 295, 975
- 17 Clardy, J., Fischbach, M. A. and Walsh, C. T. (2006) *New antibiotics from bacterial natural products*. Nat. Biotechnol. 24, 1541-1550
- 18 Linko, Y.-Y., Javanainen, P. and Linko, S. (1997) *Biotechnology of bread baking*. Trends in Food & Science Technology. 8, 339-344
- 19 Buchholz, K. and Collins, J. (2013) *The roots--a short history of industrial microbiology and biotechnology*. Appl. Microbiol. Biotechnol. 97, 3747-3762

- 20 Vogel, G. (2013) *Biomedicine. Human stem cells from cloning, finally*. Science. 340, 795
- 21 OECD. (2012) *OECD Factbook 2011-2012: Economic, Environmental and Social Statistics*. <http://www.oecd-ilibrary.org/sites/factbook-2011-en/08/01/04/index.html;jsessionid=1o9xjursobuf3.delta?contentType=/ns/StatisticalPublication,/ns/Chapter&itemId=/content/chapter/factbook-2011-71-en&containerItemId=/content/serial/18147364&accessItemIds=&mimeType=text/html>
- 22 Green, M. L. and Crutchfield, G. (1969) *Studies on the preparation of water-insoluble derivatives of rennin and chymotrypsin and their use in the hydrolysis of casein and the clotting of milk*. Biochem J. 115, 183-190
- 23 Perez-Traves, L., Lopes, C. A., Barrio, E. and Querol, A. (2012) *Evaluation of different genetic procedures for the generation of artificial hybrids in Saccharomyces genus for winemaking*. Int. J. Food Microbiol. 156, 102-111
- 24 Elander, R. P. (2003) *Industrial production of beta-lactam antibiotics*. Appl. Microbiol. Biotechnol. 61, 385-392
- 25 Avery, O. T., MacLeod, C. M. and McCarthy, M. (1944) *Studies on the chemical nature of the substance inducing transformation of pneumococcal types*. J. Exp. Med. 79, 137-158
- 26 Watson, J. D. and Crick, F. H. (1953) *Genetical implications of the structure of deoxyribonucleic acid*. Nature. 171, 964-967
- 27 Watson, J. D. and Crick, F. H. (1953) *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid*. Nature. 171, 737-738
- 28 Watson, J. D. and Crick, F. H. (1953) *The structure of DNA*. Cold Spring Harbor symposia on quantitative biology. 18, 123-131
- 29 De Leeuw, I., Delvigne, C. and Bekaert, J. (1982) *Insulin allergy treated with human insulin (recombinant DNA)*. Diabetes Care. 5 Suppl. 2, 168-170
- 30 Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, C. A., Hutchison, C. A., Slocombe, P. M. and Smith, M. (1977) *Nucleotide sequence of bacteriophage phi X174 DNA*. Nature. 265, 687-695
- 31 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *DNA sequencing with chain-terminating inhibitors*. Proc. Natl. Acad. Sci. U. S. A. 74, 5463-5467
- 32 Turner, N. J. (2009) *Directed evolution drives the next generation of biocatalysts*. Nat. Chem. Biol. 5, 567-573
- 33 Arnold, F. H. (2009) *How proteins adapt: lessons from directed evolution*. Cold Spring Harbor symposia on quantitative biology. 74, 41-46
- 34 Cobb, R. E., Si, T. and Zhao, H. (2012) *Directed evolution: an evolving and enabling synthetic biology tool*. Curr. Opin. Chem. Biol. 16, 285-291
- 35 Cobb, R. E., Sun, N. and Zhao, H. (2013) *Directed evolution as a powerful synthetic biology tool*. Methods. 60, 81-90
- 36 Ruff, A. J., Dennig, A. and Schwaneberg, U. (2013) *To get what we aim for - progress in diversity generation methods*. FEBS J. 280, 2961-2978
- 37 Shivange, A. V., Marienhagen, J., Mundhada, H., Schenk, A. and Schwaneberg, U. (2009) *Advances in generating functional diversity for directed protein evolution*. Curr. Opin. Chem. Biol. 13, 19-25

- 38 Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R. Y., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z. Q., Segall-Shapiro, T. H., Calvey, C. H., Parmar, P. P., Hutchison, C. A., 3rd, Smith, H. O. and Venter, J. C. (2010) *Creation of a bacterial cell controlled by a chemically synthesized genome*. *Science*. 329, 52-56
- 39 Dong, Y., Xie, M., Jiang, Y., Xiao, N., Du, X., Zhang, W., Tosser-Klopp, G., Wang, J., Yang, S., Liang, J., Chen, W., Chen, J., Zeng, P., Hou, Y., Bian, C., Pan, S., Li, Y., Liu, X., Wang, W., Servin, B., Sayre, B., Zhu, B., Sweeney, D., Moore, R., Nie, W., Shen, Y., Zhao, R., Zhang, G., Li, J., Faraut, T., Womack, J., Zhang, Y., Kijas, J., Cockett, N., Xu, X., Zhao, S., Wang, J. and Wang, W. *Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (Capra hircus)*. *Nat. Biotechnol.* 31, 135-141
- 40 Nordstrom, K. J., Albani, M. C., James, G. V., Gutjahr, C., Hartwig, B., Turck, F., Paszkowski, U., Coupland, G. and Schneeberger, K. *Mutation identification by direct comparison of whole-genome sequencing data from mutant and wild-type individuals using k-mers*. *Nat. Biotechnol.* 31, 325-330
- 41 Siuti, P., Yazbek, J. and Lu, T. K. (2013) *Synthetic circuits integrating logic and memory in living cells*. *Nat. Biotechnol.* 31, 448-452
- 42 Carbonell, P., Planson, A. G. and Faulon, J. L. (2013) *Retrosynthetic design of heterologous pathways*. *Methods. Mol. Biol.* 985, 149-173
- 43 Chovan, T. and Guttman, A. (2002) *Microfabricated devices in biotechnology and biochemical processing*. *Trends Biotechnol.* 20, 116-122
- 44 Niemeyer, C. M. (2001) *Nanoparticles, proteins and nucleic acids: Biotechnology meets materials science*. *Angew. Chem., Int. Ed. Engl.* 40, 4128-4152
- 45 BmBF. (2010) *Biotechnologie in Deutschland*. http://www.bmbf.de/pub/biotechnologie_in_deutschland.pdf, 14-16
- 46 Research, F. M. o. E. a. (2012) *Biotechnology in Germany - the future on your doorstep*. <http://www.biotechnikum.eu/en/biotechnology/uses-and-benefits/biotechnology-in-germany.html>
- 47 Bornscheuer, U. T., Huisman, G. W., Kazlauskas, R. J., Lutz, S., Moore, J. C. and Robins, K. (2012) *Engineering the third wave of biocatalysis*. *Nature*. 485, 185-194
- 48 Wandrey, C., Liese, A. and Kihumbu, D. (2000) *Industrial biocatalysis: Past, present and future*. *Org. Process Res. Dev.* 4, 286-290
- 49 Koszelewski, D., Tauber, K., Faber, K. and Kroutil, W. (2010) *omega-Transaminases for the synthesis of non-racemic alpha-chiral primary amines*. *Trends Biotechnol.* 28, 324-332
- 50 Munoz Solano, D., Hoyos, P., Hernaiz, M. J., Alcantara, A. R. and Sanchez-Montero, J. M. (2012) *Industrial biotransformations in the synthesis of building blocks leading to enantiopure drugs*. *Bioresour. Technol.* 115, 196-207
- 51 Jakoblinnert, A., Mladenov, R., Paul, A., Sibilla, F., Schwaneberg, U., Ansorge-Schumacher, M. B. and de Maria, P. D. (2011) *Asymmetric reduction of ketones with recombinant E. coli whole cells in neat substrates*. *Chem. Commun.* 47, 12230-12232
- 52 Castro, G. R. and Knubovets, T. (2003) *Homogeneous biocatalysis in organic solvents and water-organic mixtures*. *Crit. Rev. Biotechnol.* 23, 195-231
- 53 Nielsen, J., Larsson, C., van Maris, A. and Pronk, J. (2013) *Metabolic engineering of yeast for production of fuels and chemicals*. *Curr. Opin. Biotechnol.* 24, 398-404
- 54 Carbonell, P., Fichera, D., Pandit, S. B. and Faulon, J. L. (2012) *Enumerating metabolic pathways for the production of heterologous target chemicals in chassis organisms*. *BMC Syst. Biol.* 6, 10

- 55 Caspeta, L. and Nielsen, J. (2013) *Toward systems metabolic engineering of Aspergillus and Pichia species for the production of chemicals and biofuels*. Biotechnol. J. 8, 534-544
- 56 Wieschalka, S., Blombach, B., Bott, M. and Eikmanns, B. J. (2012) *Bio-based production of organic acids with Corynebacterium glutamicum*. Microb. Biotechnol. 6, 87-102
- 57 Ikeda, M. (2003) *Amino acid production processes*. Adv. Biochem. Eng. Biotechnol. 79, 1-35
- 58 Singh, R. K., Tiwari, M. K., Singh, R. and Lee, J. K. (2013) *From protein engineering to immobilization: promising strategies for the upgrade of industrial enzymes*. Int. J. Mol. Sci. 14, 1232-1277
- 59 Schrewe, M., Julsing, M. K., Buhler, B. and Schmid, A. (2013) *Whole-cell biocatalysis for selective and productive C-O functional group introduction and modification*. Chem. Soc. Rev.
- 60 Becker, J. and Wittmann, C. (2012) *Bio-based production of chemicals, materials and fuels - Corynebacterium glutamicum as versatile cell factory*. Curr. Opin. Biotechnol. 23, 631-640
- 61 Zhang, Y. H. (2011) *Substrate channeling and enzyme complexes for biotechnological applications*. Biotechnol. Adv. 29, 715-725
- 62 Schrittwieser, J. H., Sattler, J., Resch, V., Mutti, F. G. and Kroutil, W. (2011) *Recent biocatalytic oxidation-reduction cascades*. Curr. Opin. Chem. Biol. 15, 249-256
- 63 Hollmann, F., Arends, I. W. C. E., Buehler, K., Schallmeyer, A. and Buehler, B. (2011) *Enzyme-mediated oxidations for the chemist* Green Chem. 13, 226-265
- 64 Woodley, J. M., Bisschops, M., Straathof, A. J. J. and Ottens, M. (2008) *Perspective future directions for in-situ product removal (ISPR)*. J. Chem. Technol. Biotechnol. 83, 121-123
- 65 Johannes, T., Simurdiak, M. R. and Zhao, H. (2006) *Biocatalysis*. Encyclopedia of Chemical Processing. DOI: 10.1081/E-ECHP-120017565
- 66 Schoemaker, H. E., Mink, D. and Wubbolts, M. G. (2003) *Dispelling the myths--biocatalysis in industrial synthesis*. Science. 299, 1694-1697
- 67 Zhou, S., Yomano, L. P., Shanmugam, K. T. and Ingram, L. O. (2005) *Fermentation of 10% (w/v) sugar to D: (-)-lactate by engineered Escherichia coli B*. Biotechnol. Lett. 27, 1891-1896
- 68 Dennig, A., Lülldorf, N., Liu, H. and Schwaneberg, U. (2013) *Regioselective o-hydroxylation of monosubstituted benzenes by P450 BM3*. Angew. Chem., Int. Ed. Engl. 52, 8459-8462
- 69 Dennig, A., Marienhagen, J., Ruff, A. J., Guddat, L. and Schwaneberg, U. (2012) *Directed Evolution of P450 BM3 into a p-Xylene hydroxylase*. ChemCatChem. 4, 771-773
- 70 Whitehouse, C. J., Bell, S. G. and Wong, L. L. (2012) *P450(BM3) (CYP102A1): connecting the dots*. Chem. Soc. Rev. 41, 1218-1260
- 71 Lewis, J. C., Coelho, P. S. and Arnold, F. H. (2011) *Enzymatic functionalization of carbon-hydrogen bonds*. Chem. Soc. Rev. 40, 2003-2021
- 72 Patel, R. N. (2001) *Biocatalytic synthesis of intermediates for the synthesis of chiral drug substances*. Curr. Opin. Biotechnol. 12, 587-604
- 73 Pollard, D. J. and Woodley, J. M. (2007) *Biocatalysis for pharmaceutical intermediates: the future is now*. Trends Biotechnol. 25, 66-73
- 74 Kumar, A. and Singh, S. (2012) *Directed evolution: tailoring biocatalysts for industrial applications*. Crit. Rev. Biotechnol. 33, 365-378

- 75 Dunn, P. J. (2011) *The importance of green chemistry in process research and development*. Chem. Soc. Rev. 41, 1452-1461
- 76 Yamada, R., Hasunuma, T. and Kondo, A. (2013) *Endowing non-cellulolytic microorganisms with cellulolytic activity aiming for consolidated bioprocessing*. Biotechnol. Adv. 31, 754-763
- 77 Blank, L. M., Ebert, B. E., Buhler, B. and Schmid, A. (2008) *Metabolic capacity estimation of Escherichia coli as a platform for redox biocatalysis: constraint-based modeling and experimental verification*. Biotechnol. Bioeng. 100, 1050-1065
- 78 Blank, L. M., Ionidis, G., Ebert, B. E., Buhler, B. and Schmid, A. (2008) *Metabolic response of Pseudomonas putida during redox biocatalysis in the presence of a second octanol phase*. FEBS J. 275, 5173-5190
- 79 Sattler, J. H., Fuchs, M., Tauber, K., Mutti, F. G., Faber, K., Pfeffer, J., Haas, T. and Kroutil, W. (2012) *Redox self-sufficient biocatalyst network for the amination of primary alcohols*. Angew. Chem., Int. Ed. Engl. 51, 9156-9159
- 80 Tao, J. and Xu, J. H. (2009) *Biocatalysis in development of green pharmaceutical processes*. Curr. Opin. Chem. Biol. 13, 43-50
- 81 Bacchus, W., Weber, W. and Fussenegger, M. (2013) *Increasing the dynamic control space of mammalian transcription devices by combinatorial assembly of homologous regulatory elements from different bacterial species*. Metab. Eng. 15, 144-150
- 82 Bruggink, A., Schoevaart, R. and Kieboom, T. (2003) *Concepts of Nature in Organic Synthesis: Cascade Catalysis and Multistep Conversions in Concert*. Org. Process Res. Dev. 7, 622-640
- 83 Bucke, C. (1983) *Practicality of industrial enzymes*. Biochemical Society transactions. 11, 13-14
- 84 Suekane, M. (1982) *Immobilization of glucose isomerase*. Zeitschrift fur allgemeine Mikrobiologie. 22, 565-576
- 85 Liu, J., Peng, J., Shen, S., Jin, Q., Li, C. and Yang, Q. (2013) *Enzyme entrapped in polymer-modified nanopores: the effects of macromolecular crowding and surface hydrophobicity*. Chemistry (Weinheim an der Bergstrasse, Germany). 19, 2711-2719
- 86 Singh, R. (2011) *Facts, growth, and opportunities in industrial biotechnology*. Org. Process Res. Dev. 15, 175-179
- 87 Kirk, O., Borchert, T. V. and Fuglsang, C. C. (2002) *Industrial enzyme applications*. Curr. Opin. Biotechnol. 13, 345-351
- 88 Duvick, D. N. (2001) *Biotechnology in the 1930s: the development of hybrid maize*. Nat. Rev. Genet. 2, 69-74
- 89 Grant, W. F. and Owens, E. T. (2006) *Zea mays assays of chemical/radiation genotoxicity for the study of environmental mutagens*. Mutat. Res. 613, 17-64
- 90 Knight, J. G., Mather, D. W., Holdsworth, D. K. and Ermen, D. F. (2007) *Acceptance of GM food--an experiment in six countries*. Nat. Biotechnol. 25, 507-508
- 91 Arnold, F. H. and Moore, J. C. (1997) *Optimizing industrial enzymes by directed evolution*. Adv. Biochem. Eng. Biotechnol. 58, 1-14
- 92 Arnold, F. H. and Volkov, A. A. (1999) *Directed evolution of biocatalysts*. Curr. Opin. Chem. Biol. 3, 54-59

- 93 Daugulis, A. J. (1991) *Integrated product formation and recovery*. *Curr. Opin. Biotechnol.* 2, 408-412
- 94 Qazi, J. I. (2013) *Biotechnological potential and conservatory of extremophiles from climatically wide ranged developing countries: Lesson from Pakistan*. *Crit. Rev. Microbiol.* 39, 1-8
- 95 Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. and Sayers, E. W. (2012) *GenBank*. *Nucleic Acids Res.* 41, D36-42
- 96 Pagani, I., Liolios, K., Jansson, J., Chen, I. M., Smirnova, T., Nosrat, B., Markowitz, V. M. and Kyrpides, N. C. (2012) *The Genomes OnLine Database (GOLD) v.4: status of genomic and metagenomic projects and their associated metadata*. *Nucleic Acids Res.* 40, D571-579
- 97 Arnold, F. H. and Georgiou, G. (2003) *Directed Evolution Library Creation: Methods and Protocols*. Humana Press, Totowa.
- 98 Wong, T. S., Roccatano, D. and Schwaneberg, U. (2007) *Steering directed protein evolution: strategies to manage combinatorial complexity of mutant libraries*. *Environ. Microbiol.* 9, 2645-2659
- 99 Wong, T. S., Roccatano, D. and Schwaneberg, U. (2007) *Challenges of the genetic code for exploring sequence space in directed protein evolution*. *Biocatal. Biotransform.* 25, 229-241
- 100 Durfee, T., Nelson, R., Baldwin, S., Plunkett, G., 3rd, Burland, V., Mau, B., Petrosino, J. F., Qin, X., Muzny, D. M., Ayele, M., Gibbs, R. A., Csorgo, B., Posfai, G., Weinstock, G. M. and Blattner, F. R. (2008) *The complete genome sequence of Escherichia coli DH10B: insights into the biology of a laboratory workhorse*. *J. Bacteriol.* 190, 2597-2606
- 101 Goldsmith, M. and Tawfik, D. S. (2012) *Directed enzyme evolution: beyond the low-hanging fruit*. *Curr. Opin. Struct. Biol.* 22, 406-412
- 102 Güven, G., Prodanovic, R. and Schwaneberg, U. (2010) *Protein Engineering – An Option for Enzymatic Biofuel Cell Design*. *Electroanalysis.* 22, 765 – 775
- 103 Reetz, M. T., Bocola, M., Carballeira, J. D., Zha, D. and Vogel, A. (2005) *Expanding the range of substrate acceptance of enzymes: combinatorial active-site saturation test*. *Angew. Chem., Int. Ed. Engl.* 44, 4192-4196
- 104 Reetz, M. T. and Carballeira, J. D. (2007) *Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes*. *Nat. Protoc.* 2, 891-903
- 105 Reetz, M. T., Kahakeaw, D. and Lohmer, R. (2008) *Addressing the numbers problem in directed evolution*. *ChemBioChem.* 9, 1797-1804
- 106 Reetz, M. T., Prasad, S., Carballeira, J. D., Gumulya, Y. and Bocola, M. (2010) *Iterative saturation mutagenesis accelerates laboratory evolution of enzyme stereoselectivity: rigorous comparison with traditional methods*. *J. Am. Chem. Soc.* 132, 9144-9152
- 107 Reetz, M. T., Wang, L. W. and Bocola, M. (2006) *Directed evolution of enantioselective enzymes: iterative cycles of CASTing for probing protein-sequence space*. *Angew. Chem., Int. Ed. Engl.* 45, 1236-1241
- 108 Chica, R. A., Doucet, N. and Pelletier, J. N. (2005) *Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design*. *Curr. Opin. Biotechnol.* 16, 378-384
- 109 Trott, O. and Olson, A. J. (2010) *AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading*. *J. Comput. Chem.* 31, 455-461
- 110 Brouk, M., Nov, Y. and Fishman, A. (2010) *Improving biocatalyst performance by integrating statistical methods into protein engineering*. *Appl. Environ. Microbiol.* 76, 6397-6403

- 111 Patrick, W. M., Firth, A. E. and Blackburn, J. M. (2003) *User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries*. *Protein Eng.* 16, 451-457
- 112 Rodriguez, R., Chinae, G., Lopez, N., Pons, T. and Vriend, G. (1998) *Homology modeling, model and software evaluation: three related resources*. *Bioinformatics.* 14, 523-528
- 113 Eckert, K. A. and Kunkel, T. A. (1990) *High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase*. *Nucleic Acids Res.* 18, 3739-3744
- 114 Lin-Goerke, J. L., Robbins, D. J. and Burczak, J. D. (1997) *PCR-based random mutagenesis using manganese and reduced dNTP concentration*. *BioTechniques.* 23, 409-412
- 115 Wong, T. S., Tee, K. L., Hauer, B. and Schwaneberg, U. (2004) *Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution*. *Nucleic Acids Res.* 32, e26
- 116 Hogrefe, H. H. and Cline, J. M. (2003) *Compositions and methods for random nucleic acid mutagenesis*. United States Patent US6803216
- 117 Caldwell, R. C. and Joyce, G. F. (1992) *Randomization of genes by PCR mutagenesis*. *PCR methods and applications.* 2, 28-33
- 118 Wong, T. S., Roccatano, D., Loakes, D., Tee, K. L., Schenk, A., Hauer, B. and Schwaneberg, U. (2008) *Transversion-enriched sequence saturation mutagenesis (SeSaM-Tv+): a random mutagenesis method with consecutive nucleotide exchanges that complements the bias of error-prone PCR*. *Biotechnol. J.* 3, 74-82
- 119 Wong, T. S., Tee, K. L., Hauer, B. and Schwaneberg, U. (2005) *Sequence saturation mutagenesis with tunable mutation frequencies*. *Anal. Biochem.* 341, 187-189
- 120 Wong, T. S., Zhurina, D. and Schwaneberg, U. (2006) *The diversity challenge in directed protein evolution*. *Comb. Chem. High Throughput Screening* 9, 271-288
- 121 Kolkman, J. A. and Stemmer, W. P. (2001) *Directed evolution of proteins by exon shuffling*. *Nat. Biotechnol.* 19, 423-428
- 122 Stemmer, W. P. (1994) *DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution*. *Proc. Natl. Acad. Sci. U. S. A.* 91, 10747-10751
- 123 Lawrence, J. G. (2005) *Horizontal and vertical gene transfer: the life history of pathogens*. *Contrib. Microbiol.* 12, 255-271
- 124 Marienhagen, J., Dennig, A. and Schwaneberg, U. (2012) *Phosphorothioate-based DNA recombination: an enzyme-free method for the combinatorial assembly of multiple DNA fragments*. *BioTechniques.* 0, 1-6
- 125 Hogrefe, H. H., Cline, J., Youngblood, G. L. and Allen, R. M. (2002) *Creating randomized amino acid libraries with the QuikChange Multi Site-Directed Mutagenesis Kit*. *BioTechniques.* 33, 1158-1160, 1162, 1164-1165
- 126 Matsumura, I. and Rowe, L. A. (2005) *Whole plasmid mutagenic PCR for directed protein evolution*. *Biomol. Eng.* 22, 73-79
- 127 Firth, A. E. and Patrick, W. M. (2008) *GLUE-IT and PEDEL-AA: new programmes for analyzing protein diversity in randomized libraries*. *Nucleic Acids Res.* 36, W281-285
- 128 Reetz, M. T., Carballeira, J. D. and Vogel, A. (2006) *Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability*. *Angew. Chem., Int. Ed. Engl.* 45, 7745-7751

- 129 Ruff, A. J., Dennig, A., Wirtz, G., Blanus, M. and Schwaneberg, U. (2012) *Flow cytometer-based high-throughput screening system for accelerated directed evolution of P450 monooxygenases*. ACS Catalysis. 2, 2724-2728
- 130 Link, A. J., Jeong, K. J. and Georgiou, G. (2007) *Beyond toothpicks: new methods for isolating mutant bacteria*. Nat. Rev. Microbiol. 5, 680-688
- 131 Turner, N. J. (2006) *Agar Plate-based Assays*. Enzyme Assays: High-throughput Screening. Genetic Selection and Fingerprinting. DOI: 10.1002/3527607846, 137-161
- 132 Griffiths, A. D. and Tawfik, D. S. (2006) *Miniaturising the laboratory in emulsion droplets*. Trends Biotechnol. 24, 395-402
- 133 Mayr, L. M. and Bojanic, D. (2009) *Novel trends in high-throughput screening*. Curr. Opin. Pharmacol. 9, 580-588
- 134 Agresti, J. J., Antipov, E., Abate, A. R., Ahn, K., Rowat, A. C., Baret, J.-C., Marquez, M., Klibanov, A. M., Griffiths, A. D. and Weitz, D. A. (2010) *Ultra-high-throughput screening in drop-based microfluidics for directed evolution*. Proc. Natl. Acad. Sci. U. S. A. 107, 4004-4009
- 135 Guo, M. T., Rotem, A., Heyman, J. A. and Weitz, D. A. (2012) *Droplet microfluidics for high-throughput biological assays*. Lab on a chip. 12, 2146-2155
- 136 Lee, C., Lee, J., Kim, H. H., Teh, S. Y., Lee, A., Chung, I. Y., Park, J. Y. and Shung, K. K. (2012) *Microfluidic droplet sorting with a high frequency ultrasound beam*. Lab on a chip. 12, 2736-2742
- 137 Taly, V., Kelly, B. T. and Griffiths, A. D. (2007) *Droplets as microreactors for high-throughput biology*. ChemBioChem. 8, 263-272
- 138 Schomburg, I., Chang, A., Ebeling, C., Gremse, M., Heldt, C., Huhn, G. and Schomburg, D. (2004) *BRENDA, the enzyme database: updates and major new developments*. Nucleic Acids Res. 32, D431-433
- 139 Bernhardt, R. (2006) *Cytochromes P450 as versatile biocatalysts*. J. Biotechnol. 124, 128-145
- 140 Nelson, D. R. (1998) *Cytochrome P450 nomenclature*. Methods. Mol. Biol. 107, 15-24
- 141 Omura, T. and Sato, R. (1964) *The Carbon Monoxide-Binding Pigment of Liver Microsomes. I. Evidence for Its Hemoprotein Nature*. J. Biol. Chem. 239, 2370-2378
- 142 Graham, S. E. and Peterson, J. A. (1999) *How similar are P450s and what can their differences teach us?* Arch. Biochem. Biophys. 369, 24-29
- 143 Guengerich, F. P. (2001) *Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity*. Chem. Res. Toxicol. 14, 611-650
- 144 Guengerich, F. P. and Munro, A. W. (2013) *Unusual cytochrome p450 enzymes and reactions*. J. Biol. Chem. 288, 17065-17073
- 145 Coelho, P. S., Brustad, E. M., Kannan, A. and Arnold, F. H. (2013) *Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes*. Science. 339, 307-310
- 146 Hollmann, F., Arends, I. W. C. E., Buehler, K., Schallmeyer, A. and Bühler, B. (2011) *Enzyme-mediated oxidations for the chemist*. Green Chem. 13, 226-265
- 147 Denisov, I. G., Makris, T. M., Sligar, S. G. and Schlichting, I. (2005) *Structure and chemistry of cytochrome P450*. Chem. Rev. 105, 2253-2277

- 148 Hannemann, F., Bichet, A., Ewen, K. M. and Bernhardt, R. (2007) *Cytochrome P450 systems-biological variations of electron transport chains*. *Biochim. Biophys. Acta.* 1770, 330-344
- 149 Narhi, L. O. and Fulco, A. J. (1987) *Identification and characterization of two functional domains in cytochrome P-450BM-3, a catalytically self-sufficient monooxygenase induced by barbiturates in Bacillus megaterium*. *J. Biol. Chem.* 262, 6683-6690
- 150 Miura, Y. and Fulco, A. J. (1974) *(Omega -2) hydroxylation of fatty acids by a soluble system from bacillus megaterium*. *J. Biol. Chem.* 249, 1880-1888
- 151 Whitehouse, C. J., Bell, S. G. and Wong, L. L. (2012) *P450(BM3) (CYP102A1): connecting the dots*. *Chem. Soc. Rev.* 41, 1218-1260
- 152 Furuya, T. and Kino, K. (2010) *Genome mining approach for the discovery of novel cytochrome P450 biocatalysts*. *Appl. Microbiol. Biotechnol.* 86, 991-1002
- 153 Dodhia, V. R., Fantuzzi, A. and Gilardi, G. (2006) *Engineering human cytochrome P450 enzymes into catalytically self-sufficient chimeras using molecular Lego*. *J. Biol. Inorg. Chem.* 11, 903-916
- 154 de Visser, S. P. and Shaik, S. (2003) *A proton-shuttle mechanism mediated by the porphyrin in benzene hydroxylation by cytochrome p450 enzymes*. *J. Am. Chem. Soc.* 125, 7413-7424
- 155 Omura, T. and Sato, R. (1964) *The Carbon Monoxide-Binding Pigment of Liver Microsomes. II. Solubilization, Purification, and Properties*. *J. Biol. Chem.* 239, 2379-2385
- 156 Whitehouse, C. J., Rees, N. H., Bell, S. G. and Wong, L. L. (2011) *Dearomatisation of o-xylene by P450BM3 (CYP102A1)*. *Chemistry (Weinheim an der Bergstrasse, Germany)*. 17, 6862-6868
- 157 Loida, P. J. and Sligar, S. G. (1993) *Molecular recognition in cytochrome P-450: mechanism for the control of uncoupling reactions*. *Biochemistry*. 32, 11530-11538
- 158 Munro, A. W., Leys, D. G., McLean, K. J., Marshall, K. R., Ost, T. W., Daff, S., Miles, C. S., Chapman, S. K., Lysek, D. A., Moser, C. C., Page, C. C. and Dutton, P. L. (2002) *P450 BM3: the very model of a modern flavocytochrome*. *Trends Biochem. Sci.* 27, 250-257
- 159 Paulsen, M. D. and Ornstein, R. L. (1992) *Predicting the product specificity and coupling of cytochrome P450cam*. *J. Comput.-Aided Mol. Des.* 6, 449-460
- 160 Nazor, J., Dannenmann, S., Adjei, R. O., Fordjour, Y. B., Ghampson, I. T., Blanusa, M., Roccatano, D. and Schwaneberg, U. (2008) *Laboratory evolution of P450 BM3 for mediated electron transfer yielding an activity-improved and reductase-independent variant*. *Protein Eng., Des. Sel.* 21, 29-35
- 161 Nazor, J. and Schwaneberg, U. (2006) *Laboratory evolution of P450 BM-3 for mediated electron transfer*. *ChemBioChem.* 7, 638-644
- 162 Zilly, F. E., Taglieber, A., Schulz, F., Hollmann, F. and Reetz, M. T. (2009) *Deazaflavins as mediators in light-driven cytochrome P450 catalyzed hydroxylations*. *Chem. Commun.*, 7152-7154
- 163 Wu, J. T., Wu, L. H. and Knight, J. A. (1986) *Stability of NADPH: effect of various factors on the kinetics of degradation*. *Clin. Chem.* 32, 314-319
- 164 Müller, C. A., Akkapurathu, B., Winkler, T., Staudt, S., Hummerl, W., Gröger, H. and Schwaneberg, U. (2013) *In vitro double oxidation of n-heptane with direct cofactor regeneration*. *Advanced Synthesis & Catalysis.* 355, 1787-1798
- 165 Zhang, J. D., Li, A. T., Yu, H. L., Imanaka, T. and Xu, J. H. (2011) *Synthesis of optically pure S-sulfoxide by Escherichia coli transformant cells coexpressing the P450 monooxygenase and glucose dehydrogenase genes*. *J. Ind. Microbiol. Biotechnol.* 38, 633-641

- 166 Girvan, H. M., Dunford, A. J., Neeli, R., Ekanem, I. S., Waltham, T. N., Joyce, M. G., Leys, D., Curtis, R. A., Williams, P., Fisher, K., Voice, M. W. and Munro, A. W. (2011) *Flavocytochrome P450 BM3 mutant W1046A is a NADH-dependent fatty acid hydroxylase: implications for the mechanism of electron transfer in the P450 BM3 dimer*. Arch. Biochem. Biophys. 507, 75-85
- 167 Dudek, H. M., Popken, P., van Bloois, E., Duetz, W. A. and Fraaije, M. W. (2013) *A Generic, Whole-Cell-Based Screening Method for Baeyer-Villiger Monooxygenases*. J. Biomol. Screen. 18, 678-687
- 168 Ryan, J. D., Fish, R. H. and Clark, D. S. (2008) *Engineering cytochrome P450 enzymes for improved activity towards biomimetic 1,4-NADH cofactors*. ChemBioChem. 9, 2579-2582
- 169 English, N., Hughes, V. and Wolf, C. R. (1994) *Common pathways of cytochrome P450 gene regulation by peroxisome proliferators and barbiturates in Bacillus megaterium ATCC14581*. J. Biol. Chem. 269, 26836-26841
- 170 English, N., Palmer, C. N., Alworth, W. L., Kang, L., Hughes, V. and Wolf, C. R. (1997) *Fatty acid signals in Bacillus megaterium are attenuated by cytochrome P-450-mediated hydroxylation*. Biochem. J. 327, 363-368
- 171 Sevrioukova, I. F. and Poulos, T. L. (2011) *Structural biology of redox partner interactions in P450cam monooxygenase: a fresh look at an old system*. Arch. Biochem. Biophys. 507, 66-74
- 172 Govindaraj, S. and Poulos, T. L. (1995) *Role of the linker region connecting the reductase and heme domains in cytochrome P450BM-3*. Biochemistry. 34, 11221-11226
- 173 Govindaraj, S. and Poulos, T. L. (1996) *Probing the structure of the linker connecting the reductase and heme domains of cytochrome P450BM-3 using site-directed mutagenesis*. Protein Sci. 5, 1389-1393
- 174 Neeli, R., Girvan, H. M., Lawrence, A., Warren, M. J., Leys, D., Scrutton, N. S. and Munro, A. W. (2005) *The dimeric form of flavocytochrome P450 BM3 is catalytically functional as a fatty acid hydroxylase*. FEBS Lett. 579, 5582-5588
- 175 Joyce, M. G., Ekanem, I. S., Roitel, O., Dunford, A. J., Neeli, R., Girvan, H. M., Baker, G. J., Curtis, R. A., Munro, A. W. and Leys, D. (2012) *The crystal structure of the FAD/NADPH-binding domain of flavocytochrome P450 BM3*. FEBS J. 279, 1694-1706
- 176 Zilly, F. E., Acevedo, J. P., Augustyniak, W., Deege, A., Hausig, U. W. and Reetz, M. T. (2011) *Tuning a p450 enzyme for methane oxidation*. Angew. Chem., Int. Ed. Engl. 50, 2720-2724
- 177 Kille, S., Zilly, F. E., Acevedo, J. P. and Reetz, M. T. (2011) *Regio- and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution*. Nat. Chem. 3, 738-743
- 178 Mehareenna, Y. T., Slessor, K. E., Cavaignac, S. M., Poulos, T. L. and De Voss, J. J. (2008) *The critical role of substrate-protein hydrogen bonding in the control of regioselective hydroxylation in p450cin*. J. Biol. Chem. 283, 10804-10812
- 179 Clark, J. P., Miles, C. S., Mowat, C. G., Walkinshaw, M. D., Reid, G. A., Daff, S. N. and Chapman, S. K. (2006) *The role of Thr268 and Phe393 in cytochrome P450 BM3*. J. Inorg. Biochem. 100, 1075-1090
- 180 Whitehouse, C. J., Yang, W., Yorke, J. A., Rowlatt, B. C., Strong, A. J., Blanford, C. F., Bell, S. G., Bartlam, M., Wong, L. L. and Rao, Z. (2010) *Structural basis for the properties of two single-site proline mutants of CYP102A1 (P450BM3)*. ChemBioChem. 11, 2549-2556
- 181 Noble, M. A., Quaroni, L., Chumanov, G. D., Turner, K. L., Chapman, S. K., Hanzlik, R. P. and Munro, A. W. (1998) *Imidazolyl carboxylic acids as mechanistic probes of flavocytochrome P-450 BM3*. Biochemistry. 37, 15799-15807

- 182 Wong, T. S., Arnold, F. H. and Schwaneberg, U. (2004) *Laboratory evolution of cytochrome p450 BM-3 monooxygenase for organic cosolvents*. *Biotechnol. Bioeng.* 85, 351-358
- 183 Shoji, O., Kunimatsu, T., Kawakami, N. and Watanabe, Y. (2013) *Highly Selective Hydroxylation of Benzene to Phenol by Wild-type Cytochrome P450BM3 Assisted by Decoy Molecules*. *Angew. Chem., Int. Ed. Engl.* 52, 6606-6610
- 184 Schwaneberg, U., Schmidt-Dannert, C., Schmitt, J. and Schmid, R. D. (1999) *A continuous spectrophotometric assay for P450 BM-3, a fatty acid hydroxylating enzyme, and its mutant F87A*. *Anal. Biochem.* 269, 359-366
- 185 Li, Q. S., Schwaneberg, U., Fischer, M., Schmitt, J., Pleiss, J., Lutz-Wahl, S. and Schmid, R. D. (2001) *Rational evolution of a medium chain-specific cytochrome P-450 BM-3 variant*. *Biochim. Biophys. Acta.* 1545, 114-121
- 186 Li, H. and Poulos, T. L. (1999) *Fatty acid metabolism, conformational change, and electron transfer in cytochrome P-450(BM-3)*. *Biochim. Biophys. Acta.* 1441, 141-149
- 187 Staudt, S., Burda, E., Giese, C., Muller, C. A., Marienhagen, J., Schwaneberg, U., Hummel, W., Drauz, K. and Groger, H. (2012) *Direct oxidation of cycloalkanes to cycloalkanones with oxygen in water*. *Angew. Chem., Int. Ed. Engl.* 52, 2359-2363
- 188 Chen, M. M., Snow, C. D., Vizcarra, C. L., Mayo, S. L. and Arnold, F. H. (2012) *Comparison of random mutagenesis and semi-rational designed libraries for improved cytochrome P450 BM3-catalyzed hydroxylation of small alkanes*. *Protein Eng., Des. Sel.* 25, 171-178
- 189 Appel, D., Lutz-Wahl, S., Fischer, P., Schwaneberg, U. and Schmid, R. D. (2001) *A P450 BM-3 mutant hydroxylates alkanes, cycloalkanes, arenes and heteroarenes*. *J. Biotechnol.* 88, 167-171
- 190 Shehzad, A., Panneerselvam, S., Linow, M., Bocola, M., Roccatano, D., Mueller-Dieckmann, J., Wilmanns, M. and Schwaneberg, U. (2013) *P450 BM3 crystal structures reveal the role of the charged surface residue Lys/Arg184 in inversion of enantioselective styrene epoxidation*. *Chem. Commun.* 49, 4694-4696
- 191 Tee, K. L. and Schwaneberg, U. (2006) *A screening system for the directed evolution of epoxygenases: importance of position 184 in P450 BM3 for stereoselective styrene epoxidation*. *Angew. Chem., Int. Ed. Engl.* 45, 5380-5383
- 192 Bhunia, S., Saha, D. and Koner, S. (2011) *MCM-41-supported oxo-vanadium(IV) complex: a highly selective heterogeneous catalyst for the bromination of hydroxy aromatic compounds in water*. *Langmuir.* 27, 15322-15329
- 193 Chen, A.-J., Wong, S.-T., Hwang, C.-C. and Mou, C.-Y. (2011) *Highly efficient and regioselective halogenation over well dispersed rhenium-promoted mesoporous zirconia*. *ACS Catalysis.* 1, 786-793
- 194 Deborde, M. and von Gunten, U. (2008) *Reactions of chlorine with inorganic and organic compounds during water treatment - Kinetics and mechanisms: a critical review*. *Water Res.* 42, 13-51
- 195 Duque, A. F., Hasan, S. A., Bessa, V. S., Carvalho, M. F., Samin, G., Janssen, D. B. and Castro, P. M. (2011) *Isolation and characterization of a Rhodococcus strain able to degrade 2-fluorophenol*. *Appl. Microbiol. Biotechnol.* 95, 511-520
- 196 Shehata, M., Durner, J., Thiessen, D., Shirin, M., Lottner, S., Van Landuyt, K., Furche, S., Hickel, R. and Reichl, F. X. (2012) *Induction of DNA double-strand breaks by monochlorophenol isomers and ChKM in human gingival fibroblasts*. *Arch. Toxicol.* 86, 1423-1429
- 197 Schmidt, R. J. (2004) *Industrial catalytic processes - phenol production*. *Appl. Catal., A.* 280, 89-103

- 198 Hock, H. and Lang, S. (1944) *Autoxydation von Kohlenwasserstoffen, IX. Mittel.: Über Peroxyde von Benzol-Derivaten*. Berichte der deutschen chemischen Gesellschaft. 77B, 257-264
- 199 Santoro, S. W., Wang, L., Herberich, B., King, D. S. and Schultz, P. G. (2002) *An efficient system for the evolution of aminoacyl-tRNA synthetase specificity*. Nat. Biotechnol. 20, 1044-1048
- 200 Cambie, R. C., Rutledge, P. S., Smith-Palmer, T. and Woodgate, P. D. (1976) *Selective Iodination of Phenols in the ortho-Position*. J. Chem. Soc., Perkin Trans. 1, 1161-1164
- 201 Zhang, P., Gong, Y., Li, H., Chen, Z. and Wang, Y. (2013) *Solvent-free aerobic oxidation of hydrocarbons and alcohols with Pd@N-doped carbon from glucose*. Nat. Commun. 4, 1593
- 202 Wierckx, N. J., Ballerstedt, H., de Bont, J. A. and Wery, J. (2005) *Engineering of solvent-tolerant Pseudomonas putida S12 for bioproduction of phenol from glucose*. Appl. Environ. Microbiol. 71, 8221-8227
- 203 Farinas, E. T., Alcalde, M. and Arnold, F. H. (2004) *Alkene epoxidation catalyzed by cytochrome P450 BM-3 139-3*. Tetrahedron. 60, 525-528
- 204 Tinberg, C. E., Song, W. J., Izzo, V. and Lippard, S. J. (2011) *Multiple roles of component proteins in bacterial multicomponent monooxygenases: phenol hydroxylase and toluene/o-xylene monooxygenase from Pseudomonas sp. OX1*. Biochemistry. 50, 1788-1798
- 205 Zhou, N. Y., Jenkins, A., Chan Kwo Chion, C. K. and Leak, D. J. (1999) *The alkene monooxygenase from Xanthobacter strain Py2 is closely related to aromatic monooxygenases and catalyzes aromatic monohydroxylation of benzene, toluene, and phenol*. Appl. Environ. Microbiol. 65, 1589-1595
- 206 Whitehouse, C. J., Bell, S. G., Tufton, H. G., Kenny, R. J., Ogilvie, L. C. and Wong, L. L. (2008) *Evolved CYP102A1 (P450BM3) variants oxidise a range of non-natural substrates and offer new selectivity options*. Chem. Commun., 966-968
- 207 Whitehouse, C. J. C., Bell, S. G., Yang, W., Yorke, J. A., Blanford, C. F., Strong, A. J. F., Morse, E. J., Bartlam, M., Rao, Z. and Wong, L.-L. (2009) *A highly active single-mutation variant of P450BM3 (CYP102A1)*. ChemBioChem. 10, 1654-1656
- 208 Carmichael, A. B. and Wong, L. L. (2001) *Protein engineering of Bacillus megaterium CYP102. The oxidation of polycyclic aromatic hydrocarbons*. Eur. J. Biochem. 268, 3117-3125
- 209 Rock, D. A., Boitano, A. E., Wahlstrom, J. L., Rock, D. A. and Jones, J. P. (2002) *Use of kinetic isotope effects to delineate the role of phenylalanine 87 in P450(BM-3)*. Bioorg. Chem. 30, 107-118
- 210 Noorlander, C. W., van Leeuwen, S. P., Te Biesebeek, J. D., Mengelers, M. J. and Zeilmaker, M. J. (2011) *Levels of perfluorinated compounds in food and dietary intake of PFOS and PFOA in the Netherlands*. J. Agric. Food Chem. 59, 7496-7505
- 211 Fang, X., Feng, Y., Shi, Z. and Dai, J. (2009) *Alterations of cytokines and MAPK signaling pathways are related to the immunotoxic effect of perfluorononanoic acid*. Toxicol. Sci. 108, 367-376
- 212 Upham, B. L., Deocampo, N. D., Wurl, B. and Trosko, J. E. (1998) *Inhibition of gap junctional intercellular communication by perfluorinated fatty acids is dependent on the chain length of the fluorinated tail*. Int. J. Cancer. 78, 491-495
- 213 Blanus, M., Schenk, A., Sadeghi, H., Marienhagen, J. and Schwaneberg, U. (2010) *Phosphorothioate-based ligase-independent gene cloning (PLICing): An enzyme-free and sequence-independent cloning method*. Anal. Biochem. 406, 141-146
- 214 Shivange, A. V., Serwe, A., Dennig, A., Roccatano, D., Haefner, S. and Schwaneberg, U. (2012) *Directed evolution of a highly active Yersinia mollaretii phytase*. Appl. Microbiol. Biotechnol. 95, 405-418

- 215 Zhang, Z., He, J., Yao, B., Zhou, Y., Chen, Y. and Yi, Y. (2001) *Production of phytase and acid phosphatase by use of silkworm-bioreactor*. Patent: China (01127288.0)-A 29-SEP-2001
- 216 Lassen, S. F., Sjoeholm, C. and Skov, L. K. (2008) *Hafnia Phytase*. United States Patent Application 20080263688
- 217 Hanahan, D., Jessee, J. and Bloom, F. R. (1991) *Plasmid transformation of Escherichia coli and other bacteria*. Methods in enzymology. 204, 63-113
- 218 Inoue, H., Nojima, H. and Okayama, H. (1990) *High efficiency transformation of Escherichia coli with plasmids*. Gene. 96, 23-28
- 219 Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors*. Gene. 33, 103-119
- 220 Cline, J., Braman, J. C. and Hogrefe, H. H. (1996) *PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases*. Nucleic Acids Res. 24, 3546-3551
- 221 Owczarzy, R., Tataurov, A. V., Wu, Y., Manthey, J. A., McQuisten, K. A., Almabrazi, H. G., Pedersen, K. F., Lin, Y., Garretson, J., McEntaggart, N. O., Sailor, C. A., Dawson, R. B. and Peek, A. S. (2008) *IDT SciTools: a suite for analysis and design of nucleic acid oligomers*. Nucleic Acids Res. 36, W163-169
- 222 Eckstein, F. and Gish, G. (1989) *Phosphorothioates in molecular biology*. Trends Biochem. Sci. 14, 97-100
- 223 Dennig, A., Shivange, A. V., Marienhagen, J. and Schwaneberg, U. (2011) *OmniChange: the sequence independent method for simultaneous site-saturation of five codons*. PLoS One. 6, e26222
- 224 Siguret, V., Ribba, A. S., Cherel, G., Meyer, D. and Pietu, G. (1994) *Effect of plasmid size on transformation efficiency by electroporation of Escherichia coli DH5 alpha*. BioTechniques. 16, 422-426
- 225 Miyazaki, K. (2003) *Creating random mutagenesis libraries by megaprimer PCR of whole plasmid (MEGAWHOP)*. Methods. Mol. Biol. 231, 23-28
- 226 Kirsch, R. D. and Joly, E. (1998) *An improved PCR-mutagenesis strategy for two-site mutagenesis or sequence swapping between related genes*. Nucleic Acids Res. 26, 1848-1850
- 227 Glieder, A., Farinas, E. T. and Arnold, F. H. (2002) *Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase*. Nat. Biotechnol. 20, 1135-1139
- 228 Wong, T. S., Wu, N., Roccatano, D., Zacharias, M. and Schwaneberg, U. (2005) *Sensitive assay for laboratory evolution of hydroxylases toward aromatic and heterocyclic compounds*. J. Biomol. Screening. 10, 246-252
- 229 Glieder, A. and Meinhold, P. (2003) *High-throughput screens based on NAD(P)H depletion*. Methods. Mol. Biol. 230, 157-170
- 230 Emerson, E. (1943) *The condensation of 4-aminoantipyrine. Part II. A new color test of phenolic compounds*. J. Org. Chem. 8, 417-428
- 231 Arain, S., John, G. T., Krause, C., Gerlach, J., Wolfbeis, O. S. and Klimant, I. (2006) *Characterization of microtiterplates with integrated optical sensors for oxygen and pH, and their applications to enzyme activity screening, respirometry, and toxicological assays*. Sensors and Actuators B: Chemical. 113, 639-648
- 232 Schwaneberg, U., Sprauer, A., Schmidt-Dannert, C. and Schmid, R. D. (1999) *P450 monooxygenase in biotechnology. I. Single-step, large-scale purification method for cytochrome P450 BM-3 by anion-exchange chromatography*. J. Chromatogr., A. 848, 149-159

- 233 Laemmli, U. K. (1970) *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*. *Nature*. 227, 680-685
- 234 Bradford, M. M. (1976) *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding*. *Anal. Biochem.* 72, 248-254
- 235 Cirino, P. and Arnold, F. H. (2002) *Regioselectivity and Activity of Cytochrome P450 BM-3 and Mutant F87A in Reactions Driven by Hydrogen Peroxide*. *Advanced Synthesis & Catalysis*. 344, 932-937
- 236 Michaelis, L., Menten, M. L., Johnson, K. A. and Goody, R. S. (2012) *The original Michaelis constant: translation of the 1913 Michaelis-Menten paper*. *Biochemistry*. 50, 8264-8269
- 237 Krieger, E., Koraimann, G. and Vriend, G. (2002) *Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field*. *Proteins*. 47, 393-402
- 238 Sevrioukova, I. F., Li, H., Zhang, H., Peterson, J. A. and Poulos, T. L. (1999) *Structure of a cytochrome P450-redox partner electron-transfer complex*. *Proc. Natl. Acad. Sci. U. S. A.* 96, 1863-1868
- 239 Iyer, K. R., Jones, J. P., Darbyshire, J. F. and Trager, W. F. (1997) *Intramolecular isotope effects for benzylic hydroxylation of isomeric xylenes and 4,4'-dimethylbiphenyl by cytochrome P450: relationship between distance of methyl groups and masking of the intrinsic isotope effect*. *Biochemistry*. 36, 7136-7143
- 240 Mitchell, K. H., Rogge, C. E., Gierahn, T. and Fox, B. G. (2003) *Insight into the mechanism of aromatic hydroxylation by toluene 4-monooxygenase by use of specifically deuterated toluene and p-xylene*. *Proc. Natl. Acad. Sci. U. S. A.* 100, 3784-3789
- 241 Pikus, J. D., Studts, J. M., McClay, K., Steffan, R. J. and Fox, B. G. (1997) *Changes in the regiospecificity of aromatic hydroxylation produced by active site engineering in the diiron enzyme toluene 4-monooxygenase*. *Biochemistry*. 36, 9283-9289
- 242 Kobayashi, S. and Higashimura, H. (2003) *Oxidative polymerization of phenols revised*. *Prog. Polym. Sci.* 28, 1015-1048
- 243 Shibasaki, Y., Hoshi, K., Suzuki, E., Shiraishi, Y., Norisue, Y. and Oishi, Y. (2009) *Oxidative Coupling Copolycondensation of 2,6-Dimethylphenol with 2,5-Dimethylphenol: Highly Thermostable Poly(phenylene ether)*. *Polym. J.* 41, 1136-1143
- 244 Netscher, T. (2007) *Synthesis of vitamin E*. *Vitam. Horm.* 76, 155-202
- 245 Roy, A. and Pahan, K. (2009) *Gemfibrozil, stretching arms beyond lipid lowering*. *Immunopharmacol. Immunotoxicol.* 31, 339-351
- 246 Whitehouse, C. J., Rees, N. H., Bell, S. G. and Wong, L. L. (2011) *Dearomatisation of o-xylene by P450BM3 (CYP102A1)*. *Chemistry (Weinheim an der Bergstraße, Germany)*. 17, 6862-6868
- 247 Morawski, B., Lin, Z., Cirino, P., Joo, H., Bandara, G. and Arnold, F. H. (2000) *Functional expression of horseradish peroxidase in Saccharomyces cerevisiae and Pichia pastoris*. *Protein Eng.* 13, 377-384
- 248 Miles, C. S., Ost, T. W., Noble, M. A., Munro, A. W. and Chapman, S. K. (2000) *Protein engineering of cytochromes P-450*. *Biochim. Biophys. Acta.* 1543, 383-407
- 249 Harrelson, J. P., Atkins, W. M. and Nelson, S. D. (2008) *Multiple-ligand binding in CYP2A6: probing mechanisms of cytochrome P450 cooperativity by assessing substrate dynamics*. *Biochemistry*. 47, 2978-2988

- 250 Henne, A., Schmitz, R. A., Bömeke, M., Gottschalk, G. and Daniel, R. (2000) *Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on Escherichia coli*. Appl. Environ. Microbiol. 66, 3113-3116
- 251 Kanaya, S., Koyanagi, T. and Kanaya, E. (1998) *An esterase from Escherichia coli with a sequence similarity to hormone-sensitive lipase*. Biochem. J. 332 (Pt 1), 75-80
- 252 Zhang, N., Suen, W. C., Windsor, W., Xiao, L., Madison, V. and Zaks, A. (2003) *Improving tolerance of Candida antarctica lipase B towards irreversible thermal inactivation through directed evolution*. Protein Eng. 16, 599-605
- 253 Sehgal, A. C., Tompson, R., Cavanagh, J. and Kelly, R. M. (2002) *Structural and catalytic response to temperature and cosolvents of carboxylesterase EST1 from the extremely thermoacidophilic archaeon Sulfolobus solfataricus P1*. Biotechnol. Bioeng. 80, 784-793
- 254 Dominguez de Maria, P. and Maugeri, Z. (2011) *Ionic liquids in biotransformations: from proof-of-concept to emerging deep-eutectic-solvents*. Curr. Opin. Chem. Biol. 15, 220-225
- 255 Roccatano, D., Wong, T. S., Schwaneberg, U. and Zacharias, M. (2006) *Toward understanding the inactivation mechanism of monooxygenase P450 BM-3 by organic cosolvents: a molecular dynamics simulation study*. Biopolymers. 83, 467-476
- 256 Kuper, J., Wong, T. S., Roccatano, D., Wilmanns, M. and Schwaneberg, U. (2007) *Understanding a mechanism of organic cosolvent inactivation in heme monooxygenase P450 BM-3*. J. Am. Chem. Soc. 129, 5786-5787
- 257 Whitehouse, C. J., Bell, S. G. and Wong, L. L. (2008) *Desaturation of alkylbenzenes by cytochrome P450(BM3) (CYP102A1)*. Chemistry (Weinheim an der Bergstrasse, Germany). 14, 10905-10908
- 258 Campanelli, A. R., Domenicano, A., Ramondo, F. and Hargittai, I. (2004) *Group electronegativities from benzene ring deformations: A quantum chemical study*. J. Phys. Chem. A. 108, 4940-4948
- 259 Salonen, L. M., Ellermann, M. and Diederich, F. (2011) *Aromatic rings in chemical and biological recognition: energetics and structures*. Angew. Chem., Int. Ed. Engl. 50, 4808-4842
- 260 Fernandez, A. (2002) *Desolvation shell of hydrogen bonds in folded proteins, protein complexes and folding pathways*. FEBS Lett. 527, 166-170
- 261 Fernandez, A. and Berry, R. S. (2003) *Proteins with H-bond packing defects are highly interactive with lipid bilayers: Implications for amyloidogenesis*. Proc. Natl. Acad. Sci. U. S. A. 100, 2391-2396
- 262 Quideau, S. (2011) *Organic chemistry: Triumph for unnatural synthesis*. Nature. 474, 459-460
- 263 Rayne, S., Forest, K. and Friesen, K. J. (2009) *Mechanistic aspects regarding the direct aqueous environmental photochemistry of phenol and its simple halogenated derivatives. A review*. Environ. Int. 35, 425-437
- 264 Vione, D., Maurino, V., Minero, C., Calza, P. and Pelizzetti, E. (2005) *Phenol chlorination and photochlorination in the presence of chloride ions in homogeneous aqueous solution*. Environ. Sci. Technol. 39, 5066-5075
- 265 Galano, A., Leon-Carmona, J. R. and Alvarez-Idaboy, J. R. (2012) *Influence of the Environment on the Protective Effects of Guaiacol Derivatives against Oxidative Stress: Mechanisms, Kinetics, and Relative Antioxidant Activity*. J. Phys. Chem. B
- 266 Talawar, M. B., Jyothi, T. M., Sawant, P. D., Raja, T. and Rao, B. S. (2000) *Calcined Mg-Al hydrotalcite as an efficient catalyst for the synthesis of guaiacol*. Green Chem. 2, 266-268

- 267 Mageroy, M. H., Tieman, D. M., Floystad, A., Taylor, M. G. and Klee, H. J. (2012) *A Solanum lycopersicum catechol-O-methyltransferase involved in synthesis of the flavor molecule guaiacol*. Plant J. 69, 1043-1051
- 268 (A server providing information on substrate solubilities in aqueous solution: <http://www.pharmacy.arizona.edu/outreach/aquasol/>)
- 269 Rietjens, I. M., Soffers, A. E., Veeger, C. and Vervoort, J. (1993) *Regioselectivity of cytochrome P-450 catalyzed hydroxylation of fluorobenzenes predicted by calculated frontier orbital substrate characteristics*. Biochemistry. 32, 4801-4812
- 270 Wu, L. L., Yang, C. L., Lo, F. C., Chiang, C. H., Chang, C. W., Ng, K. Y., Chou, H. H., Hung, H. Y., Chan, S. I. and Yu, S. S. (2012) *Tuning the regio- and stereoselectivity of C-H activation in n-octanes by cytochrome P450 BM-3 with fluorine substituents: evidence for interactions between a C-F bond and aromatic pi systems*. Chemistry (Weinheim an der Bergstrasse, Germany). 17, 4774-4787
- 271 Breslow, R. (1980) *Biomimetic Control of Chemical Selectivity*. Acc. Chem. Res. 13, 170-177
- 272 Ost, T. W., Clark, J., Mowat, C. G., Miles, C. S., Walkinshaw, M. D., Reid, G. A., Chapman, S. K. and Daff, S. (2003) *Oxygen activation and electron transfer in flavocytochrome P450 BM3*. J. Am. Chem. Soc. 125, 15010-15020
- 273 Meunier, B., de Visser, S. P. and Shaik, S. (2004) *Mechanism of oxidation reactions catalyzed by cytochrome p450 enzymes*. Chem. Rev. 104, 3947-3980
- 274 Feichtinger, D. and Plattner, D. A. (2001) *Probing the reactivity of oxomanganese-salen complexes: an electrospray tandem mass spectrometric study of highly reactive intermediates*. Chemistry (Weinheim an der Bergstrasse, Germany). 7, 591-599
- 275 Gilchrist, A. and Sutton, L. E. (1952) *Relations between length and reactivity in some carbon-halogen bonds*. J. Phys. Chem. 56, 319-321
- 276 Lippert, J. L., Hanna, M. W. and Trotter, P. J. (1969) *Bonding in donor-acceptor complexes. III. Relative contributions of electrostatic, charge-transfer, and exchange interactions in aromatic halogen and aromatic-TCNE [tetracyanoethylene] complexes*. J. Am. Chem. Soc. 91, 4035-4044
- 277 Sun, H., Harms, K. and Sundermeyer, J. (2004) *Aerobic oxidation of 2,3,6-trimethylphenol to trimethyl-1,4-benzoquinone with copper(II) chloride as catalyst in ionic liquid and structure of the active species*. J. Am. Chem. Soc. 126, 9550-9551
- 278 Prokofieva, A., Prikhod'ko, A. I., Dechert, S. and Meyer, F. (2008) *Selective benzylic C-C coupling catalyzed by a bioinspired dicopper complex*. Chem. Commun., 1005-1007
- 279 Pohl, M., Lingen, B. and Muller, M. (2002) *Thiamin-diphosphate-dependent enzymes: new aspects of asymmetric C-C bond formation*. Chemistry (Weinheim an der Bergstrasse, Germany). 8, 5288-5295
- 280 Leach, B. E. (1980) *Process for the synthesis of 2,6-xylenol and 2,3,6-trimethylphenol*. Patentnumber US4283574
- 281 Hill, A. D. and Reilly, P. J. (2008) *A Gibbs free energy correlation for automated docking of carbohydrates*. J. Comput. Chem. 29, 1131-1141
- 282 Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F. and Serrano, L. (2005) *The FoldX web server: an online force field*. Nucleic Acids Res. 33, W382-388
- 283 Nanda, V. and Koder, R. L. (2010) *Designing artificial enzymes by intuition and computation*. Nat. Chem. 2, 15-24
- 284 Lambrianou, A., Demin, S. and Hall, E. A. (2008) *Protein engineering and electrochemical biosensors*. Adv. Biochem. Eng. Biotechnol. 109, 65-96

- 285 Podust, L. M., Ouellet, H., von Kries, J. P. and de Montellano, P. R. (2009) *Interaction of Mycobacterium tuberculosis CYP130 with heterocyclic arylamines*. J. Biol. Chem. 284, 25211-25219
- 286 Korzekwa, K. R., Krishnamachary, N., Shou, M., Ogai, A., Parise, R. A., Rettie, A. E., Gonzalez, F. J. and Tracy, T. S. (1998) *Evaluation of atypical cytochrome P450 kinetics with two-substrate models: evidence that multiple substrates can simultaneously bind to cytochrome P450 active sites*. Biochemistry. 37, 4137-4147
- 287 Williams, P. A., Cosme, J., Vinkovic, D. M., Ward, A., Angove, H. C., Day, P. J., Vornrhein, C., Tickle, I. J. and Jhoti, H. (2004) *Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone*. Science. 305, 683-686
- 288 Modi, S., Sutcliffe, M. J., Primrose, W. U., Lian, L. Y. and Roberts, G. C. (1996) *The catalytic mechanism of cytochrome P450 BM3 involves a 6 Å movement of the bound substrate on reduction*. Nat. Struct. Biol. 3, 414-417
- 289 Wireduaah, S., Parker, T. M. and Lewis, M. (2013) *Effects of the Aromatic Substitution Pattern in Cation- π Sandwich Complexes*. J. Phys. Chem. A. 117, 2598-2604
- 290 Carballeira, J. D., Quezada, M. A., Hoyos, P., Simeo, Y., Hernaiz, M. J., Alcantara, A. R. and Sinisterra, J. V. (2009) *Microbial cells as catalysts for stereoselective red-ox reactions*. Biotechnol. Adv. 27, 686-714
- 291 Sikkema, J., de Bont, J. A. and Poolman, B. (1994) *Interactions of cyclic hydrocarbons with biological membranes*. J. Biol. Chem. 269, 8022-8028
- 292 Julsing, M. K., Kuhn, D., Schmid, A. and Buhler, B. (2012) *Resting cells of recombinant E. coli show high epoxidation yields on energy source and high sensitivity to product inhibition*. Biotechnol. Bioeng. 109, 1109-1119
- 293 Bloom, J. D., Labthavikul, S. T., Otey, C. R. and Arnold, F. H. (2006) *Protein stability promotes evolvability*. Proc. Natl. Acad. Sci. U. S. A. 103, 5869-5874
- 294 Fasan, R., Chen, M. M., Crook, N. C. and Arnold, F. H. (2007) *Engineered alkane-hydroxylating cytochrome P450(BM3) exhibiting nativelike catalytic properties*. Angew. Chem., Int. Ed. Engl. 46, 8414-8418
- 295 Munro, A. W., Leys, D. G., McLean, K. J., Marshall, K. R., Ost, T. W., Daff, S., Miles, C. S., Chapman, S. K., Lysek, D. A., Moser, C. C., Page, C. C. and Dutton, P. L. (2002) *P450 BM3: the very model of a modern flavocytochrome*. Trends Biochem. Sci. 27, 250-257
- 296 Sulistyaningdyah, W. T., Ogawa, J., Li, Q. S., Maeda, C., Yano, Y., Schmid, R. D. and Shimizu, S. (2005) *Hydroxylation activity of P450 BM-3 mutant F87V towards aromatic compounds and its application to the synthesis of hydroquinone derivatives from phenolic compounds*. Appl. Microbiol. Biotechnol. 67, 556-562
- 297 Carrea, G. and Riva, S. (2000) *Properties and Synthetic Applications of Enzymes in Organic Solvents*. Angew. Chem., Int. Ed. Engl. 39, 2226-2254
- 298 Robinson, R. A. (1967) *The Dissociation Constants of Some Disubstituted Anilines and Phenols in Aqueous Solution at 25°C*. J. Res. Natl. Inst. Stand. Technol. 71, 213-218
- 299 Gardossi, L., Poulsen, P. B., Ballesteros, A., Hult, K., Svedas, V. K., Vasic-Racki, D., Carrea, G., Magnusson, A., Schmid, A., Wohlgemuth, R. and Halling, P. J. (2010) *Guidelines for reporting of biocatalytic reactions*. Trends Biotechnol. 28, 171-180
- 300 Ross, L., Barclay, C. and Vinqvist, M. R. (2003) *The Chemistry of Phenols - Part 2 (Ed. Zvi Rappoport)*. Book Chapter, 874
- 301 Ran, N., Knop, D. R., Draths, K. M. and Frost, J. W. (2001) *Benzene-free synthesis of hydroquinone*. J. Am. Chem. Soc. 123, 10927-10934

- 302 Yoshikawa, A., Yoshida, S. and Terao, I. (1993) *Microbial production of hydroquinone*. *Bioprocess Technol.* 16, 149-161
- 303 Kubota, A., Maeda, T., Nagafuchi, N., Kadokami, K. and Ogawa, H. I. (2008) *TNT biodegradation and production of dihydroxylamino-nitrotoluene by aerobic TNT degrader Pseudomonas sp. strain TM15 in an anoxic environment*. *Biodegradation.* 19, 795-805
- 304 Borch, T., Inskeep, W. P., Harwood, J. A. and Gerlach, R. (2005) *Impact of ferrihydrite and anthraquinone-2,6-disulfonate on the reductive transformation of 2,4,6-trinitrotoluene by a gram-positive fermenting bacterium*. *Environ. Sci. Technol.* 39, 7126-7133
- 305 Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., Djoumbou, Y., Eisner, R., Guo, A. C. and Wishart, D. S. (2011) *DrugBank 3.0: a comprehensive resource for 'omics' research on drugs*. *Nucleic Acids Res.* 39, D1035-1041
- 306 Lang-Feulner, J. and Rau, W. (1975) *Redox dyes as artificial photoreceptors in carotenoid synthesis*. *Photochem. Photobiol.* 21, 179-183
- 307 Meyer, D., Witholt, B. and Schmid, A. (2005) *Suitability of recombinant Escherichia coli and Pseudomonas putida strains for selective biotransformation of m-nitrotoluene by xylene monooxygenase*. *Appl. Environ. Microbiol.* 71, 6624-6632
- 308 Hartman, W. W. and Rahrs, E. J. (1955) *Organic Syntheses*, Coll. 3, 650
- 309 Hartman, W. W. and Rahrs, E. J. (1955) *Organic Syntheses*, Coll. 3, 652
- 310 Kinne, M., Zeisig, C., Ullrich, R., Kayser, G., Hammel, K. E. and Hofrichter, M. (2010) *Stepwise oxygenations of toluene and 4-nitrotoluene by a fungal peroxygenase*. *Biochem. Biophys. Res. Commun.* 397, 18-21
- 311 Soojhawon, I., Lokhande, P. D., Kodam, K. M. and Gawai, K. R. (2005) *Biotransformation of nitroaromatics and their effects on mixed function oxidase system*. *Enzyme Microb. Technol.* 37, 527-533
- 312 Kirby, P. J., Roberts, J. A. and Nocera, D. G. (1997) *Significant Effect of Salt Bridges on Electron Transfer*. *J. Am. Chem. Soc.* 119, 9230-9236
- 313 Fertinger, C. (2011) *Activation of Peroxides by Model Systems for Cytochrome P450*. Dissertation (<http://www.opus.ub.uni-erlangen.de/opus/volltexte/2011/2807/pdf/ChristophFertingerDissertation.pdf>)
- 314 Hamamatsu, N., Nomiya, Y., Aita, T., Nakajima, M., Husimi, Y. and Shibana, Y. (2006) *Directed evolution by accumulating tailored mutations: thermostabilization of lactate oxidase with less trade-off with catalytic activity*. *Protein Eng., Des. Sel.* 19, 483-489
- 315 Jochens, H. and Bornscheuer, U. T. (2010) *Natural diversity to guide focused directed evolution*. *ChemBioChem.* 11, 1861-1866
- 316 Brownie, J., Shawcross, S., Theaker, J., Whitcombe, D., Ferrie, R., Newton, C. and Little, S. (1997) *The elimination of primer-dimer accumulation in PCR*. *Nucleic Acids Res.* 25, 3235-3241
- 317 Herman, A. and Tawfik, D. S. (2007) *Incorporating Synthetic Oligonucleotides via Gene Reassembly (ISOR): a versatile tool for generating targeted libraries*. *Protein Eng., Des. Sel.* 20, 219-226
- 318 Peng, R. H., Xiong, A. S. and Yao, Q. H. (2006) *A direct and efficient PAGE-mediated overlap extension PCR method for gene multiple-site mutagenesis*. *Appl. Microbiol. Biotechnol.* 73, 234-240
- 319 Seyfang, A. and Jin, J. H. (2004) *Multiple site-directed mutagenesis of more than 10 sites simultaneously and in a single round*. *Anal. Biochem.* 324, 285-291

- 320 Shen, Z., Qu, W., Wang, W., Lu, Y., Wu, Y., Li, Z., Hang, X., Wang, X., Zhao, D. and Zhang, C. (2010) *MPprimer: a program for reliable multiplex PCR primer design*. BMC bioinformatics. 11, 143
- 321 Tabuchi, I., Soramoto, S., Ueno, S. and Husimi, Y. (2004) *Multi-line split DNA synthesis: a novel combinatorial method to make high quality peptide libraries*. BMC Biotechnol. 4, 19
- 322 Wang, H. H. (2010) *Synthetic genomes for synthetic biology*. J. Mol. Cell. Biol. 2, 178-179
- 323 Vroom, J. A. and Wang, C. L. (2008) *Modular construction of plasmids through ligation-free assembly of vector components with oligonucleotide linkers*. BioTechniques. 44, 924-926
- 324 Liang, J., Luo, Y. and Zhao, H. (2011) *Synthetic biology: putting synthesis into biology*. Wiley interdisciplinary reviews. 3, 7-20
- 325 Bershtein, S. and Tawfik, D. S. (2008) *Advances in laboratory evolution of enzymes*. Curr. Opin. Chem. Biol. 12, 151-158
- 326 Arango Gutierrez, E., Mundhada, H., Meier, H., Duefel, H., Bocola, M. and Schwaneberg, U. (2013) *Reengineered glucose oxidase for amperometric glucose determination in diabetes analytics*. Biosensors and Bioelectronics. 50, 84-90
- 327 Zhao, H., Giver, L., Shao, Z., Affholter, J. A. and Arnold, F. H. (1998) *Molecular evolution by staggered extension process (StEP) in vitro recombination*. Nat. Biotechnol. 16, 258-261
- 328 Coco, W. M., Levinson, W. E., Crist, M. J., Hektor, H. J., Darzins, A., Pienkos, P. T., Squires, C. H. and Monticello, D. J. (2001) *DNA shuffling method for generating highly recombined genes and evolved enzymes*. Nat. Biotechnol. 19, 354-359
- 329 Ostermeier, M., Nixon, A. E., Shim, J. H. and Benkovic, S. J. (1999) *Combinatorial protein engineering by incremental truncation*. Proc. Natl. Acad. Sci. U. S. A. 96, 3562-3567
- 330 Lutz, S., Ostermeier, M., Moore, G. L., Maranas, C. D. and Benkovic, S. J. (2001) *Creating multiple-crossover DNA libraries independent of sequence identity*. Proc. Natl. Acad. Sci. U. S. A. 98, 11248-11253
- 331 Sieber, V., Martinez, C. A. and Arnold, F. H. (2001) *Libraries of hybrid proteins from distantly related sequences*. Nat. Biotechnol. 19, 456-460
- 332 Hiraga, K. and Arnold, F. H. (2003) *General method for sequence-independent site-directed chimeragenesis*. J. Mol. Biol. 330, 287-296
- 333 Ostermeier, M. and Benkovic, S. J. (2000) *Evolution of protein function by domain swapping*. Adv. Protein. Chem. 55, 29-77
- 334 Wang, Q., Wu, H., Wang, A., Du, P., Pei, X., Li, H., Yin, X., Huang, L. and Xiong, X. (2010) *Prospecting metagenomic enzyme subfamily genes for DNA family shuffling by a novel PCR-based approach*. J. Biol. Chem. 285, 41509-41516
- 335 Villiers, B. R., Stein, V. and Hollfelder, F. (2010) *USER friendly DNA recombination (USERec): a simple and flexible near homology-independent method for gene library construction*. Protein Eng., Des. Sel. 23, 1-8
- 336 Engler, C., Gruetzner, R., Kandzia, R. and Marillonnet, S. (2009) *Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes*. PloS one. 4, e5553
- 337 Bikard, D., Julie-Galau, S., Cambray, G. and Mazel, D. (2010) *The synthetic integron: an in vivo genetic shuffling device*. Nucleic Acids Res. 38, e153
- 338 Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. and Pease, L. R. (1989) *Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension*. Gene. 77, 61-68

- 339 Bendtsen, J. D., Nielsen, H., von Heijne, G. and Brunak, S. (2004) *Improved prediction of signal peptides: SignalP 3.0*. J. Mol. Biol. 340, 783-795

7. Further Contributions to Scientific Publications

Directed evolution of a highly active *Yersinia mollaretii* phytase.

Shivange, A. V.; Serwe, A.; **Dennig, A.**; Roccatano, D.; Haefner, S.; Schwaneberg, U., Directed evolution of a highly active *Yersinia mollaretii* phytase. *Appl. Microbiol. Biotechnol.* **2012**, 95, (2), 405-18.

Abstract

Phytase improves as a feed supplement the nutritional quality of phytate-rich diets (e.g., cereal grains, legumes, and oilseeds) by hydrolyzing indigestible phytate (myo-inositol 1,2,3,4,5,6-hexakis dihydrogen phosphate) and increasing abdominal absorption of inorganic phosphates, minerals, and trace elements. Directed phytase evolution was reported for improving industrial relevant properties such as thermostability (pelleting process) or activity. In this study, we report the cloning, characterization, and directed evolution of the *Yersinia mollaretii* phytase (YmPhytase). YmPhytase has a tetrameric structure with positive cooperativity (Hill coefficient was 2.3) and a specific activity of 1,073 U/mg which is ~10 times higher than widely used fungal phytases. High-throughput prescreening methods using filter papers or 384-well microtiter plates were developed. Precise subsequent screening for thermostable and active phytase variants was performed by combining absorbance and fluorescence-based detection system in 96-well microtiter plates. Directed evolution yielded after mutant library generation (SeSaM method) and two-step screening (in total ~8,400 clones) a phytase variant with ~20% improved thermostability (58°C for 20 min; residual activity wild type ~34%; variant ~53%) and increased melting temperature (1.5°C) with a slight loss of specific activity (993 U/mg).

Flow Cytometer-Based High-Throughput Screening System for Accelerated Directed Evolution of P450 Monooxygenases.

Ruff, A. J.; **Dennig, A.**; Wirtz, G.; Blanusa, M.; Schwaneberg, U., Flow cytometer-based high-throughput screening system for accelerated directed evolution of P450 monooxygenases. *ACS Catalysis* **2012**, 2, (12), 2724-2728.

Abstract

Flow cytometry-based screening systems have successfully been used in directed evolution experiments. Herein, we report the first whole-cell, high-throughput screening platform for P450 monooxygenases based on flow cytometry. O-dealkylation of 7-benzyloxy-3-carboxycoumarin ethyl ester (BCCE) by P450 BM3 generates a fluorescence coumarin derivative. After one round of directed evolution, P450 BM3 variants with up to 7-fold increased activity (P450 M3 DM-1: R255H) could be identified at a sampling rate of 500 events s⁻¹. The reported screening platform can likely be applied to directed evolution campaigns of any P450 monooxygenase that catalyzes the O-dealkylation of BCCE.

Extending the substrate scope of a Baeyer-Villinger monooxygenase by multiple-site mutagenesis.

Dudek, H. M.; Fink, M.; Shivange, A. V.; **Dennig, A.**; Mihovilovic, M. D.; Schwaneberg, U.; Fraaije, M. W., Engineering PAMO towards activity on alicyclic ketones. *Appl. Microbiol. Biotechnol.* **2013**, [Epub ahead of print].

The Sequence Saturation Mutagenesis (SeSaM) method (Book Chapter)

Ruff, A. J., Kardashliev, T., **Dennig, A.**; Schwaneberg, U., The Sequence Saturation Mutagenesis (SeSaM) method. *Submitted in 2012*.

8. Acknowledgements

This work was accomplished within the EU funded 7th Framework project OXYGREEN (FP7-KBBE; Project Reference: 212281) that is acknowledged for financial support. BASF SE (Ludwigshafen, Germany) is acknowledged for financial support during the development of OmniChange and PTRec.

I would like to acknowledge the unlimited and at any time available support from Prof. Dr. Ulrich Schwaneberg who gave me the opportunity to perform research in his excellently equipped institute. After many hours of discussing research, preparing manuscripts (sometimes late after work) I am deeply thankful for his scientific guidance over the last three years.

I would like to thank Prof. Dr. Lars M. Blank for being the second referee and PhD defense committee member. Prof. Dr. Anett Schallmey is acknowledged for being the third examiner in the committee as well as for her critical and valuable suggestions regarding biocatalytic conversions with P450 monooxygenases. Prof. Dr. Schäffer is acknowledged for being the head of the examination committee.

All members of the OXYGREEN project are acknowledged for their valuable feedback on results and scientific input during occasional meetings.

Dr. Jan Marienhagen is acknowledged for the guidance at the beginning of my PhD work and his honesty and support also during critical situations. Furthermore I would like to thank him for giving me the opportunity to do also research at the Forschungszentrum in Jülich (IBG-1).

Dr. André Jakoblinert is acknowledged not only for the great fun in the lab and office (wall of fame) but also sparing time outside of work (DoKo). For his explanations on deciphering the (last) mystery of enzyme kinetics namely the *Benny Hill* Coefficient, I am highly thankful.

Dr. Anna Joelle Ruff is acknowledged for her invaluable support during the OXYGREEN project, having always time for discussions and to reach scientific goals as a team.

Dr. Haifeng Liu is acknowledged for spending time for discussions (mainly in front of the fume hood), suggestions on new substrates, analytical challenges and giving feedback from an organic chemist's point of view.

Tsvetan Kardashliev is acknowledged for generating the homology model structures of P450 BM3 variants used in this thesis as well as for the docking studies of substrates. In addition we shared some really good times in the lab that was making the boring parts bearable without losing focus.

Paula Bracco is acknowledged for discussions on P450s and biocatalytic conversions as well as the support for everything concerning GC analytics.

Especially I would like to thank Paula, Ljubica, Tsvetan, Haifeng, Joelle, André and Ljubica for reading through my thesis and giving me honest feedback.

Namely I would like to thank Josiane, Erik, Christian, Marcus and Paula sharing three years in Aachen in and outside of the lab.

A lot of work wouldn't have been accomplished without the support of master, bachelor and research students. In this regard, I would like to thank Lukas Guddat, Thomas Pankau, Nina Lültsdorf, Marco Grull, Moritz Gauer, Bastian Zirpel and Kübra Kömürçü for their help in the lab.

The remaining people from the Schwaneberg group (past and present) that were not named in person are also acknowledged for their support during the last three years.

I would like to thank my closest friends for their patience and support in alphabetical order: Benny, Carl, Christian, Florian, Hendrik and Michael. Despite that most of the last three years I spent working on this thesis, I never forget what is important in life and I deeply miss the times we had.

I would like to thank my whole family, my mother, my father, Jens, Carmen und Isabel for their invaluable support.

At last but not least I would like to thank Christina (*Chiquita*) to share two exciting but also very stressful years at my site. I am looking forward to the next years to share with you.

Statement

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass alle Stellen der Arbeit, die wörtlich oder sinngemäß aus anderen Quellen übernommen wurden, als solche kenntlich gemacht sind und dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt wurde.

Aachen, den

Alexander Dennig

Lebenslauf

Persönliches

Name: Alexander Dennig

Geburtsdatum: 02.04.1982

Geburtsort: Mülheim an der Ruhr

Staatsangehörigkeit: deutsch

Ausbildung

2010 –2013

Promotion an der RWTH Aachen University, Lehrstuhl für Biotechnologie

2006 –2010

RWTH Aachen University

Diplomarbeitsthema: ‘Directed Evolution of Phosphatase for Increased Thermostability’

Abschluss: Diplom Biologe

2003 –2006

Johannes Gutenberg Universität in Mainz

Vordiplom Biologie

2001 –2002

Wehrdienst: Bundeswehr - Standortsanitätszentrum Düsseldorf

1991 –2001

Städtisches Gymnasium an der Grashofstraße in Essen

Allgemeine Hochschulreife (Abitur)