Discovery and Protein Engineering of Baeyer-Villiger monooxygenases

Inauguraldissertation

zur

Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Ernst-Moritz-Arndt-Universität Greifswald

vorgelegt von
Andy Beier
geboren am 11.10.1988
in Parchim

Greifswald, den 02.08.2017

Dekan: Prof. Dr. Werner Weitschies

1. Gutachter: Prof. Dr. Uwe T. Bornscheuer

2. Gutachter: Prof. Dr. Marko Mihovilovic

Tag der Promotion: 24.10.2017

We need to learn to want what we have,
not to have what we want,
in order to get stable and steady
happiness.

- The Dalai Lama -

List of abbreviations

%	Percent	MPS	Methyl phenyl sulfide
% (v/v)	% volume per volume	MPSO	Methyl phenyl sulfoxide
% (w/v)	% weight per volume	$MPSO_2$	Methyl phenyl sulfone
°C	Degrees Celsius	MTS	Methyl *p*-tolyl sulfide
µM	µmol/L	MTSO	Methyl *p*-tolyl sulfoxide
aa	Amino acids	$MTSO_2$	Methyl *p*-tolyl sulfone
AGE	Agarose gel electrophoresis	NAD^+	Nicotinamide adenine dinucleotide, oxidized
aq. dest.	Distilled water	NADH	Nicotinamide adenine dinucleotide, reduced
BLAST	Basic Local Alignment Search Tool	$NADP^+$	Nicotinamide adenine dinucleotide phosphate, oxidized
bp	Base pair(s)	NADPH	Nicotinamide adenine dinucleotide phosphate, reduced
BVMO	Baeyer-Villiger monooxygenase	OD_{600}	Optical density at 600 nm
CHMO	Cyclohexanone monooxygenase	PAGE	Polyacrylamide gel electrophoresis
Da	Dalton	PAMO	Phenylacetone monooxygenase
DMF	Dimethyl formamide	PCR	Polymerase chain reaction
DMSO	Dimethyl sulfoxide	PDB	Protein Data Bank
$DMSO_2$	Dimethyl sulfone	rpm	Revolutions per minute
DNA	Desoxyribonucleic acid	rv	Reverse
dNTP	Desoxynucleoside triphosphate	SDS	Sodium dodecyl sulfate
E. coli	*Escherichia coli*	SOC	Super Optimal broth with Catabolite repression
ee	Enantiomeric excess	TAE	TRIS-Acetate-EDTA
FAD	Flavin adenine dinucleotide	TB	Terrific broth
Fig.	Figure	TCE	2,2,2-Trichloroethanol
FMN	Flavin adenine mononucleotide	TCEP	tris(2-carboxyethyl)phosphine
FMO	Flavoprotein monooxygenase	TEMED	Tetramethylethylenediamine
fw	Forward	TRIS	Tris(hydroxymethyl)aminomethane
GC	Gas chromatography	UV	Ultraviolet
h	Hours	x g	Times gravity of Earth
HAPMO	4-Hydroxyacetophenone monooxygenase		
$His_{(6)}$	hexahistidine tag		
IPTG	Isopropyl β-D-1-thiogalactopyranoside		
L	Liter		
LB	Lysogenic broth		
M	mol/L		
min	Minutes		

Furthermore, SI units (base, derived and prefixes) and the common notation for amino acids and nucleic acids are used.

1 U is defined as the amount of enzyme that catalyzes the depletion of 1 µmol NADPH per minute in the NADPH depletion assay.

Table of contents

1 Introduction

1.1 White Biotechnology

Chemical processes generally have to be carried out under harsh conditions and by employing organic solvents.[1] In the course of this, toxic side products are frequently formed, which can be hazardous for the environment and humans. Additionally, the chemical production of complex compounds is challenging as numerous steps of synthesis and laborious strategies applying protective groups are necessary and enantioselective reactions are hard to achieve. The White Biotechnology is a highly topical field for science and industry.[2] Here, microorganisms or enzymes are employed as biocatalysts for chemical processes (Figure 1.1).

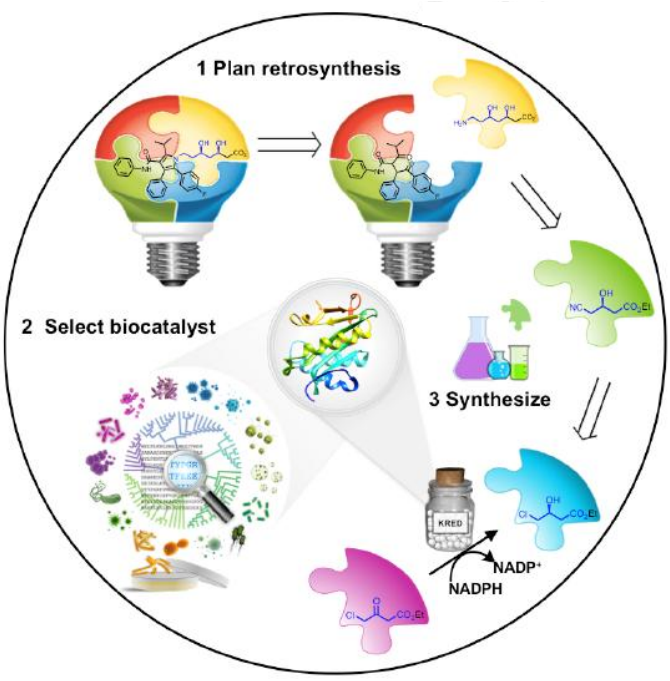


Figure 1.1: General steps to apply a biocatalyst for the synthesis of a desired compound (from Rodrigo *et al*).[3]
First, the strategy for the (retro-)synthesis needs to be planned. Then, a suitable biocatalyst from the available set able to catalyze the necessary reaction(s) is selected. Finally, the reaction is performed under environmentally friendly conditions leading to the pure product without side products.

This is an environmentally friendly and energy saving alternative for the classical organic synthesis as it can be carried out under mild conditions without the addition of toxic and expensive organic solvents. Furthermore, a high substrate specificity and the often found outstanding stereo-, chemo- and regioselectivities lead to not just higher yields of the desired products but also to a minimization or even exclusion of the formation of toxic side products.[2a, 2c, 4] Therefore, this field, also designated Industrial Biotechnology, provides an important contribution to the conservation of the environment and simplifies the establishment of the Green Chemistry.

1.2 Protein Engineering

Even though enzymes are remarkable biocatalysts the naturally found representatives (wild types) often are not suited for large-scale industrial processes. A variety of properties are important for a biocatalytic process (Figure 1.2).

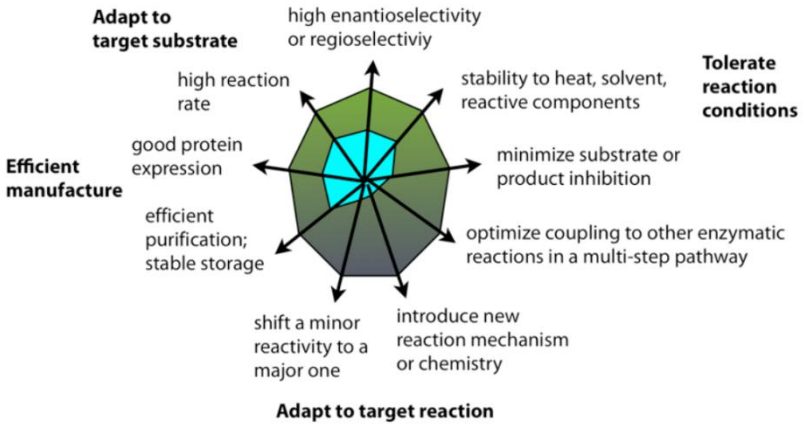


Figure 1.2: Exemplary profile of decisive properties of a biocatalyst (from Bornscheuer and Kazlauskas).[5]
The performance for every property is displayed making a comparison and development of a multi-parameter fingerprint possible.

However, with the combined tools of molecular biology and bioinformatics, optimization of enzymes on a molecular level became possible to meet the requirements of industrial applications.[6] This approach, called protein engineering, was successfully applied to alter almost every possible property such as stability, activity and substrate scope of numerous enzymes over the last years.[7] There are two general strategies for performing protein engineering – rational design and directed evolution.[6, 7b-d, 8]

For the rational design of proteins, as much information about the protein (structural and mechanistical) as possible is needed.[7d] Optimal is a 3D crystal structure with high resolution containing a substrate and possible cofactors to estimate which residues form the active site. Alternatively, it is possible to generate a homology model of a protein – a predicted structure oriented on known structures homologous to the protein.[9] The three dimensional structure is utilized to predict the most suitable mutations to alter the protein structure and consequently modify certain properties of the enzyme. Advancing progress in (bio)informatics is constantly making predictions by molecular modeling for this approach to be more precise. However, we are still far away from entirely understanding structure-function relationships of proteins and not every result from changing the amino acid sequence of a protein can be well predicted. Nevertheless, many successful rational designs of enzymes have been reported.[10] In general, amino acid residues likely determining the property to be altered need to be identified. These

are then exchanged with different amino acids predicted to alter the property into the desired direction by applying site-directed mutagenesis. The different enzyme variants are subsequently produced and examined for different characteristics (Figure 1.3).[7d]

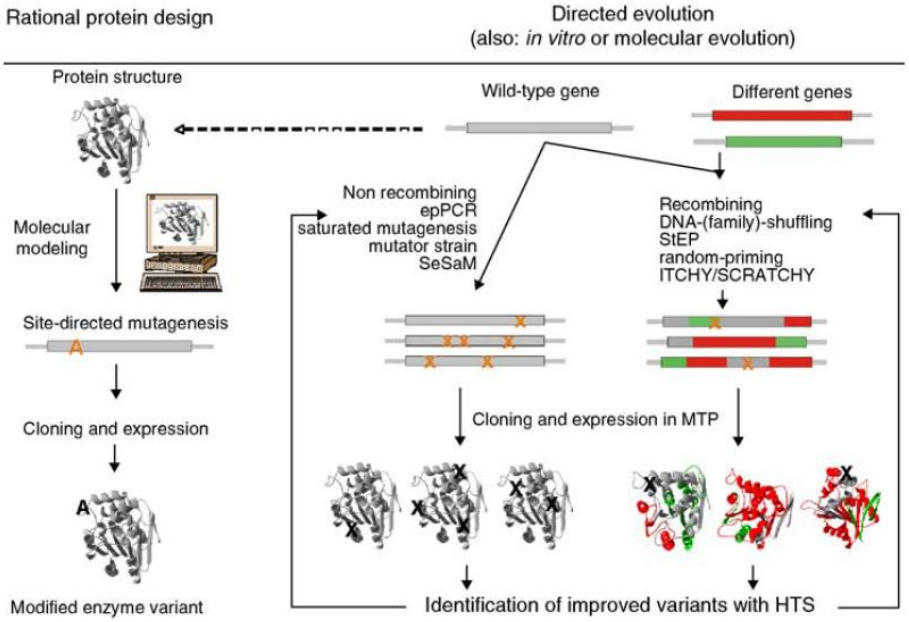


Figure 1.3: Prinicples of protein engineering via either rational design (left) or directed evolution (right) (from Bornscheuer).[8]
Rational design: On basis of the protein structure (or homology model) property determining residues are identified, then changed by site-directed mutagenesis and finally expressed and checked for altered characteristics; Directed evolution: The gene encoding for the enzyme to be evolved towards an improved property is randomly mutated by either non-recombining or recombining methods. The generated mutant libraries are produced recombinantly and are subsequently screened with high-throughput systems (both commonly in microtiter plates, MTPs).

In contrast, employing directed evolution means to randomly change the proteins by mutagenesis without using structural information.[7d] Consequently, mutant libraries of a huge numbers of enzyme variants are generated. This means that the necessary screening effort to find improved mutants is very high (Table 1.1).

Table 1.1: Number of possible variants of a protein consisting of 200 amino acids as a result of one to four simultaneous random mutations (from Bornscheuer).[7d]

Mutations (M)	**Number of variants (sequence length N=200)**
1	3800
2	7183900
3	9008610600
4	8429807368950

Number of possible variants of a protein by introduction of M substitutions in N amino acids = 19M[N!/(N-M)!M!]

Enhanced variants subsequently undergo further rounds of mutagenesis and screening until the desired degree of improvement is achieved. For using directed evolution, the gene sequence of the enzyme, efficient methods for mutagenesis and high-throughput screening are needed (Figure 1.3).[7a, 7d, 8] If applied successfully, every desired protein property can be optimized, for example the substrate range or the stereoselectivity.[7a, 11]

It is also possible to combine both approaches leading to a strategy called focused directed evolution or semi-rational design.[7d] Here, the randomization process is focused on a few areas or residues of the protein.[12] This compensates for missing information about the enzyme and less computational prediction is needed. Additionally, focused directed evolution drastically decreases the screening effort as the library size becomes much smaller since only parts of the protein undergo random mutagenesis (Figure 1.4).

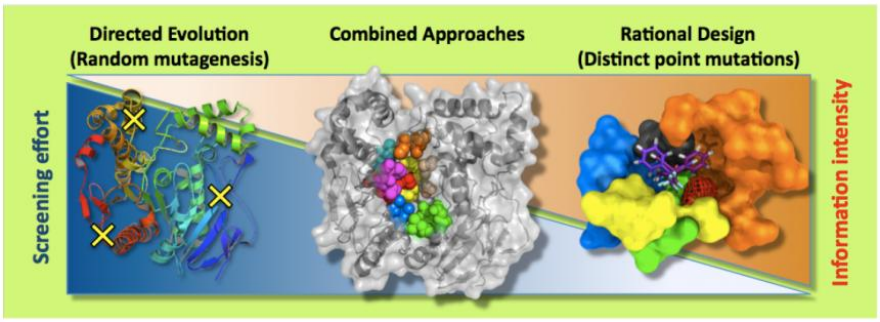


Figure 1.4: **Comparison of the protein engineering approaches directed evolution, rational design and their combination, the focused directed evolution (from Lutz and Bornscheuer).[7b]** While for directed evolution, the highest screening effort is necessary and for rational design comprehensive information about the protein structure is crucial, the semi-rational design offers a compromise of the two strategies.

Methods pursuing this approach like Combinatorial Active-site Saturation Test (CAST), Iterative Saturation Mutagenesis (ISM), Protein Sequence Activity Relationships (ProSAR) and the "consensus approach" succeeded to improve a variety of biocatalysts.[13] The concept of the latter strategy is to compare a set of enzymes homologous to the one to be evolved to identify conserved and deviating residues. Then, the amino acid exchanges at these positions are planned accordingly.[14] With a similar approach using the database 3DM, hundreds of structures (and subsequently sequences) can be aligned to find residues differing in the protein of interest and identify the most common amino acids for respective positions. Resulting "small, but smart" libraries already served to improve properties like enantioselectivity, activity and thermoactivity of enzymes.[15]

In all three approaches not only one or more amino acids can be substituted for others, also additional residues can be inserted and deleted or the location of the N- and C-termini can be altered.[7c, 16] However, protein engineering, regardless of the strategy, is not trivial.[7d]

Even if an improvement of the observed property of an enzyme is achieved, various characteristics are important for its practical application.[7d] One needs to be aware that another trait can be negatively affected in the course of the structural changes as well.[7d, 17]

It cannot be said, that one of the described strategies is better than the other.[7c] Each has its benefits and drawbacks and for every protein the best solution has to be figured out anew, also depending on the possibilities in the respective laboratory. In many cases more than one approach will meet the desired goal.[7c] Sufficient knowledge of the proteins and sophisticated analysis of all existing strategies and their outcomes will speed up protein engineering by identifying the way that gives maximum success with minimum effort.[7c] According to Kazlauskas and Bornscheuer, "protein engineering will move towards rational design" as it "allows one to reach the goal with the least effort".[7c]

1.3 Flavin-dependent monooxygenases

Flavin-dependent monooxygenases catalyze a variety of reactions. Heteroatom oxidations (i.e. sulfoxidations), Baeyer-Villiger oxidations, hydroxylations, epoxidations and halogenations are the activities found in this class of enzymes that are performed with often high regio- and stereoselectivities.[18] A non-covalently bound FAD or FMN is needed as cofactor from which the isoalloxazine moiety is firstly reduced and subsequently oxygenated then acting as the reactive species (Scheme 1.1).

Scheme 1.1: Reaction mechanism of flavin-dependent monooxygenases (adapted from van Berkel *et al.*). X: Substrate; XO: Product.[19]

Flavin is reduced by NAD(P)H and subsequently reacts with molecular oxygen forming the reactive C4α-(hydro)-peroxyflavin. Whether the peroxyflavin is protonated or not determines if a nucleophilic or electrophilic attack takes place, respectively.[18-20] Hereby, one oxygen atom is transferred to the substrate while the remaining is eliminated as water so that the flavin is recovered to its oxidized status.[18-20]

The class of flavin-dependent monooxygenases consists of eight groups, designated A-H (Table 1.2).[19]

Table 1.2: Groups of flavin-dependent monooxygenases (adapted from Van Berkel *et al*. and Huijbers *et al*.).[18-19]

Group	Cofactor	Electron donor	Fold	Reaction examples	Enzymes
A	FAD	NAD(P)H	Rossmann (1)	Hydroxylation, epoxidation	*p*-Hydroxybenzoate-hydroxylase
				Sulfoxidation	MICAL
B	FAD	NAD(P)H	Rossmann (2)	Baeyer-Villiger oxidation	Cyclohexanone monooxygenase
				Heteroatom oxygenation	Dimethylanillin monooxygenase
				N-Hydroxylation	L-Ornithine monooxygenase
				Oxidative decarboxylation	Indole-3-pyruvate monooxygenase
C	FMN	$FMNH_2$	TIM barrel	Light emission	Luciferase
				Baeyer-Villiger oxidation, epoxidation	Diketocamphane monooxygenase
				Desulfurization, sulfoxidation	Alkanesulfonate monooxygenase
				Hydroxylation	Long-chain alkane monooxygenase
D	FAD/FMN	$FADH_2$/$FMNH_2$	Acyl-CoA dehydrogenase	Hydroxylation	*p*-Hydroxyphenylacetate-3-hydroxylase
				N-Hydroxylation	KijD3 sugar N-oxygenase
E	FAD	$FADH_2$	Rossmann (1)	Epoxidation	Styrene monooxygenase
F	FAD	$FADH_2$	Rossmann (1)	Halogenation	Tryptophan 7-halogenase
G	FAD	Substrate	Rossmann (1)	Oxidative decarboxylation	Tryptophan 2-monooxygenase
H	FMN	Substrate	TIM barrel	Oxidative decarboxylation	Lactate 2-monooxygenase
				Oxidative denitration	Nitronate monooxygenase

Numbers in parentheses indicate the number of Rossmann folds; MICAL: molecule interacting with CasL.

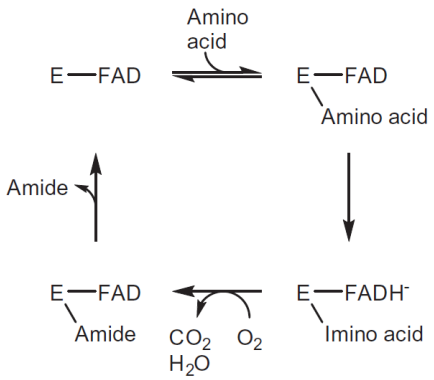
This distribution is based on structural features, protein sequence motifs, electron donor and type of reaction of the respective enzymes.[19] It includes external flavin-dependent monooxygenases which are dependent from an external electron donor (mostly NAD(P)H) and internal monooxygenases utilizing the substrate as electron donor. Enzymes from group A are participating in the microbial degradation of (poly)aromatic substances and are applied for the biosynthesis of natural products (Table 1.3).[18, 21]

Table 1.3: Examples for group A flavin monooxygenases involved in the biosynthesis of natural products (adapted from Huijbers *et al.*).[18]

Natural product	Enzyme	Reference
Violacein	Prodeoxyviolacein hydroxylase	Balibar *et al.*[22]
Pyocyanin	Phenazine-1-carboxylate hydroxylase	Greenhagen *et al.*[23]; Mavrodi *et al.*[24]
Staurosporine	7-Carboxy-K252C hydroxylase	Goldmann *et al.*[25]
Tetracycline	Tetracycline hydroxylase	Volkers *et al.*[26]; Walkiewicz *et al.*[27]
Aurachin B	Aurachin C monooxygenase	Katsuyama *et al.*[28]
Fumiquinazoline C	Fumiquinazoline F monooxygenase	Ames *et al.*[29]; Gao *et al.*[30]
Asperlicin	Asperlicin C monooxygenase	Haynes *et al.*[31]
Asukamycin	Protoasukamycin hydroxylase AsuE1	Rui *et al.*[32]
Xiamycin	Xiamycin monooxygenase	Baunach *et al.*[33]

Group B contains four subgroups: N-hydroxylating monooxygenases (NHMOs), YUCCAs, flavoprotein monooxygenases (FMOs) and Baeyer-Villiger monooxygenases (BVMOs).[18, 34] NHMOs are only found in bacteria and fungi. They are involved in the biosynthesis of siderophores by N-hydroxylation of amine groups of L-lysine, L-ornithine, cadaverine, putrescine and 1,3-diaminopropane.[18] All plants contain YUCCAs where they are involved in the biosynthesis of auxin, the primary development regulator and growth hormone of plants.[18] FMOs are found in all kingdoms of life but only a few have been investigated to date.[18, 34a] FMOs from mammals use the electrophilic C4α-hydroperoxyflavin for the oxygenation of a variety of sulfurs, halides and carbon-bound nucleophilic nitrogens.[18, 35] The natural function of these enzymes is to simplify the detoxification of xenobiotics by their conversion to the respective more hydrophilic products.[19, 36] BVMOs are referred to in detail in the following section. For group C only twelve examples are known.[18] These catalyze hydroxylations, epoxidations, sulfoxidations, desulfonations (a Baeyer-Villiger oxidation) as well as light emission (aldehyde oxidation by luciferase).[18] Enzymes in group D perform aromatic hydroxylations or N-hydroxylations.[18] Monooxygenases in group E just include styrene monooxygenases (SMOs) performing epoxidations to form (S)-styrene oxides out of the respective styrene derivatives.[18, 37] Group F flavin monooxygenases regioselectively chlorinate and brominate activated organic molecules, which is used for the synthesis of anti-tumor agents, antibiotics and other natural products.[18]

The four representatives of group G use an amino acid substrate as electron donor and thus are internal monooxygenases.[18, 38] The reductive half-reaction of their reaction mechanism probably involves the cleavage of the α-CH bond of the amino acid (Scheme 1.2).



Scheme 1.2: Proposed reaction mechanism of group G flavin-dependent monooxygenases (from Huijbers *et al.*).[18, 38]
E: Enzyme.

Subsequently, a hydride equivalent is transferred to FAD leading to the generation of the enzyme-bound imino acid.[38] Then, in the oxidative half-reaction, the formed imino acid is converted to an amide in an oxygenative decarboxylation. Lactate 2-monooxygenase (LMO) and nitronate monooxygenase (NMO) are the two known members of the last group H.[18] Like enzymes in group G they utilize a substrate to reduce their flavin cofactor. LMO oxidizes L-lactate leading to the formation of acetate, carbon dioxide and water. NMO employs molecular oxygen for the oxidation of propionate 3-nitronate, being toxic for the metabolism, and different alkyl nitronates to the respective semialdehydes and nitrite.[18]

1.3.1 Baeyer-Villiger-Monooxygenases

Baeyer-Villiger monooxygenases (BVMOs, EC 1.14.13.x) are a class of flavin-dependent monooxygenases that catalyze the oxidation of ketones to esters or lactones similar to the chemical Baeyer-Villiger oxidation reported in 1899 (Scheme 1.3).[39]

Scheme 1.3: Mechanism of the Baeyer-Villiger oxidation.[39]

Here, a peracid reacts with the carbonyl carbon of a ketone in a nucleophilic attack forming a tetrahedral Criegee intermediate.[40] The higher substituted carbon atom migrates to an oxygen of the peracid leading to the so called Criegee rearrangement.[40] Consequently, the corresponding ester is formed. Especially the required use of an unstable and expensive peracid makes the enzymatic approach more attractive.[41] BVMOs offer an environmentally friendly alternative as these enzymes just need molecular oxygen, a flavin cofactor (FMN or FAD), which as peroxyflavin fulfills the role of the peracid, and a nicotinamide cofactor (NADPH or NADH) for their activity just producing water as a side product.[42] Furthermore, these enzymes can display a high enantio- and regioselectivity.[42b]

1.3.1.1 Mechanistic aspects and enzyme properties

The mechanism of the enzymatic Baeyer-Villiger oxidation, postulated in 1982 by Reyerson *et al*., was demonstrated by investigation of the cyclohexanone monooxygenase from *Acinetobacter* sp. NCIMB 9871 ($CHMO_{Acineto}$).[43] After different crystal structures of the CHMO from *Rhodococcus* sp. HI-31 ($CHMO_{Rhodo}$) were solved, the understanding of how the reaction takes place in the enzyme was dramatically increased (Figure 1.5).[44]

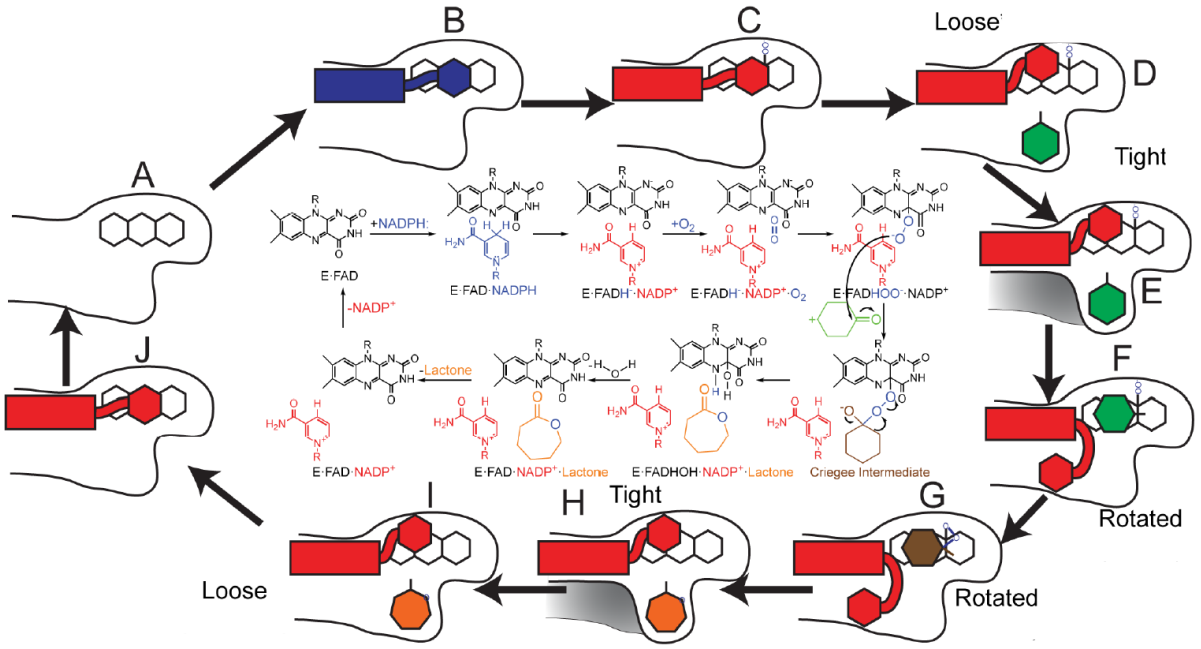


Figure 1.5: Mechanism of the enzymatic Baeyer-Villiger oxidation (from Yachnin *et al.*).[43-44]

Throughout the whole catalytic cycle, the enzyme is in a dynamic process and undergoes a number of conformational changes. It starts with a reductive half reaction in the "open" conformation (PDB code: 3GWF), where the oxidized flavin is reduced by NADPH (steps A-C). Next, the oxidative half reaction takes place in which the reduced flavin molecule forms a covalent adduct with molecular oxygen resulting in a peroxyflavin (steps C-D). Then, the ketone substrate (in this case cyclohexanone) enters the active site just before the BVMO switches to the "closed" conformation (PDB code: 3GWD) to bind it, reflected in the "loose" conformation (PDB code: 4RG4, step D).[44a, 44b] Subsequently, the enzyme switches into the "tight" conformation (PDB code: 4RG3), where substrate acceptance and stereospecificity is determined (step E).[44b] As a consequence of another conformational change in which the nicotinamide moiety of the flavin is rotated, leading to the "rotated" conformation (PDB code: 3UCL), a nucleophilic attack of the peroxyflavin towards the carbonyl carbon leads to the tetrahedral Criegee intermediate (steps F-G). As a consequence of the rearrangement, the corresponding ester or lactone (in this case ε-caprolactone) as well as the hydroxyflavin are formed (step H). The oxidized flavin is subsequently regenerated by an elimination of water, followed by the release of the product through of a series of domain movements between the "tight", "closed", "loose" and "open" conformations (steps H-A).[44a] The "loose" conformation also represents the state upon release of the product.[44b]

To date, in addition to the five crystal structures of $CHMO_{Rhodo}$, in total 24 structures of nine further BVMOs are available as well.[44-45] Sixteen are from bacterial type I Baeyer-Villiger monooxygenases – five from the phenylacetone monooxygenase (PAMO) from *Thermobifida*

fusca (PDB codes: 1W4X, 2YLR, 2YLT, 2YLS, 4OVI), six from the 2-oxo-Δ^3-4.5.5-trimethylcyclopentenylacetyl-coenzyme A monooxygenase (OTEMO) from *Pseudomonas putida* ATCC 17453 (PDB codes: 3UP5, 3UP4, 3UOZ, 3UOY, 3UOX, 3UOV), four from the steroid monooxygenase (STMO) from *Rhodococcus rhodochrous* (PDB codes: 4AP3, 4AP1, 4AOX, 4AOS) and one from the CHMO from *Thermocrispum municipal* (TmCHMO, PDB code: 5M10).[45b-d, 45g, 45k] Two eukaryotic type I BVMO structure were solved in 2016 for $BVMO_{AFL838}$ from *Aspergillus flavus* (PDB code: 5J7X) and in 2017 for the Polycyclic ketone monooxygenase from *Thermothelomyces thermophila* (PDB code: 5MQ6).[45a, 45h] Two other crystal structures were determined for the 3.6-Diketocamphane-1.6-monooxygenase (3.6-DKCMO) from *P. putida* ATCC 17453, a type II BVMO (PDB codes: 4UWM, 5AEC).[45e, 45i] The crystallization of the type III BVMO *Stenotrophomonas maltophilia* flavoprotein monooxygenases (SMFMO) was successful as well (PDB code: 4A9W).[46] Lastly, the structure (in three different forms), which resembles Group A flavin-dependent monooxygenases, belongs to the atypical type 0 BVMO MtmOIV from *Streptomyces argillaceus* (PDB codes: 3FMW, 4K5R, 4K5S).[45f, 45j]

It is worth emphasizing that in the course of the enzymatic Baeyer-Villiger oxidation, if the ketone is asymmetric, both possible regioisomers can be formed (Scheme 1.4).

Scheme 1.4: Examplary formation of both possible regioisomers from hexanone in course of the enzymatic Baeyer-Villiger oxidation.

The “normal” ester is produced by migration of the more nucleophilic and higher substituted carbon atom next to the carbonyl carbon, while the “abnormal” product is formed by rearrangement of the less stabilized and thus less favored residue. The occurrence of both reactions can be explained by a different positioning of the substrate in the active site of the enzyme. This leads to different orientations of the Criegee intermediate in the course of the reaction as well. The residue of the Criegee intermediate that is then arranged antiperiplanar to the oxygen−oxygen bond of the peroxide group of the peroxyflavin is the migrating one.[47] The possible formation of the “abnormal” product constitutes a big difference to the chemical counterpart, in which the rearrangement of the higher substituted carbon is always preferred.[41, 48] However, the regioselectivity often is not perfect and the ratio “normal”/”abnormal” ester/lactone of the resulting product mixture is different for every

substrate and again for each BVMO.[47a] This was demonstrated on aliphatic compounds like β-aminoketones and β-hydroxyketones, but mostly fused bicyclic ketones were used as model substrates.[45c, 47b, 49] For instance, $CHMO_{Acineto}$ formed the "normal" product out of (–)-carvomenthone and (–)-trans-dihydrocarvone while the "abnormal" ones were obtained from the cyclopentanone monooxygenase (CPMO).[50] It has also been shown that the regioselectivity of Baeyer-Villiger monooxygenases can be changed by semirational protein design.[47a] For the cyclohexanone monooxygenase from *Arthrobacter* sp. a complete switch in regioselectivity for the ketone (+)-trans-dihydrocarvone from 99% "abnormal" to 99% "normal" could be achieved.[47a] The three applied mutations were transferred to $CHMO_{Acineto}$ as well leading to the very same change.[47a] These mutations enabled the substrate to sterically orientate in the favorized way, meaning that the higher substituted carbon atom in the Criegee intermediate could be positioned antiperiplanar to the peroxide group of the peroxyflavin resulting in the formation of the "normal" instead of the "abnormal" lactone.[47a] Apart from the positioning of the substrate in the active site, the regioselectivity can be influenced by oxygen and substrate concentrations, which is not fully understood yet.[49d, 51] Even the formation of carbonates, through the insertion of oxygen atoms on both sides of the carbonyl function, has been reported (Scheme 1.5).[52]

Scheme 1.5: Observed occurence of two subsequent Baeyer-Villiger oxidations catalyzed by the BVMO CcsB leading to the formation of the carbonate cytochalasin Z16, an intermediate of the cytochalasin E and K synthesis in *Aspergillus clavatus* (from Hu *et al.*).[52]

Instead of attacking the substrate and thus leading to product formation, the C4α-(hydro)peroxyflavin can also decay to hydrogen peroxide in the course of the reaction cycle of BVMOs.[53] This process is called uncoupling and especially occurs in the absence of a substrate as none of the two oxygen atoms of peroxyflavin can be further transferred to another compound. Thus, the hydride transferred from NAD(P)H to the flavin yields H_2O_2 instead of oxidized product and water.[53] However, uncoupling is also possible when substrate is available, making it necessary to differentiate between activity (formation of ε-caprolactone) and uncoupling (yielding H_2O_2 from NAD(P)H).[53b, 53c] The formation of hydrogen peroxide is not only problematic as it leads to a reduced efficiency of the catalyst, but it also damages the enzyme.[53a] Mainly the sulfur containing cysteine and methionine

residues are targets for this oxidative stress.[53a] When these amino acids are located in the active site, this leads to a reduced or even nullified activity.[53a, 54] At distal sites the oxidation leads to structural changes, which can alter the enzymes' performance or decrease its stability resulting in protein denaturation.[53a, 55] Mostly, the reactive C4α-(hydro-)peroxyflavin is stabilized by the BVMO to prevent its decay to hydrogen peroxide. In course of mutagenesis experiments with PAMO, J. Cahn was confronted with an enzyme variant displaying >99% uncoupling so that cofactor consumption was completely separated from product formation.[56] A negative screening without substrate was claimed not to be possible due to activity of the enzyme towards natural compounds of the *E. coli* lysate.[56] None of the additional investigated mutations restored the coupled activity. Therefore, he concluded that the complex and not fully understood multistep electron transfer pathways found in BVMOs can easily be disturbed by mutations, especially in the cofactor binding pocket, leading to a state in which the donor-acceptor pairs are not aligned precisely enough anymore.[56] In this state the hydride originally transferred from the NADPH can be passed on to other acceptors like O_2.[56] A mutational analysis in the course of a structural exploration of $CHMO_{Rhodo}$ and cyclopentadecanone monooxygenase (CPDMO) by Yachnin *et al*. took this problem into account.[53c, 57] Under limiting NADPH concentrations the consumption of the cofactor was measured photometrically for several days. After NADPH depletion was achieved, the reactions were analyzed by GC to check for actual product formation. This way, they could demonstrate that all of their six generated mutants showed a higher degree of uncoupling than the wild type enzyme.[53c] As uncoupling can always occur when working with BVMOs, such a confirmation of product formation is needed for convincing results. Therefore, measurements based on the depletion of NAD(P)H are insufficient as they do not give information about the desired product formation. Consequently, biocatalysis with subsequent product analysis via GC and/or HPLC is required.

Regarding their cofactor dependence, BVMOs can be divided into different types.[58] The majority of known BVMOs belong to the NADPH-dependent type I BVMOs that bind FAD and consist of one polypeptide chain that catalyzes both substrate oxidation and flavin reduction. To bind NADPH and FAD, they possess two dinucleotide binding domains, called Rossmann folds, consisting of a β-α-β-α-β fold containing the characteristic sequence motif GxGxx[G/A].[59] Two more sequence motifs are characteristic for type I BVMOs, which are surrounded by both Rossmann folds. These are called fingerprints, the first with the consensus FxGxxxHxxxW[P/D] and the later identified motif [A/G]GxWxxxx[F/Y]P[G/M]xxxD.[34a, 58a, 60] The "fingerprint 2" contains the catalytically important aspartate, which coordinates FAD

as well as the catalytic arginine, and is due to its higher conservation probably better suited to identify new type I BVMOs.[61] As FMOs do not contain this sequence motif, it can also be employed to differentiate between these enzymes and BVMOs.[34a, 62] "Fingerprint 1" can be found in group B FMOs, but with the deviation of Y(K/R) instead of W(P/D).[62] In the course of my diploma thesis by employing a multiple sequence alignment consisting of the protein sequences of characterized and putative BVMOs, two new fingerprints were derived, Dx[I/L][V/I]xxTG[Y/F] and [G/D][P/A]xxYxxxxxxxxPN[L/M][W/F]xxxG, designated "fingerprint 3" and "fingerprint 4", respectively.[53b, 63] After further verification, these could be useful for the identification of type I Baeyer-Villiger monooxygenases in sequence databases as well. In contrast, type II BVMOs utilize NADH and FMN and in addition to the type I BVMOs a suitable reductase is required.[41, 64] In other words, there is no catalysis without the fitting flavin-reductase. That is probably one of the main reasons why just a few type II BVMOs could be investigated until now. The conversions of (±)-camphor and norcamphor with the type II BVMOs 2,5-Diketocamphane-1,2-monooxygenase and 3,6-Diketocamphane-1,6-monooxygenase were possible with the reductase Fre from *E. coli*, equivalent to reactions with their natural reductase Fred from *P. putida*.[64-65] This finding could enable the further discovery, characterization and especially the applicability of new type II Baeyer-Villiger monooxygenases. Recently, Willetts, *et al*. reported on a FAD-dependent monooxygenase from *Stenotrophomonas maltophilia* PML 168 (SMFMO), able to accomplish Baeyer-Villiger reactions using FAD and NADH and they proposed that this enzyme constitutes a new class named type III BVMOs.[46a] Furthermore, a fourth type of BVMO was identified.[45f] The monooxygenase MtmOIV from *Streptomyces argillaceus* is one example belonging to this as type 0 designated group.[45f] On the one hand, it has many properties in common with type I BVMOs, like the utilization of the cofactor NADPH and FAD and the fact that the enzyme consists of only one polypeptide chain. On the other hand, the typical sequence motifs, which can be found in type I BVMOs, are not present. Additionally, the protein structure is significantly different, making it necessary to separate this enzyme from the remaining BVMOs.

Baeyer-Villiger monooxygenases display rather different substrate scopes. Linear, mono-, bi-, tri- and heterocyclic ketones as well as steroids and terpenoids can be converted by members of this class of enzymes.[41, 49c, 58a, 66] BVMOs not only catalyze the oxygenation of ketones but also the epoxidation of C-C double bonds, conversion of aldehydes to either fatty acids or esters and the oxidation of heteroatoms like sulfur, nitrogen, phosphorus, boron or selenium.[58a, 67] Investigations of Ryerson *et al*. demonstrated the participation of a

4α-hydroperoxyflavin intermediate during catalysis performing S-oxygenation of thiane to thiane-1-oxide.[43] In contrast to the Baeyer-Villiger oxidation activity, where the peroxyflavin acts as a nucleophile, this reaction shows a nucleophilic displacement by the sulfur towards the terminal oxygen of the hydroperoxide and accordingly an electrophilic attack of the latter on a lone electron pair of the heteroatom.[43, 68] The oxidation of sulfides or other heteroatom containing compounds often leads to the formation of chiral products (Scheme 1.6).

Scheme 1.6: Oxidation of sulfides to chiral sulfoxides as an example for heteroatom oxygenation catalyzed by BVMOs. In case of a second oxidation, the corresponding sulfone is produced (from Bordewick).[69]

If a second oxidation occurs, the corresponding fully oxidized heteroatom is generated. In case of sulfides, which are firstly oxidized to sulfoxides, the corresponding achiral sulfone is usually formed as a side product of the heteroatom oxygenation.[48b, 67a, 70] In some cases sulfone formation can be high. BVMOAf1 from *Aspergillus fumigatus* produced 25% sulfoxide and 75% sulfone from benzyl ethyl sulfide, being the maximal reported sulfone production to date.[71] Nowadays, high sulfone formations from sulfides are undesired as the chiral sulfoxides are employed as chiral auxiliaries in organic synthesis and the production of pharmaceuticals like armodafinil or esomeprazol (Codexis, Scheme 1.7).[72]

Scheme 1.7: Synthesis of esomeprazole in Codexis using the sulfoxidation activity of $CHMO_{Acineto}$.

For substrate specificity of BVMOs no determining residues could be identified so far.[73] Every substrate able to diffuse into the catalytic center of the active site will be activated for a nucleophilic attack of the flavin-peroxide through a hydrogen bond formed between its carbonyl oxygen and a conserved catalytic arginine of the BVMO.[73] The positive charge created by this arginine together with $NADP^+$ stabilizes the resulting negatively charged Criegee intermediate as well leading to the subsequent catalytical steps and the formation of the product.[73]

Because of their broad substrate scope, Baeyer-Villiger monooxygenases can be further employed for the synthesis of a range of important other compounds.[66b] One promising industrial application of the $CHMO_{Acineto}$ is the conversion of cyclohexanone to ε-caprolactone.[74] This monomer is an important building block in the polymer synthesis of poly-ε-caprolactone. ε-Caprolactone is still produced chemically at a multi-10,000 ton scale per year despite the hazardous peracetic acid used as oxidation reagent and the modest selectivity (85–90%).[75] It was shown recently that an enzyme cascade production of the monomer as well as oligomers is possible with $CHMO_{Acineto}$ in an environmentally friendly and selective way and this could replace the chemical synthesis currently used.[76] ε-Caprolactone can also be a precursor for the production of caprolactam, which further can be polymerized to polycaprolactam (nylon-6).[77] $CHMO_{Acineto}$ and the cyclododecanone monooxygenase (CDDMO) can also be applied to synthesize *cis*/*trans*-jasmine lactone to be used as a fragrance or to produce protected β-amino acid/alcohols for the pharmaceutical industry.[49c, 78] Sigma Aldrich was employing $CHMO_{Acineto}$ in a continuous stirred tank reactor approach using whole cells to convert bicyclo[3.2.0]hept-2-en-6-one to the corresponding lactones in kg scale.[79] In a follow-up study 48% conversion could be achieved in a scale of 200 L leading to a yield of 495 g lactone product.[80]

1.3.1.2 Discovery of Baeyer-Villiger monooxygenases

1948 the first indication for Baeyer-Villiger monooxygenases was found during investigations of the degradation of steroids by fungi.[81] Later, they were discovered in bacteria, in which they are important for oxidative catabolic pathways.[82]

For instance, *Acinetobacter* sp. NCIMB 9871 contains a BVMO (cyclohexanone monooxygenase, CHMO) participating in a pathway enabling this bacterium to grow on cyclohexanol as a sole source of carbon (Scheme 1.8).

Scheme 1.8: Degradation of cyclohexanol in *Acinetobacter* sp. NCIMB 9871.
ChnB: cyclohexanone monooxygenase (adapted from Iwaki *et al*.).[83]

Cyclohexanol is first oxidized to cyclohexanone by an alcohol dehydrogenase (ChnA). Then, this ketone is converted by $CHMO_{Acineto}$ (ChnB) to ε-caprolactone. After a sequence of other enzymes acetyl-CoA and succinyl-CoA are formed fully integrating cyclohexanol into the metabolism of *Acinetobacter* sp.[83] $CHMO_{Acineto}$ was the first type I BVMO to be purified and characterized.[84] Besides its natural substrate cyclohexanone, $CHMO_{Acineto}$ displays a remarkably broad substrate scope accepting over 100 substrates, including many derivatives of cycloalkanones, but also more complex bi- or tricyclic molecules and a variety of heteroatom compounds are converted often with high enantioselectivities.[66a, 70b, 85] Another example can be found in *Pseudomonas putida*, which actually possesses three BVMOs participating in the degradation of (±)-camphor (Scheme 1.9).[86]

Scheme 1.9: Degradation of (±)-camphor in *Pseudomonas putida* NCIMB 10007 (from Kadow *et al*.).[87] Firstly, a hydroxylation of camphor (**1**) by the P450Cam monooxygenase results in the formation of 5-exo-hydroxycamphor (**2**).[88] Subsequently, an oxidation by the 5-exo-alkohol dehydrogenase leads to the corresponding diketocamphane (**3**).[89] While (+)-**3** is converted by the 2,5-diketocamphane 1,2-monooxygenase (a), (-)-**3** only can be oxygenized by the 3,6-diketocamphane 1,6-monooxygenase (b), leading to the formation of unstable lactones. Then, both lactones spontaneously form 2-oxo-Δ^3-4,5,5-trimethylcylopentenylacetic acid, which is converted to 2-Oxo-Δ^3-4,4,5-trimethyl-cyclopentenylacetyl-CoA (**4**) by a CoA-Ester synthethase. 2-oxo-Δ^3-4,5,5-trimethylcyclopentenylacetyl-coenzyme A monooxygenase (OTEMO, c) is the third BVMO participating in this degradation pathway, accepting **4** and producing the corresponding δ-lactone (**5**), which can be hydrolyzed by an esterase for its integration into the metabolism.[86-87]

The 2-oxo-Δ^3-4,5,5-trimethylcyclopentenylacetyl-coenzyme A monooxygenase (OTEMO) was the first of the three enzymes to be identified.[86] Together with the later discovered type II BVMOs 2,5-Diketocamphane-1,2-monooxygenase (2,5-DKCMO) and 3,6-Diketocamphane-1,6-monooxygenase (3,6-DKCMO) and other catabolic enzymes OTEMO constitutes a way to utilize (±)-camphor as a source of carbon and energy.[90] In 2012, it could be successfully expressed recombinantly and characterized by Kadow *et al*.[91] 2,5-DKCMO and 3,6-DKCMO are the first and to date only type II Baeyer-Villiger monooxygenases to be purified and characterized.[87, 91] Similarly, the cyclopentanone monooxygenase (CPMO) discovered in 2002 allows *Comanonas sp.* NCIMB 9872 to grow on cyclopentanol, converting the resulting cyclopentanone to the corresponding δ-valerolactone, enabling its further degradation.[51a, 92] 4-hydroxyacetophenone monooxygenase (HAPMO, designated from its natural substrate) from *Pseudomonas fluorescens* ACB was the first characterized BVMO that accepted aromatic ketones.[51a, 93] With the identification of the BVMO "fingerprints 1 and 2" in 2002 and 2012, respectively, the possibility to discover putative Baeyer-Villiger monooxygenases was dramatically increased.[34a, 58a, 60] Since then, the classical way to find BVMOs in organisms by investigating their metabolism changed to the bioinformatic approach of genome mining.[58a, 61] An outstanding example for this is the identification of more than 20 putative Baeyer-Villiger monooxygenases in the genome of *Rhodococcus jostii* RHA1.[34a, 51b] Even though, these enzymes can be frequently found, usually only a few BVMOs are present in an organism. In

2012, 22 of the identified genes were cloned, followed by heterologous expression in *E. coli* and characterization.[34a] In the thermophilic bacteria *Thermobifida fusca* and *Thermocrispum municipale* DSM 44069, PAMO and the cyclohexanone monooxygenase TmCHMO were discovered by genome mining as well.[94] These two enzymes display a high thermostability (melting temperature of 61 °C and 48 °C, respectively) in contrast to the low stability of Baeyer-Villiger monooxygenases in general (i.e. 37 °C for $CHMO_{Acineto}$).[45g, 53a, 94] PAMO is probably part of a degradation pathway of aromatic compounds, whereas TmCHMO is likely needed for the degradation of cyclohexanone.[45g, 94] In total, over 80 bacterial Baeyer-Villiger monooxygenases have been identified to date complementing each other in regard to their substrate scopes (Figure 1.6).[41, 58a, 66b, 95]

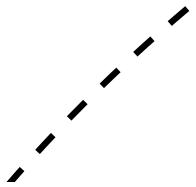


Figure 1.6: Complementing substrate scopes of four different BVMOs.
The preferred respective substrate is shown in grey. CHMO: Cyclohexanone monooxygenase from *Acinetobacter calcoaceticus*, CPMO: Cyclopentanone monooxygenase, HAPMO: 4-Hydroxyacetophenone monooxygenase, PAMO: Phenylacetone monooxygenase (from Pazmiño *et al*.).[95]

The comparison of the substrate specificities of $CHMO_{Acineto}$, CPMO, HAPMO and PAMO shows that many substrates are converted by more than one BVMO. However, some substrates are solely accepted by one enzyme and each Baeyer-Villiger monooxygenase prefers a certain class of substrate. What cannot be seen in Figure 1.6 is that there are still many compounds that are either not or barely converted by the available set of bacterial BVMOs.

This just started to change in 2011 with the identification and characterization of an eukaryotic Baeyer-Villiger monooxygenase – the cycloalkanone monooxygenase (CAMO) from *Cylindrocarpon radicicola* ATCC 11011.[96] This enzyme converts many cyclic and bicyclic as well as aliphatic ketones, even though the latter are not accepted by the majority of BVMOs, making CAMO a special catalyst. Inspired by this, six more BVMOs were found in fungi – one in *Aspergillus fumigatus* Af293, four in *Aspergillus flavus* NRRL3357 and one in *Thermothelomyces thermophila* ATCC 42464.[45h, 71, 97] BVMOAf1 from *A. fumigatus* showed excellent chemo-, regio- and stereoselectivity in the kinetic resolution of bicyclo[3.2.0]hept-2-en-6-one as well as in the oxidation of asymmetric sulfides.[71] In contrast, cyclohexanone and phenylacetone, well accepted BVMO substrates in general, were not converted. Worth mentioning is also the robustness of this enzyme; an incubation in organic solvents or a wide pH range did not decrease its activity.[71] The four enzymes $BVMO_{AFL210}$, $BVMO_{AFL456}$, $BVMO_{AFL619}$, $BVMO_{AFL838}$ from *A. flavus*, of which the last three atypically did mostly not accept cyclic ketones, readily converted aliphatic ketones.[97] In the thermophilic fungus *T. thermophila* the BVMO PockeMO was characterized.[45h] Remarkable is its high thermostability with a melting temperature (Tm) of 47 °C, making it comparably robust as PAMO (61 °C).[98] In combination with its perfect enantioselectivity and a broad substrate scope with the highest activity for bulky molecules, including steroids, it is one of the most applicable Baeyer-Villiger monooxygenases described to date.[45h] Furthermore, two BVMOs have been identified in the photosynthetic eukaryotes *Physcomitrella patens* (a moss) and *Cyanidioschyzon merolae* (a red alga).[99] Pp-BVMO and Cm-BVMO convert a large set of ketones, including aryl aliphatic, aromatic, bicyclic and aliphatic compounds of which the latter were the best substrates, even with alternating positions of the keto function.[99] Additionally, the melting temperature of Cm-BVMO was determined to be 56 °C, ranking it the second most thermostable type I BVMO.[99] To date only the described nine special eukaryotic BVMOs have been characterized. Following my diploma thesis, in the course of this exploration, there was the desire to discover more eukaryotic BVMOs in yeasts.[63] Yeasts are fungi, which mostly grow in form of single cells or pseudomycel, reproduce by budding or division and possess a smooth colony morphology. Some yeasts show an excellent performance in the degradation of *n*-alkanes and other aliphatic hydrocarbons.[100] Especially worth mentioning are members of the genus *Candida*, which consistently showed the best growth with hydrocarbons among numerous tested fungi.[100b, 100c] When the substrate of such *n*-alkane-assimilating yeasts changes from glucose to *n*-alkanes a number of enzymes become induced.[101] Among these enzymes are P450 monooxygenases initiating the aerobic degrada-

tion of the highly hydrophobic alkanes to be used as the sole source of carbon and energy.[100a, 102] The oxidation catalyzed by these enzymes can occur at one or both ends of the alkane or intramolecular leading to the formation of primary and secondary alcohols and diols.[100a, 103] In case a secondary alcohol is formed, an alcohol dehydrogenase can convert it to the corresponding ketone.[104] Subsequently, the ketone can be further metabolized to the respective ester through a Baeyer-Villiger oxidation, which is then hydrolyzed by an esterase (Scheme 1.10).[104-105]

Scheme 1.10: Subterminal degradation of *n*-alkanes, exemplary for dodecane.
After the alkane dodecane is hydroxylated by a P450 monooxygenase, the resulting secondary alcohol 2-dodecanol is oxidized to the ketone 2-dodecanone, which can be a substrate for a BVMO. After Baeyer-Villiger oxidation, an esterase hydrolyzes the formed ester to an alcohol (decanol) and a carboxylic acid (acetic acid), which subsequently can be further degraded. P450: P450 monooxygenase, ADH: alcohol dehydrogenase.

To investigate the capability of selected yeast strains to utilize alkanes and their corresponding alkenes and ketones, a plate assay with dodecane, 1-dodecene and 2-dodecanone was carried out in my diploma thesis.[63] Each yeast displayed a specific growth pattern (Table 1.4).

Table 1.4: Growth of selected yeast strains in a plate assay with dodecane (A), 1-dodecene (B) or 2-dodecanone (C) as a sole source of carbon and energy (adapted from Beier).[63]

Strain	A	B	C
Candida maltosa SBUG 700	++	++	++
Candida tropicalis SBUG 1019	++	++	+
Candida catenulata SBUG 512	++	(+)	+
Pichia guilliermondii SBUG 50	++	+	-
Lodderomyces elongisporus SBUG 400	+	-	-
Yarrowia lipolytica SBUG 1888	+++	(+)	(+)
Rhodosporidium toruloides SBUG 137	(+)	-	-
Trichosporon asahii SBUG 833	-	+	++
Candida utilis SBUG 61 (NC)	-	-	-
Saccharomyces cerevisiae SBUG 118 (NC)	-	-	-

+++: very strong growth, ++: strong growth, +: medium growth, +: weak growth, (+): very weak growth, -: no growth, NC: negative control

In total five yeast strains, among them *Candida maltosa* and *Yarrowia lipolytica*, were able to use the ketone as a sole source of carbon and energy, which indicated the presence of enzymes with Baeyer-Villiger monooxygenase activity. This was also indicated later through growth experiments in liquid medium with subsequent analysis of the metabolites. However, a subterminal hydroxylation of dodecane could be excluded as 2-dodecanol could only be detected in cultures with 1-dodecene and 2-dodecanone.[63] In a study of Lowery *et al*. from 95 fungi also *Yarrowia lipolytica*, two *Candida* strains and two *Rhodotorula glutinis* strains were able to utilize ketones as substrates.[100c] Using the BVMO fingerprint, the Rossmann fold motifs and the protein BLAST, the genome sequences of *Candida maltosa*, *Candida tropicalis*, *Candida albicans* and *Candida dubliniensis* were screened for sequences belonging to Baeyer-Villiger monooxygenases, participating in their ketone metabolism.[58a, 60, 63, 106] Hereby, 21 putative BVMOs and FMOs were identified, of which 19 are unique on the protein sequence level (Table 1.5).

Table 1.5: In the genome sequences of *C. maltosa, C. tropicalis, C. albicans* and C. *dubliniensis* identified putative BVMO/FMO protein sequences.

Organism	Protein	Sequence motifs	Accession
Candida maltosa	$BVMO_{malto}$	1, 2, R(2x)	EMG51019.1
	TMO	R	EMG47137.1
	MO_{malto1}	R	EMG45888.1
	MO_{malto2}	R	EMG50453.1
Candida tropicalis	$BVMO_{tropi}$	1, 2, R(2x)	XP_002546907.1
	MO_{tropi1}	R	XP_002547981.1
	MO_{tropi2}	R	XP_002548149.1
	MO_{tropi3}	R	XP_002550122.1
	MO_{tropi4}	2 (1M), R	XP_002545727.1
Candida albicans	$BVMO_{albi1}$	1, 2, R(2x)	XP_720980.1
	$BVMO_{albi2}$ (identical to 1)	1, 2, R(2x)	-
	$BVMO_{albi3}$	1, 2, R(2x)	XP_722567.1
	$BVMO_{albi4}$ (identical to 3)	1, 2, R(2x)	-
	$BVMO_{albi5}$	1, 2, R(2x)	KHC81419.1
	MO_{albi1}	1, 2, R(2x)	KHC81420.1
	MO_{albi2}	1, 2, R(2x)	EEQ45644.1
	MO_{albi3}	R	XP_716605.1
	MO_{albi4}	R	KHC81819.1
	MO_{albi5}	R	XP_718130.1
Candida dubliniensis	MO_{dubli1}	1, 2, R(2x)	XP_002418311.1
	MO_{dubli2}	1, 2, R(2x)	XP_002418165.1

1: “fingerprint 1”, 2: “fingerprint 2”, R: Rossmann fold, M: mutation, MO: monooxygenase, TMO: thiol-specific monooxygenase.

Nine of them contained both fingerprints, two Rossmann folds and displayed sequence identities to known Baeyer-Villiger monooxygenases of about 25-40%, what made them promising candidates for further investigations.[63] The two sequences thiol-specific monooxygenase (TMO, later designated *Candida* monooxygenase, CMO) and 45888 from *C. maltosa* showed higher sequence identities to FMOs of up to 73%, making them interesting targets for the characterization of new members of this group of enzymes.[63, 107] $BVMO_{malto}$ and CMO could already be cloned into pET28a(+) vectors and expressed in *E. coli*.[63] However, the latter was only achieved in form of inclusion bodies.[63] In another study, 107 fungal strains were also screened for BVMO activity.[108] Here, 86 strains converted bicyclo[3.2.0]hept-2-en-6-one to the corresponding lactones. *Yarrowia lipolytica* and a

Candida strain (*glabrata*-like) were again among the ketone-utilizing organisms. Additionally, Mascotti *et al*. screened genome sequences originating from all domains of life – Eukarya, Bacteria and Archaea.[35] With this approach, in addition to 92 new potential BVMOs, six putative Baeyer-Villiger monooxygenases could be identified in the genome of the yeast *Yarrowia lipolytica*.

The discovery and characterization of eukaryotic Baeyer-Villiger monooxygenases expands the set of these valuable biocatalysts for a higher diversity of substrates, other enantio- and regioselectivities and consequently a new line of products and different reaction conditions needed for industrial processes. This group of BVMOs is barely explored and thus promises enzymes greatly differing from their prokaryotic representatives, having the potential to enrich the field of the White Biotechnology.

1.3.1.3 Handling the cofactor dependency of BVMOs

1.3.1.3.1 Biocatalytic strategies

As discussed above, Baeyer-Villiger monooxygenases need the cofactor NADPH in stoichiometric amounts for their activity. For industrial applications, it would be too expensive to use such high amounts of this expensive compound. This can be circumvented by employing cofactor recycling systems, like coupling the activities of a BVMO and a glucose-, alcohol- or phosphite dehydrogenase where a cheap co-substrate is used to regenerate the NADPH. [58a] Even fusion of the two cofactor-complementing enzymes is possible, creating a self-sufficient enzyme (Figure 1.7).[109]

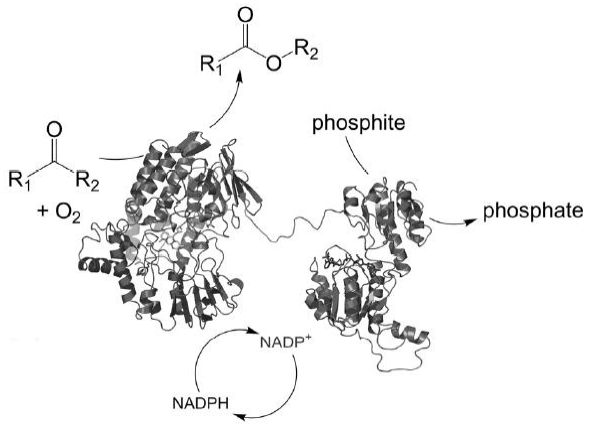


Figure 1.7: Fusion of a BVMO with a phosphite dehydrogenase, creating a self-sufficient enzyme with internal cofactor recycling (from Pazmiño *et al.*).[109a]

Another way is the application of an enzyme cascade, in which the BVMO is used together with other enzymes to create a redox neutral process.[66b, 110] This can serve for the production of valuable compounds like 6-aminohexanoic acid (nylon-6 monomer) or poly-ε-caprolactone and has the additional advantage of preventing substrate and product inhibition (Scheme 1.11).[66b, 76, 111]

Scheme 1.11: Exemplary enzyme cascades consisting of a BVMO together with other enzymes for the sake of cofactor recycling to create a redox neutral process, production of the valuable compounds 6-aminohexanoic acid (nylon-6 monomer) and poly-ε-caprolactone and prevention of substrate and product inhibition starting from cyclohexanol (adapted from Bučko *et al.*).[66b, 76, 111]

ADH: alcohol dehydrogenase, BVMO: $CHMO_{Acineto}$, CalA: *Candida antarctica* lipase A, ω-TA: ω-transaminase, AlaDH: Alanine dehydrogenase

The next strategy is the application of whole-cell systems, in which the metabolism of the enzyme expressing host cell continuously regenerates the cofactor by consuming glucose or another inexpensive co-substrate.[112] Moreover, this avoids the protein instability issues faced when working with BVMOs.[113] One example is the production of chiral carvolactone from limonene, which can easily be obtained from the waste product orange peel.[114] The chiral carvolactone can be seen as a building block for the synthesis of natural or bioactive products or as a monomer for thermoplastic polymers (Figure 1.8).[66b, 114]

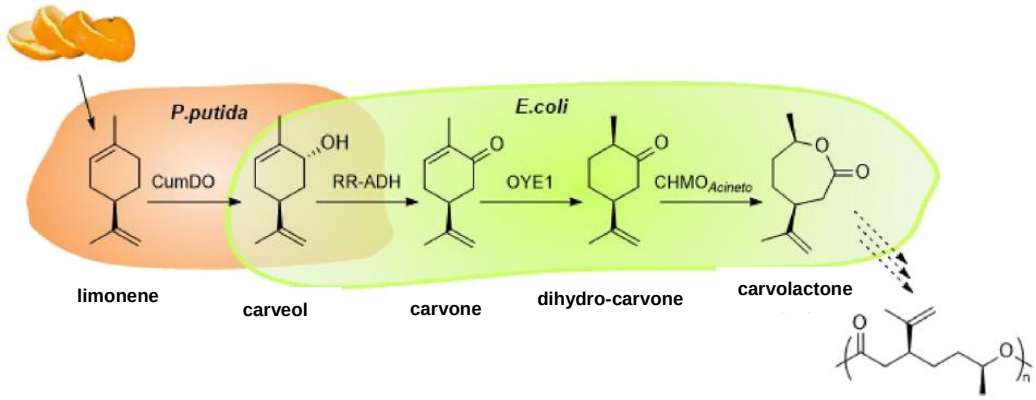


Figure 1.8: Mixed whole cell approach for the production of a chiral carvolactone (adapted from Bučko *et al.*)[66b, 114]
Limonene extracted by water from orange peel is converted to carveol in *Pseudomonas putida*. Subsequently, a reaction cascade in *E. coli* transforms carveol via carvone and dihydro-carvone directly to chiral carvolactone. Next, a ring-opening followed by polymerization can be carried out. CumDO: cumene dioxygenase from *Pseudomonas putida* PWD32, RR-ADH: alcohol dehydrogenase from *Rhodococcus ruber*, OYE1: Old Yellow Enzyme 1 – enoate reductase XenB from *Pseudomonas putida*.

1.3.1.3.2 Engineering the cofactor specificity of BVMOs

Even though all the approaches mentioned above are working, it would be highly useful to use NADH instead of NADPH, since the latter is ten times more expensive and even less stable.[115] Furthermore, recycling of NAD^+ can be performed more easily compared to $NADP^+$-recycling.[116] Even for metabolic engineering, cofactor dependence is an issue.[117] When the cofactor availability is well balanced in the host cell, product yields are improved as oxygen supply is not needed anymore, side products and carbon inefficiencies are eliminated and steady-state metabolite levels are enhanced.[118] As BVMOs usually do not display sufficient activity with NADH, the k_{cat}/K_M of $CHMO_{Acineto}$ for NADPH being 600-fold higher than for NADH, they cannot be employed with this cofactor.[119] However, with the tools protein engineering is offering, the cofactor specificity of enzymes can be changed, as reported in 1990 for a gluthatione reductase for the first time.[120] In the majority of the successfully designed enzymes, the phosphate binding site of the cofactor binding pocket was targeted.[37-38, 42a, 43, 46a, 59b, 65a, 74-75, 76b, 85, 115-116, 121] The reason is obvious – the only structural difference

between NADPH and NADH is the additional esterification of the 2'-hydroxy group of the ribose unit with phosphoric acid (Figure 1.9).

Figure 1.9: Comparison of the structures of NAD^+ and $NADP^+$. The only difference can be found at the 2'-position – a hydroxy group in NAD^+ and a phosphate group in $NADP^+$.

Examples for enzymes with engineered cofactor specificity exist for both directions, from NADPH towards NADH and *vice versa*. However, in the majority of the cases the switch in cofactor preference was accompanied with a loss of catalytic activity.[118h] This has various reasons. Even though the cofactor binding pockets of the proteins are not in proximity of their catalytic sites, they affect their activity.[118h] Mutations in that region can influence enzyme kinetics and sometimes substrate specificity.[121d, 122] Even small chemical modifications of the flavin cofactor can greatly alter the activity and the dynamic changes of the protein throughout the catalytic cycle make mutational effects hard to predict.[118h, 123] Furthermore, the diversity of structural motifs to bind the nicotinamide cofactors and the means to discriminate between NADP(H) and NAD(H) in nature are huge.[59a, 118h, 123c, 124] Apart from the Rossmann fold, FAD/NAD-binding, TIM-barrel and dihydoquinoate synthase-like (DHQS-like) and other folds can be found. Both NADPH and NADH acceptance can be found within the same fold and also within an enzyme family, which can be caused by independent evolution leading to unique determinants for cofactor specificity for each enzyme.[118h, 124a, 125] However, residues decisive for the specificity can be classified into six groups (Figure 1.10).[118h]

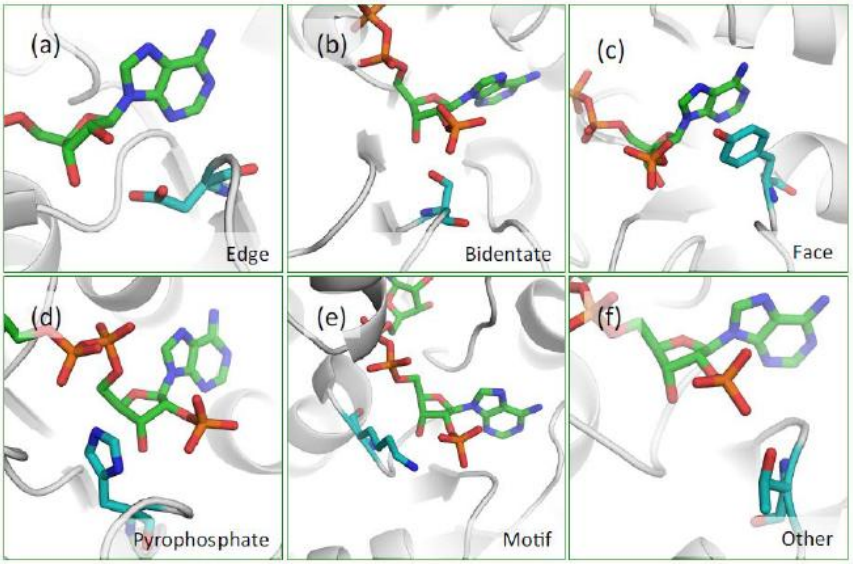


Figure 1.10: The six structural residue classes in NAD(P)H utilizing enzymes (from Cahn *et al.*).[118h]
Edge: Along and parallel to the edge of the adenine moiety, often with contact to it; Bidentate: Contact to the 2'- and 3'-oxygen; Face: cation-pi interaction with the adenine moiety as well as an often observed hydrogen bond to the 2'-oxygen; Pyrophosphate: Hydrogen bond to the 2'-oxygen and the pyrophosphate; Motif: The not conserved residues (x) in the Rossmann fold motif Gx(x)Gx(x)[G/A], occasionally contacting the cofactor molecule. The glycines show contacts to the pyrophosphate moiety; Other: Residues which cannot be grouped into the other classes; PDBs of the exemplary structures: 1VC2, 1AMO, 1AMO, 1EZ0, 1CYD, and 1CYD.

A major influence on the discrimination between NADPH and NADH seems to come from the polarity of the cofactor binding pocket.[118h] In NADPH utilizing enzymes, residues with a positive charge are dominant to bind the negative 2'-phosphate group (Figure 1.11).

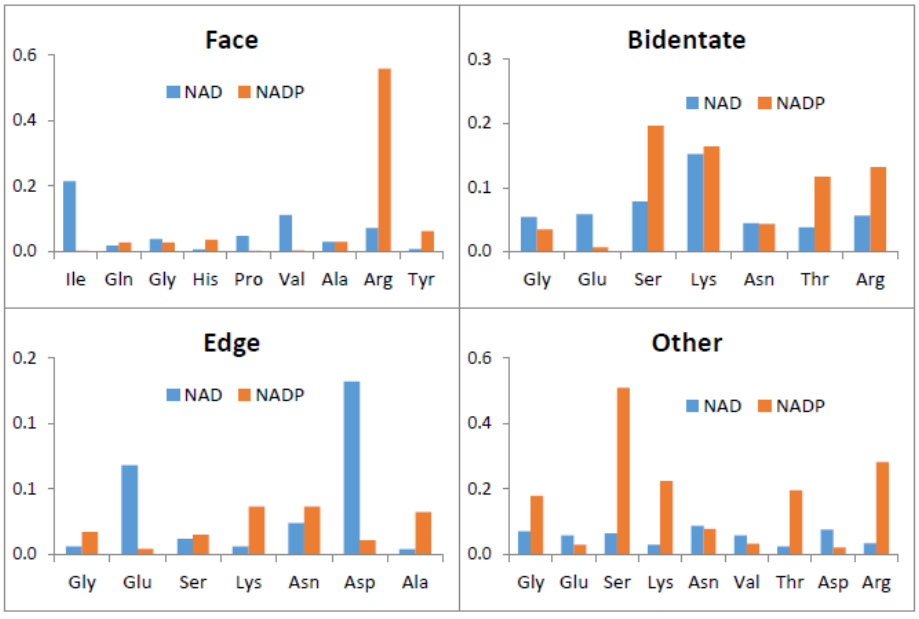


Figure 1.11: Average number of specificity-determining residues of the most influential structural classes per NAD(P)H binding pockets in NADP-bound and NAD-bound structures, assembled from 463 NADP-bound structures and 499 NAD-bound structures (from Cahn *et al.*).[118h]
Only the most frequent amino acids are shown, respectively.

Here, a conserved arginine ("face") can be found, of which the positively charged guanidinium moiety displays cation-pi stacking interactions with the adenine moiety.[118h, 121a, 126] This as well as the other residues of the binding pocket form hydrogen bonds not only with the oxygen atoms of the phosphate, but also with the 3'-hydroxyl group.[118h, 121a] In contrast, in enzymes specific for NADH, residues with a negative charge can be found (Figure 1.11).[118h] These repulse the 2'-phosphate electrostatically and display hydrogen bonds to the 2'- and 3'-hydroxyl groups of NADH.[118h] Sometimes an arginine can be found in these enzymes as well, coordinating the cofactor through cation-pi interactions with its adenine moiety.[59a, 118h] To reshape the cofactor binding pocket of a NADPH utilizing enzyme into a NADH specific one and the other way around, mutagenesis approaches have been employed.[59a] In many cases a successful remodeling proved to be difficult as various residues determine the respective specificity.[118h] Thus, simultaneous amino acid exchanges need to be carried out leading to a too high number of possible combinations.[118d, 118h] Furthermore, variants harboring multiple mutations tend to become less effective with every additional mutation due to the strong nonadditivity of mutational effects.[118d, 118h, 127] In the group headed by Frances Arnold many interesting insights into the switch of cofactor specificity of a class of oxidoreductases, the keto-acid reductoisomerases (KARIs), have been provided in the last years dealing with that very problem.[56, 118d, 118h, 121a, 121d, 124a, 125a] For instance, in 2013 Brinkmann-Chen *et al*. presented a guide to reverse the cofactor dependence from NADPH to NADH for this enzyme family.[121a] After classification of the respective KARI to be engineered, one of two sets of mutations has to be chosen. Both sets demand the mutation of the last residue of the cofactor specificity loop situated in the Rossmann fold to an aspartate to enable an interaction with the 2'- and 3'-hydroxyl groups of NADH and repulsing the phosphate group of NADPH.[121a] As such a cofactor switch mostly results in a decreased activity, which also was the case for the KARIs, a random mutagenesis was employed to restore it.[70c, 119, 128] The resulting unpredictable mutations increased the performance of the variants to levels even higher than found for the wild-type enzymes.[121a] Interestingly, some of the amino acid exchanges were placed in regions distant from the specificity determining loop. However, in the course of the random mutagenesis one determining kind of residue could be identified. It influences the binding of the flavin cofactors solely through interaction with the adenine moiety. It corresponds to hydrophobic residues of the adenine recognition site described before.[124d, 129] Due to the slightly different conformation of NADH, the pi-pi interaction between its adenine moiety and the conserved arginine is disturbed.[121a, 126] This needs to be adjusted with the right mutations. Correctly exchanging the determining hydro-

phobic residue led to a slight shift of the NAD(P)H that compensated for the modulated cofactor binding pocket, enabling the same or even more efficient electron transfer necessary for catalysis.[121a] Together with two other residues the conserved arginine also bound the 2´-phosphate group of NADPH or the 2´-hydroxy group of NADH in the mutant enzymes.[121a] In a follow-up study, Cahn *et al*. further investigated the hydrophobic residue identified by random mutagenesis interacting with the adenine moiety.[121d] Here, the activity of seven out of ten enzymes, including two KARIs, different alcohol dehydrogenases, a reductase and one oxidase, with different binding folds (Rossmann, FAD/NAD, DHQS-like) and cofactor preferences could be improved by site-directed saturation mutagenesis at corresponding positions. All this knowledge has been summarized in 2016 to generate a tool to switch the cofactor specificity of any oxidoreductase – Cofactor Specificity Reversal: Structural Analysis and Library Design (CSR-SALAD, Figure 1.12).[118h]

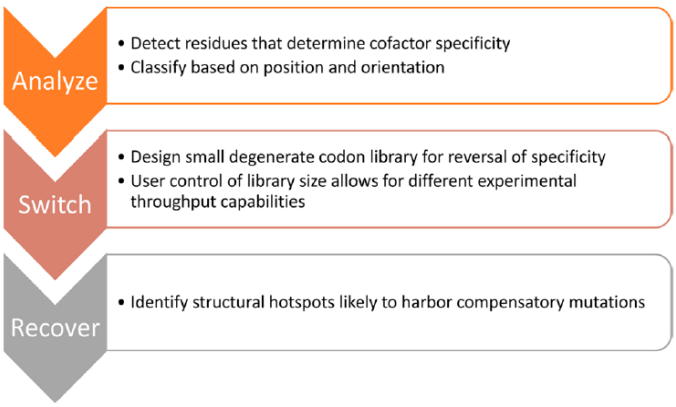


Figure 1.12: Principle of the online tool CSR-SALAD to predict mutations to switch the cofactor preference of oxidoreductases (from Cahn *et al*.).[118h]

This approach was used to successfully demonstrate the switch of the cofactor specificity of four diverse oxidoreductases and recapitulate the precise combination of mutations present in over 20 cofactor switched enzymes previously reported.[118h] This could also be applied to the published attempts to change the cofactor dependence of BVMOs. The first was reported in 2004 by Kamerbeek *et al*.[119] In this study, they analyzed the roles in cofactor specificity of three conserved, basic residues in 4-hydroxyacetophenone monooxygenase (HAPMO).[119] Mutation of R440 (equivalent to R327 in $CHMO_{Acineto}$) to alanine produced a totally inactive enzyme, and R339A (R207 in $CHMO_{Acineto}$) drastically decreased catalytic efficiencies for both NADPH and NADH. However, mutation K439N led to a sixfold increase in catalytic efficiency, and K439F led to a sevenfold increase. They also showed that transferring the mutation R439A (K326) to $CHMO_{Acineto}$ resulted in changed coenzyme specificity as well. Dudek *et al*. investigated mutants of a phenylacetone monooxygenase (PAMO) based on

sequence alignments of type I BVMOs and available structures for PAMO and the CHMO from *Rhodococcus* sp. HI-31.[70c] Their findings confirmed the central role of the conserved R217 (R207 in $CHMO_{Acineto}$) in cofactor binding, as R217 mutants showed a drastic decrease in efficiency with both cofactors. In contrast to Kamerbeek *et al*., they found decreased efficiency with NADH for the mutant K336N (K439N in HAPMO). Mutations of H220 (Q210 in $CHMO_{Acineto}$) to asparagine or glutamine increased efficiencies for NADH threefold. However, neither HAPMO, $CHMO_{Acineto}$ nor PAMO reached activities equal to the wild type with NADPH. Cahn *et al*. also attempted to switch the cofactor specificity of PAMO utilizing CSR-SALAD.[56, 118h] This resulted in a BVMO (PAMO_R217T_T218E_K336Y) displaying a higher activity with NADH than the wild type with NADPH. However, the cofactor oxidation here was completely (>99%) uncoupled so that no product formation occurred, making it a useless biocatalysator.[56] Thus, up to now there was no successfully engineered type I BVMO that efficiently uses NADH as a cofactor for catalysis. However, there are BVMOs being able to efficiently utilize NADH for catalysis, indicating that the 2'-phosphate is not essential for their activity.[56] Völker *et al*. characterized MekA, the BVMO from *Pseudomonas veronii* MEK700; this enzyme is unique as it accepts NADH with an activity of 55% compared to NADPH.[130] Also the type III BVMO SMFMO accomplishes Baeyer-Villiger oxidations with NADH.[46b] The structure of this enzyme shows homologies to type I BVMOs which could give decisive indications for the cofactor discrimination.[46b] Moreover, FMOs with Baeyer-Villiger oxidation activity have been discovered, showing a relaxed cofactor specificity as well.[46b, 62, 131] Therefore, it has to be possible to design a BVMO sufficiently catalyzing Baeyer-Villiger oxidations with NADH. In 2013, Sven Bordewick initiated the rational design of $CHMO_{Acineto}$ to switch its cofactor specificity towards NADH in his bachelor thesis.[132] Structures of the homologous $CHMO_{Rhodo}$ have been chosen as orientation for this structure-guided approach. Here, six positions (T184, R207, S208, Q210, K326 and K349) were striking due to their proximity to the 2'-phophate group of NADPH and thus possible participation in cofactor discrimination. Employing structure- and sequence alignments with diverse BVMOs (PAMO, HAPMO, MekA, SMFMO), promising mutations were selected and subsequently a set of thirteen enzyme variants was generated. Mutations T184G, R207Q/H and K349E decreased the activity with both cofactors while Q210N only led to a lower performance with NADPH. For five variants an increased activity with NADH could be detected, K326H_S208E being the best with a 3.6-fold improvement. This mutant even reached up to 100% of wild-type activity with NADPH in biocatalysis. However, in a photometric assay, no significant activity was observed for any of the variants. Therefore, further

validation of these results and eventually a more comprehensive protein engineering approach is needed to obtain the first truly cofactor-switched BVMO. This would provide insight into the mechanism of cofactor discrimination in this class of enzymes, which could then be applied to further protein engineering studies. Additionally, more cost-efficient oxidative synthesizes could be established, employing Baeyer-Villiger monooxygenases together with the ten times cheaper NADH.

2 Scope of this thesis

Baeyer-Villiger monooxygenases (BVMOs) catalyze a remarkably wide variety of oxidative reactions, which are difficult to obtain chemically. Their main activity is the oxygenation of ketones to esters or lactones by utilizing molecular oxygen and a cofactor. Additionally, epoxidations and the oxidation of heteroatoms like nitrogen, boron or sulfur is catalyzed. The majority of characterized BVMOs consists of prokaryotic enzymes. To date just nine eukaryotic BVMOs are accessible of which many show interesting substrate profiles. Therefore, the discovery and characterization of this special group of monooxygenases of eukaryotic origin and thus an exploration of a higher biocatalytic diversity and the further understanding of this class of enzymes in general were the aims of this thesis.

In order to explore new sources of BVMOs, a microbiological approach was employed, which was initiated in my diploma thesis.[63] Even though most of the known eukaryotic Baeyer-Villiger monooxygenases are of fungal origin, no yeast enzymes are known. Thus, a metabolic analysis of different yeast strains after cultivation with a ketone and the corresponding alkene should give an indication for enzymes with Baeyer-Villiger monooxygenase activity.[63] Subsequently, genome-wide analyses of selected yeasts from my diploma thesis were continued to search for respective gene sequences.[63] Expression of putative monooxygenases was investigated, of which active enzymes were characterized in regard to accepted substrates. Ultimately, a phylogenetic classification of the identified BVMOs could provide insight into their relationship to already characterized type I BVMOs.[69]

Type I BVMOs display a strong preference for NADPH. However, for industrial purposes NADH is the preferred cofactor, as it is ten times cheaper and more stable. To expand the knowledge of BVMOs in general and design an enzyme improved for biocatalysis in industrial scale, a rational protein engineering approach was employed, based on the bachelor thesis of Sven Bordewick.[132] Thus, variants of the cyclohexanone monooxygenase from *Acinetobacter* sp. NCIMB 9871 ($CHMO_{Acineto}$) were created to increase its activity and specificity towards NADH and to explore structural reasons for the preference of the two different cofactor molecules.

3 Results

3.1 Baeyer-Villiger monooxygenases participating in the metabolism of ketones in yeasts

3.1.1 Determination of metabolites from *Candida maltosa* and other yeasts from 2-dodecanone and 1-dodecene

Some yeasts show an excellent performance in the degradation of *n*-alkanes and other aliphatic hydrocarbons.[100] In my diploma thesis, I found evidence for growth of different yeast strains with 2-dodecanone and 1-dodecene and first respective products derived from these cultures.[63] These indications needed to be further studied and verified to get a detailed understanding of the metabolism of these two compounds and participating enzymes, including Baeyer-Villiger monooxygenases, in the investigated yeasts.[63]

To achieve this, the yeast species *Candida maltosa*, *Candida albicans, Candida catenulata, Candida tropicalis, Yarrowia lipolytica, Trichosporon asahii, Pichia guilliermondii, Lodderomyces elongisporus* and *Rhodosporidium toruloides* were cultivated in a mineral salt medium and 2-dodecanone as a sole source of carbon and energy. Additionally, *C. maltosa* was incubated with 1-dodecene and dodecane (control). Then, the medium was analyzed for metabolites via GC-MS. The growth pattern of the strains could be verified.[63] The yeasts *C. maltosa, C. albicans, C. catenulata, C. tropicalis, Y. lipolytica* and *T. asahii* were able to grow with 2-dodecanone as a sole source of carbon and energy. The growth was diverse, though. Members of the genus *Candida*, with exception of *C. albicans*, showed the fastest growth. In the supernatants of the culture media a variety of metabolites was formed (Scheme 3.1, see Appendix Scheme 8.1 for metabolites from 1-dodecene culture and Figure 8.1 for an exemplary GC chromatogram).

Scheme 3.1: Extracellular metabolites detected in yeast cultures containing 2-dodecanone as a sole source of carbon and energy.
Carboxylic acids were detected as methyl esters.[133]

The spectrum of products formed from 2-dodecanone was very similar among the test organisms. In every case 2-dodecanol and also (with the exception of *C. albicans*) decyl acetate was detected, indicating Baeyer-Villiger monooxygenase activity. The incubation of *C. maltosa* with 2-dodecanone led to the detection of six products, while in the incubation with 1-dodecene only five were detected. In control cultures with dodecane only two metabolites were formed, whereas it has to be highlighted that here, in contrast to 2-dodecanone, no formation of ketones, esters and secondary alcohols could be observed. In the other control approaches (with glucose, substrate control and cell control), none of the intermediates could be detected.

3.1.2 Investigations of novel BVMOs from yeasts

After the growth experiments described in chapter 3.1.1 indicated the presence of enzymes with BVMO activity in different yeast strains, the aim was to express selected ones recombinantely.

3.1.2.1 Expression of putative BVMOs from *Candida spec*.

Members of the genus *Candida*, which show the best growth with hydrocarbons among numerous tested fungi, also displayed the best performance in utilizing 2-dodecanone as a sole carbon and energy source.[63, 100b, 100c, 133] Therefore, investigations were started with identified putative enzymes from members of this genus, namely $BVMO_{albi1}$, $BVMO_{malto}$ and Candida monooxygenase (CMO, Table 3.1).

Table 3.1: Overview of identified putative BVMOs/FMOs from *C. maltosa* and *C. albicans* investigated in this thesis.

Organism	Protein	Accession	Length [aa]	Molecular weight [kDa][a]
Candida maltosa	$BVMO_{malto}$	EMG51019.1	500	62.8 (65)
	CMO	EMG47137.1	496	57.2 (59.4)
Candida albicans	$BVMO_{albi1}$	XP_720980.1	552	63.1 (65.3)

[a]: Molecular weight with $His_{(6)}$-tag in parentheses

1: “fingerprint 1”, 2: “fingerprint 2”, R: Rossmann fold, M: mutation

3.1.2.1.1 Expression of a putative BVMO from Candida albicans - *$BVMO_{albi}$*

One member of the *Candida* genus tested was *Candida albicans*. In the genome of this yeast five different putative BVMOs could be identified in the course of my diploma thesis. First, the ORF coding for $BVMO_{albi1}$ was cloned out of the genomic DNA employing classical cloning as it did not show CTG codons, for which a reassignment in this yeast genus was reported (see section 3.1.2.1.3.2). Subsequently, expression of the putative BVMO was investigated in *E. coli* BL21(DE3). In order to evaluate the expression, cultivation samples normalized to 7/OD were taken at different time points and soluble and insoluble fractions were analyzed by SDS-PAGE (Figure 3.1).

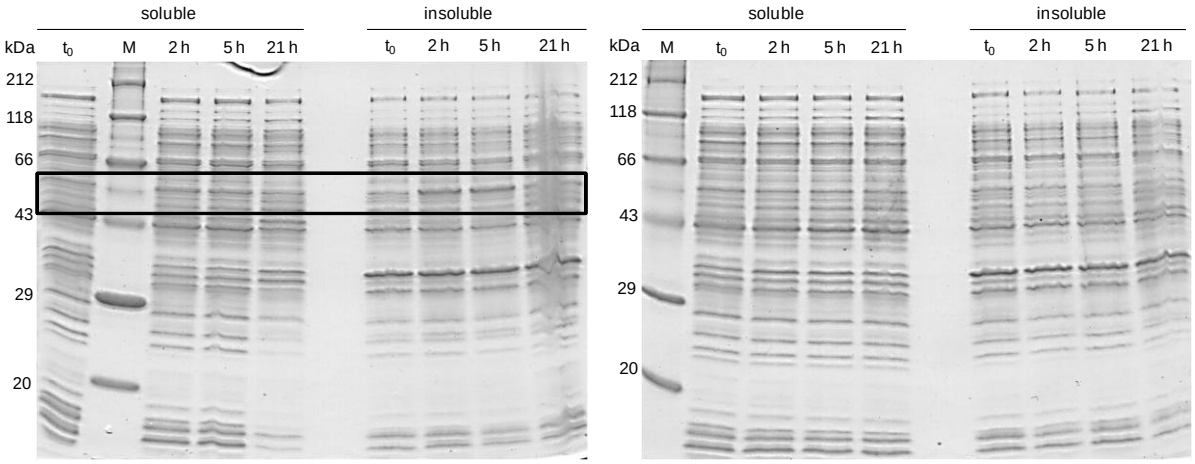


Figure 3.1: SDS polyacrylamide gel of the soluble and insoluble fraction from the expression of $BVMO_{albi1}$ (left, black frame) and empty vector pET28a(+) (right) in *E. coli* BL21(DE3) at 30 °C.
All time specifications in hours after induction (t_0), M: Roti-Mark standard.

In regard to the protein sequence with a $His_{(6)}$-tag, a band at 65.3 kDa was expected for $BVMO_{albi1}$. Only in the insoluble fraction an additional protein band at around 50 kDa was visible for the 2 h and 5 h samples. Correspondingly, a NADPH depletion assay with crude extract and the substrates cyclohexanone, 2-dodecanone, bicyclo[3.2.0]hept-2-en-6-one, progesterone, acetophenone, methyl-*p*-tolyl sulfide and methyl phenyl sulfide was performed but none of the fractions showed any activity as expected from the insoluble expression.

3.1.2.1.1.1 Expression of $BVMO_{albi1}$ with alternative vectors

In the course of a research stay in Groningen (The Netherlands) in the group of Prof. Fraaije, being specialized on flavoproteins, alternative vectors were employed for the expression of $BVMO_{albi1}$. The first choice was the construct pBAD_SUMO. By insertion of genes into this vector, a translational fusion with the small ubiquitin-related modifier (SUMO) is achieved. In eukaryotic cells SUMO is involved in various cellular processes like DNA replication, mitosis and signal transduction. After fusing a target protein to this modifier, its expression and solubility can be improved.[134] The plasmid pBAD_SUMO_$BVMO_{albi1}$ was constructed successfully and subsequently expressed in *E. coli* TOP 10 cells. The putative BVMO could be expressed solubly and even purified albeit with the loss of its flavin cofactor FAD (Appendix Figure 8.2).

An advantage of flavoproteins like BVMOs is their visibility during the purification as they lead to a yellow colorization of the column material due to absorption of their tightly bound FAD cofactor.[69, 93b] Consequently, the purified enzyme solution also displays a yellow color.[69] Due to the loss of the cofactor, no activity with cyclohexanone or 2-dodecanone could be detected with $BVMO_{albi1}$. Additionally, a spectral analysis of the purified protein in the range of 250 nm to 800 nm did not show a flavin peak. As it was possible that $BVMO_{albi1}$

was too diluted and in order to attempt a FAD restoration, it was concentrated. Afterwards, the solution had a yellow color but did not show a flavin peak in the spectrophotometrical analysis. Also no activity with NADPH alone or together with 2-dodecanone or cyclohexanone could be detected.

In section 3.1.2.1.3.6 expression of $BVMO_{albi1}$ together with a putative BVMO from *Candida maltosa* with the other alternative vector pCRE3 will be discussed as well.

3.1.2.1.2 Expression of a putative BVMO from Candida maltosa - *$BVMO_{malto}$*

For *Candida maltosa* the fastest growth with 2-dodecanone was observed indicating the presence of enzymes with good activity, including a participating BVMO, responsible for the metabolism of this ketone.[63, 133] Therefore, investigations were focused on the putative enzymes $BVMO_{malto}$ and CMO. These were identified in the genome sequence of *C. maltosa* (accession: AOGT01000000) from nineteen putative BVMOs and FMOs identified in total from different yeasts.[63]

3.1.2.1.2.1 Expression in eukaryotic systems

To increase the possibility to obtain active protein, it was decided to try the expression of $BVMO_{malto}$ in eukaryotic hosts. Not only because it is more likely to get a gene originating from a yeast expressed in yeast systems, also the possibility of post-translational modifications is given, which can be important for enzymatic activity. Therefore, expression in *Pichia pastoris* X-33 and *Yarrowia lipolytica* was intended.

3.1.2.1.2.1.1 Expression in Pichia pastoris

Pichia pastoris X-33 is a commonly used yeast for the heterologous expression of proteins. Expression rates can be higher than in *E. coli* and high cell densities can be achieved leading to a sufficient yield of protein of interest. Additionally, it is possible to have either secretion of the protein after expression or intracellular expression.

3.1.2.1.2.1.1.1 Intracellular expression

Due to their low stability and the need for cofactor regeneration, it is advantageous for BVMOs to be employed in whole cell biocatalysis. For this purpose, pPICZ_A served as expression vector. In this vector no secretion signal is present so that the protein of interest stays in the cell after expression.

Firstly, the construct pPICZ_A_$BVMO_{malto}$_L111S_L261S could be created successfully. After electroporation of *P. pastoris* X-33 with the linearized construct, transformants were isolated. Through PCRs with genomic DNA and colony PCRs of selected transformants and sequencing, the successful integration of the construct into the genome of *P. pastoris* was verified. With the nine investigated transformants and a pPICZ_A empty vector transformant a test cultivation was carried out. Samples were taken for analysis via SDS-PAGE. However, cell disruption by neither sonication, vortexing with glass beads (according to the EasySelect™ *Pichia* Expression Kit manual, Invitrogen) nor FastPrep led to an efficient cell lysis (data not shown). Still, by analyzing the protein patterns resulting from SDS-PAGE, it could

be concluded that there was no overproduction of a protein, meaning $BVMO_{malto}$ probably was not expressed.

3.1.2.1.2.1.1.2 Expression with subsequent secretion

In most cases, secretion is the method of choice when using *P. pastoris* as it leads to protein in high purity because this yeast naturally secretes only a few proteins into the medium. Therefore, the protein of interest just needs to be concentrated after removal of the cells.

The construct pPICZ_α_C_$BVMO_{malto}$_L111S_L261S was obtained through classical cloning. This construct was linearized and subsequently used for transformation of *P. pastoris* X-33. Through PCRs with genomic DNA of selected transformants, sequencing of resulting PCR products and yeast colony PCRs, the integration of the construct into the genome of *P. pastoris* was verified. With five transformants and a pPICZ_A_CAL-A transformant (positive control for expression), a test cultivation was performed. Medium samples were taken for analysis via SDS-PAGE and Western-Blot. However, just the expression of CAL-A could be detected with both methods.

3.1.2.1.2.1.1.2 Expression in Yarrowia lipolytica

Yarrowia lipolytica is a yeast that is well suited for the production of heterologous proteins. Compared to other yeasts, the handling of *Y. lipolytica* is easier as no carbon source or inductor need to be added periodically as it is the case for instance for *P. pastoris* with methanol. Additionally, *Y. lipolytica* has a highly efficient secretion pathway. For this project, this yeast is advantageous because it is closer related to *Candida*, compared to *Pichia*.[135]

Like discussed for *P. pastoris* (see Chapter 3.1.2.1.2.1.1.2), expression with subsequent secretion is the method of choice when producing proteins with *Y. lipolytica*. Thus, $BVMO_{malto}$ was cloned into the vector pSKI, which is an integrative shuttle vector for expression in *Y. lipolytica*, containing the N-terminal Lip2 prepro secretion signal.[136] Employing SLiCE, $BVMO_{malto}$_L111S_L261S was successfully cloned with and without N-terminal $His_{(6)}$-tag. After linearization of the plasmids, they were used to transform *Y. lipolytica* Po1f. After the integration of the constructs into the genome of this yeast was verified, selected transformants were cultivated. No activity with the used substrates and no expression of $BVMO_{malto}$ could be detected, neither in the untreated supernatant samples, nor after concentration.

3.1.2.1.3 Expression in E. coli

Expression in *E. coli* is much easier to perform than in eukaryotic systems. Production of proteins originating from eukaryotes can be difficult, though. However, successful expression of eukaryotic proteins in *E. coli* has been reported and there are many ways to influence the expression in this bacterium.

3.1.2.1.3.1 Influence of different expression conditions

As small changes in the expression conditions can make huge differences in the amount of correctly folded and thus active enzyme, various conditions were examined.

The influence of temperature was already investigated in my diploma thesis.[63] *E. coli* BL21(DE3) carrying the construct pET28a(+)_$BVMO_{malto}$ was cultivated in LB medium at 15, 30 and 37 °C.[63] Only at 37 °C, overexpression of $BVMO_{malto}$ could be observed. However, respective protein bands only appeared in the insoluble fractions.[63]

Next, the expression of $BVMO_{malto}$ at varying IPTG concentrations was investigated. *E. coli* BL21(DE3) transformants were cultivated in LB medium at 20 °C and 37 °C. Gene expression was induced by the addition of IPTG to a final concentration of 0.1, 0.4 or 1 mM. For all the investigated conditions, only a very low and insoluble expression of the desired protein could be determined.

3.1.2.1.3.2 Investigations about the codon reassignment in Candida *spec.*

The investigated *Candida* species show a codon reassignment meaning that for the base triplet CTG (CUG) during the translation serine instead of leucine is incorporated into the growing peptide chain. This could also influence activity, stability and solubility when translated differently in heterologous hosts so that leucine is incorporated at the respective positions.[137] For this reason, the gene sequence of $BVMO_{malto}$ was investigated for the existence of this codon in my diploma thesis.[63] This putative BVMO actually contains two CTG codons coding for serine 111 and serine 261.[63] Thus, site-directed mutagenesis was performed to exchange both triplets, leading also in *E. coli* to the native serines. After obtaining the construct pET28a(+)_$BVMO_{malto}$_L111S_L261S, expression in *E. coli* BL21(DE3) at 37 °C was investigated. The mutagenesis did not affect the solubility of $BVMO_{malto}$. Only insoluble protein without activity could be obtained.

3.1.2.1.3.3 Expression of codon optimized $BVMO_{malto}$

Due to the poor expression of $BVMO_{malto}$, a sequence with optimized codon usage for *E. coli* was ordered as a synthetic gene also containing both amino acid exchanges L111S and L261S (see Chapter 3.1.2.1.3.2) here referred to as $BVMO_{malto\_opt}$. The expression of the codon opti-

mized version of the putative BVMO resulted in a much higher amount of protein. (Figure 3.2).

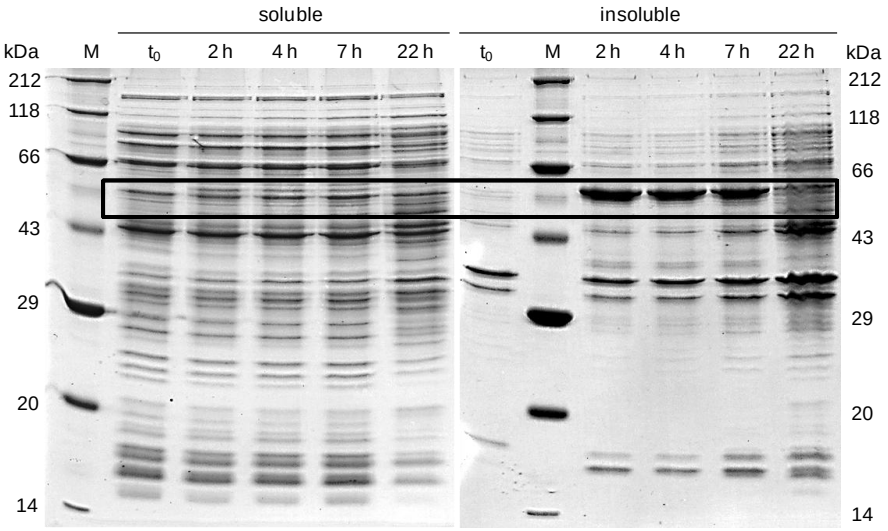


Figure 3.2: SDS polyacrylamide gel of the soluble and insoluble protein fractions of the expression of $BVMO_{malto\_opt}$ (black frame) in *E. coli* BL21(DE3) at 30 °C.
All time specifications in hours after induction (t_0), M: Roti-Mark standard.

However, the solubility was still very low and there was no activity with cyclohexanone or 2-dodecanone detectable in the NADPH depletion assay.

3.1.2.1.3.4 Coexpression of chaperones

Molecular chaperones can improve protein folding by preventing the aggregation of newly translated peptides or refolding already misfolded proteins while consuming ATP in the process. The major chaperones of *E. coli* are the DnaK-DnaJ-GrpE and GroEL-GroES chaperone systems, which are part of the heat shock response mediated by the sigma factor 32 (σ^{32}).[138] Overexpression of chaperones can significantly improve soluble expression of proteins and has already been employed successfully for the expression of Baeyer-Villiger monooxygenases.[139] For this work, the commercial TaKaRa Chaperone Set was used, which contains five plasmids with different combinations of chaperone systems (Table 6.7).[140] To employ the chaperones, *E. coli* BL21(DE3) was cotransformed with the vector construct pET28a(+)_$BVMO_{malto_opt}$ and one of the TakaRa chaperone plasmids pKJE7, pG-KJE8, pTf16, pGro7 or pG-Tf2, respectively. Then, coexpression was carried out (Figure 3.3).

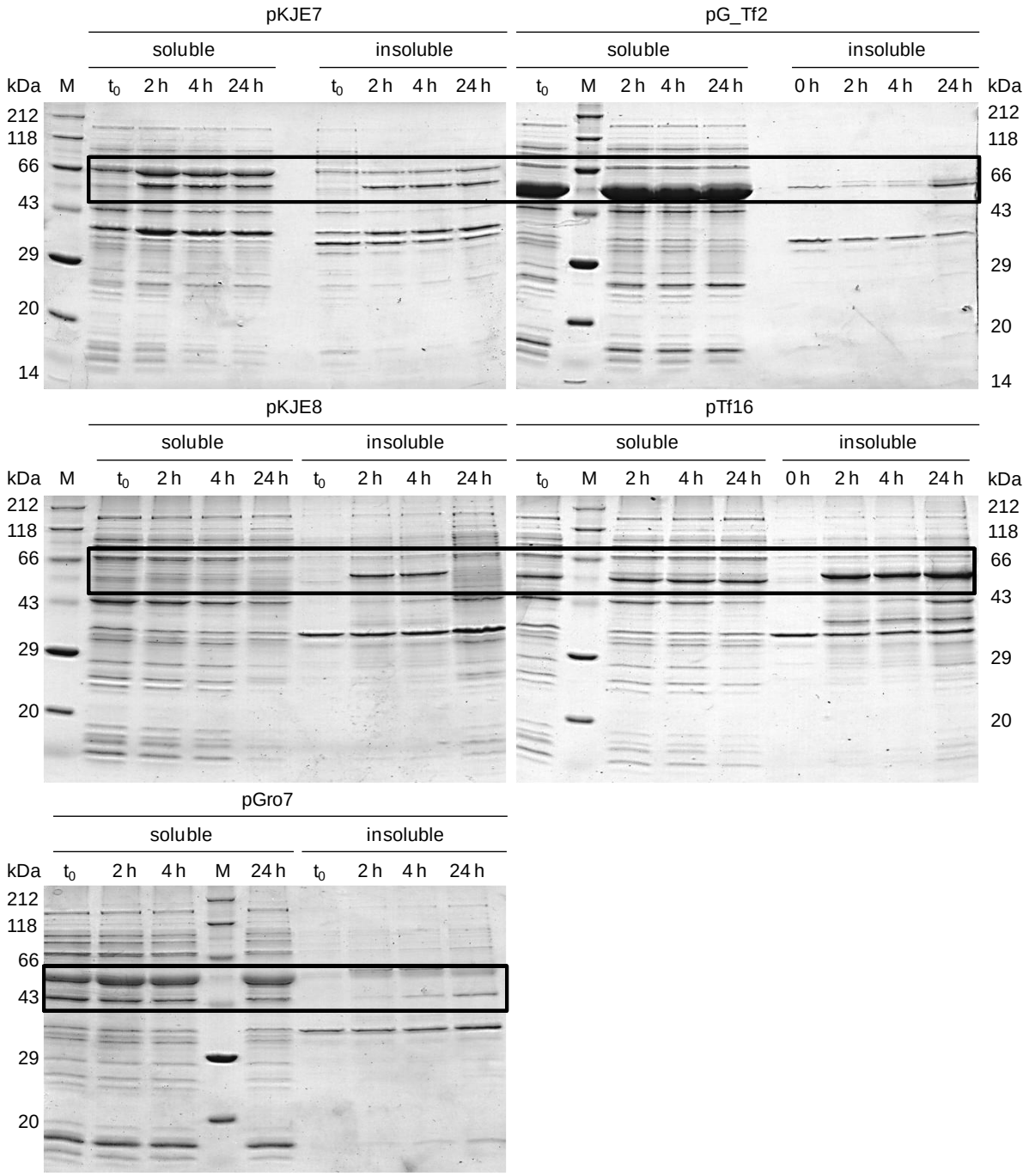


Figure 3.3: SDS polyacrylamide gel of the soluble and insoluble protein fractions of the coexpression of $BVMO_{malto\_opt}$ (black frame) and chaperones in *E. coli* BL21(DE3) at 30 °C.
All time specifications in hours after induction (t_0), M: Roti-Mark standard.

The SDS-PAGE revealed that the percentage of soluble $BVMO_{malto\_opt}$ was increased by the coexpression of the chaperone plasmid pKJE7. There was no measurable activity with cyclohexanone or 2-dodecanone in the NADPH depletion assay, though.

3.1.2.1.3.5 Addition of Riboflavin to the culture medium

It has been reported that addition of riboflavin can lead to soluble and active expression of the flavin-containing BVMOs.[99] Thus, $BVMO_{malto\_opt}$ was expressed with addition of 1µg/mL riboflavin to the cultures. The solubility and activity were not improved by this approach.

3.1.2.1.3.6 Expression of $BVMO_{malto}$ in alternative vectors

In accordance with the COST action program, some experiments could be carried out in the group of Prof. Dr. Marco Fraaije in Groningen (The Netherlands). This group possesses the alternative vectors pBAD_SUMO and pCRE3. The expression of $BVMO_{malto}$, $BVMO_{malto\_opt}$ and $BVMO_{albi1}$ was investigated using these plasmids.

Firstly, pBAD_SUMO was used. For further information about this construct and the expression of $BVMO_{albi}$ using it, see Chapter 3.1.2.1.1.1. After cloning of the putative BVMO of *C. maltosa* into this vector, expression was carried out in *E. coli* TOP10. For neither $BVMO_{malto}$ nor $BVMO_{malto\_opt}$ overexpression could be seen.

Then, cloning into pCRE3 was performed. This vector is used to create a translational fusion with a thermostable phosphite dehydrogenase (PTDH, codon optimized) which worked well with many BVMOs and increased their solubility.[71, 141] Additionally, PTDH can be used for NADPH regeneration leading to a fusion protein with internal cofactor recycling. Due to the inherent time constraints of the thesis, only the construct pCRE3_$BVMO_{malto_opt}$ could be obtained. Expression in *E. coli* BL21(DE3) resulted in no detectable amount of protein, the fraction diplaying also no activity. Additionally, purification of the fusion proteins was tried but did not result in any purified protein.

3.1.2.1.3.7 Whole cell biocatalysis with $BVMO_{malto}$ in E. coli *BL21(DE3)*

It is possible that the activities in the cell extracts of the previously described strategies were too low to be detected by the NADPH depletion assay since the consumption of the cofactor was only measured for two minutes. In contrast, employing whole cell biocatalysis enables incubation of the enzyme together with the substrate for a much longer time. Therefore, lower activities can be detected as well, using a GC. Furthermore, BVMOs in some cases just retain their activity and stability inside of cells. Therefore, whole cell biocatalysis was performed.

In a first approach, *E. coli* BL21(DE3) cells harboring the two constructs pet28a(+)_$BVMO_{malto_opti}$ and pKEJ7 (DnaK-DnaJ-GrpE) were utilized as coexpression with these plasmids resulted in the best expression of the putative BVMO from all the tested conditions (see Chapter 3.1.2.1.3.4). This resulted in neither substrate conversion nor product formation with 2-dodecanone or cyclohexanone.

In a different approach, *E. coli* BL21(DE3) cells with just the plasmid pet28a(+)_BVMO$_{malto_opti}$ were cultivated in a low temperature range 12-20 °C. Cells obtained by this procedure were used for biocatalysis with 5 mM substrate (cyclohexanone, bicyclo[3.2.0]hept-2-en-6-one, 2-dodecanone, methyl phenyl sulfide, methyl *p*-tolyl sulfide) and an equimolar amount of glucose for NADPH regeneration (Figure 3.4).

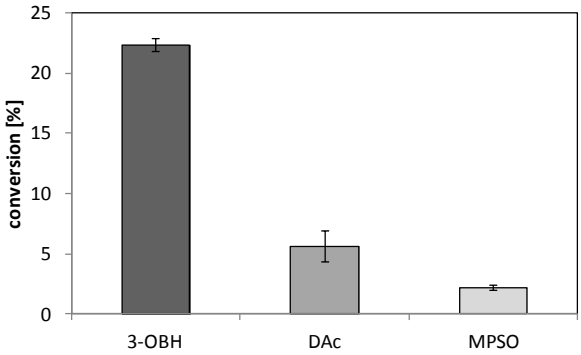


Figure 3.4: Conversion and products formed in whole cell biocatalysis with BVMO$_{malto}$.
3-OBH: 3-oxabicyclo[3.3.0]oct-6-en-2-one, DAc: decyl acetate, MPSO: methyl phenyl sulfoxide.

In total three products could be detected, namely decyl acetate, 3-oxabicyclo[3.3.0]oct-6-en-2-one and methyl phenyl sulfoxide. The best conversion of 22.4% could be achieved with bicyclo[3.2.0]hept-2-en-6-one. The conversion of 2-dodecanone and methyl phenyl sulfide was rather low, showing values of 5.6% and 2.2%, respectively. For the other substrates no product formation could be detected.

3.1.2.2 Expression of putative *Candida maltosa* monooxygenase CMO

3.1.2.2.1 Influence of different expression conditions

In my diploma thesis, a first experiment to investigate the influence of temperature on the expression of *Candida* monooxygenase (CMO) was carried out.[63] From the temperatures tested (15, 30, 37 °C), the cultivation at 37 °C resulted in the best expression of this putative monooxygenase.[63] However, almost all of the protein could be found in the insoluble fraction. Only after two hours there was a tiny amount of soluble CMO present.

As discussed earlier, in the investigated *Candida spec*. there is a codon reassignment (see Chapter 3.1.2.1.3.2). In CMO one CTG (CUG) codon could be identified coding for a serine at position 462 in its amino acid sequence. This triplet was mutated to TCT by utilizing site directed mutagenesis enabling an equal translation in *E. coli*, which did not affect the solubility any further.

3.1.2.2.2 Coexpression of chaperones

After the expression of CMO in *E. coli* BL21(DE3) resulted in just a small amount of soluble protein, it was intended to increase the solubility by the utilization of chaperones. Thus, co-transformations and subsequently coexpressions of pET28a(+)_CMO_L462S with one of the five TakaRa chaperone plasmids pKJE7, pG-KJE8, pTf16, pGro7 or pG-Tf2 were performed, respectively. There was no increase in soluble protein expression with the chaperone systems. Therefore, the two disruption methods FastPrep24® and supersonication were compared. The supersonication resulted in a much better cell disruption. As a result, much more protein and even soluble CMO_L462S were detected. Consequently, even without chaperones CMO_L462S was soluble after 4 h of cultivation. However, coexpression with pG-Tf2 resulted in an improved expression (Figure 3.5).

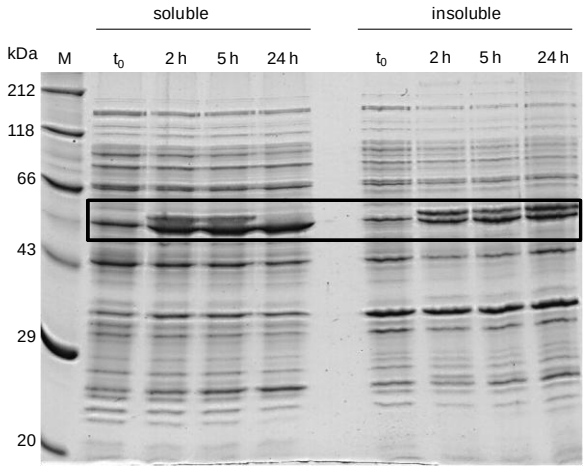


Figure 3.5: SDS polyacrylamide gel of the soluble and insoluble protein fraction of the coexpression of CMO_L462S and the chaperone pG-Tf2 in *E. coli* BL21(DE3) at 30 °C.
All time specifications in hours after induction (t_0), M: Roti-Mark standard.

By combining the positive effects of the coexpression with chaperones from pG-Tf2 and of supersonication as disruption method, a sufficient amount of soluble CMO could be obtained after two and five hours of cultivation. However, CMO_L462S did not show any activity in the NADPH depletion assay with cyclohexanone or 2-dodecanone.

3.1.2.3 Investigation of fingerprint motifs

In the course of my diploma thesis, a multiple sequence alignment consisting of the protein sequences of the nine most promising putative yeast BVMOs (including $BVMO_{albi1}$ and $BVMO_{malto}$) and the BVMOs $CHMO_{Acineto}$, CAMO, PAMO and HAPMO was employed.[64] From this, two new Baeyer-Villiger monooxygenase fingerprints sequence motifs derived, Dx[I/L][V/I]xxTG[Y/F] and [G/D][P/A]xxYxxxxxxxxPN[L/M][W/F]xxxG, designated "fingerprint 3" and "fingerprint 4", respectively.[53b, 63] In order to further investigate these motifs, 55 BVMOs originating from different organisms were compared in a multiple sequence alignment (results are summarized in Table 3.2).

Table 3.2: In a multiple sequence alignment, consisting of 55 BVMO protein sequences, investigated fingerprint motifs, identified deviations and resulting BVMO sequence motifs.

Initial motif[64]	**Deviations**	**Resulting fingerprint motif**
Dx[I/L] [V/I]xxTG[Y/F] ("fingerprint 3")	[I/L] three (V)	DxxxxxG[Y/F]
	[V/I] one (A)	
	[T] three (2x V, 1x L)	
[G/D] [P/A] xxYxxxxxx xxPN[L/M] [W/F]xxxG ("fingerprint 4")	[G/D] 13 (5xR, 1xA, 2xE, 2xQ, 3xT)	PNxxxxxP
	[P/A] ten (1xG, 7xI, 2xL)	
	[Y] 16 (1xW, 4xH, 3xF, 2xA, 4xT, 1xM, 1xV)	
	[L/M] 13 (2x W, 11x F)	
	[W/F] eight (7x L, 1x Y)	
	[G] seven (4x A, 3x V)	

For the motif Dx[I/L][V/I]xxTG[Y/F] ("fingerprint 3"), deviations were identified at three positions. At the third position instead of isoleucine and leucine, valine is also possible. In one sequence, the fourth amino acid of the motif was an alanine, for the others it was always valine or isoleucine. Threonine at the seventh spot turned out to be replaceable by valine or leucine. Thus, the less conserved motif DxxxxxG[Y/F] was determined even though the more conserved version could be found without any changes for 49 sequences. The second identified motif [G/D][P/A]xxYxxxxxxxxPN[L/M][W/F]xxxG ("fingerprint 4") could be found in 25 BVMO sequences as well. Many deviations were determined though, so that the sequence was reduced to PNxxxxxP.

3.1.2.4 Investigations of putative BVMOs from *Yarrowia lipolytica*

Yarrowia lipolytica also showed a conversion of 2-dodecanone to decyl acetate, indicating BVMO activity (see Chapter 3.1). Therefore, it was intended to investigate Baeyer-Villiger monooxygenases from this yeast as well.

3.1.2.4.1 Identification

By using the protein BLAST, nine sequences homologous to CHMO from *Acinetobacter sp.* NCIMB 9871 ($CHMO_{Acineto}$) were identified in the annotated proteins from the genome of *Y. lipolytica*, having sequence identities of around 20% to this well described monooxygenase and PAMO (Figure 3.6).[69, 106b]

	CHMO_Acineto	PAMO	YMOB	YMOI	YMOE	YMOH	BVMO_malto	YMOA	YMOC	YMOD
CHMO_Acineto		39.2%	22.2%	24.2%	21.9%	21.9%	21.7%	18.5%	18.0%	18.2%
PAMO	39.2%		21.9%	23.6%	22.4%	21.7%	21.0%	20.3%	19.5%	20.9%
YMOB	22.2%	21.9%		56.8%	46.5%	49.0%	35.3%	17.7%	18.2%	19.4%
YMOI	24.2%	23.6%	56.8%		48.2%	49.9%	39.0%	21.0%	19.8%	21.2%
YMOE	21.9%	22.4%	46.5%	48.2%		46.8%	33.1%	18.8%	19.2%	20.6%
YMOH	21.9%	21.7%	49.0%	49.9%	46.8%		35.9%	20.7%	21.9%	21.2%
BVMO_malto	21.7%	21.0%	35.3%	39.0%	33.1%	35.9%		17.9%	17.0%	18.5%
YMOA	18.5%	20.3%	17.7%	21.0%	18.8%	20.7%	17.9%		60.3%	57.9%
YMOC	18.0%	19.5%	18.2%	19.8%	19.2%	21.9%	17.0%	60.3%		55.4%
YMOD	18.2%	20.9%	19.4%	21.2%	20.6%	21.2%	18.5%	57.9%	55.4%	
YMOG	19.1%	21.6%	19.0%	20.7%	21.5%	21.2%	18.1%	62.8%	62.3%	72.8%
YMOF	17.1%	18.7%	18.6%	21.6%	18.9%	19.6%	18.1%	46.6%	43.5%	44.3%

Figure 3.6: Sequence alignment with the identified proteins from *Y. lipolytica* and $BVMO_{malto}$, $CHMO_{Acineto}$ and PAMO.

Four of the sequences (YMOB, E, H, I) show an identity of more than 30% to $BVMO_{malto}$. Surprisingly, some of the *Y. lipolytica* proteins just have around 20% sequence identity to each other. Thus, they probably display different properties, i.e. substrate spectrum and activity. The discovered sequences were subsequently identified as possible BVMOs by the presence of BVMO fingerprints and Rossmann fold motifs and designated YMOA-I (*Yarrowia* monooxygenase A-I, Table 3.3).[34a, 58a, 62]

Table 3.3: Overview of identified putative BVMOs from *Y. lipolytica*.

Protein	Accession no.	Length [aa]	Molecular weight [kDa][a]
YMOA	XM_503819	660	74.1 (76.3)
YMOB	XM_503445	536	60.6 (62.8)
YMOC	XM_499840	624	69.8 (72)
YMOD	XM_502954	636	71.1 (73.3)
YMOE	XM_503686	511	58 (60.2)
YMOF	XM_503818	691	78.1 (80.3)
YMOG	XM_504212	614	68.8 (71)
YMOH	XM_505929	496	56.3 (58.5)
YMOI	XM_505928	497	56.2 (58.4)

[a]: Molecular weight with $His_{(6)}$-tag in parentheses

As described in the introduction, six putative BVMO sequences in *Y. lipolytica* were already identified by Mascotti *et al*.[35] However, these only included YMOB, D, E, F and I, but not YMOA, C, G and H. YMOA and YMOB were investigated in the course of the master thesis of Sven Bordewick, which was co-supervised by me.[69]

3.1.2.4.2 Homologous expression of YMOA and YMOB in Yarrowia lipolytica

Homologous expression of YMOA and YMOB was intended firstly with the strain *Y. lipolytica* Po1f, as this seemed a promising strategy for high expression levels of active protein.[69]

FastCloning was employed to clone both genes from genomic DNA of *Y. lipolytica* strain 63 into the vectors pSKI and pUC_INTB to enable expression with and without subsequent secretion, respectively.[69] After successful transformation into *Y. lipolytica* Po1f, expression was carried out, which resulted in no detectable production of neither YMOA nor YMOB, proven by SDS-PAGE and NADPH depletion assay. The strategy of homologous expression in *Y. lipolytica* was abandoned at this point due to positive results obtained with heterologous expression in *E. coli*, which was investigated in parallel (see next section).[69]

3.1.2.4.3 Heterologeous expression in E. coli *BL21(DE3)*

For expression in *E. coli* BL21(DE3), the genes coding for YMOA and B were subcloned into the vector pET28a(+).[69] After investigating numerous expression conditions (temperatures, coexpression of chaperones with varying inducing parameters), sufficient enzyme production could be achieved. YMO was found to be well expressed at 25 °C for 24 hours (Figure 3.7).

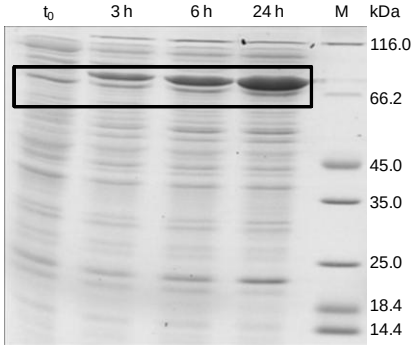


Figure 3.7: Soluble expression of YMOA (black frame) at 25 °C (adapted from Bordewick).[69]
Time specification in hours (h) after induction (t_0), M: Marker**.**

For YMOB, the best yield was found at 30 °C using 2 mg/mL L-arabinose for induction of the chaperones from pKEJ7 30 min before enzyme induction for eight hours (Figure 3.8).

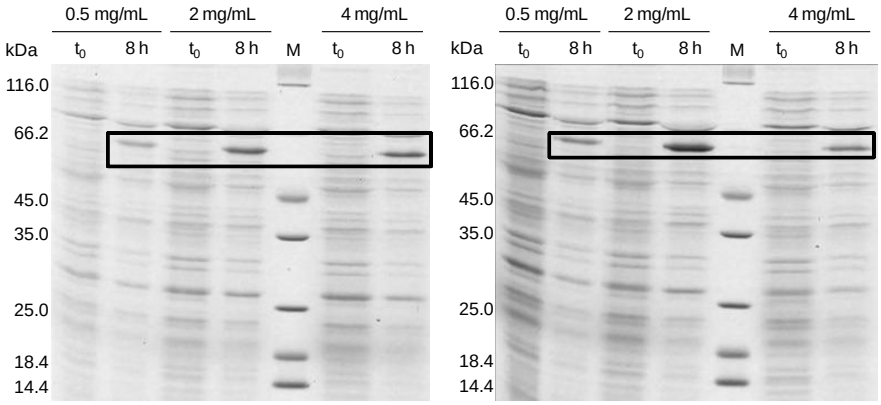


Figure 3.8: Optimization of induction time and inducer concentration for pKJE7 (adapted from Bordewick).[69]
Left: chaperones induced at start of cultivation, right: chaperones induced 30 min before induction of YMOB, Time specification in hours (h) after induction (t_0), M: Marker.

After the best expression conditions for YMOA and YMOB were determined, the remaining seven putative *Y. lipolytica* BVMOs, designated YMOC-I, were supposed to be analyzed as well. Employing classical cloning, the genes for YMOC-H could be successfully cloned into the pET28a(+) vector for expression in *E. coli* BL21(DE3). However, in some cases deviations from the genome sequences were spotted. YMOD contained five (L9S, A12P, L14S, P263L, K482E), YMOF two (P319L, I328T) and YMOH three (T66A, T92A, L247P) mutations. Further investigation is needed to determine, whether these are mutations or strain specific adaptations as the same mutations could be identified in all sequenced transformants.

Due to time limitations, expression had to be conducted with the deviating constructs. In a fast approach, in which three colonies per construct were investigated, expression was carried out at 30 °C for 6 hours in ONC tubes with 5 mL LB. SDS-PAGE was used to visualize the soluble and insoluble protein fractions (Figure 3.9).

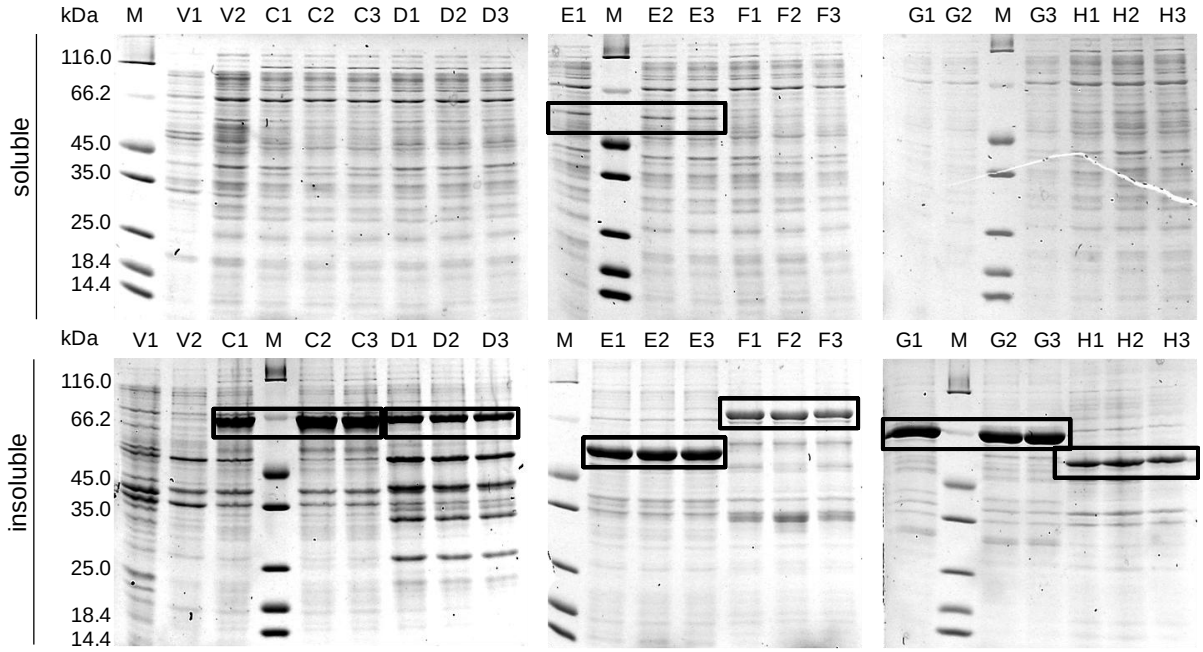


Figure 3.9: SDS polyacrylamide gel of 7/OD cultivation samples of expression of *Y. lipolytica* BVMOs YMOC-H.
V1: empty vector control, V2*: E. coli* BL21(DE3) w/o induction, M: Roti-Mark standard.

Expression of YMOC-H resulted mostly in the production of insoluble protein. Only traces of YMOE were detected in the soluble fraction. However, in all cases the overall level of expression was high (YMOC, D, F and H) up to very high (YMOE and G).

Next, soluble expression had to be optimized. Thus, the temperature was lowered to 25 °C and in parallel, coexpression of the TaKaRa chaperone plasmid pKEJ7 (DnaK-DnaJ-GrpE) was performed (Figure 3.10).

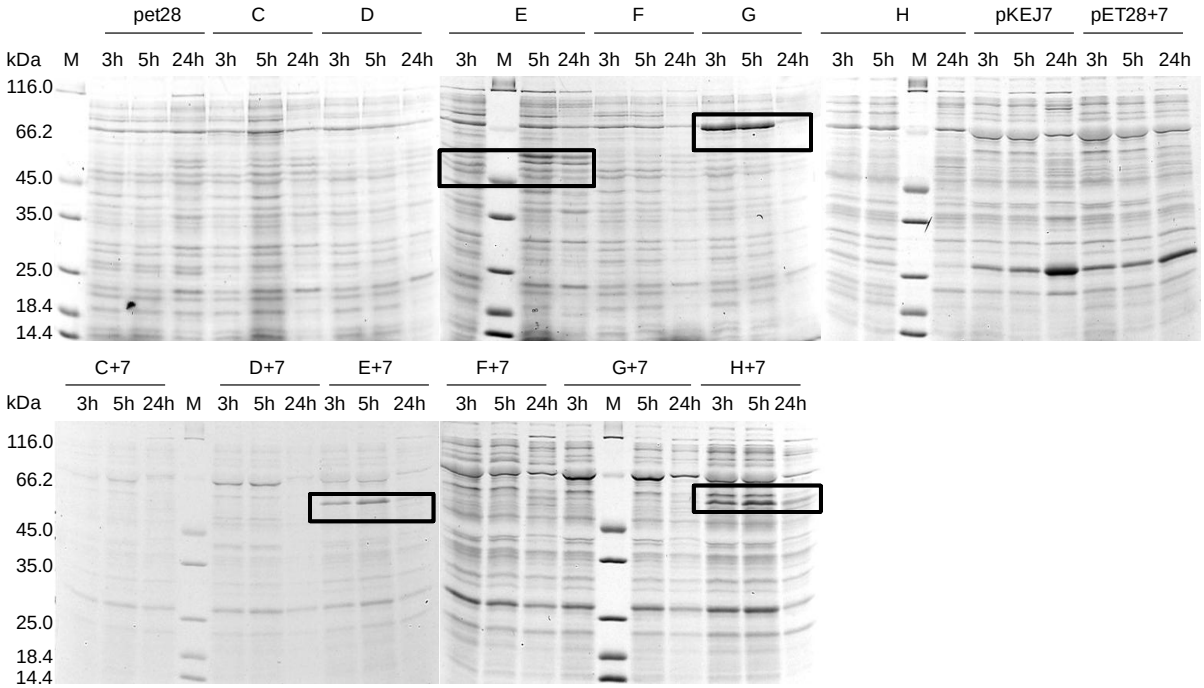
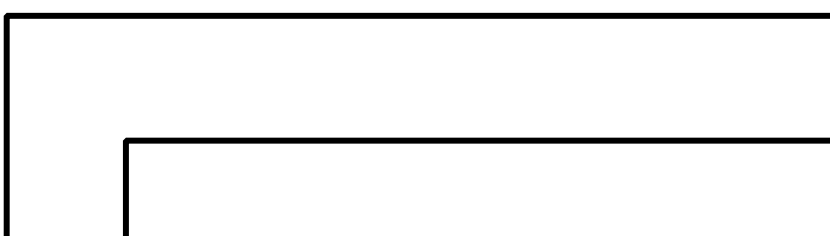


Figure 3.10: Expression optimization of YMOC-H without (pET28, C-H) and with chaperone plasmid pKEJ7 (pKEJ7, pET28/C-H+7).
The gels were visualized with TCE. Time specification in hours (h) after induction (t_0), M: Roti-Mark standard.

Simply by the decrease in temperature, YMOE and G could be obtained in soluble form. By coexpression of DnaK-DnaJ-GrpE chaperones from pKEJ7, additionally to YMOE, YMOH was present in the soluble protein fraction. Expression with and without the coexpression of chaperones at 20 °C did not lead to an improvement compared to 25 °C (data not shown).

3.1.2.4.4 Activities of YMOs in cell extract

With a set of fourteen substrates, *E. coli* cell extracts from expression under the best conditions determined for the YMOs were screened for activity (Scheme 3.2, Table 3.4).).[69]



Scheme 3.2: Overview of screened substrates (from Bordewick).[69]

Table 3.4: Substrate screening with cell extract for YMOA and YMOB with NADPH (from Bordewick).[69]

Substrate	YMOA Activity [U/mg total protein]	YMOB Activity [U/mg total protein]
Aliphatic ketones		
Acetone	n.c.	n.q.
2-Octanone	n.c.	n.q.
2-Dodecanone	n.c.	0.006 ± 0.001
Cyclic, non-aromatic ketones		
Cyclohexanone	n.c.	n.q.
rac-Bicyclo[3.2.0]hept-2-en-6-one	n.c.	n.c.
Aromatic ketones		
Acetophenone	n.c.	n.c.
4-Hydroxyacetophenone	n.c.	n.c.
Methoxyphenylacetone	n.c.	n.q.
Sulfides		
Methyl phenyl sulfide (MPS)	0.103 ± 0.003	n.c.
Methyl *p*-tolyl sulfide (MTS)[a]	0.081 ± 0.009	0.004 ± 0.001
L-Methionine	n.c.	n.c.
Sulfoxides		
rac-Methyl phenyl sulfoxide (MPSO)	0.086 ± 0.014	n.q.
Dimethyl sulfoxide (DMSO)	0.055 ± 0.014	n.c.
N-Heterocycle		
Indole	n.c.	n.c.

n.c.: not converted: Activity below limit of detection (see Chapter 6.9.1.1 for definition)

n.q.: not quantifiable: Activity below limit of quantification (see Chapter 6.9.1.1 for definition)

[a]: 1 mM instead of 2 mM substrate because the low solubility was interfering with photometric measurements

Interestingly, YMOA converted none of the used ketones. However, it showed sulfoxidation activity with two sulfides and also with two sulfoxides with comparable activities.[69] YMOB accepted two substrates (2-dodecanone and MTS).[69] Furthermore, it showed activities with five other substrates, which were below the limit of quantification and thus were not classified as substrates for YMOB. By using the cell extracts of YMOC-H in the NADPH depletion assay, no activity could be detected with any of the used substrates.

In several attempts to purify the BVMOs, YMOA lost its FAD cofactor and displayed a very low stability and YMOB did not bind significantly to the Ni^{2+} column.[69] Due to these findings, purification was abandoned for YMOC-H and all biocatalytic reactions were performed with whole cells expressing the Baeyer-Villiger monooxygenases YMOA-H.[69]

3.1.2.4.5 Whole cell biocatalysis

3.1.2.4.5.1 Mutational study of YMOA

Using a homology model of YMOA, created with YASARA and subsequently validated by the web tool MolProbity, a rational mutational study was performed to further investigate this BVMO, which showed apparently only sulfoxidation activity.[69, 142] Based on this model as well as structure- and sequence alignments consisting of BVMOs with proven sulfoxidation activity (except for MekA) together with the knowledge from literature, the importance of residues for the activity towards ketones, discrimination between sulfoxide/sulfone formation and stereoselectivity was intended to be explored (Appendix Figure 8.3,Figure 8.4 and Figure 8.5).[69, 130] Residues at the FAD- and NADPH-binding sites and the active site were targeted for mutagenesis leading to the generation of seven enzyme variants. Both the wild type (WT) and the enzyme variants of YMOA were used in whole cell biocatalysis. This was performed in deep well plates sealed with an oxygen permeable membrane for 4 h at 25 °C.[69] Varying concentrations of substrates were used; 5 mM of sulfides MTS and MPS (Table 3.5, Figure 3.11), 10 mM of sulfoxides MTSO and MPSO (Figure 3.12) and a substrate mix containing 2 mM of ketones 2-dodecanone, cyclohexanone and acetophenone, together with equimolar amounts of glucose for cofactor recycling, respectively.[69]

Table 3.5: Summary of YMOA whole cell biocatalysis with the aromatic sulfides (5 mM) and sulfoxides (10 mM, from Bordewick).[69]

Variant	Methyl *p*-tolyl sulfide (MTS)						*rac*-Methyl *p*-tolyl sulfoxide (MTSO)			
	Conversion [%]			SO_2/SO	ee [%]		Conversion [%]	ee [%]		*E*
	SO	SO_2	Total				SO_2			
WT	16	15	31	0.9	95	(*S*)	100	n.d.		-
K274R_R275S	17	11	28	0.6	96	(*S*)	28	37	(*S*)	3
R367K	6	1	7	0.2	91	(*S*)	2	3	(*S*)	1
V121T	7	26	33	3.5	93	(*S*)	100	n.d.		-
Y477P	10	9	19	1.0	15	(*S*)	49	20	(*S*)	2
Y479G	10	27	37	2.7	>99	(*R*)	98	n.d.		-
C480F	1	0	1	0.0	n.d.		1	1	(*S*)	1
A483L	1	0	1	0.1	n.d.		1	< 1	(*S*)	1
Variant	**Methyl phenyl sulfide (MPS)**						***rac*-Methyl phenyl sulfoxide (MPSO)**			
	Conversion [%]			**SO_2/SO**	**ee [%]**		**Conversion [%]**	**ee [%]**		***E***
	SO	**SO_2**	**Total**				**SO_2**			
WT	5	24	29	4.9	1	(*R*)	63	52	(S)	3
K274R_R275S	9	10	19	1.1	27	(*R*)	1	4	(*R*)	5
R367K	2	1	3	0.2	n.d.		1	1	(*R*)	2
V121T	0	55	55	only SO_2	n.d.		36	15	(*R*)	2
Y477P	16	5	21	0.3	40	(*R*)	31	10	(*R*)	2
Y479G	1	50	51	68.1	n.d.		14	14	(*R*)	10
C480F	1	0	1	0.0	n.d.		1	2	(*R*)	2
A483L	1	0	1	0.0	n.d.		0	4	(*R*)	3

Conversion SO: Conversion to the sulfoxide (MTSO or MPSO, respectively)

Conversion SO_2: Conversion to the sulfone ($MTSO_2$ or $MPSO_2$, respectively)

n.d.: The enantiomeric excess could not be accurately determined for samples with very low sulfoxide concentrations

SO_2/SO: Ratio of sulfone to sulfoxide formation

ee: Enantiomeric excess

E: Enantiomeric ratio = $\ln[(1 - C)(1 - ee_s)]/\ln[(1 - C)(1 + ee_s)]$ | C: Conversion; ee_s: enantiomeric excess of substrates

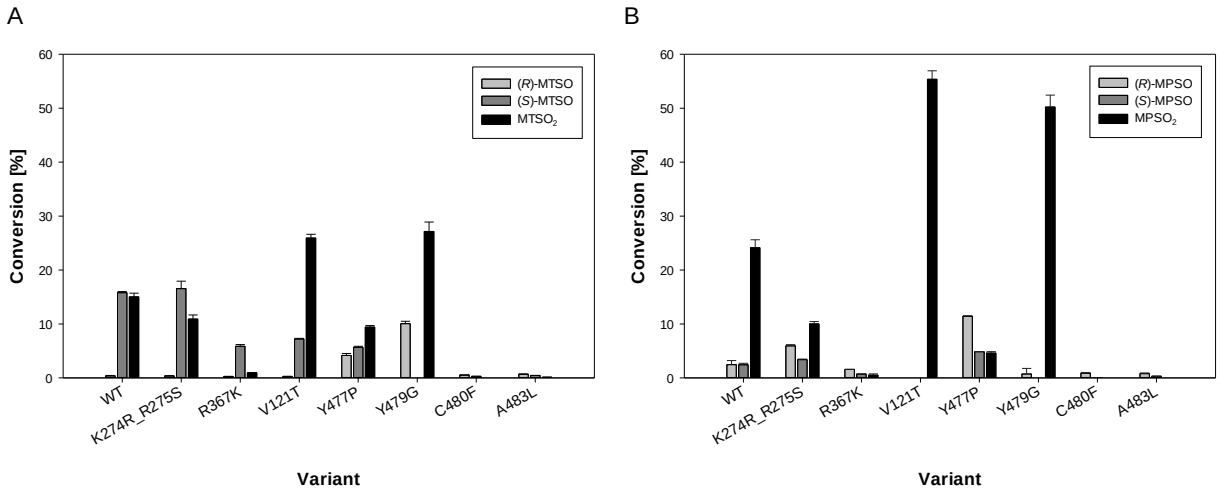


Figure 3.11: YMOA whole cell biocatalysis with 5 mM MTS (left) and 5 mM MPS (right) for 4 h at 25 °C. (from Bordewick).[69]
The conversion to the different products is shown.

In the formation of (*S*)-MTSO, wild-type YMOA showed a high enantiomeric excess of 95% ee, while MPS was converted practically nonselectively (1% ee).[69] Sulfoxide yields were rather low, only 16% for MTSO and 5% for MPSO. The only variants which produced significant amounts of (*R*)-MTSO were Y477P and Y479G (4% and 10%, respectively). However, Y477P retained its (*S*)-selectivity with 15% ee, even though the formation of (*S*)-MTSO was decreased by 65%. In contrast, a complete inversion of enantioselectivity was achieved with Y479G to over 99% (*R*)-MTSO while retaining 63% of MTS conversion observed for the WT. K274R_R275S and Y477P showed an improved enantiomeric excess with (*R*)-MPSO of 27% ee and 40% ee, respectively. Additionally, they displayed increased total MPSO yields. The only variant with a higher sulfoxide formation than the WT was K274R_R275S with MTS. Even though, the improvement was rather small (17% vs. 16%), it presented the highest sulfoxide yields obtained. With V121T and Y479G higher conversions with both substrates were achieved, from which the almost doubled performance from 29% to 55% and 51% with MPS is especially noteworthy. This was mainly caused by doubling of the sulfone formations. Consequently, the sulfone/sulfoxide ratios (SO_2/SO) were increased, resulting in a more than 10-fold higher SO_2/SO ratio of 68 for Y479G (4.9 for WT). With V121T, no sulfoxide was detected, leading to the highest SO_2/SO ratio and the highest conversion of 55%. V121T and Y479G also showed a decreased sulfoxide yield with MTS and therefore the sulfone/sulfoxide ratio was increased from 0.9 to 3.7 and 2.7, respectively. The remaining variants showed lower conversions and sulfone formations with both substrates.[69] The mutations affected the sulfone formation in kinetic resolutions with the sulfoxides MTSO and MPSO as well (Table 3.5, Figure 3.12).[69]

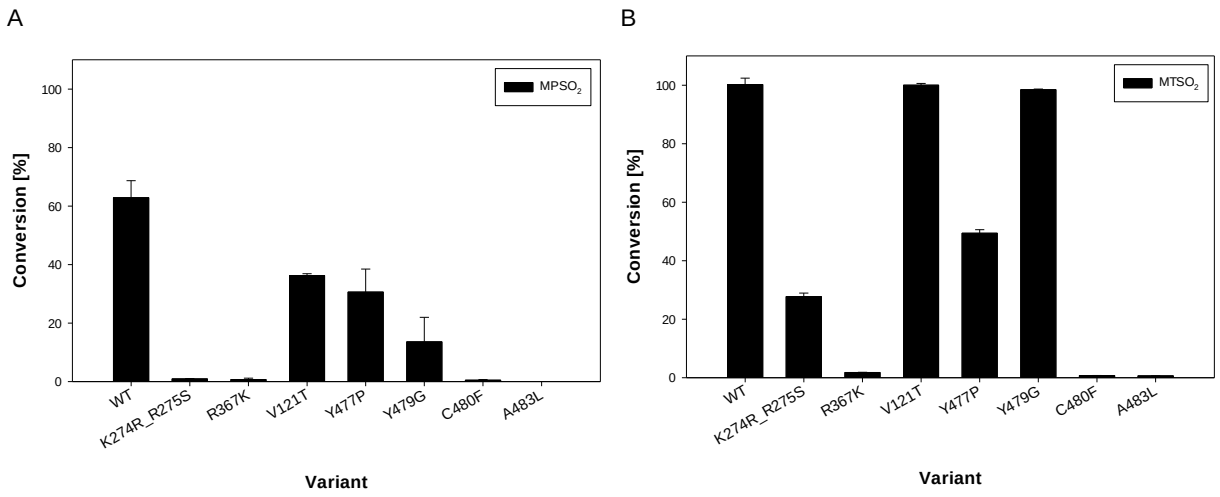


Figure 3.12: YMOA whole cell biocatalysis with 10 mM MTSO (left) and 10 mM MPSO (right) for 4 h at 25 °C (from Bordewick).[69]
The conversion to the sulfone is shown.

With the sulfoxides higher conversions of up to 100% could be achieved. For MTSO, the reached conversions were similar to the results obtained with its corresponding sulfide MTS.[69] Variants V121T, Y479G and the WT reached 100% $MTSO_2$ formation. The other mutations (K274R_R275S, V477P) reduced or completely destroyed the activity of YMOA. In contrast, the results obtained with MPSO differed from the ones for MPS. V121T, Y479G and V477P showed lower conversions of 40%, 14% and 31%, respectively. With the remaining variants only 1% or less MPSO was converted. The enantioselectivity could just be increased for the formation of $MTSO_2$. Wild-type YMOA displayed an (*S*)-selectivity in the kinetic resolution of MPSO with an enantiomeric excess of 52% ee, while determination for MTSO was not possible due to very low sulfoxide concentrations. With K274R_R275S the conversion of (*R*)-MTSO to $MTSO_2$ was preferred, thus leaving more of the (*S*)-enantiomer, showing with 37% ee the highest value for this reaction.

With YMOA the oxidation of 5 mM DMSO to $DMSO_2$ was also investigated (Figure 3.13).[69]

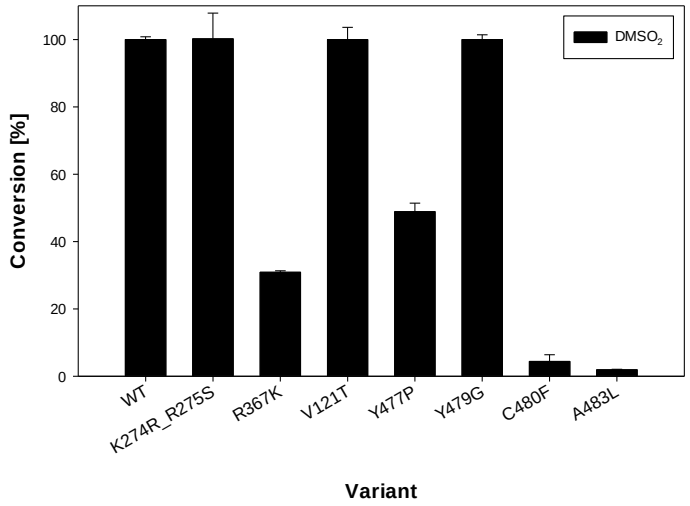


Figure 3.13: YMOA whole cell biocatalysis with 5 mM DMSO for 4 h at 25 °C (from Bordewick).[69] The conversion to the sulfone is shown.

The WT already reached a complete conversion of DMSO. Likewise, variants V121T, Y479G and K274_R275S fully converted this sulfoxide. The other mutations decreased the formation of $DMSO_2$.

From the three selected ketones, none was converted by YMOA or one of its variants.

3.1.2.4.5.2 YMOB

Whole cell biocatalysis with cells expressing YMOB was performed equally to YMOA.[69] YMOB showed low conversions with both sulfoxides and MTS (<1%). From the tested ketones, only 2-dodecanone was oxygenated to decyl acetate.

3.1.2.4.5.3 YMOC-H

Due to time restrictions, whole cell biocatalysis with all six remaining YMOs was performed with the best conditions found up to this point. Therefore, all enzymes were expressed at 25 °C for five hours. Only YMOH was coexpressed with pKEJ7 (DnaK-DnaJ-GrpE). Biocatalysis was carried out equally to $BVMO_{malto}$ (Chapter 3.1.2.1.3.7) with 5 mM substrate (cyclohexanone, bicyclo[3.2.0]hept-2-en-6-one, 2-dodecanone, methyl phenyl sulfide, methyl-*p*-tolyl sulfide) and an equimolar amount of glucose for NADPH regeneration at 25 °C for 24 h. Product formation did not occur in any of the reactions.

3.2 Switch of the cofactor specificity of the cyclohexanone monooxygenase from *Acinetobacter calcoaceticus* NCIMB 9871

NADH and NADPH are ubiquitous organic cofactors acting as electron carriers. Their only structural difference is the additional esterification of the 2'-hydroxy group with phosphoric acid present in NADPH, yet they fulfill considerably different roles in nature.[143] While NAD^+ is primarily involved in catabolism as an electron acceptor, NADPH is generally used as an electron donor for reductive biosynthesis.[144] In most cases enzymes have evolved to be highly selective towards NADH or NADPH, indicated by their contrasting roles in metabolism. Although there are examples for successful attempts to change the cofactor specificity in other enzyme classes through protein engineering, progress for BVMOs has been limited.[37-38, 42a, 43, 46a, 59b, 65a, 74-75, 76b, 85, 115-116, 121] Regarding that topic only the 4-hydroxyacetophenone monooxygenase (HAPMO), phenylacetone monooxygenase (PAMO) and cyclohexanone monooxygenase ($CHMO_{Acineto}$) have been investigated – with limited success.[118h, 145] It was possible to increase the activity ratio NADH/NADPH but mainly by destruction of the native activity with NADPH. Thus, the authors concluded that their "results indicate that the function of NADPH in catalysis cannot be easily replaced by NADH".[145] As a consequence, no type I BVMO could yet be engineered to efficiently use NADH as a cofactor. To address this important challenge, it was aimed to rationally design the cofactor specificity of $CHMO_{Acineto}$. This was started already in the Bachelor thesis of Sven Bordewick by identification of the residues in proximity of the phosphate group and the introduction of some mutations at these positions.[132] However, the activities of the variants obtained had to be verified and further improved.

3.2.1 Mutations of the phosphate recognition site

Firstly, a homology model was created to identify important residues for cofactor discrimination. As the only difference of NADPH and NADH is the additional phosphate group, residues in its proximity were targeted first. This led to the identification of six residues (Figure 3.14).[53b, 132]

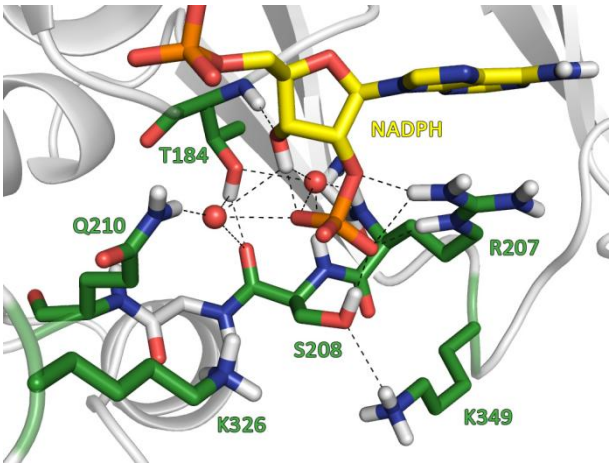


Figure 3.14: Residues of CHMO$_{Acineto}$ that are in the proximity of the phosphate group of NADPH. The protein scaffold is shown in gray, important residues are displayed as green sticks, NADPH is highlighted in yellow, oxygen in red, phosphorus in orange, nitrogen in blue and hydrogen in white. Water molecules are represented as red spheres. Dashed lines indicate hydrogen bonds or salt bridges. The structure was modeled based on the structure of the CHMO from *Rhodococcus* sp. (pdb-code 3GWD).[53b, 132]

T184 is involved in a hydrogen bond to S208. Through bridging water molecules it also shows hydrogen bonds to the 2'-oxygen of NADPH, another contact to the adenosine moiety of the cofactor and also to the residues R207 and S208. In the same way there is an interaction with Q210, indicating a very important and complex role of T184. R207, S208 and Q210 are located on the loop closest to the 2'-phosphate. The phosphate group forms hydrogen bonds with R207 (in this case two salt bridges) and S208 directly and with Q210 indirectly through a bridging water molecule. The guanidino group of R207 interacts with the adenine base via stacking interactions. K326 and K349 just point at the phosphate group without having direct interactions with it. K326 is especially interesting because of the adjacent R327. R327 points directly at the catalytic center, is essential for the catalytic cycle and is strictly conserved among type I BVMOs.[61, 73] K349 displays a hydrogen bond to S208.[132]

Additionally, structure and sequence alignments including typical (i.e. PAMO, HAPMO) and atypical (MekA, SMFMO) BVMOs were performed to predict the best possible mutations for the identified residues.[53b, 132] This sequence alignment contained 37 known BVMOs and 20 NADH employing enzyme sequences to compare the NAD(P)H binding regions (Appendix Figure 8.6).[53b]

The six identified positions close to the phosphate group of NADPH and the mutations derived from this analysis are summarized in Table 3.6.

Table 3.6: **Residues in the proximity of the phosphate group of NADPH in $CHMO_{Acineto}$ and their respective occurring deviations in the sequence alignment.**

Position	Deviations
T184	A, *D*, *E*, G, I, N, *S*, V
R207	H, *K*, *N*, Q
S208	*A*, *D*, E, H, *N*, T
Q210	*A*, *C*, *D*, *H*, *I*, *K*, N[46b], *P*, S
K326	F[118h], *G*, H[118h, 145], N[118h, 145], *Q*, R
K349	*A*, E, *G*, *L*, *N*, *P*, *Q*, R, *T*, *V*

Red: frequent residue in the alignment, green: residue from NADH oxidases, blue: mutations reported in literature, *italic*: residues occurring in the alignment, underlined: chosen mutations.[53b]

As the most frequent amino acid at position T184 among BVMOs was asparagine, this mutation was chosen as well as alanine in order to mimic the structure in NADH oxidases.[53b, 132] The SMFMO displays a glycine here and thus the mutation T184G was chosen, too.

A structural alignment with SMFMO suggested that R209/T210 of the CHMO from *Rhodococcus sp*. HI-31 (equivalent to R207/S208 in $CHMO_{Acineto}$) align with Q193/H194 from SMFMO (Figure 3.15). The structure alignment was performed with $CHMO_{Rhodo}$ instead of $CHMO_{Acineto}$ as this did not result in a good alignment with the latter.[132]

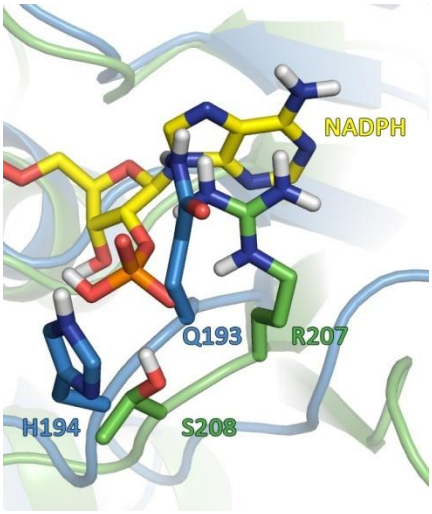


Figure 3.15: Structure alignment of $CHMO_{Rhodo}$ (3GWD) and SMFMO for the residues R209 and T210. $CHMO_{Rhodo}$ is shown in green, SMFMO in blue, NADPH is highlighted in yellow, oxygen in red, phosphorus in orange, nitrogen in blue and hydrogen in white. The respective residues are displayed as sticks.[53b, 132]

The sequence alignment (Figure 3.16) of $CHMO_{Acineto}$ and SMFMO rather suggests that R207/S208 align with H194/E195.[132] This is plausible since acidic residues are a common binding motif for NADH.

	200										208
$CHMO_{Acineto}$	P	L	A	K	H	L	T	V	F	Q	R	S
SMFMO	T	V	A	E	T	T	W	I	T	Q	H	E

Figure 3.16: Sequence alignment of $CHMO_{Acineto}$ and SMFMO for the residues R207 and S208. R207 and S208 of $CHMO_{Acineto}$ align with H194 with E195 of SMFMO.[132]

As aspartic acid is structurally similar to glutamic acid, its effect was investigated. Furthermore, tyrosine is the most conserved amino acid for position 208 among BVMOs (74%) and thus was chosen as well. Because of these apparent contradictions, five point mutations were selected: R207Q, S208D, S208H, S208E and S208T.

Mutating H220 in PAMO (equivalent to Q210 in $CHMO_{Acineto}$) to glutamine or asparagine proved to be beneficial for the activity with NADH.[70c] Since $CHMO_{Acineto}$ already exhibits a glutamine, the mutation Q210N was attempted to elucidate its role in this BVMO.[132] Moreover, the most frequent residue in the alignment, serine, was chosen, together with aspartic acid to investigate different amino acid types at this position.[53b]

In general, basic residues are of special interest concerning cofactor specificity because of their possible interaction with the negatively charged 2'-phosphate group.[132] At the position corresponding to K326, MekA displays a histidine. This is unusual since the lysine at this

position is conserved among type I BVMOs. Mutation of this lysine to histidine or asparagine was advantageous for the activity with NADH in HAPMO but disadvantageous in PAMO.[70c, 119] The change to phenylalanine at this position was reported to enhance the activity with NADH in HAPMO as well and thus was selected, too.[53b, 119] The most frequent residue in the alignment for this position was arginine and therefore interesting as well. Thus, the mutations K326F, K326H, K326N and K326R were chosen to examine their effects in $CHMO_{Acineto}$.

Position 349 is not strictly conserved among the typical type I BVMOs.[132] Both $CHMO_{Acineto}$ and MekA exhibit a lysine at this position. However, PAMO, HAPMO and SMFMO all display small, hydrophobic amino acids. A mutation to glutamate was selected, because its opposite polarity might induce a big disturbance at this position.[132] This approach was chosen to examine the overall importance of this residue. Additionally, it was mutated to arginine as that was the most frequent and additionally a basic residue in the alignment.[53b]

3.2.1.1 Activity of phosphate binding site variants

Activities of the variants were measured in the NADPH depletion assay (Figure 3.17, Appendix Figure 8.8).

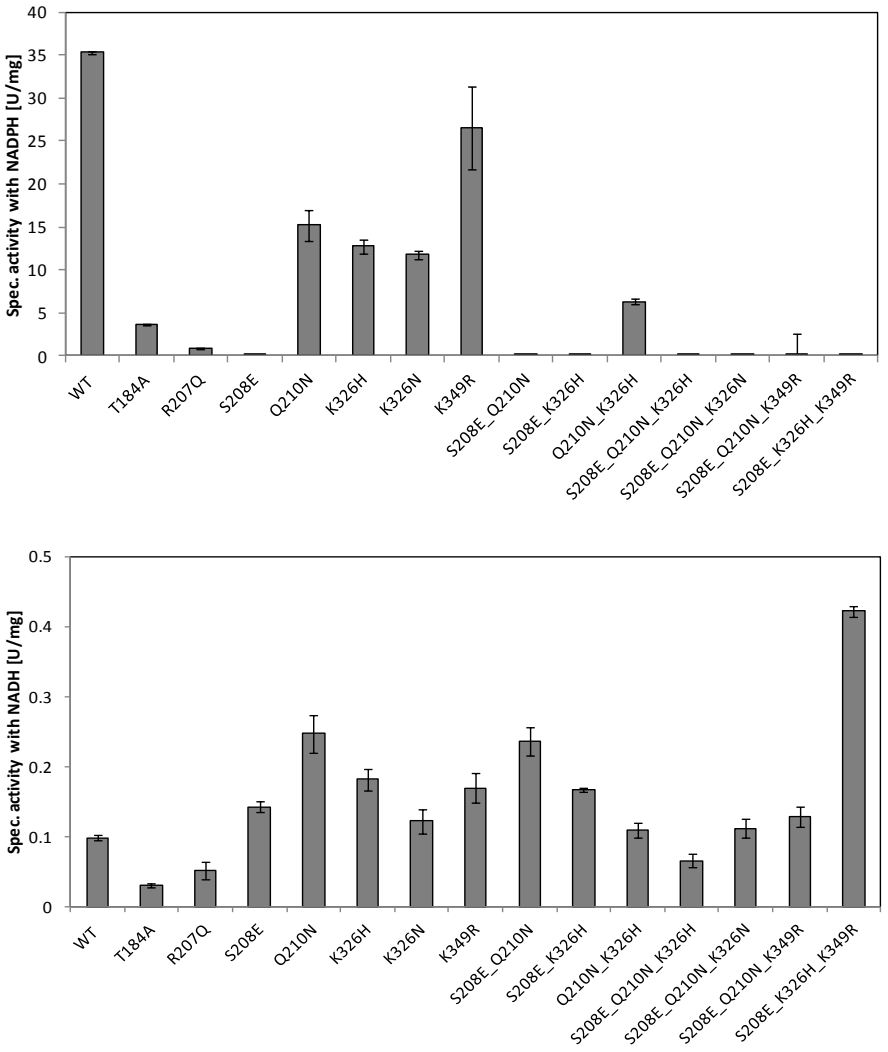


Figure 3.17: Specific activity of $CHMO_{Acineto}$ wild type (WT) and variants (mutations in the phosphate recognition site) using NADPH (top graph) or NADH (bottom graph) as cofactors.

Mutations in proximity to the phosphate group of NADPH that led to a substantially increased activity with NADH were S208D, S208E, Q210N, Q210S, K326H, K326N and K349R.

By combination of the positive mutations of the phosphate binding site, the variants S208E_K326H_K349R, S208E_Q210N and S208E_K326H were obtained with an improved activity with NADH compared to the WT (Figure 3.17). Additionally, they displayed an increased activity ratio NADH/NADPH of 7.60 (S208E_K326H_K349R), 7.49 (S208E_Q210N) and 3.33 (S208E_K326H) in comparison to the WT (0.003, Figure 3.18). Activity ratios of variants R50L, L55R, T139L, I182V, T184A/G, S186P, V189I, S208H/T, Q210N/S, K326H/N/R, V253Y, D341C, K349E/R, D347V, F380Y, W490Y and Q210N_K326H were too low to be presented here (Appendix Figure 8.7).

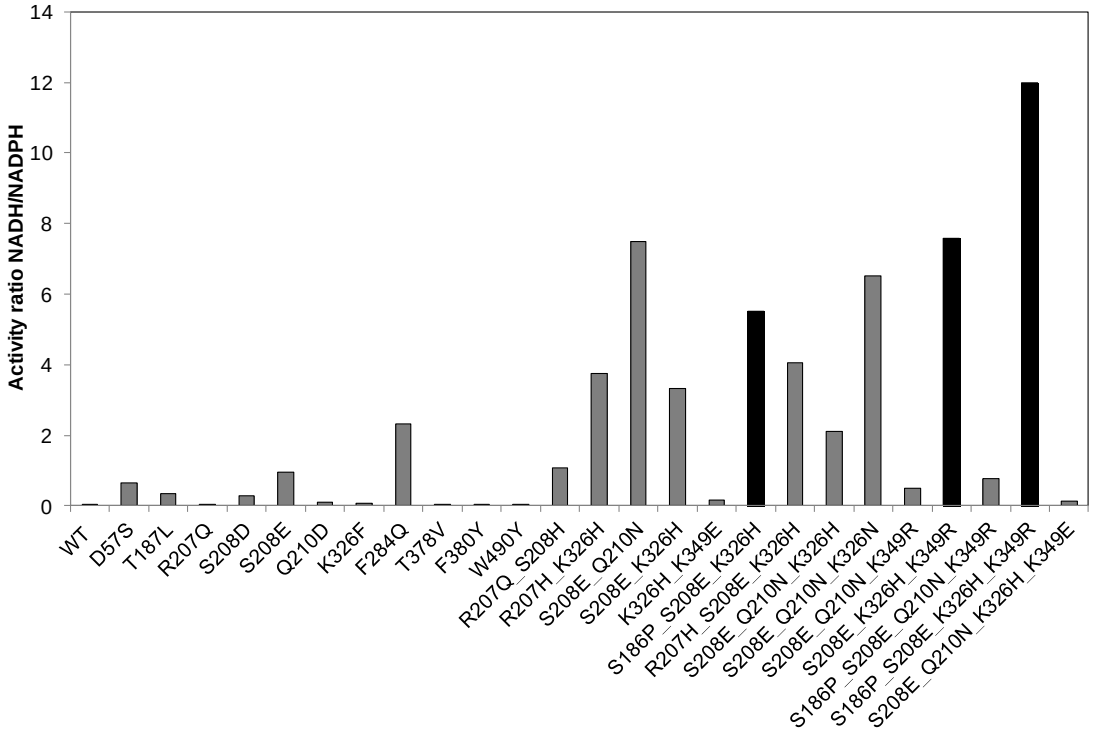


Figure 3.18: Specific activity ratio for the utilization of NADH over NADPH with enzyme variants of $CHMO_{Acineto}$.
Black bars: mutants with a specific activity with NADH >0.4 U/mg.

3.2.2 Investigating residues in proximity of NAD(P)H

In order to further enhance the activity of the variants, a more detailed investigation of the homology model was performed. 24 residues were identified in the proximity of the cofactor, of which six (L55, D57, S186, T187, F380 and W490) showed direct or indirect hydrogen bonds towards the cofactor molecule, which could result in a more flexible interaction with NADH after mutagenesis (Scheme 3.3, Appendix Table 8.2).

Scheme 3.3. Simplified illustration of the NADPH binding pocket of $CHMO_{Acineto}$.
Only residues are shown that have direct or indirect contact to NADPH via hydrogen bonds (dashed lines).

L55 and D57 are situated in the "fingerprint 2" [A/G]GxWxxxx[F/Y]P[G/M]xxxD.[34a] F380 and T378 were also considered as both are located in another conserved region, named "fingerprint 3" (Appendix Figure 8.9). S186 and T187 are located within the Rossmann fold. Directly adjacent to this fold are I182 and V189. As the Rossmann fold is responsible for binding the NADPH, these residues were considered good targets as well. Three more residues are present in this conserved region: G183, G185 and G188, which were not mutated in the end because they are completely conserved (G183, G185 with 100%) or did just show either glycine or alanine within the alignment (G188). Using the alignment (Appendix Figure **8.6**), the most frequent amino acids and those occurring in the included NADH utilizing enzymes at these positions were selected for the mutational studies (Table 3.7).

Table 3.7. Positions identified in the proximity of NADPH and their respective occurring deviations in the sequence alignment.

Position	Deviations
L55	A, G, *H*, M, Q, R, *V*
D57	**S**
I182	F, *M*, T, V**(+)**
S186	A, *C*, G, N, P
T187	I, L**(+)**, *R*, S, V
V189	A, H, I**(+)**, L, N, Q, *T*
T378	*S*, V
F380	*L*, Y
W490	*F*, *N*, *T*, Y

Red: frequent residue in the alignment, green: residue from NADH oxidases, *italic*: residue occurring in the alignment, underlined: chosen mutations.

Moreover, another six mutations (R50L, T139L, V253Y, F284Q, D341C and D347V) were chosen as they were determined to be beneficial for the BVMO BmoF1 with respect to either conversion or enantioselectivity.[146] All these residues were investigated within the sequence alignment and, except for T139 and V253, showed a good conservation (R50: 55%, T139: 18%, V253: 16%, 284: F53%, D341: 37%, D347: 58%) and thus were deemed to be targets for mutations.

After introduction of these mutations into the $CHMO_{Acineto}$, variants L55R, S186P, T187L, V253Y, F284Q, D341C and W490Y showed an improved activity with NADH (Figure 3.19).

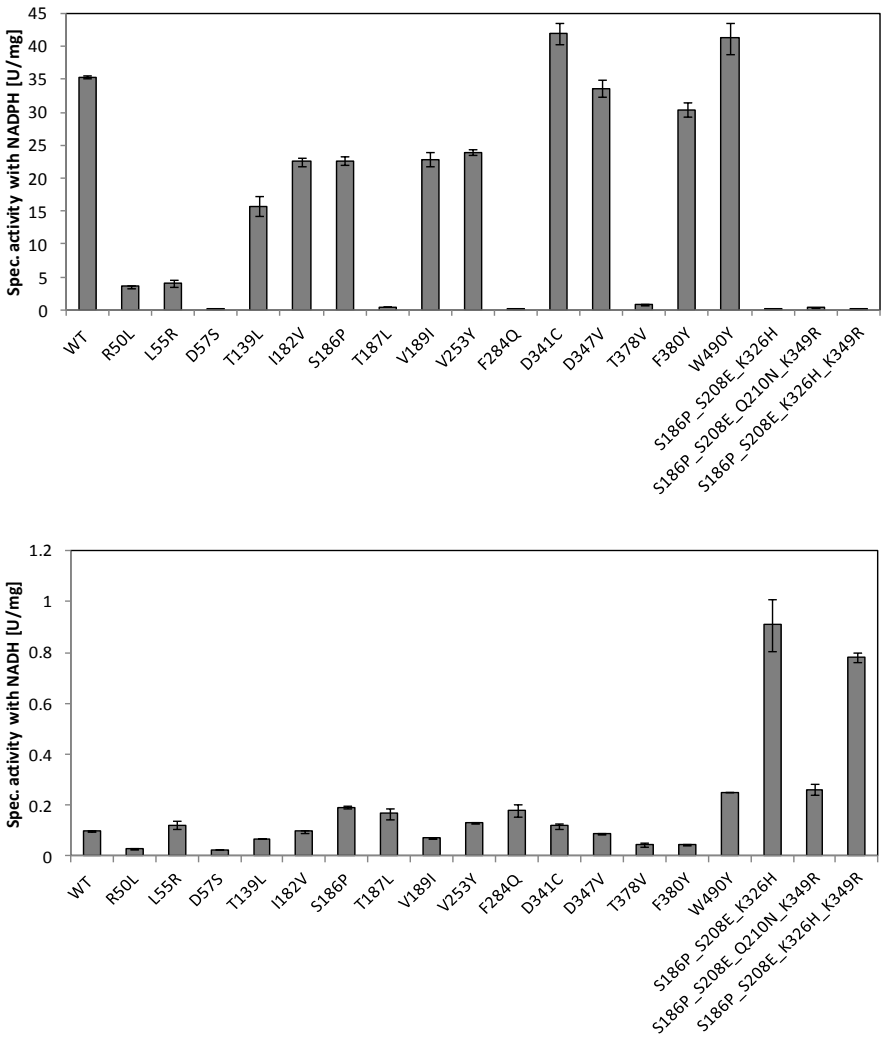


Figure 3.19: Specific activities of enzyme variants targeting residues in proximity of NAD(P)H of $CHMO_{Acineto}$ and mutations transferred from the BmoF1 from *P. fluorescens* DSM 50106 using NADPH (top) and NADH (bottom) as cofactors.

Different combinations of the best mutations revealed S186P_S208E_K326H and S186P_S208E_K326H_K349R to be the best ones, having a 9- and 8-fold increased activity with NADH compared to the wild type, respectively. Regarding their NADH specificity, these two variants show an increased activity ratio (NADH/NADPH) of 1,920 and 4,170-fold, respectively, having not only an increased activity with NADH, but also being specific for it (Figure 3.18).

The mutant S186P_S208E_K326H_K349R_W490Y was constructed, but it could not be purified as it did not bind to the Co^{2+} column material and thus it was not investigated.

3.2.3 Determination of kinetic parameters

In addition, kinetic parameters of the wild type and the variant S186P_S208E_K326H were determined to further validate the influence of the introduced mutations (Table 3.8).

Table 3.8: Kinetic parameters determined for WT $CHMO_{Acineto}$ and its variant S186P_S208E_K326H (3M) with NAD(P)H with and without substrate to differentiate between activity (in the formation of ε-caprolactone) and uncoupling (yielding H_2O_2 from NAD(P)H).

Variant	K_M [μM]	k_{cat} [s^{-1}]	k_{cat}/K_M [$mM^{-1}*s^{-1}$]
WT_NADPH	5.9 ±0.75	41 ±1	6,979
WT_NADPH_unc[a]	-	0.9 ±0.03	–
3M_NADPH	2,259 ±259	5.3 ±0.13	2.2
3M_NADPH_unc	820 ±83	0.7 ±0.01	0.9
WT_NADH	1,733 ±239	2.3 ±0.12	1.3
WT_NADH_unc	1,485 ±187	1.9 ±0.05	1.3
3M_NADH	681 ±57	7.1 ±0.1	10.4
3M_NADH_unc	57 ±9	1.1 ±0.02	19.2

[a] K_M value was not determinable due to instant saturation at extremely low cofactor concentrations (1.56 μM)

The catalytic efficiency with NADH for the variant S186P_S208E_K326H is 8-fold higher, the K_M is 2.5-fold lower and the k_{cat} is 3-fold higher compared to the wild type when looking at the activity values. Furthermore, when comparing the activities without substrate, uncoupling – formation of hydrogen peroxide from NAD(P)H instead of ε-caprolactone formation – with both cofactors is reduced for the triple mutant from 0.9 s^{-1} and 1.9 s^{-1} to 0.7 s^{-1} and 1.1 s^{-1} for NADPH and NADH, respectively. Because such a large difference in uncoupling was observed, activity tests without substrate with more variants were performed (Table 3.9).

Table 3.9: Uncoupling of enzyme variants of $CHMO_{Acineto}$ with NADH. The percentage of uncoupling using NADH is shown.

Variant	Spec. activity[a] [U/mg]	Spec. activity[b] [U/mg]	Uncoupling[c] [%]
WT	0.16	0.18	115
S186P	0.32	0.34	107
S208E	0.19	0.27	144
Q210N	0.19	0.30	157
K326H	0.28	0.38	137
D341C	0.19	0.23	126
K349R	0.22	0.19	83
W490Y	0.35	0.54	155
S208E_Q210N	0.32	0.30	97
S208E_K326H	0.32	0.38	121
Q210N_K326H	0.17	0.27	153
S186P_S208E_K326H	0.91	0.67	74
S208E_Q210N_K349R	0.15	0.21	134
S208E_K326H_K349R	0.59	0.60	103
S186P_S208E_K326H_K349R	0.90	0.49	54

[a]Using 1 mM cyclohexanone as substrate; [b]In the absence of cyclohexanone; [c]in case of >100% uncoupling, activity with NADH alone was higher than with the substrate

For the most variants, including the wild type, uncoupling was higher than production of ε-caprolactone. Additionally, the majority of the variants displayed an increased uncoupling. However, variants S186P, K349R, S208E_Q210N, S186P_S208E_K326H, S208E_K326H_K349R and S186P_S208E_K326H_K349R showed a decreased uncoupling.

3.2.4 Biocatalysis with variants of $CHMO_{Acineto}$

Next, the top variants were investigated in biocatalytic reactions using cyclohexanone as substrate to confirm the observed increase in NADH preference (Figure 3.20).

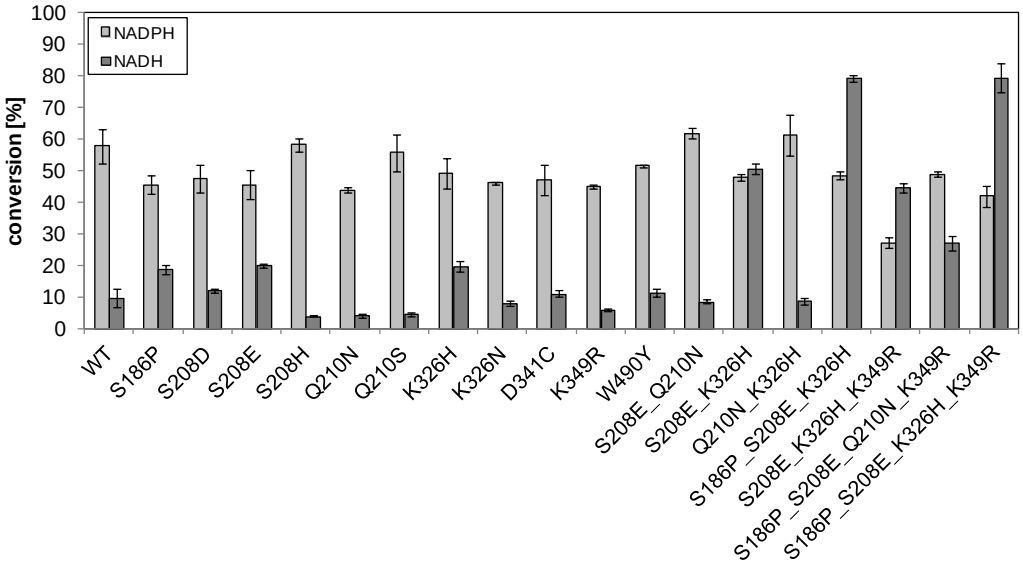


Figure 3.20: ε-Caprolactone formation from 5 mM cyclohexanone after 2 h using purified variants of the $CHMO_{Acineto}$ with 5 mM NADPH or NADH as cofactors in TrisHCl buffer at pH 9.

The NADPH concentration led to inhibiting effects of the WT enzyme.

This revealed that most mutants still reached 50-60% conversion using NADPH, except for the triple mutant S208E_K326H_K349R (27% conversion). Several variants did not enable higher conversions with NADH as expected from the data shown in FiguresFigure 3.17 andFigure 3.19, which is due to higher uncoupling rates with this cofactor as described in the previous section. However, S186P, S208E and K326H showed significantly improved activity with NADH, whereas S208D, D341C and W490Y just showed a slightly increased conversion. The combinatorial variants (except for S208E_Q210N and Q210N_K326H), showed much higher activities. The triple and quadruple mutants S186P_S208E_K326H and S186P_S208E_K326H_K349R resulted in >79% conversion with NADH, exceeding the values determined for the WT using NADPH.

4 Discussion

4.1 Baeyer-Villiger monooxygenases participating in the metabolism of ketones in yeasts

4.1.1 Determination of metabolites from yeasts from 2-dodecanone and 1-dodecene

In this project, which was initiated in my diploma thesis, the metabolism of ketones and alkenes in cultures of growing yeast cells was investigated.[63, 133] Detected degradation products from 1-dodecene included 2-dodecanone, 2-dodecanol, decanoic acid, dodecanoic acid and hexanedioic acid (Appendix Scheme 8.1) and additionally degradation products from 2-dodecanone, namely decyl acetate, decanol, octanedioic acid and decanedioic acid (Scheme 3.1) were detected. The production of some of these metabolites provided evidence of the induction of ketone degrading enzymes after growth on 2-dodecanone or 1-dodecene, but not on dodecane, in the investigated hydrocarbon oxidizing yeasts.[133]

4.1.1.1 Metabolism of 2-dodecanone

After cultivation of *C. maltosa*, *C. tropicalis*, *C. catenulata*, *C. albicans*, *Y. lipolytica* and *T. asahii* with 2-dodecanone or 1-dodecene as sole carbon and energy source, the culture media were used for the determination of formed metabolites to study the metabolism of aliphatic ketones and alkenes.[133]

In culture media from *C. maltosa* with 2-dodecanone as substrate, a total of six metabolites were detected (Scheme 3.1). Although in quantity fewer compounds were secreted by some other yeasts, the products were the same. The conversion of 2-dodecanone was either initiated by a keto reductase or a BVMO (Scheme 4.1).[133]

Scheme 4.1: Suggested metabolism of 2-dodecanone, 1-dodecene and dodecane in *Candida maltosa*.[133]
A black frame indicates intermediates detected from cultures of *C. maltosa* with 2-dodecanone (1), 1-dodecene (2) or dodecane (3) as sole carbon and energy source. ADH: Alcohol dehydrogenase, ALDH: Aldehyde dehydrogenase.

The first degradation pathway, the formation of decyl acetate, initiated by a BVMO, was identified for all investigated yeast strains with exception of *C. albicans*.[133] Due to the slow growth of *C. albicans* with 2-dodecanone, the accumulation of metabolites like decyl acetate may not have been sufficient to detect them in the cultures of this yeast. That could be the reason why just 2-dodecanol was found as an intermediate of *C. albicans*. The enzymatic Baeyer-Villiger oxidation can lead to both regioisomeric products of the corresponding ketone.[41, 47a, 48, 58a, 99, 147] In this case the "normal" product is decyl acetate and undecanoic acid methyl ester would be the "abnormal" one. It seems that *C. maltosa* as a whole cell catalyst rather formed the "normal" product decyl acetate than the "abnormal" undecanoic acid methyl ester as the latter could not be detected. The preferred formation of the "normal" product from a ketone by a BVMO can be observed in many cases like in the conversion of different cyclic ketones with $CHMO_{Brevi1}$ from *Brevibacterium epidermidis* HCU.[49a, 53b, 148] However, the "abnormal" ester can be the preferred product as well like it is the case for the cyclohexanone monooxygenase of *Arthrobacter* sp. and $CHMO_{Acineto}$ with (+)-trans-dihydrocarvone.[47a] In addition, decanol was detected, which was probably formed by conversion of decyl acetate via a hydrolase.[133] A similar reaction was reported by Forney *et al.*, in which *Pseudomonas multivorans* converted 2-tridecanone to undecyl acetate and subsequently hydrolysed it to undecanol and acetate.[149]

In addition, the formed decanoic acid was probably hydroxylated at its ω-end resulting in the formation of 10-hydroxydecanoic acid, which was further oxidized to the respective dicarboxylic acid decanedioic acid.[133] Decanedioic acid was degraded by β-oxidation, which was confirmed by detection of its C_2 shortened intermediates until hexanedioic acid. Therefore, a diterminal degradation took place, which has previously been reported for the filamentious fungus *Mortierella isabellina*, the bacteria *Rhodococcus rhodochrous* and *Corynebacterium* spp. as well as for the yeasts *Candida guilliermondii* and *Candida* spp.[35, 103e-g] At least some of the intermediates of the β-oxidation of this dicarboxylic acid were detected in the cultures of all investigated yeasts.[133] The formation of dicarboxylic acids ranging from six to twelve carbons from dodecane for *C. tropicalis* and from twelve to sixteen carbons from the respective alkanes for *Y. lipolytica* has already been reported.[150]

Another degradation pathway is proposed by the detection of 2-dodecanol.[133] Therefore, the yeasts have to contain at least one keto reductase that can reduce 2-dodecanone to this secondary alcohol. The reduction of ketones like 2-hexanone, methyl acetoacetate and α-tetralone to the respective alcohols was described for other yeasts such as *Saccharomyces cerevisiae* and *Candida viswanathii*.[63, 151] However, further investigation is required to see

whether 2-dodecanol can be converted over another pathway than oxidation back to 2-dodecanone with its subsequent degradation.

4.1.1.2 Metabolism of 1-dodecene

In *C. maltosa* cultures with 1-dodecene as sole source of carbon and energy, five compounds were detected (Appendix Scheme 8.1).[133] One of the products was 2-dodecanone. For the formation of this compound, 1-dodecene possibly reacted via 2-decyloxirane to 2-dodecanol and subsequently to 2-dodecanone (Scheme 4.1). Also, *C. tropicalis* and *Y. lipolytica* were able to form 2-tetradecanol and 2-tetradecanone from tetradecene.[152]

In addition, dodecanol was also possibly formed via 1,2-epoxidodecane from 1-dodecene and oxidized to dodecanal and subsequently to dodecanoic acid.[133] This acid was further metabolized via β-oxidation. Some of these reactions were already described for *Y. lipolytica*, *C. tropicalis* and *Candida* spec as well.[152-153] By ω-oxidation of dodecanoic acid, the dioic acid was formed and transformed by β-oxidation to hexanedioic acid.[133] The reason why there was no formation of dodecanol from 1-dodecene detectable, could be that this primary alcohol was metabolized too fast to decanoic acid and the subsequent products.

In contrast, in *C. maltosa* culture media with dodecane neither decyl acetate nor decanol nor 2-dodecanol were formed. Instead, the degradation was accomplished by a monoterminal oxidation.[133] Thus, dodecanoic acid was detected. In addition, a diterminal oxidation was observed supported by the detection of hexandioic acid. A similar study with *Candida rugosa* and decane indicated a mono- and diterminal oxidation as well.[154]

In order to investigate the differences in product formation, the induction of ketone converting enzymes was studied in biotransformations with either 2-dodecanone, 1-dodecene or dodecane with resting cells of *C. maltosa*, obtained from cultivation with one of these substrates in my diploma thesis.[63, 133] While 2-dodecanone was immediately converted after pre-cultivation with 1-dodecene, cells pre-cultured with dodecane barely transformed the ketone. This matched the principle of simultaneous adaptation, after which the catabolism of all formed intermediates is fully induced by incubation with the initial substrate, if the participating enzymes of the pathway are inducible.[155] As 1-dodecene was converted to 2-dodecanone, when the former was used as a carbon and energy source, the participating enzymes in the corresponding catabolism of *C. maltosa* readily converted 2-dodecanone. In addition, an enzyme with BVMO activity and a keto reductase were induced in cells pre-cultured with 2-dodecanone or 1-dodecene as those formed decyl acetate and its fission product decanol as well as 2-dodecanol in the whole cell biocatalysis with 2-dodecanone. In contrast, in dodecane pre-cultured cells incubated with 2-dodecanone none of these inter-

mediates were formed, confirming the lack of conversion of the alkane to its corresponding ketone during cultivation.
It was shown that *C. maltosa* can convert 1-dodecene to 2-dodecanone and this aliphatic ketone to decyl acetate resulting from a BVMO activity, which can also be found in all of the here investigated hydrocarbon oxidizing yeasts.[133] The reduction of 2-dodecanone to its corresponding secondary alcohol 2-dodecanol was additionally shown.

In conclusion, the ketone metabolism in yeasts is comparable to that of bacteria.[133] Ketones can either be oxygenated to esters or reduced to the corresponding alcohol. These pathways seem to be quite ubiquitary among several hydrocarbon oxidizing yeasts. *C. albicans* was the only yeast not being able to form an ester out of the ketone in a detectable amount. Beyond this, alkenes can be converted to ketones and thus induce the production of ketone-degrading enzymes, whereas this pathway seems to be invalid for alkanes.

To further validate the degradation of the intermediates, they could be used as substrates for the cultivation of the yeasts. Additionally, cells precultured with ketone and alkene could be used in biocatalytic reactions with these intermediates. If growth occurs with the intermediates as well and an immediately high conversion of them takes place in biocatalysis, the proposed degradation pathways would be further validated. To mark the substrates with isotopes would simplify to follow their degradation pathway and would also verify the obtained results.

4.1.2 Investigations of novel BVMOs from yeasts

4.1.2.1 Putative Baeyer-Villiger monooxygenase from *C. albicans* - $BVMO_{albi1}$

Of the nineteen putative BVMOs and FMOs identified in my diploma thesis, $BVMO_{albi1}$ from *C. albicans* and $BVMO_{malto}$ and CMO from *C. maltosa* were investigated first, as members of the genus *Candida* show the best growth with hydrocarbons among fungi, verified also by growth experiments with the ketone 2-dodecanone as a sole carbon and energy source.[63, 100b, 100c, 133]

After trying expression using the vector pET28a(+), which solely resulted in the production of insoluble protein in form of inclusion bodies, the system was changed to pBAD_SUMO. By having a translational fusion with SUMO, soluble expression was actually achieved like reported before for other proteins.[134] The level of expression was low but the putative BVMO could still be purified out of the cell extract from *E. coli*. However, the respective protein fraction did not show a yellow color typical for flavin-containing solutions, indicating the loss of the flavin-cofactor FAD during the purification procedure. This can happen when purifying BVMOs and leads to colorless and inactive solutions as reported for $CHMO_{Arthro}$ F299 variants.[47a] A spectral analysis in the range of 250 nm to 800 nm, in which no flavin peak could be observed, confirmed this assumption. An attempt to restore the FAD in the potential BVMO did not succeed. Apparently, $BVMO_{albi1}$ is an enzyme that cannot be restored after the loss of the cofactor. For $CHMO_{Arthro}$ it was reported that cofactor restoration was possible, but after that the regioselectivity and activity changed, thus making it indispensable to maintain the bound FAD in the BVMO.[47a]

Additionally, according to the protein sequence of $His_{(6)}$-$BVMO_{albi1}$, a molecular weight of 65.3 kDa was expected, but was around 50 kDa according to the analysis via SDS-PAGE. For this procedure, it is assumed that all proteins just move in accordance to their mass/charge ratio in an electrical field what has been reported several times already.[156] Due to different factors, deviations from the estimated values can occur, though.[157] As the tertiary structure of $BVMO_{albi1}$ is unknown, a more compact structure of this protein is also possible, blocking the SDS from some regions of the protein and making it more positive and thus migrating faster.[158] A slight proteolysis could not be excluded at this point either. Proteins present in form of inclusion bodies are less susceptible to the degradation by proteases, but it still can occur.[159] Thus, apparently wrong migration distances can be observed if the protein does not act “normal” in the electrophoresis.

Since expression of $BVMO_{albi1}$ was rather difficult and $BVMO_{malto}$ from *C. maltosa* was the more attractive target for research as this yeast displayed the best growth with 2-dodecanone,

investigations were focused on this enzyme (see the following chapter). However, further investigation of $BVMO_{albi1}$ in the future should be interesting and promising as it represents an uncharacterized eukaryotic BVMO.

4.1.2.2 Putative Baeyer-Villiger monooxygenase from *C. maltosa* - $BVMO_{malto}$

C. maltosa proved to be the yeast with the highest growth rate with 2-dodecanone under the investigated strains leading to the assumption to find a highly active BVMO in this alkane-oxidizing eukaryote.[63]

4.1.2.2.1 Expression in eukaryotic systems

As *C. maltosa* is a yeast, expression of $BVMO_{malto}$ in another yeast was regarded as the best strategy to get a high amount of soluble and active protein due to both the similar codon usage and possible post-translational modifications.

Two yeasts were selected for heterologous expression: *Pichia pastoris* and *Yarrowia lipolytica*. In the first, expression with and without subsequent secretion was investigated, in the latter due to time issues only with secretion. However, no expression of $BVMO_{malto}$ could be observed in any of the approaches and protein samples never showed any activity in NADPH depletion assays. Thus, it was obvious that this putative BVMO was not expressed in these yeasts. This can have different reasons. The enzyme could have been degraded instantly after translation or the step of protein synthesis could have been disrupted by a terminating structure in the mRNA like a hairpin. However, due to the failed expression of $BVMO_{malto}$ in *P. pastoris* and *Y. lipolytica*, the strategy to use eukaryotic expression systems was changed to employ a prokaryotic one, *E. coli*.

4.1.2.2.2 Expression in E.coli

4.1.2.2.2.1 Influence of different expression conditions

For cloning, a pET system was chosen as in this, BVMOs have already been successfully expressed.[71, 87, 96] After integration of the gene $bvmo_{malto}$ into the vector pET28a(+), its expression was investigated. First, the influence of temperature on the soluble expression was tested.[63] At 15 and 30 °C no expression was detectable. At 37 °C a protein band indicating overexpression of $BVMO_{malto}$ could be seen. However, it was faint and solely in the insoluble fraction.[63] Likewise, different IPTG concentrations and the addition of riboflavin to the culture medium, did not influence the solubility or activity of the desired protein. The expression in the alternative vectors pBAD_SUMO and pCRE3 also resulted in levels of $BVMO_{malto}$ which could not be detected. By design of a codon optimized variant, $BVMO_{malto}$ overexpression could be increased a lot, but still the soluble fraction of this protein was very small (Figure 3.2) also showing no activity in a NADPH depletion assay. The correct folding of BVMOs presumably is an even more complex process than for many other proteins. This is clearly evident from former work on BVMOs like the enzymes from *P. putida* and *P. fluorescens*, which could only be expressed solubly after coexpression of chaperones.[139]

Actually, $BVMO_{malto}$ was coexpressed with the chaperone systems pKJE7, pG-KJE8, pTf16, pGro7 or pG-Tf2 as well (Figure 3.3). This indeed increased its solubility in case for pKJE7, but there was still no measurable activity in the NADPH depletion assay. Thus, the chaperones probably did not enable folding of the putative BVMO into its functional form. Also the chaperones could have remained bound to $BVMO_{malto}$, interfering with its activity.

4.1.2.2.2.2 Whole cell biocatalysis with $BVMO_{malto}$ in E. coli *BL21(DE3)*

Even though many different approaches had been tested to increase the solubility of $BVMO_{malto}$, cell extracts never showed activity. It has been reported in some cases that BVMOs are only active inside cells due to stabilizing effects, like it was the case for the enzyme from *Pseudomonas fluorescens* DSM 50106.[160] Thus, whole cell biocatalysis was employed. After a first approach failed in which the enzyme was coexpressed with the chaperone system DnaK-DnaJ-GrpE from the plasmid pKEJ7 at 30 °C, a different approach was applied. Here, expression was performed in TB medium without coexpression in a temperature range of twelve to twenty degree. The selection of substrates used in biocatalysis contained a broader spectrum with representatives of different substrate classes accepted by BVMOs, ranging from cyclic (cyclohexanone) and bicyclic (bicyclo[3.2.0]hept-2-en-6-one) over aliphatic (2-dodecanone) ketones to sulfides (methyl phenyl sulfide [MPS], methyl *p*-tolyl sulfide [MTS]). With that approach, $BVMO_{malto}$ displayed activity with three of the substrates, namely bicyclo[3.2.0]hept-2-en-6-one, 2-dodecanone and MPS leading to the formation of the products 3-oxabicyclo[3.3.0]oct-6-en-2-one, decyl acetate and methyl phenyl sulfoxide, respectively. A relatively high conversion of 22.4% with bicyclo[3.2.0]hept-2-en-6-one in comparison to 5.6% and 2.2% with 2-dodecanone and MPS, respectively (Figure 3.4), indicates a specialization of the enzyme towards bulkier, maybe polycyclic, ketones. The structures of MTS, cyclohexanone and MPS are quite similar but only the latter was converted, maybe due to its slightly more space filling sulfide group without the additional methyl group at the phenyl ring MTS is displaying (Scheme 4.2).

Scheme 4.2: Comparison of the similar structures of the substrates cyclohexanone (CH), methyl phenyl sulfide (MPS) and methyl *p*-tolyl sulfide (MTS).

However, the activities were probably too low to be detected by the NADPH depletion assay, as it was limited to just two minutes, especially when thinking about the decreased sensitivity

due to the background activity of enzymes in the cell extract. This explains the missing activity of the cell lysate. Actually, detecting products of $BVMO_{malto}$ after it never showed activity in many previous attempts was quite surprising. It is assumed that the cultivation temperature gave the highest impact in getting an active enzyme in the end. The gene expression was carried out between 12 and 20 °C in a time period of 5 h. At higher temperatures, the expression rate is generally higher, but in many cases insoluble inclusion bodies are then formed.[161] The improvement in solubility of different proteins like subtilisin E, bacterial luciferase or β-lactamase by decreasing the cultivation temperature has been reported several times already.[161c] The aggregation of proteins to inclusion bodies generally is increased at higher temperatures as the strong temperature dependency of hydrophobic interactions determines the aggregation reaction.[162] A direct consequence of a lower temperature is the decrease of heat shock proteases, which are induced at overexpression conditions.[163] Furthermore, the activity and expression of some chaperones in *E. coli* is increased at reduced temperatures.[164] The van't Hoff equation is one of the determining factors of the described and further processes, also of the protein synthesis. This equation says that a rise in temperature of 10 °C leads to a 2 to 4-fold increased reaction velocity.[165] Of course, in an organism this is only possible in its range of tolerance. When proteins are produced more rapidly because of a rise in temperature, the folding machinery also needs to fold them into the right conformation. At some point this is not working anymore, so that the proteins cannot be folded correctly any longer leading to an increased formation of inclusion bodies.[161b, 161c] Expression of $BVMO_{malto}$ between 12 and 20 °C apparently created conditions suitable for the folding process and machinery leading to an enzyme capable of converting the three substrates bicyclo[3.2.0]hept-2-en-6-one, 2-dodecanone and MPS (Scheme 4.3).

Scheme 4.3: Overview of Baeyer-Villiger oxidations catalyzed by $BVMO_{malto}$.

Bicyclo[3.2.0]hept-2-en-6-one is accepted by the majority of the BVMOs and thus is used as a standard substrate to screen for activity of new BVMOs, but also to test well known ones and respective enzyme variants.[99, 166] Considering this, the comparatively high activity of $BVMO_{malto}$ with this compound could be explained, but also indicates the discussed preference for conversion of bulkier substrates. It is worth emphasizing, that only one of the two possible regioisomers was formed from bicyclo[3.2.0]hept-2-en-6-one, making it an even more selective biocatalyst. This observation strengthens the statement of chapter 4.1.1.1 that this BVMO from *C. maltosa* prefers the production of the "normal" oxygenation products. This, of course, can vary from substrate to substrate.[41, 48, 58a, 99, 147] In a different study, it has been shown that the regioselectivity can be changed by mutation of residues being necessary for substrate positioning in the active site.[47a] In that way, the selectivity could be completely changed from the production of the abnormal ester to the normal one in $CHMO_{Arthro}$ and $CHMO_{Acineto}$ for (+)-trans-dihydrocarvone.[167] It would be interesting to investigate the effects of these mutations in this new type I BVMO.

Concerning the reactions with MPS and MTS, it needs to be said, that due to the volatility of these sulfides they were not available for the whole time of biocatalysis. In fact, they might have evaporated nearly completely after just some hours. After 24 h not even traces were detectable any more, even without product formation. Constantly decreasing substrate concentrations of the sulfides led to a limited substrate availability, reducing the reaction velocity. Periodically refeeding MPS and MTS could circumvent this problem and thus a higher conversion of MPS and product formation in case of MTS might be achieved.

The substrate scope of BVMOs can vary a lot from very broad in case of $CHMO_{Acineto}$ with over 100 identified substrates to quite narrow when talking about PAMO preferably accepting small aromatic ketones and sulfides.[45a, 66a, 70e, 85, 94] $BVMO_{malto}$ with the activities detected so far, seems to group inside the well described type I BVMOs known to date. However, more structurally different substrates should be tested to give a better picture of the specialization of this eukaryotic enzyme and to see whether the supposed preferred conversion of bulkier compounds proves to be true.

4.1.2.3 Putative monooxygenase from *C. maltosa* - CMO

The putative monooxygenase CMO was investigated as well, even though it did not contain the "fingerprints 1 and 2" but showed sequence similarities to FMOs.[63] After cloning the gene *cmo* into the vector pET28a(+), its expression in *E. coli* was investigated. First the influence of temperature was examined to determine the best conditions for producing this protein.[63] At all temperatures tested, 15, 30 and 37 °C, no big difference in expression and solubility could be seen. However, slightly more CMO was produced at 37 °C. The solubility was low, though. Only after two hours a small percentage of the putative monooxygenase could be detected in the soluble fraction.[63] By coexpression of the TaKaRa chaperone plasmids pKJE7, pG-KJE8, pTf16, pGro7 or pG-Tf2 the solubility could not been improved either. Only after cell disruption using supersonication instead of FastPrep24®, soluble CMO could be obtained. FastPrep24® is a mechanical disruption method, in which glass beads that are shaken with a high frequency by a ball mill are used to set cells under frictional and shear forces. This results in an efficient disruption of the cell membranes, but leads to protein denaturation in many cases as well. In contrast, supersonication is a more gentle method. Here, alternating current voltage is transformed into a high frequent form. With this, vibration of a resonator is induced and transferred into the sample containg the cells. Hereby, microscopic air bubbles are generated, which implode immediately again. This phenomenon called cavitation causes pressure changes that can break up cell membranes. During both methods, heat is generated that has to be compensated by cooling the sample. In case of FastPrep24® this cannot be done during the disruption process while it is possible for supersonication, making the latter again a more sensitive method. Thus, cell lysis by FastPrep24® was probably too harsh and led to denaturation of the putative monooxygenase. By employing supersonication, CMO was detectable in the soluble fraction even without coexpression of chaperones, but in higher amounts when folding was supported by pG-Tf2. However, in any case no activity could be detected in the NADPH depletion assay. Its sequence identity to other FMOs of up to 73% and the missing BVMO fingerprints could be an explanation for this. If this protein is indeed a FMO, for a spectrophotometrical assay it needs to be purified and tested with NADH as cofactor as the background in the cell extract for a measurement with this cofactor is too high. A biocatalysis would be applicable without purification, though. However, the enzyme might need a reductase for its activity, supplying reduced FAD for the oxidative reaction of the monooxygenase.[18] With sequence motifs Maria Kadow already used to identify the reductase Fre from *E. coli*, the genome of *C. maltosa* could be screened as well.[64] Either Fre or a putative reductase for CMO from *C. maltosa* needs to be expressed together or separately

and then, in a purified form, be used in a NADH depletion assay and/or in biocatalysis. By using a broad spectrum of BVMO and FMO substrates, it could be possible to detect activity of this putative monooxygenase.

4.1.2.4 Investigations about the codon reassignment in *Candida* spec.

In a distinct group of yeasts, the universal codon CUG standing for leucine is translated as serine in 97% of the cases and just with a likelihood of 3% as leucine.[137] From 78 investigated *Candida* yeasts, only in eleven leucine was built in the proteins when a CUG was present in the respective mRNA.[168] *C. albicans* and *C. maltosa* proved to belong to the group of yeasts with this special codon reassignment. As heterologous expression of CTG containing genes from these yeasts would lead to the wrong insertion of leucine at the respective positions, the identified monooxygenases were analyzed for the presence of this codon.[63] Hereby, one (CMO) and two ($BVMO_{malto}$) CTG codons could be identified, $BVMO_{albi1}$ did not contain any. With an increasing GC content of the genome in general more CTG codons can be found.[137a, 137b] Thus, it is rare in *S. cerevisiae* (GC: 40%), *C. albicans* (GC: 34%) and *C. maltosa* (GC: 34.2%).[169] Therefore, it is not surprising that at most two CTG codons were found in the ORFs of the monooxygenases.[63] Intriguingly, the probability that a leucine is inserted at such a site can be increased from 3% to 28% without affecting the growth of the yeast.[170] However, a study of Miranda *et al*. showed that the reversion of the genetic code in respective yeasts had many consequences regarding among others cell morphology, altered gene expression and secretion of hydrolytic enzymes.[171] Employing site-directed mutagenesis in order to exchange the CTG codons with TCT codons to enable insertion of serines at these positions, did not improve the solubility of either $BVMO_{malto}$ or CMO. However, due to the changes observed in the native hosts when manipulating the CTG sites, causing the exchange of leucines to serines and *vice versa* at the corresponding positions of all proteins, it is likely that it somehow affected $BVMO_{malto}$ or CMO.[170] To that point, only the activity of $BVMO_{malto}$_L111S_L261S could be determined with the substrates 2-dodecanone, bicyclo[3.2.0]hept-2-en-6-one and MPS (see Chapter 4.1.2.2.2.2). Thus, it would be inter-esting to compare the activity and the substrate spectrum, but also properties like stability and temperature and pH optima between both enzyme variants.

4.1.2.5 Investigations of BVMO fingerprints

By analysis of a multiple sequence alignment consisting of thirteen protein sequences of BVMOs, next to the known sequence motifs GxGxx[G/A] (Rossmann fold), FxGxxxHxxxW(P/D) ("fingerprint 1") and [A/G]GxWxxxx[F/Y]P[G/M]xxxD ("fingerprint 2"), I identified additional conserved regions in my diploma thesis.[58a, 60, 62-63] Consequently, these were designated "fingerprint 3" (Dx[I/L][V/I]xxTG[Y/F]) and "fingerprint 4" ([G/D][P/A]xxYxxxxxxxxPN[L/M][W/F]xxxG). The conservation was further verified in this thesis by employing a multiple sequence alignment of 56 BVMO sequences originating from different organisms, also containing $BVMO_{malto}$ (see Chapter 4.1.2.2).[63] However, there were some deviations determined so that both motifs were reduced to DxxxxxG[Y/F] and PNxxxxxG, respectively to have a conservation of 100% within the alignment (Table 4.1).

Table 4.1: Sequence motifs in BVMO protein sequences.

Rossmann fold 1	**Rossmann fold 2**
GxGxxG[58a, 60]	GxGxx[G/A][60, 62]
"fingerprint 1"	**"fingerprint 2"**
FxGxxxHxxxW(P/D)[60]	[A/G]GxWxxxx[F/Y]P[G/M]xxxD[34a, 58a]
"fingerprint 3"	**"fingerprint 4"**
Dx[I/L][V/I]xxTG[Y/F][63] (DxxxxxG[Y/F])[a]	[G/D][P/A]xxYxxxxxxxxPN[L/M][W/F]xxxG[63] (PNxxxxxG)[a]

[a] "fingerprint 3 and 4" in parentheses represent the respective reduced motifs

However, "fingerprint 3" was found to be conserved in 49 of the 55 sequences (89.1%), making it a motif suited to identify the majority of BVMOs in sequence databases. Additionally, this motif could already be published in the course of another project (see Chapter 4.2).[53b] "Fingerprint 4" could only be found in roughly the half of the included sequences so that it was necessary to adapt it and use the reduced form to employ it for identification purposes. However, Fraaije *et al.* just used fourteen sequences of the flavoprotein monooxygenase super family to designate the "fingerprint 1" and assign the Rossmann fold to BVMOs.[40a] Thus, it was necessary to analyze them in a larger alignment. In a later study, in which 24 sequences were compared, the conserved residues, except for the glycine and tryp-

tophane, of the motif FxGxxxHxxxW(P/D) showed deviations, even though all the used sequences were originating from the same organism, *Rhodococcus jostii* RHA1.[34a] In this study, they also designated the "fingerprint 2" by just using these 24 BVMO sequences, which also should be verified in much larger alignments. In a publication by Rebehmed *et al*., 116 type I BVMO sequences were compared in multiple sequence alignments to study the conservation of important residues.[61] Here, "fingerprint 1 and 2" were conserved in 88.8% and 99.1% of the sequences, demonstrating their high, but not full conservation. Not further specified but listed as conserved were residues from "fingerprint 3 and 4" as well (Table 4.2).

Table 4.2: Residues of "fingerprint 3 and 4" which were reported as conserved by Rebehmed *et al.*[61]

"fingerprint 3"	
Residue[a]	**Conservation [%]**
D374	100
T380	90.5
G381	100
F/Y382	88.8/11.2
"fingerprint 4"	
Residue[a]	**Conservation [%]**
G409	89.7
P422	100
N423	100
F425	97.4
G429	85.3

[a]Residue numbering is refered to $CHMO_{Rhodo}$

Combining both the sequence analysis of this work and the one of Rebehmed *et al*., a high variety of sequences is given, with which the motifs DxxxxxTG[Y/F] ("fingerprint 3") and PNxFxxxG ("fingerprint 4") can be assigned to be suited to identify type I BVMOs out of databases.

Furthermore, another sequence motif containing the second (more central located) Rossmann fold could be designated in principle as the sequence region surrounding this fold exhibits a high conservation (Figure 4.1).

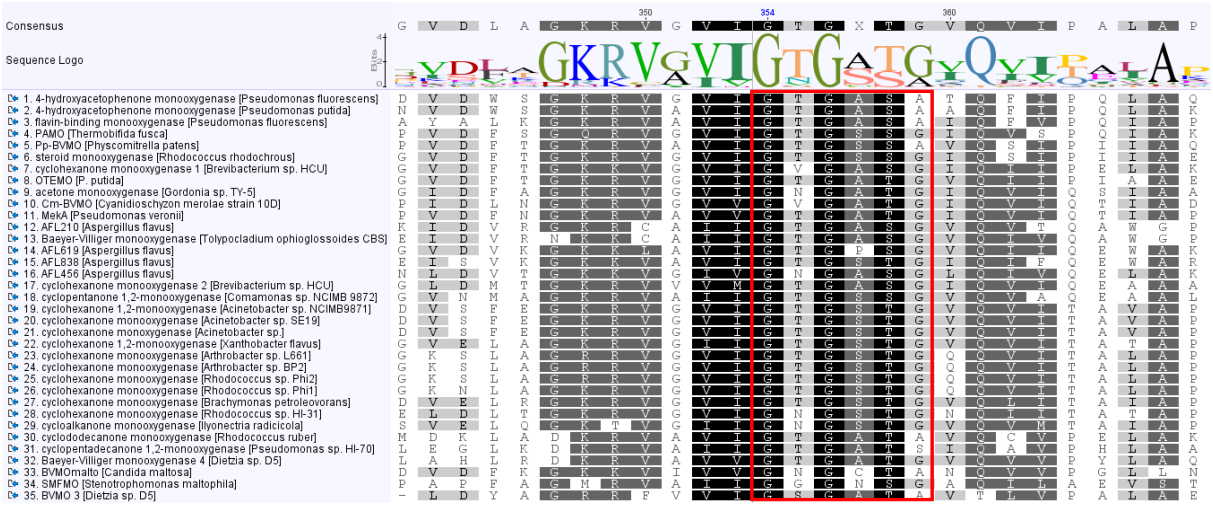


Figure 4.1: Multiple sequence alignment containing 35 BVMO protein sequences to investigate the second (more central located) Rossmann fold.
A red frame indicates the motif GxGxx[G/A].

The motif [R/K]Vx[V/I][I/V]G[T/V]Gx[S/T][G/A]xQxx[P/T/Q]xx[A/G] for this region is conserved in more than 90% of the sequences used in this thesis and also of the 116 sequences analyzed by Rebehmed *et al*.[61] Therefore, it could be useful for genome mining and the discovery of further type I BVMOs.

To get a better understanding of the importance and function of the conserved residues in the new fingerprint motifs, they need to be further investigated. First, one could try finding a relation between the presence of the initially found and then reduced forms of the motifs in different BVMOs to their properties like activity and substrate specificity. Secondly, a mutational analysis of the conserved residues is required to verify the presumptions and to understand the role of the single residues and also the motifs as a whole.

4.1.2.6 Putative monooxygenases from *Y. lipolytica* – YMOs

Because *Yarrowia lipolytica* forms decyl acetate from 2-dodecanone, it definitely contains enzymes with BVMO activity (see Chapter 4.1.1.1). Utilizing CHMO from *Acinetobacter sp.* NCIMB 9871 ($CHMO_{Acineto}$) for a protein alignment using BLAST in the annotated proteins from the genome of *Y. lipolytica*, nine homologous sequences, designated *Yarrowia* monooxygenases (YMOs) A-I, with sequence identities of about 20% could be identified (Figure 3.6).[106b] Furthermore, the sequences contain the BVMO fingerprints and two Rossmann fold motifs making them likely to be type I BVMOs.[34a, 58a, 62]

4.1.2.6.1 Homologous expression of YMOA and YMOB in Yarrowia lipolytica *Po1f*

In the master thesis of Sven Bordewick, which was co-supervised by me, the investigations were started with YMOA and B (Table 3.3) by performing homologous expression as this is in general the best choice to obtain a decent amount of correctly folded protein.[69] However, both expression with and without subsequent secretion did not yield any of the desired proteins. This could have several reasons already laid out in Chapter 4.1.2.2.1. Therefore, the strategy to express the YMOs in *Y. lipolytica* was abandoned for all the identified putative *Yarrowia* BVMOs and the work was focused on the expression in *E. coli* (see next section).

4.1.2.6.2 Heterologous expression and purification of YMOA-H

The genes for YMOA-H were successfully cloned into pET28a(+) vectors from the genomic DNA of *Yarrowia lipolytica* due to the absence of introns.[69] 25 °C proved to be the best temperature for expressing YMOA (Figure 3.7).[69] With a reduction in temperature from 30 °C to 25 °C, YMOE and G could be obtained in soluble form as well. The lower temperature probably led to an improved soluble expression through decreased proteolysis, and a slower protein expression enabled proper folding, like discussed above (see Chapter 4.1.2.2.2.2).[69, 172] In order to express YMOB and H solubly, coexpression of the chaperone system DnaK-DnaJ-GrpE from the pKJE7 plasmid was necessary.[140] YMOC, D and F could not be expressed solubly within the scope of this thesis.

In an attempt to purify the YMOs, YMOA could be purified, albeit just yielding a low amount of functional enzyme due to the high inactivation during purification.[69] Additionally, the enzyme, either in cell extract or in the purified form, was unstable even when stored on ice.[69] YMOB could not be purified at all, as it did not bind to the Ni^{2+} column material.[69] The $His_{(6)}$-tag was probably not accessible for binding either, because the N-terminus with the tag might have been located inside the protein or chaperones bound to YMOB blocked the $His_{(6)}$-tag.[69, 173] Also a misfolding caused by the tag is imaginable.[69] As purification did not work

well for YMOA and B, it was skipped for the remaining YMOs in favor of utilizing them in whole cell biocatalysis reactions.

4.1.2.6.3 Substrate spectrum of the YMOs

4.1.2.6.3.1 YMOA

In the master thesis of Sven Bordewick, fourteen substrates (Scheme 3.2) were tested in NADPH depletion assays and whole cell biocatalysis to characterize YMOA and YMOB.[69] A variety of ketones was included, ranging from the simplest ketone, acetone, to linear and cyclic ketones and finally bicyclic and aromatic ketones.[69] Surprisingly, none of the eight ketone substrates were converted. In contrast, YMOA displayed activity towards the sulfides MTS, MPS and their corresponding sulfoxides MTSO, MPSO and additionally DMSO. L-Methionine was the only sulfide that was not converted, probably due to the large structural differences (Scheme 3.2).[69] Intriguingly, YMOA accepted only the tested sulfides/sulfoxides even though the structure of the used ketones acetophenone, 4-hydroxyacetophenone and acetone is very similar (Scheme 4.4).[69]

Scheme 4.4: Comparison of ketone and sulfide/sulfoxide substrates with similar structure (from Sven Bordewick).[69]

In a study by Orru *et al*. about the catalytical mechanism in PAMO, no distinctive residues for substrate recognition could be found.[73] Thus, it was concluded, that BVMOs are mainly "oxygen-activating and "Criegee-stabilizing" catalysts that act on any chemically suitable substrate that can diffuse into the active site and reach the catalytic center where the flavin-peroxide and oxyanion hole are positioned".[69, 73] Therefore, it is highly intriguing that the ketones structurally similar to the sulfides/sulfoxides were not converted. However, in other BVMOs there is also no clear correlation between the activity towards ketones and their sulfide analogues. While HAPMO showed similar catalytic efficiencies with MTS/MPS and their respective ketones, *p*-methylacetophenone and acetophenone, this was not the case for

PAMO.[51a, 70e] These discrepancies might be caused by the differences that can be found when comparing structurally resembling ketones and sulfides/sulfoxides. Ketones are sp_2-hybridized and thus planar and both sulfides and sulfoxides show a tetrahedral geometry due to their sp_3-hybridization.[72e] Additionally, the mechanism of a Baeyer-Villiger oxidation differs from a sulfoxidation. [68] In a recent work, the catalytic arginine R327 of the Ar-BVMO was mutated to alanine, which completely destroyed its Baeyer-Villiger activity but 84% of its sulfoxidation activity were retained.[67e] Although in a similar study, the corresponding mutation overall inactivated PAMO, it is obvious that there is a difference in the mechanism of the oxygenation of ketones and the oxidation of heteroatoms like found in sulfides and sulfoxides.[174] The ability of YMOA for specific sulfoxidation while apparently missing Baeyer-Villiger oxidation activity was further investigated in a mutational study.[69] Here, the largest influence was observed when mutating residues in the active site of the enzyme (Figure 4.2).

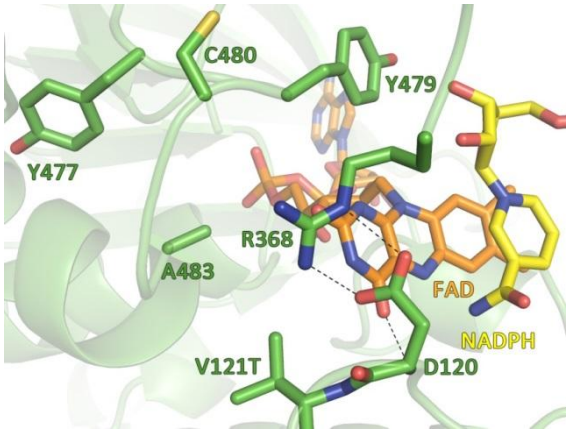


Figure 4.2: Overview of active site residues of YMOA (from Sven Bordewick).[69]
The mutated residues are shown along the conserved residues R368 and D120 as green sticks.The protein scaffold is shown in green. NADPH is highlighted in yellow, FAD in orange, oxygen in red, nitrogen in blue. Dashed lines indicate hydrogen bonds or salt bridges.

Mutations V121T and Y479G increased the sulfone production from both MTS and MPS 2-fold (Figure 3.11), also leading to a largely increased total conversion for MPS (90% and 76%) and a slight increase for MTS (6% and 19% respectively).[69] Y479G showed a significant effect on the enantioselectivity as well as in case of MTS complete inversion of the enantiopreference could be achieved (Table 3.5).[69] Zhang *et al*. investigated effects of active site mutations on sulfoxidations as well.[68] By the generation of enzyme variants of PAMO and respective docking simulations, they found explanations for increased activities and enantioselectivities. Due to the mutations, the shape of the substrate binding pocket was either directly or indirectly reshaped leading to a rotated binding mode of the substrate. In the generated quadruple mutant I67Q/P440F/A442N/L443I (corresponding to residues

V121/Y477/Y479/C480 in YMOA, Figure 4.2) this even inverted the enantioselectivity as the other enantiotopic lone electron pair of the sulfur was closer to the hydroperoxide group and thus attacked. A relation between longer O-S distances (hydroperoxide–substrate) and a higher sulfone formation was found as well, without an elucidation, though.[68] It can be assumed that in YMOA the mutations resulted in a different shape of the substrate binding pocket as well.[69] T121 could form a hydrogen bond to the neighboring and catalytically important D120, which is not the case for V121 in the wild-type enzyme (Figure 4.2).[61, 69] This together with the different chemical properties of a polar threonine in comparison to a hydrophobic valine could have changed the substrate binding site.[69] The change of one of the largest amino acids, tyrosine, at position 479 to the smallest amino acid, glycine, certainly caused structural changes enabling an electrophilic attack on the opposite lone electron pair of MTS causing the inversion in enantioselectivity.[69] Mutation of K274 and R275, which are part of the phosphate recognition site in YMOA, also had a big impact on its sulfoxidation activity (Figure 4.3).[69]

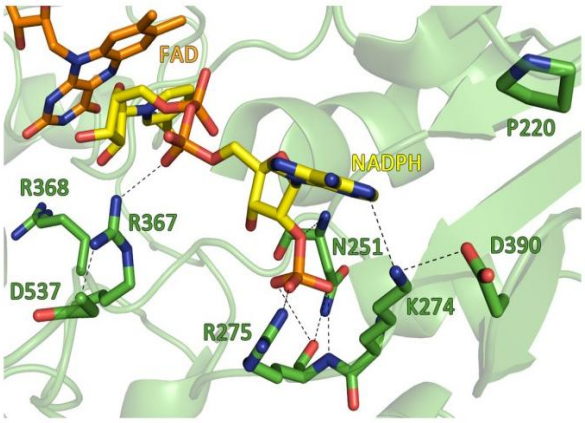


Figure 4.3: NADPH recognition site of YMOA (from Sven Bordewick).[69]
The protein scaffold and important residues are shown in green. NADPH is highlighted in yellow, FAD in orange, oxygen in red, nitrogen in blue. Dashed lines indicate hydrogen bonds or salt bridges.

These two positions are usually conserved among type I BVMOs with arginine and threonine/serine for positions 274 and 275, respectively. Thus, the double mutant K274R_R275S was investigated, in which sulfoxide formation was increased by 6% (MTS) and 80% (MPS) and sulfone formation was decreased by 27% and 98%, respectively.[69] This could have been caused by a change of the NADPH orientation due to structural changes induced by the mutations.[69] The shape of the phosphate recognition site in YMOA is significantly different from typical type I BVMOs, displaying the residues N251, K274, R275 and V277 in contrast to the expected ones threonine, arginine, threonine/serine and glutamine/asparagine, respectively, like found i.e. in $CHMO_{Acineto}$, PAMO and HAPMO (see Chapter 4.2).[53b, 61, 69, 70c, 119, 132]

Changing the atypical residues to the conserved ones, could have caused a different positioning of the NADPH molecule, more similar to the one found in other type I BVMOs.[69] Therefore, the activity profile of K274R_R275S could have changed to a more "typical" one of type I BVMOs with sulfone as a by-product instead of a main product.[69] All in all, the selected mutations in YMOA influenced its overall activity, sulfone-to-sulfoxide ratio and stereoselectivity.[69] However, none of the variants was capable of a Baeyer-Villiger oxidation and thus further studies are necessary to see whether reshaping of this monooxygenase to a "typical" type I BVMO by further mutations is possible, only a suitable substrate is missing or it is incapable of converting ketones at all.

It is noteworthy as well that YMOA, while displaying a preferred formation of (*R*)-MPSO with a very low enantioselectivity of 1% ee, produced (*S*)-MTSO with a high enantioselectivity of 95% ee.[69] These results are comparable to the finding with $CHMO_{Acineto}$ that preferred these two enantiomers, too.[70a] Ottolina *et al*. explained these different preferences with an active site model build from the determined enantioselectivities of 30 sulfides with $CHMO_{Acineto}$ (Figure 4.4).[67a]

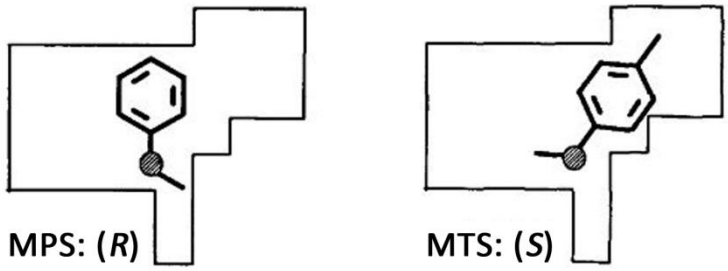
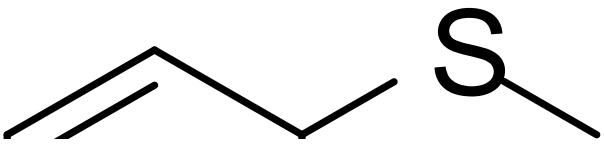


Figure 4.4: Active site model of CHMO showing the different binding modes and resulting enantioselectivities for MPS and MTS (adapted from Ottolina *et al*., taken from Sven Bordewick).[67a, 69]

Because of its *p*-methyl substituent, MTS is bound differently at the substrate binding site than MPS leading to an altered orientation of the sulfide group.[69] Therefore, the opposite enantiomer is formed. This probably is also the reason for the contrasting enantioselectivities towards MTS and MPS in YMOA.[69]

All in all, even though it showed a sequence identity to $CHMO_{Acineto}$ of about 20% and displays all type I BVMO sequence motifs, YMOA did not catalyze a Baeyer-Villiger oxidation. To find a ketone to be accepted by this monooxygenase necessitates a more comprehensive substrate screening. However, sulfides and sulfoxides proved to be suited substrates for YMOA, making it a valuable catalyst (Scheme 4.5).



Scheme 4.5: Overview of sulfoxidations catalyzed by YMOA (from Bordewick).[69]
Abbreviations for the compounds are given in parentheses

This activity profile of YMOA resembles the one found for group B FMOs, which do not convert ketones either, but display heteroatom oxygenation.[18, 69] Thus, it could fulfill an analogous role in *Y. lipolytica* in detoxifying compounds.[18, 69]

4.1.2.6.3.2 YMOB

In contrast, YMOB showed a low, typical type I BVMO activity.[69] In the NADPH depletion assay the conversion of five out of eight ketones could be detected (Table 3.4).[69] Also the sulfide MTS and the sulfoxide MPSO were accepted. In a whole cell biocatalysis this could only be partially verified.[69] Here, 2-dodecanone, MTS, its corresponding sulfoxide MTSO and MPSO were converted, although less than 1% product formation were detected. The remaining apparent substrates identified in the NADPH depletion assay, could indicate uncoupling (see Chapter 4.2.3) and thus could have been false-positive results.[53b, 69] Conversely, the performance of YMOB in the biocatalysis might have been very low so that it could not be detected by GC.[69] The overall low activity of this BVMO could be caused by incorrect folding due to its N-terminal $His_{(6)}$-tag or it might exhibit a naturally complex folding, which cannot be easily obtained in *E. coli*.[69] Only when chaperones were coexpressed, a soluble state was achieved.[69] However, the chaperones again could have remained bound to YMOB, interfering with the structure and/or the activity of the enzyme. Like described above, the chaperones might have folded insoluble YMOB into a soluble but not (fully) functional one as well. These assumptions would fit to the unsuccessful purification of this BVMO (see Chapter 4.1.2.6.2).[69]

4.1.2.6.3.3 YMOC-H

The six putative *Yarrowia* BVMOs YMOC-H were investigated in parallel in the course of this thesis. In the NADPH depletion assay and whole cell catalysis, a variety of different substrates was tested (cyclohexanone, bicyclo[3.2.0]hept-2-en-6-one, 2-dodecanone, MPS, MTS). However, the YMOs showed in neither approach activity with any of the substrates. In case of YMOC, D and F this was due to their insoluble expression. In case that there were soluble traces of one of the putative BVMOs, the found deviations for YMOD (L9S, A12P, L14S, P263L, K482E) and YMOF (P319L, I328T) could be destructive mutations derived from the cloning procedure, leading to inactive versions of both enzymes. However, for YMOE, G and H, product formation of at least one substrate was expected as they were in a soluble form. Nevertheless, even though five quite different substrates were used for the investigation, it is possible that none of them was a suitable substrate for any of these YMOs. There are BVMOs with quite narrow substrate spectra like PAMO, $BVMO_{Brevi2}$ and $CPMO_{Coma}$, making it likely for the new enzymes to be specific for different compounds like steroids.[94, 175] In case of YMOH there are two additional possibilities why no product was formed. Firstly, it only was soluble when coexpressing pKEJ7. As discussed for YMOB, a chaperone can fold a protein into a soluble but not active state or after the folding process it still can be bound to the protein, hindering its activity. Secondly, three deviations (T66A, T92A, L247P) were identified within the cloned sequence of *ymoh* which also could be destructive mutations. Additionally, the reason for the absence of activity for all six YMOs could be the N-terminal $His_{(6)}$-tag as well.

4.1.2.7 Classification of the identified yeast BVMOs

In order to further characterize, valuate and compare the identified yeast monooxygenases with each other and other type I BVMOs, a phylogenetic tree from a sequence alignment of 85 BVMO sequences was constructed (Figure 4.5).[69]

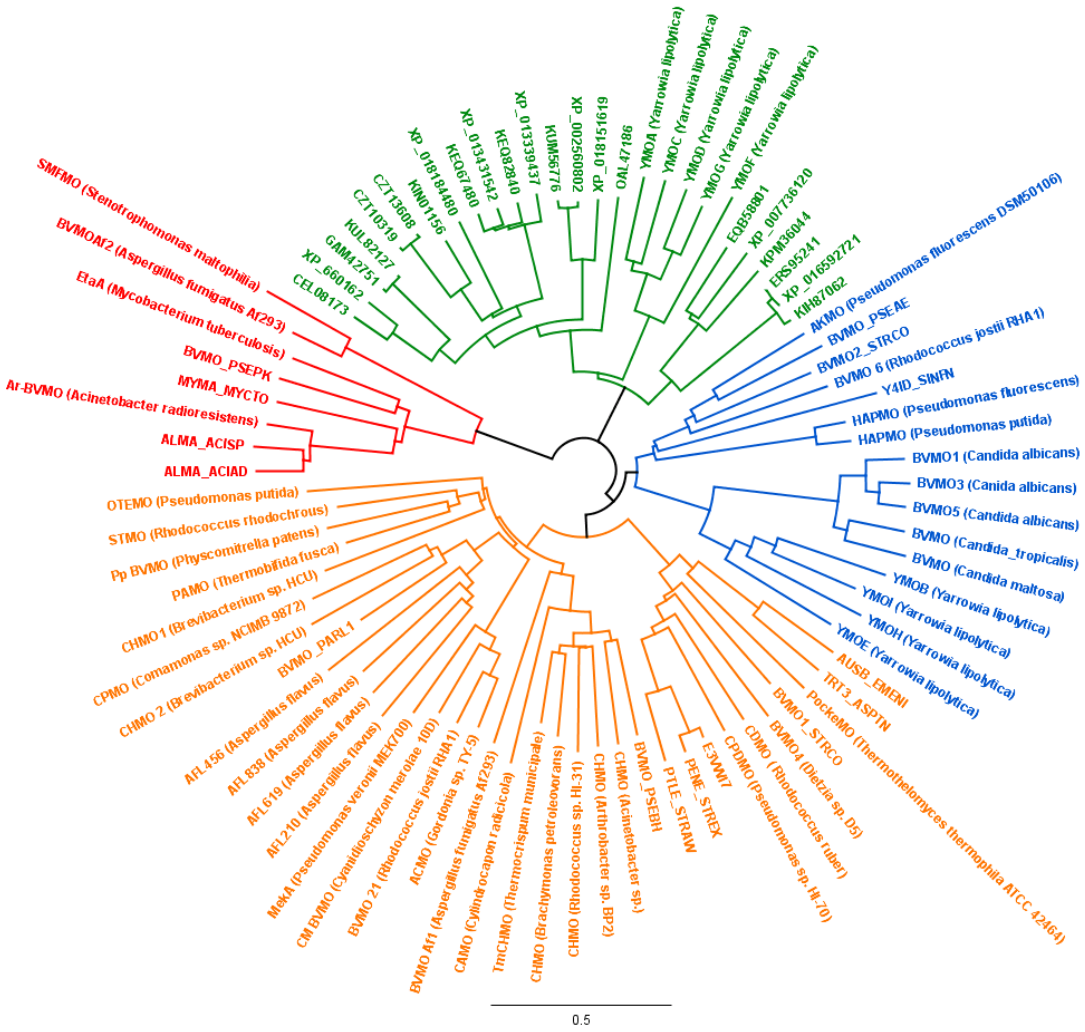


Figure 4.5: Phylogenetic tree of 85 BVMO sequences including $BVMO_{albi1}$, $BVMO_{malto}$ and YMOA-I (adapted from Sven Bordewick).[69]
Characterized BVMOs are shown with name and source organism, uncharacterized ones without source organism, except for the *Candida* and *Yarrowia* enzymes and BVMOAf2.[63] Green: Fungi branch; Blue: HAPMO branch; Orange: Main BVMO branch; Red: EtaA branch.

In this phylogenetic analysis, the BVMO protein sequences were split into four groups.[69] The "EtaA branch" consisted of six bacterial BVMOs, EtaA, Ar-BVMO and four putative ones. EtaA and Ar-BVMO show a quite similar substrate scope and thus it is not surprising that they are clustered together.[176] Moreover, the atypical SMFMO and the uncharacterized BVMOAf2 were placed in this evolutionary line. The majority of the characterized BVMOs was grouped in the "Main BVMO branch".[69] This included the most bacterial BVMOs and all nine eukaryotic BVMOs that have been described before. In contrast, all *Candida* enzymes

including $BVMO_{albi1}$, $BVMO_{malto}$ and the *Yarrowia* monooxygenases YMOB, E, H and I were placed in the "HAPMO branch" and YMOA and the remaining YMOs in the "Fungi branch". This puts all yeast BVMOs in evolutionary lines, which are separated from the other eukaryotic BVMOs. Even the ones originating from other fungi like the *Aspergillus flavus* enzymes were placed into the "Main branch". The "HAPMO branch" could be further splitted as there is one clear bacterial and one fungal line, containing on one side the two known HAPMOs, BVMO6 and AKMO plus three putative BVMOs and on the other side $BVMO_{albi1}$, $BVMO_{malto}$ and YMOB, E, H and I.[69] Inside the "Fungi branch", except for the remaining YMOs, only putative fungal BVMOs are placed. A special characteristic for this branch was an additional C-terminal sequence of about 30 residues with twelve highly conserved residues (the same residue in more than 70% of the sequences).[69] The YMOs in this branch proved to be special due to their variation from the universally conserved arginine at position R274 (YMOA).[69] The clear phylogenetic separation of these YMOs from the ones in the "HAPMO branch" was very interesting as well, especially when keeping in mind that they are originating from the same organism – *Yarrowia lipolytica*. YMOB and $BVMO_{malto}$ are the only fungal BVMOs of the "HAPMO branch" with proven activity towards ketones, making it reasonable that they are grouped together with type I BVMOs like HAPMO. In contrast, YMOA of the "Fungi branch" did not show activity towards ketone substrates, but the formation of sulfones as a main product from the sulfoxidation of sulfides and the conversion of DMSO was detectable (see Chapter 4.1.2.6.3.1). Due to this striking substrate spectrum and its sequence aberrations, already before this classification it was assumed that YMOA might belong to a different group of monooxygenases.[69] This phylogenetic tree verifies this assumption and could explain the abnormal properties by putting YMOA into a group distinctive from all type I BVMO.[69] Furthermore, one could presume that YMOC, D, F and G display a similar activity profile when characterized. As long as no ketone is converted by these monooxygenases they have to be assigned to a distinctive, new class of BVMOs, even though they contain the typical BVMO sequence motifs and sequence similarities to $CHMO_{Acineto}$ of about 20% were determined.[69]

In summary, $BVMO_{malto}$ and YMOB can be clearly classified as type I BVMOs.[69] It is likely that this can be applied to YMOE, H and I as well, even though the acceptance of ketone substrates still needs to be demonstrated.

4.1.2.8 Outlook

The apparently new group of fungal BVMOs with deviating characteristics from type I BVMOs is highly intriguing. Thus, YMOA and the other YMOs of the "Fungi branch" deserve further investigations. Especially a more exhausting substrate screening with YMOA is necessary to see whether it actually does not accept ketone substrates and therefore indeed belongs to a special group of BVMOs.[69] In that case the continuation of the mutational study of YMOA is also needed to find key residues that might be responsible for the absent BVMO activity.[69] In the course of that, structural reasons for the balance of the formation of sulfoxides and sulfones from sulfides could additionally be identified.[69] These findings could be applied to future protein engineering projects. As the formation of sulfones from sulfides is not desired, because chiral sulfoxides are the more valuable compounds, it would be beneficial to be able to switch off this unwanted activity by rational protein design.[69] Furthermore, the ability of YMOA to efficiently oxidize DMSO to $DMSO_2$ (=Methylsulfonylmethane, MSM) is an useful reaction as MSM can be used as a drug for the treatment of osteoarthritis, snoring and seasonal allergic rhinitis and is used as a food supplement, especially in the fitness industry as it has pain releasing properties and reduces oxidative stress and inflammation.[121b, 121c, 143-145, 177] Because of its polarity and thermal stability, MSM is also used in the industry as a high-temperature solvent for both inorganic and organic synthesis reactions.[148] Detailed characterizations of YMOB-H are needed as well. Previous to further investigations, the deviations found in YMOD, F and H have to be adapted to the genomic sequences by employing site-directed mutagenesis to exclude destructive mutations. For $BVMO_{albi1}$, YMOC, E, G and H expression has to be optimized, so that soluble protein without coexpression of chaperones can be obtained. The possibility to get correctly folded and active enzymes is quite high, especially when considering that this worked well for YMOA and B. If necessary, also expression with a C-terminal instead of a N-terminal or even without a $His_{(6)}$-tag needs to be attempted. Subsequently, a large variety of possible substrates has to be utilized to identify substrates for the yeast BVMOs. This could be done in parallel not just with $BVMO_{albi1}$ and all YMOs, including YMOA and B, but also $BVMO_{malto}$ to determine further accepted compounds. It can be expected that all these new BVMOs differ substantially from each other due to quite low sequence identities from about 20-30% to each other leading to the discovery of unique enzymes with interesting activities and overall properties.

As further 16 putative BVMOs and FMOs have been identified in my diploma thesis, they should be investigated as well.[63] Additionally, more genomes of yeasts can be screened with

high chances to find even more potential BVMOs to characterize. Especially interesting should be enzymes from the "Fungi branch" but also the other branches should be further explored. Now that there are examples to successfully express yeast BVMOs, this step will be shortened and simplified a lot. Then, the major goal to screen the enzymes with a broad range of potential substrates needs to be performed to find applications for this class of enzymes. These should range from small aliphatic ketones like acetone and butanone, over longer-chained ones (i.e. 2-dodecanone) to cyclic (i.e. cyclohexanone) and bicyclic (i.e. bicyclo[3.2.0]hept-2-en-6-one) as well as more complex ketones like steroids (i.e. progesterone). As an apparent specialization for heteroatom oxidation is possible, as seen for YMOA and FMOs, different substrates containing sulfur, nitrogen and phosphor have to be included as well.[177] This way, even more interesting yeast BVMOs can be obtained and characterized. Therefore, this work provides intriguing examples and a "how-to" work-flow that could be highly useful for the discovery of eukaryotic BVMOs with outstanding properties (Figure 4.6).

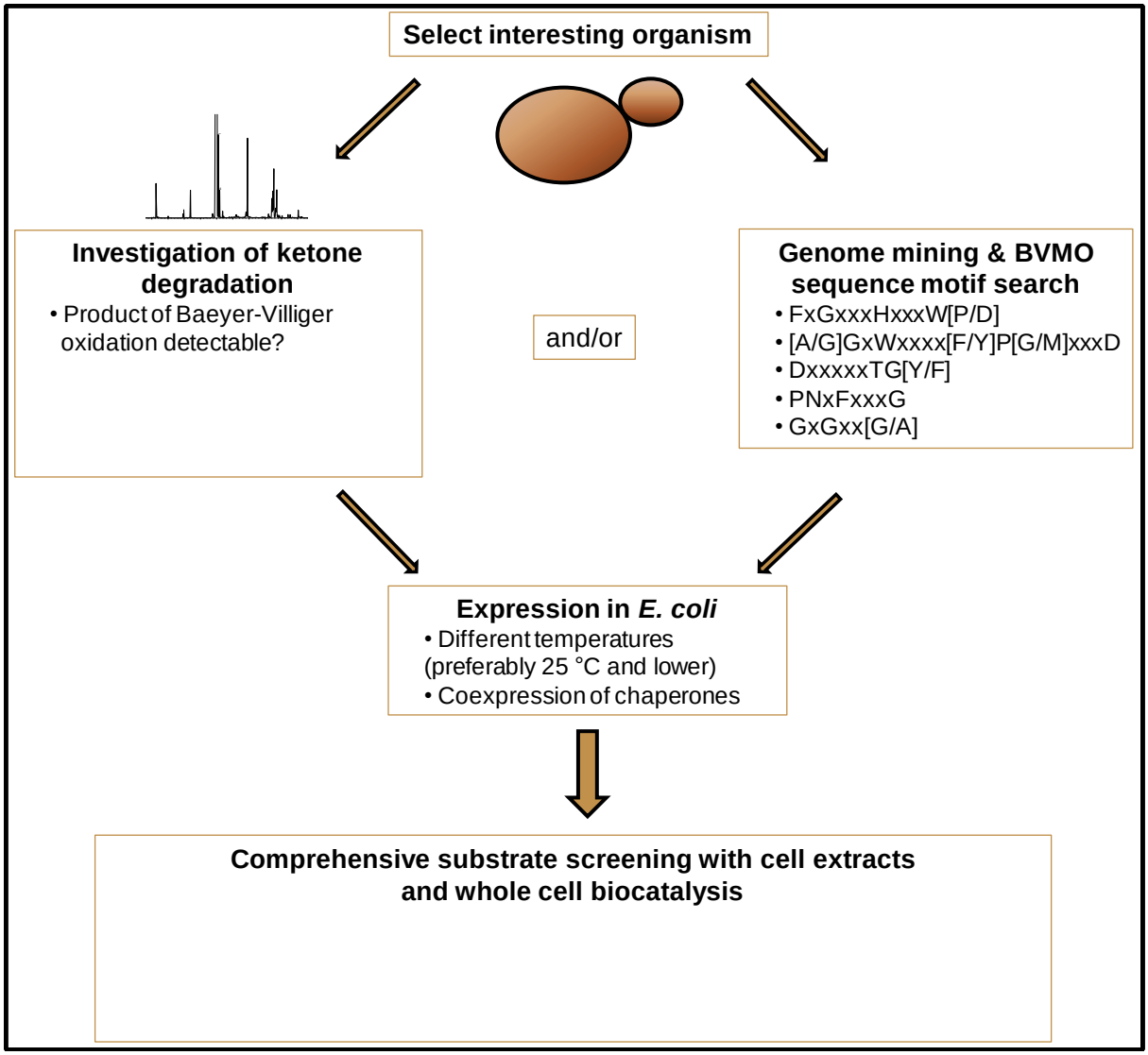


Figure 4.6: Work-flow for the discovery of new BVMOs.
First, an organism with interesting characteristics such as a high temperature optimum or fast growth with ketones needs to be selected, promising a catalyst with outstanding properties. By growth experiments and subsequent metabolite analysis, a Baeyer-Villiger oxidation can be detected, indicating the presence of BVMO(s) in the organism. In parallel or as a different strategy, genome mining can be applied to identify protein sequences in databases homologous to known BVMOs. These sequences can be further checked for the presence of BVMO sequence motifs to validate them as putative BVMOs. Subsequently, expression in bacterial and/or eukaryotic systems for the production of an active enzyme needs to be investigated. Finally, the Baeyer-Villiger monooxygenase is tested for activity in cell extracts and due to the often found instability in whole cell biocatalysis with a variety of possible substrates.

4.2 Switch of the cofactor specificity of $CHMO_{Acineto}$

Examples for engineering the cofactor specificity exist for each direction, from NADPH specific to a preference for NADH and *vice versa*.[37-38, 42a, 43, 46a, 59b, 65a, 74-75, 76b, 85, 115-116, 121] Unfortunately, at the present time there is no generalized approach for cofactor protein design and successful results from one enzyme class are often difficult to transfer to another.[121c] Especially progress for BVMOs has been limited and so far no type I BVMO has yet been engineered to efficiently use the cofactor NADH for catalysis. Only HAPMO, PAMO and $CHMO_{Acineto}$ have been investigated – with limited success.[118h, 145] Therefore, a rational protein engineering approach was employed to change the cofactor specificity of $CHMO_{Acineto}$.

4.2.1 Mutation of the phosphate recognition site

By the use of a homology model of $CHMO_{Acineto}$, six residues in proximity to the phosphate group of NADPH were identified (T184, R207, S208, Q210, K326 and K349; Figure 3.14).[53b, 132] Utilizing structure and sequence alignments, the best possible mutations for the identified residues were predicted and 28 variants of this BVMO were generated (Table 3.6).[53b]

T184 was mutated to asparagine, alanine as well as glycine. All three mutations had a negative impact on the activity; T184N did not even show activity at all (Figure 3.17). The decrease in activity with both cofactors for the T184 variants and the inactive T184N variant illustrate the importance of this residue. The homology model of $CHMO_{Acineto}$ suggests a quite complex role of T184. Interactions via direct or indirect hydrogen bonds can be found to the 2'-oxygen and the adenosine moiety of NADPH and to the residues R207, S208 and Q210. This complex interaction is probably disturbed by any kind of mutation concerning T184. Rebehmed *et al*. also found this position to be conserved with 94%, not allowing much leeway for deviations.[47a]

For R207, the mutation to glutamine was chosen. R207Q displayed decreased activities with both NADPH (97% of WT) and NADH (53% of WT, Figure 3.17). The importance for both cofactors is explained by the stacking interactions of the guanidino group with the adenine moiety of NADPH (Figure 3.14). Mutation R207Q disrupts the hydrogen bonds with the 2'-phosphate, but also decreases binding of both cofactor variants since the important stacking interactions are disturbed. Rebehmed *et al*. found arginine at this position to be universally conserved with 100%.[61] Mutations of this position to alanine or leucine in PAMO and alanine in HAPMO also led to large decreases of the catalytic efficiency with both cofactors (up to 2,800,000-fold for R339A of HAPMO with NADPH).[118h, 145]

S208 has been changed to aspartic acid, glutamic acid, histidine and tyrosine. The positive results for S208D and S208E indicate that the strategy of introducing acidic residues for getting a negative charge at the phosphate recognition site worked (Figure 3.17). The acidic carboxylate group of either aspartate or glutamate seems to successfully hinder the use of NADPH as cofactor, possibly by electrostatically repulsing the 2'-phosphate. Brinkmann-Chen *et al*. published a general approach for the reversal of cofactor specificity of keto-acid reductoisomerases from NADPH to NADH.[121a] Apart from other mutations surrounding the 2'-phosphate, an aspartate was introduced to fulfill the important role of the conserved acidic residue in NADH-dependent enzymes. Conversely, Didierjean, Rahuel-Clermont *et al*. reversed the specificity of glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* from NAD^+ and $NADP^+$, although the resulting mutant displayed a lower activity with $NADP^+$ than the wild type with NAD^+.[59b] Interestingly, they solved crystal structures of their mutants with both cofactors, NAD^+ and $NADP^+$, enabling the direct comparison of the resulting structural features. The mutation with the most beneficial effect on $NADP^+$ activity was D32G, which removed the important acidic residue interacting with the 2'- and 3'-hydroxy groups of NADH. This residue was limiting for NADPH as it repulsed the similarly negative 2'-phosphate. Thus, it can be concluded that acidic residues at the cofactor recognition site are necessary for an efficient use of NADH. Dudek *et al*. only mutated the corresponding residue T218 in PAMO to an alanine.[70c] This did not lead to great changes in enzyme performance. Only a 2-fold increase in both the K_M and the k_{cat} with NADH could be observed. Even for NADPH the k_{cat} was improved 1.2-fold while the K_M increased 2-fold as well. Thus, it was concluded that this residue just plays a marginal role in coenzyme recognition.[70c] However, different amino acid exchanges would be necessary for a detailed analysis of the function of this residue in PAMO. It is likely that an exchange with an acidic residue would create a repulsing effect towards NADPH, making the enzyme more specific for NADH.

When H220 in PAMO (equivalent to Q210 in $CHMO_{Acineto}$) was changed to glutamate or aspartate, this also created a repulsion of the phosphate group, even though performance for H220E and H220D was decreased with both cofactors.[70c] Consequently, the position for the introduction of acidic residues is crucial. Not any residue in proximity of the phosphate group can be turned into an acidic one to efficiently switch the specificity towards NADH.

Residue Q210 of $CHMO_{Acineto}$ was mutated to asparagine, serine and aspartic acid. Q210 was apparently the most influential position for increasing activity with NADH. Q210N increased the activity with NADH in the depletion assay by about 1.5-fold. The specific activity with

NADPH was decreased 2.3-fold. Also mutations of H220 in PAMO to asparagine or glutamine increased efficiencies for NADH 3-fold.[145] However, it turned out that Q210N does not exhibit improved activity but an increased uncoupling, the production of hydrogen peroxide, instead of ε-caprolactone (see Chapter 4.2.3), which was not investigated for PAMO.

For K326, mutations to phenylalanine, histidine, asparagine and arginine were selected. In general, basic residues are of special interest concerning cofactor specificity, because of their possible interaction with the negatively charged 2'-phosphate group.[121e] Indeed, K326 proved to be an important residue concerning cofactor specificity and mutation to histidine was the best choice to increase the specificity towards NADH. One could assume that the structure of MekA is being mimicked by doing this amino acid exchange as this unique BVMO, being able to efficiently use NADPH and NADH, displays a histidine at this position, which could be a reason for its cofactor promiscuity (Figure 4.7).[130] Actually, this theory also applies to HAPMO but does not seem to be general, as activity of PAMO variant K336H towards NADH was decreased.[70c, 119]

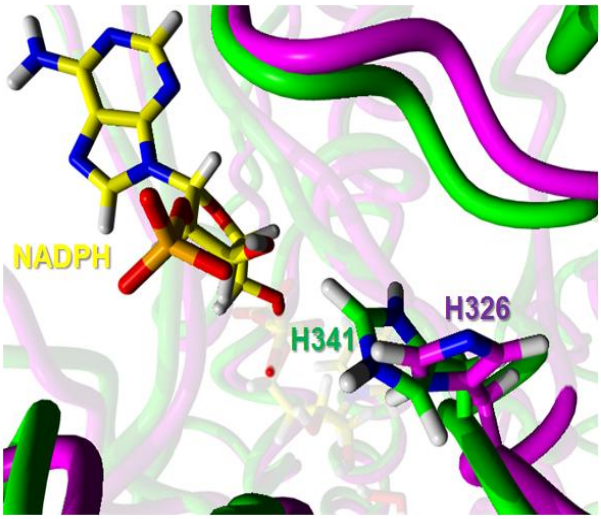


Figure 4.7: Structure alignment of CHMO$_{Acineto}$ variant K326H and the atypical BVMO MekA.[178] The protein scaffold of K326H is shown in magenta, of MekA in green, residues H326 (K326H) and H341 (MekA) were displayed as sticks, NADH is highlighted in yellow, oxygen in red, phosphorus in orange, nitrogen in blue and hydrogen in white.

The terminal amino group of K326 points towards the 2'-phosphate although with an O-N distance of 5.9 Å (Figure 3.14). According to Kumar and Nussinov, this classifies as a long-range ion pair, which is generally destabilizing.[34a] It is possible that the interaction between K326 and the 2'-phosphate is still important for correctly positioning the loop on which K326 is situated. This is especially important since its neighbor, R327, is one of the most important residues for the catalytic cycle.[18, 73] R327 points at the active site, is participating in the catalysis and is strictly conserved among type I BVMOs (see also Chapter 4.1.2.6.3).[18, 73]

When R440 in HAPMO (equivalent to R327 in $CHMO_{Acineto}$) was mutated to alanine, this led to a complete loss of activity.[118h] The interaction of K326 and the 2'-phosphate is not possible with NADH, giving K326 the opportunity to form other interactions and thus changing the conformation of the loop. Thus, R327 cannot efficiently function, leading to a decreased activity with NADH in comparison to NADPH. Mutation of K326 to histidine removes the long range interaction to the phosphate group of NADPH and thus destroys the interaction of this residue with the cofactor, creating the same effect as with NADH in the wild type but also making the specificity more flexible.

K349 was mutated to glutamate and arginine. The latter turned out to be one of the most influential mutations in this project. The activity with NADH in the depletion assay was increased 1.7-fold (Figure 3.17) even though in biocatalysis conversion was apparently not increased (see Chapter 4.2.4). However, one of the two final mutants contained this mutation making it a catalyst quite specific for the cofactor NADH indicating synergistic effects of the introduced mutations (see Chapter 4.2.4). Thus, the strategy of taking the most frequent residue for this position worked out. As the majority of BVMOs displayed an arginine, this had to provide an evolutionary advantage. It might be that the slightly increased size of arginine in comparison to lysine enables more interactions with the residues of the cofactor recognition site, leading to a better organization and stabilization of this region. Actually, when comparing the homology models of the wild type and K349R, one can see that R349 is close enough to R207 and S208 to form hydrogen bonds with them, while this is not the case for K349 in the wild-type enzyme (Figure 4.8).

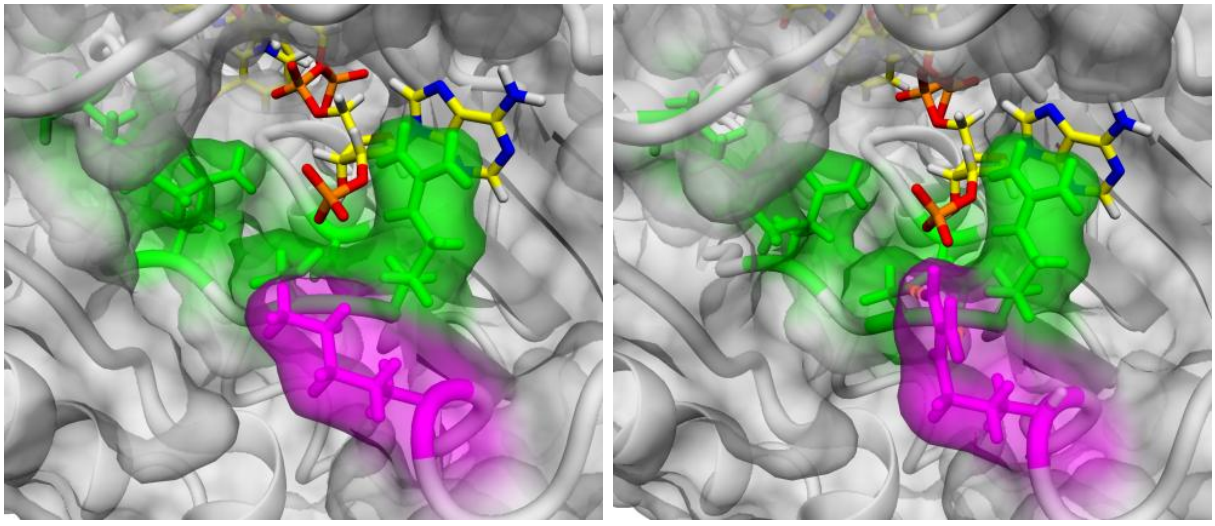


Figure 4.8: View at the phosphate-binding site of $CHMO_{Acineto}$ wild type (left) and K349R (right). The protein scaffold and its molecular surface are shown gray, residues T184, R207, S208, Q210 and K326 were displayed as sticks in green, K349(R) in magenta, NADH is highlighted in yellow, oxygen in red, phosphorus in orange, nitrogen in blue and hydrogen in white. Yellow dashed lines indicate hydrogen bonds of R349 with R207 and S208.

Moreover, a clear difference regarding the NADPH binding tunnel can be seen. R349 is slightly decreasing the diameter of the tunnel that probably serves as the entry for NADPH. This could be the reason for an increased specificity towards the smaller NADH. This residue was mutated neither in HAPMO nor PAMO so that the influences of arginine and other amino acids at this position need to be investigated for further BVMOs to make general conclusions.

To sum this up, the four mutations S208D, S208E, K326H and K349R proved to positively influence the activity and specificity of $CHMO_{Acineto}$ towards NADH.

The combinatorial variants S208E_K326H_K349R, S208E_Q210N and S208E_K326H showed increased activities over the wild type in the NADH depletion assay (Figure 3.17). Interestingly, even though both double variants had a lower activity than their parent variants K326H and Q210N, the activity ratio NADH/NADPH is much better for these variants, being 7.60 for S208E_K326H_K349R, 7.49 for S208E_Q210N and 3.33 for S208E_K326H in comparison to 0.01 (K326H), 0.02 (Q210) and 0.003 (WT). This is equivalent to a more than 2,600-fold improvement in specificity for NADH (Figure 3.18). Here, the described influences of the mutations resulted in synergistic effects leading to a further decrease in activity with NADPH, increasing the specificity for the other cofactor. However, in the case of S208E_K326H_K349R, substrate conversion using NADH was additionally increased. All the other combinatorial mutants showed either activity comparable to the WT or even lower, indicating a high disturbance of the protein functionality through the introduction of multiple mutations. For PAMO, two combinatorial variants have been designed to increase cofactor specificity.[70c] H220Q_K336H/N (corresponding to Q210_K326H/N) displayed activity ratios NADH/NADPH of 0.35 and 0.36. In comparison to 0.13 of the wild-type enzyme this was just a 2.69-fold improvement in specificity. With H220D they could achieve a value of 0.63, meaning a 4.85-fold improvement. In other words, the improvement of $CHMO_{Acineto}$ to that point in this thesis was more than 2,555-fold better, emphasizing the well selected mutations and combinations here.

4.2.2 Mutation of residues in proximity of NADPH

In order to increase the activity of $CHMO_{Acineto}$ with NADH to the level of the wild type with NADPH, more residues in proximity of the cofactor were mutated. Out of 24 residues identified, six (L55, D57, S186, T187, F380 and W490) showed direct or indirect hydrogen bonds towards the cofactor molecule from which all, except W490, are situated in conserved regions of the enzyme. Likewise, the three residues I182, V189, and T378 were chosen as they are directly adjacent to the Rossmann fold (I182, V189) or placed in the "fingerprint 3" (Appendix Figure 8.9). Using the multiple protein sequence alignment (Appendix Figure 8.6), rational mutations were chosen for these residues, so that twelve further enzyme variants were generated (Table 3.7).

Mutation of these residues turned out to be a good strategy as the variants L55R, S186P, T187L and W490Y displayed an increased activity with NADH (Figure 3.19). The change of residue L55, which shows a hydrogen bond to the nicotinamide function of NADPH, to arginine seemed to make the cofactor specificity more flexible by removing this interaction. D57 is a conserved residue among BVMOs, being a key residue in the catalytic cycle.[61] Thus, probably any change of this aspartate is destroying the activity of the enzyme like it happened here with $CHMO_{Acineto}$. Mutations directly within the Rossmann fold (S186P, T187L) increased the activity with NADH. S186P leads to small conformational changes of the Rossmann fold and in turn to a different positioning of the NADH resulting in higher flexibility and likely a better electron transfer. It also affects positioning of G185 by bringing it 0.45 Å closer to the phosphodiester bond of NADH for a more stable hydrogen bond through a bridging water molecule. All these changes also lead to a decreased uncoupling of the BVMO with NADH, making the enzyme a more efficient catalyst (see Chapter 4.2.3). T187 displays two salt bridges to the second phosphate group inside the ester bond of NADPH. In T187L one of them is destroyed, increasing the flexibility of the cofactor and widening the acceptance for NADH. Brinkmann-Chen *et al*. reported an approach to reverse the cofactor dependence of ketol-acid reductoisomerases (KARIs) from NADPH to NADH also by mutations of the Rossmann fold.[121a] By mutation of two serines to aspartates, they could increase the catalytic efficiency of the Ll_KARI 8.6-fold. In KARIs only this conserved region is responsible for recognition of the cofactor, though. Thus, it is not surprising that exchanging residues of that fold leads to a switch in cofactor specificity. However, currently no positive mutation within the Rossmann fold of BVMOs has been reported, making these findings rather unique. Interestingly, variant W490Y showed an increased activity with both cofactors. However, later investigations revealed that the increased activity of W490 mainly

comes from a higher uncoupling rate (see Chapter 4.2.3). W490Y destroys the indirect hydrogen bond to the phosphodiester bond of NADPH but leads to reorientations so that a new indirect hydrogen bond through two bridging water molecules to the phosphate group of the cofactor is established. As this interaction works with two water molecules, it probably is quite unstable leading to disturbances of the structure and the electron transfer. However, the detailed consequences of the reorientations will stay elusive.

Six mutations (R50L, T139L, V253Y, F284Q, D341C and D347V) were transferred from BmoF1 as they proved to be beneficial for this BVMO.[146] These residues, except T139 and V253, apparently are quite conserved among BVMOs (R50: 55%, T139: 18%, V253: 16%, 284: F53%, D341: 37%, D347: 58%). Thus, it is likely that they play an important role for the structure and/or activity of the enzyme like it is the case for residues situated in the Rossmann fold and the fingerprint motifs, even though perhaps not to the same extent.[34a, 58a, 60, 62] The mutations V253Y, F284Q and D341C resulted in improved activity with NADH and thus were beneficial for $CHMO_{Acineto}$, too (Figure 3.19). Mutation V253Y increased the enantioselectivity of BmoF1, but increased the activity of this CHMO. A higher overall activity, mainly due to an increased uncoupling (see Chapter 4.2.3), was found for D341C as well, pointing out that the effects of the mutations taken from BmoF1 are non-transferable and rather vary for different BVMOs.[146] They have to be applied to more BVMOs to give more concrete patterns. Because of their conservation, the residues V253, F284 and D341 must be important for the functionality of BVMOs. However, the function of these residues is not known yet and is difficult to rationally predict as they are situated at the surface of the structure (Figure 4.9).

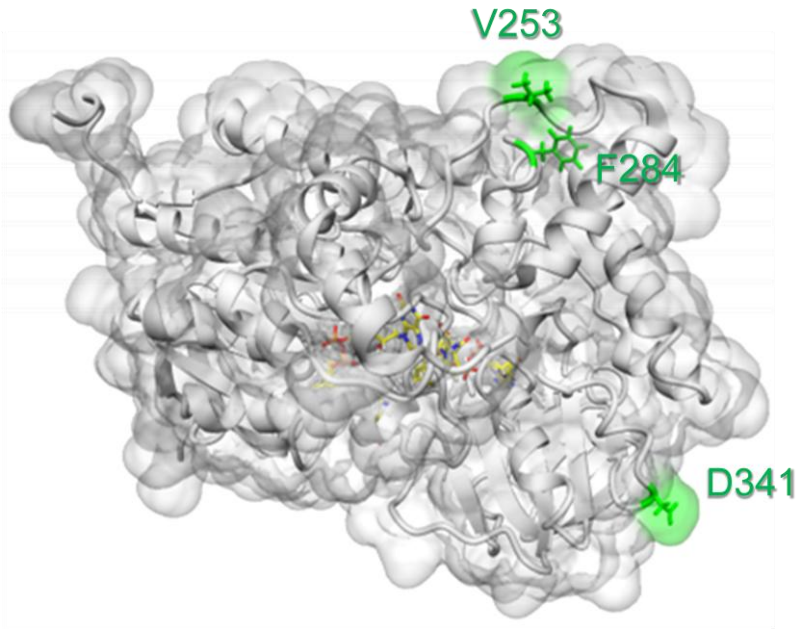


Figure 4.9: View at the homology model of $CHMO_{Acineto}$.
The protein scaffold and its surface are shown gray, residues V253, F284 and D31C were displayed as sticks in green, NADH and FAD are highlighted in yellow, oxygen in red, phosphorus in orange, nitrogen in blue and hydrogen in white.

In contrast, the function of residues mutated in HAPMO and PAMO is better understood, but still some mutations led to differing results in $CHMO_{Acineto}$ as described above.[53b, 118h, 145] F284Q, which increased conversion with NADPH in BmoF1, did not have the same effect in CHMO.[146] In contrast, conversion with NADH was improved. However, expression of this mutant was very poor and thus the mutation seems to destabilize the structure of this enzyme.

Next, various mutations were combined and resulting variants tested for increased activity. In the course of that, the two best variants were discovered – S186P_S208E_K326H and S186P_S208E_K326H_K349R. They displayed a 9- and 8-fold improvement in activity with NADH in the depletion assay, respectively (Figure 3.19). When looking at NADH specificity, a 1,920 and 4,170-fold increased activity ratio (NADH/NADPH) could be achieved. Consequently, the additional mutation S186P added a further 1570-fold improvement in specificity towards NADH in comparison to the best combinatorial variant S208E_K326H_K349R discussed above (see Chapter 4.2.1). This means, that the effects caused by this mutation fit well in this variant, synergistically leading to a rather specific enzyme for the new cofactor (Figure 3.18). Compared to the best example in literature for a "cofactor-switched" BVMO, variant H220D from PAMO, a more than 4,165-fold higher improvement could be reached with S186P_S208E_K326H_K349R, illustrating the achievement accomplished here.[70c]

4.2.3 Kinetics and uncoupling of $CHMO_{Acineto}$

To further validate the introduced mutations, kinetic values and uncoupling – formation of hydrogen peroxide from NAD(P)H instead of ε-caprolactone formation – were compared between the WT and selected variants. Kinetics of the wild type and the triple variant S186P_S208E_K326H differ substantially (Table 3.8). The catalytic efficiency with NADH for this variant is 8-fold higher, the K_M is 2.5-fold lower and the k_{cat} is 3-fold higher making NADH a well-accepted cofactor for $CHMO_{Acineto}$. For HAPMO, a 6.7-fold increase in catalytic efficiency, 4.8-fold lower K_M and a 1.4-fold higher k_{cat} was obtained with mutation of K439 (corresponding to K326 in $CHMO_{Acineto}$) to phenylalanine.[119] Kinetic analysis of PAMO mutants revealed variant H220N (equivalent to Q210N) with 3.3-fold improved catalytic efficiency, 2.2-fold decreased K_M and a 1.5-fold increased k_{cat}.[70c] In conclusion, as catalytic efficiency is the most convincing value for an enzyme, for variant S186P_S208E_K326H of $CHMO_{Acineto}$ the biggest improvement could be accomplished, especially when keeping in mind the total values of 8 (HAPMO_K439), 2.3 (PAMO_H220N) and 10.4 $mM^{-1} \times s^{-1}$ (CHMO_ S186P_S208E_K326H).

Furthermore, the major part of the activity of the WT with NADH comes from uncoupling, k_{cat}_unc being 83% of k_{cat} in that case. Thus, the improvement of the mutant is even higher, showing a decreased uncoupling of just 15%. When comparing the formation of H_2O_2 of other selected variants of $CHMO_{Acineto}$, one can clearly see the effects of the mutations introduced into this BVMO (Table 3.9). The majority of the variants showed a higher uncoupling compared to the wild type (115%) under the conditions tested. In these cases, the ratio of the activity with NADH alone to the activity with the substrate cyclohexanone was even higher than for the wild type. This explains the apparently increased activity of the variants Q210N, Q210S, K326N, Q210N_K326H with NADH in the depletion assay (Figure 3.17 andFigure 3.19). Interestingly, formation of hydrogen peroxide from NADH in general was higher than the production of ε-caprolactone from cyclohexanone, except for K349R, S208E_Q210N, S186P_S208E_K326H and S186P_S208E_K326H_K349R. Actually, due to these data, one could assume that the activity with NADH for the wild type and most of the variants is completely uncoupled. However, this was refuted by biocatalysis in which even the wild type produced ε-caprolactone with NADH, using higher cofactor and substrate concentrations (see Chapter 4.2.4). Nevertheless, S186P, K349R, S208E_Q210N, S186P_S208E_K326H, S208E_K326H_K349R and S186P_S208E_K326H_K349R showed decreased uncoupling rates. Especially the quadruple mutant exhibiting 54% uncoupling is improved as it showed 61% less uncoupling than the wild type. For HAPMO and PAMO, this was not investigated,

making their increased activities with NADH questionable.[70c, 119] As BVMOs can convert NAD(P)H solely by producing H_2O_2, it always needs to be checked whether the conversion of the substrate or just the uncoupling is increased.[53b] For instance, by introduction of the three mutations R217T, T218E and K336Y, Jackson Cahn successfully engineered PAMO to consume NADH faster than NADPH.[56] However, this activity was found to be nearly completely uncoupled. Thus, one needs to be aware of a possible increase in uncoupling when introducing mutations into BVMOs. After finding positive variants, they need to be checked for a higher formation of hydrogen peroxide instead of product formation, ideally by investigation of their activity in biocatalysis like discussed in the following chapter.

4.2.4 Biocatalysis with $CHMO_{Acineto}$

Next, the best variants, according to the activities seen in the NADH assay, were investigated in regard to their performance in biocatalysis to validate their applicability in industrial processes (Figure 3.20).

In contrast to the specific activities seen in the NADPH depletion assay (Figure 3.17 and Figure 3.19), most mutants still gave, like the wild type, 50-60% conversion using NADPH except for the triple mutant S208E_K326H_K349R (27% conversion). Apparently, the low activities of the mutants in the NADPH depletion assay are due to higher K_M values as they were as active as the wild type in the biocatalysis with 5 mM NADPH and cyclohexanone while in the spectrometric assay just 0.3 mM and 1 mM were used, respectively. The increased NADPH concentration also led to inhibiting effects of the WT but not of its variants, which can be considered as another improvement. However, some variants did not show improved conversions with NADH as expected from the NADH depletion assay (FiguresFigure 3.17 andFigure 3.19), indicating uncoupling and thus verifying the uncoupling observed before by activity tests without substrate (Table 3.8 andTable 3.9, see previous Chapter 4.2.3). The significantly improved activities and increased NADH preference of S186P and especially of the combinatorial variants S186P_S208E_K326H, S208E_K326H_K349R and S186P_S208E_K326H_K349R correspond to these findings. With the triple and quadruple mutants S186P_S208E_K326H and S186P_S208E_K326H_K349R >79% conversion with NADH could be achieved (cf. 10% WT), exceeding even the values determined using NADPH. This is the first example of a type I BVMO that could be engineered to prefer NADH over NADPH, especially when keeping in mind that for the wild-type enzyme the activity for NADH was substantially lower. Both S186P_S208E_K326H and S186P_S208E_K326H_K349R can be considered the best variants in different points of view. They showed an equally high conversion of cyclohexanone when utilizing NADH. The triple

mutant retained an activity with NADPH comparable to the wild type (48% vs. 58% conversion) making it the BVMO being useful for a broader variety of reactions. The quadruple variant is the more specialized enzyme, as it displayed 42% conversion with NADPH, coming closer to the goal of designing a type I BVMO only accepting NADH as cofactor.

Intriguingly, K349R and S208E_Q210N did not show an improved conversion even though they displayed a reduced uncoupling. Thus, stability issues could limit their activity. If the stability of the enzyme is disturbed by the introduced mutations, the incubation of two hours could have been enough to at least partially inactivate it, decreasing the total product formation. Inhibiting effects of the higher cofactor and substrate levels in biocatalysis are also conceivable. For S208E, K326H and the respective double mutant S208E_K326H, not a decreased uncoupling but solely the structural reasons discussed above (see Chapter 4.2.1) and the higher substrate and cofactor concentrations have to be the reason for their higher conversion of cyclohexanone using NADH as cofactor. Likewise, S208D, D341C and W490Y showed a slightly increased conversion even though formation of H_2O_2 was increased (n.d. for S208D). Dudek *et al*. also tested their PAMO variants with improved specificities towards NADH in biocatalysis.[70c] With H220N, showing a 3.3-fold improved catalytic efficiency in the NADH depletion assay, a higher conversion of MPS could be achieved as well (37% vs. 14% wild-type performance). Conversion of 3-methyl-4-phenylbutan-2-one was decreased from 32% to 17%, though. However, to determine kinetics they used the substrate phenylacetone and in biocatalyis the two other substrates, making a clear comparison of both approaches rather difficult. As mentioned before, uncoupling was not investigated in their study, which is a hindrance for a good comparison to the $CHMO_{Acineto}$ variants. Thus, this is the first example for a detailed analysis of mutations affecting uncoupling and substrate conversion and establishing a link between the two.

4.2.5 Structural investigation of S186_S208E_K326H

As the variant S186_S208E_K326H showed a 1,920-fold improved specificity for NADH (see Chapter 4.2.2), a reduced uncoupling of 15-74% (depending on the conditions), a 8-fold improved catalytic efficiency (see Chapter 4.2.3) and an excellent performance in biocatalysis with NADH, while still retaining about wild-type activity with NADPH (see Chapter 4.2.4), it was considered the best $CHMO_{Acineto}$ mutant. The good performance of this variant can be explained by its extensive network of hydrogen bonds enabled by the right combination of mutations (Figure 4.10).

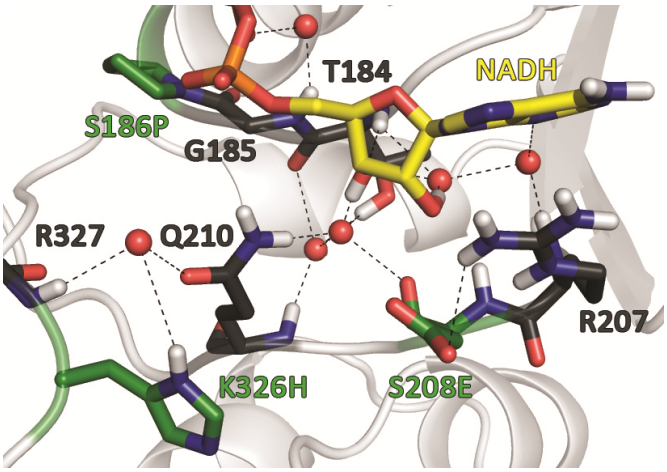


Figure 4.10: View at the NADH binding pocket of the triple mutant S186P_S208E_K326H in the homology model of $CHMO_{Acineto}$.
The protein scaffold is shown gray, important residues were displayed as sticks (black: WT, green: mutated), NADH is highlighted in yellow, oxygen in red, phosphorus in orange, nitrogen in blue and hydrogen in white. Water molecules are represented as red spheres. Dashed lines indicate hydrogen bonds.

Both hydroxyl groups of NADH are positioned through water molecules that are coordinated either directly or through bridging water molecules by the residues T184, R207, E208 and Q210. There is also a contact between E208 and R207 that stabilizes the binding pocket. Additionally, Q210 and H326 interact with the catalytically important R327 through a water molecule. As mentioned before (see Chapter 4.2.2), S186P is a crucial mutation as it changes the conformation of the Rossmann fold, which positions the NADH in a slightly different angle. Presumably, this compensates to some degree for the slightly different conformations of NADPH and NADH enabling a better electron transfer, which is a must for a cofactor-switched BVMO to function. In a study of an isocitrate dehydrogenase, high-resolution crystal structures demonstrated that a changed adenine binding, caused by subtle chemical modifications of the adenine ring of NADPH, caused a misalignment of the nictotinamide, leading to a significantly decreased turnover.[123a] In 2013, by investigation of aldo-keto reductases, it

was also concluded that the arrangement of the cofactor and substrate was the most important factor influencing the catalytic activity.[123d]

It is possible to find common properties of this model to the structure of the cofactor switched Se_KARI^{DDV} (Figure 4.11).[121a]

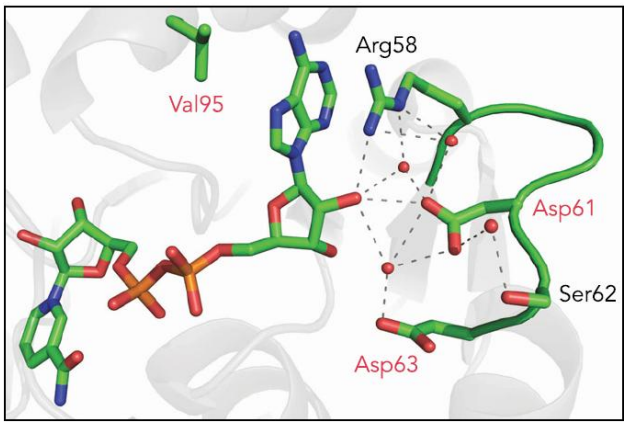


Figure 4.11: Crystal structure of variant Se_KARI^{DDV} with cocrystallized NADH (from Brinkmann-Chen).[121a]
Side chains involved in defining cofactor-specificity are shown as sticks. Introduced mutations (Ser61Asp, Ser63Asp, and Ile95Val) are shown with red labels.

The two mutated serines, S61D and S63D, are positioned like E208 and Q210, enabling direct and indirect hydrogen bonds to the 2'-OH group of NADH and within the whole cofactor recognition site. Additionally, the introduced aspartates result in electrostatic repulsion to the 2'-phosphate of NADPH. R58 takes over the role of R207, forming cation-pi stacking interactions with the adenine moiety. The discussed shift of the cofactor, achieved with the mutation S186P in CHMO$_{Acineto}$, is accomplished in Se_KARI^{DDV} by changing I95 to valine.[121a] Thus, it can be concluded that the cofactor binding in different enzymes is comparable and also similarities between cofactor-switched enzymes can be found. However, to create an enzyme to efficiently use another cofactor, involved residues and respective mutations have to be validated for every single enzyme as they are in many cases difficult to transfer from one to another.

The in 2016 published automated approach "Cofactor Specificity Reversal−Structural Analysis and Library Design" (CSR-SALAD) has been employed on the homology model of CHMO$_{Acineto}$ for the purpose of comparison (Table 4.3).[118h]

Table 4.3: Analysis results for the homology model of $CHMO_{Acineto}$ with CSR-SALAD.

Residue	Type	Codon	Amino acids
T184	Motif	RMC	ADNT
R207	Face	HVC	CHNPRSTY
S208	Simple	RRC	DGNS
Q210	Nonsimple	SAA	EQ
K326	Peripheral	RWK	DEIKMNV

A maximum library size of 3000 was selected for the determination of the mutations for a cofactor switch; the suggested library size then was 2304. Medium priority residues for activity recovery by site-saturation mutagenesis were N148, P150, N151, I152, T378 and N495; low priority was designated to K349.

Some of the applied mutations could be found by the online tool as well. However, only T184A, S208D and K326N are among the predicted mutations, the first even being a deleterious mutation. Especially worth mentioning again is the residue S186, which could not be identified by CSR-SALAD, making the "manual efforts" presented in this thesis even more valuable. Not all predicted mutations were tried though, so that this would be necessary to compare the manual selection efforts and predictions by the tool in a more comprehensive way. Also suggested by this application is the saturation mutagenesis of the four positions N148, P150, T378 and N495, which are close to the adenine moiety of NAD(P)H. These residues were among the 24 identified positions in proximity of the cofactor (Appendix Table 8.2) and their mutagenesis could further boost the activity of $CHMO_{Acineto}$ (see next section).

4.2.6 Outlook

By random mutagenesis and/or mutations of the adenine binding pocket, the activity of S186_S208E_K326H could be further improved. Brinkmann-Chen *et al*. reported about the introduction of random mutations in cofactor switched KARIs, which actually boosted the activity with NADH of some enzymes to levels higher than wild-type activity with NADPH.[121a] These were surface mutations making their effects difficult to predict rationally. Additionally, mutations of the adenine binding pocket were described to recover the wild-type activity of KARIs and other enzymes with the naturally disfavored cofactor by readjusting the catalytically active nicotinamide moiety of NAD(P)H to take on a more favorable position for electron transfer.[118h, 121a, 121d] Actually, residues likely being involved in binding the adenosine moiety of NAD(P)H in $CHMO_{Acineto}$ were identified in this thesis (N148, P150, T378, N495), but not mutated due to inherent time constrictions (Figure 4.12, Appendix Table 8.2).

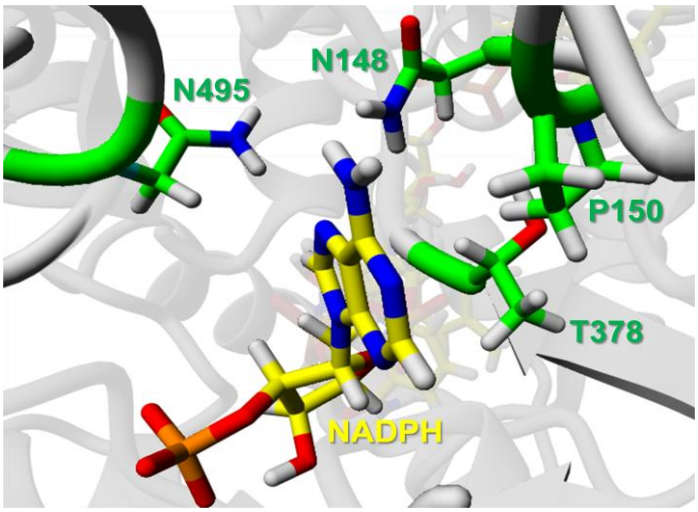


Figure 4.12: View at the NADPH binding pocket of CHMO$_{Acineto}$ with focus on the adenosine binding. The protein scaffold is shown gray, important residues were displayed as sticks in green, NADH is highlighted in yellow, oxygen in red, phosphorus in orange, nitrogen in blue and hydrogen in white.

Thus, this project could continue by creating saturation libraries (either single-site or multiple-site) targeting these residues, possibly supported by random mutagenesis to further boost the activity of this BVMO with NADH.

5 Summary

Baeyer-Villiger monooxygenases (BVMOs) are versatile biocatalysts for the synthesis of numerous compounds. The further exploration of the little known eukaryotic representatives of this class of enzymes offers great opportunities to find highly useful candidates for the biocatalytic toolbox of industrial applications.

In this PhD thesis, it was confirmed that *Candida maltosa* can convert 1-dodecene to 2-dodecanone and this aliphatic ketone to decyl acetate resulting from a BVMO activity, which can also be found in the most of the here investigated hydrocarbon oxidizing yeasts.[63, 133] The reduction of 2-dodecanone to its corresponding secondary alcohol 2-dodecanol was additionally shown.[133]

Of nine in my diploma thesis identified putative BVMOs, two – $BVMO_{albi1}$ and $BVMO_{malto}$, originating from the yeasts *Candida albicans* and *Candida maltosa*, respectively – were expressed heterologously.[63] The latter could be obtained in an active form. In a whole cell biocatalysis it displayed activity towards bicyclo[3.2.0]hept-2-en-6-one, 2-dodecanone and methyl phenyl sulfide (MPS). The much higher conversion of bicyclo[3.2.0]hept-2-en-6-one (22.4%) in comparison to the other substrates (5.6% and 2.2%, respectively) indicated a specialization of $BVMO_{malto}$ for bicyclic or maybe even bulkier ketones. In the genome of *Yarrowia lipolytica*, further nine potential BVMOs, designated *Yarrowia* monooxygenases (YMOs) A-I, were identified, of which eight were cloned and recombinantly expressed in a heterologous host – *E. coli*. Two of them, YMOA and B, were investigated in a master thesis, which was co-supervised by me.[69] YMOA did not act as a typical type I BVMO as it did not convert any of the tested ketones, but displayed promiscuous sulfoxidations of methyl *p*-tolyl sulfide (MTS), MPS and DMSO with unexpectedly high sulfone/sulfoxide ratios.[69] By a structure- and sequence-oriented protein engineering approach, key residues for manipulation of these ratios and inversion of the enantioselectivity for MTS were identified.[69] YMOB, on the contrary, showed typical type I BVMO activity by accepting 2-dodecanone, MTS, its corresponding sulfoxide methyl *p*-tolyl sulfoxide and methyl phenyl sulfoxide.[69] YMOC, D and F were only obtained in insoluble but YMOE, G and H in soluble form. However, with YMOC-H no conversion of the tested substrates could be detected.

To classify the newly discovered enzymes, a phylogenetic analysis containing 85 BV MO protein sequences was employed.[69] According to this and their respective substrate spectra $BVMO_{malto}$ and YMOB are type I BVMOs.[69] YMOE, H and I were grouped together with them, but suitable substrates still need to be found for these enzymes. In contrast, the branch in which YMOA, C, D, F and G were placed in, was classified as a putative new class of fungal BVMOs with an apparent absence of Baeyer-Villiger oxygenation activity.[69]

Moreover, the putative FMO *Candida* monooxygenase (CMO) from *C. maltosa* could be obtained as soluble protein after expression optimization and switching the method for cell lysis from FastPrep to supersonication. When the chaperone plasmid pG-Tf2 was coexpressed even higher amounts of CMO were successfully produced. However, without optimization of the assay conditions, no activity could be detected yet.

To expand the knowledge about the interesting class of Baeyer-Villiger monooxygenases, conserved motifs to be found within their sequences were analyzed in detail. A multiple sequence alignment of 56 BVMO sequences originating from different organisms verified the conservation of the "fingerprints 3 and 4" identified in my diploma thesis. By comparison to an alignment with 116 sequences of Rebehmed *et al*., the motifs were set to DxxxxxTG[Y/F] and PNxFxxxP, respectively and proved to be suitable for the identification of the majority of BVMOs in databases.[61] This could lead to a more reliable discovery of even more representatives of these interesting monoxygenases.

Furthermore, in order to investigate the strict cofactor preference of the cyclohexanone monooxygenase from *Acinetobacter* sp. NCIMB 9871 ($CHMO_{Acineto}$), a combinatorial mutagenesis approach of structure analysis and sequence alignments together with literature data was applied leading to the generation of 46 enzyme variants. This strategy enabled to substantially alter the cofactor specificity of this BVMO from NADPH to the much cheaper and more easily recyclable NADH. Interestingly, out of the 21 positions targeted, mutations at the phosphate binding site of NADPH alone were not sufficient to significantly increase the affinity for NADH. Therfore, the strict preference of type I BVMOs for NADPH was overcome for the first time. This could provide new insights into the structural reasons to engineer cofactor specificity. These findings should be highly useful for the protein engineering of other NAD(P)H-dependent enzymes as well.

6 Material and Methods

6.1 Equipment

Table 6.1: Equipment.

Category	Name	Manufacturer
Agarose gel electrophoresis	Mini-Sub Cell GT	Bio-Rad (München)
Autoclave	V-120	Sys Tec (Bergheim-Glessen)
	Laboklav	SHP Steriltechnik AG (Detzel)
Balances	Explorer E14130	Ohaus (Pine Brook, NJ, USA)
	MC1 Analytic AC 120S	Sartorius (Göttingen)
	PCB2500-2	Kern & Sohn GmbH (Balingen)
	PCB350-3	Kern & Sohn GmbH (Balingen)
Cell disruption	Sonoplus HD2070	Bandelin (Berlin)
	FastPrep24®	MP Biomedicals (Illkirch Cedex, Frankreich)
Centrifuges	Galaxy 16DH	VWR (Darmstadt)
	Heraeus Biofuge pico	Thermo Scientific (Waltham, MA, USA)
	Heraeus Fresco 17	Thermo Scientific (Waltham, MA, USA)
	Heraeus Labofuge 400R	Thermo Scientific (Waltham, MA, USA)
	Heraeus Multifuge 3S-R	Thermo Scientific (Waltham, MA, USA)
	Sorvall RC-5B Plus Refrigerated Superspeed Centrifuge	Thermo Scientific (Waltham, MA, USA)
Cleanbench	HeraSafe KS15	Thermo Scientific (Waltham, MA, USA)
GC	GC-2010 plus	Shimadzu (Duisburg)
GC-MS	7890A GC System + 5975C inert XL EI/CI MSD with Triple-Axis detector	Agilent (Böblingen)
GC columns	FS-Hydrodex β-3P	Macherey-Nagel (Düren)
	FS-Hydrodex ß-TBDAC	Macherey-Nagel (Düren)
	SolGel-WAX	Trajan Scientific Europe Ltd (UK)
	Zebron® DB5 MS	Phenomex (Aschaffenburg)
Incubation shakers	Minitron	Infors AG (Bottmingen)
	Multitron	Infors AG (Bottmingen)
	Unitron	Infors AG (Bottmingen)
Incubators	Friocell	MMM Medcenter-Einrichtungen GmbH (Gräfeling)
	Incucell	MMM Medcenter-Einrichtungen GmbH (Gräfeling)
pH meter	pH 211 Microprocessor	Hanna Instruments (Kehl am Rhein)

Photometers	NanoDrop 1000	Thermo Scientific (Waltham, MA, USA)
	UVmini-1240	Shimadzu (Duisburg)
	V-550	Jasco (Easton, MD, USA)
Plate reader	Infinite® 200 PRO series	Tecan Group Ltd. (Männedorf, CH)
PCR cyclers	T-Personal Thermal Cycler	Biometra (Göttingen)
	Thermocycler Progene	Techne (Cambridge, UK)
	Touchgene Gradient	Thermo Electron (USA)
	PXE 0.2	Thermo Electron (USA)
Rotary evaporator	Vacuubrand® PC 3001 Vario™	Vacuubrand GmbH und Co KG (Wertheim)
SDS-PAGE	Minigel-Twin	Biometra (Göttingen)
Thermocyclers	FlexCycler	Analytik Jena (Jena)
	T.Personal	Biometra GmbH (Göttingen)
Thermoshaker	Thermomixer comfort	Eppendorf (Wessling-Berzdorf)
Ultrapure water	Milli-Q Reference	Merck Millipore (Billerica, MA, USA)
UV table	Benchtop UV Transilluminator	UVP (Upland, CA, USA)
vortex	7-2020Vortex-Genie®	neoLab (Heidelberg)

6.2 Chemicals

Unless stated otherwise all chemicals were purchased from Sigma-Aldrich (Steinheim, Germany), Fluka (Buchs, Switzerland), New England Biolabs (Ipswich, MA, USA) or Merck (Darmstadt, Germany). Primers were synthesized by Invitrogen and Eurofins MWG Operon. Sequencing was done at Eurofins MWG Operon.

6.3 Buffers, growth media and solutions

Table 6.2: Buffers and solutions.

Name	Preparation
mineral salt medium[179]	5 g $NH_4H_2PO_4$ 2.5 g KH_2PO_4 1 g $MgSO_4 \times 7\ H_2O$ 0.02 g $Ca(NO_3)_2 \times 7\ H_2O$ 2.0 mg $FeCl_3 \times 6\ H_2O$ 0.5 mg H_3BO_3 0.4 mg $MnSO_4 \times 5\ H_2O$ 0.4 mg $ZnSO_4 \times 7\ H_2O$ 0.2 mg Na_2MoO_4 0.1 mg $CuSO_4 \times 5\ H_2O$ 0.1 mg $CoCl_2$ 0.1 mg KI in 1 L aq. dest Adjust to pH pH 5.4 autoclaved at 120 °C for 20 min

Name	Preparation
Vitamine solution[180]	200 mg myo-Inositol 100 mg thiamin hydrochloride 40 mg nicotinic acid 40 mg pantothenic acid, Ca-salt 40 mg pyridoxin hydrochlorid 20 mg riboflavin 20 mg *p*-Aminobenzoic acid 0,2 mg biotin 0,2 mg folic acid in 1 L aq. dest sterile filtered
LB medium: Luria Bertani or Lysogenic Broth[181]	Purchased from Sigma-Aldrich, 20 g in 800 mL aq. dest, autoclaved at 120 °C for 20 min
LB-agar	1.5 % (w/v) agar in LB medium, autoclaved for 20 min at 120 °C
10x SOC stock[182]	100 mg KCl 1 g $MgCl_2$ 1 g $MgSO_4$ 2 g glucose for 50 mL in aq. dest., sterile filtered
LB-SOC medium	LB medium with 10 % (v/v) SOC stock
TB medium (modified)	Purchased from Fluka, 38.08 g + 6.4 mL glycerol in 800 mL, autoclaved at 120 °C for 20 min
YPD medium	9 g yeast extract 9 g peptone in 800 mL aq. des., autoclaved for 20 min at 120 °C, afterwards add 100 mL sterile glucose solution (100 g/L)
YPD agar	Prepared like YPD medium with 1.5 % (w/v) agar
Minimal methanol medium	1.34% YNB $4 * 10^{-5}$% biotin 0.5% methanol
Chloramphenicol stock solution	50 mg/mL in 70 % (v/v) ethanol, sterile filtered
Kanamycin stock solution	50 mg/mL in aq. dest., sterile filtered
Ampicillin stock solution	100 mg/mL in aq. dest., sterile filtered
Zeocine stock solution	100 mg/mL in aq. dest., sterile filtered
IPTG stock solution	1 mol/L in aq. dest., sterile filtered
Sodium phosphate buffer (50 mM, pH 7.5)	1.311 g $NaH_2PO_4 \cdot H_2O$ 14.5 g $Na_2HPO_4 \cdot 12\ H_2O$ adjust to pH 7.5 with NaOH for 1 L solution in aq. dest.

Name	Preparation
Disruption buffer Sodium phosphate buffer (50 mM, pH 7.5) 100 mM NaCl 10 % (w/v) glycerol	1.311 g $NaH_2PO_4 \cdot H_2O$ 14.5 g $Na_2HPO_4R \cdot 12\ H_2O$ 5.844 g NaCl 100 g glycerol adjust to pH 7.5 with NaOH for 1 L solution in aq. dest.
Washing buffer Sodium phosphate buffer (50 mM, pH 7.5) 100 mM NaCl 10 % (w/v) glycerol 5 mM imidazole	1.311 g $NaH_2PO_4 \cdot H_2O$ 14.5 g $Na_2HPO_4 \cdot 12\ H_2O$ 5.844 g NaCl 100 g glycerol 0.34 g imidazole adjust to pH 7.5 with NaOH for 1 L solution in aq. dest.
Elution buffer Sodium phosphate buffer (50 mM, pH 7.5) 100 mM NaCl 10 % (w/v) glycerol 500 mM imidazole	1.311 g $NaH_2PO_4 \cdot H_2O$ 14.5 g $Na_2HPO_4 \cdot 12\ H_2O$ 5.844 g NaCl 100 g glycerol 34.04 g imidazole adjust to pH 7.5 with NaOH for 1 L solution in aq. dest.
Acrylamide solution for SDS-PAGE	30 % acrylamide (w/v), 0.8 % (w/v) *N*,*N*'-Methylenebisacrylamide in aq. dest.
APS for SDS-PAGE	10% (w/v) ammonium persulfate (APS) in aq. dest.
Upper-TRIS buffer for SDS-PAGE	6 g TRIS (1.25 M), 0.1 g SDS in 100 ml aq. dest., adjust pH to 6.8
Lower-TRIS buffer for SDS-PAGE	18.2 g TRIS (1.8 M), 0.1 g SDS in 100 ml aq. dest., adjust pH to 8.8
Loading buffer for SDS-PAGE (Laemmli)[183]	20 % (w/v) glycerol, 6 % (w/v) 2-mercaptoethanol, 0.0025 % bromophenol blue in Upper-TRIS buffer
10x running buffer for SDS-PAGE	30.3 g TRIS 144 g Glycine 10 g SDS in 1 L aq. dest.
Staining solution for SDS-PAGE	1 g Coomassie Brilliant Blue G250 100 mL acetic acid 300 mL ethanol 600 mL aq. dest.
Destaining solution for SDS-PAGE	100 mL acetic acid 300 mL ethanol 600 mL aq. dest.
50 x TAE buffer for AGE	242 g TRIS 57.1 mL acetic acid 18.6 g EDTA for 1 L in aq. dest. adjust to pH 8.5
1x TAE buffer for AGE	2 % (v/v) 50x TAE buffer in aq. dest.

Name	**Preparation**
Loading buffer for AGE	3 mL glycerol 0.025 g bromophenol blue 2 mL 50x TAE buffer 5 mL aq. dest.
Agarose solution for AGE	1 % (w/v) agarose in 1x TAE buffer
NADPH stock solution	300 mM in aq. dest.
NADH stock solution	300 mM in aq. dest.
TRIS-HCl buffer (50 mM, pH 9.0)	6.1 g TRIS base adjust to pH 9.0 with HCl for 1 L solution in aq. dest.
RF 1 buffer	100 mM RbCl 50 mM MnCl2 x 4 H_2O 30 mM Kaliumacetat 10 mM $CaCl_2$ x 6 H_2O 15% Glycerol adjust to pH 5.8 with acetic acid in aq. dest., sterile filtered
RF 2 buffer	10 mM RbCl 75 mM $CaCl_2$ x 6 H_2O 10 mM MOPS 15% Glycerol adjust to pH 7 with NaOH in aq. dest., sterile filtered

6.4 Kits / markers / enzymes

Table 6.3: Kits / markers.

Purpose	Name	Manufacturer
Chaperone coexpression	TaKaRa Chaperone Plasmid Set	Takara Bio USA, Inc. (Mountain View, CA, USA)
Plasmid preparation	innuPREP Plasmid Mini Kit	Analytik Jena (Jena)
Isolation of genomic DNA	innuSPEED Bacteria/Fungi DNA Kit	Analytik Jena (Jena)
PCR purification	NucleoSpin Gel and PCR Clean-up	Macherey-Nagel (Düren)
	PCR purification Kit	Roche (Basle, Switzerland)
Protein marker for SDS-PAGE	Pierce Unstained Protein MW Marker	Thermo Scientific (Waltham, MA, USA)
DNA marker for AGE	1 kbp ladder	Carl Roth (Karlsruhe)
Staining for AGE	Roti GelStain	Carl Roth (Karlsruhe)

Table 6.4: Used commercial enzymes.

Enzyme	Manufacturer
*Dpn*I restriction endonuclease	New England Bioloabs (Ipswich, MA, USA)
Opti*Taq* polymerase	EURX, Roboklon (Berlin)
Taq DNA Polymerase	EurX (Roboklon (Berlin)
Pfu+ DNA-Polymerase	EurX (Roboklon (Berlin)
restriction endonucleases	Thermo Fisher Scientific (Waltham, MA, USA)
T4-DNA-Ligase	Thermo Fisher Scientific (Waltham, MA, USA)
Proteinase K	New England Biolabs (Beverly, MA, USA) Analytik Jena (Jena)

6.5 Strains, plasmids and primers

Table 6.5: Used strains.

Strain	Genotype	Manufacturer/Source
Escherichia coli BL21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm]* Δ*hsdS*	New England Biolabs (Beverly, MA, USA)
Escherichia coli TOP10	F´(*lac*Iq, Tn*10*(TetR)) *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lacZ*ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Invitrogen (Carlsbad, CA, USA)
Pichia pastoris X-33	Wild type	Invitrogen (Carlsbad, CA, USA)
Yarrowia lipolytica Strain 63	Wild type	Prof. em. Dr. Frieder Schauer (Greifswald)
Candida maltosa SBUG 700	Wild type	Prof. em. Dr. Frieder Schauer (Greifswald)
Candida utilis SBUG 61	Wild type	Prof. em. Dr. Frieder Schauer (Greifswald)
Candida tropicalis SBUG 1019	Wild type	Prof. em. Dr. Frieder Schauer (Greifswald)
Candida catenulata SBUG 512	Wild type	Prof. em. Dr. Frieder Schauer (Greifswald)
Pichia guilliermondii SBUG 50	Wild type	Prof. em. Dr. Frieder Schauer (Greifswald)
Lodderomyces elongisporus SBUG 400	Wild type	Prof. em. Dr. Frieder Schauer (Greifswald)
Yarrowia lipolytica SBUG 1888	Wild type	Prof. em. Dr. Frieder Schauer (Greifswald)
Rhodosporidium toruloides SBUG 137	Wild type	Prof. em. Dr. Frieder Schauer (Greifswald)
Trichosporon asahii SBUG 833	Wild type	Prof. em. Dr. Frieder Schauer (Greifswald)
Saccharomyces cerevisiae SBUG 118	Wild type	Prof. em. Dr. Frieder Schauer (Greifswald)
Candida albicans SBUG 5121	Wild type	Prof. em. Dr. Frieder Schauer (Greifswald)

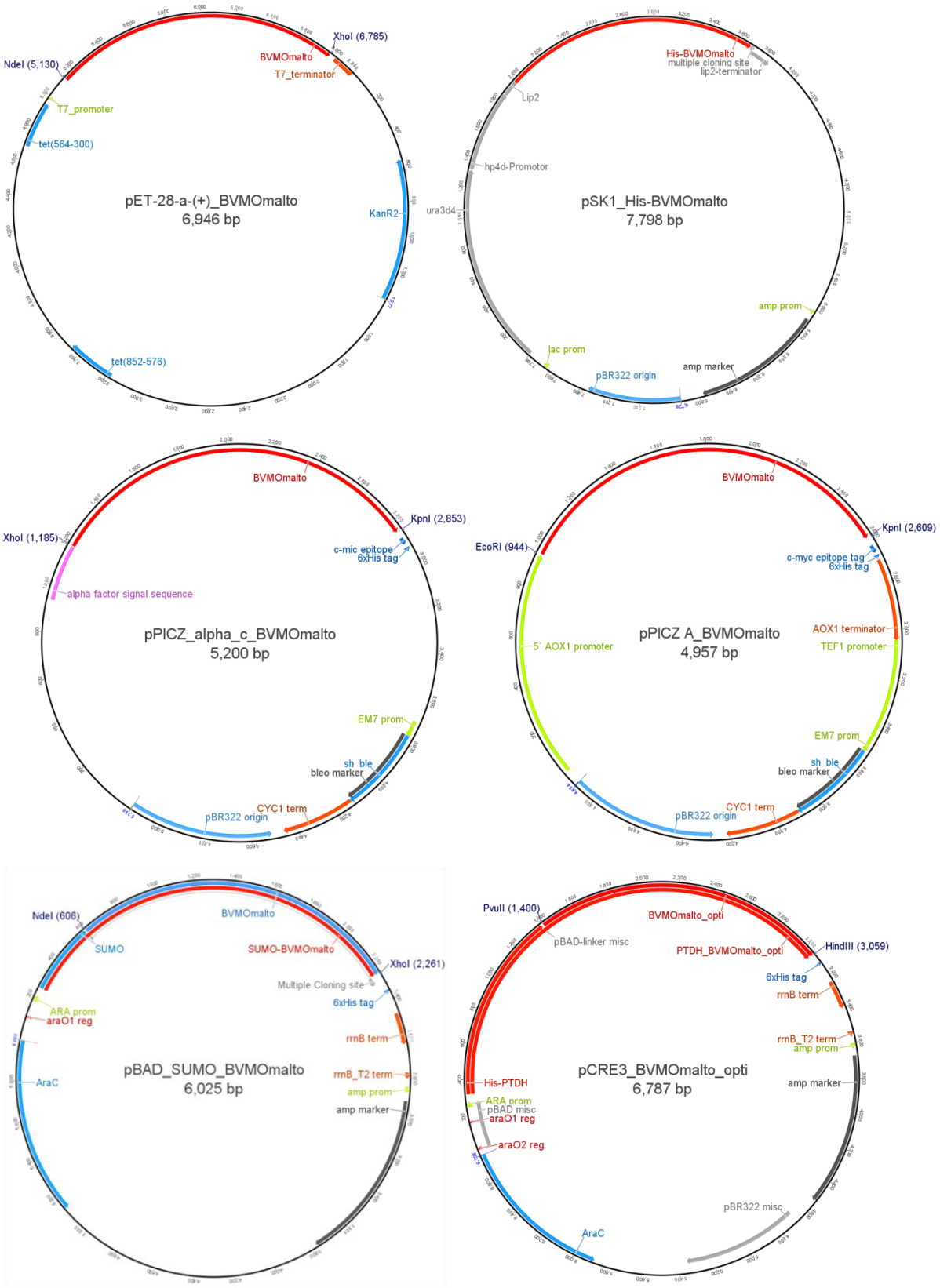


Figure 6.1: Exemplary vector maps for pET28, pSK1, pPICZa/α, pBAD_SUMO and pCRE3 constructs. The pSK1 vector was kindly provided by Prof. Kolmar (Darmstadt).

All primers were designed with the in the program Geneious integrated software Primer 3.[184]

Table 6.6: Used primers.

Name[a]	Sequence (5'->3')
Yeast BVMO project	
19.55_fw_new_NdeI	GCGAATTCCATATGGCGGCGTCTATGTCTGTCATTACATTAACAAAA-GAG
19.55_rv_new_BamHI	CGCGGATCCTAAGGTCATCTCCTCTTTTTGTCTTCG
malto_pSKI_fw	TCACTCCTTCTGAGGCCGCAGTTCTCCAGAAGCGAGTGATGCCAGT-TATCACCTTAACTAAAGAATCG
malto_pSKI_rv	TAGGTACCGGATCCGCCTAGGTTACATATGAC-TAGTCTAATCTTTTTTAATAGATTTTGTATCATACTCC
His-tag-malto_pSKI_fw	TCACTCCTTCTGAGGCCGCAGTTCTCCAGAAGCGAGTGATGGGCAG-CAGCCATCATCAT
malto_pPICZA_fw_EcoRI	CCGGAATTCTATGCCAGTTATCACCTTAACTAAAGAATCG
malto_pPICZA_rv_KpnI	CGCGGTACCGGAATCTTTTTTAATAGATTTTGTATCATACTCC
malto_pPICZ_c_α_fw_ XhoI	CCGCTCGAGAAGAGAATGCCAGTTATCACCTTAACTAAAGAATCG
malto_pPICZ_c_α_rv_ KpnI	CGCGGTACCGGATCTTTTTTAATAGATTTTGTATCATACTCC
malto_fw_NdeI_pBAD_ SUMO	GGGAATTCCATATGCCAGTTATCACCTTAACTAAAGAATCG
malto_fw_XhoI_pBAD_ SUMO	CCGCTCGAGCTAATCTTTTTTAATAGATTTTGTATCATACTCC
BVMOmalto_opti_fw_ PvuII_pCRE3	TCTGCAGCTGGTATGCCGGTCATTACCCTGACG
BVMOmalto_opti_rv_ HindIII_pCRE3	GCCCAAGCTTTTAGTCTTTTTTGATTGATTTCGTGTCG
malto_opt_fw_NdeI_ pBAD_SUMO	GGGAATTCCATATGCCGGTCATTACCCTGACG

malto_opt_rv_XhoI_ pBAD_SUMO	CCGCTCGAGTTAGTCTTTTTTGATTGATTTCGTGTCG
albi_fw_NdeI_pBAD_ SUMO	GGGAATTCCATATGTCTGTCATTACATTAACAAAAGAG
albi_rv_XhoI_pBAD_ SUMO	CCGCTCGAGTCATCTCCTCTTTTTGTC
G210_L111S_fw	CTTATTCATTC**TCT**CCAGTTTCAAACTGGAGTAG
G210_L111S_rv	CCAGTTTGAAACTGG**AGA**GAATGAATAAGAATACC
G210_L261S_fw	GAAACATTATGTC**TCT**CCACCTGTCCCTAAAGC
G210_L261S_rv	GGACAGGTGG**AGA**GACATAATGTTTCGATC
TMO_L462S_fw	GGTCAAGTT**TCT**GAAAAAATTAAAGTGGGTAG
TMO_L462S_rv	CCACTTTAATTTTTTC**AGA**AACTTGACCTAAATTGAC
YMOC_fw_NdeI	ACGAATTCCATATGCCCTCAATTGATCCTTCC
YMOC_rv_NotI	AAGGAAAATTGCGGCCGCTTACGCCCTGATGTCAACCC
YMOD_fw_NdeI	ACGAATTCCATATGCTGGCTGTCTACATATACTCC
YMOD_rv_NotI	AACCAATTACGCGGCCGCTCATCTTGCAACCACCCTGG
YMOE_fw_NdeI	ACGAATTCCATATGTCGACAGTATTTGCAGACG
YMOE_rv_NotI	AAGGAAAATTGCGGCCGCTTACTTTGTTTTCTTCTCCCCGTTAC
YMOF_fw_NdeI	AGCTATTACATATGCGGAGATTGAAATACAACTTTC
YMOF_rv_NotI	AAGGAAAATTGCGGCCGCTTAGAGACTGACCTCCTTGGC
YMOG_fw_NdeI	ACCTATTCCATATGTCGAGCAAAAACGGAACTG
YMOG_rv_NotI	AAGGAAAATTGCGGCCGCTCACTTTGAATGAGGGGCAC
YMOH_fw_NdeI	ACGAATTCCATATGACTGAGCTCTACTCTCACTC
YMOH_rv_NotI	AAGGAAAATTGCGGCCGCTTAATGTCGGATGATATCTTTCTGGT

Cofactor project	
D57S_fw	GCATTGACGT**CT**ACAGAAACCCACCTCTACTGC
D57S_rv	GGTTTCTGT**AG**ACGTCAATGCACCTGGGTAACGG
D341C_fw	CTTTAACCGT**TG**CAATGTCCGTTTAGAAGATGTG
D341C_rv	CGGACATTG**CA**ACGGTTAAAGGTGTTGTAGTAAC
D347V_fw	CGTTTAGAAG**TG**GTGAAAGCCAATCCGATTGTTG
D347V_rv	GGCTTTCAC**CA**CTTCTAAACGGACATTGTCACG
F284Q_fw	GTTTGAAACT**CAG**GGTGATATTGCCACCAATATG
F284Q_rv	GCAATATCACC**CTG**AGTTTCAAACATGAAACGG
F380Y_fw	GATATGTGCCACAGGTT**AC**GATGCCGTCGATGGC
F380Y_rv	GTTGCCATCGACGGCATC**GT**AACCTGTGGCAC
I182V_fw	CGTGTCGGCGTG**GTG**GGTACGGGTTCCACCGG
I182V_rv	GGAACCCGTACC**CAC**CACGCCGACACGTTTAC
K326F_fw	GGATTTGTATGCA**TTT**CGTCCGTTGTGTGACAG
K326F_rv	CACACAACGGACG**AAA**TGCATACAAATCCTGTG
K326H_fw	GATTTGTATGCA**C**ATCGTCCGTTGTGTGACAG
K326H_rv	CACACAACGGACGAT**G**TGCATACAAATCCTGTG
K326N_fw	GGATTTGTATGCAAA**C**CGTCCGTTGTGTGACAG
K326N_rv	CACACAACGGACG**G**TTTGCATACAAATCCTGTG
K326R_fw	GGATTTGTATGCA**CGC**CGTCCGTTGTGTGACAG
K326R_rv	CACAACGGACG**GCG**TGCATACAAATCCTGTGGC
K349E_fw	GTCCGTTTAGAAGATGTG**G**AAGCCAATCCGATTG
K349E_rv	CAACAATCGGATTGGCTT**C**CACATCTTCTAAAC
K349R_fw	GTTTAGAAGATGTG**CGC**GCCAATCCGATTGTTG

K349R_rv	CAATCGGATTGGC**GCG**CACATCTTCTAAACGGAC
L55R_fw	CCAGGTGCA**CG**TACGGATACAGAAACCCACCTC
L55R_rv	CTGTATCCGTA**CG**TGCACCTGGGTAACGGTTCC
Q210D_fw	CCAGCGTTCTGCA**G**ATTACAGCGTTCCAATTGG
Q210D_rv	GGAACGCTGTAAT**C**TGCAGAACGCTGGAAGAC
Q210N_fw	CAGCGTTCTGCAA**A**A**C**TACAGCGTTCCAATTGG
Q210N_rv	GGAACGCTGTA**G**T**T**TGCAGAACGCTGGAAGAC
Q210S_fw	CTTCCAGCGTTCTGC**AAG**CTACAGCGTTCCAATTG
Q210S_rv	GGAACGCTGTAG**CTT**GCAGAACGCTGGAAGACAG
R50L_fw	GTACTGGAACC**TG**TACCCAGGTGCATTGACGGATAC
R50L_rv	CACCTGGGTA**CA**GGTTCCAGTACCAAGTACCTGCG
S186P_fw	GGTACGGGT**C**CGACCGGTGTTCAGGTTATTACGG
S186P_rv	GAACACCGGTCG**G**ACCCGTACCAATCACGCCGAC
S208D_fw	CTGTCTTCCAGCGT**GA**TGCACAATACAGCGTTC
S208D_rv	GCTGTATTGTGCA**TC**ACGCTGGAAGACAGTGAG
S208E_fw	GTCTTCCAGCGT**GAG**GCACAATACAGCGTTCC
S208E_rv	GCTGTATTGTGC**CTC**ACGCTGGAAGACAGTGAG
S208E_Q210N_fw	CTTCCAGCGT**GAA**GCA**A**A**C**TACAGCGTTCCAATTG
S208E_Q210N_rv	GCCAATTGGAACGCTGTA**G**T**T**TGC**TTC**ACGCTG
S208T_fw	CTGTCTTCCAGCGT**A**C**C**GCACAATACAGCGTTCC
S208T_rv	CGCTGTATTGTGC**G**G**T**ACGCTGGAAGACAGTGAG
T139L_fw	GTTTCCTCATC**CTG**GCTTTAGGCTTATTGTCTGCG
T139L_rv	GCCTAAAGC**CAG**GATGAGGAAACGCGCCGTGTAC
T184A_fw	GGCGTGATTGGT**G**CGGGTTCCACCGGTGTTCAG

T184A_rv	GTGGAACCCG**C**ACCAATCACGCCGACACGTTTAC
T184N_fw	GCGTGATTGGTA**AC**GGTTCCACCGGTGTTCAGG
T184N_rv	GTGGAACC**GT**TACCAATCACGCCGACACGTTTAC
T187L_fw	GGTACGGGTTCC**CTG**GGTGTTCAGGTTATTACG
T187L_rv	CCTGAACACC**CAG**GGAACCCGTACCAATCACGC
T378V_fw	CTGATATGTGCC**GTG**GGTTTTGATGCCGTCGATG
T378V_rv	GGCATCAAAACC**CAC**GGCACATATCAGCATGTC
V189I_fw	CCACCGGT**A**TTCAGGTTATTACGGCTGTGGCAC
V189I_rv	GTAATAACCTGAA**T**ACCGGTGGAACCCGTACCAATC
V253Y_fw	GAAAGCACA**TAT**CCAGCAATGAGCGTATCAGC
V253Y_rv	CATTGCTGG**ATA**TGTGCTTTCATTCAGGCCAAAG
W490Y_fw	CTAAAGCGCAATCCT**AT**ATTTTTGGTGCGAATATCC
W490Y_rv	CACCAAAAAT**AT**AGGATTGCGCTTTAGGGAATAAG
Sequencing primers	
T7	TAATACGACTCACTATAGGG
pET-RP	CTAGTTATTGCTCAGCGG

[a]the "XnumberZ" primers are mutagenic primer for QuikChange site-directed mutagenesis (Chapter 6.7.9). The **bold** bases indicate the mutation sites, underlined bases indicate restriction sites of the used endonucleases. fw: forward; rv: reverse.

6.6 Microbiological methods

6.6.1 Strain maintenance

Glycerol stocks of *E. coli*, *Pichia pastoris* and *Y. lipolytica* strains were prepared by mixing 500 µL of overnight culture with 500 µL sterile 60 % (v/v) glycerol and subsequently kept at -80 °C for long-term storage.

For short-term storage *E. coli* strains were kept on LB-agar plates with addition of the corresponding antibiotic(s) at 4 °C. Yeast cells were kept on malt or YPD agar in tubes or plates, if needed with addition of antibiotics, at 4 °C. Prior to an experiment, these were transferred to new agar plates and incubated overnight at 30 °C.

6.6.2 Cultivation and expression *Pichia pastoris* X-33

Cells of *Pichia pastoris* X-33 from an YPD agar plate were dispersed in 20 mL of YPD medium in 250 ml erlenmeyer flasks with baffles and incubated overnight at 30 °C and 180 rpm. Then, the culture was centrifuged for 20 min at 4500 x g / 4 °C. The supernatant was removed and the cell pellet was frozen at -20 °C until extraction of the genomic DNA.

To investigate the expression of $BVMO_{malto}$ encoded on pPICZ_A and pPICZ_α in *P. pastoris* X-33, nine transformants were cultivated according to the instructions of the EasySelect™ *Pichia* Expression Kit manual (Invitrogen) in 20 ml minimal methanol medium in 100 ml erlenmeyer flasks without baffles at 30 °C. 1 mL samples have been taken after 0, 6, 21, 30, 46, 53, 69 and 104 h.

6.6.3 Cultivation of *Yarrowia lipolytica*

6.6.3.1 *Y. lipolytica* 63

Cells of *Y. lipolytica* Strain 63 from an YPD agar plate were dispersed in 50 mL of YPD medium and incubated overnight at 30 °C and 180 rpm. 15 mL of culture were centrifuged for 20 min at 4500 x g / 4 °C. The supernatant was removed and the cell pellet was frozen at -20 °C until extraction of the genomic DNA.

6.6.3.2 Cultivation and expression in *Y. lipolytica* Po1f

50 mL YPD were inoculated with two loops of respective *Y. lipolytica* cells from an YPD agar plate and incubated at 30 °C and 180 rpm for five days. Samples (1 mL for supernatant investigation, normalized to ($7/OD_{600}$) mL for cell protein investigation) were taken after 0, 6, 28, 53, 72 and 100 hours. In case of the expression with subsequent secretion of $BVMO_{malto}$,

30 mL of the remaining supernatants after 100 h were concentrated to 1 mL using a centricon® with a size exclusion of 30 kDa.

6.6.4 Cultivation and expression in *E. coli* BL21(DE3)

6.6.4.1 Overnight cultures

Cells of *E. coli* were dispersed in 5 ml of LB medium with addition of the needed antibiotic and incubated overnight at 30 °C and 180 rpm.

6.6.4.2 Enzyme production

All pET28a(+), pBAD_SUMO and pCRE3 constructs were expressed in *Escherichia coli* BL21(DE3). Precultures were prepared by inoculating 5 mL LB medium (with 50 µg/mL kanamycin for pet28 constructs or 100 µg/mL ampicillin for pBAD_SUMO and pCRE3 constructs) with 5 µL of a glycerol stock and subsequent incubation at 37 °C/180 rpm overnight. For expression, TB or LB medium supplemented with 50 µg/mL kanamycin or 100 µg/mL ampicillin in a sterile shaking flask was inoculated with a preculture (1/100). The flasks were incubated at 37 °C and 180 rpm until OD_{600} 1 – 1.5 (TB) or 0.4 – 0.9 (LB) was achieved. IPTG (0.1 – 1 mM for pET constructs) or arabinose (0.02% for pBAD_SUMO and pCRE3 constructs) was added for induction followed by incubation at 15 – 37 °C at 180 rpm. Samples normalized to $7/OD_{600}$ were collected during cultivation and centrifuged at 4500 x g, 4 °C for 10 minutes before freezing the pellets at -20 °C.

For purification, the cultivations were harvested by centrifugation at 4500 x g, 4 °C for 20 min after the best time determined for the respective enzyme, respectively. The pellet was then frozen at -20 °C until purification.

For the production of resting cells, the cultivation was harvested by centrifugation at 4000 x *g*, 4 °C for 20 min after the best time determined for the respective enzyme, respectively. The pellet was washed with 20 mL cold sodium phosphate buffer (50 mM, pH 7.5). After removal of the supernatant by another centrifugation step, the cells were stored overnight at 4 °C.

To obtain resting cells for whole cell biocatalysis with $BVMO_{malto}$, the construct pet28a(+)_$BVMO_{malto_opti}$ was expressed in *E. coli* BL21(DE3). Cultivation of the cells was performed in TB medium at 37 °C until an OD_{600} of 1 was reached. Then, the culture was incubated at 12 °C overnight. Subsequently, gene expression was induced with 0.1 mM IPTG and the temperature was increased to 20 °C. After 5 h cells were harvested and treated further like described above.

6.6.4.2.1 $BVMO_{albi1}$

$BVMO_{albi1}$ encoded on pET28a(+) was expressed in LB medium supplemented with kanamycin in 250 mL shaking flasks with baffles at 30 °C at 180 rpm. After reaching an OD_{600} of about 0.4, gene expression of $BVMO_{albi1}$ was induced by addition of IPTG to a final concentration of 0.4 mM.

$BVMO_{albi1}$ encoded on pBAD_SUMO was expressed in TB medium supplemented with ampicillin in 250 mL shaking flasks with baffles at 17 °C at 180 rpm for 48 h.

6.6.4.2.2 $BVMO_{malto}$ and CMO

For investigation of the best expression temperature for $BVMO_{malto}$ and CMO encoded on pET28a(+), they were expressed in LB medium supplemented with kanamycin in 250 mL shaking flasks with baffles at 15, 30 and 37 °C at 180 rpm. The induction of gene expression was achieved by addition of IPTG to a final concentration of 0.1 mM.

For the investigation of different IPTG concentrations BVMOmalto was expressed at 20 °C and 37 °C. Gene expression was induced by the addition of IPTG to a final concentration of 0.1, 0.4 or 1 mM.

The codon-optimized BVMO $BVMO_{malto\_opt}$ encoded on pET28a(+) was expressed in LB medium supplemented with kanamycin in 250 mL shaking flasks with baffles at 30 °C at 180 rpm. After reaching an OD_{600} of about 0.4, gene expression was induced by addition of IPTG to a final concentration of 0.4 mM.

For investigation of the influence of riboflavin addition to the culture medium, $BVMO_{malto\_opt}$ was expressed like described above but with addition of 1µg/mL riboflavin to the LB-medium.

$BVMO_{malto}$ and $BVMO_{malto\_opt}$ encoded on pBAD_SUMO was expressed in TB medium supplemented with ampicillin in 250 mL shaking flasks with baffles at 17 °C at 180 rpm for 48 h.

The vector pCRE3 vector is used to create a translational fusion with a thermostable phosphite dehydrogenase (PTDH, codon optimized) what worked well with many BVMOs and increased their solubility.[71, 141] Additionally, PTDH can be used for NADPH regeneration leading to a fusion protein with internal cofactor recycling making a biocatalysis more efficient. Expression of $BVMO_{malto\_opt}$ encoded on pCRE3 was carried out equally to the expression from the pBAD_SUMO vector.

6.6.4.2.3 YMOC-H

YMOC-H, encoded on pET28a(+), were expressed in 5 mL LB medium supplemented with kanamycin at 30 °C at 180 rpm in ONC tubes for 6 hours. Induction was achieved with 0.2 mM IPTG after an overnight incubation at 20 °C. This approach served for a fast overview of the expression of the YMOs from three different transformants, respectively.

YMOC-H, encoded on pET28a(+), were expressed in TB medium supplemented with kanamycin at 25 °C and 20 °C at 180 rpm. Induction of *ymoc-h* was achieved with 0.1 mM IPTG at OD_{600} 1.

6.6.4.3 Expression of wild-type $CHMO_{Acineto}$ and enzyme variants.

E. coli BL21(DE3) cells containing the plasmids encoding for the mutants of $CHMO_{Acineto}$ were precultured in 5 mL LB medium (with 50 µg/mL kanamycin) at 37 °C and 180 rpm overnight. For cultivation, TB medium supplemented with 50 µg/mL kanamycin was inoculated with a preculture (1/100). The cultures were incubated at 37 °C and 180 rpm until OD_{600} of 0.7 – 0.9 was achieved. IPTG (0.1 mM) was added for induction followed by incubation at 30 °C at 180 rpm. The cells were harvested after approximately 20 h by centrifugation at 4500 x *g*, 4 °C for 20 min.

6.6.4.4 Chaperone coexpression

The coexpression of chaperones was realized with the TaKaRa Chaperone Plasmid Set (Takara Bio USA, Inc.) containing five plasmids with different chaperone combinations (Table 6.7).[140]

Table 6.7: Plasmids contained in the TaKaRa chaperone set.

Plasmid	Chaperone	Molecular weight [kDa]	Inducer
pKJE7	DnaK DnaJ GrpE	70 40 22	L-Arabinose
pG-KJE8	DnaK DnaJ GrpE	70 40 22	L-Arabinose
	GroES GroEL	60 10	Tetracycline
pTf16	Tf	56	L-Arabinose
pGro7	GroES GroEL	60 10	L-Arabinose
pG-Tf2	GroES GroEL Tf	60 10 56	Tetracycline

Tf: trigger factor[140b]

To employ the chaperones, *E. coli* BL21(DE3) was cotransformed with the respective vector construct and one of the chaperone plasmids. Precultures were prepared by inoculating 5 mL LB medium (with 50 µg/mL kanamycin and chloramphenicol) with 5 µL of the glycerol stock containing cells with the respective cotransformed plasmid and subsequent incubation at 37 °C/180 rpm overnight.

For expression, LB or TB medium supplemented with 50 µg/mL kanamycin and chloramphenicol in a sterile shaking flask was inoculated with a preculture (1/100). The chaperones were induced at the beginning of the cultivation by adding 1 mg/mL L-arabinose and/or 5 ng/mL tetracycline dependent on the respective plasmid (Table 6.7). Then, cultures were treated equally to the ones without chaperone coexpression (see section 6.6.4.2).

Samples normalized to 7/OD_{600} were collected during cultivation and centrifuged at 4500 x *g*, 4 °C for 10 minutes before freezing the pellets at -20 °C.

6.7 Molecular biological methods

6.7.1 Determination of DNA concentration

The concentration of DNA solutions was determined photometrically utilizing a NanoDrop device. Hereby, the quotient abs_{260nm} /abs_{280nm} gives information about the contamination with proteins. The sample is relatively pure if this value is around 1.8. The quotient abs_{260nm}/abs_{230nm} gives indication about the contamination with buffer components (i.e. EDTA), carbohydrates and phenols. A value of 2.2 stands for a sample which is relatively free of these components.

6.7.2 Isolation of genomic DNA from yeasts

Genomic DNA from yeasts was extracted with the innuSPEED *Bacteria*/*Fungi* DNA Kit (Analytik Jena) according to the manufacturer's instructions with the exception that the elution buffer was heated to 60 °C prior to use to yield better elution.

6.7.3 Plasmid preparation

For plasmid isolations, 5 mL of overnight culture were centrifuged at 4500 x g, 4 °C for 15 minutes and the pellet used further. All isolations were performed using the innuPREP Plasmid Mini Kit (Analytik Jena) according to the manufacturer's instructions with the exception that the elution buffer was heated to 60 °C prior to use to yield better elution. The principle of the kit is to bind DNA on a silica membrane, the subsequent removal of unwanted

cell components, like proteins and single nucleotides, through washing with an ethanol solution and the final elution step of the bound DNA with water (or buffer).

6.7.4 Agarose gel electrophoresis

Agarose gel electrophoresis is a method for the separation of DNA samples according to their length. It was used to evaluate all PCR-based and to cloning-related experiments. To create an agarose gel, 1% agarose in 1x TAE was heated in the microwave to liquefy it, after which 30 mL were poured into a glass vessel. When the solution became lukewarm, 1.5 µL of Roti-Safe GelStain (Carl Roth) were added. After stirring, the solution was poured into the casting tray (with an inserted well comb) and air bubbles were removed. The gel was solid after approximately 45 minutes at which point the well comb was removed. The casting tray was then inserted into the gel box of the Mini-Sub Cell GT (Bio-Rad) and 1x TAE-buffer was added according to the markings on the inside of gel box.
The samples were prepared by mixing 5 µL of sample with 2 µL loading buffer and pipetted into the wells of the agarose gel. One well was loaded with 3 µL of the DNA marker (1 kbp DNA ladder from Carl Roth) for a later estimation of the size of the sample DNA. Gels were run at 110 V for approximately 20 minutes. Afterwards the gel was put on a UV tray (312 nm) to visualize the separated DNA bands. The detection of DNA is based on a large increase of fluorescence of the RotiSafe GelStain dye when it intercalates into DNA.

6.7.5 Sequencing

All sequencing was performed by Eurofins MWG GmbH (Ebersberg) or GATC Biotech AG (Konstanz).

6.7.6 Cloning

6.7.6.1 Classical cloning

For the classical cloning approach, the insert (gene of interest) is amplified using primers to introduce overhangs at both ends of the insert containing a specific recognition site for an endonuclease, respectively. Afterwards, the insert is purified with the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). The amplified insert and the vector are then separately digested utilizing the same endonucleases. Afterwards, they are purified with the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). The purified insert and vector are mixed together with a ligase to obtain the ligated vector construct. This is then transformed into com-

petent *E. coli* TOP10 cells to obtain the desired clones (analogous to Chapter 6.7.7.4). The success is confirmed by sequencing of the plasmids of by colony PCR selected transformants. Endonucleases or restriction enzymes are employed in molecular biology for the sequence specific restriction of DNA. The restriction can happen either in the middle of the four to eight long recognition sequence leading to blunt ends or two to four nucleotides away from it producing sticky ends. The latter are especially interesting for cloning experiments as by restriction of the vector with the same enzymes as used for the insert, its desired orientation in the final construct can be guaranteed. The connection of two DNA, strands like the insert with the linearized vector, is catalyzed by ligases. The T4-DNA-ligase is commonly used in molecular biology. This enzyme is obtained from *E. coli* cells which are infected with the bacteriophage T4. It catalyzes the formation of phosphodiester bonds between the free 5´-phosphate and 3´-hydroxy group of the desoxyribose under consumption of ATP.

6.7.6.1.1 $BVMO_{albi1}$ and YMOC-H into pET28

$BVMO_{albi}$ (*C. albicans*) and YMOC-H (*Y. lipolytica*) were cloned in the same way from the genomic DNA of the respective yeasts. Amplification was performed according to the following protocol (Table 6.8):

Table 6.8: PCR for insert amplification.

Insert PCR program

Step	Cycles	Temperature [°C]	Time [min]
Initial denaturation	1	95	5
Denaturation		95	1
Annealing	35	55	1
Extension		72	2
Final Extension	1	72	10
Cooling		15	hold

Insert PCR reaction mixture

Component	Volume [µL]
Template[a]	3
Primer fw1[b]	1
Primer rv2[b]	1
dNTPs	1
Pfu buffer C (10 x)	5
Pfu$^+$ polymerase	0.3
Sterile MilliQ water	Fill up to 50 µL

[a]: Cloning from gDNA: 1 µL gDNA (~ 500 ng).

[b]: Primers: Table 6.6

The endonuclease digest was performed according to the following protocol (Table 6.9):

Table 6.9: Endonuclease digest.

digest reaction mixture

Component	Volume [µL]
vector/insert[a]	15/10
NdeI	2
BamHI/NotI[b]	2
fast digest buffer (10 x)	5
Sterile MilliQ water	Fill up to 50 µL

[a]: amplified insert (15 µL), pET28a (10 µL)

[b]: BamHI for cloning of $BVMO_{albi1}$ and NotI for cloning of YMOC-H

The digest reaction mixture was incubated for two hours at 37 °C with a final step of ten minutes at 80 °C for enzyme inactivation. The ligation was performed acoording to the following protocol (Table 6.10):

Table 6.10: ligation of insert and vector

ligation program

Step	Temperature [°C]	Time [h]
1	20	2
2	16	4
3	14	3
4	12	3
5	10	2
Denaturation	72	0.167

ligation reaction mixture

Component	Volume [µL]
vector[a]	1-3
insert[b]	2
T4 ligase buffer (10 x)	0.7
T4 ligase	0.5
Sterile MilliQ water	Fill up to 7 µL

[a]: pET28a

[b]: $BVMO_{albi1}$/YMOC-H to reach a ratio of DNA concentration (vector/insert) of 1:2 and 1:10

Clones obtained after transformation of competent *E. coli* TOP10 cells with the complete ligation approach were checked by colony PCR. Plasmids of positive clones were sequenced what verified the correct insertion of $BVMO_{albi}$/YMOC-H into pET28a(+) .

6.7.6.1.2 $BVMO_{malto}$ into Pichia *expression vectors pPICZ_A and pPICZ_α*

$BVMO_{malto}$ was subcloned from a previously obtained pET28 construct into pPICZ_A and pPICZ_α vectors.[63] Amplification was performed according to the following protocol (Table 6.11):

Table 6.11: PCR for insert amplification.

Insert PCR program			
Step	**Cycles**	**Temperature [°C]**	**Time [min]**
Initial denaturation	1	95	5
Denaturation		95	1
Annealing	30	58/60	1
Extension		72	2
Final Extension	1	72	10
Cooling		15	hold

Insert PCR reaction mixture	
Component	**Volume [µL]**
Template[a]	1
Primer fw1[b]	1
Primer rv2[b]	1
dNTPs	1
Pfu buffer C (10 x)	5
Pfu$^+$ polymerase	0.2
Sterile MilliQ water	Fill up to 50 µL

[a]: Subcloning from pET28_$BVMO_{malto}$ construct: 1 µL pET28 construct (~ 50 ng)

[b]: Primers: Table 6.6

The endonuclease digest was performed according to the following protocol (Table 6.12):

Table 6.12: Endonuclease digest.

digest reaction mixture	
Component	**Volume [µL]**
vector/insert	15[a]/10[b]
EcoRI	2
KpnI	2
Fast digest buffer (10 x)	5
Sterile MilliQ water	Fill up to 50 µL

[a]: amplified insert (1863 ng)

[b]: pPICZ_A/pPICZ_α_c (465 ng)

The digest reaction mixture was incubated for two hours at 37 °C with a final step of ten minutes at 80 °C for enzyme inactivation. The ligation was performed according to the following protocol (Table 6.13):

Table 6.13: ligation of insert and vector

ligation program		
Step	**Temperature [°C]**	**Time [h]**
1	20	2
2	16	4
3	14	3
4	12	3
5	10	2
Denaturation	72	0.167

ligation reaction mixture	
Component	**Volume [µL]**
vector[a]	0.2
insert[b]	2
T4 ligase buffer (10 x)	0.7
T4 ligase	0.5
Sterile MilliQ water	Fill up to 7 µL

[a]: pPICZ_A (4.22 ng)

[b]: $BVMO_{malto}$ (142.6 ng)

Clones obtained after transformation of competent *E. coli* TOP10 cells with the ligation approach grown on Zeocin containing low salt LB agar were checked by colony PCR. Plasmids of positive clones were sequenced what verified the correct insertion of $BVMO_{albi}$/YMOC-H into pET28a(+) and of $BVMO_{malto}$ into pPICZ_A/α.

6.7.6.1.3 $BVMO_{malto}$, $BVMO_{malto\_opti}$ and $BVMO_{albi1}$ into pBAD_SUMO

$BVMO_{malto}$, $BVMO_{malto\_opti}$ and $BVMO_{albi1}$ were cloned in the same way from previously obtained pET28 constructs. Amplification was performed according to the following protocol (Table 6.14):

Table 6.14: PCR for insert amplification.

Insert PCR program			
Step	**Cycles**	**Temperature [°C]**	**Time [min]**
Initial denaturation	1	95	5
Denaturation	35	95	1
Annealing		55	1
Extension		72	2
Final Extension	1	72	10
Cooling		15	hold

Insert PCR reaction mixture	
Component	**Volume [µL]**
Template[a]	4
Primer fw1[b]	1
Primer rv2[b]	1
Phusion Pol Hot Start Mix	25
Sterile MilliQ water	Fill up to 50 µL

[a]: Cloning from pET28 constructs

[b]: Primers: Table 6.6

The endonuclease digest was performed according to the following protocol (Table 6.15):

Table 6.15: Endonuclease digest.

digest reaction mixture	
Component	**Volume [µL]**
vector/insert[a]	25/10
NdeI	2
XhoI	2
Cut smart buffer (10 x)	5
Sterile MilliQ water	Fill up to 50 µL

[a]: insert (25 µL): 2500 ng, pBAD_SUMO (4 µL)

The digest reaction mixture was incubated for two hours at 37 °C with a final step of twenty minutes at 80 °C for enzyme inactivation. The ligation was performed according to the following protocol (Table 6.16):

Table 6.16: ligation of insert and vector

ligation program		
Step	**Temperature [°C]**	**Time [h]**
1	20	2
2	16	4
3	14	3
4	12	3
5	10	2
Denaturation	72	0.167

ligation reaction mixture	
Component	**Volume [µL]**
vector[a]	0.3
insert[b]	2
T4 ligase buffer	0.7
T4 ligase	0.5
Sterile MilliQ water	Fill up to 7 µL

[a]: pBAD_SUMO_NdeI_XhoI : 0.57 ng

[b]: $BVMO_{malto}$: 160.2 ng, $BVMO_{malto\_opti}$: 141.8 ng, $BVMO_{albi1}$: 180.2 ng

Clones obtained after transformation of competent *E. coli* TOP10 cells with the complete ligation approach were checked by colony PCR. Plasmids of positive clones were sequenced what verified the correct insertion of $BVMO_{malto}$/$BVMO_{malto\_opti}$/$BVMO_{albi1}$ into pBAD_SUMO.

6.7.6.1.4 $BVMO_{malto\_opti}$ into pCRE3

$BVMO_{malto\_opti}$ was cloned from the previously obtained pET28 construct. Amplification was performed according to the following protocol (Table 6.17):

Table 6.17: PCR for insert amplification.

Insert PCR program			
Step	**Cycles**	**Temperature [°C]**	**Time [min]**
Initial denaturation	1	95	5
Denaturation	35	95	1
Annealing		55	1
Extension		72	2
Final Extension	1	72	10
Cooling		15	hold

Insert PCR reaction mixture	
Component	**Volume [µL]**
Template[a]	5
Primer fw1[b]	1
Primer rv2[b]	1
Phusion Pol Hot Start Mix (2x)	25
Sterile MilliQ water	Fill up to 50 µL

[a]: Cloning from pET28a_$BVMO_{malto_opti}$

[b]: Primers: Table 6.6

With the vector pCRE3 an endonuclease digest was performed with SmaI first to cut the PA-MO gene in the construct according to the following protocol (Table 6.18):

Table 6.18: Endonuclease digest.

digest reaction mixture	
Component	**Volume [µL]**
vector[a]	5
SmaI	2
Cut smart buffer (10 x)	5
Sterile MilliQ water	Fill up to 50 µL

[a]: pCRE3: 1000 ng

The endonuclease digest was performed according to the following protocol (Table 6.19):

Table 6.19: Endonuclease digest.

digest reaction mixture

Component	Volume [µL]
vector/insert[a]	25/30
PvuII	2
HindIII	2
NEB buffer 2 (10 x)	5
Sterile MilliQ water	Fill up to 50 µL

[a]: amplified insert (25 µL): 2500 ng, pCRE3_SmaI: 519 ng

The digest reaction mixture was incubated for two hours at 37 °C with a final step of twenty minutes at 80 °C for enzyme inactivation. The ligation was performed acoording to the following protocol (Table 6.20):

Table 6.20: ligation of insert and vector

ligation program

Step	Temperature [°C]	Time [h]
1	20	2
2	16	4
3	14	3
4	12	3
5	10	2
Denaturation	72	0.167

ligation reaction mixture

Component	Volume [µL]
vector[a]	0.5
insert[b]	2.5
Quick ligase buffer (10 x)	3.5
Quick ligase	0.5
Sterile MilliQ water	Fill up to 7 µL

[a]: pCRE3_SmaI_PvuII_HindIII: 5.1 ng

[b]: $BVMO_{malto\_opti}$: 158.25 ng

Clones obtained after transformation of competent *E. coli* TOP10 cells with the complete ligation approach were checked by colony PCR. Plasmids of positive clones were sequenced what verified the correct insertion of $BVMO_{malto\_opti}$ into pCRE3 .

6.7.6.2 SLiCE cloning

Cloning into the *Yarrowia lipolytica* expression vectors pSKI and pUC_INTB had to be performed employing the Seamless Ligation Cloning Extract (SLiCE) cloning method due to a missing multiple cloning site in these vectors.[130] $BVMO_{malto}$ was subcloned from a previously obtained pET28 construct into pSKI and pUC_INTB vectors.[63] In a first step, the insert (gene of interest) is amplified using primers to introduce overhangs at both ends of the insert corresponding to the position in the vector while the vector is digested with endonucleases to show these overlapping sequences at both ends. Afterwards, insert and vector are purified with the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). The purified insert and vector are mixed in different ratios (1:2, 1:10) and are incubated with the so called SLiCE extract. Then, the approaches are transformed into competent *E. coli* TOP10 cells to obtain the desired clones (analogous to Chapter 6.7.7.4). The success is confirmed by sequencing of the plasmids of by colony PCR selected transformants.

6.7.6.2.1 Cloning into pSKI

$BVMO_{malto}$ was cloned into the vector pSKI which is an integrative shuttle vector for expression in *Y. lipolytica* containing the N-terminal Lip2 prepro secretion signal.[136] The encoded 34 amino acids long signal peptide directs the protein to the secretion pathway and is eventually cleaved by the Xpr6p endoprotease, which is a Kex2-like endoprotease. Cloning of $BVMO_{malto}$_L111S_L261S with and without N-terminal $His_{(6)}$-tag was performed simultaneously.

Table 6.21: PCR for insert amplification.

Insert PCR program

Step	Cycles	Temperature [°C]	Time [min]
Initial denaturation	1	95	5
Denaturation		95	1
Annealing	35	55	1
Extension		72	2
Final Extension	1	72	10
Cooling		15	hold

Insert PCR reaction mixture

Component	Volume [µL]
Template[a]	5
Primer fw1[b]	1
Primer rv2[b]	1
dNTPs	1
Taq buffer B (10 x)	5
Opti*Taq* polymerase	0.3
Sterile MilliQ water	Fill up to 50 µL

[a]: Subcloning from pET28_$BVMO_{malto}$ construct: 1 µL pET28_$BVMO_{malto}$_S111L_S261L construct (~ 50 ng)

[b]: Primers: Table 6.6

Table 6.22: First endonuclease digest.

digest reaction mixture	
Component	**Volume [µL]**
vector[a]	14
SpeI	1
Cut smart buffer (10 x)	2
Sterile MilliQ water	Fill up to 20 µL

[a]: pSKI vecor: 1 µL pSKI (~ 300 ng)

The digest reaction mixture was incubated for 30 minutes at 37 °C with a final step of 20 minutes at 80 °C for enzyme inactivation.

The digestions and PCR were purified and measured by NanoDrop:

- pSKI_SpeI_THO: 77,5 ng/µL
- SLiCE insert $BVMO_{malto}$: 198 ng/µL
- SLiCE insert $His_{(6)}$-$BVMO_{malto}$:178 ng/µL

Table 6.23: Second endonuclease digest.

digest reaction mixture	
Component	**Volume [µL]**
vector[a]	17
BglII	1
buffer 3.1 (10 x)	2

[a]: pSKI_SpeI digest product: 1 µL pSKI_SpeI (~ 77.5 ng)

The digest reaction mixture was incubated for 30 minutes at 37 °C with a final step of 20 minutes at 80 °C for enzyme inactivation.

Table 6.24: SLiCE reaction for the construct pSKI_$BVMO_{malto}$.

SLiCE reaction mixture	
Component	**Volume [µL]**
vector[a]	5.65
Insert[b]	0.27/1.35
T4 ligase buffer (10 x)	1
SLiCE extract	2
Sterile MilliQ water	Fill up to 10 µL

[a]: pSKI_SpeI_BglII digest product: 100 ng

[b]: $BVMO_{malto}$: 53.2 ng for molar ratio 1:2, 267 ng for molar ratio 1:10

Table 6.25: SLiCE reaction for the construct pSKI_His$_{(6)}$-BVMO$_{malto}$.

SLiCE reaction mixture	
Component	**Volume [µL]**
vector[a]	5.65
Insert[b]	0.31/1.55
T4 ligase buffer (10 x)	1
SLiCE extract	2
Sterile MilliQ water	Fill up to 10 µL

[a]: pSKI_SpeI_BglII digest product: 100 ng

[b]: BVMO$_{malto}$: 55.32 ng for molar ratio 1:2, 276 ng for molar ratio 1:10

The SLiCE reaction mixture was incubated for 1 hour at 37 °C with a final step of 20 minutes at 80 °C for enzyme inactivation.

Clones obtained after transformation of competent *E. coli* TOP10 cells via electroporation with the complete SLiCE approaches were checked by colony PCR. Plasmids of positive clones were sequenced what verified the correct insertion of (His$_{(6)}$-)BVMO$_{malto}$ into pET28a(+) .

6.7.6.2.2 Cloning into pUC_INTB

For the expression of BVMO$_{malto}$ without secretion in *Yarrowia lipolytica*, it was cloned into the vector pUC_INTB that does not contain a secretion signal and thus the produced protein stays inside of the cell.

Table 6.26: PCR for insert amplification.

Insert PCR program			
Step	**Cycles**	**Temperature [°C]**	**Time [min]**
Initial denaturation	1	95	5
Denaturation		95	1
Annealing	35	55	1
Extension		72	2
Final Extension	1	72	10
Cooling		15	hold

Insert PCR reaction mixture	
Component	**Volume [µL]**
Template[a]	5
Primer fw1[b]	1
Primer rv2[b]	1
dNTPs	1
Taq buffer B (10 x)	5
Opti*Taq* polymerase	0.3
Sterile MilliQ water	Fill up to 50 µL

[a]: Subcloning from pET28_BVMO$_{malto}$ construct: 1 µL pET28_BVMO$_{malto}$_S111L_S261L construct (~ 50 ng)

[b]: Primers: Table 6.6

Table 6.27: First endonuclease digest.

digest reaction mixture	
Component	**Volume [µL]**
vector[a]	12
PmeI	1
Cut smart buffer (10 x)	2
Sterile MilliQ water	Fill up to 20 µL

[a]: pUC_INTB vecor: 1 µL pUC_INTB (~ 300 ng)

The digest reaction mixture was incubated for 2.5 hours at 37 °C with a final step of 20 minutes at 80 °C for enzyme inactivation.

The second digest was performed with SfiI by adding 1 µl of SfiI to the first digest approach. The digest was performed for 2.5 hours at 50 °C with a final step of 20 minutes at 80 °C. The digest reaction mixture was incubated for 2.5 hours at 50 °C with a final step of 20 minutes at 80 °C for enzyme inactivation.

As other projects were handled in parallel that seemed more promising as the expression in *Y. lipolytica*, this one was canceled at this point in time.

6.7.7 Transformation

6.7.7.1 Transformation of *Pichia pastoris* X-33

Prior to transformation of *Pichia pastoris* X-33 with the construct, linearization of the plasmid DNA had to be performed according to the following protocol (Table 6.28):

Table 6.28: Endonuclease digest.

digest reaction mixture	
Component	**Volume [µL]**
vector	10[a]
SacI	2
Fast digest buffer (10 x)	5
Sterile MilliQ water	Fill up to 50 µL

[a]: pPICZ_A (465 ng)

The digest was performed for three hours at 37 °C with a final step of ten minutes at 80 °C for enzyme inactivation.

Afterwards the approaches were purified using the PCR purification Kit (Roche). With these, electroporation of *P. pastoris* X-33 was performed according to the EasySelect™ *Pichia* Ex-

pression Kit manual (Invitrogen). After three days of incubation at 30 °C some colonies of transformants became visible. Because some have been very small they have been spread on new zeocin containing YPDS plates so that more cell material could be obtained.

6.7.7.2 Transformation of *Yarrowia lipolytica*

An YPD plate with *Y. lipolytica* Po1f cells is incubated overnight. From this, one loopful of cells is taken and resuspended into 1 mL TE in a sterile 2 mL reaction tube. This is then centrifuged for 1 min at 10000 x g and supernatant is discarded. The cells are resuspended into 600 µL lithium acetate (0.1 M, pH 6). The solution is incubated for 1 hour at 28 °C in a water bath without shaking and subsequently centrifuged for 2 min at 3000 x g. The supernatant is discarded and the cells are resuspended softly into 80-120 µL 0.1 M lithium acetate, pH 6.0 (for 2-3 transformations). The transformation is then carried out according to the following protocol:

- Take 40 µL of competent cells + 5 µL : 2 (or 3) µL carrier DNA
  + 3 (or + 2) µL transforming DNA
- Incubate 15 min at 28 °C in a water bath.
- Add 350 µL of PEG 4000 – lithium acetate (0.1 M, pH 6)
- Add 16 µL of 1 M DTT (40 mM final)
- Incubate 1 hour at 28 °C in a water bath, without shaking.
- Add 40 µL of DMSO (nearly 10% final
- Heat shock 10 min at 39 °C.
- Add 600 µL lithium acetate (0.1 M, pH 6)
- Streak on 5 plates of selective medium (YNB-N_{5000} containing 500 mg/L leucine) per transformation (200 µL per plate) and incubate at 28 °C. Transformants that integrated the pSKI or pUC_INTB vector into their genome, were able to compensate for their uracil auxotrophy.

Prior to transformation of *Yarrowia lipolytica* with the construct, linearization of the plasmid DNA had to be performed according to the following protocol (Table 6.29):

Table 6.29: Endonuclease digest.

digest reaction mixture

Component	**Volume [µL]**
vector	12[a]
SacII	2
Cut smart buffer (10 x)	2
Sterile MilliQ water	Fill up to 20 µL

[a]: pSKI (~ 3000 ng)

The digest was performed for 30 minutes at 37 °C with a final step of twenty minutes at 80 °C for enzyme inactivation.

6.7.7.3 Production of chemically competent *E. coli* cells (RbCl method)

The production of chemically competent *E. coli* cells was achieved through the rubidium chloride method.[185] 200 mL LB medium were inoculated with an overnight culture of the respective strain and incubated at 37 °C until an OD600 of 0.4. Then, the culture was incubated on ice for 15 min and centrifuged for 20 min at 4000 x g at 4 °C. All following solutions were cooled on ice. The cell pellet was carefully resuspended in 20 mL RF 1 buffer, again incubated on ice for 15 min and centrifuged for 20 min at 4000 x g at 4 °C. Finally, the cell pellet was resuspended in 8 mL RF 2 buffer, incubated on ice for 15 min and aliquoted in sterile, pre-cooled 1.5 reaction containers. The 50 µL aliquots were then flash-freezed in liquid nitrogen and kept at -80 °C.

6.7.7.4 Heat-shock transformation of *E. coli* with vector constructs

Chemocompetent *E. coli* TOP10 and BL21(DE3) were provided as 50 µL aliquots frozen at -80 °C. They were thawed on ice for 15 min. 1 – 2 µL of plasmid solution or 7 µL ligation approaches were added followed by incubation on ice for 15 min. Afterwards a heat shock was applied by immersing the samples for 30 s in a 42 °C water bath allowing the plasmids to permeate the cell membrane. After subsequent cooling on ice for 2 minutes, 200 µL of LB-SOC were added followed by incubation at 37 °C / 180 rpm for 1 h, allowing the cells containing the plasmid to express their resistance. 30 – 200 µL were plated on LB-agar plates containing 50 µg/mL kanamycin (pET constructs) or 100 µg/mL ampicillin (pSK1 constructs) for selection purposes. The plates were incubated overnight at 37 °C and afterwards stored at -4 °C.

6.7.7.5 Electroporation of *E. coli* with vector constructs

Electrocompetent *E. coli* TOP10 and BL21(DE3) were provided as 50 µL aliquots frozen at -80 °C. They were thawed on ice for 15 min. 1 – 2 µL of plasmid solution or 7 µL ligation approaches were added followed by incubation on ice for 15 min. The mixture was added into pre-cooled 2-mm electroporation cuvettes (BTX Cuvettes Plus, Harvard Apparatus, US) and 2.5 kV were applied for 5 ms (MicroPulser, Biorad, GER). Subsequently, 1 mL of LB medium was added directly after the pulse and the cells were recovered at 37 °C for 60 min, allowing the cells containing the plasmid to express their resistance. 30 – 200 µL were plated on LB-agar plates containing 50 µg/mL kanamycin (pET constructs) or 100 µg/mL ampicillin (pSK1 constructs) for selection purposes. The plates were incubated overnight at 37 °C and afterwards stored at -4 °C.

6.7.7.6 Cotransformation with chaperone plasmids

Chemocompetent *E. coli* BL21(DE3) were provided as 50 µL aliquots frozen at -80 °C. They were thawed on ice for 15 min. For each chaperone plasmid (Table 6.7), 5 µL of chaperone plasmid and 5 µL of the respective pET28a construct were added followed by incubation on ice for 15 min. Afterwards a heat shock was applied by immersing the samples for 30 s in a 42 °C water bath allowing the plasmids to permeate the cell membrane. After subsequent cooling on ice for 2 minutes, 200 µL of LB-SOC were added followed by incubation at 37 °C/180 rpm for 1 h, allowing the cells containing the plasmid to express their resistance. 250 µL were plated on LB-agar plates containing 50 µg/mL kanamycin and 50 µg/mL chloramphenicol for selection purposes. The plates were incubated overnight at 37 °C and afterwards stored at -4 °C. Precultures were prepared by inoculating 5 mL LB medium (with 50 µg/mL kanamycin and 50 µg/mL chloramphenicol) with cells from the LB agar plates and subsequent incubation at 37 °C/180 rpm overnight. Glycerol stocks were prepared as described in Chapter 6.6.1).

6.7.8 Colony PCR

6.7.8.1 Colony PCR with *E. coli* cells

To check a large number of transformants for whether they contain the correct construct or not, a colony PCR can be performed. In this PCR based approach, a part of the construct (mostly the insert) is amplified. Firstly, a master mix was prepared (Table 6.30). The master mix was divided into 6 µl aliquots. 12 colonies were picked and suspended in 10 µl *A. Bidest*, respectively. For cell lysis these suspensions were incubated at 95 °C for five minutes after streaking a small amount of the suspensions on an agar plate with the appropriate antibiotic. 1 µl cell lysate was added to the PCR approaches. Reaction was performed with the following program.

Table 6.30: Colony PCR.

Insert PCR program			
Step	**Cycles**	**Temperature [°C]**	**Time [min]**
Initial denaturation	1	95	5
Denaturation		95	1
Annealing	40	60[a]	1
Extension		72	2
Final Extension	1	72	10
Cooling		15	hold

colony PCR master mix (for 12 colonies)	
Component	**Volume [µL]**
Primer fw1[b]	1.32
Primer rv2[b]	1.32
dNTPs	1.32
DMSO	0.48
Taq buffer B (10 x)	8.4
Taq polymerase	0.84
Sterile MilliQ water	Fill up to 78.5 µL

[a]: Annealing temperature varies depending on the optimal temperature for the used primers, 60 °C for amplification of $BVMO_{malto}$ from pPICZ_A, 55 °C for amplification of $BVMO_{malto}$ from pSKI

[b]: Primers: Table 6.6

6.7.8.2 *Pichia* colony PCR

In order to optimize the investigation of *P. pastoris* transformants, an approach similar to colony PCR was carried out.

First of all, 5 ml cultures of *Pichia* transformants were prepared in YPD. After incubation overnight at 30 °C, the cultures were centrifuged at 4000 x g for 10 min and the pellets were resuspended in 1 ml sterile *A. dest*. These were frozen at -20 °C and afterwards incubated at 95 °C for ten minutes. This freeze-and-thaw cycle was repeated two times to obtain solutions containing DNA of the transformants. Those were used for PCR.

Table 6.31: *Pichia* colony PCR.

PCR program			
Step	**Cycles**	**Temperature [°C]**	**Time [min]**
Initial denaturation	1	95	5
Denaturation		95	1
Annealing	35	55	1
Extension		72	2
Final Extension	1	72	10
Cooling		15	hold

PCR reaction mixture	
Component	**Volume [µL]**
Template[a]	5
Primer fw1[b]	1
Primer rv2[b]	1
dNTPs	1
Taq buffer B (10 x)	5
Opti*Taq* polymerase	0.3
Sterile MilliQ water	Fill up to 20 µL

[a]: DNA solution from one *P. pastoris* transformant

[b]: Primers: Table 6.6

6.7.8.3 *Yarrowia* colony PCR

First of all, each colony is resuspended in 3 µL 20 mM NaOH. On top, 2 droplets paraffin can be added with a subsequent incubation at 100 °C for 10 min. Then, a master mix needs to be prepared for n+1 colonies, from which 50 µL are added to the cell suspensions. With the final PCR solutions, the amplification can be performed (Table 6.32).

Table 6.32: *Yarrowia* colony PCR.

Insert PCR program

Step	Cycles	Temperature [°C]	Time [min]
Initial denaturation	1	95	5
Denaturation		95	0.5
Annealing	35	55	0.5
Extension		72	2
Final Extension	1	72	4
Cooling		15	hold

colony PCR mixture (for 1 colony)[a]

Component	Volume [µL]
Primer fw1[b]	1
Primer rv2[b]	1
dNTPs	1
Taq buffer C (10 x)	5
Taq polymerase	0.5
Sterile MilliQ water	Fill up to 50 µL

[a]: it is recommended to prepare the mix for n+1 colonies

[b]: Primers: Table 6.6

6.7.9 Site-directed mutagenesis

Employing site-directed mutagenesis, directed point mutations, insertions and deletions can be introduced. For this purpose, two complementary primers are designed which contain the desired nucleotide exchange in their middle. To avoid primer insertions, they can be elongated on one end and shortened on the other, to obtain a not completely complementary primer pair. These are then used in a PCR with a long elongation step (1 min per kbp of the vector), so that the polymerase can replicate the whole plasmid. Consequently, the replicated plasmids contain the mutation of the primers. Subsequently, the restriction enzyme *Dpn*I is utilized, which only cuts methylated DNA. *In vitro* no methylation of the by PCR synthesized DNA is occurring and therefore only the parental vector without the desired mutation is digested.

Site-directed mutagenesis was performed using a QuikChange™ protocol in which the whole plasmid is amplified with primers containing the desired mutation. A 50 µL PCR master mix was divided into 10 µL aliquots and a PCR with temperature gradient was performed and analyzed by agarose gel electrophoresis to identify a suitable annealing temperature (Table 6.33).

Table 6.33: QuikChange™ PCR for site-directed mutagenesis.

QuikChange™ PCR program

Step	Cycles	Temperature [°C]	Time [min]
Initial denaturation	1	95	3
Denaturation		95	0.75
Annealing	30	55 - 70 °C[a]	1
Extension		72	8
Final Extension	1	72	15
Cooling		15	hold

QuikChange™ PCR master mix

Component	Volume [µL]
vector	0.3
Primer fw[b]	1
Primer rv[b]	1
dNTPs	1
Opti*Taq* buffer C (10 x)	5
Opti*Taq* polymerase	0.3
Sterile MilliQ water	Fill up to 50 µL

[a]: Temperature gradient.

[b]: Primers: Table 6.6

To reactions that contained amplified plasmid underwent *Dpn*I digestion to remove the template DNA. This was achieved by adding 0.2 µL *Dpn*I and incubating the approaches for 2 h at 37 °C with a subsequent step at 95 °C for 15 min for enzyme inactivation. Afterwards, 4 µL of reaction solution were transformed into *E. coli* TOP10 analogously to Chapter 6.7.7.4 and the success was subsequently confirmed by sequencing.

6.7.10 *Dpn*I digestion

1 µL of *Dpn*I was added to each reaction vessel followed by incubation at 37 °C for 2 hours. *Dpn*I was denatured by heat shock at 80 °C for 10 min.

6.8 Biochemical methods

6.8.1 Cell disruption

6.8.1.1 Ball mill

A fast cell disruption method for small amounts of cells is the cell homogenizator Fastprep24® (MP Biomedicals). For this, the cell pellets were resuspended in 500 µL ice-cold disruption buffer. Beads with a size of 0.1 mm (Lysing Matrix B, MP Biosystems) were added and the disruption of the samples was achieved through a shaking amplitude of 4 m/s for 2x 30 seconds. Between the two disruption steps samples were kept on ice for 5 min. After lysis, the samples were centrifuged for 10 min at 10500 x g at 4 °C to separate the supernatant containing the soluble proteins and the cell debris together with insoluble proteins. The pellet-bead mixture was washed with 1 ml disruption buffer and subsequently resuspended in 500 µl

disruption buffer. A sample consisting of 15 µL was taken from both fractions to be analyzed by SDS-PAGE.

6.8.1.2 Sonication

The pellets from the cultivation samples were resuspended in 500 µL disruption buffer. Samples were disrupted by sonication using the Sonoplus HD2070 (Bandelin) for 30 s (50 % power and cycle) and centrifuged at 17000 * g, 4 °C for 15 min. The supernatant was collected as the soluble fraction for SDS-PAGE. The pellet was washed once with 500 µL disruption buffer and centrifuged again. After resuspension in 500 µL of the same buffer, 20 µL were collected as the insoluble fraction for SDS-PAGE.

For the production of cell extract, the pellet was resuspended in 3 mL disruption buffer with 100 µM FAD. Sonication was performed using the Sonoplus HD2070 (Bandelin) for 3 min (50% power and cycle) with subsequent centrifugation (10000 x g for 20 min at 4 °C). The supernatant was filtered using a 0.45 µm filter and the resulting cell free extract was used for substrate screening.

To disrupt the have a rigid cell wall of *P. pastoris*, supersonication was performed five times for 30 seconds. This method was described before to work with yeast cells.[186]

6.8.1.3 Vortexing with glass beads

Vortexing with glass beads is the recommended cell disruption method for *P. pastoris* according to the EasySelect™ *Pichia* Expression Kit manual (Invitrogen). Cell pellets obtained from a cultivation (see section 6.6.2) were resuspended in 500 µL NaPP buffer (pH7.5). Glass beads were added to the cell suspensions and these were vortexed for 30 seconds and afterwards placed on ice for 30 seconds, for a total of eight cycles.

6.8.2 Enzyme purification

For this work, immobilized-metal affinity chromatography (IMAC) was used for a one-step purification. The cell extract is applied to a column containing Ni^{2+} or Co^{2+} ions chelated by nitrilotriacetic acid (NTA) which is bound to a solid support. The hexahistidine sequence of the enzymes coordinates Ni^{2+}/Co^{2+} and is retained on the column while the cell proteins flow through the column.[173] After washing the column, the enzyme can be eluted with imidazole which competitively displaces the imidazole groups of the hexahistidine sequence.

6.8.2.1 Purification of yeast BVMOs

All buffers for purification were cooled to 4 °C before usage. The cells were resuspended in 20 mL disruption buffer. After addition of 10 µM FAD, sonication was performed using the Sonoplus HD2070 (Bandelin) for 7.5 min (50% power and cycle) with a subsequent centrifu-

gation (10000 x g for 45 min at 4 °C). The supernatant was filtered using a 0.45 µm filter. Then, it was applied to a column containing Rotigarose His Beads (Carl Roth). The column was washed with three volumes of disruption buffer and after that with three volumes of washing buffer. The enzyme was subsequently eluted with elution buffer.

6.8.2.1.1 FAD restoration

As it was possible that $BVMO_{albi1}$ was too diluted and to attempt a FAD restoration, a concentration using an Amicon® stirred cell (10 kDa) was employed. Firstly, 10 µM FAD were added to the solution, followed by incubation on ice for 2 h. Afterwards, the solution was concentrated to 1 mL. To wash out the excess of FAD, the solution was filled up to 10 mL with TrisHCl (pH 7.5) and the process was repeated.

6.8.2.1.2 Spectral analysis

For the spectral analysis of the fractions obtained from yeast protein purification, absorption was measured in the range of 250 nm to 800 nm. Buffer (50 mM TrisHCl, pH 7.5) served as blank.

6.8.2.2 Purification of wild-type $CHMO_{Acineto}$ and enzyme variants

Cells were resuspended in 20 mL TrisHCl buffer, pH 9, containing 100 mM NaCl and 10% glycerol (v/v). After addition of 10 µM FAD, sonication was performed for 7.5 min with a subsequent centrifugation (10 000 x *g* for 45 min at 4 °C). The supernatant was filtered using a 0.45 µm filter. Then, it was applied to Co^{2+}-sepharose column material. The column was washed with three volumes of the resuspension buffer and after that with three volumes of the same buffer containing additional 5 mM imidazole. Elution of the enzyme was performed by using the buffer containing 500 mM imidazole. The excess of imidazole was removed by applying 2 mL of enzyme to an EconoPac 10-DG desalting column that has been pre-incubated with the resuspension buffer. Subsequently it was eluted using this buffer once more. The enzyme solution was divided into fractions of 300 µL, quick-frozen using liquid N_2 and stored at –80ºC.

6.8.3 Determination of protein concentration

The concentration of protein solutions was determined using the BC Assay kit (Interchim) according to the manufacturer's instructions for the microtiter plate based assay. The assay uses a purple colored Cu^{+}-protein complex, which enables photometric measurement at 562 nm. All measurements were conducted in triplicates using the Tecan plate reader. Samples were diluted appropriately to fit into the range of the standard curve.

6.8.4 SDS-PAGE

The sodium dodecyl sulfate polyacrylamide gel electrophoresis was performed using the Minigel-Twin (Biometra) system using a 12.5 % separation gel and a 4% stacking gel (Table 6.34).

Table 6.34: Composition of stacking and separation gels.

Component	12.5 % separation gel	4 % stacking gel
Lower TRIS buffer	2 mL	-
Upper TRIS buffer	-	1 mL
Acrylamide solution (30%)	3.33 mL	0.53 mL
Aq. dest.	2.67 mL	2.47 mL
TEMED[b]	4 µL	4 µL
APS solution	40 µL	40 µL
TCE[a]	75 µL	-

[a]: TCE: 2,2,2-Trichloroethanol (if desired)

[b]: N,N,N‘,N‘-Tetramethylethyldiamin

Samples were prepared by mixing 15 µL sample and 15 µL loading buffer followed by heating to 95 °C for 5 min. 15 µL were loaded onto the gel and separated for approximately 1 h at 180 V and 25 mA (50 mA if two gels were separated simultaneously). 4 µL of Pierce Unstained Protein MW Marker (Thermo Scientific) were used as a molecular weight marker. In the most cases, the gel was stained with Coomassie afterwards. For Coomassie staining, the gels were stained with Coomassie staining solution overnight and subsequently destained with destaining solution until the protein bands were clearly distinguishable from the background.
If the application required it, the separation gels contained 1% TCE for the fluorescent visualization of the protein bands. Tryptophans in proteins undergo an ultraviolet light-induced with TCE producing fluorescence, thus enabling a stainless detection of the separated proteins with a detection limit of approx. 2 µg globular protein.[187] After separation, the gels are put onto an UV table (λ = 312 nm) which produces a stable fluorescent signal after 2 min, which can then be photographed.

6.8.5 Biocatalysis

6.8.5.1 Whole cell biocatalysis with growing yeast cells with 2-dodecanone and 1-dodecene

Yeast cells (*Candida maltosa*, *Candida albicans*, *Candida catenulata*, *Candida tropicalis*, *Yarrowia lipolytica*, *Trichosporon asahii*, *Pichia guilliermondii*, *Lodderomyces elongisporus* and *Rhodosporidium toruloides*) from a malt agar plate were dispersed in 100 mL of a mineral salt medium supplemented with 1% (v/v) vitamine solution and 1% (v/v) of the respective substrate (2-dodecanone, 1-dodecene). Cultures grew until an OD_{600nm} of 3 at 30 °C and 250 rpm for 32 h (1-dodecene), 40 h (2-dodecanone, *C. maltosa*), 290 h (*C. albicans*), 72 h (*C. catenulata*) 47 h (*C. tropicalis*), 189 h (*Y. lipolytica*) and 381 h (*T. asahii*). Cultures were centrifuged at 4 °C and 10000 × g for 15 min (Sorvall RC-5B Plus Refrigerated Superspeed Centrifuge). The supernatant was analyzed for secreted products (see section 6.9.2.1). Controls contained either 1% (w/v) dodecane, glucose, no substrate or no cells.

6.8.5.2 Whole cell biocatalysis with $BVMO_{malto}$ and YMOC-H in *E. coli* BL21(DE3)

Whole cell biocatalysis experiments were performed in 2 mL glass vials sealed with an oxygen permeable membrane. Resting cells from cultivation were resuspended in sodium phosphate buffer (50 mM, pH 7.5; 10 mL per gram cell pellet). All reactions were performed in sodium phosphate buffer (50 mM, pH 7.5) in reaction volumes of 400 µL with 40 µL of the cell solutions (0.004 g resting cells, 0.01 g/mL resting cells) and 5 mM substrate (2-dodecanone, cyclohexanone, bicyclo[3.2.0]hept-2-en-6-one, MTS, MPS). All reactions contained glucose for cofactor recycling with a concentration equimolar to the substrate concentration. A set of control samples were taken. A t_0 sample was immediately frozen at -20 °C and used to confirm the initial substrate concentration. A control reaction without resting cells, the substrate control, was used to account for substrate loss during the reaction and autoxidation. Control reactions with substrate and resting cells containing an empty pET28a vector were used to account for a possible conversion of the respective substrate by *E. coli* enzymes. The reactions were started with the addition of the resting cells and were incubated for 24 h at 25 °C and 750 rpm on a Thermomixer comfort (Eppendorf). The approaches were frozen immediately at -20 °C until extraction.

6.8.5.3 Biocatalysis with purified $CHMO_{Acineto}$ in *E. coli* BL21(DE3)

Biocatalysis with purified enzyme was performed in a reaction volume of 1 mL in air-tight 2 mL glass vials (“GC vials”) with 5 mM cyclohexanone, 5 mM NAD(P)H, 0.25 mg/mL de-

salted enzyme eluate and TrisHCl buffer (50 mM, pH 9.0) to reach the reaction volume. The reaction was started by adding the desalted enzyme eluate. After 2 h reaction time, 600 µL samples were taken. A control sample (t_0) containing buffer instead of desalted enzyme eluate was immediately frozen at the start of the reaction to assess the actual substrate concentration.

6.9 Analytical methods

6.9.1 Determination of activity

6.9.1.1 Activity in cell extracts

Enzymatical activity in cell extracts was measured for a faster way to access the performance of the enzymes without a protein purification step. Stock solutions for almost all substrates (Scheme 3.2) were prepared with a concentration of 100 mM in DMF (HPLC grade). The activities were determined by the depletion of NADPH. NADPH, but not its reduced counterpart $NADP^+$, has an absorbance maximum at 340 nm and can therefore be determined photometrically. Measurements were performed in cuvettes. Each reaction volume (1 mL in natrium phosphate buffer, 50 mM, pH 7.5) contained 1 mM substrate, 0.3 mM NADPH and 20 – 100 µL cell extract depending on the apparent activity. The reactions were started with the addition of NADPH and thoroughly mixed. The change in absorbance at 340 nm was measured for two minutes with the V-550 photometer (Jasco). Measurements were performed in triplicate; additionally blanks with no substrate were also determined in triplicate. The limit of detection was defined as the average change of absorbance of the blank measurements plus three times its standard deviation. Analogously, the limit of quantification was set as the average blank plus ten times its standard deviation.[188] A standard curve for NADPH from 0 to 375 µM was prepared in triplicate. The determined molar absorbance coefficient for NADPH was $\varepsilon_{NADPH;\,340\,nm} = 4.90\ mM \cdot 10^3\ L \cdot mol^{-1} \cdot cm^{-1}$

6.9.1.2 Activity of purified $CHMO_{Acineto}$

All $CHMO_{Acineto}$ activity measurements were performed in TrisHCl buffer (50 mM, pH 9.0), with 1 mM cyclohexanone, 0.3 mM NADPH/0.15 mM NADH and 3.19 mg desalted enzyme eluate. Uncoupling was determined without addition of cyclohexanone. The Michaelis–Menten constants, k_{cat} and K_M, for NADPH and NADH were determined by varying the concentrations of the coenzyme in the presence of 1 mM cyclohexanone. The calculations were performed with SigmaPlot 12.0. Enzyme activities were determined spectrophotometrically at 25 °C at 340 nm by measuring the decrease in absorption caused by NAD(P)H consumption. The extinction coefficients were determined to be $4.90 \cdot 10^3\ L \cdot mol^{-1} \cdot cm^{-1}$ at 340 nm and

$0.35 \cdot 10^3$ $L \cdot mol^{-1} \cdot cm^{-1}$ at 390 nm (NADPH) and $6.05 \cdot 10^3$ $L \cdot mol^{-1} \cdot cm^{-1}$ at 340 nm and $0.34 \cdot 10^3$ $L \cdot mol^{-1} \cdot cm^{-1}$ at 390 nm (NADH).

6.9.2 Gas chromatography

6.9.2.1 Identification of metabolites after yeast cultivation

Before extraction, the pH-value of the supernatants was first adjusted to pH 9 by addition of 25% NaOH (v/v). The samples were extracted three times with 50 mL diethyl ether (basic extracts). Next, the pH-value was set to pH 2 by addition of 32% HCl (v/v). Then, samples were extracted again as mentioned above (acidic extracts). Samples were dried over anhydrous sodium sulphate, concentrated in a rotary evaporator and desiccated with nitrogen. For GC/MS analysis the samples were dissolved in 500 µl hexane. The acidic extracts were derivatized by methylation using diazomethane to convert barely volatile acids to more volatile methyl ethers.

For the methylation reaction, a methylation device (Aldrich) was utilized. To produce diazomethane, 2 mL KaOH (40%), 1 mL carbitol (2-(2-ethoxyethoxy)-ethanol), 1 ml diethylether and 1 – 2 tips of a spatula of diazald (*N*-methyl-*N*-nitro-*p*-toluensulfonamide) were added to the methylation device. The resulting diazomethane was directly channeled into the sample through a Pasteur pipette. A yellow coloration of the sample indicated a complete methylation, what not always was clearly visible. The methylated samples were stored at -20 °C until analysis by GC/MS.

Extracts of the alkaline samples were measured with a concentration of 2.5% whereas the extracts of the acidic samples were measured either without dilution or with a concentration of 10%. Dilution was carried out with hexane.

GC/MS analyses were performed on an Agilent gas chromatograph 7890A GC system (Waldbronn, Germany) equipped with a 30 m HP-5 ms column (0.25 mm by 0.25 µm film) and linked to a mass selective detector 5975C inert XL EI/CI MSD with a quadrupole mass spectrometer to identify metabolites formed in course of the cultivation of yeasts. For separation of products a temperature program was used, starting with 5 min at 60 °C followed by a ramp from 60–120 °C at 20 °C/min. The 120 °C were maintained for 5 min and then followed by heating the column to 200 °C at 3 °C/min, to 280 °C at 20 °C/min and lastly 5 min at 280 °C. For quantification of products formed, commercially available standard substances were measured with concentrations from 0.5-10 mM for the creation of calibration curves. Analytical data of the products can be found in the Appendix (Table 8.1).

6.9.2.2 Determination of activity after whole cell biocatalysis

400 µL samples from the biocatalysis reactions were extracted with 400 µL dichloromethane containing 2 mM acetophenone as an internal standard. After vortexing and centrifuging for 1 min each, 400 µL of the bottom, organic phase were collected and dried with Na_2SO_4. After centrifuging for 5 min, 200 µL of the supernatant were collected in a GC vial with inlet. 1 µL per sample were injected by the auto-injector of the GC 2010 (Shimadzu) and separated on a FS-Hydrodex β-3P column (25 m x 0.25 mm ID, Macherey-Nagel) with an injector temperature of 200 °C, a detector temperature of 220 °C and a column temperature program with 60 °C for 10 min, 10 °C/min gradient up to 160 °C and 160 °C for 10 min.

6.9.2.3 Determination of $CHMO_{Acineto}$ activity

Samples (600 µL) from the biocatalysis reactions were extracted with dichloromethane (600 µL) containing 2 mM acetophenone as an external standard. After vortexing and centrifuging for 1 min, 400 µL of the bottom as well as the organic phase were collected and dried with anhydrous Na_2SO_4. After centrifuging for 5 min, 200 µL of the supernatant were transferred to a GC vial with inlet. 1 µL per sample were injected by the auto-injector of the GC 2010 (Shimadzu) and separated on a FS-Hydrodex β-3P column (25 m x 0.25 mm ID, Macherey-Nagel) with an injector temperature of 200 °C, a detector temperature of 220 °C and a column temperature program with 60 °C for 10 min, 10 °C/min gradient up to 160 °C and 160 °C for 10 min.

6.10 Bioinformatical methods

6.10.1 Homology modeling

A homology model was built using YASARA[9b-e] based on a 2.3 Å resolution X-ray structure of the CHMO from *Rhodococcus* sp. (PDB code: 3GWD) in the closed conformation with bound FAD and NADPH.[44c, 189] Among the aligned residues, the sequence identity is 57% and the sequence similarity is 73%. The structural refinement was carried out by energy minimization and molecular dynamics simulation in a water box. The resulting model was qualified as "good" with a Z-score of −0.529 by YASARA.

6.10.2 Sequence alignments

These were carried out using the Geneious software.[190]

7 Literature

[1] D. Voet, J. G. Voet, C. W. Pratt, A. Beck-Sickinger, U. Hahn, *Lehrbuch der Biochemie*, Wiley VCH Verlag GmbH, **2010**.

[2] a) U. Bornscheuer, G. Huisman, R. Kazlauskas, S. Lutz, J. Moore, K. Robins, *Nature* **2012**, *485*, 185-194; b) P. Lorenz, J. Eck, *Nature Reviews Microbiology* **2005**, *3*, 510-516; c) S. Wenda, S. Illner, A. Mell, U. Kragl, *Green Chemistry* **2011**, *13*, 3007-3047.

[3] R. O. M. A. de Souza, L. S. M. Miranda, U. T. Bornscheuer, *Chem. Eur. J.* **2017**, *accepted*.

[4] a) J. M. Berg, J. L. Tymoczko, L. Stryer, B. Häcker, A. Held, *Biochemie*, Spektrum Akademischer Verlag, **2007**; b) K. Buchholz, V. Kasche, U. T. Bornscheuer, *Biocatalysts and enzyme technology*, John Wiley & Sons, **2012**.

[5] U. T. Bornscheuer, R. J. Kazlauskas, *unpublished*.

[6] T. Yoshida, *Enzyme Technology*, Wiley Biotechnology Series, **2015**.

[7] a) D. Quaglia, M. C. Ebert, P. F. Mugford, J. N. Pelletier, *PLoS ONE* **2017**, *12*, e0171741; b) S. Lutz, U. T. Bornscheuer, *Protein Engineering Handbook*, Wiley-VCH, Weinheim,, **2009**; c) R. J. Kazlauskas, U. T. Bornscheuer, *Nature chemical biology* **2009**, *5*, 526-529; d) U. T. Bornscheuer, *Protein Engineering*, KBB, **2017**; e) U. T. Bornscheuer, R. J. Kazlauskas, *Angewandte Chemie International Edition* **2004**, *43*, 6032-6040; f) U. Bornscheuer, R. J. Kazlauskas, *Current Protocols in Protein Science* **2011**, 26.27. 21-26.27. 14; g) T. Davids, M. Schmidt, D. Böttcher, U. T. Bornscheuer, *Current opinion in chemical biology* **2013**, *17*, 215-220.

[8] U. Bornscheuer, in *Handbook of Hydrocarbon and Lipid Microbiology*, Springer, **2010**, pp. 2929-2938.

[9] a) L. A. Kelley, M. J. Sternberg, *Nature protocols* **2009**, *4*, 363-371; b) E. Krieger, K. Joo, J. Lee, J. Lee, S. Raman, J. Thompson, M. Tyka, D. Baker, K. Karplus, *Proteins: Structure, Function, and Bioinformatics* **2009**, *77*, 114-122; c) A. A. Canutescu, A. A. Shelenkov, R. L. Dunbrack, *Protein science* **2003**, *12*, 2001-2014; d) A. C. Marko, K. Stafford, T. Wymore, *Journal of chemical information and modeling* **2007**, *47*, 1263-1270; e) U. Mückstein, I. L. Hofacker, P. F. Stadler, *Bioinformatics* **2002**, *18*, S153-S160.

[10] a) D. A. Estell, T. P. Graycar, J. A. Wells, *Journal of Biological Chemistry* **1985**, *260*, 6518-6521; b) S. Bartsch, R. Kourist, U. T. Bornscheuer, *Angewandte Chemie International Edition* **2008**, *47*, 1508-1511; c) A. O. Magnusson, M. Takwa, A. Hamberg, K. Hult, *Angewandte Chemie International Edition* **2005**, *44*, 4582-4585; d) Y. Terao, Y. Ijima, K. Miyamoto, H. Ohta, *Journal of Molecular Catalysis B: Enzymatic* **2007**, *45*, 15-20.

[11] a) J. D. Carballeira, P. Krumlinde, M. Bocola, A. Vogel, M. T. Reetz, J.-E. Baeckvall, *Chemical Communications* **2007**, 1913-1915; b) D. Zha, S. Wilensek, M. Hermes, M. T. Reetz, *Chemical Communications* **2001**, 2664-2665; c) M. Wada, C.-C. Hsu, D. Franke, M. Mitchell, A. Heine, I. Wilson, C.-H. Wong, *Bioorganic & medicinal chemistry* **2003**, *11*, 2091-2098; d) F. H. Arnold, G. Georgiou, *Directed enzyme evolution: screening and selection methods*, Springer Science & Business Media, **2003**; e) S. Brakmann, K. Johnsson, *Directed molecular evolution of proteins, Vol. 1*, Wiley-VCH, Weinheim, **2002**; f) F. Arnold, G. Georgiou, *Directed evolution library creation: methods and protocols* Humana Press, Totawa, **2003**; g) S. Brakmann, A. Schwienhorst, *Evolutionary Methods in Biotechnology: Clever Tricks for Directed Evolution*, Wiley-VCH, Weinheim, **2004**; h) D. Böttcher, U. T. Bornscheuer, *Current opinion in microbiology* **2010**, *13*, 274-282; i) C. Neylon, *Nucleic acids research*

2004, *32*, 1448-1459; j) M. T. Reetz, *Angewandte Chemie International Edition* **2011**, *50*, 138-174; k) N. J. Turner, *Trends in biotechnology* **2003**, *21*, 474-478.

[12] G. A. Behrens, A. Hummel, S. K. Padhi, S. Schätzle, U. T. Bornscheuer, *Advanced Synthesis & Catalysis* **2011**, *353*, 2191-2215.

[13] a) M. T. Reetz, J. Sanchis, *ChemBioChem* **2008**, *9*, 2260-2267; b) M. T. Reetz, P. Soni, L. Fernández, Y. Gumulya, J. D. Carballeira, *Chemical Communications* **2010**, *46*, 8657-8658; c) D. M. Weinreich, N. F. Delaney, M. A. DePristo, D. L. Hartl, *Science* **2006**, *312*, 111-114; d) R. J. Fox, S. C. Davis, E. C. Mundorff, L. M. Newman, V. Gavrilovic, S. K. Ma, L. M. Chung, C. Ching, S. Tam, S. Muley, *Nature Biotechnology* **2007**, *25*, 338-344; e) M. Lehmann, L. Pasamontes, S. Lassen, M. Wyss, *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology* **2000**, *1543*, 408-415; f) M. T. Reetz, J. D. Carballeira, *Nature protocols* **2007**, *2*, 891-903.

[14] E. Vazquez-Figueroa, V. Yeh, J. Broering, J. Chaparro-Riggers, A. Bommarius, *Protein Engineering Design and Selection* **2008**, *21*, 673-680.

[15] a) H. Jochens, D. Aerts, U. T. Bornscheuer, *Protein Engineering Design and Selection* **2010**, *23*, 903-909; b) H. Jochens, U. T. Bornscheuer, *ChemBioChem* **2010**, *11*, 1861-1866; c) R. Kourist, H. Jochens, S. Bartsch, R. Kuipers, S. K. Padhi, M. Gall, D. Böttcher, H. J. Joosten, U. T. Bornscheuer, *ChemBioChem* **2010**, *11*, 1635-1643; d) R. K. Kuipers, H. J. Joosten, W. J. van Berkel, N. G. Leferink, E. Rooijen, E. Ittmann, F. van Zimmeren, H. Jochens, U. Bornscheuer, G. Vriend, *Proteins: Structure, Function, and Bioinformatics* **2010**, *78*, 2101-2113.

[16] Z. Qian, S. Lutz, *Journal of the American Chemical Society* **2005**, *127*, 13466-13467.

[17] M. Schmidt, D. Hasenpusch, M. Kähler, U. Kirchner, K. Wiggenhorn, W. Langel, U. T. Bornscheuer, *ChemBioChem* **2006**, *7*, 805-809.

[18] M. M. E. Huijbers, S. Montersino, A. H. Westphal, D. Tischler, W. J. H. van Berkel, *Archives of biochemistry and biophysics* **2014**, *544*, 2-17.

[19] W. Van Berkel, N. Kamerbeek, M. Fraaije, *Journal of biotechnology* **2006**, *124*, 670-689.

[20] D. T. Pazmino, M. Winkler, A. Glieder, M. Fraaije, *Journal of biotechnology* **2010**, *146*, 9-24.

[21] S. Montersino, D. Tischler, G. T. Gassner, W. J. van Berkel, *Advanced Synthesis & Catalysis* **2011**, *353*, 2301-2319.

[22] C. J. Balibar, C. T. Walsh, *Biochemistry* **2006**, *45*, 15444-15457.

[23] B. T. Greenhagen, K. Shi, H. Robinson, S. Gamage, A. K. Bera, J. E. Ladner, J. F. Parsons, *Biochemistry* **2008**, *47*, 5281-5289.

[24] D. V. Mavrodi, R. F. Bonsall, S. M. Delaney, M. J. Soule, G. Phillips, L. S. Thomashow, *Journal of bacteriology* **2001**, *183*, 6454-6465.

[25] Peter J. Goldman, Katherine S. Ryan, Michael J. Hamill, Annaleise R. Howard-Jones, Christopher T. Walsh, Sean J. Elliott, Catherine L. Drennan, *Chemistry & Biology* **2012**, *19*, 855-865.

[26] G. Volkers, G. J. Palm, M. S. Weiss, G. D. Wright, W. Hinrichs, *FEBS Letters* **2011**, *585*, 1061-1066.

[27] K. Walkiewicz, M. Davlieva, G. Wu, Y. Shamoo, *Proteins: Structure, Function, and Bioinformatics* **2011**, *79*, 2335-2340.

[28] Y. Katsuyama, K. Harmrolfs, D. Pistorius, Y. Li, R. Müller, *Angewandte Chemie International Edition* **2012**, *51*, 9437-9440.

[29] B. D. Ames, S. W. Haynes, X. Gao, B. S. Evans, N. L. Kelleher, Y. Tang, C. T. Walsh, *Biochemistry* **2011**, *50*, 8756.

[30] X. Gao, Y.-H. Chooi, B. D. Ames, P. Wang, C. T. Walsh, Y. Tang, *Journal of the American Chemical Society* **2011**, *133*, 2729.

[31] S. W. Haynes, X. Gao, Y. Tang, C. T. Walsh, *Journal of the American Chemical Society* **2012**, *134*, 17444.
[32] Z. Rui, M. Sandy, B. Jung, W. Zhang, *Chemistry & Biology* **2013**, *20*, 879-887.
[33] M. Baunach, L. Ding, T. Bruhn, G. Bringmann, C. Hertweck, *Angewandte Chemie International Edition* **2013**, *52*, 9040-9043.
[34] a) A. Riebel, H. M. Dudek, G. de Gonzalo, P. Stepniak, L. Rychlewski, M. W. Fraaije, *Appl Microbiol Biotechnol* **2012**, *95*, 1479-1489; b) P. Ballou David, B. Entsch, in *Handbook of Flavoproteins Volume 2 Complex Flavoproteins, Dehydrogenases and Physical Methods*, **2013**.
[35] M. L. Mascotti, W. J. Lapadula, M. J. Ayub, *PLoS ONE* **2015**, *10*, e0132689.
[36] a) J. R. Cashman, J. Zhang, *Annu. Rev. Pharmacol. Toxicol.* **2006**, *46*, 65-100; b) D. M. Ziegler, *Drug metabolism reviews* **2002**, *34*, 503-511.
[37] a) A. Kantz, F. Chin, N. Nallamothu, T. Nguyen, G. T. Gassner, *Archives of biochemistry and biophysics* **2005**, *442*, 102-116; b) K. Otto, K. Hofstetter, M. Röthlisberger, B. Witholt, A. Schmid, *Journal of bacteriology* **2004**, *186*, 5292-5302.
[38] P. F. Fitzpatrick, *Archives of biochemistry and biophysics* **2010**, *493*, 13-25.
[39] A. von Baeyer, V. Villiger, *Berichte der deutschen chemischen Gesellschaft* **1899**, *32*, 3625-3633.
[40] R. Criegee, *European Journal of Organic Chemistry* **1948**, *560*, 127-135.
[41] H. Leisch, K. Morley, P. C. Lau, *Chemical reviews* **2011**, *111*, 4165-4222.
[42] a) D. Sheng, D. P. Ballou, V. Massey, *Biochemistry* **2001**, *40*, 11156-11167; b) M. D. Mihovilovic, in *Enzyme Catalysis in Organic Synthesis*, Third ed. (Eds.: K. Drauz, H. Groger, O. May), Wiley-VCH, **2012**, pp. 1439-1485.
[43] C. C. Ryerson, D. P. Ballou, C. Walsh, *Biochemistry* **1982**, *21*, 2644-2655.
[44] a) B. J. Yachnin, T. Sprules, M. B. McEvoy, P. C. Lau, A. M. Berghuis, *Journal of the American Chemical Society* **2012**, *134*, 7788-7795; b) B. J. Yachnin, M. B. McEvoy, R. J. MacCuish, K. L. Morley, P. C. Lau, A. M. Berghuis, *ACS chemical biology* **2014**, *9*, 2843-2851; c) I. A. Mirza, B. J. Yachnin, S. Wang, S. Grosse, H. Bergeron, A. Imura, H. Iwaki, Y. Hasegawa, P. C. Lau, A. M. Berghuis, *Journal of the American Chemical Society* **2009**, *131*, 8848-8854.
[45] a) F. M. Ferroni, C. Tolmie, M. S. Smit, D. J. Opperman, *PLoS ONE* **2016**, *11*, e0160186; b) E. Malito, A. Alfieri, M. W. Fraaije, A. Mattevi, *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 13157-13162; c) H. Leisch, R. Shi, S. Grosse, K. Morley, H. Bergeron, M. Cygler, H. Iwaki, Y. Hasegawa, P. C. Lau, *Applied and environmental microbiology* **2012**, *78*, 2200-2212; d) S. Franceschini, H. L. van Beek, A. Pennetta, C. Martinoli, M. W. Fraaije, A. Mattevi, *Journal of Biological Chemistry* **2012**, *287*, 22626-22634; e) M. N. Isupov, E. Schröder, R. P. Gibson, J. Beecher, G. Donadio, V. Saneei, S. A. Dcunha, E. J. McGhie, C. Sayer, C. F. Davenport, *Acta Crystallographica Section D: Biological Crystallography* **2015**, *71*, 2344-2353; f) M. P. Beam, M. A. Bosserman, N. Noinaj, M. Wehenkel, J. Rohr, *Biochemistry* **2009**, *48*, 4476; g) E. Romero, J. Castellanos, A. Mattevi, M. W. Fraaije, *Angewandte Chemie* **2016**, *128*, 16084-16087; h) M. J. Fürst, S. Savino, H. M. Dudek, J. R. Gomez Castellanos, C. Gutiérrez de Souza, S. Rovida, M. W. Fraaije, A. Mattevi, *Journal of the American Chemical Society* **2017**; i) E. McGhie, M. Isupov, E. Schröder, J. Littlechild, *Acta Crystallographica Section D: Biological Crystallography* **1998**, *54*, 1035-1038; j) M. A. Bosserman, T. Downey, N. Noinaj, S. K. Buchanan, J. Rohr, *ACS chemical biology* **2013**, *8*; k) C. Martinoli, H. M. Dudek, R. Orru, D. E. Edmondson, M. W. Fraaije, A. Mattevi, *ACS catalysis* **2013**, *3*, 3058.

[46] a) A. Willetts, I. Joint, J. A. Gilbert, W. Trimble, M. Mühling, *Microbial biotechnology* **2012**, *5*, 549-559; b) C. N. Jensen, J. Cartwright, J. Ward, S. Hart, J. P. Turkenburg, S. T. Ali, M. J. Allen, G. Grogan, *ChemBioChem* **2012**, *13*, 872-878.

[47] a) K. Balke, S. Schmidt, M. Genz, U. T. Bornscheuer, *ACS chemical biology* **2016**, *11*, 38-43; b) D. R. Kelly, C. J. Knowles, J. G. Mahdi, I. N. Taylor, M. A. Wright, *Journal of the Chemical Society, Chemical Communications* **1995**, 729-730; c) I. Polyak, M. T. Reetz, W. Thiel, *Journal of the American Chemical Society* **2012**, *134*, 2732-2741.

[48] a) D. R. Kelly, C. J. Knowles, J. G. Mahdi, M. A. Wright, I. N. Taylor, D. E. Hibbs, M. B. Hursthouse, A. K. Mish'al, S. M. Roberts, P. W. H. Wan, G. Grogan, A. J. Willetts, *Journal of the Chemical Society, Perkin Transactions 1* **1995**, 2057-2066; b) D. R. Kelly, C. J. Knowles, J. G. Mahdi, M. A. Wright, I. N. Taylor, S. M. Roberts, P. W. H. Wan, G. Grogan, S. Pedragosa-Moreau, A. J. Willetts, *Journal of the Chemical Society, Chemical Communications* **1996**, *20*, 2333.

[49] a) M. D. Mihovilovic, P. Kapitán, *Tetrahedron Letters* **2004**, *45*, 2751-2754; b) R. Snajdrova, G. Grogan, M. D. Mihovilovic, *Bioorganic & medicinal chemistry letters* **2006**, *16*, 4813-4817; c) J. Rehdorf, A. Lengar, U. T. Bornscheuer, M. D. Mihovilovic, *Bioorganic & medicinal chemistry letters* **2009**, *19*, 3739-3743; d) J. Rehdorf, M. D. Mihovilovic, U. T. Bornscheuer, *Angewandte Chemie International Edition* **2010**, *49*, 4506-4508.

[50] G. de Gonzalo, M. D. Mihovilovic, M. W. Fraaije, *ChemBioChem* **2010**, *11*, 2208-2231.

[51] a) N. M. Kamerbeek, A. J. Olsthoorn, M. W. Fraaije, D. B. Janssen, *Applied and environmental microbiology* **2003**, *69*, 419-426; b) C. Szolkowy, L. D. Eltis, N. C. Bruce, G. Grogan, *ChemBioChem* **2009**, *10*, 1208-1217.

[52] Y. Hu, D. Dietrich, W. Xu, A. Patel, J. A. Thuss, J. Wang, W.-B. Yin, K. Qiao, K. N. Houk, J. C. Vederas, *Nature chemical biology* **2014**, *10*, 552-554.

[53] a) D. J. Opperman, M. T. Reetz, *ChemBioChem* **2010**, *11*, 2589-2596; b) A. Beier, S. Bordewick, M. Genz, S. Schmidt, T. van den Bergh, C. Peters, H. J. Joosten, U. T. Bornscheuer, *ChemBioChem* **2016**, *17*, 2312-2315; c) B. J. Yachnin, P. C. Lau, A. M. Berghuis, *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2016**, *1864*, 1641-1648.

[54] a) A. Suemori, M. Iwakura, *Journal of Biological Chemistry* **2007**, *282*, 19969-19978; b) C. Miyazaki-Imamura, K. Oohira, R. Kitagawa, H. Nakano, T. Yamane, H. Takahashi, *Protein engineering* **2003**, *16*, 423-428; c) L.-L. Lin, H.-F. Lo, W.-Y. Chiang, H.-Y. Hu, W.-H. Hsu, C.-T. Chang, *Current microbiology* **2003**, *46*, 0211-0216.

[55] a) A. Slavica, I. Dib, B. Nidetzky, *Applied and environmental microbiology* **2005**, *71*, 8061-8068; b) L. Perry, R. Wetzel, *Protein engineering* **1987**, *1*, 101-105; c) Y. H. Kim, A. H. Berry, D. S. Spencer, W. E. Stites, *Protein engineering* **2001**, *14*, 343-347.

[56] J. K. Cahn, PhD thesis, University of California (California), **2016**.

[57] H. Iwaki, S. Wang, S. Grosse, H. Bergeron, A. Nagahashi, J. Lertvorachon, J. Yang, Y. Konishi, Y. Hasegawa, P. C. Lau, *Applied and environmental microbiology* **2006**, *72*, 2707-2720.

[58] a) K. Balke, M. Kadow, H. Mallin, S. Saß, U. T. Bornscheuer, *Organic & biomolecular chemistry* **2012**, *10*, 6249-6265; b) J. Rehdorf, U. T. Bornscheuer, *Encyclopedia of Industrial Biotechnology: Bioprocess, Bioseparation, and Cell Technology* **2009**.

[59] a) O. Carugo, P. Argos, *Proteins: Structure, Function, and Bioinformatics* **1997**, *28*, 10-28; b) C. Didierjean, S. Rahuel-Clermont, B. Vitoux, O. Dideberg, G. Branlant, A. Aubry, *Journal of molecular biology* **1997**, *268*, 739-759; c) G. Kleiger, D. Eisenberg, *Journal of molecular biology* **2002**, *323*, 69-76.

[60] M. W. Fraaije, N. M. Kamerbeek, W. J. van Berkel, D. B. Janssen, *FEBS letters* **2002**, *518*, 43-47.
[61] J. Rebehmed, V. Alphand, V. d. Berardinis, A. G. de Brevern, *Biochimie* **2013**.
[62] A. Riebel, G. de Gonzalo, M. W. Fraaije, *Journal of Molecular Catalysis B: Enzymatic* **2013**, *88*, 20-25.
[63] A. Beier, Diploma thesis, Ernst-Moritz-Arndt University Greifswald (Greifswald), **2013**.
[64] M. Kadow, S. Saß, K. Balke, M. Schmidt, U. T. Bornscheuer, *ChemBioChem* **2012**.
[65] a) M. Kadow, K. Balke, A. Willetts, U. T. Bornscheuer, J.-E. Bäckvall, *Appl Microbiol Biotechnol* **2014**, *98*, 3975-3986; b) H. Iwaki, S. Grosse, H. Bergeron, H. Leisch, K. Morley, Y. Hasegawa, P. C. Lau, *Applied and environmental microbiology* **2013**, *79*, 3282-3293.
[66] a) J. D. Stewart, *Current Organic Chemistry* **1998**, *2*, 195-216; b) M. Bučko, P. Gemeiner, A. Schenkmayerová, T. Krajčovič, F. Rudroff, M. D. Mihovilovič, *Appl Microbiol Biotechnol* **2016**, *100*, 6585-6599.
[67] a) G. Ottolina, P. Pasta, G. Carrea, S. Colonna, S. Dallavalle, H. L. Holland, *Tetrahedron: Asymmetry* **1995**, *6*, 1375-1386; b) S. Colonna, V. Pironti, P. Pasta, F. Zambianchi, *Tetrahedron letters* **2003**, *44*, 869-871; c) S. Colonna, N. Gaggero, G. Carrea, G. Ottolina, P. Pasta, F. Zambianchi, *Tetrahedron Letters* **2002**, *43*, 1797-1799; d) M. J. Moonen, A. H. Westphal, I. M. Rietjens, W. J. van Berkel, *Advanced Synthesis & Catalysis* **2005**, *347*, 1027-1034; e) G. Catucci, I. Zgrablic, F. Lanciani, F. Valetti, D. Minerdi, D. P. Ballou, G. Gilardi, S. J. Sadeghi, *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2016**, *1864*, 1177-1187; f) S. Colonna, V. Pironti, G. Carrea, P. Pasta, F. Zambianchi, *Tetrahedron* **2004**, *60*, 569-575; g) M. T. Reetz, F. Daligault, B. Brunner, H. Hinrichs, A. Deege, *Angewandte Chemie* **2004**, *116*, 4170-4173; h) F. M. Ferroni, C. Tolmie, M. S. Smit, D. J. Opperman, *ChemBioChem* **2017**, *18*, 515-517.
[68] Z.-G. Zhang, R. Lonsdale, J. Sanchis, M. T. Reetz, *J. Am. Chem. Soc.* **2014**, *136*, 17262-17272.
[69] S. Bordewick, Master thesis, University of Greifswald (Greifswald), **2016**.
[70] a) G. Carrea, B. Redigolo, S. Riva, S. Colonna, N. Gaggero, E. Battistel, D. Bianchi, *Tetrahedron: Asymmetry* **1992**, *3*, 1063-1068; b) S. Colonna, N. Gaggero, P. Pasta, G. Ottolina, *Chemical Communications* **1996**, 2303-2307; c) H. M. Dudek, D. E. T. Pazmiño, C. Rodríguez, G. de Gonzalo, V. Gotor, M. W. Fraaije, *Appl Microbiol Biotechnol* **2010**, *88*, 1135-1143; d) G. de Gonzalo, D. E. T. Pazmino, G. Ottolina, M. W. Fraaije, G. Carrea, *Tetrahedron: Asymmetry* **2006**, *17*, 130-135; e) G. de Gonzalo, D. E. T. Pazmiño, G. Ottolina, M. W. Fraaije, G. Carrea, *Tetrahedron: Asymmetry* **2005**, *16*, 3077-3083.
[71] M. L. Mascotti, M. J. Ayub, H. Dudek, M. K. Sanz, M. W. Fraaije, *AMB Express* **2013**, *3*, 33.
[72] a) Y. K. Bong, M. D. Clay, S. J. Collier, B. Mijts, M. Vogel, X. Zhang, J. Zhu, J. Nazor, D. J. Smith, S. Song, Google Patents, **2016**; b) Y. Bong, M. Clay, S. Collier, B. Mijts, M. Vogel, X. Zhang, J. Zhu, J. Nazor, D. Smith, S. Song, *PCT Int Appl WO2011071982A2* **2011**; c) H. B. Kagan, in *Organosulfur Chemistry in Asymmetric Synthesis*, Wiley-VCH Verlag GmbH & Co. KGaA, **2009**; d) T. Matsui, Y. Dekishima, M. Ueda, *Appl Microbiol Biotechnol* **2014**, *98*, 7699-7706; e) I. Fernández, N. Khiar, *Chemical reviews* **2003**, *103*, 3651-3706; f) M. C. Carreño, G. Hernández-Torres, M. Ribagorda, A. Urbano, *Chemical Communications* **2009**, 6129-6144.
[73] R. Orru, H. M. Dudek, C. Martinoli, D. E. T. Pazmiño, A. Royant, M. Weik, M. W. Fraaije, A. Mattevi, *Journal of Biological Chemistry* **2011**, *286*, 29284-29291.

[74] S. Schmidt, M. Genz, K. Balke, U. T. Bornscheuer, *Journal of biotechnology* **2015**, *214*, 199-211.
[75] K. Weissermel, H. J. Arpe, *Industrial Organic Chemistry, Fourth Edition* **2003**, 145-192.
[76] a) S. Schmidt, H. C. Büchsenschütz, C. Scherkus, A. Liese, H. Gröger, U. T. Bornscheuer, *ChemCatChem* **2015**, *7*, 3951-3955; b) S. Schmidt, C. Scherkus, J. Muschiol, U. Menyes, T. Winkler, W. Hummel, H. Gröger, A. Liese, H. G. Herz, U. T. Bornscheuer, *Angewandte Chemie International Edition* **2015**, *54*, 2784-2787.
[77] a) M. Labet, W. Thielemans, *Chemical Society Reviews* **2009**, *38*, 3484-3504; b) F. J. v. Natta, J. W. Hill, W. H. Carothers, *Journal of the American Chemical Society* **1934**, *56*, 455-457.
[78] M. J. Fink, F. Rudroff, M. D. Mihovilovic, *Bioorganic & medicinal chemistry letters* **2011**, *21*, 6135-6138.
[79] S. D. Doig, P. J. Avenell, P. A. Bird, P. Gallati, K. S. Lander, G. J. Lye, R. Wohlgemuth, J. M. Woodley, *Biotechnology progress* **2002**, *18*, 1039-1046.
[80] C. V. Baldwin, R. Wohlgemuth, J. M. Woodley, *Organic Process Research & Development* **2008**, *12*, 660-665.
[81] G. E. Turfitt, *Biochem J* **1948**, *42*, 376-383.
[82] A. Willetts, *Trends Biotechnol* **1997**, *15*, 55-62.
[83] H. Iwaki, Y. Hasegawa, M. Teraoka, T. Tokuyama, H. Bergeron, P. C. Lau, *Applied and environmental microbiology* **1999**, *65*, 5158-5162.
[84] N. A. Donoghue, D. B. Norris, P. W. Trudgill, *European Journal of Biochemistry* **1976**, *63*, 175-192.
[85] M. D. Mihovilovic, B. Müller, P. Stanetty, *European Journal of Organic Chemistry* **2002**, *2002*, 3711-3730.
[86] H. J. Ougham, D. G. Taylor, P. W. Trudgill, *Journal of bacteriology* **1983**, *153*, 140-152.
[87] M. Kadow, S. Sass, M. Schmidt, U. T. Bornscheuer, *AMB Express* **2011**, *1*, 13.
[88] B. Unger, I. Gunsalus, S. G. Sligar, *Journal of Biological Chemistry* **1986**, *261*, 1158-1163.
[89] H. Koga, E. Yamaguchi, K. Matsunaga, H. Aramaki, T. Horiuchi, *Journal of biochemistry* **1989**, *106*, 831-836.
[90] K. H. Jones, R. T. Smith, P. W. Trudgill, *Microbiology* **1993**, *139*, 797-805.
[91] M. Kadow, K. Loschinski, S. Saß, M. Schmidt, U. T. Bornscheuer, *Appl Microbiol Biotechnol* **2012**, *96*, 419-429.
[92] H. Iwaki, Y. Hasegawa, S. Wang, M. M. Kayser, P. C. Lau, *Applied and environmental microbiology* **2002**, *68*, 5671-5684.
[93] a) N. M. Kamerbeek, M. J. Moonen, J. G. Van der Ven, W. J. Van Berkel, M. W. Fraaije, D. B. Janssen, *European Journal of Biochemistry* **2001**, *268*, 2547-2557; b) N. M. Kamerbeek, D. B. Janssen, W. J. van Berkel, M. W. Fraaije, *Advanced Synthesis & Catalysis* **2003**, *345*, 667-678.
[94] M. W. Fraaije, J. Wu, D. P. Heuts, E. W. van Hellemond, J. H. Spelberg, D. B. Janssen, *Appl Microbiol Biotechnol* **2005**, *66*, 393-400.
[95] D. T. Pazmiño, M. W. Fraaije, *Future directions in biocatalysis, Elsevier Sciences* **2007**, 107-128.
[96] F. Leipold, R. Wardenga, U. T. Bornscheuer, *Appl Microbiol Biotechnol* **2012**, *94*, 705-717.
[97] F. Ferroni, M. Smit, D. Opperman, *Journal of Molecular Catalysis B: Enzymatic* **2014**, *107*, 47-54.
[98] H. L. van Beek, G. de Gonzalo, M. W. Fraaije, *Chemical Communications* **2012**, *48*, 3288-3290.

[99] E. Beneventi, M. Niero, R. Motterle, M. Fraaije, E. Bergantino, *Journal of Molecular Catalysis B: Enzymatic* **2013**, *98*, 145-154.

[100] a) F. Schauer, M. Gesell, S. Mundt, U. Lindequist, *Bodden* **1997**, *5*, 45-56; b) M. J. Klug, A. J. Markovetz, *Applied microbiology* **1967**, *15*, 690-693; c) C. E. Lowery, Jr., J. W. Foster, P. Jurtshuk, *Archiv fur Mikrobiologie* **1968**, *60*, 246-254.

[101] a) S. Fukui, A. Tanaka, *Trends in Biochemical Sciences* **1979**, *4*, 246-249; b) M. Gallo, J. Bertrand, B. Roche, E. Azoulay, *Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism* **1973**, *296*, 624-638; c) J.-M. Lebeault, E. T. Lode, M. J. Coon, *Biochemical and biophysical research communications* **1971**, *42*, 413-419; d) M. Osumi, F. Fukuzumi, Y. Teranishi, A. Tanaka, S. Fukui, *Archives of Microbiology* **1975**, *103*, 1-11.

[102] Y. Ji, G. Mao, Y. Wang, M. Bartlam, *Frontiers in microbiology* **2013**, *4*.

[103] a) R. Blasig, S. Mauersberger, P. Riege, W.-H. Schunck, W. Jockisch, P. Franke, H.-G. Müller, *Appl Microbiol Biotechnol* **1988**, *28*, 589-597; b) O. Käppeli, *Microbiological reviews* **1986**, *50*, 244; c) S. Mauersberger, M. Ohkuma, W.-H. Schunck, M. Takagi, in *Nonconventional Yeasts in Biotechnology*, Springer, **1996**, pp. 411-580; d) W.-H. Schunck, S. Mauersberger, J. Huth, P. Riege, H.-G. Müller, *Archives of microbiology* **1987**, *147*, 240-244; e) R. J. Watkinson, P. Morgan, in *Physiology of Biodegradative Microorganisms*, Springer, **1991**, pp. 79-92; f) M. J. Coon, *Biochemical and biophysical research communications* **2005**, *338*, 378-385; g) A. Kester, J. Foster, *Journal of bacteriology* **1963**, *85*, 859-869; h) F. W. Forney, A. J. Markovetz, *Journal of bacteriology* **1970**, *102*, 281-282; i) H. Rehm, I. Reiff, *Mechanisms and occurrence of microbial oxidation of long-chain alkanes*, Springer, **1981**.

[104] T. Kotani, H. Yurimoto, N. Kato, Y. Sakai, *Journal of bacteriology* **2007**, *189*, 886-893.

[105] F. Forney, A. Markovetz, *Biochemical and biophysical research communications* **1969**, *37*, 31-38.

[106] a) C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T. Madden, *BMC bioinformatics* **2009**, *10*, 421; b) S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *Journal of molecular biology* **1990**, *215*, 403-410; c) B. Dujon, D. Sherman, G. Fischer, P. Durrens, S. Casaregola, I. Lafontaine, J. De Montigny, C. Marck, C. Neuvéglise, E. Talla, *Nature* **2004**, *430*, 35-44.

[107] S. Eswaramoorthy, J. B. Bonanno, S. K. Burley, S. Swaminathan, *Proceedings of the National Academy of Sciences* **2006**, *103*, 9832-9837.

[108] L. Butinar, M. Mohorčič, V. Deyris, K. Duquesne, G. Iacazio, M. Claeys-Bruno, J. Friedrich, V. Alphand, *Phytochemistry* **2015**, *117*, 144-153.

[109] a) D. E. Torres Pazmiño, R. Snajdrova, B. J. Baas, M. Ghobrial, M. D. Mihovilovic, M. W. Fraaije, *Angewandte Chemie* **2008**, *120*, 2307-2310; b) D. E. Torres Pazmiño, A. Riebel, J. de Lange, F. Rudroff, M. D. Mihovilovic, M. W. Fraaije, *ChemBioChem* **2009**, *10*, 2595-2598.

[110] J. Muschiol, C. Peters, N. Oberleitner, M. D. Mihovilovic, U. T. Bornscheuer, F. Rudroff, *Chemical Communications* **2015**, *51*, 5798-5811.

[111] J. H. Sattler, M. Fuchs, F. G. Mutti, B. Grischek, P. Engel, J. Pfeffer, J. M. Woodley, W. Kroutil, *Angewandte Chemie International Edition* **2014**, *53*, 14153-14157.

[112] L. M. O'Sullivan, S. Patel, J. M. Ward, J. M. Woodley, S. D. Doig, *Enzyme and microbial technology* **2001**, *28*, 265-274.

[113] M. D. Mihovilovic, *Current Organic Chemistry* **2006**, *10*, 1265-1287.

[114] N. Oberleitner, A. Ressmann, K. Bica, P. Gärtner, M. Fraaije, U. Bornscheuer, F. Rudroff, M. Mihovilovic, *Green Chemistry* **2016**.

[115] J. T. Wu, L. H. Wu, J. A. Knight, *Clinical chemistry* **1986**, *32*, 314-319.

[116] A. Weckbecker, H. Gröger, W. Hummel, in *Biosystems Engineering I*, Springer, **2010**, pp. 195-242.
[117] Y. Wang, K.-Y. San, G. N. Bennett, *Current Opinion in Biotechnology* **2013**, *24*, 994-999.
[118] a) A. Matsushika, S. Watanabe, T. Kodaki, K. Makino, H. Inoue, K. Murakami, O. Takimura, S. Sawayama, *Appl Microbiol Biotechnol* **2008**, *81*, 243-255; b) O. Bengtsson, B. Hahn-Hägerdal, M. F. Gorwa-Grauslund, *Biotechnology for Biofuels* **2009**, *2*, 9; c) S. Hasegawa, K. Uematsu, Y. Natsuma, M. Suda, K. Hiraga, T. Jojima, M. Inui, H. Yukawa, *Applied and environmental microbiology* **2012**, *78*, 865-875; d) S. Bastian, X. Liu, J. T. Meyerowitz, C. D. Snow, M. M. Chen, F. H. Arnold, *Metabolic engineering* **2011**, *13*, 345-352; e) H. Tamakawa, S. Ikushima, S. Yoshida, *Bioscience, biotechnology, and biochemistry* **2011**, *75*, 1994-2000; f) C. R. Shen, E. I. Lan, Y. Dekishima, A. Baez, K. M. Cho, J. C. Liao, *Applied and environmental microbiology* **2011**, *77*, 2905-2915; g) Z. Xiao, C. Lv, C. Gao, J. Qin, C. Ma, Z. Liu, P. Liu, L. Li, P. Xu, *PLoS ONE* **2010**, *5*, e8860; h) J. K. Cahn, C. A. Werlang, A. Baumschlager, S. Brinkmann-Chen, S. L. Mayo, F. H. Arnold, *ACS Synthetic Biology* **2016**.
[119] N. M. Kamerbeek, M. W. Fraaije, D. B. Janssen, *European Journal of Biochemistry* **2004**, *271*, 2107-2116.
[120] N. S. Scrutton, A. Berry, R. N. Perham, *Nature* **1990**, *343*, 38.
[121] a) S. Brinkmann-Chen, T. Flock, J. K. Cahn, C. D. Snow, E. M. Brustad, J. A. McIntosh, P. Meinhold, L. Zhang, F. H. Arnold, *Proceedings of the National Academy of Sciences* **2013**, *110*, 10946-10951; b) B. Nakhostin-Roohi, S. Barmaki, F. Khoshkhahesh, S. Bohlooli, *Journal of Pharmacy and Pharmacology* **2011**, *63*, 1290-1294; c) L. Kim, L. Axelrod, P. Howard, N. Buratovich, R. Waters, *Osteoarthritis and Cartilage* **2006**, *14*, 286-294; d) J. K. Cahn, A. Baumschlager, S. Brinkmann-Chen, F. H. Arnold, *Protein engineering, design & selection : PEDS* **2016**, *29*, 31-38; e) T. M. Penning, J. M. Jez, *Chemical reviews* **2001**, *101*, 3027-3046.
[122] D. J. Maddock, W. M. Patrick, M. L. Gerth, *Protein Engineering Design and Selection* **2015**, *28*, 251-258.
[123] a) A. D. Mesecar, B. L. Stoddard, D. E. Koshland, *Science* **1997**, *277*, 202-206; b) C. E. Naylor, S. Gover, A. K. Basak, M. S. Cosgrove, H. R. Levy, M. J. Adams, *Acta Crystallographica Section D: Biological Crystallography* **2001**, *57*, 635-648; c) K. L. Kavanagh, M. Klimacek, B. Nidetzky, D. K. Wilson, *Biochem J* **2003**, *373*, 319-326; d) E. Campbell, S. Chuang, S. Banta, *Protein Engineering Design and Selection* **2012**, gzs095; e) B. V. Plapp, *Archives of biochemistry and biophysics* **2010**, *493*, 3-12; f) G. A. Khoury, H. Fazelinia, J. W. Chin, R. J. Pantazes, P. C. Cirino, C. D. Maranas, *Protein science* **2009**, *18*, 2125-2138.
[124] a) J. K. Cahn, S. Brinkmann-Chen, T. Spatzal, J. A. Wiig, A. R. Buller, O. Einsle, Y. Hu, M. W. Ribbe, F. H. Arnold, *Biochem J* **2015**, *468*, 475-484; b) J. R. Wallen, C. Paige, T. C. Mallett, P. A. Karplus, A. Claiborne, *Biochemistry* **2008**, *47*, 5182; c) K. A. Denessiouk, V. V. Rantanen, M. S. Johnson, *Proteins: Structure, Function, and Bioinformatics* **2001**, *44*, 282-291; d) S. L. Moodie, J. B. Mitchell, J. M. Thornton, *Journal of molecular biology* **1996**, *263*, 486-500.
[125] a) S. Brinkmann-Chen, J. Cahn, F. H. Arnold, *Metabolic engineering* **2014**, *26*, 17-22; b) E. Di Luccio, R. A. Elling, D. K. Wilson, *Biochem J* **2006**, *400*, 105-114; c) R. Hatrongjit, K. Packdibamrung, *Enzyme and microbial technology* **2010**, *46*, 557-561; d) C. N. Jensen, S. T. Ali, M. J. Allen, G. Grogan, *Journal of Molecular Catalysis B: Enzymatic* **2014**, *109*, 191-198.

[126] a) J. P. Gallivan, D. A. Dougherty, *Proceedings of the National Academy of Sciences* **1999**, *96*, 9459-9464; b) S.-H. Wong, T. G. Lonhienne, D. J. Winzor, G. Schenk, L. W. Guddat, *Journal of molecular biology* **2012**, *424*, 168-179.
[127] a) A. Rodríguez-Arnedo, M. Camacho, F. Llorca, M.-J. Bonete, *The protein journal* **2005**, *24*, 259-266; b) A. Rosell, E. Valencia, W. F. Ochoa, I. Fita, X. Parés, J. Farrés, *Journal of Biological Chemistry* **2003**, *278*, 40573-40580.
[128] a) C. L. Elmore, T. D. Porter, *Journal of Biological Chemistry* **2002**, *277*, 48960-48964; b) K. Kristan, J. Stojan, J. Adamski, T. L. Rižner, *Journal of biotechnology* **2007**, *129*, 123-130; c) B. Petschacher, S. Leitgeb, K. L. Kavanagh, D. K. Wilson, B. Nidetzky, *Biochem J* **2005**, *385*, 75-83; d) N. Shiraishi, C. Croy, J. Kaur, W. H. Campbell, *Archives of biochemistry and biophysics* **1998**, *358*, 104-115; e) M. J. Rane, K. Calvo, *Archives of biochemistry and biophysics* **1997**, *338*, 83-89.
[129] I. Nobeli, R. A. Laskowski, W. S. Valdar, J. M. Thornton, *Nucleic acids research* **2001**, *29*, 4294-4309.
[130] A. Völker, A. Kirschner, U. T. Bornscheuer, J. Altenbuchner, *Appl Microbiol Biotechnol* **2008**, *77*, 1251-1260.
[131] A. Riebel, M. J. Fink, M. D. Mihovilovic, M. W. Fraaije, *ChemCatChem* **2014**, *6*, 1112-1117.
[132] S. Bordewick, Bachelor thesis, University of Greifswald (Greifswald), **2013**.
[133] A. Beier, V. Hahn, U. T. Bornscheuer, F. Schauer, *AMB Express* **2014**, *4*, 75.
[134] R. J. Peroutka III, S. J. Orcutt, J. E. Strickler, T. R. Butt, *Heterologous Gene Expression in E. coli: Methods and Protocols* **2011**, 15-30.
[135] S. Diezmann, C. J. Cox, G. Schönian, R. J. Vilgalys, T. G. Mitchell, *Journal of clinical microbiology* **2004**, *42*, 5624-5635.
[136] a) C. Madzak, J.-M. Beckerich, in *Yarrowia lipolytica*, Springer, **2013**, pp. 1-76; b) G. Pignède, H.-J. Wang, F. Fudalej, M. Seman, C. Gaillardin, J.-M. Nicaud, *Applied and environmental microbiology* **2000**, *66*, 3283-3289.
[137] a) Y. Kawaguchi, H. Honda, J. Taniguchi-Morimura, S. Iwasaki, *Nature* **1989**, *341*, 164-166; b) T. Ohama, T. Suzuki, M. Mori, S. Osawa, T. Ueda, K. Watanabe, T. Nakase, *Nucleic acids research* **1993**, *21*, 4039-4045; c) G. Pesole, M. Lotti, L. Alberghina, C. Saccone, *Genetics* **1995**, *141*, 903; d) R. Rocha, P. J. B. Pereira, M. A. Santos, S. Macedo-Ribeiro, *Proceedings of the National Academy of Sciences* **2011**, *108*, 14091-14096; e) M. A. Santos, M. F. Tuite, *Nucleic acids research* **1995**, *23*, 1481-1486.
[138] J. G. Thomas, A. Ayling, F. Baneyx, *Applied biochemistry and biotechnology* **1997**, *66*, 197-238.
[139] a) J. Rehdorf, A. Kirschner, U. T. Bornscheuer, *Biotechnology letters* **2007**, *29*, 1393-1398; b) A. Kirschner, J. Altenbuchner, U. T. Bornscheuer, *Appl Microbiol Biotechnol* **2007**, *73*, 1065-1072.
[140] a) K. Nishihara, M. Kanemori, M. Kitagawa, H. Yanagi, T. Yura, *Applied and environmental microbiology* **1998**, *64*, 1694-1699; b) K. Nishihara, M. Kanemori, H. Yanagi, T. Yura, *Applied and environmental microbiology* **2000**, *66*, 884-889.
[141] W.-D. Fessner, V. Helaine, *Current Opinion in Biotechnology* **2001**, *12*, 574-586.
[142] a) V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66*, 12-21; b) I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall III, J. Snoeyink, J. S. Richardson, *Nucleic acids research* **2007**, *35*, W375-W383.
[143] J. I. Morton, B. V. Siegel, *Proceedings of the Society for Experimental Biology and Medicine* **1986**, *183*, 227-230.

[144] J. Kocsis, S. Harkaway, R. Snyder, *Annals of the New York Academy of Sciences* **1975**, *243*, 104-109.
[145] E. M. Debbi, G. Agar, G. Fichman, Y. B. Ziv, R. Kardosh, N. Halperin, A. Elbaz, Y. Beer, R. Debi, *BMC complementary and alternative medicine* **2011**, *11*, 50.
[146] A. Kirschner, U. T. Bornscheuer, *Appl Microbiol Biotechnol* **2008**, *81*, 465-472.
[147] a) V. Alphand, A. Archelas, R. Furstoss, *Tetrahedron Letters* **1989**, *30*, 3663-3664; b) Y. Zhang, U. Werling, W. Edelmann, *Nucleic acids research* **2012**, *40*, e) 55-e55.
[148] G. Hareau, P. Kocienski, P. Jankowski, S. Marczak, J. Wicha, in *Encyclopedia of Reagents for Organic Synthesis*, John Wiley & Sons, Ltd, **2001**.
[149] a) F. Forney, A. Markovetz, R. Kallio, *Journal of bacteriology* **1967**, *93*, 649-655; b) F. Forney, A. Markovetz, *Journal of bacteriology* **1968**, *96*, 1055-1064.
[150] a) S. Picataggio, T. Rohrer, K. Deanda, D. Lanning, R. Reynolds, J. Mielenz, L. D. Eirich, *Nature Biotechnology* **1992**, *10*, 894-898; b) M. S. Smit, M. M. Mokgoro, E. Setati, J.-M. Nicaud, *Biotechnology letters* **2005**, *27*, 859-864.
[151] R. Schoevaart, F. van Rantwijk, R. A. Sheldon, *J. Org. Chem.* **2000**, *65*, 6940-6943.
[152] M. Terasawa, J. Takahashi, *Agricultural and Biological Chemistry* **1981**, *45*, 2433-2441.
[153] a) J. Bruyn, *Koninkl. Ned. Acad. Wetenschap. Proc Ser C* **1954**, *57*, 41-44; b) M. J. Klug, Dissertation thesis, University of Iowa **1969**; c) M. J. Klug, A. J. Markovetz, *Journal of bacteriology* **1967**, *93*, 1847-1852; d) M. J. Klug, A. J. Markovetz, *Journal of bacteriology* **1968**, *96*, 1115-1123.
[154] M. Mizuno, Y. Shimojima, T. Iguchi, I. Takeda, S. Senoh, *Agricultural and Biological Chemistry* **1966**, *30*, 506-510.
[155] R. Stanier, *Journal of bacteriology* **1947**, *54*, 339.
[156] a) A. Dunker, R. R. Rueckert, *Journal of Biological Chemistry* **1969**, *244*, 5074-5080; b) A. L. Shapiro, J. V. Maizel Jr, *Analytical biochemistry* **1969**, *29*, 505-514; c) A. L. Shapiro, E. Viñuela, *Biochemical and biophysical research communications* **1967**, *28*, 815-820; d) K. Weber, M. Osborn, *Journal of Biological Chemistry* **1969**, *244*, 4406-4412.
[157] A. Matagne, B. Joris, J.-M. Frère, *Biochem. J* **1991**, *280*, 553-556.
[158] A. Rath, M. Glibowicka, V. G. Nadeau, G. Chen, C. M. Deber, *Proceedings of the National Academy of Sciences* **2009**, *106*, 1760-1765.
[159] X. Carbonell, A. Villaverde, *Biotechnology letters* **2002**, *24*, 1939-1944.
[160] A. Kirschner, PhD thesis, University Greifswald (Greifswald), **2007**.
[161] a) E. García-Fruitós, R. Sabate, N. S. de Groot, A. Villaverde, S. Ventura, *FEBS Journal* **2011**, *278*, 2419-2427; b) R. Sabate, N. S. de Groot, S. Ventura, *Cellular and molecular life sciences* **2010**, *67*, 2695-2715; c) H. Sørensen, K. Mortensen, *Microbial cell factories* **2005**, *4*, 1; d) H. P. Sørensen, K. K. Mortensen, *Journal of biotechnology* **2005**, *115*, 113-128.
[162] T. Kiefhaber, R. Rudolph, H.-H. Kohler, J. Buchner, *Nature Biotechnology* **1991**, *9*, 825-829.
[163] J. A. Chesshyre, A. R. Hipkiss, *Appl Microbiol Biotechnol* **1989**, *31*, 158-162.
[164] a) M. Ferrer, T. N. Chernikova, M. M. Yakimov, P. N. Golyshin, K. N. Timmis, *Nature Biotechnology* **2003**, *21*, 1266-1267; b) A. Mogk, M. P. Mayer, E. Deuerling, *Chembiochem* **2002**, *3*, 807-814.
[165] A. F. Holleman, E. Wiberg, N. Wiberg, *Lehrbuch der anorganischen Chemie*, Walter de Gruyter, **1995**.
[166] S. Bisagni, R. Hatti-Kaul, G. Mamo, *AMB Express* **2014**, *4*, 23.
[167] K. Balke, M. Bäumgen, U. Bornscheuer, *ChemBioChem* **2017**.
[168] T. Sugita, T. Nakase, *Systematic and Applied Microbiology* **1999**, *22*, 79-86.

[169] a) S. Aota, T. Gojobori, F. Ishibashi, T. Maruyama, T. Ikemura, *Nucleic Acids Res.* **1988**, *16*, 315-402; b) A. J. P. Brown, G. Bertram, P. J. F. Feldmann, M. W. Peggie, R. K. Swoboda, *Nucleic Acids Res.* **1991**, *19*, 4298.
[170] A. C. Gomes, I. Miranda, R. M. Silva, G. R. Moura, B. Thomas, A. Akoulitchev, M. Santos, *Genome Biol* **2007**, *8*, R206.
[171] I. Miranda, R. Rocha, M. C. Santos, D. D. Mateus, G. R. Moura, L. Carreto, M. A. S. Santos, *PLOS ONE* **2007**, *2*, e996.
[172] a) R. S. Donovan, C. W. Robinson, B. Glick, *Journal of Industrial Microbiology & Biotechnology* **1996**, *16*, 145-154; b) K. Terpe, *Appl Microbiol Biotechnol* **2006**, *72*, 211.
[173] H. Block, B. Maertens, A. Spriestersbach, N. Brinker, J. Kubicek, R. Fabis, J. Labahn, F. Schäfer, *Protein expression and purification* **2011**.
[174] D. E. Torres Pazmiño, B.-J. Baas, D. B. Janssen, M. W. Fraaije, *Biochemistry* **2008**, *47*, 4082-4093.
[175] a) P. C. Brzostowicz, K. L. Gibson, S. M. Thomas, M. S. Blasko, P. E. Rouvière, *Journal of bacteriology* **2000**, *182*, 4241-4248; b) M. Griffin, P. W. Trudgill, *European journal of biochemistry / FEBS* **1976**, *63*, 199-209.
[176] a) D. Minerdi, I. Zgrablic, S. J. Sadeghi, G. Gilardi, *Microbial biotechnology* **2012**, *5*, 700-716; b) M. W. Fraaije, N. M. Kamerbeek, A. J. Heidekamp, R. Fortin, D. B. Janssen, *Journal of Biological Chemistry* **2004**, *279*, 3354-3360.
[177] a) P. Usha, M. Naidu, *Clinical drug investigation* **2004**, *24*, 353-363; b) T. A. Pagonis, P. A. Givissis, A. C. Kritis, A. C. Christodoulou, *International Journal of Orthopaedics* **2014**, *1*, 19-24; c) S. Barmaki, S. Bohlooli, F. Khoshkhahesh, B. Nakhostin-Roohi, *Journal of Sports Medicine and Physical Fitness* **2012**, *52*, 170; d) E. Barrager, J. R. Veltmann Jr, A. G. Schauss, R. N. Schiller, *The Journal of Alternative & Complementary Medicine* **2002**, *8*, 167-173; e) D. S. Kalman, S. Feldman, A. R. Scheinberg, D. R. Krieger, R. J. Bloomer, *Journal of the International Society of Sports Nutrition* **2012**, *9*, 46; f) S. W. Jacob, *The miracle of MSM: the natural solution for pain*, Penguin, **1999**; g) F. Liebke, *MSM–eine Supersubstanz der Natur*, VAK Verlag, **2001**.
[178] A. S. Konagurthu, J. C. Whisstock, P. J. Stuckey, A. M. Lesk, *Proteins: Structure, Function, and Bioinformatics* **2006**, *64*, 559-574.
[179] S. Hornei, M. Kohler, H. Weide, *Zeitschrift fur allgemeine Mikrobiologie* **1972**, *12*, 19-27.
[180] J. Van der Walt, A. E. van Kerken, *Antonie van Leeuwenhoek* **1961**, *27*, 81-90.
[181] S. Luria, J. Adams, R. Ting, *Virology* **1960**, *12*, 348-390.
[182] J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular cloning, Vol. 2*, Cold spring harbor laboratory press New York, **1989**.
[183] U. Laemmli, *Nature* **1970**, *227*, 680-685.
[184] a) Untergrasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, R. SG, *Nucleic acids research* **2012**, *40*, e115; b) T. Koressaar, M. Remm, *Bioinformatics* **2007**, *23*, 1289-1291.
[185] D. Hanahan, *Journal of molecular biology* **1983**, *166*, 557-580.
[186] F. Okungbowa, A. Ghosh, R. Chowdhury, P. Chaudhuri, A. Basu, K. Pal, *World Journal of Medical Sciences* **2007**, *2*, 101-104.
[187] C. L. Ladner, J. Yang, R. J. Turner, R. A. Edwards, *Analytical biochemistry* **2004**, *326*, 13-20.
[188] D. A. Armbruster, T. Pry, *Clin Biochem Rev* **2008**, *29*, S49-52.
[189] E. Bermúdez, O. N. Ventura, L. A. Eriksson, P. Saenz-Méndez, *Computational biology and chemistry* **2014**, *49*, 14-22.

[190] A. Drummond, B. Ashton, S. Buxton, M. Cheung, A. Cooper, C. Duran, M. Field, J. Heled, M. Kearse, S. Markowitz, **2011**.

8 Appendix

8.1 Baeyer-Villiger monooxygenases participating in the metabolism of ketones in yeasts

8.1.1 Determination of metabolites from yeasts from 2-dodecanone and 1-dodecene

Scheme 8.1: Extracellular metabolites detected in yeast cultures containing 1-dodecene as a sole source of carbon and energy.
Carboxylic acids were detected as methyl esters.[133]

Table 8.1: Overview of extracellular compounds detected in culture media with 2-dodecanone or 1-dodecene as sole source of carbon and energy.

No.	Compound (MW [g/mol])	R_t [min]	Fragmentation m/z, relative intensity [%]
1	2-Dodecanone[1] (184)	15.836	39 [5], 41,05 [18], 41,95 [5], 43 [65], 55 [14], 56,05 [5], 57,05 [13], 57,95 [100], 59 [35], 67 [4], 69 [6], 70,05 [4], 70,95 [44], 72 [3], 81,05 [4], 82,1 [7], 83,05 [5], 84,05 [4], 85,05 [15], 95,05 [6], 96,05 [7], 97,05 [6], 98,05 [3], 113,05 [3], 124,05 [7], 126,1 [9], 127,1 [4], 169,1 [4], 184,2 [9]
2	1-dodecene[2] (168)	10.105	39,1 [26], 41,05 [86], 42,05 [22], 43,05 [71], 53,1 [11], 54,05 [16], 55,1 [100], 56,05 [77], 57,05 [55], 67,1 [18], 68,1 [15], 69,05 [82], 70,1 [73], 71,1 [23], 81,1 [6], 82,1 [17], 83,1 [65], 84,1 [41], 85,1 [12], 96,1 [7], 97,1 [49], 98,1 [20], 110,1 [3], 111,1 [16], 112,1 [7], 125,1 [6], 126,1 [3], 140,05 [4], 168,1 [9],
3	Dodecane (170)	10.264	39,1 [7], 41,1 [34], 42,1 [9], 43,1 [62], 55,1 [17], 56,1 [17], 57,1 [100], 58,1 [4], 69,1 [8], 70,1 [15], 71,1 [66], 72,1 [4], 83,1 [4], 84,1 [9], 85,1 [41], 86,1 [3], 98,1 [8], 99,1 [8], 112,1 [5], 113,1 [5], 127,1 [4], 170,2 [6]
4	Decyl acetate[1] (200)	16.379	39 [6], 41 [31], 42 [14], 43 [100], 44 [4], 53 [3], 54 [5], 55 [43], 56 [38], 57 [22], 58 [4], 61 [32], 67 [7], 68 [12], 69 [35], 70 [43], 71 [10], 73 [6], 81,1 [3], 82 [13], 83 [33], 84 [21], 85,1 [6], 96 [4],

			97,1 [22], 98 [12], 111 [12], 112,1 [13], 116 [4], 140,1 [4]
5	Decanol[1] (158)	11.860	39 [18], 40 [6], 41 [76], 42,1 [30], 43 [75], 44 [13], 45 [3], 53 [7], 54 [10], 55 [100], 56 [80], 57 [44], 67 [13], 68 [28], 69,1 [72], 70 [86], 71 [15], 73 [3], 81 [4], 82 [23], 83 [63], 84 [40], 85 [7], 96 [6], 97 [30], 98 [11], 111,1 [13], 112 [20], 125 [6]
6	2-Dodecanol[1] (186)	16.176	39 [4], 41 [19], 42 [6], 43 [23], 43,95 [6], 45 [100], 55 [23], 56 [12], 57 [21], 58 [3], 67 [3], 69 [19], 70 [12], 71 [11], 82 [4], 83 [17], 84,05 [9], 85,10 [6], 97,05 [19], 98,1 [6], 111,05 [8], 112,05 [3], 125,1 [3], 140 [4]
7	Decanoic acid[1] (186)	13.381	39 [4], 41 [14], 42 [4], 43 [17], 55 [17], 56 [3], 57 [6], 59 [10], 69 [8], 71 [3], 74 [100], 75 [10], 83 [4], 84 [3], 87 [54], 88 [5], 97 [3], 100,95 [7], 128,95 [4], 143,05 [15], 155,05 [8], 157,05 [3]
8	Hexanedioic acid[1] (174)	11.064	39 [12], 41 [24], 42 [17], 43 [31], 44 [3], 44,95 [7], 52,95 [5], 53,95 [8], 55 [80], 55,95 [10], 57 [4], 58 [7], 59 [100], 59,95 [4], 67,95 [4], 69 [8], 71 [5], 71,95 [8], 72,95 [30], 73,95 [43], 82 [12], 83 [24], 84 [4], 85 [5], 86,95 [13], 88,05 [3], 97 [4], 99 [3], 100,95 [65], 103 [3], 110 [3], 111 [60], 112 [5], 114 [82], 114,95 [16], 141,95 [11], 142,95 [46], 144,05 [4]
9	Octanedioic acid[1] (202)	17.628	39 [20], 39,95 [3], 41 [56], 42 [16], 43 [51], 44 [3], 45 [10], 53 [8], 54 [6], 55 [100], 55,95 [23], 57 [12], 57,9 [4], 58,95 [79], 67 [14], 67,95 [23], 69 [98], 69,95 [11], 71 [6], 71,95 [9], 73 [9], 73,95 [96], 74,95 [6], 81 [11], 82 [17], 83 [51], 83,95 [12], 85 [7], 86,95 [46], 87,95 [4], 93 [4], 96 [5], 96,95 [87], 97,95 [6], 98,9 [3], 100,95 [10], 110 [22], 111 [27], 112 [6], 112,95 [18], 114 [12], 116 [3], 126,9 [4], 127,95 [4], 129 [91], 129,95 [8], 136,9 [3], 138 [83], 139 [18], 140 [6], 140,95 [17], 142 [11], 170 [3], 171 [54], 172 [5]
10	Decanedioic acid[1] (230)	25.210	39 [12], 40 [3], 41 [44], 41,95 [15], 43 [37], 44 [3], 45 [7], 52,95 [6], 54 [5], 55 [100], 56 [12], 57 [8], 58,95 [55], 67 [13], 68 [6], 69 [38], 69,95 [6], 71 [4], 73,05 [10], 74 [93], 75 [6], 79 [7], 79,95 [5], 81 [14], 81,95 [6], 83 [39], 84 [46], 85 [6], 86,95 [34], 87,95 [4], 93 [4], 94 [4], 95 [9], 96 [12], 97 [48], 98 [65], 99 [5], 100,95 [9], 107 [5], 109 [3], 109,95 [8], 111 [10], 112 [6], 114,95 [8], 119,95 [3], 121 [10], 122,95 [3], 123,95 [9], 125 [60], 126 [5], 137 [3], 138 [35], 138,95 [13], 143,95 [4], 148 [5], 157,05 [29], 157,95 [3], 166 [25], 167 [3], 170 [4], 199,1 [36], 200,1 [4]

MS-data of *C. maltosa* cultures 1: with 2-dodecanone (**1**), 2: with 1-dodecene (**2**); carboxylic acids were detected as methyl esters due to a previous derivatisation

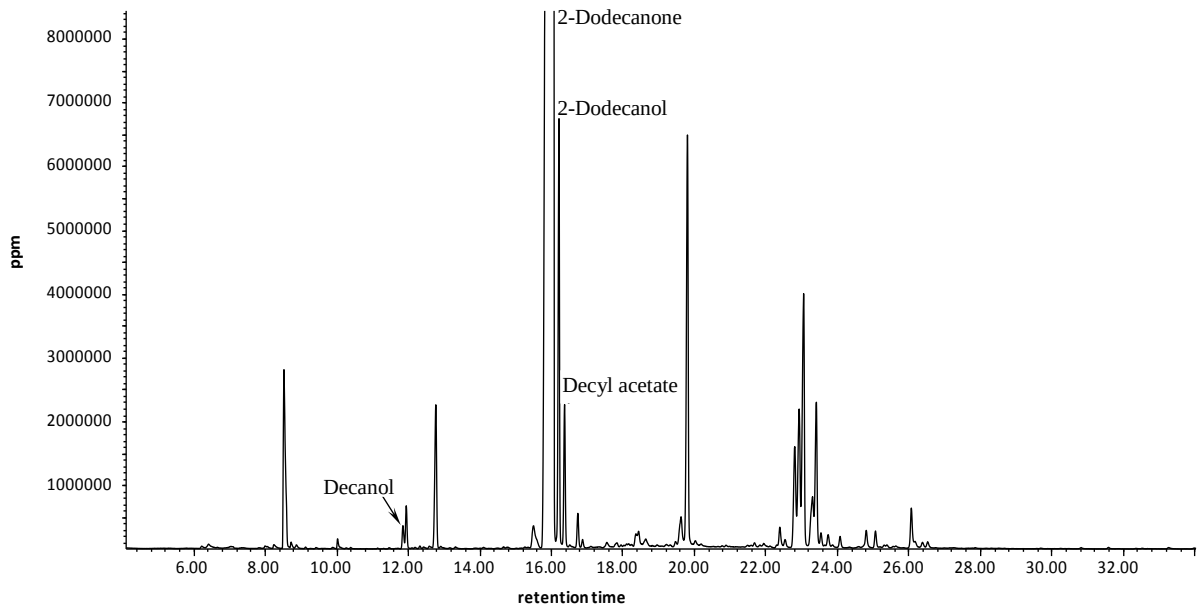


Figure 8.1: GC-chromatogram of an alkaline extract of a culture with *C. maltosa* and 2-dodecanone.

8.1.2 Investigations of novel BVMOs from yeasts

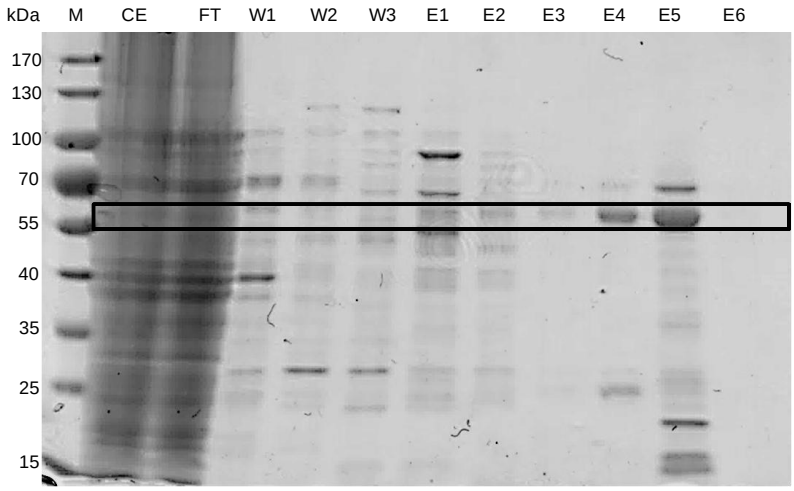


Figure 8.2: SDS-PAGE gel of the purification of $BVMO_{albi1}$.
CE: cell extract, FT: flow through, W1: washing fraction without imidazole, W2: washing fraction with 10 mM imidazole, W3: washing fraction with 20 mM imidazole, E1: first elution fraction with 50 mM imidazole, E2: second elution fraction with 50 mM imidazole, E3: third elution fraction with 50 mM imidazole, E4: elution fraction with 100 mM imidazole, E5: first elution fraction with 500 mM imidazole, E6: second elution fraction with 500 mM imidazole

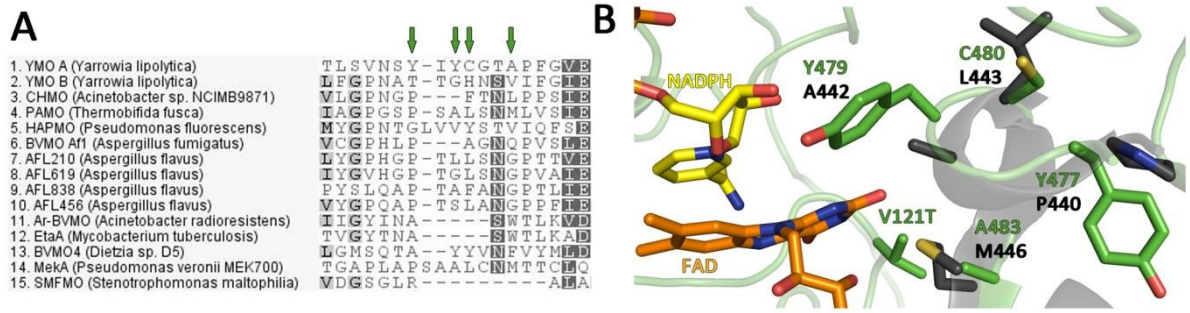


Figure 8.3: Multiple sequence alignment (left) and structural alignment (right) of Y477/Y479/C480/A483 in YMOA with PAMO. V121 was included for reference (from Bordewick).[69]
Orange: FAD; Yellow: NADPH; Green: YMOA; Black: PAMO (2YLR).

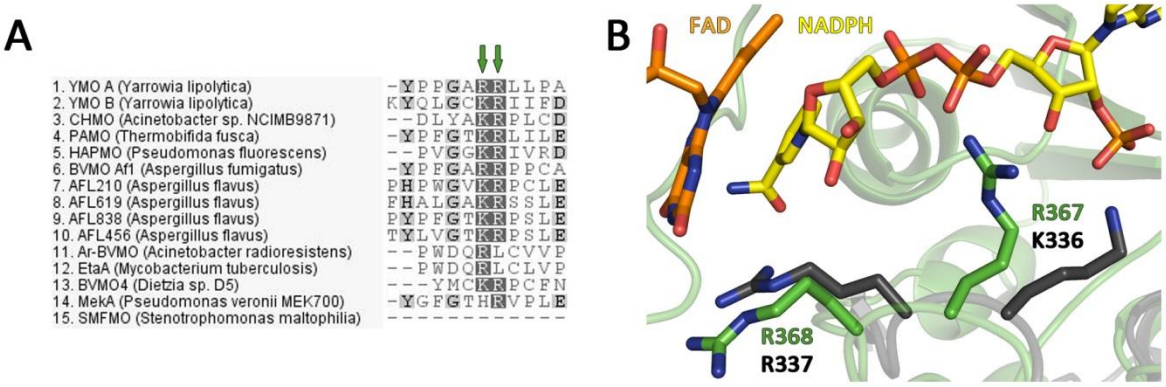


Figure 8.4: Multiple sequence alignment (left) and structural alignment (right) of R367/R368 in YMOA with PAMO (from Bordewick).[69]
Orange: FAD; Yellow: NADPH; Green: YMOA; Black: PAMO (2YLR).

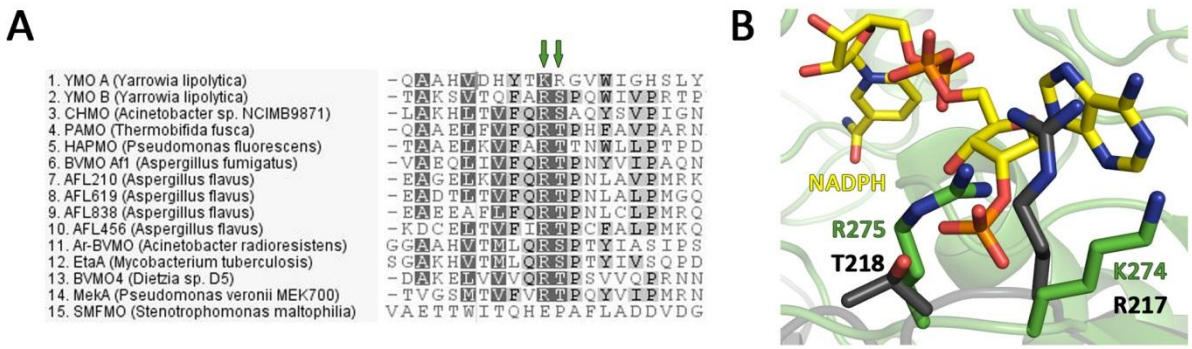


Figure 8.5: Multiple sequence alignment (left) and structural alignment (right) of K274/R275 in YMOA with PAMO (from Bordewick).[69]
Orange: FAD; Yellow: NADPH; Green: YMOA; Black: PAMO (2YLR).

Complete nucleotide- and amino acid sequences of the putative BVMOs/FMOs from *C. maltosa*, *C. albicans* and *Y. lipolytica* described in this thesis:

$BVMO_{malto}$:

Nucleotide sequence:

ATGCCAGTTATCACCTTAACTAAAGAATCGCATACTCATCCGGAGAAATTCCAAATCAAATAT-
GAAGATGCTGTCAAAATCTTAGGTCCTGATCGTAAAGATCGTTTCGTTTACAATGAATCAT-
TACCTACTCTCACCACCTCATCCACCGTGTCTATTGTCGGTGCTGGATTTGGAGG-
TATGGCAACTGCCATCAAGACCATTGAAGATTTACATGAAGAAGATGTTGTTATTTTTGAAAGACA
TGATAATTTCGGAGGTACTTGGTACGCTAACCATTTCCCAAATAGTGCTTCTGA-
TATCCCAAGTTTATGGTATTCTTATTCATTCCTGCCAGTTTCAAACTGGAGTAGAGTTCAAC-
CACCACAATATGAAATGGAAGAATATTTGTTGAGAGTGGCTGAAACTTATAAATTGAGA-
GAAAAAGCTAGATTCCAAACTGAAATTAATAAAGTTGAATGGATTGATGAGGACAATGTTTGGAA
ATTACATGCTCGTAACGTGGTTACCGGTCAAAAATGTATCCATACCAGTAAGATAT-
TAGTTGCTTGTCAAGGGGGTTTAGTCCATCCTTCCCAGTTGAAAGCTGAAGGATTA-
GAAAATTTCAAAGGTGAATATATGCATTCGGCAATTTGGAATGATGATGTTGATTTCAAAGG-
TAAGAAAGTTATCGTTGTTGGTAACGGATGTACTGCCAATCAAGTTGTCCCTGGTTTATTGAATAAT
CCAAAATATGGTGTCAAATCTTTGACCCAGATTGTTAGATCGAAACAT-
TATGTCCTGCCACCTGTCCCTAAAGCCCTTTTCTATTTGTATAAAT-
TACTTTCTTTTAATTTTTTTGGTTTGATTTTCGTTCGTTGGTTGGTTATTGCTATTGGTGAAT-
CAAGATTCCCATTGTTTAAAGGAGTTGGATTGATTAACAGATTTGTTCGTTGGGTTAATACCAGAGT
TGCGGTTAATTATATGAAGAAGAAAGCACCTAAAAAATTCCATGATTTGATTATTCCAGATTA-
CAAAATTGGTTGCAAAAGGTTAATTTTTGATCATGATTATATTCCAACGTTGAATGACCCAA-
GAATTGATCTTAAAGATAGTCCAATCGATAAAGTTGTGGAAAATGGTATTTTATTAAA-
GAGTGGAGAATTGATTGAAGCCGATATCATTGTTGCTTGTACTGGTTACAATGTGACTCAAAGTTTT
TTCAATTACGAAATCATTGGCCGTAACAAAACAAGCATCACTAAACTTTGGAAAGAAGATG-
GACCAAGTGCTTACAGAACATTATTGGTTAAAGAATGTCCAAACATGTGGAT-
GATTGCTGGTCCAAATTCAGCTACTGGTCACGCATCTGTGGTTATGGCTATT-
GAAAATGGGGTTGATTATTACACGAAAACCGCTAAACCAATTATACAAAACAAAGCTGCATCCGTT

AGAGTCAAGAATGAAGCTTATGATAATTGGTTTACAATTATTCAAGCGGAATTGAAAA-
GATGTGTTTTTGGTACACCTTTCGGAGGTTGTATTTCATGGTATGCTAATGAAAAAGT-
GAATGCCACTGCATACCCATGGAGTCAAAGTAATTACTGGTGGGTCACCCATCATCCAAATTA-
TAAAGATTTGGAGTATGATACAAAATCTATTAAAAAAGATTAG

Amino acid sequence:

MPVITLTKESHTHPEKFQIKYEDAVKILGPDRKDRFVYNESLPTLTTSSTVSIVGAGFGGMA-
TAIKTIEDLHEEDVVIFERHDNFGGTWYANHFPNSASDIPSLWY-
SYSFSPVSNWSRVQPPQYEMEEYLLRVAETYKLREKARFQTEINKVEWI-
DEDNVWKLHARNVVTGQKCIHTSKILVACQGGLVHPSQLKAEGLENFKGEYMHSAIWNDDVDFKGK
KVIVVGNGCTANQVVPGLLNNPKYGVKSLTQIVRSKHYVSPPVPKALFY-
LYKLLSFNFFGLIFVRWLVIAIGESRFPLFKGVGLINRFVRWVNTRVAV-
NYMKKKAPKKFHDLIIPDYKIGCKRLIFDHDYIPTLNDPRIDLKDSPIDKVVENGILLKSGE-
LIEADIIVACTGYNVTQSFFNYEIIGRNKTSITKLWKEDGPSAYRTLLVKECPNMWMIAGPNSATGHASV
VMAIENGVDYYTKTAKPIIQNKAASVRVKNEAYDNWFTIIQAELKRCVFGTPFGGCISWYA-
NEKVNATAYPWSQSNYWWVTHHPNYKDLEYDTKSIKKD

CMO:

Nucleotide sequence:

ATGCCATTTTTGACAGATTTGGATTATGAAAATCCAGTAACTATTTCATCCCAACAAA-
TAAAAACTATTGCTATTATTGGAGGTGGTGCTTCAGGAGCAATTATTTTAGATAGTTTACT-
TAAAGAACCTTCTGGTATTGAAAAAATTGTTATCTTTGAAAGACAGAAT-
GAATTGGGTGGGGTATGGTTTTTCAATAAAGATATCAGATCAACACCAAATGAGTTGATTAAATCA
GGTAATTCTCATCTTGATAATGATCCACAATTACCTAACCCATTTCATGACCACACCGATAAA-
GATAAACTTGTTTTACCGAAAAACAATCAAGAACGGTTTATACAAACTCCAAGTTAT-
TATGGGATCAAAACTAATATCATTGAAAACATGATGACATACAGTGATAACAAAA-
GATGGGAAGTTGAAGGAGATGAAGAACAAAGAAAGTATGTTGAAGGAAGTGTCGTGCAGAAATA
CATTGAGAAATATATCAACAAAAACTTGGATGATCCAAGGGTTGATTTAAGATTAAACTCGA-
CAGTTGAAGATGTTGAAAGAATTGACCGTGATGACGATGATGCTGAATTACCATACAGATT-
TAAAGTTACTGTTCGTACTCCACACGATGATCACAATGATGCATGGTACCAAGAA-
GATTTCGATTCTATTGTTGTTGCTACTGGCCATTACCATGTTCCACATATTCCTCATGTACCGGGTTT
GAAAAAAATACAAGAAACTTTCCCGGAAAAAGTCCAACATGCCAAATTTTATAGAGAAT-
CAAGTCAATACAAGGGGAAGAAAGTGGTTGTGGTTGGGTCTCGAG-
CATCTGGTGCTGATTTGACTAAATTTGTGGCTAGAGAACCTGGAACTACCGTCTACCAAT-
CAATTAGAAATTATGAAAACACCAAAGTTTTATCTGCTCAAACCAATGTTTTCAAAAAACCGGCCA
TTAAAAATTATGAAATTGTCAATGATAAGGTTAAAGTGATTTTCGAAGATGATTCTGTTATT-
GAAGATCCAGATGTTATTATTTATTGTACTGGATATTTGTTTTCCTATCCGTATTTGAA-
CAGGTTGACAAATCATAAAATCACCGAAGGGATAACCATTCCAAACTTATACCAACA-
CACTTTCCTTATCAATGAACCATTGATTACCATAATCGGAGTCCCAATCGATGGTATTTCATTTAGA
GTTTTTGAATACCAAGCGGTGTTGGTCACTCGATACTTGACCGGGAAAATTAGTTTACCAT-
CAAGAAAGGAACAAAGTGAATGGGTTAATAAACGATACGAAGAAAAGAAAAGCACGA-
GAGCGTTCCACACCATTGGCGTTATTGATGCGTTTGATTATTCCAATGGTTTAGTCAATT-

TAGGTCAAGTTCTGGAAAAAATTAAAGTGGGTAGAGAATTCCCAAAAATAACTGCTGAAGAGA-
TAAAGGTTTATAGAGAAGCTGGTGAGAAGTTACGTAAATTTTGGGATGAGAGATAA

Amino acid sequence:

MPFLTDLDYENPVTISSQQIKTIAIIGGGASGAIILDSLLKEPSGIEKIVIFERQ-
NELGGVWFFNKDIRSTPNELIKSGNSHLDNDPQLPNPFHDHTDKDKLVLPKNNQER-
FIQTPSYYGIKTNIIENMMTYSDNKRWEVEGDEEQRKYVEGSVVQKYIE-
KYINKNLDDPRVDLRLNSTVEDVERIDRDDDDAELPYRFKVTVRTPHDDHNDAWYQEDFDSIVVATGH
YHVPHIPHVPGLKKIQETFPEKVQHAKFYRESSQYKGKKVVVVGSRASGADLTKFVA-
REPGTTVYQSIRNYENTKVLSAQTNVFKKPAIKNYEIVNDKVKVIFEDDSVIEDPDVIIYCT-
GYLFSYPYLNRLTNHKITEGITIPNLYQHTFLINEPLITIIGVPIDGISFRVFEY-
QAVLVTRYLTGKISLPSRKEQSEWVNKRYEEKKSTRAFHTIGVIDAFDYSNGLVNLGQVSEKIKVGREF
PKITAEEIKVYREAGEKLRKFWDER

$BVMO_{albi1}$:

Nucleotide sequence:

ATGTCTGTCATTACATTAACAAAAGAGTCTCACAAGAATCCAGAAAAGTTTAGAATTAAACAT-
CAGGATGAGGTGGAGATATTAGGTCCACATCGTAAAGACCGTTTTGCCATCAATGAA-
GACTTGCCAACCATAACCACCACTTCTAAAATTGCCATTCTCGGAGCCGGTTTTGGAGG-
TATGGCAAGTGCAATCAAGACAATGCAAAAATACAATGAGCAAGATATTAAAATTTTTGAAAGAC
ATGACAACTTTGGTGGTACTTGGTATGCCAACACTTACCCAGGATGTGCCAGTGA-
TATTCCCGCTTTATGGTATTCATTTTCGTTTGCATTGACATCCAACTGGAGTAGAGTTCAAC-
CACCACAGTATGAGATGGAAGAATACATTTTACGAGTTGCCGAACAATTCAAATTAAGAGA-
GAAAACTAGATTTCAAACTGAAATCAACAAGTTTGAGTGGGATGATGTGAATGGTGAGTGGACCTT
GTATGCACATGATGTTAAAACTGGTCAAAGAATCCTCCACAAAAG-
TAAGCTTTTACTTGCCTGTCAAGGTGGGTTAGTTCATCCTTTGCAATTACAAGCCGAAG-
GATTGGAAAACTTCAAAGGGGCATACATGCACTCGGCTCTTTGGGAT-
CATTCTGTTGACTTCAAAGGGAAAAAAGTCATTGTGATTGGTAATGGGTGTAGTGCTAATCAAACT
GTTCCTGCGTTACTCAACAACCCTGATTACAGTGTCGGTTCGTTGACTCAGATTTCAAGATC-
CAAGCATTACATTTTGAAACCCCTCCCTAGAATACTTTACATACTTTACCGTTTATTGTCATT-
CAACTTTATTGCATTATACTTTGTTCGTTTAATTGTTGTTTTTGGTGCTGAAATGAGGGTAC-
CATTGTTCAAAGGTGATGGGTTTATCTCCAAAATTGTTCGTTGGATAAACACAACTGCTTCCGTTAG
CTATATGAAAGGTAATGCTCCTGAGAAATTCCATGATATGATTATTCCTAATTA-
CAAAATTGGATGTAAAAGATTAATTTTTGACTATAATTATATTCCATCGTTAAATGACCCAA-
GAGTTGACATCAAGAATCAAGGAATTGATAGAGTTGTTGAAAATGGAATATTATTGAAAAATG-
GAGAACACATTGAAGCTGATATTATTGTTGCATGTACTGGGTATAATTTGAGCAAAAGTTACTTTA
ATTTTGAAATTGTTGGTCGTAATGGAGCAAATATCTCTGAAGAATGGAAGAAA-
GATGGTCCAAGTGCTTATAGAACGCTTTTAGTCAAACAAAGTCCTAATCTTTGGA-
CAATTGGGGGTACCAATTCAGCTACTGGACATGCATCTGTTGTCATGGCAATTGA-
GAATGGTGTTGATTATTTCCTCAAGACTGCCAAACCAATCATTGAAGGAAAAGCTAAATCAGTTAG
AGTTACCGACGAAGCTTACGACAACTGGCTTACAACTATTCAAAAAGAATT-
GAAAAAATCTGTCTTTGGTACTCCATTCGGAGGTTGTGTTTCCTGGTATTCTGATG-

CAAAGGTCAATTCAACTGTCTACCCTTGGAGTCAATTTCATTATTGGTGGATTACA-
CATTTCCCAAATTATAAAGATTTAGTATATGAGCCATTAAACGAAGACAAAAAGAGGAGATGA

Amino acid sequence:

MSVITLTKESHKNPEKFRIKHQDEVEILGPHRKDRFAINEDLPTITTTSKIAILGAGFGGMA-
SAIKTMQKYNEQDIKIFERHDNFGGTWYANTYPGCASDIPAL-
WYSFSFALTSNWSRVQPPQYEMEEYILRVAEQFKLREKTRFQTEINKFEWDDVNGEW-
TLYAHDVKTGQRILHKSKLLLACQGGLVHPLQLQAEGLENFKGAYMHSALWDHSVDFKGKKVIVIGN
GCSANQTVPALLNNPDYSVGSLTQISRSKHYILKPLPRILYILYRLLSFNFIALYF-
VRLIVVFGAEMRVPLFKGDGFISKIVRWINTTASVSYMKGNAPEKFHDMIIPNYKIGCKRLIF-
DYNYIPSLNDPRVDIKNQGIDRVVENGILLKNGEHIEADIIVACTGYNLSKSYFN-
FEIVGRNGANISEEWKKDGPSAYRTLLVKQSPNLWTIGGTNSATGHASVVMAIENGVDYFLKTAKPIIE
GKAKSVRVTDEAYDNWLTTIQKELKKSVFGTPFGGCVSWYSDAKVNSTVYPWSQFHYW-
WITHFPNYKDLVYEPLNEDKKRR

YMOA:

Nucleotide sequence:

ATGACTATCTCCAAGCCTCCATCGCCAAACAACCTCGACAATGCCGGAATCACCTCCTCCAGC-
CAGGCCGGTAGGGGACATACCAACGTGACCGGCGTGGACAAAGAAGCCCTCTGGAAC-
GAGTTTGACTTTCTCAAAAACCTCGAGCCTCCAAGTGAATGGGCCGAAACCATTCTCAACCGA-
GAGTACCACGGCCGACGGCCCGTCAAGGTGGTCATTTCGGGCGCTGGGCTGTCTGGAATCACCACC
GGTATCCTCATCAACGGCAAAGTGGACGATGTCGATCTGACCATCCTGGAGCGAAACGAA-
GAGGCCGGAGGAGTGTGGTTCAAGAACACT-
TACCCGGGCGTGCGATGCGATGTTCCCTCCCACTCGTACCAACTCTCCTTCGATCCCAAAACA-
GACTGGAAGAGCGTCTATGCCTACGGAGAAGACATCAAAAAGTACTGGCAGAGTCGAGCTGAAAA
GTACGGCATCTCAGATAAGATCAAAACCCAACAGAACATCCTCGAAGCTAAATGGGACCAG-
GAGGACGGCCAATGGCACATTCTGGTAGAAGATCTCACCAAACCCCATCAGGATCAGTACA-
CAGTCAAGGCCGACTTCTTCATCTCGTCCTCAGGAACCCTCAACCAGCCCAGATACCCTCCCA-
CACAGCCCGGATACGACAAGTTCAAGGGCGAAAAGTTCCACCCAGTCAATTGGCCCAAGGGACTT
TCTCTTGAGGGCAAACGGGTGGCTTTGATCGGAAATGGAGCCACTGGTGTA-
CAGCTCCTCCCCCAGATCGCACTCCAGGCTGCTCATGTTGATCACTACACCAAAC-
GAGGTGTTTGGATCGGCCATTCTCTATATGGATCTCGAGTTCCTGGATACGTGGATTACACG-
CAGGAAGAGATTGACGAGATTCAACAGTCTTCCGAGTACCACAAGTTCCGAAAACAGCTGGACGA
GGCTCTTTTGGGTAACTACGGCGGCTCCTTCTTTGGAACAGAATCCTATAAGGGCCTCA-
TAAAGGAGCTACTGGCAATCATGTTTATTCGAGTGGGTAAGGACCTGGAGCTGTTCAA-
GAAGGTGGTACCTAATTACCCCCCTGGAGCCCGACGTCTTCTTCCTGCTCCCGGA-
TATCTCGAGGCTCTGACTCGAGAAAATGTTTCTTATCATCTTGGAGACATTCAGGAGTTTACTGAAA
AGGGTATCATTGGTCCTGATGGCGTAGAACGAGAAGTTGATGTCAT-
TATTGCCTCTACTGGCTATGTTCGAGATGATGGAGCTGGAGTGACCCCCAACTACGAAATC-
TACGGCCAGGAAGGATACACTCTTCGCCAGCATTTCAACCCTCCCGAATCCAAGCTTGGA-
TATTCCGCGTGCTACCTTGGACTGGCAGCTCCTCACTTTCCCAACTTCTTCTACACACTGTCCGTCA
ACTCTTACATTTACTGTGGAACAGCTCCGTTTGGAGTCGAGTTGCAGGCCACATA-

CATCGCCAAGGCTATTCGAAAGGCCCAGCTGGAAGACAT-
CAAGTCGCTTGTTCCCTCGGTGCGAGCCTCTGTGCTCTTCAACCGACGAATCAAT-
GAGTTCTCCAAGACGTCATGTGTGTGTCGAGGCATTGATGGATACTACACTGAGCGAGACACT-
GAGGGCAATGTGCGTCTCAAGGGTTCTTGGCCCGGAACCATGACCCATGCTCTGAGTATGCTGCGA
GAGCCTCGATGGGAAGATTACGACTACGAGTACCTCAATCCTGAT-
GATCCCTTCTCGTACTTTGGATCTGGAAAGACGTGGATTGATGACCATGATGGAGACAA-
GACTTTCTACCTCACCGAGCCCGGTAAGGTGTCTGTTCGTAACGTCCAT-
GAGGGCTGGGTGTCATTGTCTCGGCATCATGCCCCCAACTGTTCTCACAATGCTGACGAACATATT
GAAGATGGACCCAAGGCCAATGGGCATGTTAATGGTCTCAAATCGAAGGTTAATGGGGTCTAG

Amino acid sequence:

MTISKPPSPNNLDNAGITSSSQAGRGHTNVTGVDKEALWNEFDFLKNLEPPSEWAETIL-
NREYHGRRPVKVVISGAGLSGITTGILINGKVDDVDLTILER-
NEEAGGVWFKNTYPGVRCDVPSHSYQLSFDPKTDWKSVYAYGEDIKKYWQSRAEKYGISD-
KIKTQQNILEAKWDQEDGQWHILVEDLTKPHQDQYTVKADFFISSSGTLNQPRYPPTQPGYDKFKGEKF
HPVNWPKGLSLEGKRVALIGNGATGVQLLPQIALQAAHVDHYTKRGVWIGHSLYGSRVPGYV-
DYTQEEIDEIQQSSEYHKFRKQLDEALLGNYGGSFFGTESYKGLIKELLAIM-
FIRVGKDLELFKKVVPNYPPGARRLLPAPGYLEALTRENVSYHLGDIQEFTEKGIIGPDGVE-
REVDVIIASTGYVRDDGAGVTPNYEIYGQEGYTLRQHFNPPESKLGYSACYLGLAAPHFPNFFYTLSVN
SYIYCGTAPFGVELQATYIAKAIRKAQLEDIKSLVPSVRASVLFNRRINEFSKTSCVCRGID-
GYYTERDTEGNVRLKGSWPGTMTHALSMLREPRWEDY-
DYEYLNPDDPFSYFGSGKTWIDDHDGDKTFYLTEPGKVSVRNVHEGWVSLSRHHAPNCSHNA-
DEHIEDGPKANGHVNGLKSKVNG

YMOB:

Nucleotide sequence:

ATGAAACCGTTATATATAACGGCGGATATATTTGATTTTATTACATCTTACACCAACATC-
GATCTTTATCGACAACCAAACACCATGACTAAGCTTCATTCTCAAGTTCTTATTGTAG-
GAGGGGGGTTCTCAGGAATAGCCACCTCCATCAAACTGCTCAAAGACTGGAAGGTGACC-
GATTTCCATGTCTACGACCGTAACGAGAAGTTTGGAGGCACTTGGGCTGCCAACACTTACCCCGGA
GCTGCCTCCGACATCCCTGCAGTCTGGTACTGTCTAGCCAGTGATCCCAAGATC-
GATTGGGAGTCTGCCTATCCATCGCAGCAGGAGCTGTCTGAATACATTGCAGGAGTTGTCGA-
CAAATACGGTCTCAAGTCATTTGCAACCTTCAATTCTGAGATCGAGCGAATCGAATG-
GATTCCTAACGAGCGCTTGTGGAAGGCTACCATCGCTCACAATGGCAACACCATCACACACACTGC
TCGAGTTCTCTTCATGGGCCAGGGATGCCTGGTAACACCTAACCATGTCAAGAT-
CAAGGGCATGGAAGACTTTCAGGGTCCCATTATGCATACAGCCGAGTGGAAGCCGTTTGAT-
TACGACAACAAGGACGTTGTTGTTATTGGAAACGGATGTTCTGCTGCTCAGGTCACCAGT-
GAGGTGGCCAAAACAGCGAAGTCAGTCACTCAGTTCGCCCGTTCTCCACAGTGGATTGTTCCACGG
ACTCCAGTCACTATTGGGCCTATTTTGCGGAC-
CATTCTCAGTTGGTTCCCGTTCTTGATCCCTGTCCTGCGATTCTTCGTATTTTGCTTACTGGAAATGA
ATTGGAATATGTTCCGAGGAGGCTGGTGGTCCGACTTCGATCGGTCTATGCGAAC-
CAAGGTGGCTGTGAAAATTGCCAAGAAGAACATGCCTGAAAAGTACCACGAAACAGC-

CATCCCCAAGTATCAGCTCGGATGTAAACGAAT-
CATCTTCGATTGCGGCTACTGGGCAGCCTTGAAAATGGAGAGTGTGCTTCTGACCTTTGA-
CAAGCTGGTTGAAGTTGGAAAGAACTCAGTCAAGGACATCAATGGCAACCAG-
TATCCGGCTGATCTGATTGTGGACGCTACGGGGTTCAACATTGGACGATCAATGACCTCTGTGGAT
GTGATTGGAGAGAACGGAATGCCTTTATCTGAATTCTGGGATGGCAAGGTCGCTGCGTACGA-
GACGGTCATGGTGCCTAACTACCCCAACATGTTCATGCTGTTCGGTCCCAATGCCAC-
GACTGGACATAACTCAGTTATCTTCGGTATTGA-
GAACGGTCTCAAGTTTGTGGAGTCGGTTGCTTCCGATGTCATCCAGGGACGAAGTGACTACGTGAC
AGTGAAGCCCCAGGCATACGACCAGTGGGTGCAGCGAATCCAGGCTGCTATCAAACAGAC-
CAACTTTGCCACGGGTGGGTGCGTGTCGTGGTACATGTCTGTTGGAGAAGCCACTCA-
CAATGCCGTTTCCTATCCTTGGACCCAGCTGAGATTCTGGTGGAGAGCTCGATTCCCCCAT-
TACGACGACATTTATGTTGAAAACAAATCTCAGAAGGTTACTCCAGGCAAAGCGATCAAGGCCGA
GTAG

Amino acid sequence:

MKPLYITADIFDFITSYTNIDLYRQPNTMTKLHSQVLIVGGGFSGIAT-
SIKLLKDWKVTDFHVYDRNEKFGGTWAANTYPGAASDIPAVWYCLASDPKIDWE-
SAYPSQQELSEYIAGVVDKYGLKSFATFNSEIERIEWIPNERLWKATIAHNGNTITH-
TARVLFMGQGCLVTPNHVKIKGMEDFQGPIMHTAEWKPFDYDNKDVVVIGNGCSAAQVTSEVAKTAK
SVTQFARSPQWIVPRTPVTIGPILR-
TILSWFPFLIPVLRFFVFCLLEMNWNMFRGGWWSDFDRSMRTKVAVKIAKKNMPEKYHE-
TAIPKYQLGCKRIIFDCGYWAALKMESVLLTFDKLVEVGKNSVKDINGNQYPADLIV-
DATGFNIGRSMTSVDVIGENGMPLSEFWDGKVAAYETVMVPNYPNMFMLFGPNATTGHNSVIFGIENG
LKFVESVASDVIQGRSDYVTVKPQAYDQWVQRIQAAIKQTNFATGGCVSWYMSVGEATHNAV-
SYPWTQLRFWWRARFPHYDDIYVENKSQKVTPGKAIKAE

YMOC:

Nucleotide sequence:

ATGCCCTCAATTGATCCTTCCAACGTGAAACTCGGAGTGGGCCACACGGACA-
CAACGGGGGTGGATAAGGAGGCGCTGATTGCGCAGTTTGAGCAGTTCAA-
CAACCTGGAGGCCAGCGAATGGGCCAAGAAGATCCTGGAGCGGCCATATCTCGGTA-
GACGTGCCGTCAGCGTCATCATCTCAGGCGCTGGTCTCGCCGGAATCACCACAGCTATCCTGCTGT
CTCAGAAAGTCGACAATTTGACTCTGACTGTTCTGGATC-
GAAACTCCAAGGTCGGAGGTGTTTGGGCTACAAACGAGTATC-
CAGGCGTGCGATGTGACGTTCCCAGTCACTCATACCAGCTTACTTTTGATCCCAAAACA-
GACTGGCAGAGCGCCTATGCCTTCGGAAAGGATATCCAGGCTTACTGGCAGAGCAGAGTGGAGAA
ACACGGGCTTGCAGACAAGTTCAGACTGAACCATTCCATCAAGGAGGCCAAATGGGATGAAC-
CAACTCACCAATGGCATGTCAGAGTGGAACACAACGGAAAAGAG-
GAGGTTCTCAAGTCGGACTTTTTCATTTCGTCATCTGGATCTCTCCA-
GACTGCCAAGTTTCCCGTCCAACCCGGCTTCGAGTCCTTCAAGGGACCCAAATTCCACCCTGTCAA
CTGGCCTAAAGATCTGAACCTCAAGGGAAAGCGGGTGGCATTGCTGGGGAATGGAGCCA-
CAGGGGTCCAGATTCTGCCAGAGCTCATCAAACAGGGTGCAGCTCACGTGGACCAC-

TACGTCAAAAGAGGAGCTTGGATAGGCCACACAC-
TATTCGGCGTGAAAGCTCCGGGTTACGTCGACTACACTCAGGAAGAAATCAACGCCATC-
CAGTCCTCTGAAGAGTACCACAAGTACCGTTCTTCTCTTGACTCCAAGCTGCATGGCAAGTAC-
GAAGCCACTTTGTTTGGAACTCCGGGCTACAAGGCGGGCATCAAGGAACTGCTGGCGCTGATGTAC
CTTCGGGTCGGCGAAAACGACGAACTGTTTGAAAAGGTTGTGCCTCATTTCACTCCGGGACC-
CAGACGTCTCCTTCCAGCTCCTGGGTATCTCGAGGCTCTAGCTCTTCCCAACGTGGACTACTA-
CAAGGGCGACATTGAGCGGTTCACTGAGAATGGTATCGTATTTGACGGAGAA-
GAACGTCCAGTTGATGTCATCATTTCTTCGACAGGATACGTTCGAGGTAACGGATATGGAGCCACT
CCCAATTATGAAATCATCGGATCCGACGGGTACACTCTACGAACA-
CATTTCTCACCTCTTGAGTCCAA-
GAAGGGCTACTCTCTGTCCTACCTTGGAGTTTCGGCGCCCGGCTTTCCCAACTTTTTCTA-
CACCCTCTCGGTAAACTCCTATCTCTACTGTGGAACTCCTCCCATAACAGTAGAGCAGCAGTCTAGC
TACATTGCCAAGGTGATCCGAAAGGCGCAGTTTGAAGACATTGCTTCCATTGATCCGAAGGA-
GAAACCCTCCGAGTCCTTTTCCAGACGTATCTGGGAACTGTCCCAGGCCTCGTCAATCAC-
CAAGGGCGGAATTGGAGGTTACTTTACCGAGATTGATCGCAACGGTGA-
CACCCGGGTGCGAATTTCGTGGCCTGGAACTATCTCTCATGCCATTTCTGTTCTACGAGAACCCAGA
TGGGAAGATTTCAACTACCAATATCTCAACCCAGACGATCGCTTTGGA-
TACTGGGGTAACGGCAAGACCTGGATTGATGATCATCCCGGAGACAAGA-
CATTCTATCTTTCCAAGCCGGGCTCTGTCAAAGTAAGGAATCTACAT-
GAAGGCTGGATTTCGCTTCCCAGAGATGGTCCTCCCGAGATGGTTCCCCTCGGGGTTGACATCAGG
GCGTAA

Amino acid sequence:

MPSIDPSNVKLGVGHTDTTGVDKEALIAQFEQFNNLEASEWAKKILERPYLGRRAVSVIIS-
GAGLAGITTAILLSQKVDNLTLTVLDRNSKVGGVWAT-
NEYPGVRCDVPSHSYQLTFDPKTDWQSAYAFGKDIQAYWQSRVEKHGLADKFRLNHSI-
KEAKWDEPTHQWHVRVEHNGKEEVLKSDFFISSSGSLQTAKFPVQPGFESFKGPKFHPVNWPKDLNLK
GKRVALLGNGATGVQILPELIKQGAAHVDHYVKRGAWIGHTLFGVKAPGYVDYTQEEI-
NAIQSSEEYHKYRSSLDSKLHGKYEATLFGTPGYKAGIKELLALMYLRVGENDEL-
FEKVVPHFTPGPRRLLPAPGYLEALALPNVDYYKGDIERFTEN-
GIVFDGEERPVDVIISSTGYVRGNGYGATPNYEIIGSDGYTLRTHFSPLESKKGYSLSYLGVSAPGFPNFF
YTLSVNSYLYCGTPPITVEQQSSYIAKVIRKAQFEDIASIDPKEKPSESFSRRIWELSQAS-
SITKGGIGGYFTEIDRNGDTRVRISWPGTISHAISVLREPRWEDFNY-
QYLNPDDRFGYWGNGKTWIDDHPGDKTFYLSKPGSVKVRNLHEGWISLPRDGPPEMVPLGVDIRA

YMOD:

Nucleotide sequence:

ATGCTGGCTGTCTACATATACTCCTTGTGCTACGCCCTCTTACCATTCACCTTTCA-
CAGTGCAACTACTAACATGTCTTCCAAAAACGGAGTTGGCCACACAGACACCTCGTCCGTCGA-
CAAGGAAGAACTCATCTCTCAATTCCCCCATTTGGACCGTCTCGACCCGTCTCCATGGGCTCA-
GAAGATCCTTTCTCGTCCTCCTCTCGGCCGTGAGAAGATCAGTGTTGTTCTCTCGGGGGCAGGTCTT
GCTGGTATTTCAACCGGAATCATCCTTTCTCAAAAGGTAGACAACATTGACCTCAC-

CATTCTGGAGCGGTCTCCGGACTTTGGGGGCGTTTGGTTCGACAACAGCTACC-
CAGGTGTCCAATGCGATGTTCCTGTTCACGCCTACCAAT-
TATCTTTTGACCCCAAGCGCGACTGGGACAGACCGTACGCCTACGGAAACGATATCAAGCAG-
TATTGGGGTGACAGGGCAAAGAAGTATCAACTAAACGAGAAGACCAGGTTTGGACACAACATTCT
AGAGGCCAAGTTCAACAAGAATACCAGTCAGTGGGTGATTCAGGTGGAAACCGTGGCCGATAA-
GAAACGTTCAGAGATTCGTGCTGACGTATTTATCGCCACCAGTGGAGCTCTTAA-
CAACCCCAGGTATCCTCCTACTCAACCGGGATTCGACTCCTTCCAGGGGAT-
CAAGTTTCATCCTCAACAGTGGCCTGAGGGACTGGACTTGACCGGCAAACGAGTGGCTCTGATAGG
AAACGGAGCCACTGGAGTCCAGATCTTGCCTCAGATTGTCGAACAAGCGGCCCACGTGGACCAC-
CAC-
TACGCCAAGTCCTCTTCTTGGATTGGTCATGCTCTGTTTGGACCCGGAGTCCCCGGCTACGTGGAGT
ACTCGAGGGACGACATTGAGTCTATCAAGTCCGACAAAGACTACTTGGAGTTCCAGAAG-
GAGTTGCACAGAAACATTGGAGGCAAGTATGATTTTTTTTTTTACGGAACTCCTGCGTTTAGA-
GAACTCACCAAGGAACTACTGGCTGTGGCGTGGATTCGGGTTGGCAAA-
GACCCCAAACTGTTCCGGAAAGTGGTACCCACGTACCCCTTTGGAGCCAAACGACTTCTACCTGCC
CCTGGATACCTCGAGGCTCTCACCCGACCAAATGTTGACTACCTCCTCGGTGACGTGAAG-
GAGTTTACCAAGAACGGTATCATTGGAGCTGATGGTGTTGAAAGACAAGTGGATGTTAT-
TATTGCCGCCACGGGCTATCCTCTGACCAACGGAAATGGATTCACCCCCAACTATGAAAT-
TATTGGCACCGATGGGTACAGTTTGAGACAACACTTCTCCCCTCTGGAGTCCAGACTTGGCTATTCA
GCATCCTACCTCGGTCTAGCTGCCCCGGGGTTCCCCAACTTCTTCTA-
CACCCTTTCTGTCAACTCGTACATCACCAAGAGCACTCCTGCTGA-
GACTGTGGAGCTGCAAGCTGCTTATATTGCCAGAGCTATCCGAAAGAAACAGCTTGAGAAAAT-
CAAGTCATTGGAGCCGTCCCTCAAGGCGACAGTGTCGTTCAACAGGAGAATCACTGAGCTGTCTAA
AGCGATTTCGGTCACTAAAGGCAACGGATTCTTCAACGAAGTAACCAAAGACGGTACTAAAC-
GATCCAAGGTCGACTGGCCCGGTTCCGTGTCCCATGCTATTGCCGTGCTTCGAGAACCAC-
GATGGGAGGACTTTGAGTATAGGTATGAAGACAATGAC-
GATCCGTTTGCCTACTTTGGGTCTGGCAAGACGTGGATCGACGATCATGATGGTGATAAGACGTTT
TACATCACTCAGTCAGCTACAGTAGCTGCCAAAGTTCACGAAGGCTGGATTTCTTTGCCTTCA-
GATGGGCCCCCCAGGGTGGTTGCAAGATGA

Amino acid sequence:

MLAVYIYSLCYALLPFTFHSATTNMSSKNGVGHTDTSSVDKEELISQFPHLDRLDPSPWAQ-
KILSRPPLGREKISVVLSGAGLAGISTGIILSQKVDNIDLTI-
LERSPDFGGVWFDNSYPGVQCDVPVHAYQLSFDPKRDWDRPYAYGNDIKQYWGDRAKKYQL-
NEKTRFGHNILEAKFNKNTSQWVIQVETVADKKRSEIRADVFIATSGALNNPRYPPTQPGFDSFQGIKFH
PQQWPEGLDLTGKRVALIGNGATGVQILPQIVEQAAHVDHYAKSSSWIGHALFGPGVPGY-
VEYSRDDIESIKSDKDYLEFQKELHRNIGGKYDFFFYGTPAFRELTKELLAVA-
WIRVGKDPKLFRKVVPTYPFGAKRLLPAPGYLEALTRPNVDYLLGDVKEFTKNGIIGADG-
VERQVDVIIAATGYPLTNGNGFTPNYEIIGTDGYSLRQHFSPLESRLGYSASYLGLAAPGFPNFFYTLSVN
SYITKSTPAETVELQAAYIARAIRKKQLEKIKSLEPSLKATVSFNRRITELS-
KAISVTKGNGFFNEVTKDGTKRSKVDWPGSVSHAIAVLREPRWEDFEY-
RYEDNDDPFAYFGSGKTWIDDHDGDKTFYITQSATVAAKVHEGWISLPSDGPPRVVAR

YMOE:

Nucleotide sequence:

ATGTCGACAGTATTTGCAGACG-
GATCGTTGTTTTCCAACGTGTTGATTCTCGGCACCGGGTTTTCCGGCCTCGCCACCTCCAT-
CAAGCTCAGAACCTCCTGGAAAGAGGCCGACTTCCATCTCTATGATCGAGAT-
CACTCCTGGGGAGGAACCTGGGCGGCCAACACGTACCCTGGCTGTGGATCGGACATTCCTGCCATC
TGGTACTGTCTCACAAGCGACCCCAAGGGCGATTGGAGCAAGGCCTTCCCTCCCAGAGACGA-
GATCTACGACTACATCCAGAAGCTGGTGGCCAAGTACGAGTTGCGGCA-
CATGGCGACTTTCCGTACGGAAATCGAGGGCTGCAAGTGGAATGCTGACGA-
GAAGTTGTGGTATGTGAACGTGCGATCTCTCGAAACCGGCAAGAAGTGGGTGCACAAGTGTCGCGT
CCTGTTCACATGCAAGGGCGGACTTGTGGAGCCCAACCGCGTGCAGATTGAGGGTCTGTA-
CACCGACTTTAAGGGTCCTGTGATGCACACAGCCCGATGGGACCATTCTGTCGACTACACCAA-
CAAGAATGTCGTTGTTATCGGCAACGGATGTTCTGCCATCCAGGTAATTGCGGCTATTCAT-
GACCAGACAAAGACTCTCACCCAGTTTGCTCGAACTCCACAGTGGATCACTCCTCGTCCCGAGTTT
GTGCCCGGACGAATCACCCGTTTCATTTTCACCCGGTTCCCGTTTGTGCTCCATCTTTTGCGCAC-
CATTGTCTTCTTTGTCATTGAGGCCGCCTATCCCATGTTTAA-
GAAGGGCTGGTTGGGCACCTTTATCCGACGAATTCGAGCCCATCACGCGACTCAAAATATCCGAAA
GAAATCTCCCGAAAAGTACTGGAAGGTCCTTAAGCCCGACTACGAGTTTGCGTGCAAACGACT-
CATCTTCGACTGCGGATACCTGGGGCCTGCACTCAATAACCCAAACATGGAGCTGA-
CATTCGACCGGGTCGTCAAGGTTGAGTCCAACAAGGTGATCACCAAGGACGGCAACTCTTACC-
CAGCCGACATCATCATTGATGCCACGGGCTTTGATCTTAGTGGAGGCTTTACTAACATTCCCTTCAT
CGGCGAACATGGAGTCAGCCTGGAAGACTTCTGGGCCAACGGACGAGTTTCTGCCTATGA-
GACCGTCCTGGTTCCCAACTTCCCAAACAATTTCCTCATCTTTGGACC-
CAACTCTGCCACCGGACATAACTCGGTGCTGTTTGCCATCGAGAACGCCAT-
CAAGTTCTGCGAGTCCGTGGGTGTCAAGAAGCTCATTTCGGGTGAAACAGACTACATTGGAGTGCG
AGCCGAGGCCTACGACCGGTGGATTCAT-
GATGTGGATTCTTCCCATGCCCAGGGTCTGTTGCAACAGGGAGGCTGCCAGT-
CATGGTACCTGGCTGATAACGGCAGAAACGCCACCACC-
TACCCCTGGTCTCAGCTGACCGCTTGGCTGCGAGCTCGATGGATCGACAAGAACGCCATTGTGATT
GGTAACGGGGAGAAGAAAACAAAGTAA

Amino acid sequence:

MSTVFADGSLFSNVLILGTGFSGLATSIKLRTSWKEADFHLYDRDHSWGGTWAAN-
TYPGCGSDIPAIWYCLTSDPKGDWSKAFPPRDEIYDYIQKLVAKYELRHMATFRTEIEGCK-
WNADEKLWYVNVRSLETGKKWVHKCRVLFTCKGGLVEPNRVQIEG-
LYTDFKGPVMHTARWDHSVDYTNKNVVVIGNGCSAIQVIAAIHDQTKTLTQFARTPQWITPRPEFVPGR
ITRFIFTRFPFVLHLLRTIVFFVIEAAYPMFKKGWLGTFIRRIRAHHATQNIRKKSPE-
KYWKVLKPDYEFACKRLIFDCGYLGPALNNPNMELTFDRVVKVESNKVITKDGNSYPADIII-
DATGFDLSGGFTNIPFIGEHGVSLEDF-

WANGRVSAYETVLVPNFPNNFLIFGPNSATGHNSVLFAIENAIKFCESVGVKKLISGETDYIGVRAEAYD
DYIGVRAEAYDRWIHDVDSSHAQGLLQQGGCQSWYLADNGRNATTYPWSQLTAWLRARWIDK-
NAIVIGNGEKKTK

YMOF:

Nucleotide sequence:

ATGCGGAGATTGAAATACAACTTTCAGTCAGTATATATACAGACGCGTCTCCGCCAATTCAAA-
CACATATCACCCTCTACCACTCCGATCGGTTACATTTGGCTTGCGCATCACGTGCCAGAT-
CATCCAACCCATCCTTCTGGAAACCCACGACTCGTTATTTACACGTTTTCCACAAC-
CAAAATGCCTACTGGACAGGAGCTCTACGACTCGCTGCCAAATCTGCACTTCCCCGATGAGCCCAA
CACCGCTGCAGAGGAGATTCTCAAGCGAGACTTTGCTGGTCACCGAAAGGTGA-
GAGCGGTGGTAGTCGGAGGAGGCCTTGCAGGAATCACCCTGGGAAC-
CATTCTGCCCCGAAAGCTCGACAACCTGGATCTTGTCATCTACGAGCGATGGCCTGA-
GACTGGAGGTGTGTGGCACCGAAACACCTACCCCGGAGTCAAGTGTGACATTCCATCTCACAATTA
CCAGCTTTCCTTTGATCCAAAGACCGACTGGAGTGCCACCTATGCCCCCGGTCAGGAGATCAA-
GAGTTACTGGCAGGGCATTGAGAAGAAGTACGGCGTCGACAAGCTGATCAAGACCAACCACGA-
CATCCAGAGTGCCGACTGGGACGCTGAGAAGGGCAAGTGGATTTTCAAGATCAAGGATCTCAA-
CACCAACACCGAGTTCACCGACGAGGCCGAGTTCTTCATCCAGGCAACCGGTATTCTCAACAACGC
ACGGTACCCTCCTTACCAGCCTGGCTTTGAC-
GATTTCCAGGGTCCCAAGTTCCACCCTTCTCAGTGGCCCAAGGACC-
TATCTCTCAAGGGCAAGCGGGTCGCGTTGATTGGTAACGGTGCTTCAGGTGTCCA-
GATTCTTCCCCAGCTGCTTGCCCAGGGTGTCTCTCACGTTGACCACTACGCCAAGCGAGGAACCTG
GATCTCTCAGCACGTTTTCGGAAAGCACCTTCCTCCCCACCGAGAGTACTCTCCCGAAGA-
GATTGCAGAGCTGCGAAACACGGAAAAGTACCACAAGTTCCGTAAGGACCTCGA-
GACTCGAGGCCAGGGAAACATTGCCTCCGACGTCTACGGATCTGAGCAGAACCGACAG-
CAGCTCAATGCCTTTCTGCTGCTCATGTACGAGCGGCTTGGCGGTGATGAGGAACTCTTCAAGAAG
GTTGTTCCCGACTATGCTCCCGGATCTCGAC-
GATTTCTGCCTGCTCCCGGCTACCTGGAGGCTCTGACTGACCCTCGAGTTTCATAC-
CATCTGGGAACCGTCAAGTCGTTCACCAAGACTGGTGTTGTTGGTGCCGACGACGTTGAGC-
GACCTACAGATATCATTGTGGCCTCCACCGGATACACCCGAGCCAACGGAGAGTCGCATGCCCCCA
ACTTTGAGGTCACTGGTCTCGACGGTACCAATCTCAAGGAG-
CACTTTTCTGGAGCTGGCTCCAAGCTCGGATACACCAACAACTACTACGGAAT-
CACTTCTCCCCATTTTCCCAACTACTTTTACGTGCTGGCTCA-
GAACTCCTATCTGTTCTGCGGCCCTGCTCCTATCGCTGCCGAGCTGTGGTCCACCTACATCTCCAAG
GTGATCCGAAAGGTGCAGCTGGAAAACATCAAGTCACTGGTTGTGTCCGA-
GAAGGCTGCGCTTGGTTTCTCTCGAGTAGTCACCGAGCTATC-
CAAGGCCTCGTCTACTTCTCGAGGTATCGATGGCTTCTTCGTCGAGAAAACCAAGGATGGA-
GAATACAGAATTGCCCTTGCCTGGCCAGGAACTATCACCCATGCCGTAACTCTGCTGCGAGAGCCT
CGATGGGAGGATTACGAGTACGAATACCTGGACAACGA-
CAACCCCTTCTCCTTCTTTGGAAACGGACACACCTTCCTGGACTTTGCCCCCAAGGGAGACAA-
GACCTTCTTCGTGCAGACCGGTGTGCCTCCCAAGCTCCTCCACGAGGAGTATCTGAC-

TATCCCTCGAGACCATGTTGCCGAAGGATACGAATATGATGGAACCGGTGACTTCCTCAAGAACCA
GAACCACCATGTGGGACTAGATGACGACGAAGACGTCGAGCAGGCCAAGGAGGTCAGTCTCTAA

Amino acid sequence:

MRRLKYNFQSVYIQTRLRQFKHISPSTTPIGYIW-
LAHHVPDHPTHPSGNPRLVIYTFSTTKMPTGQELYDSLPNLHFPDEPN-
TAAEEILKRDFAGHRKVRAVVVGGGLAGITLGTILPRKLDNLDLVIYERW-
PETGGVWHRNTYPGVKCDIPSHNYQLSFDPKTDWSATYAPGQEIKSYWQGIEKKYGVDKLIKTNHDIQ
SADWDAEKGKWIFKIKDLNTNTEFTDEAEFFIQATGILNNARYP-
PYQPGFDDFQGPKFHPSQWPKDLSLKGKRVALIGN-
GASGVQILPQLLAQGVSHVDHYAKRGTWISQHVFGKHLPPHREYSPEEIAELRNTE-
KYHKFRKDLETRGQGNIASDVYGSEQNRQQLNAFLLLMYERLGGDEELFKKVVPDYAPGSRRFLPAPG
YLEALTDPRVSYHLGTVKSFTKTGVVGADDVERPTDIIVASTGYTRANGESHAPN-
FEVTGLDGTNLKEHFSGAGSKLGYTNNYYGITSPHFPNYFYVLAQNSYLFCGPA-
PIAAELWSTYISKVIRKVQLENIKSLVVSEKAALGFSRVVTELS-
KASSTSRGIDGFFVEKTKDGEYRIALAWPGTITHAVTLLREPRWEDYEYEYLDNDNPFSFFGNGHTFLD
FAPKGDKTFFVQTGVPPKLLHEEYLTIPRDHVAEGYEYDGTGDFLKNHHVGLDDDEDVEQAKEVSL

YMOG:

Nucleotide sequence:

ATGTCGAGCAAAAACGGAACTGGACACACTGACACTTCTGGGGTGGACAAGGAG-
GAGCTCATCTCCCGATTTGAGCACCTCAACAACCTCGAGCCTTCTGAATGGGCCGAAAA-
GATTCTTTCTCGTCCACCTCTGGGTCGAGACGCCGTCAAAGTCGTGA-
TATCTGGAGCTGGACTCGCAGGTATCACCACTGGCATCATTCTGTCCAACAAAGTTGACAACATTG
ATCTGACGATTCTGGAGCGAAGCCCCGAGTCTGGCGGTGTTTGGTTCGACAATCATTATC-
CAGGCGTCGCCTGCGATGTCCCTTCTCACGCGTACCAGCTTTCGTTTGACCCCAAGAAG-
GACTGGAATCGGGCCTACGCCAAAGGACCCGACATCAAACGGTACTGGCAATCTCGAGCCAA-
GAAGTACGGTCTGGAGAACAAGATCAAGTTCCGCCACAACATTGACGAGGCCAAATGGGACGAAA
AGACTCATCAGTGGGTTCTCCAGGTGGAAGAACTGGAAGCCAGGAAGAAGAGTGAGATTCGAA-
CAGACATCTTCATCTCGTCTTCGGGTTCCCTTAACAACCCCCGATACCCTCCCCAC-
CAGCCGGGCTTTGACTCGTTCCAGGGAATCAAGTTCCATCCTCAGAAGTGGCCTGAAGGGCTA-
GACTTGACTGGTAAACGGGTTGCACTGATTGGAAATGGAGCTACTGGAGTCCAGATCCTTCCTCAG
ATTGCCGAGAAGGCTGCTCATGTCGACCACTACGCCAAGTCCGCCACCTGGATTGGACA-
CACTCTGTATGGCAAGGGCGTGCCTGGATACGTGGATTACACCGATGAAGAAAT-
TAAGGCGATCGAGACAGACGAGGAATACCACAAGTTTCGAAAGGCTCTCCATACCGA-
GATTGGAGGCAAGTACAATTACTTCTTCTACGGAACTCCAGCCTTCAGAGAGGGTATCAAAGAGCT
CCTTGCTATTGCGTGGCTGCGAGTTGGTAAAGATCCCGAGCTGTTCAA-
GAAGGTCATTCCCCTCTACACCCCGGGTCCGA-
GACGTCTTCTGCCTGCTCCTGGCTATCTGGAGGCTCTGACAAGATCCAATGTTGACTAT-
CATCTGGGAGAGGTCAAGGAGTTCACCAAGGACGGAGTGATTGGCTTTGATGGCGTTGAACGGAA
GGTAGACGTTGTCATTGCTGCCACTGGATACATCAAGTCCAATGGACAGGGATT-
CACCCCCAACTTTGACATCATTGGTCAAGACGGATACACTCTGAGAGAA-

CACTTTTCTCCTCTCGAGTCCAAGCTGGGTTACTCGGCCTCA-
TACCTTGGTCTATCTGCTCCAGGCTTCCCCAATTTCTTCTACACGCTCTCTGTCAACTCGTA-
CATTCCCGAGACTACTGCTCCTGTGACCGCCGAAGTCCAGGCCTCGTACATTGCTCGAGT-
GATTCGAAAGAAGCAGCTTGAGAAGATCTTGTCCATCGTGCCGTCTCTGGAAGCTACCAAGGCGTT
CAATCGACGTCTTGCCGAGCTGTCCGAAGCCGTCTCTCTTACCAAGGGAACAGGTATTTA-
CAGCGAGAGAACTCGAGATGGAGACAGTCGATTGAA-
GATTGCCTGGCCCGGATCTGTGTCTCATGCTGTGGCTGTTCTGCGTGAGCCTCGCTGGGAA-
GACTACGACTATGAGTACGAGGACAATGATGATCCGTTTGCGTACTTTGGGTCTGGCAAGACCTGG
ATTGATGATCACGAGGGCGACAGGACGTTTTACTTGTCTGAGCCGGGCTCAAT-
CACTGCCCGTAATCTCCACGAGGGTTGGATTTCCGTACCCTCA-
GACGGTCCTCCCAGTGCCCCTCATTCAAAGTGA

Amino acid sequence:

MSSKNGTGHTDTSGVDKEELISRFEHLNNLEPSEWAEKILSRPPLGRDAVKVVISGAGLA-
GITTGIILSNKVDNIDLTILERSPESGGVWFDNHYPG-
VACDVPSHAYQLSFDPKKDWNRAYAKGPDIKRYWQSRAKKYGLENKIKFRHNIDEAKW-
DEKTHQWVLQVEELEARKKSEIRTDIFISSSGSLNNPRYPPHQPGFDSFQGIKFHPQKWPEGLDLTGKRV
ALIGNGATGVQILPQIAEKAAHVDHYAKSATWIGHTLYGKGVPGYVDYTDEEIKAIET-
DEEYHKFRKALHTEIGGKYNYFFYGTPAFREGIKELLAIAWLRVGKDPELFKKVIP-
LYTPGPRRLLPAPGYLEALTRSNVDYHLGEVKEFTKDGVIGFDGVERKVDVVIAAT-
GYIKSNGQGFTPNFDIIGQDGYTLREHFSPLESKLGYSASYLGLSAPGFPNFFYTLSVNSYIPETTAPVTAE
VQASYIARVIRKKQLEKILSIVPSLEATKAFNRRLAELSEAVSLTKGTGIY-
SERTRDGDSRLKIAWPGSVSHAVAVLREPRWEDYDYEYEDNDDPFAYFGSGKTWIDDHEGDRT-
FYLSEPGSITARNLHEGWISVPSDGPPSAPHSK

YMOH:

Nucleotide sequence:

ATGACT-
GAGCTCTACTCTCACTCTCTCATTGTCGGCGCAGGCTTCTCTGGCGTGGCCACTGCAATCAAG-
CACATCAAGGAGTGGAACAACCCCGACTTCCACATCTACGACAGAGACTCGGCCTTTGGAGG-
CACCTGGAAGGCCAATACCTATCCTGGCTGTGCCTCGGACGTTCCTGCCATCTTTTATTGCCTTACC
TCAGACCCCAAGATTGATTGGAGCCACATGTACCCCTTCCAGAACGAGCTGTTCCAG-
TACTTTCAGGATGTTGCCACCAAGTACGGACTTCCCGACAAGAGCACTCTCAACACC-
GAAATTGTGGAGATGCGATGGAACGAGAAAACCAAGGAATATACCACAACTCTAC-
GAAACGTCAAGACTGGAAACACCCACACTCACCGATCCAAGGTGGTGTTTGTGGGCCGAGGATGT
CTCGTGGCCCCTAACAAGCTTAACCTGCCTGGTCTCGAGACCTTTAAGGGACCCGTCATGCA-
CACTGCCCAATGGGACCATAAGAACTCCATTGTCAACAA-
GAACGTGGTTGTTGTTGGCCACGGTTGCTCTGCTGTGCAGGTTGTGTCTGATATTGCAC-
CAAAGTGTAAGACCCTGACCCAGTTTGCCCGGTCTCCACAGTGGATTATTCCTCGAATCGAAAAGA
TTCTTCATCCTGGCTTCATGAAGTTCCTATCTTACATTCCAGGAGCAGTCCAAC-
TAACTCGGCTGGTGCTCTTCTTCCTGCTCGAGTACTCGTGGAC-
CATGTTCTCTGGAGCTTGGTGGTCAAAGCTTGACAACAGAATCAAGTCCACTCTGACATCCA-

GATGGATGCGATCTAAAGTGCCCGCCAAATATCAT-
GACGTCATTGTGCCCAAGTTCTCCATGGGCTGCAAGCGAAC-
CATCTTTGATCCCGGTTACCTCAAGCAGCTGTGGCTTCCCACCATG-
GAATTGTCCTTTGACCCTATTGTCAAGGTCAAGGAGCATTCAGTGGTGTCCCGAAACGGCCTCGAG
TACCAGGCCGACGTCATTATCGACGCCACTGGCTTTGATATCCCTAAGTCAATCTCAGGACT-
CAAGGTCATTGGAGAGAATGGAATCGAACTGGACCAGTTTTGGAACGGAAGAGTGTCTGCA-
TACGAGACCGTCCAAGTGCCTAACTTCCCCAACCTCTTTTTCATCTTCGGACC-
CAACGCCTTGACCGGTCACAACTCAGTCACCTTTGCTATCGACAATTCTCTGGTCTATGTTGACAAG
GTTGCCAGAGATCTTGTGTCTGGTAAGCCCAACTGCACGTACGTGTCAGTGTCCGAG-
GAGGCCTACGACAAGTGGGTTGACGACGTCCAGGAGGCCACTTCTAAGAC-
CACCTTTGGATCGGGAGGCTGTGCTTCCTGGTACCTCGGTGCCAACAAATACAACGGCAC-
CACCTATCCCTGGACTCAGATTCGAGCTTGGTGGCACTCGCATTTTCCCAACCAGAAAGATATCATC
CGACATTAA

Amino acid sequence:

MTELYSHSLIVGAGFSGVATAIKHIKEWNNPDFHIYDRDSAFGGTWKANTYPGCASDVPAI-
FYCLTSDPKIDWSHMYPFQNELFQYFQDVATKYGLPDKSTLNTEIVEMRWNEKT-
KEYTTTLRNVKTGNTHTHRSKVVFVGRGCLVAPNKLNLPGLETFKGPVMHTAQWDHKNSIVNKNVVV
VGHGCSAVQVVSDIAPKCKTLTQFARSPQWIIPRIEKILHPGFMKFLSYIP-
GAVQLTRLVLFFLLEYSWTMFSGAWWSKLDNRIKSTLTSRWMRSKVPAKYHD-
VIVPKFSMGCKRTIFDPGYLKQLWLPTMELSFDPIVKVKEHSVVSRNGLEYQADVII-
DATGFDIPKSISGLKVIGENGIELDQFWNGRVSAYETVQVPNFPNLFFIFGPNALTGHNSVTFAIDNSLVY
VDKVARDLVSGKPNCTYVSVSEEAYDKWVDDVQEATSKTTFGSGGCASWYLGAN-
KYNGTTYPWTQIRAWWHSHFPNQKDIIRH

YMOI:

Nucleotide sequence:

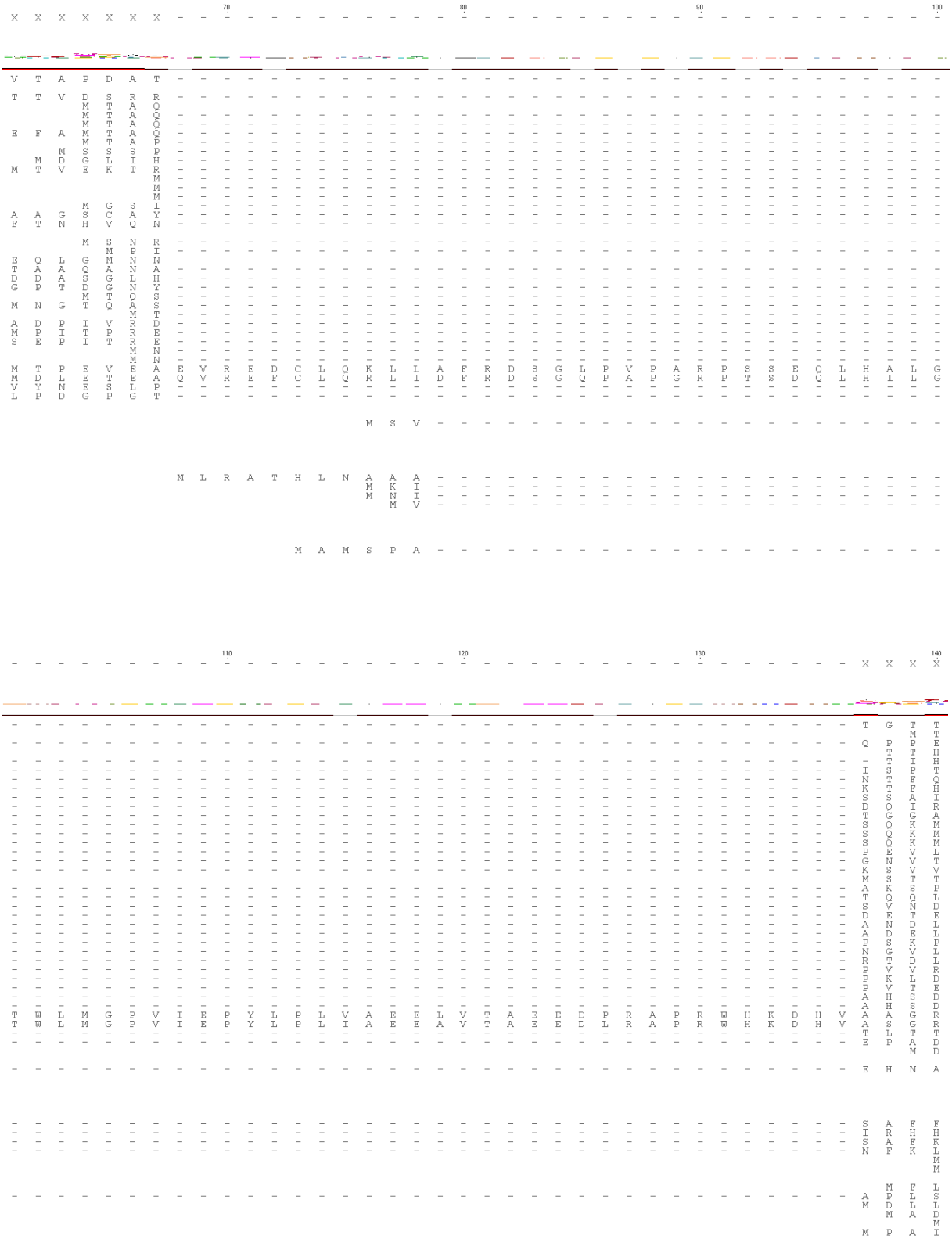
ATGGACTCTTCTAACCACACAAATATGAATCGACACACT-
CACGCCCTCATTGTTGGTGCTGGCTTCTCTGGCCTGGCATCAGCCATCAAGCTCCAGACC-
GACTGGAACACCACCGACTACCAAATCTACGACCGAGATTCCGAGTTTGGCGGCACTTGGCAA-
CAAAACACTTATCCTGGAGCTGCGTCCGACATTCCTGCACTGTGGTACTGTCTCGCTAGCGATCCCA
AGGTCGACTGGAAAGAGCCCTACCCTTCTCAGGAAGAGCTGAGACAGTACATCAAG-
GACGTGGCAGAGAAATACAACCTACGAAAGCGAGCTACTTTTGGAGCTGAGATCGA-
GAAGGTGGAGTGGCTGGCCGACCAGCAAATGTGGAAGGCATCCATCAAG-
GACGTGGCCACCGGTAACAAGTATACCCACACCTCGCGGGTGGTTTTTATGGGCAAGGGCTGTCTT
GTTGTTCCCAACAAGTTCAAGACTGCCGGAATTGAGGATTTCAAGGGGCCTATCATGCA-
TACCGCCCAATGGGACCACTCTGTCGACTACAAGGGCAA-
GAACGTGGTTGTGATTGGCAACGGCTGCTCTGCTGTCCAGGTGTGTGCTGCTATTGCACCCGAGGT
AGGTTCTCTGACCCAGTTTGCACGAACTCCTCAGTGGATGGTGCCCCGACCTGAATG-
GAAGTGGCTCAAGAC-
GATGGGTGAAACCTTCCCCTTCATGCTTGGATTTGTGCGATTTCTTATGTTCCTGA-
CACTCGAGGCCAACTTTTCGCTCTTCCGAGGTGGCTGGTATGCGCGAGCCGACCGAGCTGTCCGAA

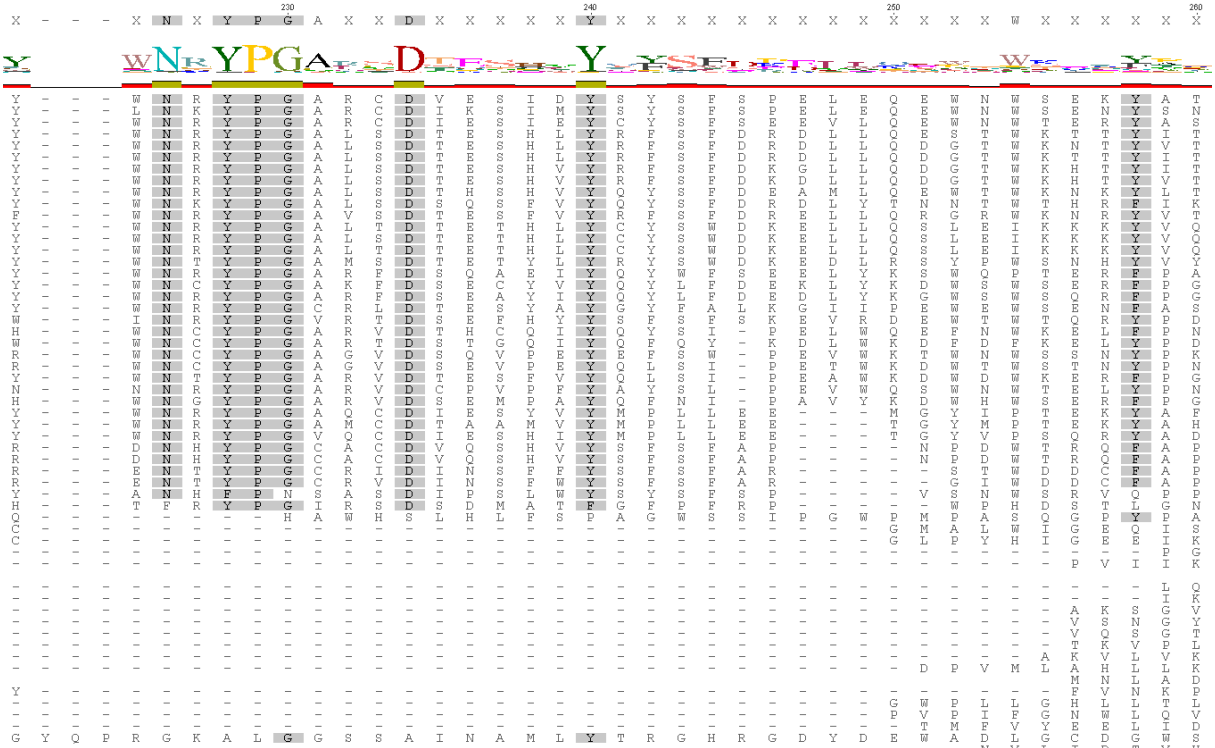
CCTGGGTCAGCACCATGCTCCTCAAGTGGCACTTGCCCAAGAAGTACCACGA-
GAACTCCATCCCCAAATACGAGCTTGGCTGTAAGAGAATCATCTATGATTGTGGA-
TACTGGAAAGCCCTCCACCAGAAGAATGTTGAACTGACTTACGATCCCATTGTTAGAATGAC-
CAGCAATAGTGTTATCACCAAGAGTGGCCAGGAGTGGCCAGCCGATATTGTCATTGATGCTACTGG
ATTTAATGTCGCCGCTTCCATGGGAGGTCTGGAGATTGTGGGCGAGACTGGCGA-
GAACCTGGTGGACTTCTGGAACGGCAAGGTCTCCGCTTACGAGACAGTTATGGTCGCCAAC-
TACCCCAACATGTTCTTCCTATTTGGCCCCAACGCCACCACTGGCCA-
CAACTCTGTCATTTTCGCAATTGAGAACGCTCTCAAGTGGATCGAGAATGTCGCTTCTGATCTCGTC
ACCGGATCTGCCACCTACGTCACAGTCAAGAACGAGGCCTACGACTCTTGGACCCA-
GAAGGTGCACGAGGCTTCCAAAAAGATGGCTTTCTCCACTGGAGGATGTGTTTCCTGGTATAT-
GAGTGCCTCTGGAGCTGGACACAACGGTGTTACCTACCCCTGGACT-
CAGTTTACTGCTTGGTGGAGAGCCCGATTCCCCGTCAAGAGCGATATGATTGTCAAATCCAAGAAG
GACGAGTAA

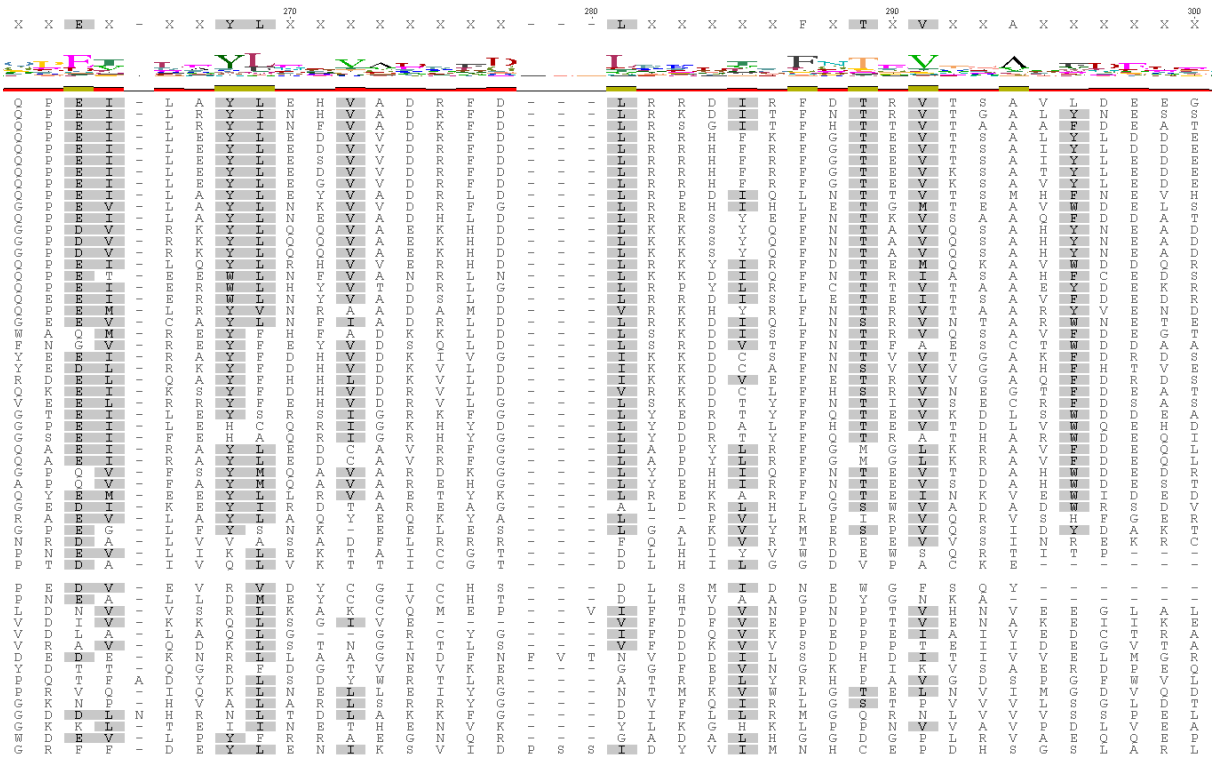
Amino acid sequence:

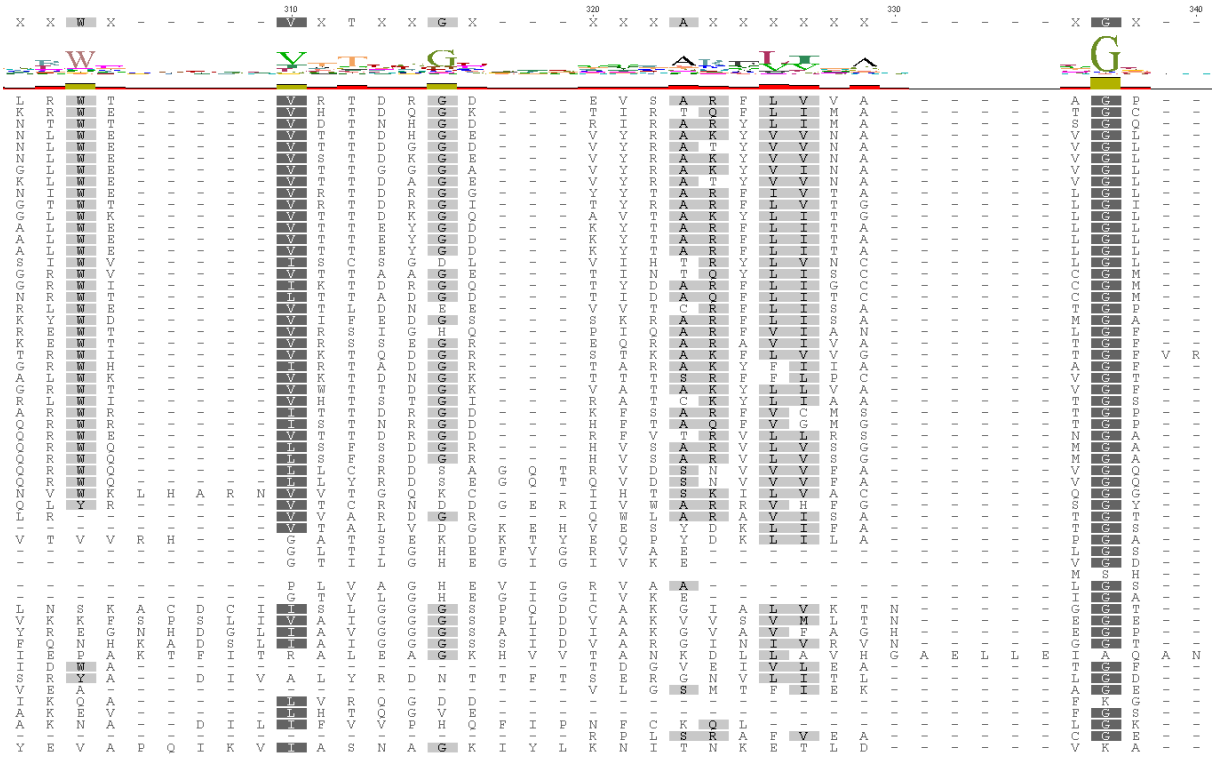
MNRHTHALIVGAGFSGLASAIKLQTDWNTTDYQIYDRDSEFGGTWQQNTYPGAASDIPALWYC-
LASDPKVDWKEPYPSQEELRQYIKDVAEKYNLRKRATFGAEIEKVEWLADQQMWKASIKD-
VATGNKYTHTSRVVFMGKGCLVVPNKFKTAGIEDFKGPIMH-
TAQWDHSVDYKGKNVVVIGNGCSAVQVCAAIAPEVGSLTQFARTPQWMVPRPEWKWLKTMGETFPF
MLGFVRFLMFLTLEANFSLFRGGWYARADRAVRTWVSTMLLKWHLPKKYHEN-
SIPKYELGCKRIIYDCGYWKALHQKNVELTYDPIVRMTSNSVITKSGQEWPADIVI-
DATGFNVAASMGGLEIVGETGENLVDFWNGKVSAYETVMVANYPNMFFLFGPNATTGHNSVI-
FAIENALKWIENVASDLVTGSATYVTVKNEAYDSWTQKVHEASKKMAFSTGGCVSWYMSASGAGHN
GVTYPWTQFTAWWRARFPVKSDMIVKSKKDE

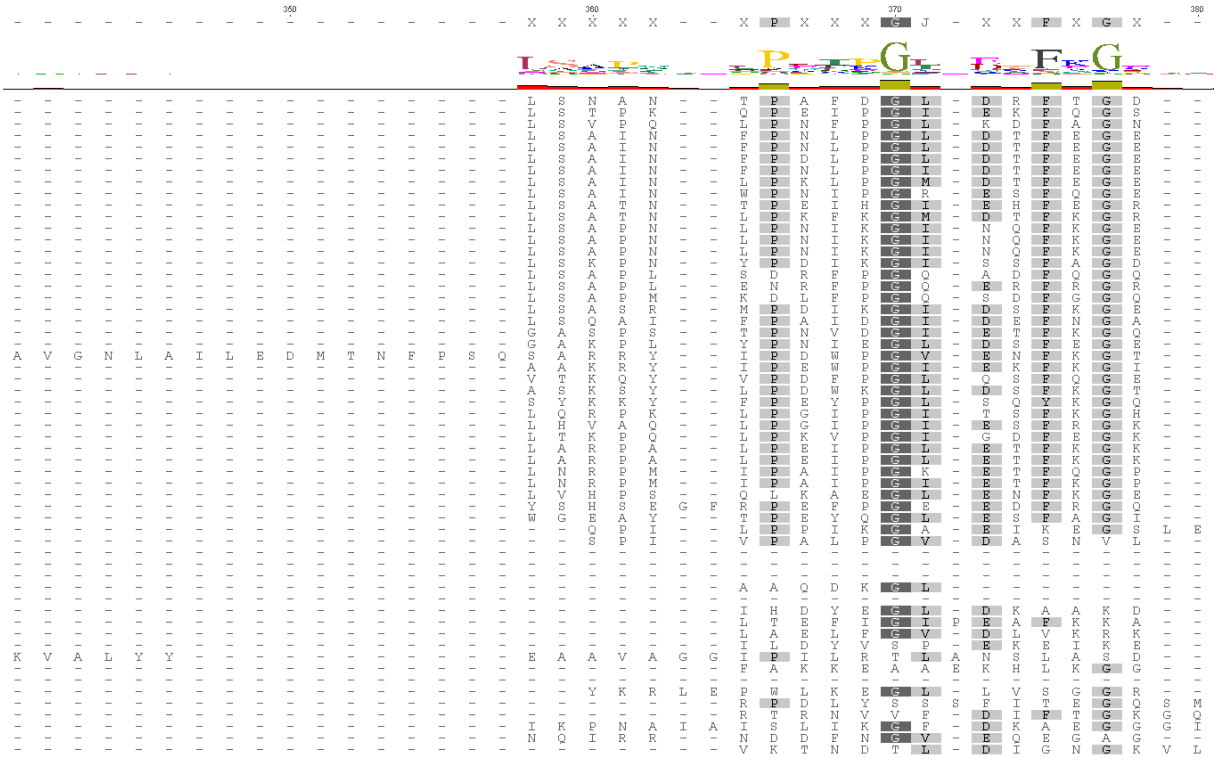
8.2 Switch of the cofactor specificity of $CHMO_{Acineto}$

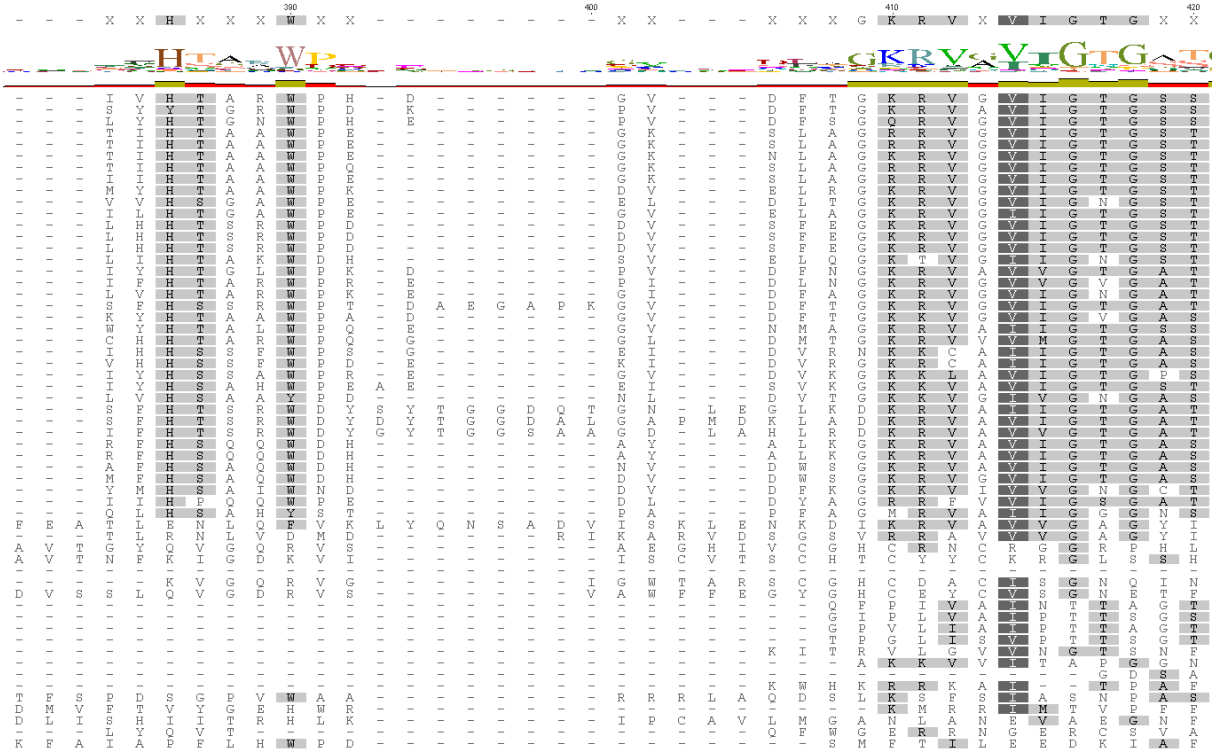


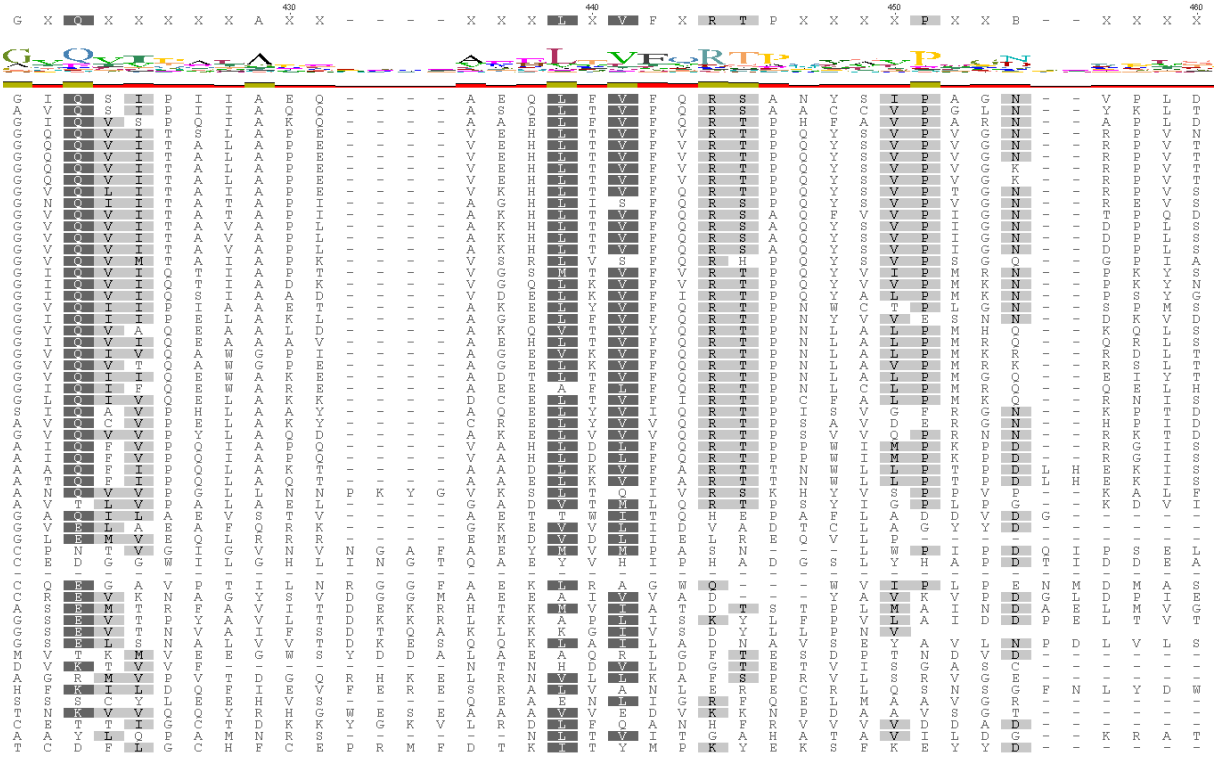

230
240
250
260

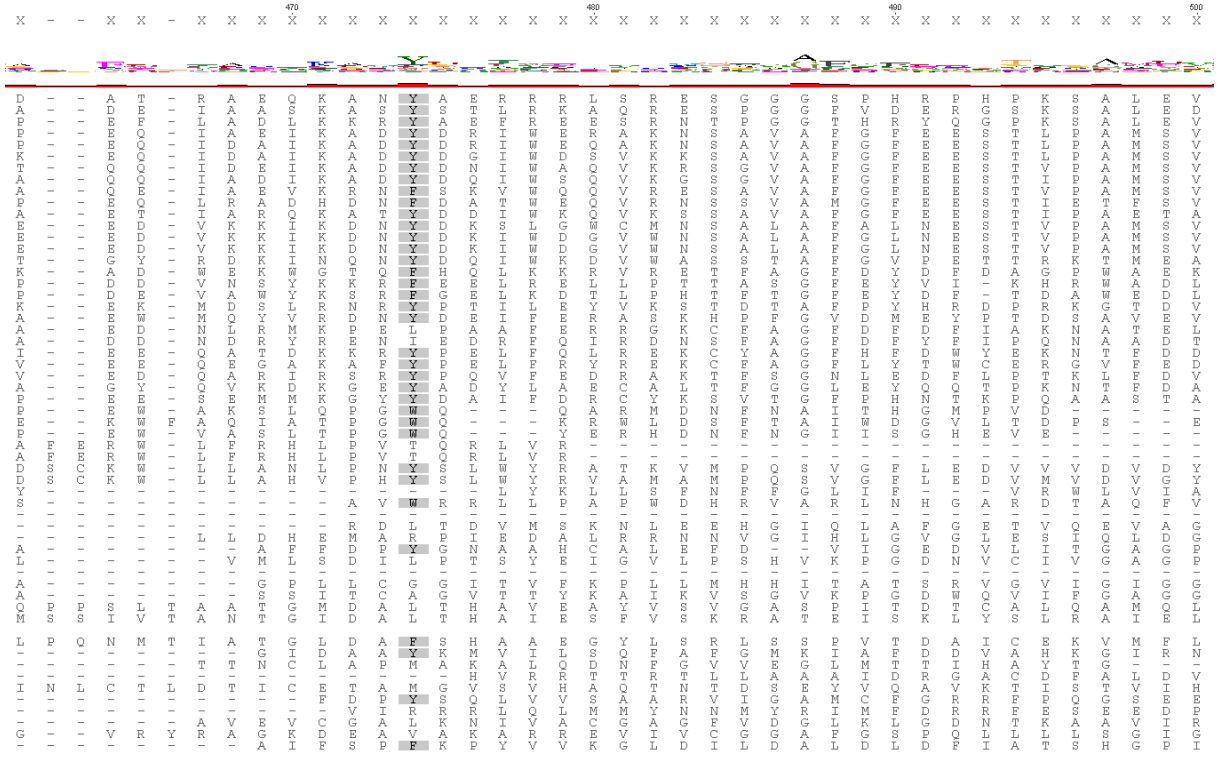

270
280
290
300

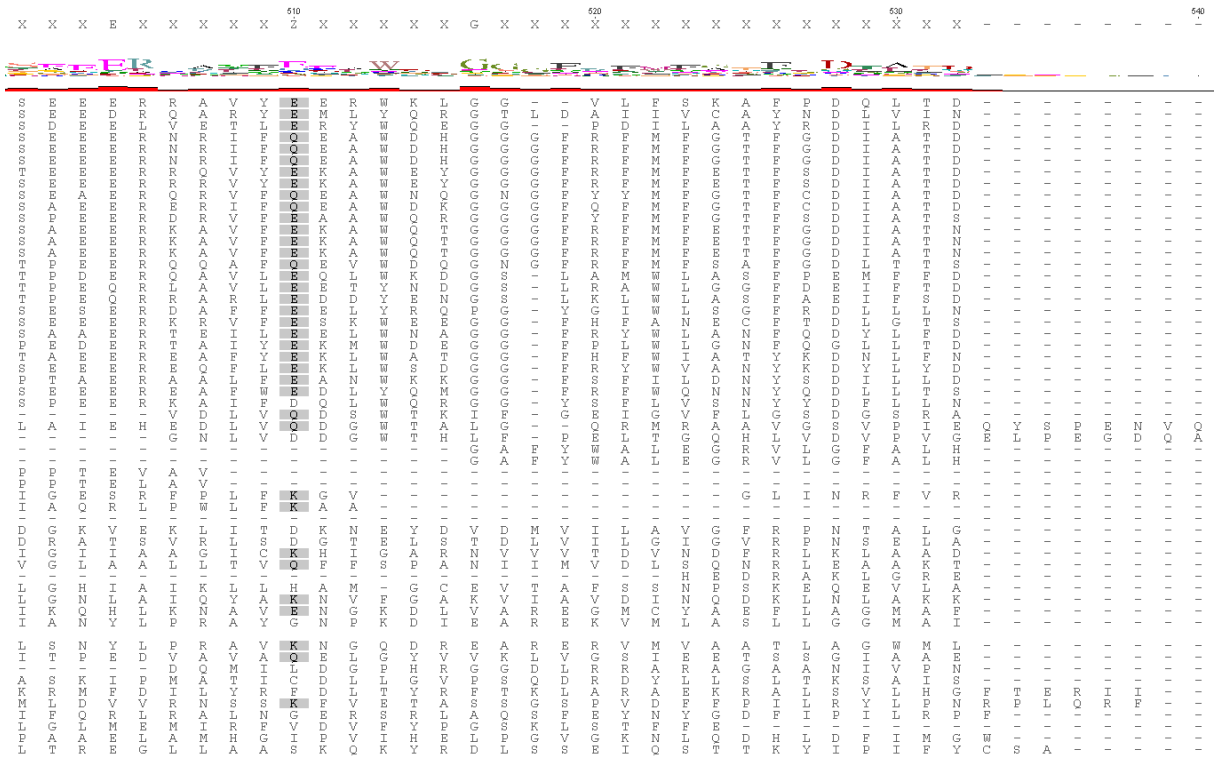


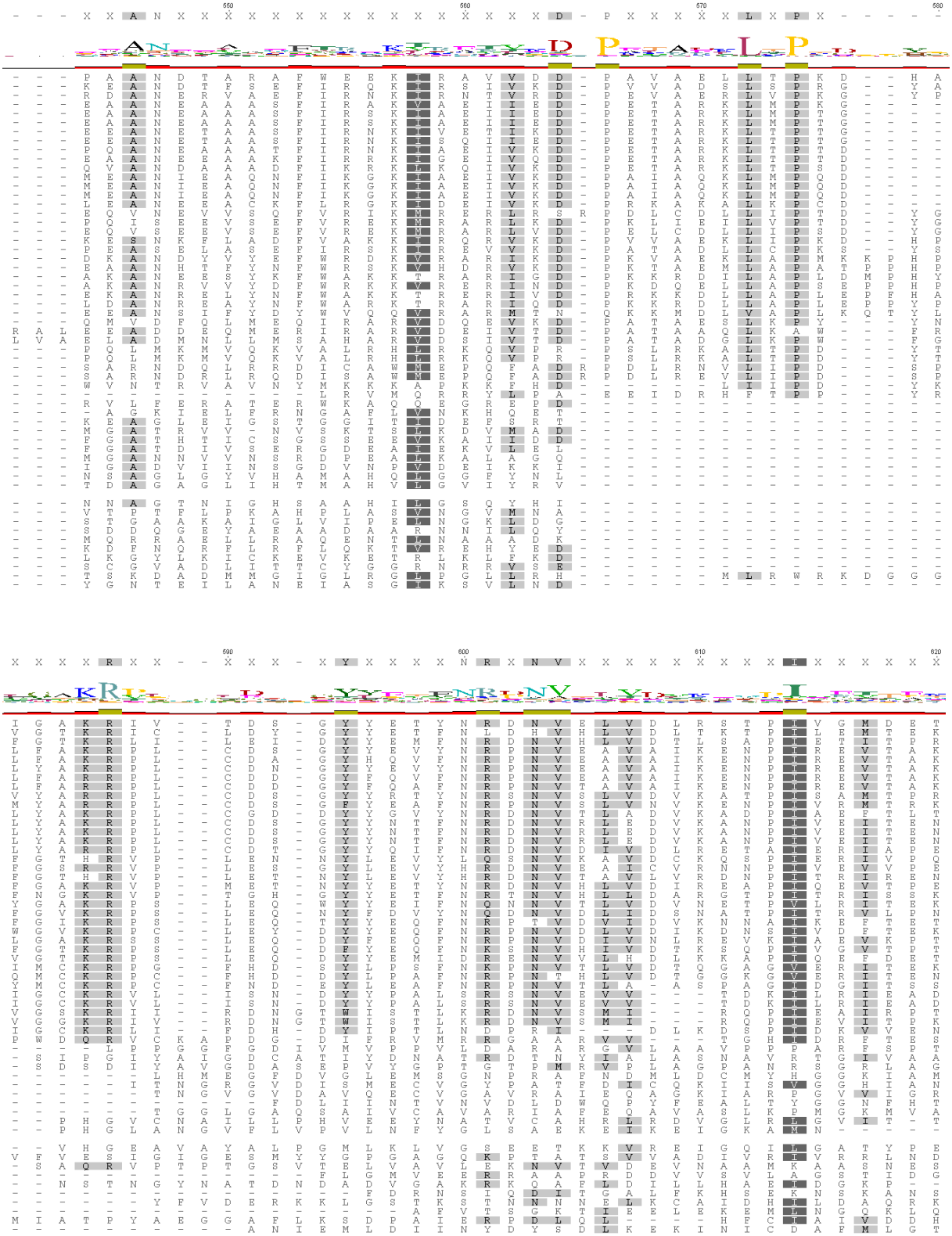


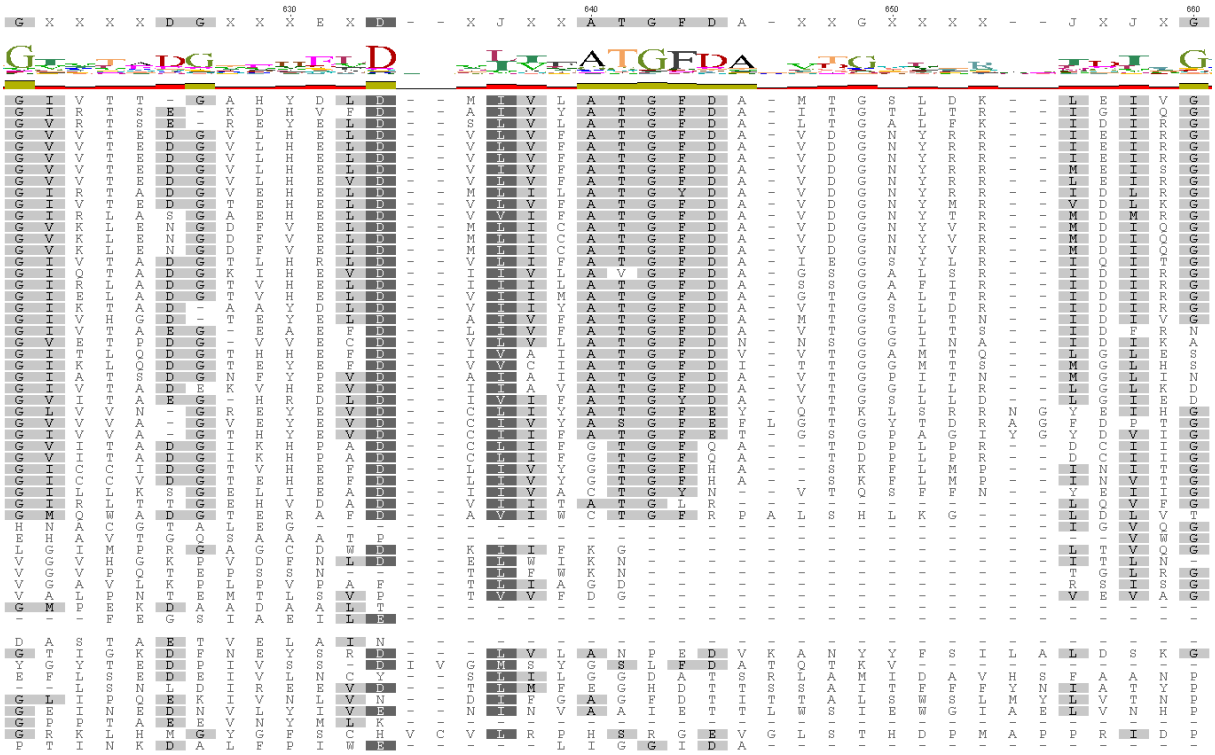


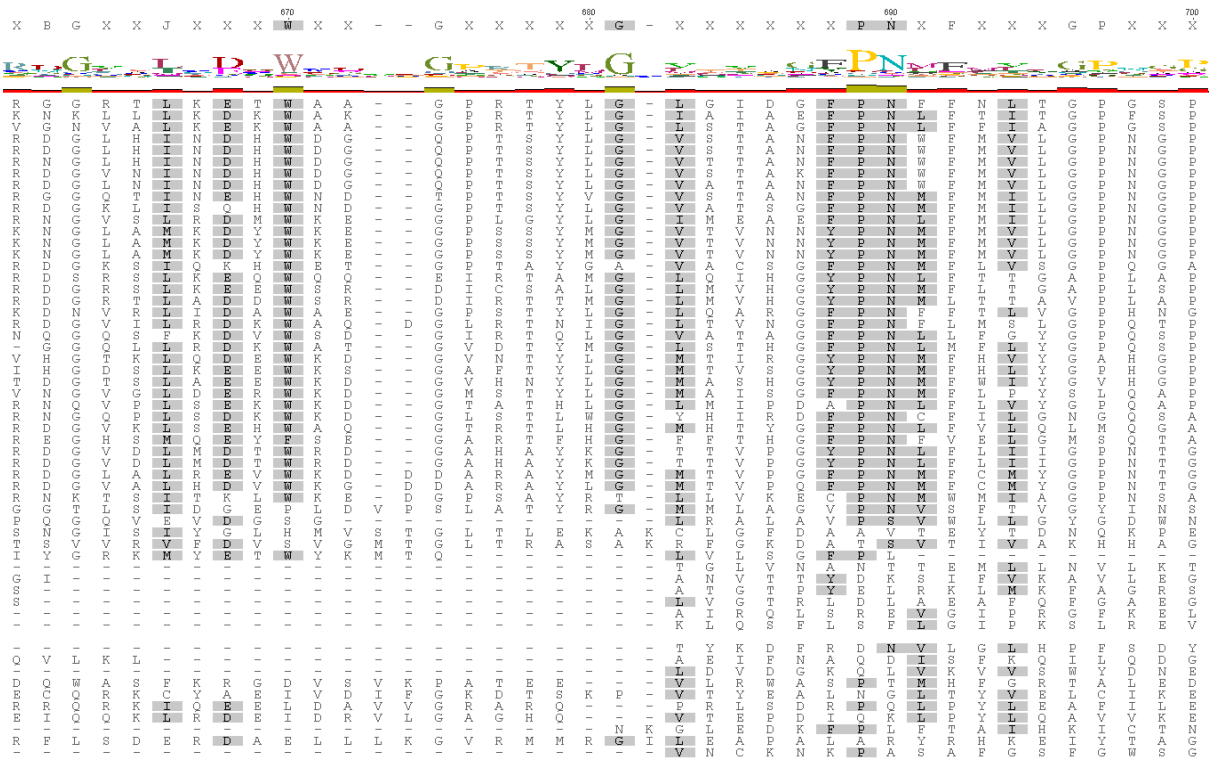


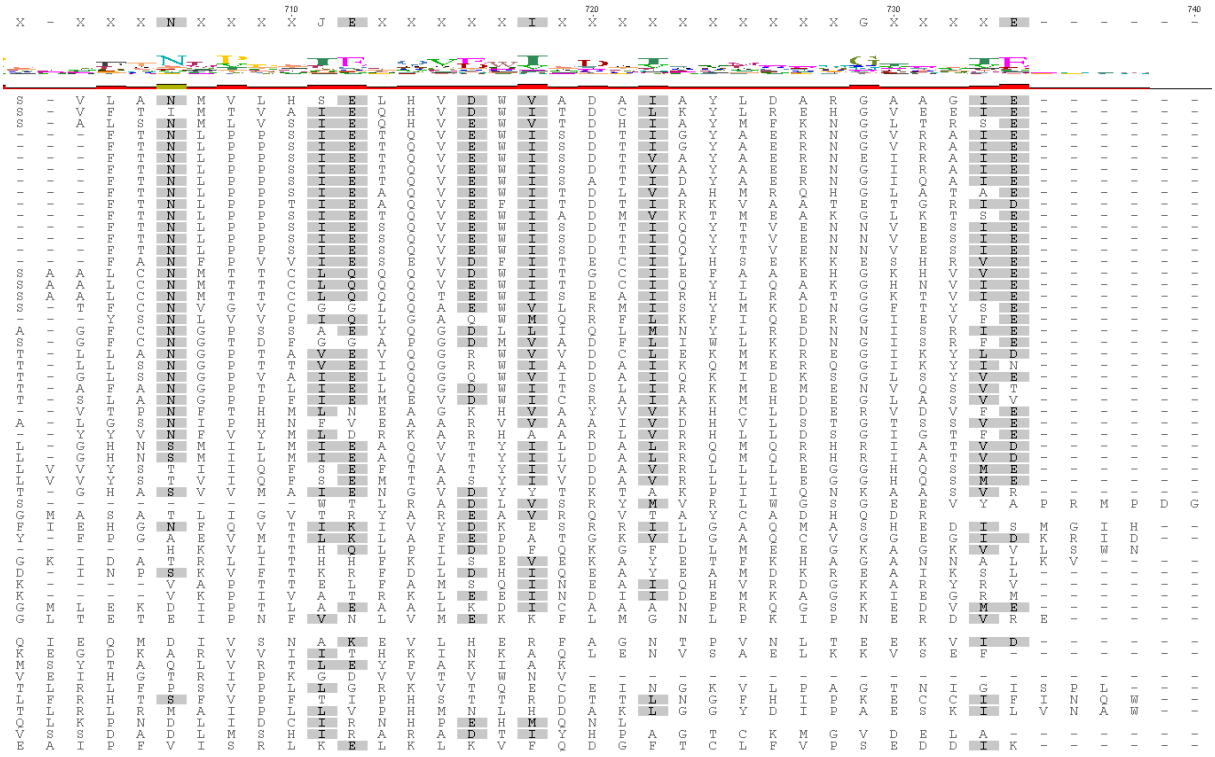

470
480
490
500

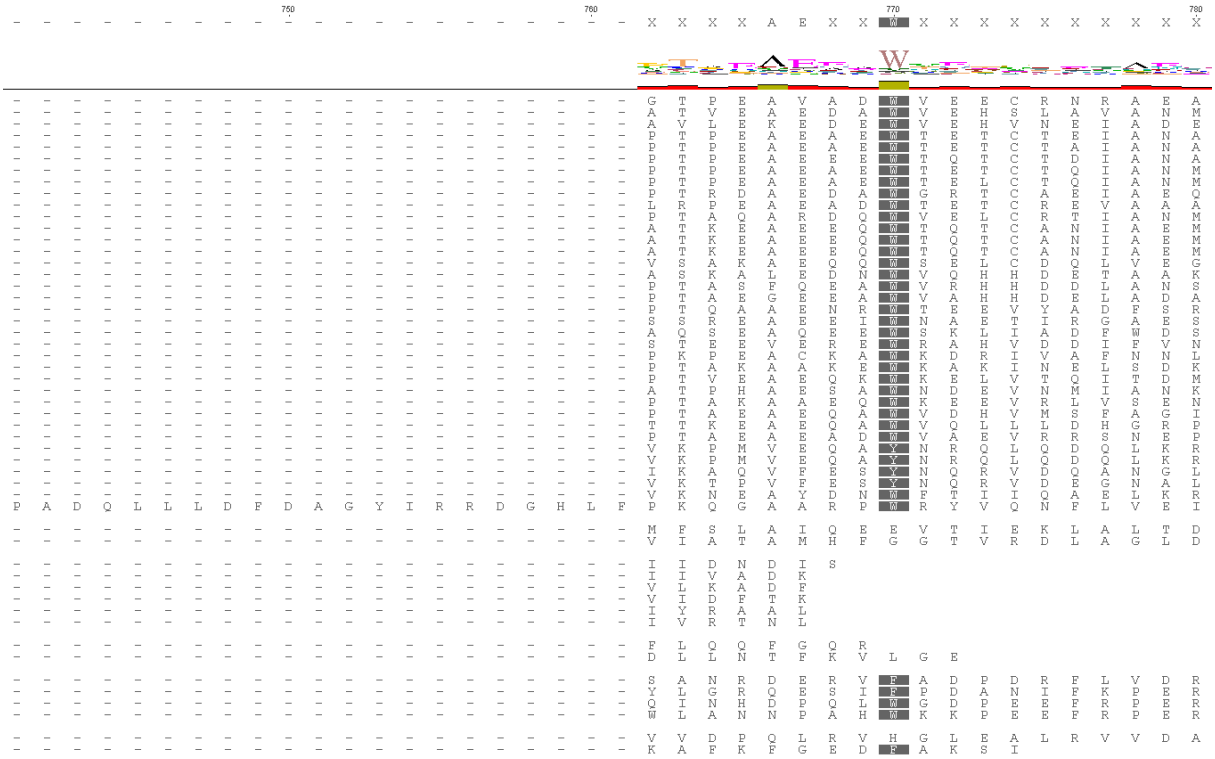

510
520
530
540

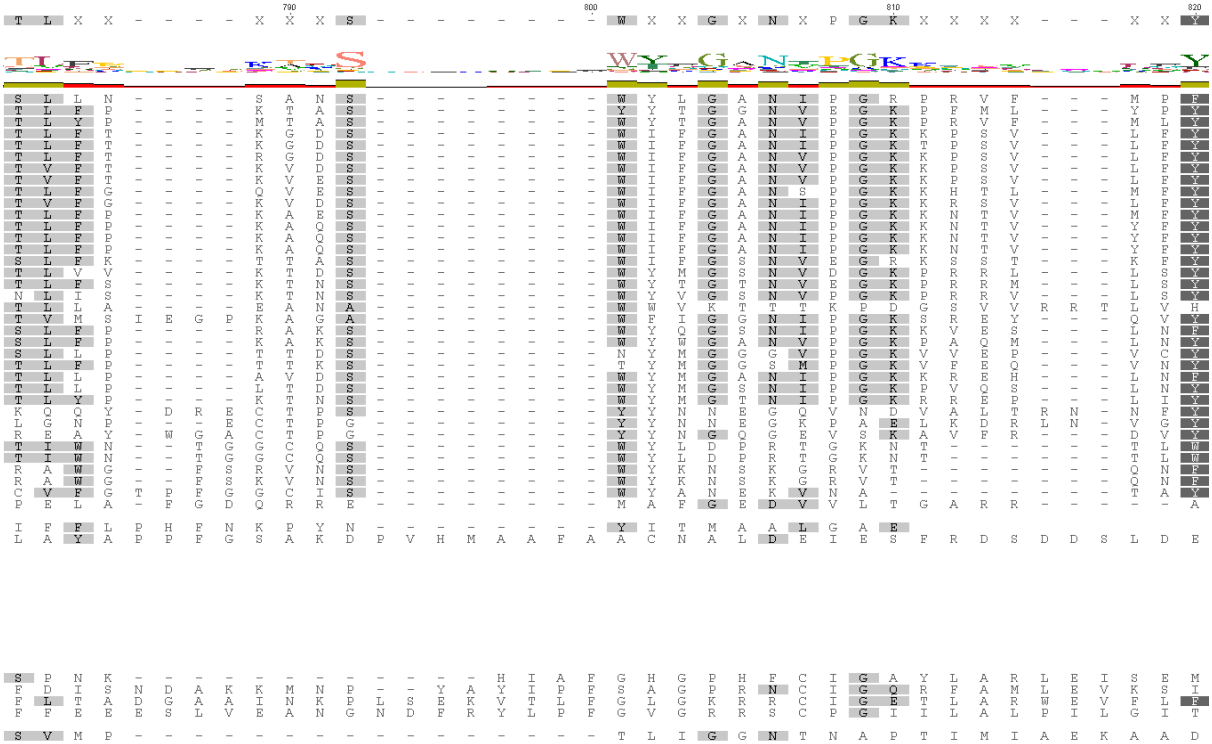



630
640
650
660


670
680
690
700







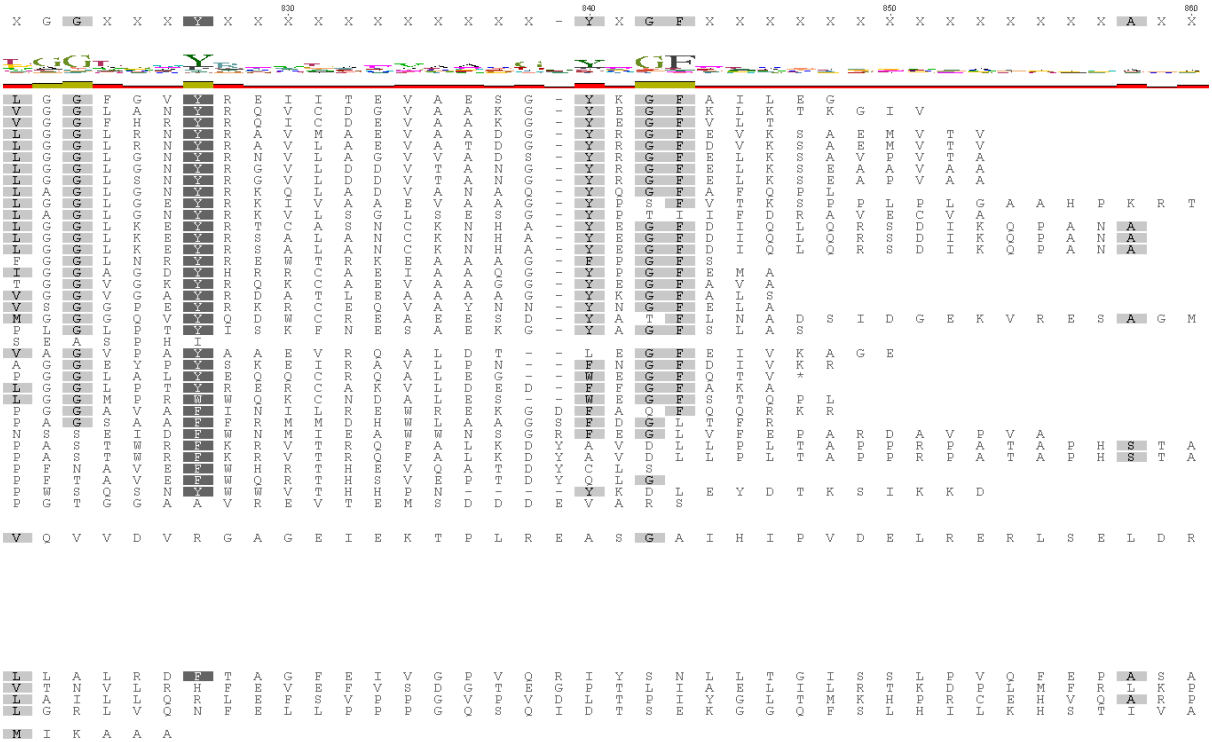


Figure 8.6: Multiple protein sequence alignment of 37 BVMOs and 20 NADH employing enzymes.

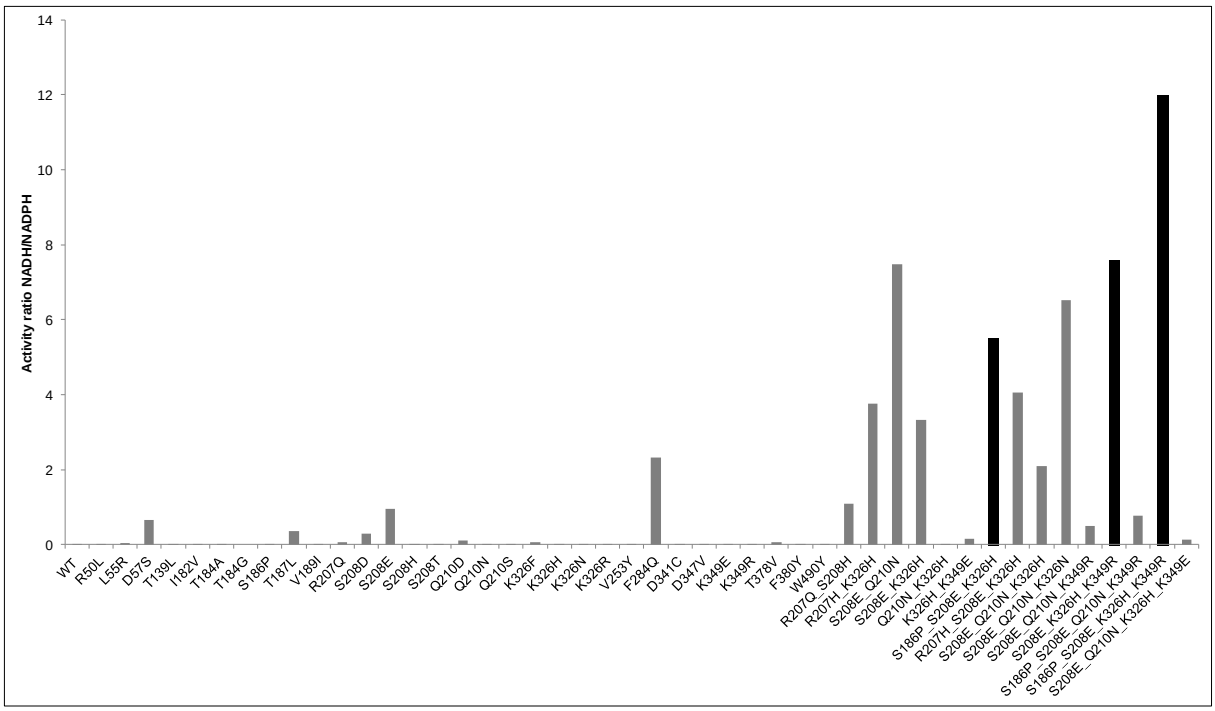


Figure 8.7: Specific activity ratio for the utilization of NADH over NADPH with enzyme variants of $CHMO_{Acineto}$.
Black bars: mutants with a specific activity with NADH >0.4 U/mg.

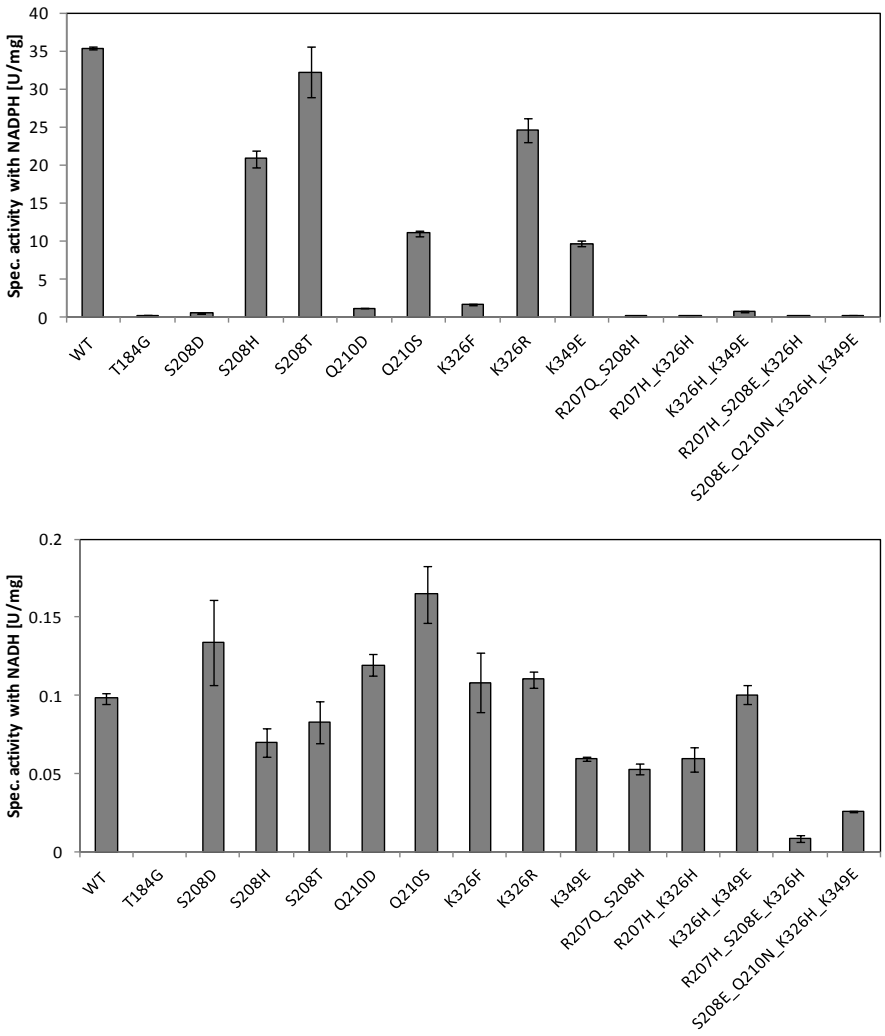


Figure 8.8: Specific activity of $CHMO_{Acineto}$ wild type (WT) and variants (mutations in the phosphate recognition site) using NADPH (top graph) or NADH (bottom graph) as cofactors.

Table 8.2: Residues in the proximity of the cofactor NADPH in $CHMO_{Acineto}$.

Residues	Characteristics
Y51	in “fingerprint 2”
L55	H-bond to NADPH; in “fingerprint 2”
D57	catalytically important; H-bond to NADPH; in “fingerprint 2”
L144	-
N148	Proximity to adenine moiety of NADPH
P150	Proximity to adenine moiety of NADPH
I182	next to Rossmann fold
G183	in Rossmann fold
G185	in Rossmann fold
S186	H-bond to NADPH; in Rossmann fold
T187	H-bond to NADPH; in Rossmann fold
G188	in Rossmann fold
V189	next to Rossmann fold
Q190	-
D322	-
R327	catalytically important; H-bond to NADPH
L329	-
V348	-
T378	in “fingerprint 3”; Proximity to adenine moiety of NADPH
F380	in “fingerprint 3”
Q488	-
S489	-
W490	H-bond to NADPH
N495	Proximity to adenine moiety of NADPH

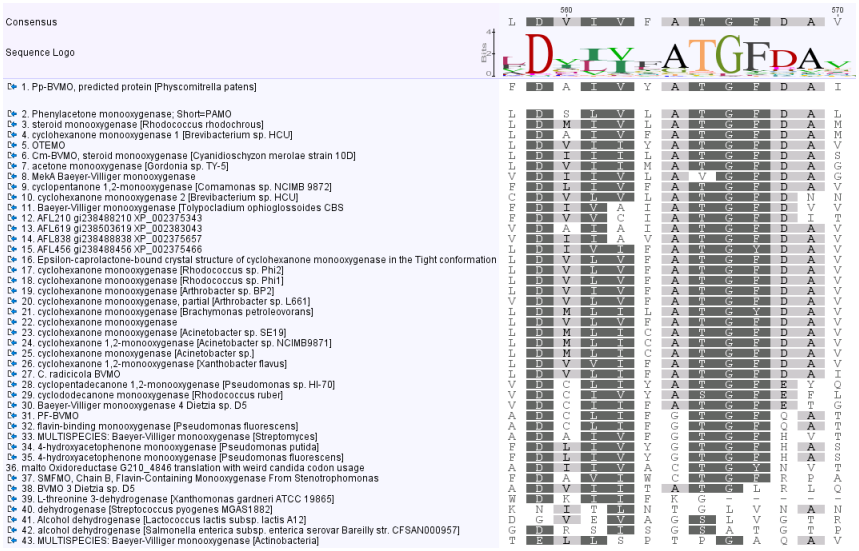


Figure 8.9: BVMO "fingerprint 3" Dx[I/L][V/I]xxTG[Y/F] in a multiple protein sequence alignment of 39 BVMOs and 4 NADH employing enzymes.

Hiermit erkläre ich, dass diese Arbeit bisher von mir weder an der Mathematisch-Naturwissenschaftlichen Fakultät der Ernst-Moritz-Arndt-Universität Greifswald noch einer anderen wissenschaftlichen Einrichtung zum Zwecke der Promotion eingereicht wurde. Ferner erkläre ich, dass ich diese Arbeit selbständig verfasst und keine anderen als die darin angegebenen Hilfsmittel und Hilfen benutzt und keine Textabschnitte eines Dritten ohne Kennzeichnung übernommen habe.

Greifswald, _____ _____

Ort, Datum Unterschrift

Publications

[1] Beier, A., Hahn, V., Bornscheuer, U. T., Schauer, F. (**2014**), Metabolism of alkenes and ketones by *Candida maltosa* and related yeasts, *AMB Express*, 4, 75.

Author contributions:

- F.S. and U.B. initiated the project. A.B. performed the experiments with support by V.H. All authors analyzed the data, wrote and approved the final manuscript.

[2] Beier, A., Bordewick, S., Genz, M., Schmidt, S., van den Bergh, T., Peters, C., Joosten, H.-J., Bornscheuer, U. T. (**2016**), Switch in cofactor specificity of a Baeyer-Villiger Monooxygenase, *ChemBioChem*, 17, 2312-2315.

Author contributions:

- A.B. developed the concept for this work. S.B. and A.B. planned and conducted all experiments. S.S. created and validated the homology model. T.B. and H.J.J. provided a 3DM database for BVMOs and related sequences. M.G. and U.T.B. gave advice throughout the work. A.B. drafted the manuscript, which was revised by S.B., M.G. and U.T.B.

[3] Bordewick, S., Beier, A., Balke, K., Bornscheuer, U. T. (**2017**), Baeyer-Villiger monooxygenases from *Yarrowia lipolytica* catalyze preferentially sulfoxidations, *Enzyme Microb. Technol.*, accepted.

Author contributions:

- A.B. and U.T.B. developed the concept for this work. S.B. and A.B. planned experiments and S.B. conducted all experiments. A.B., K.B. and U.T.B gave advice throughout the work. K.B. drafted the manuscript, which was revised by all authors.

[4] Balke, K., Beier, A.; Bornscheuer, U. T. (**2017**), Hot spots for the protein engineering of Baeyer-Villiger monooxygenases, *Biotechnol. Adv.*, submitted.

Author contributions:

- K.B. and A.B. contributed equally in drafting the manuscript, which was revised by all authors.

Acknowledgment

First to mention is Uwe. I'm very grateful for the possibilities you gave me – starting from allowing me to work on a PhD thesis in your group over valuable advices you provided for my projects up to interesting conversations about more than just work-related topics.

I also would like to thank Prof. Dr. Schauer for giving me permission to investigate the metabolism of ketones in yeasts in his lab and his dedication concerning this study leading to an interesting scientific article.

Further I want to thank without exception the whole group Biotechnology & Enzyme catalysis for the nice working atmosphere and helpfulness. I would like to mention some colleagues here, which made my stay in the group especially enjoyable.

Javier helped me in the first year of my PhD. Thank you for the motivation, sympathy and for being an amazing and very funny person!

With Micha, Sandy, Ghandi, Daniel, Ioannis, Thomas, Anders, Alberto, Jennifer, Isabel, Ayad and Lukas I enjoyed the many funny conversations we had and an overall good interpersonal relationship.

The meetings of the BVMO Crew starring Sandy, Maika, Sven, Marcus and Kathleen were very productive, pleasant and of big assistance. Maika guided them for the first one and a half years. She also supported me in the creation of our journal article.

I would like to express my gratitude to Kathleen for many fruitful conversations, excellent and competent advice and ideas, correcting my thesis and for just being a delightful person.

Sven was my Master student, but not just that – he became a friend and in my opinion we formed a great team and achieved a lot in the course of our combined efforts. With his knowledge and excellent scientific understanding we also produced a really good publication. He carefully corrected my thesis as well, for which I am very thankful. You are awesome!

Ina and Angelika deserve recognition for their help in analytical questions, especially regarding the gas chromatography, and managing the chemicals and other lab ware, respectively.

I would like to thank Anne and Veronika of the group Applied Microbiology & Biotechnology for their encouragement in my growth and metabolic analyses and their very pleasant personalities.

Marco Fraaije accepted a PhD exchange program to his group in Groningen (The Netherlands) in which I gained a lot of useful experience. Furthermore, I want to thank him and his research group, here especially worth mentioning Hugo van Beek and Elvira Romero, for the assistance with my experiments and for the very friendly work environment.

Bilge, Mattia, Mateus and the others from my exchange group in Groningen made my stay unforgettable. I hope to see you soon again, somewhere in the world!

Franklyn, Antje, Laura and the rest of the Latin dance community here I want to thank for the great experiences in salsa cubana, bachata, reggeaton and tango and their friendship.

Lorena, Gago, Cristina, Oriana, Ixchell and the others from the Alemañol group, where I became the president during my PhD thesis – I'm thankful for the nice time, in which I met so many nice people, had unforgettable moments and improved my Spanish a lot.

Thilo, Ulli, Barbara and Piet deserve to be mentioned for extraordinary experiences in Karate-Do, numerous funny lunch times and for being such good friends, which are there whenever I need them.

Soraia I want to thank for the nice times we had and especially for always listening to my problems and having sympathy plus giving advice that helped me a lot to endure difficult times.

Thank you Gleyder for being the best bro one could imagine to have. We´re not just sharing the same roots but also the same kind of humor and the same open-minded attitude. Thank you for everything, let it be on funny events, enjoyable meetings or when I had problems.

To find my relatives in Cuba and experience their very friendly and affectionate personality during my PhD time gave me a lot of strength and thus they deserve a lot of thankfulness as well.

I’m also grateful to be part of the local Erasmus Student Network (ESN) group LEI (local erasmus initiative), because in course of the related activities I gained so great and memorable experiences, learned many organizatorial skills and met some of my best friends – Laura, Martina, Jasmin, Vittoria, Alan, Valen and Gyöngyi – and my girlfriend Mirka. I also want to deeply thank them for making my time in Greifswald just awesome.

Mirka, you’re giving me so much strength and encouragement all the time since we are together. I’m so happy with you and looking forward to many further beautiful moments with you. I love you.

My Mother deserves the biggest appreciation. You supported me all the time and you made it possible for me to go my way and achieve everything I wanted to. I’m so grateful for everything you’ve done for me.