

**Protein expression and purification of moxY and of other  
Baeyer - Villiger monooxygenases**

**Master's Thesis**

Julia Pitzer

Institute of Molecular Biotechnology  
Graz, University of Technology, Austria

Biocrystallography Laboratory,  
Department of Biology and Biotechnology "Lazzaro Spallanzani"  
University of Pavia, Italy

Supervisor: Ao.-Univ.-Prof. Mag. Dr.rer.nat. Anton Glieder

Prof. Andrea Mattevi

Graz, November 2011 to December 2012

## Acknowledgements

First of all, I want to say THANK YOU to several people, without their help this thesis would not have been possible...

...to my parents for fully supporting me not only during my years of study but at any time of my life. Special thanks are due to my whole family and my sisters for always having an open ear and for being welcome at any time.

...to my supervisor Ao.-Univ.-Prof. Mag. Dr.rer.nat. Anton Glieder for the interesting project which was full of variety and for giving me the possibility to make this thesis in cooperation with the lab in Pavia.

...to the glieder group for the great working atmosphere and the entertaining lunch breaks. A big thanks goes to Martina for her support in the lab as well as during the writing of the thesis and for answering all my questions. Thanks also to Christian Schmid for his help with the dot blot and for taking care of the shipments between Graz and Italy.

...to Andrea Mattevi for his motivating supervision and his support during my stay in Italy. Moreover, I want to say thank you to the whole lab group in Pavia. Thank you for everything I learned there and of course for giving me an idea of the Italian "dolce vita". A huge thanks to Christian, who supported me at any point in the lab as well as in everyday life. I will never forget the great months in Italy.

...to the Oxygreen project for the financial support.

...and last but not least to all my friends who helped me to manage the good as well as the sometimes hard and stressful periods. Thank you for numerous great evenings and weekends together and for making my years of study unforgettable.

THANK YOU ALL!

## Abstract

Baeyer-Villiger monooxygenases (BVMOs) are able to catalyze a large variety of selective oxidation reactions. Therefore most of them require FAD as cofactor and use NADPH as electron donor. Their catalytic versatility and high selectivity make BVMOs promising biocatalysts with a high potential for industrial applications.

The main focus of this work lied on moxY, a mainly uncharacterized BVMO which is involved in the aflatoxin biosynthesis of several *Aspergillus* species. For the expression of this enzyme the methylotrophic yeast *Pichia pastoris* was chosen and a MBP-tag was fused to the protein to allow the subsequent purification of moxY. To identify strains with an increased level of expression, a high-throughput screening system via dot blot analysis was established for MBP-tagged proteins in *P. pastoris*. The cultivation of the cells was successfully up-scaled from deep well plates to a fermenter. Various methods were tested for the following purification of moxY, including affinity chromatography as well as size exclusion and anion exchange chromatography. First crystallization trials were performed to resolve the three-dimensional structure of the protein.

Additionally, the phenylacetone monooxygenase PAMO from the thermophilic bacterium *T.fusca* was expressed in *E.coli*, purified with anion exchange and size exclusion chromatography and used for crystallization experiments. Based on this well established working procedure, a similar protocol for the expression and purification of a new BVMO, named BVMO24 from *R. jostii*, was developed.

## Kurzfassung

Baeyer-Villiger Monooxygenasen (BVMOs) katalysieren eine große Vielfalt an selektiven Oxidationsreaktionen. Die meisten von ihnen benötigen dazu FAD als Cofaktor und verwenden NADPH als Elektronendonator. Ihre vielseitige katalytische Einsetzbarkeit und ihre hohe Selektivität machen BVMOs zu Erfolg versprechenden Biokatalysatoren mit großem Potential, die auch für industrielle Anwendungen geeignet sind.

Der Schwerpunkt des Projekts bezog sich auf moxY, eine bisher noch wenig bearbeitete BVMO, die an der Aflatoxin-Biosynthese verschiedener *Aspergillus* Spezies beteiligt ist. Für die Expression dieses Enzyms wurde die methylotrophe Hefe *Pichia pastoris* verwendet. Es wurde ein MBP-tag an das Protein angehängt um anschließend die Reinigung von moxY zu ermöglichen. Um Stämme mit einem hohen Level an Proteinexpression zu identifizieren, wurde ein high-throughput Screening System mittels Dot-blot Analyse für Proteine mit MBP-tag in *P. pastoris* entwickelt. Weiters konnte der Maßstab der Kultivierung der Zellen, angefangen von Deep Well Platten bis hin zu einem Fermenter, erfolgreich vergrößert werden. Die darauffolgende Reinigung von moxY erfolgte mittels Affinitäts-, Größenausschluss- und Anionenaustauschchromatographie. Erste Kristallisationsversuche wurden durchgeführt, um die drei-dimensionale Struktur des Proteins bestimmen zu können.

Zusätzlich wurde die Phenylaceton Monooxygenase PAMO aus dem thermophilen Bakterium *T. fusca* in *E. coli* exprimiert, mittels Anionenaustausch- und Größenausschlusschromatographie gereinigt und für Kristallisationsexperimente verwendet. Basierend auf diesem bereits gut etablierten Arbeitskonzept konnte ein ähnliches Protokoll für die Expression und Reinigung einer neuen BVMO, namens BVMO24 aus *R. jostii*, entwickelt werden.

## Contents

Acknowledgements .....	II
Abstract .....	III
Kurzfassung.....	IV
Abbreviations.....	VIII
1. Introduction.....	1
1.1. OXYGREEN and Green Chemistry .....	1
1.2. Flavoproteins.....	2
1.2.1. Monooxygenases.....	3
1.3. Baeyer - Villiger Oxidations.....	3
1.3.1. Chemical Baeyer - Villiger reactions.....	3
1.3.2. Enzymatic Baeyer - Villiger Reactions .....	4
1.3.3. moxY.....	8
1.3.4. Phenylacetone monooxygenase (PAMO).....	12
1.3.5. BVMO24.....	18
1.4. <i>Pichia pastoris</i> .....	19
1.4.1. Why choosing this expression system? .....	19
1.4.2. Methanol metabolism .....	20
1.4.3. Phenotypes and strains .....	22
1.4.4. Secretion .....	22
1.4.5. Fermentation .....	25
1.5. Protein Purification .....	26
1.5.1. Affinity chromatography .....	26
1.5.2. Anion exchange chromatography .....	27
1.5.3. Size exclusion chromatography.....	28
1.6. Crystallography .....	28
2. Objectives .....	31
3. Chapter I: Secretion study with moxY .....	32
3.1. Background.....	32
3.2. Materials and methods.....	33
3.2.1. Instruments and devices.....	33
3.2.1.1. Centrifuges.....	33
3.2.1.2. Shakers and incubators .....	33
3.2.1.3. PCR cyclers .....	33
3.2.1.4. Photometers and plate readers .....	33
3.2.1.5. Gel electrophoresis and associated materials and devices .....	33
3.2.1.6. Electroporation devices.....	34
3.2.1.7. Reaction tubes .....	34
3.2.1.8. Pipettes and pipette tips .....	34
3.2.1.9. Microplates.....	35
3.2.1.10. Other materials and devices .....	35
3.2.2. Media and chemicals.....	35
3.2.2.1. Media and antibiotics for the cultivation of <i>E. coli</i> .....	35
3.2.2.2. Media and solutions for the cultivation of <i>P. pastoris</i> .....	35

3.2.2.3.	Other buffers, solutions and chemicals .....	36
3.2.3.	Enzymes.....	36
3.2.3.1.	Restriction enzymes and ligase.....	37
3.2.3.2.	DNA polymerases.....	37
3.2.4.	Software and web tools.....	37
3.2.4.1.	Software .....	37
3.2.4.2.	Web tools .....	37
3.2.5.	Standard molecular biological techniques .....	38
3.2.5.1.	Agarose gel electrophoresis.....	38
3.2.5.2.	SDS-PAGE .....	38
3.2.5.3.	MS - analysis .....	39
3.2.5.4.	Plasmid Miniprep Kit.....	39
3.2.5.5.	Purification of DNA .....	39
3.2.5.7.	Ligation .....	40
3.2.5.8.	PCR.....	40
3.2.5.9.	Primers .....	41
3.2.5.10.	Transformation of <i>E. coli</i> Top10 cells .....	42
3.2.5.11.	Sequencing.....	42
3.2.5.12.	Transformation of <i>P. pastoris</i> cells.....	43
3.2.5.13.	Micro - scale cultivation of <i>P. pastoris</i> in 96 - deep well plates .....	43
3.2.5.14.	Cultivation of <i>P. pastoris</i> in 2 L baffled shake flasks .....	43
3.2.5.15.	Bradford assay for determination of the protein concentration.....	44
3.2.5.16.	TCA precipitation .....	44
3.2.5.17.	Spectrophotometric measurements of fluorescence and OD <sub>600</sub> .....	44
3.2.5.18.	Fluorescence microscopy.....	44
3.3.	Experimental.....	45
3.3.1.	Fluorescence microscopy study .....	45
3.3.2.	Secretion study .....	46
3.3.2.1.	Preparation of the <i>moxY</i> /citrine constructs with the $\alpha$ - factor signal sequence 46	
3.3.2.2.	Preparation of <i>moxY</i> /citrine constructs with the CBH2 signal peptide: .....	48
3.4.	Results and discussion .....	53
3.4.1.	<i>In silico</i> predictions of the localization of <i>moxY</i> .....	53
3.4.2.	Fluorescence microscopy study .....	55
3.4.3.	Secretion study .....	58
3.4.3.1.	Screening of the secretion of <i>moxY</i> .....	58
3.4.3.2.	Screening of the secretion of citrine .....	60
3.5.	Conclusion .....	65
4.	Chapter II: Expression and purification of <i>moxY</i> .....	66
4.1.	Background.....	66
4.2.	Materials and methods.....	68
4.2.1.	Instruments .....	68
4.2.2.	Media and recipes .....	68
4.3.	Experimental.....	70
4.3.1.	Optimization of the MBP gene for <i>P. pastoris</i> .....	70
4.3.2.	Preparation of MBP <i>moxY</i> constructs: .....	70

4.3.3.	Screening methods and purification pretests .....	74
4.3.4.	Large scale protein production .....	77
4.3.5.	Protein Purification.....	79
4.3.6.	Crystallization .....	81
4.4.	Results and discussion .....	81
4.4.1.	Optimization of the MBP gene for <i>P. pastoris</i> .....	81
4.4.2.	Screening of MBP citrine constructs.....	83
4.4.3.	Screening of MBP moxY constructs and purification pretests.....	84
4.4.4.	Purification of moxY.....	89
4.4.5.	Crystallization .....	98
4.5.	Conclusion .....	99
5.	Chapter III: Expression and purification of PAMO.....	100
5.1.	Background.....	100
5.2.	Experimental.....	100
5.2.1.	Cloning, cultivation and purification .....	101
5.2.2.	Crystallization .....	103
5.3.	Results and discussion .....	103
5.3.1.	Purification of PAMO .....	103
5.3.2.	Crystallization .....	106
5.4.	Conclusion .....	106
6.	Chapter IV: Expression and purification of BVMO24.....	107
6.1.	Background.....	107
6.2.	Experimental.....	108
6.2.1.	Cultivation and purification .....	108
6.2.2.	Crystallization .....	109
6.3.	Results and discussion .....	110
6.3.1.	Purification .....	110
6.3.2.	Crystallization .....	114
6.4.	Conclusion .....	114
7.	References.....	116
8.	Figures.....	123
9.	Tables.....	126
10.	Appendix .....	127
A.	Sequences of CBH2.....	127
B.	Sequences of citrine.....	128
C.	Sequences of moxY long.....	128
D.	Sequences of MBP and MBP gene optimization.....	129
E.	Results of MS - analysis of MBP moxY constructs .....	134
F.	Vector charts: .....	137
G.	Strain collection.....	144

## Abbreviations

96 DWP	96-well footprint deep well plate
AFB <sub>1</sub>	aflatoxin B <sub>1</sub>
Ala	alanine
AOX	alcohol oxidase gene
AOX1 TT	transcription terminator of the <i>AOX1</i> gene of <i>Pichia pastoris</i>
Arg	arginine
ARG4	argininosuccinate lyase gene 4
ATP	adenosine triphosphate
AVR	averufin
BMD	buffered minimal glucose medium
BMM	buffered minimal methanol medium
BV	Baeyer - Villiger
BVMO	Baeyer - Villiger monooxygenase
BVMO24	Baeyer - Villiger monooxygenase number 24
CBH2	gene for 1,4- $\beta$ -cellobiohydrolase
CO <sub>2</sub>	carbon dioxide
CV	column volume
ddH <sub>2</sub> O	double distilled water
dH <sub>2</sub> O	deionized water
DHA(S)	dihydroxyacetone (synthase)
DKR	dynamic kinetic resolution
DMSO	dimethylsulfoxid
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid disodium salt dihydrate
ER	endoplasmic reticulum
eq	equivalent
EtBr	ethidiumbromide
EtOH	ethanol
FAD	flavin adenine dinucleotide
FBP	fructose 1,6-bisphosphate
FMN	flavin mononucleotide
FMO	flavin monooxygenase
FPLC	fast protein liquid chromatography
G6P	glucose-6-phosphate
G6PDH	glucose-6-phosphate dehydrogenase
GAP	glyceraldehyde-3-phosphate
GBP	glycerol batch phase
GFP	glycerol fed batch phase
Glu	glutamic acid
HAPMO	4-hydroxyacetophenone monooxygenase
HEPES	4-(2-hydroxyethyl)1-piperazineethanesulfonic acid
His	histidine
HVN	hydroxyversicolorone



LB	Luria Bertani
Lys	lysine
m.c.	multi copy
MAD	multiple wavelength anomalous dispersion
mAU	milli absorption units
MBP	maltose binding protein
MeOH	methanol
MES	2-(N-morpholino)ethanesulfonic acid
MFP	methanol fed batch phase
MGY	minimal glycerol medium
MS	mass spectrometry
MUT	methanol utilization
Mut <sup>-</sup>	methanol utilization minus phenotype
Mut <sup>+</sup>	methanol utilization plus phenotype
Mut <sup>S</sup>	methanol utilization slow phenotype
NAD(P) <sup>+</sup>	nicotinamide adenine dinculeotide (phosphate), oxidized
NAD(P)H	nicotinamide adenine dinculeotide (phosphate), reduced
NMO	N-hydroxylating monooxygenase
NMR	nuclear magnetic resonance
OD	optical density
oePCR	overlap extension polymerase chain reaction
PAMO	phenylacetone monooxygenase
PCR	polymerase chain reaction
PEG	polyethylenglycol
P <sub>i</sub>	inorganic phosphate
PMSF	phenylmethanesulfonyl fluoride
PPB	potassium phosphate buffer
PTDH	phosphite dehydrogenase
RNA	ribonucleic acid
SAD	single wavelength anomalous dispersion
s.c.	single copy
SDS-PAGE	sodium dodecyl sulfate – polyacrylamide gel electrophoresis
SUMO	small ubiquitin - like modifier
TAE buffer	tris base, acetic acid, EDTA buffer solution
TB	Terrific Broth
TCA	tricarboxylic acid
TEV	tobacco etch virus
TMH	trans membrane helix
Tris	tris(hydroxymethyl)-aminomethane
VHA	versiconal hemiacetal acetate
VOAc	versiconol acetate
VONE	versicolorone
WT	wild type
YNB	yeast nitrogen base
YPD	yeast extract peptone dextrose
Zeo	Zeocin <sup>TM</sup>

## 1. Introduction

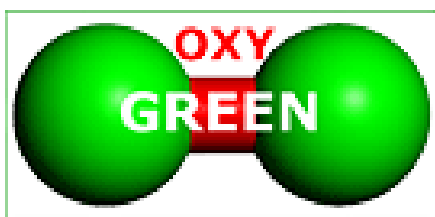
### 1.1. OXYGREEN and Green Chemistry

Green chemistry is defined as:

“The invention, design and application of chemical products and processes to reduce or to eliminate the use and generation of hazardous substances.” (1)

Green chemistry aims at an environmentally friendly way of production and a sustainable development and is gaining more and more importance in the society, the industry and also the academic research. The main principles of green chemistry are to prevent the production of waste, use substances with little or no toxicity, minimize the energy requirements and use, if possible, renewable raw materials and feed stocks. Furthermore, auxiliary substances like for example solvents should be avoided as well as substances with high potentials for accidents, fire or explosion. Preferably the produced substances should not persist in the environment after their use and analytical methods for the control and monitoring need to be optimized before hazardous compounds are formed. Most of the new pathways and methods that facilitate cleaner chemistry involve the use of catalysts. Especially biocatalysts represent a very promising approach. In contrast to chemical catalysts, enzymes often do not require toxic heavy metals and frequently show a superior degree of selectivity (2), (3).

The OXYGREEN project



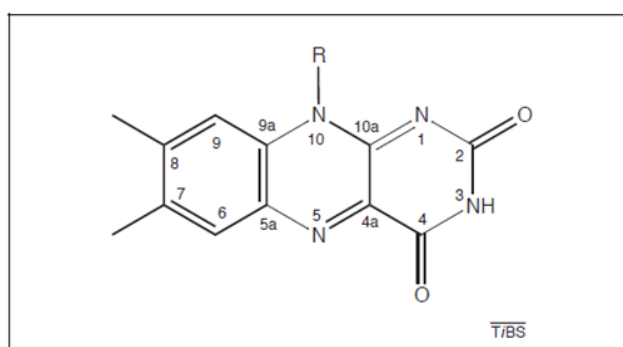
The aim of the OXYGREEN project, on which this research work is based, is the effective redesign of various oxidative enzymes for green chemistry. Apart from the laboratories in Graz and Pavia, several other research groups in The Netherlands, Germany, Poland and France participate at this European project funded by the EC. The focus lies on oxygenating enzymes to catalyze the specific and selective incorporation of oxygen to allow the clean synthesis of bioactive compounds. These substances are highly valuable and applied in different areas like medicine, agriculture and food industry. The enzymes should be designed in such a way, that the principles of green chemistry are fulfilled. Special interest lies thereby on three enzyme classes: cytochrome P-450 monooxygenases, non - heme iron dioxygenases and Baeyer - Villiger monooxygenases (BVMOs) (4).

For this thesis, BVMOs, which are of high interest for industrial and biocatalytic applications, are investigated in detail. They facilitate the fast and selective production of intermediates for biologically active compounds from a large variety of substrates with good yields. The reactions can be performed under environmentally friendly conditions in aqueous solutions and at moderate temperatures and pH values. In organic synthesis, on the other hand, strong oxidizing agents are often needed with lower degrees of regio-, chemo- and enantioselectivity (5). The implementation of these biocatalysts saves time, material and money and is therefore not only due to environmental but also because of economic reasons a highly interesting approach for new ways of industrial production.

This work is focused on three different BVMOs which depend on the cofactor flavin adenine dinucleotide (FAD) and are consequently classified as flavoproteins. These enzymes are very well suited for applications in green chemistry.

## 1.2. Flavoproteins

Flavin-dependent enzymes are able to catalyze a large variety of reactions and are therefore a highly valuable class of enzymes. Apart from electron transfer reactions and dehydrogenations they are for example also involved in halogenations, DNA repair, light sensing and emission as well as protein folding (6), (7). Depending on the type of reaction, the reactivity of the reduced flavin towards molecular oxygen can vary significantly, making flavoenzymes efficient biocatalysts for all organisms. Around 1 - 3 % of prokaryotic as well as eukaryotic genes are expected to encode flavin adenine dinucleotide (FAD) or flavin mononucleotide (FMN) binding proteins. The chemical structure of the isoalloxazine ring, the main reactive group of the flavin cofactors, is shown in Figure 1. (6)



**Figure 1: The chemical structure of the reactive group, the isoalloxazine ring, of the flavin cofactors FAD and FMN (6).**

Flavoproteins belong to the enzyme family of oxidoreductases, which are a diverse group of proteins catalyzing for example the oxygenation of C-H, C-C and C=C bonds as well as the

removal or addition of hydrogen atom equivalents.  $\text{NAD}^+$  and  $\text{NADP}^+$  are coenzymes that are often required by oxidoreductases for example for carbonyl reduction or alcohol oxidation. A subclass of oxidoreductases are oxygenases, which can further be divided into mono- and dioxygenases, depending on whether they incorporate one or two oxygen atoms into the substrate (8). Their Enzyme Commission (EC) numbers are 1.13 and 1.14.

### 1.2.1. Monooxygenases

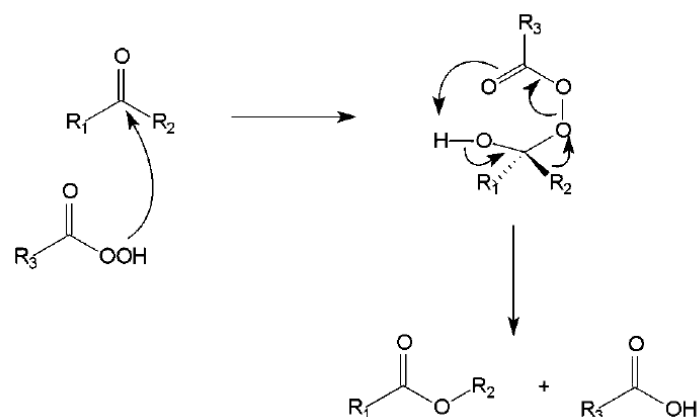
In this thesis, the focus lies on monooxygenases which use flavin-cofactors to catalyze Baeyer - Villiger oxidations or heteroatom oxidations. Apart from flavin, monooxygenases can also possess other prosthetic groups. Hydroxylation or epoxidation reactions for example are facilitated by metal-dependent, cytochrome P-450 monooxygenases (9), (10). In general, monooxygenases can be subdivided in six different groups according to the type of cofactor they use. Next to heme and flavin also copper, non-heme iron and pterin-dependent monooxygenases are known. Furthermore, other cofactors, like S-adenosyl-L-methionine, can be used by monooxygenases, although they are uncommon. Even cofactor independent monooxygenases, that do not require any cofactor, are known (5).

Next to so called external monooxygenases, in which the electrons required for the activation of oxygen or the reduction of the cofactor are derived from external donors (like NADPH), also internal monooxygenases exist. The latter class is quite rare and receives the electrons from the substrate itself. An example for an internal, cofactor independent monooxygenase is lactate monooxygenase. They need only oxygen for their activity and use the substrate as reducing agent. In this way, expensive cofactors and coenzymes can be avoided but the substrate acceptance of these enzymes is limited (5), (11).

## 1.3. Baeyer - Villiger Oxidations

### 1.3.1. Chemical Baeyer - Villiger reactions

In 1899 the conversion of ketones into esters and of cyclic ketones into lactones has been detected by Adolf Baeyer and his student Victor Villiger (12), (13). In the Baeyer - Villiger (BV) reaction a peroxy acid acts as a nucleophile and attacks the carbonyl group of the ketone [see Figure 2]. This leads to the formation of the so called Criegee intermediate. Upon rearrangement of this unstable, tetrahedral intermediate species, which includes the migration of a C-C bond, the corresponding ester and acid are obtained as products (13).



**Figure 2: Reaction mechanism of the Baeyer - Villiger oxidation (13).**

Advantages of this reaction are that various functional groups are tolerated so that not only ketones but also other carbonyl groups can be converted; phenols are for example produced from benzaldehydes and anhydrides from carboxylic acids. Furthermore, the migration preference and regiochemistry as well as the stereoselectivity can be predicted. Additionally, many different oxidants are applicable (14).

On the other side, however, serious problems are connected with the application of BV reactions. First of all, organic peracids are hazardous, shock sensitive, difficult to transport or store and expensive. During the reaction, one equivalent of the corresponding carboxylic acid salt is formed as waste and has to be disposed of. What's more, halogenated solvents are used in some cases, which are detrimental to the environment. Considering these drawbacks on the one hand and the numerous reactions and applications permitted by BV oxidations on the other hand, it becomes obvious that a 'greener' way of performing this reactions is needed (13), (14).

### 1.3.2. Enzymatic Baeyer - Villiger Reactions

A very promising approach for solving the problems connected with chemical BV oxidations consists in applying monooxygenases to catalyze this type of reactions. Additionally, they provide superior chemo-, enantio- and regioselectivity (13). Biological BV oxidations were discovered for the first time in 1948 by Turfitt during degradation of steroids by fungi (15). Apart from degradation pathways, BVMOs are also involved in biosynthetic metabolism steps. Examples are the fungal aflatoxin biosynthesis, the synthesis of toxins in shellfish or of steroids and iridoids in plants (13), (9).

## Classification and identification of BVMOs

Intensive studies of numerous sequences that encode for BVMOs revealed a common sequence motif, FXGXXXHXXXW(P/D). This amino acid fingerprint allows the identification of BVMOs and the analysis of protein sequence databases. Furthermore, it distinguishes BVMOs from the two other classes of FAD-dependent monooxygenases, the flavin containing monooxygenases (FMO) and N-hydroxylating monooxygenases (NMO) (16). FMOs are eukaryotic enzymes that oxygenate various amines or other heteroatom compounds and are well suited for the detoxification of for example organic sulfur xenobiotics, nicotine or alkaloids (17). The sequence motif typical for FMOs is very similar to that of BVMOs as it varies only by the last residue, FXGXXXHXXX(Y/F). In NMOs, on the other hand, only the histidine residue, which seems to play an essential role in the catalysis of all FAD dependent monooxygenases, is strictly conserved (16). Microbial NMOs facilitate N-hydroxylations of primary, long chain amines and show a relatively low affinity to FAD (18).

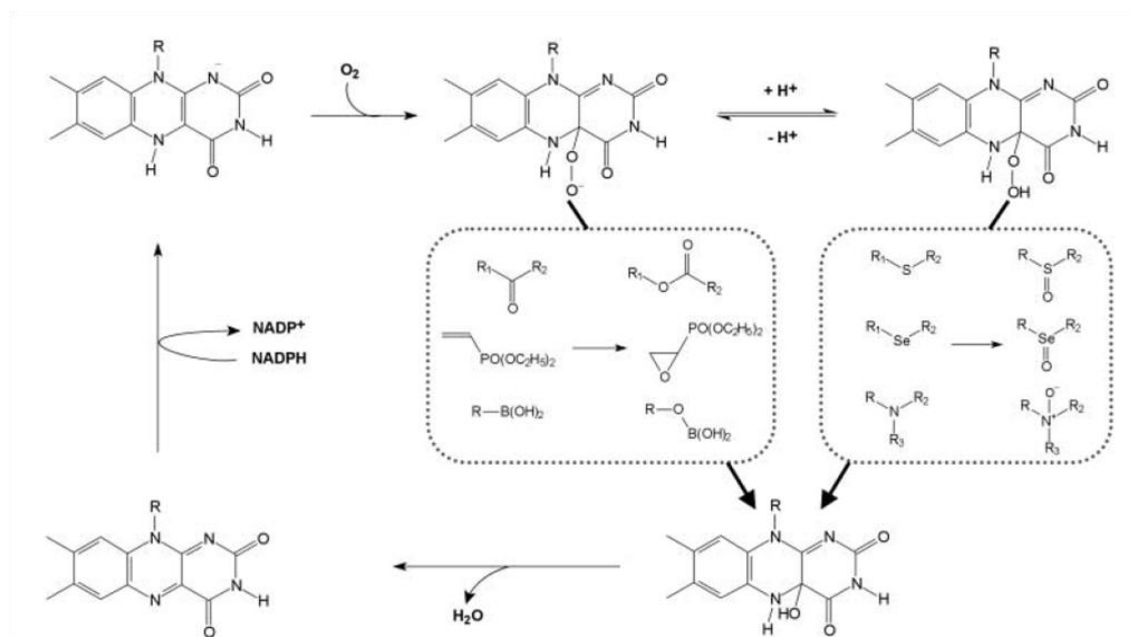
Genome mining using the specific BVMO protein sequence motif revealed a diverse distribution of these enzymes among organisms. Especially bacteria are with an average of one BVMO per genome very rich in BVMOs. Type I BVMOs are most common within the BVMO family. They occur in bacteria and fungi but in the genomes of humans, plants and animals, no Type I BVMO has been detected. These monooxygenases use tightly bound FAD as cofactor and depend on NADPH. They consist of a single polypeptide chain and contain two Rossmann motifs with the sequence GxGxxG, which represent the two dinucleotide binding sites for FAD and NADPH (19). Apart from producing esters or lactones by inserting oxygen into the corresponding ketones, Type I BVMOs can also oxidize for example nitrogen, sulfur, and boron atoms (20), (13).

Type II BVMOs, however, utilize the cofactor FMN and NADH as electron donor. Furthermore, they are less widespread and consist of  $\alpha_2\beta$  trimers (21). The only sequence that was available for *in silico* studies at NCBI was the sequence from limonene monooxygenase, which revealed only 12 bacterial sequences and 5 sequences of the Sargasso Sea database with a sequence identity > 40 % (19).

According to recent studies, additionally to these two classes other, new types of BVMOs exist. Sometimes they are referred to as atypical or Type 0 BVMOs. One example is the bacterial flavoprotein monooxygenase MtmOIV, which is involved in the biosynthesis of mithramycin, an important pharmaceutical substance as it is used as an anticancer drug and calcium lowering agent (19), (11). Furthermore eukaryotic P450 enzymes with Baeyer - Villiger activity are known, for example a heme-containing BVMO of the cytochrome P450 superfamily was discovered in plants, which converts a specific plant steroid (19).

## Mechanism

The mechanism for the enzymatic oxygenation of organic substrates is quite complex. In order to facilitate the otherwise spin forbidden reaction between molecular oxygen and organic carbon substrates, oxygen must be activated first (22). In the case of flavin-dependent monooxygenases this is achieved by the formation of a flavin C4a-(hydro)peroxide, which subsequently allows the insertion of an oxygen atom into the substrate (6).



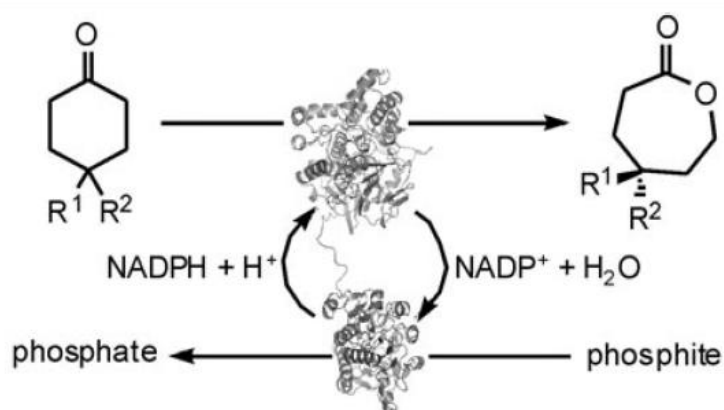
**Figure 3: The mechanism of BVMOs using flavin as cofactor for oxidations (13).**

The first step in the catalytic process of BVMOs is the reduction of the protein bound FAD by NADPH [see Figure 3]. Thereby, a reduced, binary enzyme-NADP<sup>+</sup> complex is formed which reacts with oxygen. The produced flavin peroxide acts as a nucleophile and attacks the carbonyl group of the substrate, thereby forming a Criegee intermediate, as it is found in the chemical mechanism. Through the rearrangement of this intermediate, the corresponding products, like esters or lactones, are formed as well as a flavin hydroxide. After the elimination of water the oxidized FAD is obtained again and NADP<sup>+</sup> is released. BV reactions, boron oxidations and epoxidations involve this type of nucleophilic peroxyflavin. The electrophilic hydroperoxyflavin, on the other side, facilitates sulfoxidations of thioethers and catalyzes the oxidation of amines or selenides (13).

## Coenzyme regeneration

Since NAD(P)H is essential for the activity of BVMOs, but too expensive to be applied stoichiometrically, it is necessary to regenerate the redox coenzyme (23). When recombinant whole-cell conversions are performed, the cellular machinery can be exploited to recycle nicotinamide using cheap substrates as for example glucose. However, whole cell concepts can have several drawbacks like unwanted side reactions or lower activities per unit dry weight (5). Furthermore, they are not always applicable, for instance when the used substrates or products are toxic for the cell or cannot pass the membrane. In that case isolated BVMO enzymes are applied, which are often coupled to a second enzymatic system for the regeneration of NADPH (23).

For the enzymatic coenzyme regeneration dehydrogenases are mainly used, which reduce  $\text{NADP}^+$  while simultaneously oxidizing a suited co-substrate, like glucose or formate. Various enzymatic regeneration systems exist, applying for example an alcohol dehydrogenase, formate-, glucose- or glucose-6-phosphate dehydrogenase (5). Recently, phosphite dehydrogenase (PTDH) was connected to BVMOs as a fusion partner. During the oxidation of phosphite to phosphate,  $\text{NADP}^+$  is reduced by PTDH [see Figure 4]. It turned out, that the fusion of PTDH facilitates the regeneration of NADPH without altering the catalytic properties, the stereoselectivity or the thermostability of the BVMO (23)



**Figure 4: Mechanism of a BVMO (enzyme above) and a fused PTDH (below) for the regeneration of NADPH (23).**

Another option is the chemical regeneration of NADPH, for example by using the rhodium complex  $[\text{Cp}^*\text{Rh}(\text{bpy})(\text{H}_2\text{O})]^{2+}$  as catalyst. In this case formate can be used as a cheap substrate, which is oxidized to volatile  $\text{CO}_2$ , thereby providing the electrons which are transferred to  $\text{NADP}^+$  by the catalyst. This rhodium complex is also applicable for photochemical and electrochemical ways of NAD(P)H regeneration. Especially the photochemical regeneration is very environmental friendly, as it only requires the light absorption of a photosensitive material to induce electron transfers (5). An interesting coenzyme regeneration system was investigated for the oxidation of ketones by a mutant of

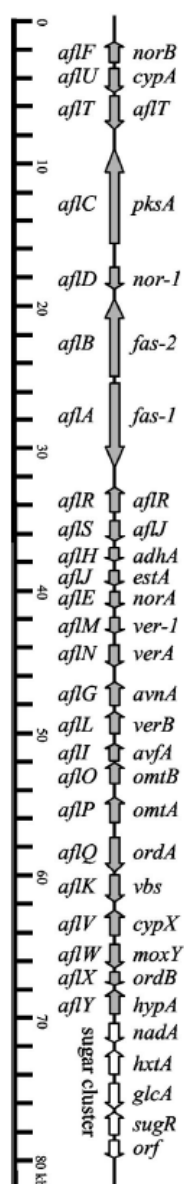


the phenylacetone monooxygenase (PAMO P3) using solar light. In that case, EDTA was applied as electron donor and an additional flavin was used as electron mediator. Adding small amounts of NADP<sup>+</sup> allowed successful conversions with very good enantioselectivities (23).

### 1.3.3. *moxY*

Characteristics of this enzyme

Analysis of the aflatoxin biosynthesis gene cluster in *Aspergillus flavus* and *Aspergillus parasiticus* revealed that several oxidative steps and at least two monooxygenases are involved in this pathway. These two monooxygenases are *cypX*, a cytochrome P-450 monooxygenase and *moxY*, a monooxygenase capable of catalyzing a Baeyer - Villiger oxidation (24), (25).

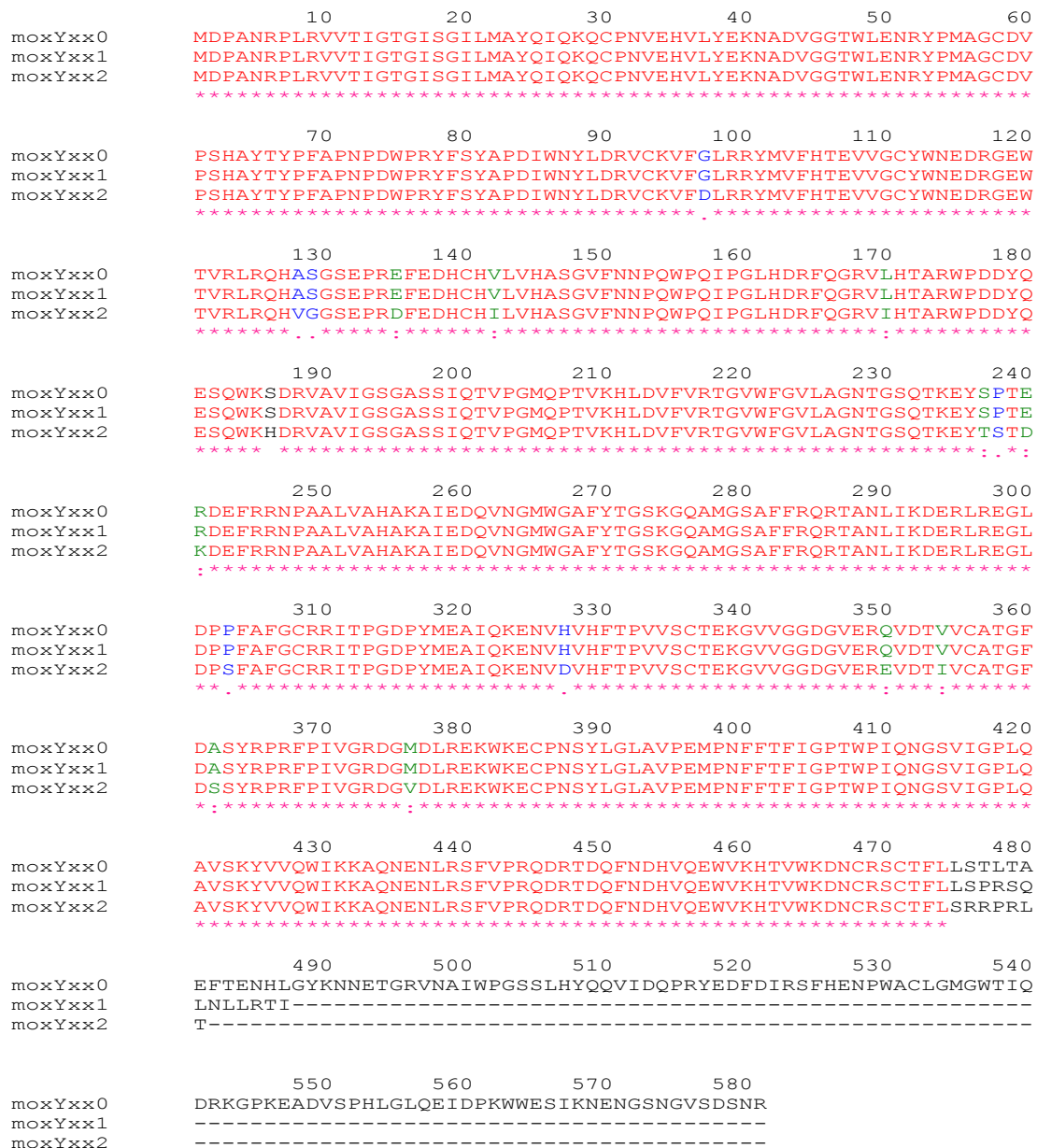


The aflatoxin pathway gene cluster consists of 25 genes or open reading frames within a 70-kb DNA sequence. Two of the genes, *aflR* and *aflS* (*aflJ*) are involved in the regulation of the pathway. *aflR* is a positive regulatory gene that is responsible for the activation of the transcription. The genes involved in aflatoxin biosynthesis were originally named according to their substrate or their enzymatic function. In order to ensure uniformity, a new system for the nomenclature of the genes was suggested, starting with the three letter code “*afl*” to indicate the aflatoxin production and a capital letter from A to Y for all 25 identified genes [see Figure 5]. According to this nomenclature, *moxY* corresponds to *aflW* and is homologous to the gene *stcW* of the sterigmatocystin pathway gene cluster in *Aspergillus nidulans* (26), (24). The homology between *stcW* and *moxY* at the genomic DNA level is significantly lower (about 40%) than the amino acid sequence identity (69%) (24). However, in this case the name *moxY* is used throughout the whole thesis and not changed to *aflW*.

The *moxY* gene identified in *A. parasiticus* contained a 1446 bp long coding sequence without introns that encode a 481 amino acid long protein. The typical AFLR binding motif was found in the promotor as well as some possible TATA boxes. The results of a BLAST search revealed homologies of more than 50% with monooxygenases (24).

**Figure 5: The aflatoxin pathway gene cluster with the old gene names on the right and the new names on the left side. The transcription direction is indicated by the arrows and the size of the genes in kilobases (kb) is shown by the ruler on the left. The aflatoxin biosynthesis cluster is followed by a sugar utilization cluster (26).**

In the NCBI database, two different sequences were found for *moxY* from *A. flavus*, one with 487 amino acid residues, which was called *moxY\_short* and one with 581 amino acid residues named *moxY\_long* (see <http://www.ncbi.nlm.nih.gov>). Sequence alignment of both proteins together with the *E. coli* codon optimized sequence of *moxY\_IT* from *A. parasiticus* (obtained from the research group of Mattevi in Pavia, unpublished experiments), is shown in Figure 6.



**Figure 6: Alignment of various *moxY* protein sequences. *moxYxx0* stands for the so-called *moxY\_long* from *A. flavus*, *moxYxx1* stands for *moxY\_short* from *A. flavus* and *moxYxx2* shows the sequence of *moxY\_IT* from *A. parasiticus*. Red letters indicate that all three proteins have the same amino acid residue at this position, whereas deviations within the sequence are shown in black, blue or green.**

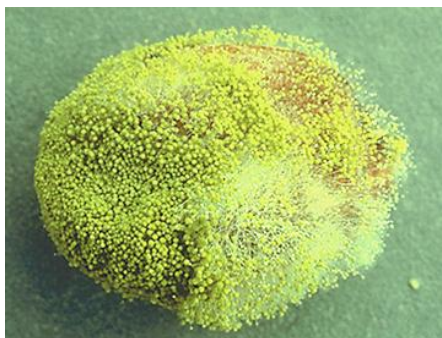
The present study focuses on the expression of *moxY\_long* which has a molecular mass of 66.2 kDa. The gene sequence of *moxY\_long* was isolated from *A. flavus* whereas it was not possible to isolate *moxY\_short*. In preliminary studies Martina Geier discovered that *moxY\_short* may be caused by possible sequencing mistake (personal communication with Martina Geier). Comparison of the DNA sequences of *moxY\_long* and *moxY\_short* revealed the insertion of a T and C at position 1428 in the case of *moxY\_short* [see **Figure 7**]. This insertion may cause a frame shift, resulting in an earlier stop codon and therefore shorter protein sequence.

	1301				1350
long	AGAATGAGAA	TCTCCGTAGC	TTCGTGCCGC	GACAGGACCG	CACGGATCAA
short	AGAATGAGAA	TCTCCGTAGC	TTCGTGCCGC	GACAGGACCG	CACGGATCAA
Consensus	AGAATGAGAA	TCTCCGTAGC	TTCGTGCCGC	GACAGGACCG	CACGGATCAA
	1351				1400
long	TTCAACGATC	ATGTCCAGGA	GTGGGTGAAG	CACACGGTGT	GGAAAGATAA
short	TTCAACGATC	ATGTCCAGGA	GTGGGTGAAG	CACACGGTGT	GGAAAGATAA
Consensus	TTCAACGATC	ATGTCCAGGA	GTGGGTGAAG	CACACGGTGT	GGAAAGATAA
	1401				1450
long	CTGCCGAAGC	TGTACGTTCC	TTCTCTC..C	ACGCTCACAG	CTGAATTTAC
short	CTGCCGAAGC	TGTACGTTTC	TTCTCTCTCC	ACGCTCACAG	CTGAATTTAC
Consensus	CTGCCGAAGC	TGTACGTTcC	TTCTCTC..C	ACGCTCACAG	CTGAATTTAC

**Figure 7: Alignment of the DNA sequences of *moxY\_long* and *moxY\_short*. The insertion of a T and C in the sequence of *moxY\_short* is highlighted in blue.**

## Aflatoxins

Aflatoxins are toxic secondary metabolites that are derived from polyketides and mostly produced by the molds *A. parasiticus* and *A. flavus* (25), (27). They occur on various crops and agricultural commodities, pre- as well as post-harvest. These contaminations can lead to serious economic problems and difficulties concerning the food safety (26). For humans and



animals aflatoxins are carcinogenic and can lead to immunotoxicity, teratogenicity and even death (25). The most common aflatoxins in nature are B<sub>1</sub>, B<sub>2</sub>, G<sub>1</sub> and G<sub>2</sub>. The aflatoxin B<sub>1</sub> (AFB<sub>1</sub>) is one of the most potent natural carcinogens and contaminations of feed crops and food like corn, cotton, peanuts or wheat led to extensive studies about this mycotoxin (28), (25).

**Figure 8: Growth of an *Aspergillus* species on a peanut.**

## Role of *moxY* in the aflatoxin biosynthesis

The production of aflatoxins involves the following substances: acetate → polyketides → anthraquinones → xanthenes → aflatoxins (26). In order to identify the detailed role of *moxY* within this biosynthesis pathway, mutants were developed, in which the *moxY* gene in *A. parasiticus* was deleted via double crossover recombination. Instead of producing aflatoxins, this mutant accumulated hydroxyversicolorone (HVN) and versicolorone (VONE). Feeding experiments showed, that while *moxY* deleted mutants were not able to convert averufin (AVR), HVN and VONE into aflatoxins, it was able to produce aflatoxins from versiconal hemiacetal acetate (VHA) and versiconol acetate (VOAc). Therefore it can be concluded, that *moxY* acts as a monooxygenase which catalyzes the conversion of HVN to VHA as well as from VONE to VOAc [see Figure 9]. These two reactions are Baeyer - Villiger oxidations since they are characterized by the insertion of an oxygen atom into a C-C bond next to the carbonyl group of a ketone. Consequently, *moxY* was identified as a BVMO which are known to use FAD as cofactor (25), (29).

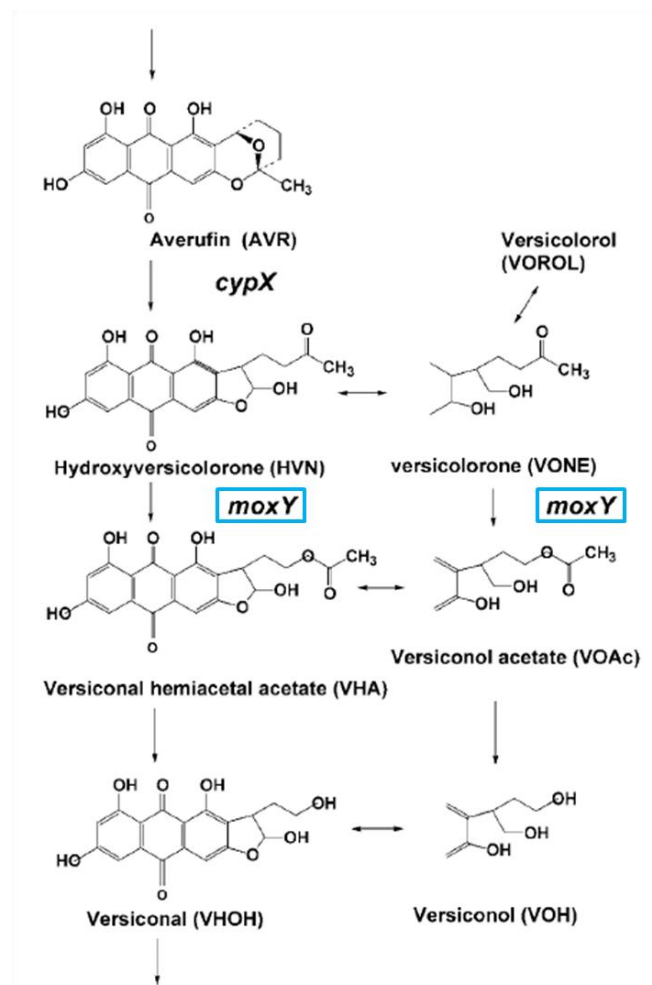


Figure 9: Parts of the metabolic pathway for the biosynthesis of aflatoxins in *Aspergillus* species, including the reaction catalyzed by *moxY* (25).

#### 1.3.4. Phenylacetone monooxygenase (PAMO)

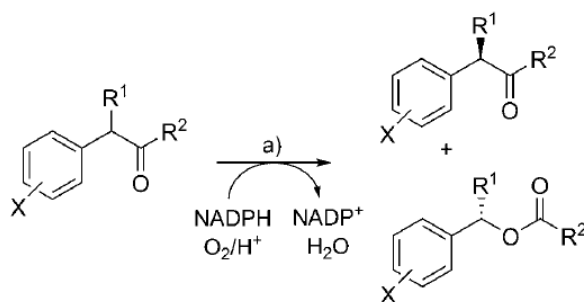
Sequence alignment studies of several bacterial BVMOs by M.W. Fraaje *et al.* in 2002 revealed a conserved motif for this enzyme class (16). With the help of this specific protein sequence motif a BVMO was identified in the genome of the actinomycete *Thermobifida fusca*, a moderately thermophilic soil bacterium usually growing at 50-60 °C. Although this enzyme showed sequence identities of 53 % with steroid monooxygenase and 40 % with cyclohexanone monooxygenase, it did not convert steroids like progesterone, but mainly aromatic ketones as well as some aliphatic ketones and sulfides. The highest activity was obtained when phenylacetone was used as substrate and converted to phenylacetate, therefore this BVMO was named phenylacetone monooxygenase (30).

Hints about the physiological function of PAMO were obtained by investigation of the neighboring genes. As many other BVMOs, PAMO seems to be involved in a degradation pathway since the enzyme is flanked by a gene encoding an esterase, which can hydrolyze the esters produced by the BVMO (19). Furthermore, a putative alcohol dehydrogenase and reductase are located downstream of *pamo*. The dehydrogenase may produce the aromatic ketones which are consequently converted by PAMO (30). Recently it was proposed that *T. fusca* is involved in the degradation of lignin, indicating that the aromatic compounds which are degraded by PAMO and the surrounding enzymes may be derived from lignin (23).

#### Applications of PAMO

The aim of most oxidation reactions catalyzed by BVMOs is to produce biologically active, high value compounds - either via intermediate building blocks or directly. In the pharmaceutical industry BVMOs are applied for the synthesis of drugs, for example for modifications of steroids, which are used as anticancer or antihypercholesterolemic compounds (23).

Optically active compounds, which can be obtained by resolution of racemic substrates, are important in organic synthesis. PAMO was found to be able to perform kinetic-resolution processes of racemic benzylketones with excellent enantioselectivities, producing the (*S*)-benzylesters while leaving the (*R*)-benzylketones unaltered [see Figure 10]. Even at substrate concentrations close to 20 g/L the selectivity of PAMO for the oxidations remained high (23), (31).

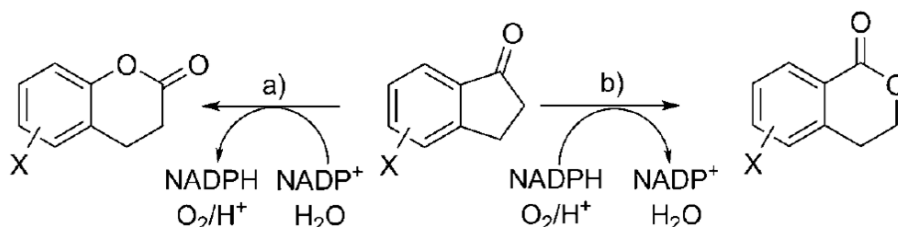


**Figure 10: Kinetic resolution of racemic benzylketones, yielding optically active (*S*)-benzylesters and (*R*)-benzylketones. a) 50 mM Tris/HCl, PAMO, G6P/G6PDH (23).**

A disadvantage of this method is that the theoretical yields of the substrate as well as product are limited to 50%, but this can be overcome via a dynamic kinetic resolution (DKR) process, with *in situ* racemisation of the substrate. By applying DKR and optimization of the reaction media via addition of water-immiscible solvents, the enantioselectivity could be modified and enantiopure (*R*)-cyanoesters were produced by PAMO (23).

Furthermore, PAMO facilitates the synthesis of chiral aromatic sulfoxides, by oxidation of for example alkyl benzyl sulfides. The degrees of selectivity as well as the enantioselectivity vary significantly depending on the type of substrate (32). In organic chemistry chiral sulfoxides are important auxiliaries in asymmetric synthesis and enantioselective catalysis and used for the preparation of biologically active compounds (23). PAMO is also capable of performing nitrogen and boron oxidations (32).

The fact that PAMO can tolerate organic solvents implies many advantages. Organic media can improve the solubility of hydrophobic substrates and recycling of the enzyme, avoid side reactions and increase the regio-, chemo- or enantioselectivity. Depending on the biocatalyst and the co-solvent added to the reaction media the oxidation of for example 1-indanone to both regioisomeric lactones can be achieved [see Figure 11] (23).



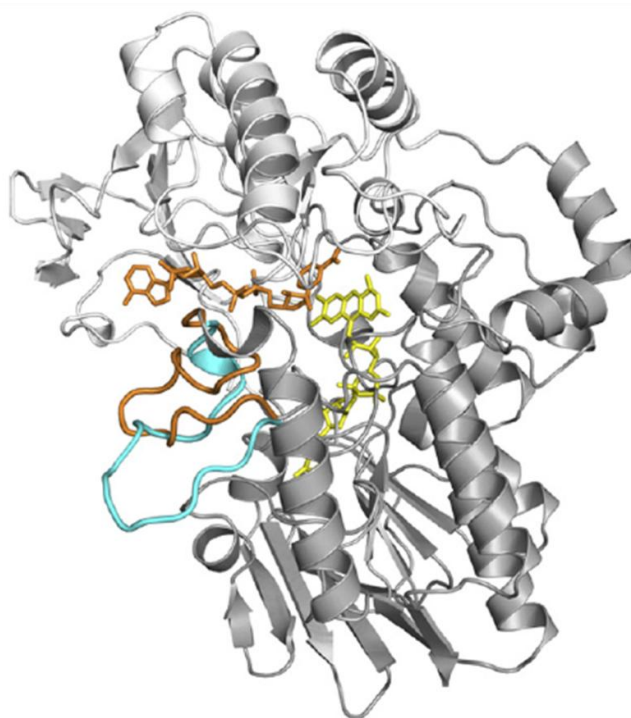
**Figure 11: The regiodivergent BV oxidation of substituted 1-indanone using a) HAPMO with 5% hexane as co-solvent or b) the PAMO mutant M446G with 5% methanol as co-solvent (23).**

Another interesting phenomenon is the inversion of the enantioselectivity upon addition of different co-solvents, which was observed for example for the oxidation of ethyl phenyl sulfides. When 30% methanol was added, the (*R*)-sulfoxides were obtained instead of the (*S*)-enantiomers (23).

### Structure and mechanism of PAMO

The crystal structure of PAMO was resolved in 2004 by the group of Mattevi and was the first three dimensional structure of a BVMO that was described [see Figure 12]. For structure analysis, the recombinant protein was expressed in *E. coli*, purified and crystallized using the vapor diffusion method (33).

PAMO consists of two domains with the characteristic dinucleotide binding fold which are responsible for the binding of FAD and NADP, respectively. The active site is located at the bottom of the cleft in between these two domains, where the flavin ring binds (33).



**Figure 12: Crystal structure of PAMO in complex with NADP<sup>+</sup> (orange). The NADP binding domain is depicted in light gray, the FAD binding domain in gray and the flavin in yellow. The loop conformation (residues 495-515) of the unligated enzyme is shown in cyan, whereas the rearranged loop of the NADP<sup>+</sup> bound enzyme is shown in orange (34).**

Recently, the complex mechanism of the catalysis by BVMOs was analyzed in detail using microspectrophotometric, kinetic and structural investigations. Crystals of PAMO in complex with NADP<sup>+</sup> could be obtained and revealed that the main difference between the ligand free and the NADP<sup>+</sup> bound structure consists in a shift of the loop of the FAD binding domain (amino acid residues 495 - 515). This shift enables the interaction with NADP<sup>+</sup> and narrows

the cleft between the two domains. As a result, the funnel shaped substrate binding site is defined by NADP<sup>+</sup> and the surrounding protein residues. Since there are no specific substrate recognition sites, the active site seems to be mainly architected for the stabilization of the flavin peroxide to activate oxygen and to stabilize the Criegee intermediate. Consequently, the substrate specificity, which seems to depend on the ability of chemically suited substrates to diffuse to the active site, can be modified quite easily by mutations and provide immense potentials in biocatalytic applications (34).

Based on the structure of PAMO, enzyme engineering methods were tested in order to improve the substrate spectrum and enantioselectivity. This was for example achieved by site-directed mutagenesis experiments resulting in the M446G mutant, which is able to convert different, new ketones as well as indigo. Met-446 is located in a loop, which is crucial for substrate specificity, at the bottom of the substrate site (23), (34), (35). In deletion mutants the active site was enlarged and the substrate acceptance could consequently be expanded towards bulkier ketones (23).

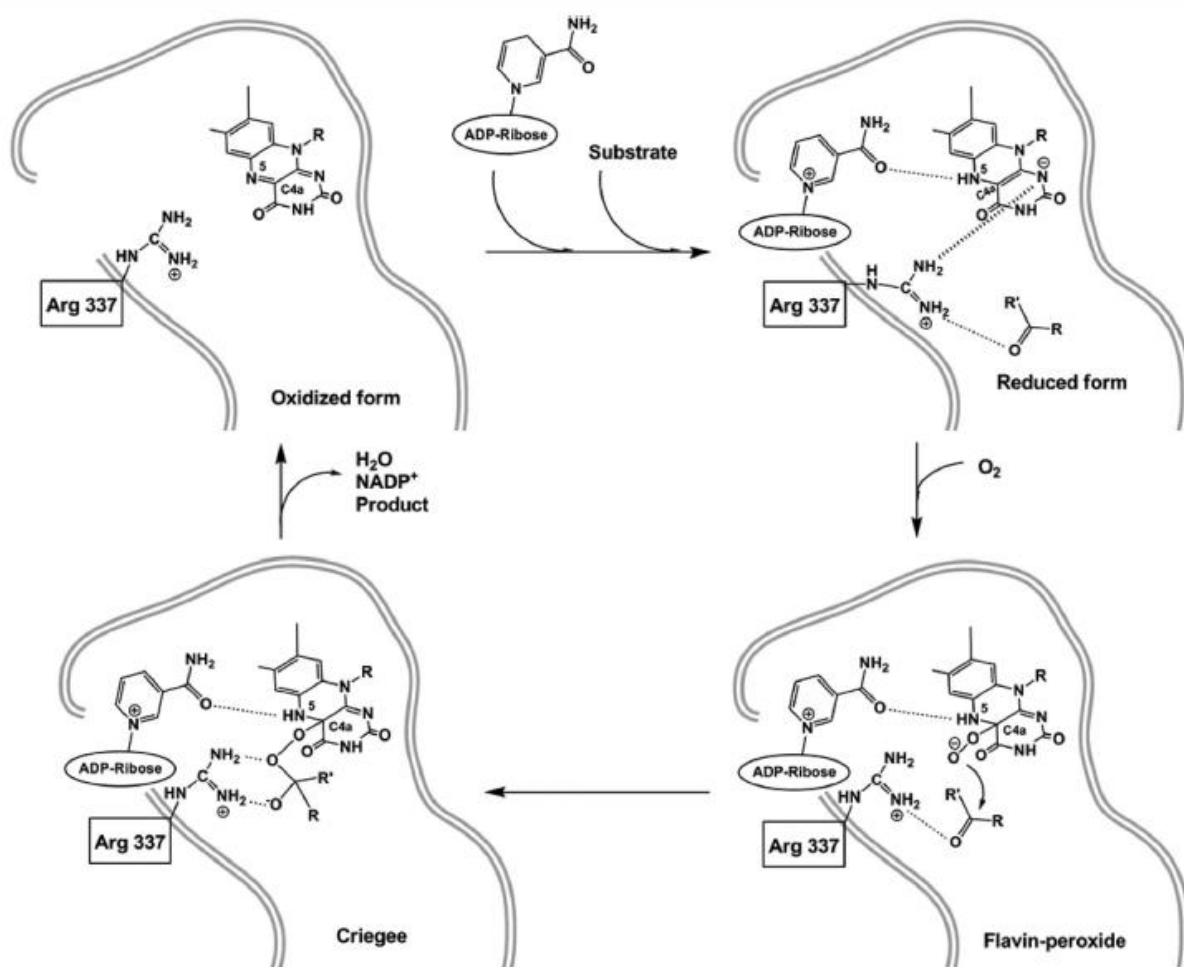


Figure 13: Detailed look at the active site of PAMO during the catalysis of a Baeyer Villiger reaction (34).



With the help of crystals of PAMO in complex with NADP<sup>+</sup>, the role of NADP<sup>+</sup> and the residue Arg-337, which are both essential elements for the catalytic mechanism, was investigated [see Figure 13]. The fundamental role of Arg-337 could be demonstrated by mutagenesis experiments. Arg-337 mutants are no longer catalytically active since they are not able to insert an oxygen atom into the organic substrate.

When the enzyme is in the oxidized form, this Arg residue forms H-bonds with the carboxamide group of NADP<sup>+</sup> and the residue Asp-66. In the reduced form, Arg-337 is rearranged and shifts from NADP<sup>+</sup> towards the flavin ring. An electrostatic interaction between the positive guanidinium group of Arg and the negative charge of the reduced flavin cofactor is generated. Simultaneously to the shift of Arg-337, the carboxamide group of NADP<sup>+</sup> is slightly rotated and an H-bond to the N5-atom of the flavin is established. As a result, the protection of the N5 proton of the reduced flavin is ensured, which is required for the stability of the intermediate and consequently for the oxygen activation.

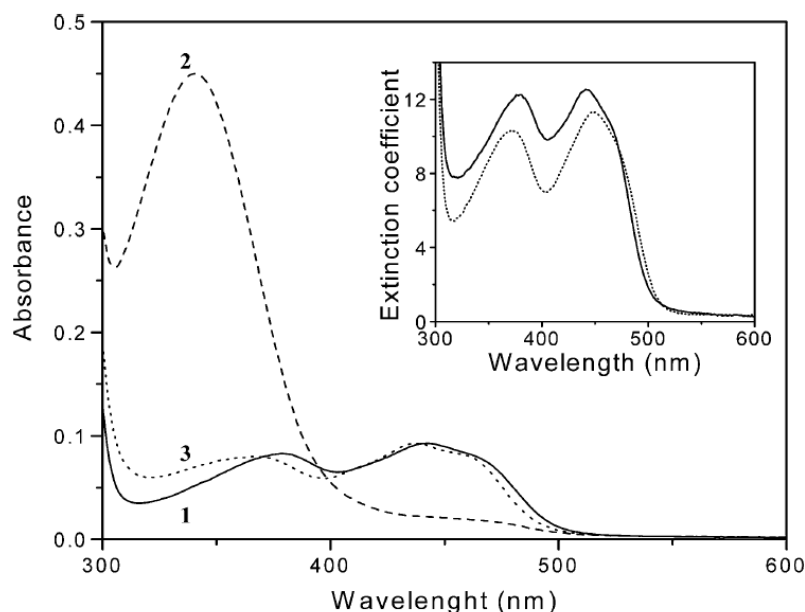
In order to interact with the negatively charged reduced flavin, the conformation of Arg-337 is extended, thereby preventing access of the substrate. Only after reaction with dioxygen and formation of the flavin-peroxide followed by the loss of the negative charge of the flavin, Arg-337 is repositioned to the conformation for interaction with NADP<sup>+</sup>, allowing access of the substrate.

What's more, this Arg-residue is important for the anchoring and proper binding of the ketone substrate, because it establishes an H-bond to the carbonyl oxygen, thereby increasing the tendency for a nucleophilic attack by the peroxy-flavin. This nucleophilic attack of the flavin-peroxide leads to the formation of the Criegee-intermediate. Arg-337 and NADP<sup>+</sup> compensate for the negatively charged Criegee intermediate, further confirming the important catalytic role of these two compounds.

The catalytic cycle is closed by the formation of the product and the release of H<sub>2</sub>O and NADP<sup>+</sup>, so that finally the oxidized form of the enzyme is obtained again (34).

#### Redox activity

PAMO crystals with bound NADP<sup>+</sup> were not only essential for a better understanding of the function of BVMOs, but also used for redox experiments. By measuring the UV-Vis absorbance spectrum of the protein, it could be demonstrated that the crystalline enzyme is redox active. PAMO crystals are very robust, they can be reduced with dithionite or x-rays and can afterwards be reoxidised again (34).



**Figure 14: UV-Vis spectra of the native PAMO (1) and upon reduction with NADPH. Spectrum 2 of the fully reduced flavin was taken 1 min after mixing with NADPH in large excess and spectrum 3 after 30 min. The inset shows the spectra of native PAMO (dark line) and the with SDS unfolded protein (dotted line) (30).**

The UV-Vis spectrum of purified, yellow PAMO shows two absorption maxima at 380 nm and 441 nm, which indicate a non-covalently bound flavin cofactor [see Figure 14]. By adding NADPH in large excess, the flavin cofactor is reduced and an absorption maximum at 340 nm due to NADPH is obtained, while a bleaching of the peak at 441 nm can be observed. The slow oxidation of NADPH by the enzyme results in a flavin spectrum which is not completely identical to the one of native PAMO, but shifted towards lower wavelengths. The extinction coefficient of PAMO at 441 nm ( $\epsilon_{441} = 12.4 \text{ mM}^{-1} \text{ cm}^{-1}$ ) was determined using the flavin spectra of the purified and the unfolded protein [see inset in Figure 14] (30).

### 1.3.5. BVMO24

Genome analysis of the bacterium *Rhodococcus jostii* RHA1 revealed that this organism is especially rich in oxidative enzymes and contains more than twenty putative BVMO genes. Recently, 22 of these Type I BVMOs could be obtained in a soluble form, after optimization of the expression conditions and the vectors. Their substrate profiles, determined with a novel activity screening assay, showed broad substrate scopes for several BVMOs, whereas for others the substrate acceptance was very limited or no activity at all was measured with the tested substrates. A great variety in regio- and enantioselective behavior was detected, depending on the different BVMOs and the substrates (36), (37), (23).

An outstanding candidate of this monooxygenases is the BVMO number 24 (BVMO24), with the accession number ro05323. Together with the other BVMOs it contains two GxGxxG Rossmann fold motifs, but the specific BVMO sequence motif is slightly modified (FxGxxxSxxxWN). Remarkably, instead of the usually conserved histidine residue, BVMO24 contains a serine in the middle of the sequence motif and an asparagine at the end. The protein consists of 564 amino acids.

As all other Type I BVMOs it uses FAD as cofactor and requires NADPH for reduction of the flavin. The conversion of various ketones, amines and sulfides was tested. It turned out, that BVMO24 shows a very broad substrate acceptance (25 of 39 substrates were converted) and good efficiencies. As many others, BVMO24 catalyzes the enantioselective oxidation of aromatic sulfides, but in contrast to all other investigated BMVOs it shows the opposite enantiopreference [see Table 1] (36).

**Table 1: Substrate conversion of various BVMOs from *R. jostii* (given as percentage %); the enantioselectivities are given as enantiomeric excess in % (36).**

Substrate	BVMO							
	3	8	9	14	15	20	21	24
Phenylacetone	99	27	99	99	92	45	99	38
2-Indanone	<3	<3	<3	<3	<3	<3	<3	62
Bicycloheptenone	97	<3	36	97	90	97	97	93
Thioanisole	87	<3	21	99	76	24	55	80
<i>Enantioselectivity</i>	90 (S)	n.d.	61 (S)	n.d.	45 (S)	5 (S)	63 (S)	82 (R)
Benzylethyl sulfide	60	<3	<3	72	14	4	92	91
<i>Enantioselectivity</i>	75 (R)	n.d.	n.d.	90 (R)	20 (R)	n.d.	93 (R)	>99 (S)

Because of the diverse substrate spectrum, the good activity and enantioselectivity, BVMO24 seems to be a very promising candidate for biocatalytic applications.

## 1.4. *Pichia pastoris*

### 1.4.1. Why choosing this expression system?

*P. pastoris* is a methylotrophic yeast and next to *Hansenula*, *Candida* and *Torulopsis*, *Pichia* is one of four genera capable of using methanol as the sole source for carbon and energy (38). This ability was detected for the first time more than forty years ago by K. Ogata (39). In the 1970's the Phillips Petroleum Company (Bartlesville, OK, USA) developed the first protocols and media for the continuous cultivation of *P. pastoris* on methanol to cell densities of > 130 g/L dry cell weight (40). The original aim was to convert abundant methanol into a high protein animal feed. However, as a consequence of the oil crisis and cheap alternative sources like soybeans, the production of single cell protein (SCP) based on methanol was not economical or competitively viable. Nevertheless, extensive studies about the characteristics of *P. pastoris* as an expression system for recombinant proteins were started and resulted in the establishment of *P. pastoris* as a model eukaryote in cell biology and a very efficient system for the expression of heterologous proteins (41), (42). Phillips Petroleum sold in 1993 the patent for the *P. pastoris* expression system to Research Corporation Technologies (Tucson, AZ, USA) and gave the license for selling components of the system to Invitrogen Corporation (Carlsbad, CA, USA) (42), (43).

There are various reasons that explain why *Pichia pastoris* is frequently selected by researchers as the host of choice for producing a recombinant protein. First of all, strong and regulatable promoters are available from genes of the methanol utilization pathway (44). The promoter from the alcohol oxidase I gene allows the tight regulation of the protein expression since induction is only facilitated by the addition of methanol, whereas no protein is expressed during cell growth on glycerol. In methanol grown cells the percentage of AOX can account for > 30 % of the total soluble protein (45). The expression of AOX is controlled at the transcription level and the enzyme is located in the peroxisome. In the peroxisomes hydrogen peroxide, which is toxic for the cells and produced during the oxidation of methanol, is further metabolized (46).

A major advantage of *Pichia* is that this yeast strongly prefers the respiratory growth over the fermentative metabolism. Consequently the production of high levels of ethanol, which occurs during fermentation, can be avoided and very high cell densities (500 OD<sub>600</sub> U/mL) and biomass yields can be reached (42), (43).

Furthermore, the techniques which are used for genetic manipulations of *P. pastoris* at the molecular level are similar to those applied for *Saccharomyces cerevisiae* and therefore well known and characterized (43).

In contrast to bacterial expression systems, *Pichia* is capable of performing post translational modifications like glycosylation, disulfide bond formation, folding or processing of signal sequences, which are typical for higher eukaryotic cells. This allows the production of functional and correctly folded proteins that otherwise end up as inactive inclusion bodies when bacteria like *E. coli* are used for their expression (42), (43). Compared to mammalian or insect cells, *P. pastoris* generally has the potential to produce proteins in a faster, less complex and cheaper way (41).

Next to the efficient production of intracellular proteins, very high levels of secretion can be obtained by using either the native signal peptide of the protein of interest or by adding for example the  $\alpha$  - factor prepro signal sequence from *S. cerevisiae*, which was successfully used in many cases. Since there are only few endogenous proteins secreted by *P. pastoris*, the extracellular expression can already serve as a first purification step and the product can account for > 80 % of the total protein in the medium (47), (42), (43).

Additionally, as a result of the extensive knowledge about this yeast, *P. pastoris* functions as a model system for modern research fields in cell biology and is used to investigate for example the import or degradation of peroxisomes and the secretion pathway in eukaryotes in detail (43).

For the production of pharmaceutical proteins it is important that the products of methylotrophic yeasts do not contain endotoxins and viral or oncogenic DNA. Some yeasts have been used in the production of food since centuries and are in the meanwhile classified as GRAS organisms (Generally Regarded As Safe) (48).

#### 1.4.2. Methanol metabolism

The expression of the enzymes of the methanol utilization (MUT) pathway is tightly regulated via a repression/induction mechanism. The absence of carbon sources like glucose, glycerol or ethanol and the addition of methanol are both required for an effective expression using the AOX promoter (48). The first part of the methanol metabolism takes place in the peroxisomes. Induction with methanol leads to a significant increase of these organelles, so that finally up to 80 % of the cytoplasmic space may consist of peroxisomes (48), (49).

First of all, methanol is oxidized to formaldehyde and hydrogen peroxide in the peroxisomes [see Figure 15]. This reaction is catalyzed by the alcohol oxidase AOX. The low affinity of this protein to oxygen explains the high expression levels of AOX, since the cell tries to compensate the low activity with larger amounts of protein. Two genes encode this enzyme, AOX1 and AOX2, with a homology of > 90 % in the protein coding sequence, but non-

homologous 5' and 3' flanking regions and an identical regulation mechanism (46), (50). However, the main part of the alcohol oxidase expression is regulated by the *AOX1* gene, which is responsible for 85 % of the activity (51).

The produced hydrogen peroxide is subsequently degraded to water and molecular oxygen by a catalase in the same cell compartment. One part of the formaldehyde enters the cytosol where it is further metabolized to formate and finally to carbon dioxide. Two enzymes are involved in this so called dissimilation pathway, the formaldehyde dehydrogenase and the formate dehydrogenase. The thereby regenerated NADH can subsequently be used for energy production. The second part of formaldehyde that is remaining in the peroxisome is used for an assimilatory pathway and reacts with xylulose-5-phosphate. This condensation reaction is catalyzed by dihydroxyacetone synthase, a transketolase and results in the production of glyceraldehyde-3-phosphate and dihydroxyacetone. These two substances are further metabolized in the cytosol to regain xylulose-5-phosphate. Furthermore, for every three cycles one molecule glyceraldehyde-3-phosphate is regenerated, which can then be used in the central metabolism and for the production of cell constituents and biomass (42), (44), (48), (51).

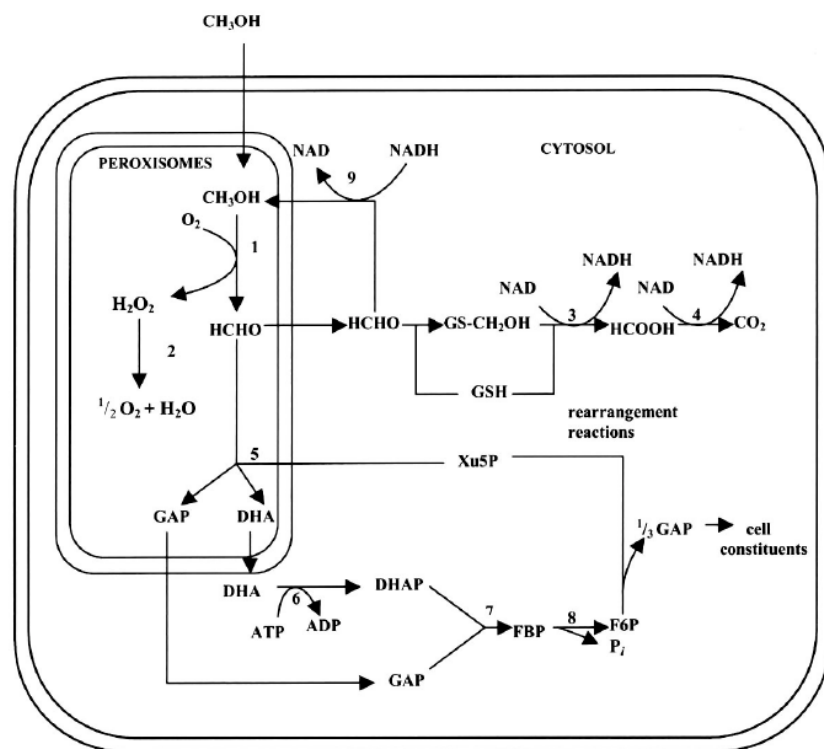


Figure 15: The methanol utilization pathway in *P. pastoris* (52). 1: alcohol oxidase, 2: catalase, 3: formaldehyde dehydrogenase, 4: formate dehydrogenase, 5: dihydroxyacetone synthase, 6: dihydroxyacetone kinase, 7: fructose-1,6-bisphosphate aldolase, 8: fructose-1,6-bisphosphatase, 9: formaldehyde reductase

### 1.4.3. Phenotypes and strains

There are three different phenotypes of host strains available for the expression of foreign proteins. Depending on their methanol utilization, the Mut<sup>+</sup>, Mut<sup>S</sup> and Mut<sup>-</sup> strains can be distinguished. The Mut<sup>+</sup> strain, which requires high levels of methanol, shows the wild type growth rate on methanol and is most commonly used (41). This phenotype is obtained when an expression cassette is inserted in the *P. pastoris* genome at the *HIS4* or *AOX1* locus via single crossover integration. A double crossover gene transplacement at the *AOX1* locus, on the other hand, results in the deletion of the *AOX1* gene and gives the Mut<sup>S</sup> phenotype (e.g. the strain KM71, see Table 2) (51). The slow growth rate of this strain on methanol due to the weaker *AOX2* gene facilitates in some cases a better production of the recombinant protein. Furthermore, less methanol is required, which may represent a considerable fire hazard in large scale cultures (43). Double knockout strains, where both *AOX* genes are deleted, are not able to grow on methanol and classified as Mut<sup>-</sup> phenotype (e.g. the strain MC100-3) (41), (42), (46). The common *P. pastoris* expression strains are derivatives from NRRL-Y 11430 (Northern Regional Research Laboratories, Peoria, IL, USA). Additionally, many more strains are available, for example protease deficient strains like the strain SMD1163 (42).

**Table 2: Different *P. pastoris* strains and their genotype as well as phenotype.**

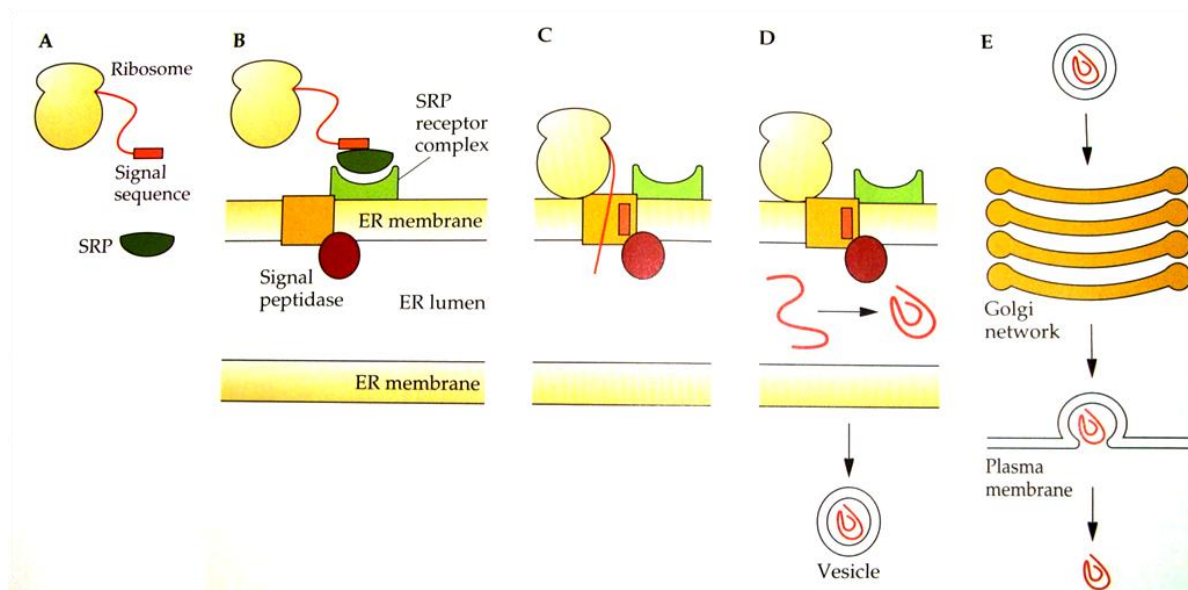
strain	genotype	phenotype	reference
NRRL-Y-11430 (=CBS-7435)	wild-type	Mut <sup>+</sup>	(42)
CBS-7435 mutS BT3445	$\Delta$ aox1	Mut <sup>S</sup>	(53)
GS115	$\Delta$ his4	Mut <sup>+</sup> His <sup>-</sup>	(54)
KM71	aox1 $\Delta$ ::SARG4 arg4 his4	Mut <sup>S</sup> His <sup>-</sup>	(55)
MC100-3	aox1 $\Delta$ ::SARG4 aox2 $\Delta$ ::Phis4 arg4 his4	Mut <sup>-</sup> His <sup>-</sup>	(46)
CBS-7435 $\Delta$ ku70 BT3499	$\Delta$ ku70	Mut <sup>+</sup>	(53)

Various selectable markers can be used, for example auxotrophic markers from genes of biosynthesis pathways like mutants with deficiencies in the histidinol dehydrogenase (*HIS4*) or argininosuccinate lyase (*ARG4*) genes. Alternatively, genes conferring drug resistance can be used, for example the zeocin resistance gene (Zeo<sup>R</sup>) (41), (42).

### 1.4.4. Secretion

There are two options for the expression of heterologous proteins in *P. pastoris*. They can be produced intracellularly or extracellularly. The secretory pathway of all eukaryotic organisms is similar [see Figure 16] and requires a signal sequence, which is bound by a signal recognition particle (SRP) during translation of the protein. This particle binds to a receptor on the endoplasmic reticulum (ER) and thereby facilitates the transport of the secretory

protein through the membrane into the lumen of the ER. Here, the folding and if necessary the core-glycosylation of the protein takes place and the signal peptide is cleaved off by a signal peptidase. The processed protein is then packed into a vesicle and transported to the *cis* membrane of the Golgi network, where further posttranslational modifications and glycosylations occur. Subsequently, the secretory protein is again encapsulated in a transport vesicle, which fuses with the plasma membrane and thereby releases the protein to the extracellular environment. This release may occur continuously or only after a certain signal, like the depolarization of the membrane or hormones (56).

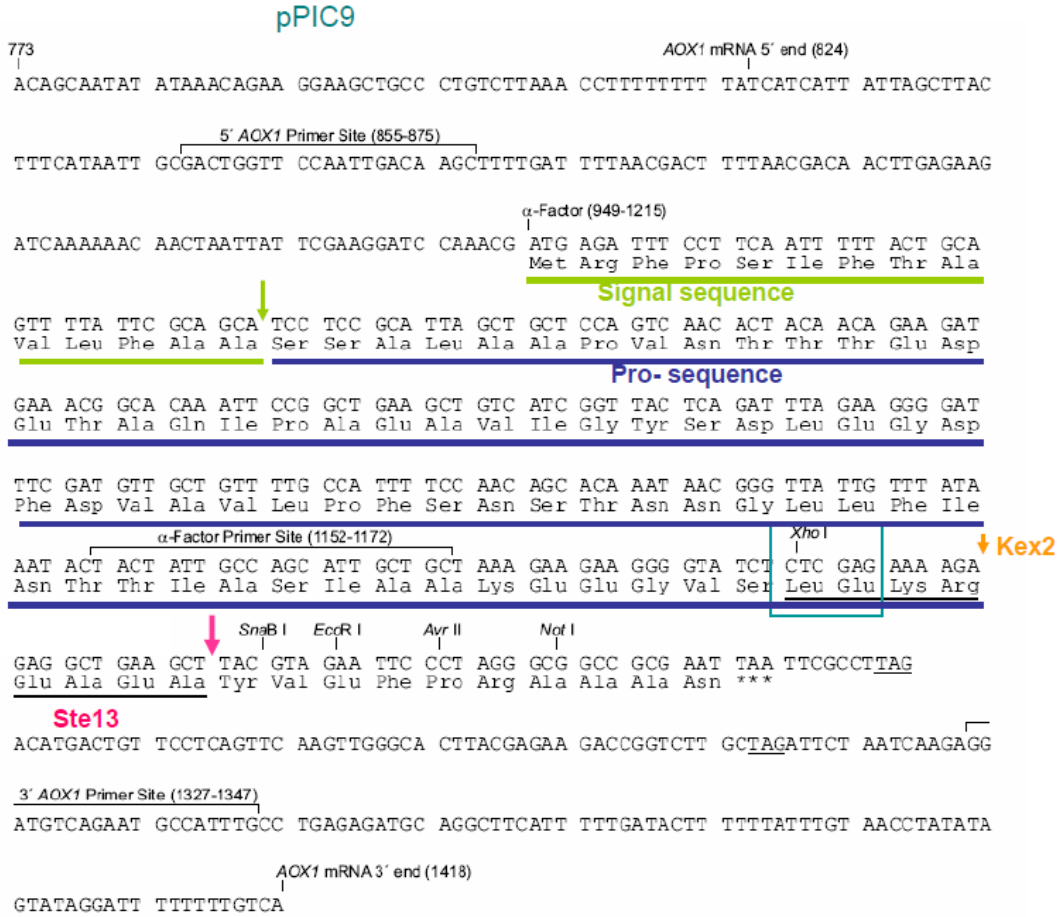


**Figure 16: The secretion pathway in eukaryotes via the endoplasmic reticulum and the Golgi apparatus (56).**

For secretion it is usually necessary that the protein of interest is secreted in the native host as well because of folding and stability requirements. Expression cassettes can be made, where genes of foreign proteins are fused to the secretion sequences of the native protein, the  $\alpha$ -factor prepro signal derived from *S. cerevisiae* or for example the acid phosphatase 1 signal (PHO1) from *P. pastoris*. Thereby, the correct reading frames have to be maintained in the course of cloning (41), (42).

So far, very good results were achieved with the frequently used  $\alpha$ -factor prepro peptide, which in some cases proved to be more efficient than the native signal sequence. It consists of a signal pre-sequence with a length of 19 amino acids and a 66 amino acid long pro-sequence with a dibasic Kex2 endopeptidase processing site (42), (57). In the first step of processing, the pre-signal is removed in the endoplasmic reticulum by signal peptidase. Subsequently, the Kex2 endopeptidase cleavage occurs at the Arg - Lys residues in the pro signal sequence. Immediately afterwards the Ste13 protein cleaves at the Glu - Ala repeats [see Figure 17] (42), (58).





**Figure 17: The Kex2 and Ste13 cleavage sites for processing of the α - factor signal sequence (59).**

The efficiency of the cleavage can be influenced by the amino acids surrounding the cleavage sites as well as by the tertiary structure of the produced protein. However, not all proteins can be expressed extracellularly with this signal sequence and secretion with the α-factor prepro signal frequently results in a variable number of amino acids at the N-terminus (42). As a result, alternatives were searched and various other strategies were tested, like creating synthetic leaders (60) or using for example the secretion signal of the pGKL killer protein (61). The PHA-E phytohemagglutinin signal sequence derived from *Phaseolus vulgaris* was successfully used for the expression of plant lectins and with the help of these alternative signal sequences, secreted proteins with correctly processed N-termini were obtained (62).

Various other options have been investigated to enhance secretion and the final product yield. Commonly used strategies to increase protein expression are to optimize the codon usage and to increase the gene dosage. Sometimes, however, exactly the opposite is reached since overexpression may result in the saturation of the secretory pathway and incorrect folding in the endoplasmic reticulum (63). One method to ensure proper folding is the overexpression of protein disulfide isomerase (PDI), an enzyme that facilitates the

rearrangement of incorrect disulfide bonds (64). Efficient processing and folding are prerequisites for a successful secretion to avoid aggregated and unfolded proteins that are otherwise targeted for degradation in the proteasome (63). Up to now, there is no method to predict in advance which secretion signal will lead to the best results. Trial and error experiments are necessary to find out whether a protein can be secreted and which signal is suited best for the individual foreign proteins (65).

#### 1.4.5. Fermentation

*P. pastoris* is very well suited for the large scale production of proteins since the used media are well defined and not expensive. The levels of protein expression and cell density that can be reached in fermentation cultures lie at >100 g/L dry cell weight and >400 g/L wet cell weight and are much higher than those obtained in shake-flask cultures. Prerequisites therefore are the optimization of the culturing conditions and the tight control of several parameters, like the pH, temperature, aeration and carbon source feed rate (65).

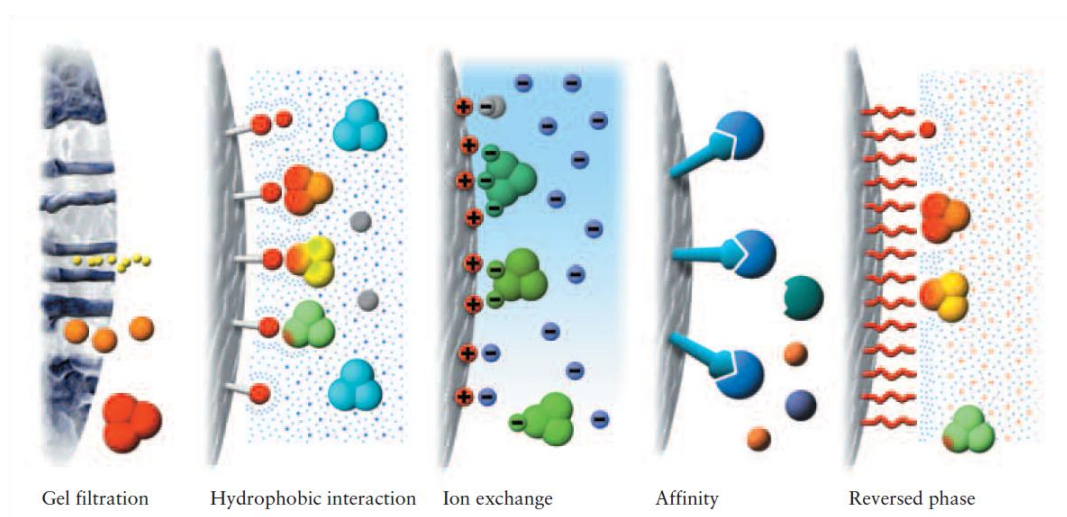
Once the successful expression of the protein of interest and the feasibility and stability of the selected host strain are proven, the scale-up of the culturing system, which requires the consideration of several factors, can be started. Parameters like the composition of the media, the pH and the temperature of the fermentation influence the solubility of the produced protein as well as the productivity and have to be optimized (66). By lowering the temperature during the induction phase of *P. pastoris* with methanol from 30 °C to 25 °C a fourfold improvement in the yield of galactose oxidase could be achieved (67). This may be explained by the poor stability and incorrect folding of some proteins at higher temperatures or by the increase of cell lysis and therefore of released proteases. Another possibility to reduce protein degradation is to modify the pH in such a way, that it is no longer optimal for protease cleavage, but still within the relatively broad range for growth of *P. pastoris* between pH 3 and 7 (41). Since the fermentation protocol affects the impurity profile and solubility of the protein, it influences at the same time the downstream processing as well as the yield and quality of the end product (66).

The production of heterologous proteins is usually divided into three phases. The first stage is called glycerol batch phase (GBP). The cells are grown on glycerol, a non-fermentable carbon source, before the glycerol fed batch phase (GFB) is started where the substrate is added continuously. In order to increase the biomass to the desired level, the glycerol is depleted first and then added again at a growth limiting rate. Thereby the cells are prepared and derepressed before induction. The induction of the protein expression with methanol occurs in the third stage, the methanol fed batch phase (MFB). The feeding rate is increased

stepwise to obtain the optimal production rate of the protein. This rate and the viability of the cells may be further improved by mixed feeding, where glycerol and methanol are added simultaneously. In this case it is important to tightly control the level of glycerol, since too high concentrations of glycerol repress the *AOX1* promoter and promote the production of acetate and ethanol (65), (51).

## 1.5. Protein Purification

The purification of biomolecules is based on various differences in specific properties [see Figure 18]. Gel filtration allows the separation according to the sizes of the molecules, while differences in the charges of proteins are used for ion exchange chromatography. Affinity chromatography on the other hand is based on biorecognition of specific ligands. Furthermore, hydrophobic interaction and reversed phase chromatography can be applied if the separation according to the hydrophobicity of the proteins is desired (68), (69), (70).

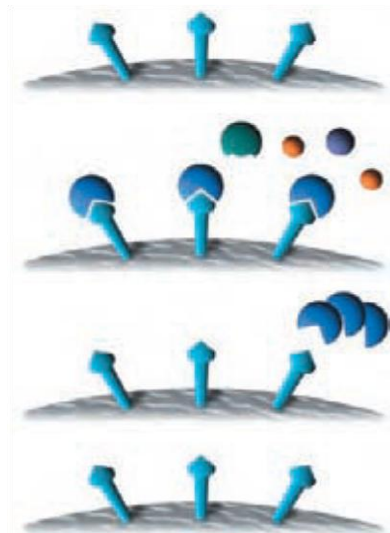


**Figure 18: Overview of various strategies available for protein purification (69).**

### 1.5.1. Affinity chromatography

The basis of affinity chromatography is the reversible interaction between a protein and a specific ligand, which is coupled to the chromatographic matrix. Reasons for the popularity of this technique are its good selectivity and resolution and the high capacity of the columns. Furthermore, high yields can be obtained, native forms can be separated from denatured ones and even small amounts out of large sample volumes can be isolated and concentrated. Affinity chromatography is the only method that allows the purification of proteins according to their specific chemical structure or their biological function. A

prerequisite of this method is the reversibility of the interactions between the ligand and the target biomolecule, which can result from H-bonds, hydrophobic or electrostatic interactions or van der Waals forces. By adding competitive ligands or by a change of the pH, the polarity or the ionic strength these interactions can be reversed and the target molecule is eluted (68). The individual steps of affinity chromatography are the following:



1. Equilibration of the affinity medium.

2. The sample is applied, only the target molecule binds specifically to the ligand while all other sample components are washed out.

3. Elution of the target molecule by competitive ligands or modification of pH, ionic strength etc.

4. Re-equilibration with binding buffer.

**Figure 19: Concept of affinity chromatography (68).**

In the present study, the maltose binding protein (MBP) is used as an affinity tag and fused to moxY. The MBP fusion protein binds to the amylose resin of the column, whereas all other substances in the sample are washed out and discarded in the flow through. Due to the strong affinity of MBP to maltose, the addition of 10 mM maltose facilitates the recovery of the MBP fusion protein in a small volume of elution buffer.

### 1.5.2. Anion exchange chromatography

Ion exchange chromatography (IEX) is applied to separate biomolecules according to differences in the net surface charges, which depend on the pH. Charged sample molecules reversibly interact with an oppositely charged IEX medium. In the case of anionic exchange chromatography, the IEX medium consists of a porous matrix of spherical particles with positively charged ionic groups. When the pH is higher than the isoelectric point of a protein it will bind to the positive IEX medium. For equilibration the buffer is chosen in such a way, that the sample binds while as much of the contaminants as possible do not interact. After loading the sample and washing the column the buffer conditions are modified for a controlled release of the interactions. Elution of the protein is typically achieved by increasing the ionic strength or by changing the pH. The salt ions compete for the ionic groups on the IEX medium and gradually replace the bound sample molecules. Those molecules with the

highest net charge at a selected pH are eluted at the end since they are retained most strongly. By applying a gradient for elution, purified and concentrated protein fractions can be obtained. The method can be used for micro scale experiments as well as for large scale applications and provides not only high resolution but also large loading capacities (69).

### 1.5.3. Size exclusion chromatography

In contrast to the two other methods, no binding between the sample and the matrix occurs when size exclusion chromatography, also called gel filtration, is applied. The matrix is a porous filtration medium that is packed in the column consisting of spherical, chemically and physically stable and inert particles. This method is based on a separation of the molecules according to their different sizes. Larger molecules diffuse not as far into the pores of the filtration medium, therefore their way through the column is shorter. As a result, large molecules are eluted prior to small ones. The elution can be performed isocratically, without needing to change the buffer. Gel filtration is a robust, simple and flexible method. It allows the separation of biomolecules which are sensitive to harsh conditions or fast changes in pH or buffer components. The buffer, ionic strength or temperature can be chosen freely depending on the actual requirements of the experiment and cofactors, detergents or important ions can be added, if needed. By high resolution fractionation the isolation of desired compounds, the refolding of denatured enzymes as well as the separation of monomers and other aggregates are possible. To sum up, homogenous samples can be obtained by the removal of protein contaminants with a different size or tertiary structure (70).

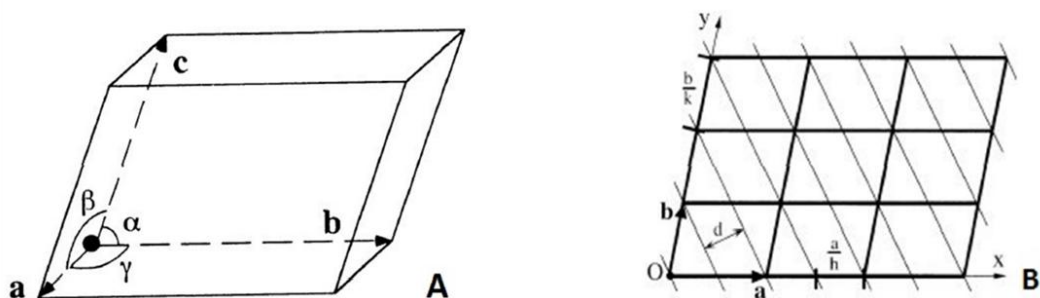
### 1.6. Crystallography

The two main methods for resolving the chemical structure of macromolecules are crystallography and NMR (nuclear magnetic resonance). Thereby crystallography allows in contrast to NMR the analysis of very large macromolecules and proteins with more than 200 amino acids or viruses, but it requires the growth of crystals.

Crystallography is based on the interrogation of macromolecules by electromagnetic waves. Therefore X-rays are used because their wavelength  $\lambda$  of  $\sim 1 \text{ \AA}$  corresponds to the desired resolution. Otherwise, it would not be possible to resolve atoms which are usually approximately  $1 - 2 \text{ \AA}$  apart from each other. X-rays are generated when electrons are accelerated or decelerated. One of the most powerful sources for X-rays is a synchrotron. Because of the high intensity of synchrotron radiation, it is possible to investigate even very tiny, poorly diffracting crystals or crystals with large unit cells (71). When electrons are hit by the incoming X-rays, this leads to scattering, so called diffraction, which is connected to the

interference between two or more scattered electromagnetic waves. Since the scattering of the electrons of an individual molecule is too weak to be measured, crystals are needed to amplify the signal.

Crystals are made of atoms and molecules which are arranged in a three - dimensionally periodic structure, consisting of identical unit cells which are repeated (72). The unit cell can be characterized by the unit cell parameters  $a$ ,  $b$  and  $c$  and the corresponding angles  $\alpha$ ,  $\beta$  and  $\gamma$ , which represent the lattice constants [see Figure 20 (A)].



**Figure 20: (A) The unit cell of a crystal with the unit cell parameters  $a$ ,  $b$ ,  $c$  and the angles  $\alpha$ ,  $\beta$ ,  $\gamma$ . (B) A two - dimensional lattice with lattice planes in a distance  $d$  (here  $h = 2$  and  $k = 1$ ) (72), (73).**

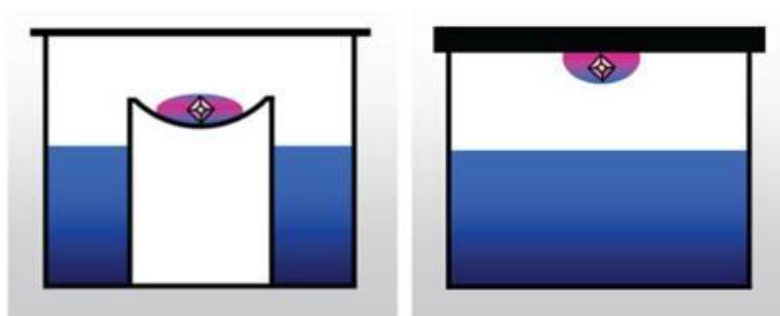
Crystals amplify only the diffraction of those scattering planes that fulfill the Laue conditions. Regularly spaced sets of planes can be drawn through the points of the crystal lattice, which are designated by the three Miller indices  $h$ ,  $k$ ,  $l$  [see Figure 20 (B)]. These indices describe into how many parts  $a$ ,  $b$  and  $c$  are cut by the sets of planes. In order to fulfill the Laue conditions and Bragg's law, these indices have to be integer numbers to ensure periodicity and to allow diffraction. In other words, constructive interference of scattered waves from parallel planes occurs when the path difference of these X-rays is an integer number of the wavelength. Therefore, the Bragg's law has to be fulfilled, with  $d$  being the distance between the scattering planes, the scattering angle  $\theta$ , the integer  $n$  and the wavelength  $\lambda$ :

$$2 d \sin \theta = n \lambda$$

The aim of crystallography is to visualize the structure of a molecule by calculation of the electron density at every point of the unit cell, so at every position  $x$ ,  $y$ ,  $z$ . Therefore it is necessary to measure the structural factors, which are the Fourier transform of the electron density. Consequently, all structural factors of all  $h$ ,  $k$ ,  $l$  - so the amplitudes and phases of every individual wave - have to be measured. While the amplitude can be determined by the intensity of the wave at the detector, it is not possible to measure the phases. In order to solve the so called phase problem, several methods have been established, including direct methods, heavy metals isomorphous replacement, single and multiple wavelength anomalous dispersion (SAD/MAD) methods or molecular replacement.

The growth of crystals can be a serious bottleneck. Crystals are very fragile, soft and easy to crush. They contain a large amount of solvent, which can range from 25 - 90 %, and are sensitive to dehydration. The solubility of a protein is a function of the pH, ionic strength, temperature and the dielectric constant. Crystals can be formed when a solution is saturated, so when there are no longer enough solvent molecules to maintain the dissolved state. Once the energy barrier for nuclei formation is overcome, the crystal growth can take place. The so-called salting out can be reached by adding high salt concentration since these ions compete with the protein for the solvent molecules. Salting in, on the other hand, means the increase in solubility due to the addition of small salt concentrations. Removing the salt leads to the precipitation of the protein (71). Salts like  $(\text{NH}_4)_2\text{SO}_4$  and  $\text{MgSO}_4$  or precipitants like PEG are frequently used for the crystallization of proteins. The solubility of proteins can also be influenced by changing the pH (e.g. within pH 4 - 9) or varying the temperature (e.g. between 4 - 25 °C). The more you know about the protein you want to crystallize the better but the main prerequisite for crystallization experiments is a very pure, homogenous protein sample with relative high concentrations (up to ~ 20 mg/mL).

There are several crystallization methods like vapor diffusion, batch crystallization, liquid-liquid diffusion or dialysis. Although the batch method, where the precipitant is instantaneously added, is the easiest and cheapest, vapor diffusion is most commonly used. It's a gentle and automated method which requires only small amounts of the protein sample. At the hanging drop method, a drop of the protein solution and the same volume of precipitant solution are placed on a siliconized glass slip. The glass is put over a depression in a tray containing the precipitant mother solution [see Figure 21]. After sealing of the system it tends towards equilibrium by vapor diffusion between the mother solution and the drop. Alternatively, the sitting drop method can be applied, for example when the surface tension of the protein solution is low. The principle is the same, but here the protein and precipitant solutions are placed on a well within the depression for the precipitant reagent (71).



**Figure 21: Sitting drop and hanging drop crystallization. The precipitant solution is shown in blue, the drop containing the protein and precipitant solutions is shown in pink (74).**

## 2. Objectives

The aim of this work was to develop a successful strategy for the expression and purification of certain Baeyer - Villiger monooxygenases, in order to obtain reasonable amounts of pure protein that finally allow crystallization experiments for structural characterizations. The main focus lied thereby on moxY, a monooxygenase involved in the aflatoxin biosynthesis pathway of *Aspergillus* species. Additionally, two other BVMOs were investigated, namely the well characterized phenylacetone monooxygenase PAMO and the brand new BVMO Number 24.

In the first part of the thesis the research work concerning moxY is described. Since so far the exact localization of this enzyme is not known, fluorescence microscopy experiments were performed at the beginning to get hints whether the protein is membrane associated or part of the cytosolic fraction of the cell. Since the secretion of a protein by *Pichia pastoris* can already serve as a first purification step, it was subsequently analyzed whether it is possible to secrete this monooxygenase using two different secretion signals. Next to the  $\alpha$ -factor signal peptide the secretion signal of the 1,4- $\beta$ -cellobiohydrolase CBH2 was tested.

For the purification of moxY via affinity chromatography, different constructs of the enzyme together with a MBP-tag were prepared. After the expression of these MBP-moxY fusion proteins in deep well plates and small shake flasks as well as in a fermenter, different strategies for the cell breakage and protein purification were analyzed to finally obtain the pure enzyme which could be used for first crystallization trials.

The second part of the thesis focuses on PAMO and BVMO24. A lot of information is already available about PAMO allowing the efficient expression and purification of the protein for subsequent crystallographic experiments.

For the recently identified BVMO24 a method for the cultivation and purification of the enzyme was established in order to obtain the protein in its pure and soluble form. Again first crystallization trials were started with the aim of determining the three - dimensional structure of this monooxygenase.



### 3. Chapter I: Secretion study with moxY

#### 3.1. Background

The localization of a protein has a large influence on the applied working and purification procedure. Therefore, it is important to clarify already at an early stage whether the enzyme is soluble and part of the cytosolic fraction in the cell, membrane bound or secreted. In the case of moxY, the localization of the protein is not known. *In silico* sequence analysis tools as well as fluorescence microscopy experiments were performed to get some information.

The secretion of proteins can be used as a convenient method to obtain the protein in a relatively pure form in the supernatant of the fermentation medium and can be applied alternatively to a first chromatographic purification step, since only few endogenous enzymes are secreted by the *P. pastoris* host itself. In the course of a secretion study, the  $\alpha$ -factor signal sequence from *Saccharomyces cerevisiae* was tested. In order to analyze also a different secretion sequence, the signal peptide of the 1,4- $\beta$ -cellobiohydrolase CBH2, also called exoglucanase II, was used which proved to be a good secretion signal in *Pichia* (75). From the CBH2 gene the signal peptide, the CBM1 binding domain and the linker were used, whereas the catalytic domain was replaced by the gene encoding the protein of interest [see Figure 22 and Appendix, chapter 10].

**Sequence annotation (Features)**

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>					
<input type="checkbox"/> Signal peptide	1 – 24	24	<a href="#">Ref.3</a>		
<input type="checkbox"/> Chain	25 – 471	447	Exoglucanase 2		PRO_0000007911
<b>Regions</b>					
<input type="checkbox"/> Domain	26 – 62	37	CBM1		
<input type="checkbox"/> Region	66 – 106	41	Linker		
<input type="checkbox"/> Region	107 – 471	365	Catalytic		

**Figure 22: Sequence annotation of the CBH2 gene of the exoglucanase 2 (76).**

The signal peptide of CBH2 was chosen due to the fact that record levels of secretion in *P. pastoris* were reached in experiments of A. Mellitzer *et al.* (75),(77). Thereby the codon optimized CBH2 gene was expressed with its native signal sequence under the control of a synthetic promotor. The results indicated that the signal peptide of CBH2 is very well suited for secretion in *P. pastoris*. Therefore the aim was to test, whether moxY can be efficiently secreted as well, when the protein is fused to the CBH2 signaling sequence.

## 3.2. Materials and methods

### 3.2.1. Instruments and devices

#### 3.2.1.1. Centrifuges

- 5810 R Centrifuge; Eppendorf AG (Hamburg, Germany)
- 5415 R Centrifuge; Eppendorf AG (Hamburg, Germany)
- J-20 XP Avanti™ centrifuge; Beckman Coulter™, Inc (Vienna, Austria)
- JA-10 Rotor; Beckman Coulter™, Inc (Vienna, Austria)
- Optima™ LE-80K Ultracentrifuge; Beckman Coulter™, Inc (Vienna, Austria)
- Ti-70 Rotor, 70 000 rpm; Beckman Coulter™, Inc (Vienna, Austria)

#### 3.2.1.2. Shakers and incubators

- Titramax 1000, 1,5 mm, for microtiter plates; Heidolph Instruments (Schwabach, Germany)
- Thermomixer comfort, 1.5 mL; Eppendorf AG (Hamburg, Germany)
- RS 306 rotary shaker, 50 mm; Infors AG (Bottmingen Basel, Switzerland)
- Binder - drying oven; Binder GmbH (Tuttlingen, Germany)
- Infors HT Multitron incubator shaker; Infors AG (Bottmingen Basel, Switzerland)
- GFL-3013 shaker; Gesellschaft für Labortechnik GmbH (Burgwedel, Germany)

#### 3.2.1.3. PCR cyclers

- GeneAmp® 2720 Thermal Cycler; Applied Biosystems (Foster City, CA, USA)
- Mastercycler® personal; Eppendorf AG (Hamburg, Germany)

#### 3.2.1.4. Photometers and plate readers

- Spectrophotometer DU 800; Beckman coulter Inc (Fullerton, CA, USA)
- Plus 384 Spectramax; Molecular Devices (München, Germany)
- Semi-micro cuvette, polystyrene, 10x4x45 mm; Sarstedt AG & Co. (Nümbrecht, Germany)
- Gemini XS Spectramax; Molecular Devices (München, Germany)
- Greiner Bio-One Plate 96W; Greiner Bio-One (Frickenhausen, Germany)
- NanoDrop Spectrophotometer 2000c; peqlab Biotechnologie GmbH (Polling, Austria)
- G:BOX imaging system; Syngene (Cambridge, UK)

#### 3.2.1.5. Gel electrophoresis and associated materials and devices

##### Agarose gel electrophoresis

- Biozym LE Agarose; Biozym Biotech Trading GmbH (Vienna, Austria)
- GeneRuler™ 1kb DNA-Ladder; Fermentas GmbH (St. Leon-Rot, Germany)
- 6x DNA Loading Dye; Fermentas GmbH, (St. Leon-Rot, Germany)
- Sub-cell GT; Bio-Rad Laboratories (Vienna, Austria)
- Power Pac™ Basic; Bio-Rad Laboratories (Vienna, Austria)

- Chroma 43 mittelwellig 302 nm; Laborgeräte Vetter GmbH (Wiesloch, Germany)
- GelDoc-It™ Imaging Systems, Benchtop2UV™ Transilluminator; UVP® (Cambridge, UK)
- BioImaging Systems Gel HR Camera 6100; UVP® (Cambridge, UK)

## SDS-PAGE

- NuPage® 4-12% Bis-Tris Gel, 1.0 mm x 15 well; Invitrogen™ life technologies (Lofer, Austria)
- PageRuler™ Prestained Protein Ladder; Fermentas GmbH (St. Leon-Rot, Germany)
- NuPage® MOPS SDS Running Buffer (20x); Invitrogen™ life technologies (Lofer, Austria)
- NuPage® Sample Reducing Agent (10x); Invitrogen™ life technologies (Lofer, Austria)
- NuPage® LDS Sample Buffer (4x); Invitrogen™ life technologies (Lofer, Austria)
- XCell SureLock™; Invitrogen™ life technologies (Lofer, Austria)
- Power Ease 500; Invitrogen™ life technologies (Lofer, Austria)

### 3.2.1.6. Electroporation devices

- Gene Pulser™; Bio-Rad Laboratories (Vienna, Austria)
- Electroporation cuvettes, 2 mm; Bridge Bioscience™ (Rochester, NY, USA)

### 3.2.1.7. Reaction tubes

- Microcentrifuge tubes ,1.5 mL volume with lid; Greiner Bio GmbH (Frickenhausen, Germany)
- Polypropylene - tubes, sterile, cap, 12 mL volume; Greiner Bio GmbH (Frickenhausen, Germany)
- Polypropylene - tubes, sterile, 15 mL volume; Greiner Bio GmbH (Frickenhausen, Germany)
- Polypropylene - tubes, sterile, with and without support skirt, 50 mL volume; Greiner Bio GmbH (Frickenhausen, Germany)

### 3.2.1.8. Pipettes and pipette tips

- Pipetman, single channel pipettes (adjustable volume 0.2 - 1000 µL); Gilson, Inc. (Middleton, WI, USA)
- Eppendorf Research® pipette (adjustable volume 0.1-2.5 µL); Eppendorf AG, (Hamburg, Germany)
- Pipette tips (10, 200 and 1000 µL) with and without filter; GBO Greiner Bio-One GmbH (Frickenhausen, Germany)
- Biohit Proline® multichannel pipettor, 8 channels (5-50 µL); Biohit Plc. (Helsinki, Finland)
- Biohit Proline® multichannel electronic pipettor, 8 channels (50-1200 µL); (Biohit Plc., Helsinki, Finland)
- Biohit Tips 300 µL, Single Tray; (Biohit Plc., Helsinki, Finland)
- Biohit Tips 1200 µL, Bulk; (Biohit Plc., Helsinki, Finland)

### 3.2.1.9. Microplates

- Polystyrene- microplate 96-well, flat bottom; Greiner Bio-One GmbH (Frickenhausen, Germany)
- 96-well deep well plate and cover, Polypropylene; Bel-Art Products (Pequannock, NJ, USA)

### 3.2.1.10. Other materials and devices

- Leica DFC 350 FX microscope; Leica Microsystems (Wetzlar, Germany)
- Vortex-Genie 2; Scientific Industrie Inc (Bohemia, NY, USA)
- inoLab® pH720 pH meter, WTW (Weilheim, Germany)
- Polyplast Temp Din pH electrode; Hamilton Company (Reno, NV, USA)
- MR 3000 and MR 2002 magnetic stirrers; Heidolph (Schwabach, Germany)
- Sartorius BL 120S scale; Sartorius Stedim Biotech GmbH (Göttingen, Germany)
- PG12001-S Delta Range Scale; Mettler - Toledo, Inc (Columbus, OH, USA)
- arium® basic ultrapure water system; Sartorius Stedim Biotech GmbH (Göttingen, Germany)
- Cell homogenizer MSK; P. Braun Biotech International GmbH (Melsunge, Germany)
- CO<sub>2</sub> gas for cooling; Linde Gas GmbH (Stadl Paura, Austria)
- Diaphragm pump; Vacuubrand GmbH + Co (Wertheim, Germany) and driving device ABM; Antriebstechnik GmbH (Marktredwitz, Germany)

## 3.2.2. Media and chemicals

### 3.2.2.1. Media and antibiotics for the cultivation of *E. coli*

- LB media: low salt Luria Bertani: 10 g/L Bacto™ tryptone, 5 g/L Bacto™ yeast extract, 5 g/L NaCl, autoclave sterilization
- LB agar: 35 g LB Agar (Lennox, Roth) were weight in and filled up to 1 L with dH<sub>2</sub>O before autoclaving it. If needed, antibiotics were added to the warm solution to reach a final concentration of 25 mg/L zeocin™ (250 µL).
- SOC medium: 5 g/L Bacto™ yeast extract, 20 g/L Bacto™ tryptone, 0.58 g/L NaCl, 2 g/L MgCl<sub>2</sub>, 2.46 g/L MgSO<sub>4</sub>, 0.18 g/L KCl, 3.81 g/L α-D(+)-glucose monohydrate, autoclave sterilization
- ampicillin stock [concentration of 100 mg/mL]: 5 g ampicillin / 50 mL dH<sub>2</sub>O, filter sterilized
- zeocine stock [concentration of 100 mg/mL]: 2 g zeocine / 20 mL dH<sub>2</sub>O, filter sterilized

### 3.2.2.2. Media and solutions for the cultivation of *P. pastoris*

- 500x B: 10 mg / 50 mL d-Biotin, filter sterilization
- 10x D: 20 % glucose, 220 g/L α-D(+)-glucose monohydrate, autoclave sterilization
- 10x YNB: 1,34 % YNB, 134 g/L Difco™ yeast nitrogen base w/o amino acids, autoclave sterilization
- 10x PPB: 30.0 g/L K<sub>2</sub>HPO<sub>4</sub>, 118.0 g/L KH<sub>2</sub>PO<sub>4</sub>. The pH of the buffer was set to 6.0 using conc. KOH, autoclave sterilization
- YPD media: Yeast extract, peptone, dextrose: 10 g/L Bacto™ yeast extract and 20 g/L Bacto™ peptone were filled up to 900 mL with dH<sub>2</sub>O. The media was sterilized via autoclavation before 100 mL 10x D were added.

- YPD agar: 10 g/L agar were added to the YPD media before autoclaving it. If needed, antibiotics were added to the warm solution to reach a final concentration of 100 mg/L zeocin™ (1 mL).

**Table 3: Composition of BMD 1 %, BMM 2 and BMM 10 media (V = 1 L).**

	BMD 1 %	BMM 2	BMM 10
10 x PPB	200 mL	200 mL	200 mL
10 x YNB	100 mL	100 mL	100 mL
10 x D	50 mL	-	-
MeOH	-	10 mL	50 mL
500 x B	2 mL	2 mL	2 mL
dH <sub>2</sub> O	650 mL	690 mL	650 mL

### 3.2.2.3. Other buffers, solutions and chemicals

- dNTP mix, 10 mM each, dATP, dTTP, dCTP, dGTP; Fermentas GmbH (St. Leon - Rot, Germany)
- 50x TAE buffer for 1 L: 242 g Tris base, 57.1 mL glacial acetic acid and 100 mL 0.5 M EDTA were filled up to 1 L with dH<sub>2</sub>O
- 1M DTT: 0,62 g dithiothreitol were freshly dissolved in 4 mL ddH<sub>2</sub>O and filter sterilized through a 0,2 µm filter
- BEDS: 10 mM bicine (1,63 g bicine were dissolved in 920 mL dH<sub>2</sub>O), the pH was set to 8.3 with 2 M NaOH, addition of 3 % (v/v) ethylene glycol (30 mL), 5 % (v/v) DMSO (50 mL), 1 M sorbitol (182 g), filter sterilization through 0,2 µm filters, freezing of aliquots at - 20 °C
- SDS gel staining solution: 2 g Coomassie Blue R, 500 mL EtOH (abs.), 100 mL acetic acid (100 %), 400 mL dH<sub>2</sub>O
- SDS gel destaining solution: containing 10 % absolute EtOH (abs.) and 10 % acetic acid
- Y-Per®: Yeast Protein Extraction Reagent #78990 (Pierce Biotechnology Inc., Rockford in IL, USA)
- Yeast lysis buffer: Triton™ X-100 20 µL/mL, 5 M NaCl 20 µL/mL, 10 % SDS 100 µL/mL, 1 M Tris, pH 8, 10 µL/mL and 0.5 M EDTA 2 µL/mL were dissolved in dH<sub>2</sub>O
- BioRad, BCA Protein Assay Kit, Thermo Scientific (Rockford, IL, USA)

### 3.2.3. Enzymes

The enzymes and buffers for the restriction cuts and ligations are listed below. They were used as described in the manufacturer's recommendations, in case of deviations, they are mentioned. The restriction enzymes, with concentrations of 10 u/µL or 1 FDU/µL (FastDigest® unit), were purchased from Fermentas GmbH (St. Leon-Rot, Germany) and applied according to the manual ([www.fermentas.com](http://www.fermentas.com)).

### 3.2.3.1. Restriction enzymes and ligase

- *EcoRI*, 5'G<sup>^</sup>AATTC3'
- *SpeI* (*BcuI*), 5'A<sup>^</sup>CTAGT3'
- *NotI*, 5'GC<sup>^</sup>GGCCGC3'
- *XhoI*, 5'C<sup>^</sup>TCGAG3'
- *SmaI* (*Swal*), 5'ATT<sup>^</sup>AAAT3'
- *BglII*, 5'A<sup>^</sup>GATCT3'
- *BamHI*, 5'G<sup>^</sup>GATTC3'
- FastDigest *EcoRI*, 5'G<sup>^</sup>AATTC3'
- FastDigest *SpeI* (*BcuI*), 5'A<sup>^</sup>CTAGT3'
- FastDigest *NotI*, 5'GC<sup>^</sup>GGCCGC3'
- FastDigest *XhoI*, 5'C<sup>^</sup>TCGAG3'
- FastDigest *BamHI*, 5'G<sup>^</sup>GATTC3'
- FastDigest *SmaI* (*Swal*), 5'ATT<sup>^</sup>AAAT3'
- FastDigest *BglII*, 5'A<sup>^</sup>GATCT3'
- T4 DNA Ligase (5 u/μL), Thermo Scientific Fermentas (Rockford, IL, USA)

### 3.2.3.2. DNA polymerases

- Phusion F - 530L DNA-Polymerase (2 u/μL), 5x HF reaction buffer (7,5 mM MgCl<sub>2</sub>); Finnzymes (Oy, Espoo, Finland)

### 3.2.4. Software and web tools

#### 3.2.4.1. Software

- Lasergene 7.0.0, EditSeq, SeqMan Pro, DNASTAR Inc. (Madison, WI, USA)
- Gene Designer, DNA 2.0 Inc. (Menlo Park, CA, USA)
- Vector NTI Suite 8.0, Invitrogen Corporation (Carlsbad, CA, USA)

#### 3.2.4.2. Web tools

- Sequence alignment with ClustalW2: <http://www.ebi.ac.uk/Tools/msa/clustalw2/>
- Expasy tool to translate DNA sequences: <http://expasy.org/tools/dna.html>
- Expasy tool to compute pI/Mw: [http://expasy.org/tools/pi\\_tool.html](http://expasy.org/tools/pi_tool.html)
- Expasy tool to calculate the molar extinction coefficient: <http://web.expasy.org/protparam/>
- EMBOSS 6.3.1 freak for base/residue frequency plots or tables: <http://mobyle.pasteur.fr/cgi-bin/portal.py?form=freak>
- GeneBee, prediction of RNA secondary structures: <http://www.genebee.msu.su/>
- NCBI BLAST: <http://www.ncbi.nlm.nih.gov/blast>
- Primer3Plus: <http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>
- Conversion of DNA strand to antiparallel: <http://www.fr33.net/seqedit.php>
- SignalP 4.0 Server, prediction of signal peptides: <http://www.cbs.dtu.dk/services/SignalP/>
- TMHMM Server, v. 2.0, prediction of transmembrane helices in proteins: <http://www.cbs.dtu.dk/services/TMHMM/>
- PSORT II, protein subcellular localization prediction: <http://psort.hgc.jp/form2.html>
- PrediSi, signal peptide prediction: <http://www.predisi.de/>

### 3.2.5. Standard molecular biological techniques

#### 3.2.5.1. Agarose gel electrophoresis

For the preparation of 1 % agarose gels, 2 g Biozym agarose were added to 200 mL 1 x TAE buffer and heated up in the microwave for 4 min. After cooling down, a drop of EtBr ( $\geq 98$  % EtBr) was added and the gel was poured in the chamber. 6 x DNA Loading Dye was added to the samples before they were pipetted in the slots of the gel. 5 - 10  $\mu$ L of the GeneRuler™ 1 kb DNA Ladder were loaded on each gel to allow the estimation of size and concentration of the samples [see Figure 23 (A)].

The voltage was set to 120 V and the electrophoresis was run for  $\sim$  50 min before the gels were checked using the GelDoc-It™ Imaging System. In case of preparative gels, the running time was extended to 2 h and the voltage was decreased to 90 V. At the Chroma 43, the desired bands were identified and cut out of the gel with a scalpel.

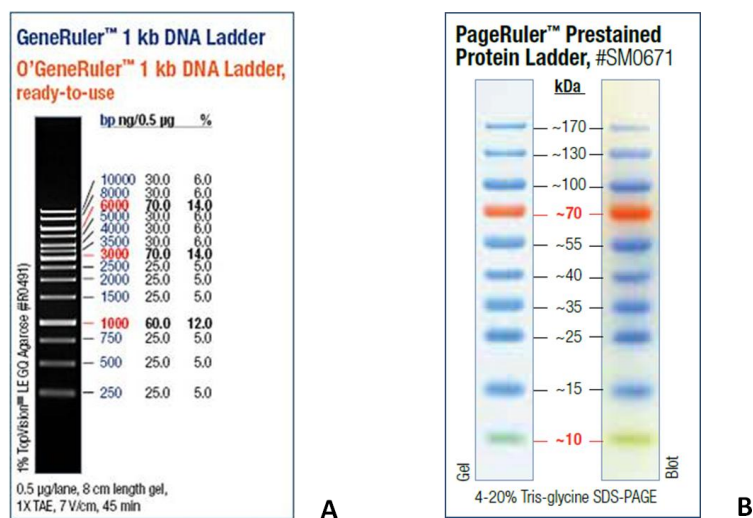


Figure 23: (A) Bands of the GeneRuler™ 1 kb DNA Ladder with the corresponding sizes given in basepairs ([www.fermentas.com](http://www.fermentas.com)); (B) Bands of the PageRuler™ Prestained Protein Ladder #SM0671 with the corresponding protein sizes given in kDa ([www.fermentas.com](http://www.fermentas.com)).

#### 3.2.5.2. SDS-PAGE

SDS-PAGE allows the separation of proteins according to their electrophoretic mobility which depends on their size and their charge. SDS, an anionic detergent, is used to denature and linearize the protein. After applying a voltage, the negatively charged molecules migrate to the anode, the positive pole, and cause a separation of the proteins due to their different molecular masses since the negative charges are distributed all over the proteins and have no influence anymore. A polyacrylamid gel, consisting of a stacking and a separating layer, serve as support medium. Smaller sized polypeptides are less restrained by the gel and can

move through it more easily. To visualize the protein bands, Coomassie blue or silver staining are frequently used methods to determine the molecular mass of proteins and their relative abundance within a sample (78).

For sample preparation, 4 x Loading Dye and 10 x Reducing Agent from Invitrogen™ life technologies (Lofer, Austria) were added to 10 µL protein sample and heated up to 70 °C for 10 min to denature the proteins. Precast gels and MOPS buffer (both from Invitrogen™ life technologies) were put in the electrophoresis chamber. The samples were loaded in the slots next to 5 µL of PageRuler™ Prestained Protein Ladder [see Figure 23 (B)]. The gel was run for 1 h and 10 min at 200 V and stained for 20 min with Coomassie blue SDS gel staining solution after shortly heating the gel in the microwave. For destaining the gel was either put in SDS gel destaining solution over night or heated shortly in the microwave and incubated on a shaker.

#### 3.2.5.3. MS - analysis

To identify protein bands after SDS-PAGE, the bands of interest were cut out of the gel with a scalpel and stored in 10 % ethanol. The samples were then sent for MS - analysis to the ZMF (Center for Medical Research) in Graz to determine the amino acid sequence of the proteins.

#### 3.2.5.4. Plasmid Miniprep Kit

For the isolation of purified plasmid DNA the GeneJet™ Plasmid Miniprep Kit from Fermentas GmbH (St. Leon-Rot, Germany) was used as described in the manufacturer's protocol. The following deviations were applied: cells were collected with a tooth pick from agar plates and after resuspension, cell lysis and neutralization the sample was centrifuged for 10 min instead of 5 min. For the elution of the DNA at the end ddH<sub>2</sub>O was used instead of elution buffer.

#### 3.2.5.5. Purification of DNA

The Wizard® SV Gel and PCR Clean-Up System from Promega GmbH (Mannheim, Germany) was used to purify DNA after PCRs, restriction cuts and preparative agarose gels. The protocol was followed as recommended by the manufacturer with small deviations: the incubation time for binding of the DNA as well as the centrifugation time to remove residual



ethanol before elution were extended to 5 min. For eluting the DNA the sample was incubated for 2 min with 40 µL ddH<sub>2</sub>O, which was warmed up to 45 °C and centrifuged.

#### 3.2.5.6. Dephosphorylation of DNA

For the dephosphorylation of vector DNA Shrimp Alkaline Phosphatase (SAP) was used as described in the producer's protocol ([www.thermoscientific.com/fermentas](http://www.thermoscientific.com/fermentas)). Five reaction mixtures were prepared for every vector, each containing 1 µg linear DNA:

**Table 4: Composition of the dephosphorylation reaction mixture with SAP.**

Components	Amount
linear DNA	1 µg
10 x reaction buffer	2 µL
SAP	1,5 µL
ddH <sub>2</sub> O	fill up to a total volume of 20 µL

The samples were mixed carefully, spinned down and incubated for 60 min at 37 °C. The reaction was stopped by heating the samples to 65 °C for 15 min. Afterwards the reaction mixtures were pooled and the DNA could be purified as described in 3.2.5.5.

#### 3.2.5.7. Ligation

For ligations 100 ng vector DNA and a threefold molar excess of the insert were used. The amount of the insert was calculated with the following formula:

$$\frac{ng\ vector * kb\ size\ of\ insert}{kb\ size\ of\ vector} * molar\ ratio\ of\ \frac{insert}{vector} = ng\ insert$$

T4 DNA Ligase and 10x T4 DNA Ligase Reaction Buffer were used as recommended by the manufacturer ([www.thermoscientific.com/fermentas](http://www.thermoscientific.com/fermentas)). 0.2 µL ligase and 2 µL of buffer were pipetted into a reaction mixture with insert and vector DNA and filled up to a total volume of 20 µL with dH<sub>2</sub>O. In most cases, ligation was performed at 16 °C over night. Alternatively the samples were incubated at 22 °C for 1 h followed by an incubation for 3 - 4 h at 16 °C. To stop the reaction, the ligase was inactivated at 65 °C for 10 min. Ligation controls containing only the vector but no insert were prepared as well.

#### 3.2.5.8. PCR

For PCR reactions Phusion DNA polymerase and primers in a final concentration of 200 nM were used (2 µL of 5 µM stocks). dNTP stock solutions (2 mM) were prepared, containing

dCTP, dGTP, dATP and dTTP in same amounts and added to a final concentration of 200  $\mu$ M (5  $\mu$ L of 2 mM or 1  $\mu$ L of 10 mM dNTP stocks). For each PCR reaction, four mixtures were prepared to ensure that enough DNA was obtained in the end. The annealing temperature was set at 58°C, at least 4 °C below the primer melting temperature, which was determined with EditSeq software (DNASTAR). The Phusion polymerase elongates 1 kb of plasmid DNA in 15 sec and 1 kb of genomic DNA in 30 sec. Therefore the elongation times were chosen depending on the length of the PCR product.

### 3.2.5.9. Primers

All primers that were used during the experiments are listed below together with their length and properties like melting temperature and GC - content.

**Table 5: A list of all primers including their number, name, sequence and properties according to the data sheets.**

Number	Name	Sequence [5' → 3']	T <sub>m</sub> [°C]	bp	GC [%]
P08445	AOXseq_rv	TCCCAAACCCCTACCAACAAG			
P08511	EcoRI-Cit-F	GAATTCGAAACGATGGCTAGCAAAGGAG			
P10120	moxYshort_seq_fw	GACGACTACCAAGAGTCACAATGG	57.2	24	50.0
P11144	CBH2_Linker_Citrine_rev	CTCCAGTGAAAAGTTCTTCTCCTTTGCTAGCACCG ACGGGTGGGACTCGGGTAG	71.0	54	55.5
P11145	CBH2_Linker_Citrine_fw	CTACCCGAGTCCCACCCGTCGGTGCTAGCAAACG AGAAGAACTTTTCACTGGAG	71.0	54	55.5
P11146	CBH2_Linker_moxY_rev	CGCAACGGGCGGTTGGCCGGGTCACCGACGGGT GGGACTCGGGTAG	77.6	46	73.9
P11147	CBH2_Linker_moxY_fw	CTACCCGAGTCCCACCCGTCGGTGACCCGGCCAA CCGCCCGTTGCG	77.6	46	73.9
P11148	CBH2_SpeI_fw	ACTTACTAGTAAAACGATGATTGTGGGAATTTGA CCACGCTC	64.8	43	41.8
P11149	moxY_NotI_rev	ATATGCGGCCGCTAGCGGTTACTGTCAGAAACT CCATTGG	69.1	41	53.6
P11150	alpha_F_moxY_XhoI_fw	TATACTCGAGAAGAGAGAGGCCGAAGCTGACCCG GCCAACCGCCCG	73.3	46	63.0
P11151	Citrine_NotI_rev	ATATGCGGCCGCTTACTTGTACAATTCATCCATGC CATGTG	66.6	41	46.3
P11152	alpha_F_Citrine_XhoI_fw	ATATCTCGAGAAGAGAGAGGCCGAAGCTGCTAGC AAAGGAGAAGAACTTTTCACTG	68.3	56	46.4
P11153	MBP_no_linker_NotI_rev	ATATGCGGCCGCTTAGGTTTGTGCATCCTTCAAAG CCTC	68.2	39	51.2
P11154	Spe_moxY_fw	ATATACTAGTGCCACCATGGACCCGGCCAACCGC CCG	71.6	37	62.1
P11155	moxY_MBP_connection_rev	GATCCAAATCACTAACTTACCTTCTCGATCTTGC GGTFACTGTGCAAACTCCATTGGAACC	68.3	63	44.4
P11156	moxY_MBP_connection_fw	GGTTCCAATGGAGTTTCTGACAGTAACCGCAAGAT CGAGGAAGGTAAGTTAGTGATTTGGATC	68.1	63	44.4
P11157	MBP_no_linker_moxY_rev	CAACGGGCGGTTGGCCGGGTCGGTTTGTGCATCC TTCAAAGCCTCGTC	74.1	48	62.5
P11158	MBP_no_linker_moxY_fw	GACGAGGCTTTGAAGGATGCACAAACCGACCCGG CCAACCGCCCGTTG	74.1	48	62.5
P11159	MBP_linker_moxY_rev	CAACGGGCGGTTGGCCGGGTCAGCTTGAAAGTAA AGGTTCTCAGGACCCTG	72.7	51	58.8

P11160	MBP_linker_moxY_fw	CAGGGTCCTGAGAACCTTTACTTTCAAGCTGACCC GGCCAACCGCCCGTTG	72.7	51	58.8
P11161	SpeI_HIS_MBP_fw	ATATACTAGTGCCACCATGCACCATCACCACCATC ACAAGATCGAGGAAGGTAAGTTAGTGATTTGGATC	69.3	70	44.2
P11162	SpeI_MBP_fw	ATATACTAGTGCCACCATGAAGATCGAGGAAGGTA AGTTAGTGATTTGG	64.8	49	40.8
P11178	MBP_seq_fw1	GACGGAGGTTATGCTTTCAAGTAC	55.6	24	45.8
P11179	MBP_seq_fw2	GCCGCTACTATGGAGAACGCAC	60.1	22	59.0
P11180	MBP_seq_rev1	CAGTGTGACGCTTCATGTGCTTG	58.9	23	52.1
P11181	MBP_seq_rev2	TCCAACCTGTGACGGTGCTCG	59.6	21	57.1
P11206	MBP_linker_Citrin_fw	CAGGGTCCTGAGAACCTTTACTTTCAAGCTGCTAG CAAAGGAGAAGAACCTTTCACTGGAG	68.7	61	45.9
P11207	MBP_linker_Citrin_rev	CTCCAGTGAAAAGTTCTTCTCCTTTGCTAGCAGCT TGAAAGTAAAGTTCTCAGGACCCTG	68.7	61	45.9
P11208	MBP_no_linker_Citrin_fw	GACGAGGCTTTGAAGGATGCACAAACCGCTAGCA AAGGAGAAGAACCTTTCACTGGAG	69.6	58	48.2
P11209	MBP_no_linker_Citrin_rev	CTCCAGTGAAAAGTTCTTCTCCTTTGCTAGCGGTT TGTGCATCCTTCAAAGCCTCGTC	69.6	58	48.2
P11210	Citrin_MBP_connection_fw	GATTACACATGGCATGGATGAATTGTACAAGAAGA TCGAGGAAGGTAAGTTAGTGATTTGGATC	66.6	64	39.0
P11211	Citrin_MBP_connection_rev	GATCCAAATCATAACTTACCTTCTCGATCTTCTT GTACAATTCATCCATGCCATGTGTAATC	66.6	64	39.0
P12060	SpeI_Citrin_fw	ATATACTAGTGAAACGATGGCTAGCAAAGGAGAAG AACTTTTAC	63.1	45	37.7
P12061	XhoI_alpha_His_MBP_fw	ATATCTCGAGAAGAGAGAGGCCGAAGCTCACCAT CACCACCATCACAAGATCGAGGAAGGTAAGTTAGT GATTTGG	70.3	76	47.3
P12062	XhoI_alpha_MBP_fw	ATATCTCGAGAAGAGAGAGGCCGAAGCTAAGATC GAGGAAGGTAAGTTAGTGATTTGG	67.4	58	44.8

### 3.2.5.10. Transformation of *E. coli* Top10 cells

100 µL of electro-competent *E. coli* Top10 cells were slowly thawed on ice and carefully mixed with 5 µL of the ligation reaction. The mixture was transferred into cooled electroporation cuvettes and incubated on ice for at least 5 min. After pulsing the cells with a voltage of 2.5 kV, 200 Ω and 25 µF, 900 µL SOC medium were added directly after the electro shock. To regenerate the cells, they were incubated at 37 °C and 700 rpm for 1 h. After the regeneration, 100 µL of the cells as well as the resuspended rest were plated on LB - agar plates containing the for the selection required antibiotic (25 µg/mL zeocin) and incubated over night at 37 °C.

### 3.2.5.11. Sequencing

To confirm the correct sequence of the prepared constructs and to identify positive clones, the plasmid DNA of obtained *E. coli* clones was isolated and sent for sequencing to LGC Genomics GmbH (Berlin, Germany). Therefore, 10 µL of DNA with a concentration of ~100 ng/µL were needed as well as 4 µL of the respective primer with a concentration of 5 µM. Primers were added either directly to the sample or provided by LGC Genomics GmbH.

#### 3.2.5.12. Transformation of *P. pastoris* cells

The preparation of competent *P. pastoris* cells and the subsequent transformation was done according to the condensed protocol by Lin-Cereghino *et al.* (79) to prepare competent cells. First of all, a preculture with 50 mL YPD - media was inoculated with cells of the *P. pastoris* Mut<sup>S</sup> strain and incubated over night at 28°C and 110 rpm. The next day, the OD of the preculture at 600 nm was measured in 1:20 diluted samples and 50 mL YPD - media of the main culture were inoculated with the preculture to a final OD<sub>600</sub> of 0.2. The main culture was grown at 28 °C and 110 rpm for ~ 4 h until an OD<sub>600</sub> of 0.8 - 1.0 was reached. The cells were harvested via centrifugation at 500 x g for 5 min and constantly kept on ice. The supernatant was discarded and the cells were resuspended in 9 mL ice cold BEDS solution and 1 mL DTT (1M). The cells were carefully shaken by hand for 5 min and afterwards centrifuged for 5 min at 500 x g. Finally the cells were resuspended in 1 mL ice cold BEDS solution and aliquots of 100 µL were made. For the transformation, 1 µg DNA was added to 100 µL of cells. The mixture was transferred into electroporation cuvettes and incubated on ice for at least 5 min. The cells were pulsed with an electro shock of 1.5 kV, 200 Ω and 25 µF. Immediately afterwards 500 µL 1 M sorbitol as well as 500 µL YPD - media were added. The cells were regenerated at 28 °C and 110 rpm for 2 h. 50 µL, 100 µL, 200 µL and the rest of the cells were plated on YPD - zeocin agar plates and incubated at 28 °C for two days.

#### 3.2.5.13. Micro - scale cultivation of *P. pastoris* in 96 - deep well plates

Micro - scale cultivation of *P. pastoris* was performed in 96 - deep well plates as described by Weis *et al.* (80). Each well was filled with 250 µL BMD 1 % media and inoculated with individual colonies from agar plates. Two wells were left out to have sterile controls, whereas two wells were inoculated with the empty *P. pastoris* Mut<sup>S</sup> strain as negative controls and other two with positive controls, if available. The plates were incubated at 28 °C, 320 rpm and 80 % humidity for 60 h before the protein expression was started by the addition of 250 µL BMM 2. Before inducing the enzyme expression the strains were stamped on quadratic YPD zeocin agar plates to preserve all clones. 12 h, 24 h and 48 h after starting the induction, 50 µL BMM 10 were added. 24 h after the last induction, the cells were either harvested via centrifugation of the plate at 3000 x g for 10 min or directly used for further investigations.

#### 3.2.5.14. Cultivation of *P. pastoris* in 2 L baffled shake flasks

200 mL BMD 1 % media were filled in 2 L baffled shake flasks, inoculated with a single colony taken from an agar plate and incubated on a shaker for 60 h at 28 °C, 80 % humidity

and 110 rpm. 20 mL BMM 10 were added to induce the protein expression and 1 mL MeOH was added 12 h, 24 h and 48 h after starting the induction. After another 24 h the cells were harvested via centrifugation.

#### 3.2.5.15. Bradford assay for determination of the protein concentration

The assay (from Bio-Rad Laboratories GmbH, Vienna, Austria) was performed according to the manufacturer's recommendations. The Bio-Rad Bradford solution was diluted 1:5 and 200  $\mu\text{L}$  of the solution were added to 10  $\mu\text{L}$  of the sample in transparent micro titer plates. After an incubation time of at least 5 min the absorbance at 595 nm was measured with a spectrophotometer. If possible, double determinations were carried out. Albumin protein standards with concentrations of 0,0625 ng/ $\mu\text{L}$ , 0,125 ng/ $\mu\text{L}$ , 0,25 ng/ $\mu\text{L}$ , 0,5 ng/ $\mu\text{L}$ , 1,0 ng/ $\mu\text{L}$  and 2,0 ng/ $\mu\text{L}$  were used to determine the calibration curve and  $\text{dH}_2\text{O}$  served as a blank.

#### 3.2.5.16. TCA precipitation

For protein precipitation, TCA solution (~ 100 % v/v) was added to the protein sample to a final TCA concentration of 10 %. After incubation on ice for 1 h the samples were centrifuged at max. speed for 15 min at 4 °C. The pellet was washed twice with ice cold acetone and centrifuged for 5 min after each step. After evaporation of residual acetone, 7  $\mu\text{L}$  SDS-PAGE sample buffer were added and the samples were heated up to 70 °C for 10 min.

#### 3.2.5.17. Spectrophotometric measurements of fluorescence and $\text{OD}_{600}$

The  $\text{OD}_{600}$  of the cell suspensions was measured spectrophotometrically at a wavelength of 600 nm. To avoid saturation of the signal, the samples were diluted correspondingly.

This applied also for determination of the fluorescence of citrine in the cell suspensions or supernatants after centrifugation. For fluorescence measurements with the platereader the following settings were used: excitation wavelength 500 nm, emission wavelength 550 nm, cut off filter 530 nm, sensitivity: 13, automix before measurements: 5 sec.

#### 3.2.5.18. Fluorescence microscopy

500  $\mu\text{L}$  fresh cells were centrifuged, washed with  $\text{ddH}_2\text{O}$  and resuspended in the same volume as before. 1  $\mu\text{L}$  of washed cells was fixed on a microscope slide with a cover glass. Before the microscope could be used, the lamp had to be turned on 30 min in advance. Phase contrast as well as fluorescence images were made and saved in Corel Photo-Paint X3. Therefore the filter set I3 with the excitation wavelength filter BP 450 – 490 and the

suppression wavelength LP 515 were used and the rotating wheel was set to pH3 for the 100x pH3 objective.

### 3.3. Experimental

#### 3.3.1. Fluorescence microscopy study

In order to analyze the localization of *moxY* already available strains were used, where *moxY* is connected to citrine to enable the visualization via fluorescence. Therefore, the following *P. pastoris* strains were cultivated in baffled 2 L shake flasks as described in chapter 3.2.5.14.

- CBS 7435 mut<sup>S</sup> (strain collection nr. 3445) (empty strain as negative control)
- *P.p.* pPic ZB\_citrine m.c. (strain collection nr. 3300) (citrine for comparison)
- *P.p.* mut<sup>S</sup> B1\_moxYlong\_citrine\_his
- *P.p.* mut<sup>S</sup> B1\_his\_citrine\_moxYlong

Further strains were cultivated in a smaller scale in DWPs (see chapter 3.2.5.13) and phase contrast as well as fluorescence images were taken with a microscope.

- *P.p.* mut<sup>S</sup> B1\_moxYshort\_citrine\_strep
- *P.p.* mut<sup>S</sup> moxY\_IT\_citrine\_his
- *P.p.* mut<sup>S</sup> B1\_citrine

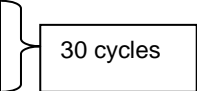
Since citrine was only available in the pPIC vector, a new construct was prepared where citrine was cloned into the B1 vector, the vector which was subsequently used for all other experiments. Preparation of the B1 citrine construct: A new primer [P12060, see Table 5] with a *SpeI* restriction site, the Kozak consensus sequence and the first bases of the citrine sequence was designed for the amplification of the citrine insert via PCR, the cloning and transformation to *E. coli*.

**Table 6: PCR reaction mixture to amplify citrine.**

	PCR for B1 citrine	
DNA template	B1_moxYlong_citrine [10 ng/μL]	1 μL
fw primer	P12060: <i>SpeI</i> _Citrin_fw	2 μL
rev primer	P11151 Citrine_ <i>NotI</i> _rev	2 μL
	Phusion polymerase	0.5 μL
	dNTP 2 mM	5 μL
	HF buffer 5 x	10 μL
	d H <sub>2</sub> O	29.5 μL

**Table 7: Temperature program for the PCR to amplify citrine.**

	<b>temperature</b>	<b>time</b>
initial denaturation	98 °C	30 sec
denaturation	98 °C	10 sec
annealing	58 °C	20 sec
extension	72 °C	15 sec
final extension	72 °C	7 min
	4 °C	∞



The components of the PCR reaction [Table 6Table 1] were pipetted in small tubes, mixed carefully and placed in the PCR machine using the above mentioned temperature program [Table 7]. After making a preparative agarose gel and purification, the DNA was cut with 2  $\mu$ L *SpeI*, 3  $\mu$ L *NotI* and 5.7  $\mu$ L Buffer G for 5 h at 37 °C. The cut insert was purified via a column and ligated with the *SpeI* and *NotI* cut, dephosphorylated B1 vector over night at 16 °C as described in chapter 3.2.5.7.

The ligated plasmid DNA was transformed into competent *E. coli* Top10 cells (see 3.2.5.10). Positive clones were identified via *SpeI/NotI* FastDigest restriction cuts followed by control agarose gels and confirmed via sequencing by LGC Genomics, using the 5AOX primer provided by LGC Genomics. The plasmid DNA was linearized with *BglII* (3  $\mu$ L *BglII* FD and 5.5  $\mu$ L FD buffer 10x white, incubation at 37 °C for 3 h) and purified prior to transformation into *P. pastoris* cells. *P. pastoris* clones were cultivated in 96 - deep well plates and the clone with the highest value of fluorescence in the supernatant was identified with the fluorescence spectrophotometer (see 3.2.5.17). This clone was finally analyzed under the fluorescence microscope together with the other *moxY\_citrine* strains.

### 3.3.2. Secretion study

The aim was to test, whether *moxY* can be secreted with the  $\alpha$ -factor signal sequence from *S. cerevisiae* or the CBH2 secretion signal. In order to compare the efficiency of both, constructs with citrine instead of *moxY* were prepared as well.

#### 3.3.2.1. Preparation of the *moxY/citrine* constructs with the $\alpha$ -factor signal sequence

First of all, primers were designed to facilitate the cloning of *moxY* and citrine into a pPpT4 $\alpha$ S vector [see Figure 24]. The forward primers contained a *XhoI* cleaving site (green) followed by the nucleotides coding for the amino acids Lys-Arg-Glu-Ala-Glu-Ala, which represented the specific Kex2 and Ste13 cleavage sites (light blue). Afterwards, the first amino acids encoding the protein of interest were added without their own start codon, since the

methionine, where the expression of the whole reading frame started, was already provided at the vector. For the reverse primer, the non-coding strand was used in the opposite reading direction and a *NotI* cleavage site (red) was added after the protein of interest to allow cloning with *XhoI* and *NotI* restriction enzymes.

P11150: alpha\_F\_moxY\_XhoI\_fw

TATACTCGAGAAAGAGAGAGGCCGAAGCTGACCCGGCCAACCGCCCG

P11149: moxY\_NotI\_rev

ATATGCGGCCGCCTAGCGGTTACTGTCAGAACTCCATTGG

P11152 alpha\_F\_Citrine\_XhoI\_fw

ATACTCGAGAAAGAGAGAGGCCGAAGCTGCTAGCAAAGGAGAAGAAGACTTTTCACTG

P11151 Citrine\_NotI\_rev

ATATGCGGCCGCCTACTTGTACAATTCATCCATGCCATGTG

At the 5'-end four nucleotides consisting of A and T were added to each primer, to increase the cleavage efficiency of the restriction enzymes. Furthermore, either a G or C was positioned at the 3'-end for stability reasons. The melting temperatures of the primers were set in such a way that  $T_m$  was at least 4 °C above the annealing temperature during PCR which is 58 °C and that the melting temperatures of the forward and reverse primers were similar. The GC content was kept at 45 - 55 % and, if possible, the formation of interfering domains or hairpins within the primer sequence was avoided. An amount of 100 nmol was ordered for each primer, purified via SDS - PAGE.

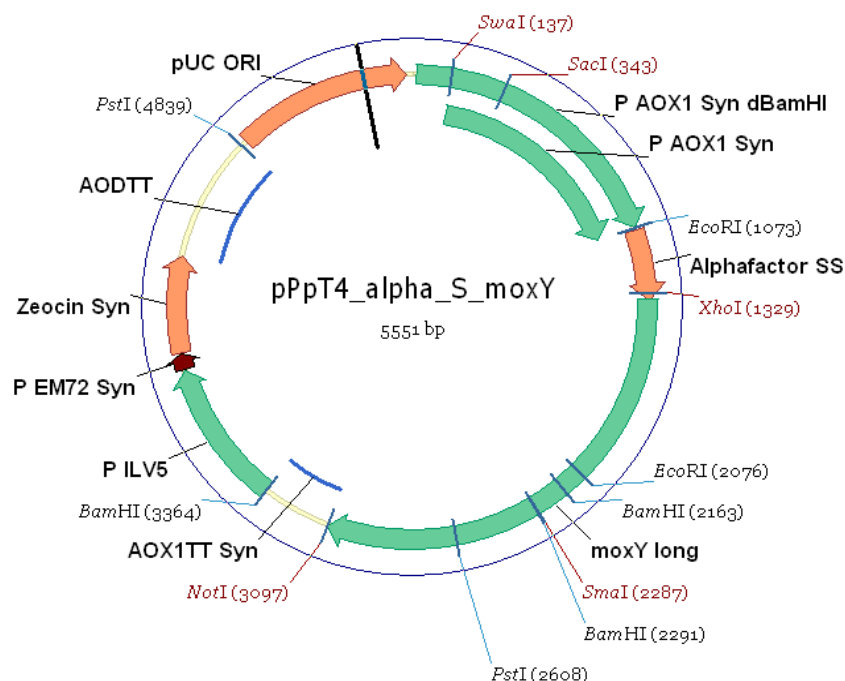
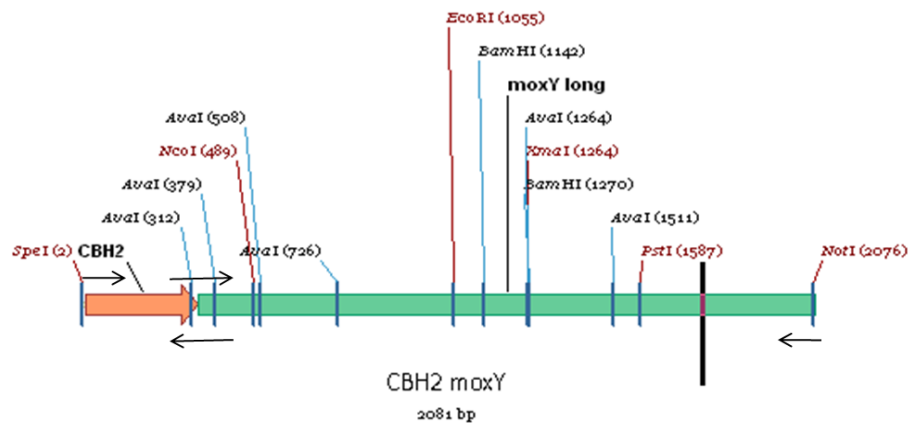


Figure 24: The vector pPpT4 with the  $\alpha$ -factor signal sequence and the gene encoding for moxY long.



### 3.3.2.2. Preparation of *moxY*/citrine constructs with the CBH2 signal peptide:

The second construct prepared for the secretion study consisted of *moxY* fused to the CBH2 signal peptide. Primers have been designed to amplify the CBH2 signal sequence and *moxY* and to link both genes afterwards via overlap extension PCR (oePCR). The construct was cloned into the vector T4 Smi with the restriction sites *SpeI* and *NotI* [see Figure 25]. The same applied for the construct with citrine.



**Figure 25:** The construct for the secretion of *moxY* with the CBH2 signal sequence. The primers necessary for the preparation and cloning of the construct are indicated as arrows.

Therefore, a primer was needed with a *SpeI* cleavage site at the 5'-end before a Kozak consensus sequence followed by the sequence encoding for CBH2, starting with methionine. Forward and reverse primers were designed, which overlapped at the connecting region between CBH2 and the protein of interest, so either *moxY* or citrine [see primers P11144 to P11148 in Table 5]. In order to guarantee that the binding is strong enough, from each site of the overlapping region at least 20 base pairs were taken which led to primers with a length of ~ 50 bp and a melting temperature of ~ 90 °C.

Preparation of the vectors and inserts:

The templates for the T4 $\alpha$ S and the T4 Smi vector as well as for CBH2 (using the codon optimized CBH2 variant V31 obtained from Andrea Mellitzer) were streaked out on LB-zeocin plates and incubated over night at 37 °C. After isolation of the plasmid DNA (see 3.2.5.4) the vectors were cut with the respective restriction enzymes as described in Table 8.

**Table 8: Composition of the reaction mixtures for the restriction cuts of the T4 $\alpha$ S and the T4 Smi vector.**

	$\alpha$ - factor signal sequence		CBH2	
	template	T4 $\alpha$ S vector	47 $\mu$ L	T4 Smi vector
restriction enzyme	<i>XhoI</i>	3 $\mu$ L	<i>SpeI</i> (= <i>BcuI</i> )	2 $\mu$ L
	<i>NotI</i>	2 $\mu$ L	<i>NotI</i>	3 $\mu$ L
buffer	Buffer O (10x)	5.7 $\mu$ L	Buffer G (10x)	5.7 $\mu$ L

The samples were incubated at 37 °C for at least 5 h and loaded on a preparative agarose gel for electrophoresis. Compared with the size of the total vector of 3.7 kbp, the 37 bp region that was cut out between the two restriction sites was too small to be identified. In order to avoid religation of only on one site cut vectors, the DNA samples were dephosphorylated. The bands were cut out of the gel and the DNA was purified as described in 3.2.5.5 and dephosphorylated with SAP (see 3.2.5.6). After purification of the DNA, the concentration was determined with the NanoDrop and a control gel was done via agarose gel electrophoresis to check the size of the vector. For the preparation of the secretion constructs with the  $\alpha$ -factor signal sequence and CBH2 signal sequence the PCR reactions for *moxY* [Table 9] and citrine [Table 10] were performed using the temperature programs described in Table 11.

**Table 9: PCR reactions for the secretion constructs of *moxY* with the  $\alpha$ -factor signal sequence and CBH2 signal peptide.**

	$\alpha$ -factor signal sequence <i>moxY</i>		CBH2 <i>moxY</i>			
	reaction 1		reaction 2		reaction 3	
DNA template	B1_ <i>moxY</i> long_citrine [10 ng/ $\mu$ L]	1 $\mu$ L	B1_CBH2_V31 [10ng/ $\mu$ L]	1 $\mu$ L	B1_ <i>moxY</i> long_citrine [10 ng/ $\mu$ L]	1 $\mu$ L
fw primer	P11150: alpha_F_ <i>moxY</i> _ <i>XhoI</i> _fw	2 $\mu$ L	P11148: CBH2_ <i>SpeI</i> _fw	2 $\mu$ L	P11147: CBH2_Linkers_ <i>moxY</i> _fw	2 $\mu$ L
rev primer	P11149: <i>moxY</i> _ <i>NotI</i> _rev	2 $\mu$ L	P11146: CBH2_Linkers_ <i>moxY</i> _rev	2 $\mu$ L	P11149: <i>moxY</i> _ <i>NotI</i> _rev	2 $\mu$ L
	Phusion polymerase	0.5 $\mu$ L	Phusion polymerase	0.5 $\mu$ L	Phusion polymerase	0.5 $\mu$ L
	dNTP 10 mM	1 $\mu$ L	dNTP 10 mM	1 $\mu$ L	dNTP 10 mM	1 $\mu$ L
	HF buffer 5 x	10 $\mu$ L	HF buffer 5 x	10 $\mu$ L	HF buffer 5 x	10 $\mu$ L
	d H <sub>2</sub> O	33.5 $\mu$ L	d H <sub>2</sub> O	33.5 $\mu$ L	d H <sub>2</sub> O	33.5 $\mu$ L

**Table 10: PCR reactions for the secretion constructs of citrine with the  $\alpha$ -factor signal sequence and CBH2 signal peptide.**

	$\alpha$ -factor signal sequence citrine		CBH2 citrine			
	reaction 4		reaction 5		reaction 6	
DNA template	B1_ <i>moxY</i> long_citrine [10 ng/ $\mu$ L]	1 $\mu$ L	B1_CBH2_V31 [10ng/ $\mu$ L]	1 $\mu$ L	B1_ <i>moxY</i> long_citrine [10 ng/ $\mu$ L]	1 $\mu$ L
fw primer	P11152 alpha_F_Citrine_ <i>XhoI</i> _fw	2 $\mu$ L	P11148: CBH2_ <i>SpeI</i> _fw	2 $\mu$ L	P11145: CBH2_Linkers_Citrine_fw	2 $\mu$ L
rev primer	P11151 Citrine_ <i>NotI</i> _rev	2 $\mu$ L	P11144: CBH2_Linkers_Citrine_rev	2 $\mu$ L	P11151 Citrine_ <i>NotI</i> _rev	2 $\mu$ L
	Phusion polymerase	0.5 $\mu$ L	Phusion polymerase	0.5 $\mu$ L	Phusion polymerase	0.5 $\mu$ L
	dNTP 10 mM	1 $\mu$ L	dNTP 10 mM	1 $\mu$ L	dNTP 10 mM	1 $\mu$ L
	HF buffer 5 x	10 $\mu$ L	HF buffer 5 x	10 $\mu$ L	HF buffer 5 x	10 $\mu$ L
	d H <sub>2</sub> O	33.5 $\mu$ L	d H <sub>2</sub> O	33.5 $\mu$ L	d H <sub>2</sub> O	33.5 $\mu$ L

**Table 11: Temperature program for the PCR reactions to prepare the secretion constructs of moxY and citrine with the  $\alpha$ -factor signal sequence and the CBH2 signal peptide.**

	Temperature	Time
initial denaturation	98 °C	30 sec
denaturation	98 °C	10 sec
annealing	58 °C	20 sec
extension	72 °C	30 sec (reaction 1 and 3: moxY 1,7 kbp)
		15 sec (reaction 4 and 6: citrine 700 bp)
		10 sec (reaction 2 and 5: CBH2 300 bp)
final extension	72 °C	7 min
	4 °C	$\infty$

30 cycles

The samples were loaded on a preparative agarose gel and the bands corresponding to moxY (1.7 kbp), citrine (700 bp) and the CBH2 secretion signal (300 bp) were cut out and purified (see 3.2.5.1 and 3.2.5.5). The insert for the constructs with the  $\alpha$ -factor signal sequence could already be cloned into the vector but the inserts with CBH2 had to be connected via oePCR beforehand. The obtained PCR products were diluted to a concentration of 1 ng/ $\mu$ L and served as templates and primers for the first round of the oePCR [see Table 12 and Table 13]. To provide the same molar ratios, the amount of the various components was chosen depending on their length; since moxY is approximately five times larger than CBH2, 5 ng of moxY and 1 ng of CBH2 were used.

**Table 12: Reaction mixtures for the oePCRs to prepare the secretion constructs of moxY and citrine with the CBH2 signal peptide, using the products of the PCRs in Table 9 and Table 10.**

CBH2 moxY		CBH2 citrine	
oePCR 1		oePCR 2	
CBH2 product of reaction 2 (1 ng)	1 $\mu$ L	CBH2 product of reaction 5 (2.5 ng)	2.5 $\mu$ L
moxY product of reaction 3 (5 ng)	5 $\mu$ L	citrine product of reaction 6 (5 ng)	5 $\mu$ L
Phusion polymerase	0.5 $\mu$ L	Phusion polymerase	0.5 $\mu$ L
dNTP 10 mM	1 $\mu$ L	dNTP 10 mM	1 $\mu$ L
HF buffer 5 x	10 $\mu$ L	HF buffer 5 x	10 $\mu$ L
d H <sub>2</sub> O	32.5 $\mu$ L	d H <sub>2</sub> O	31 $\mu$ L

**Table 13: Temperature program for the oePCR reactions to prepare the secretion constructs of moxY and citrine with the CBH2 signal peptide.**

	temperature	time
initial denaturation	98 °C	30 sec
denaturation	98 °C	10 sec
annealing	58 °C	20 sec
extension	72 °C	30 sec (oePCR 1: CBH2 moxY 2 kbp)
	72 °C	15 sec (oePCR 2: CBH2 citrine 1 kbp)
final extension	72 °C	7 min
	4 °C	$\infty$

15/20 cycles

After 15 cycles, a mastermix solution was added, containing the forward and reverse primers which allowed the amplification of the whole constructs [see Table 14]. The PCR was run for another 20 cycles. After completing the oePCR, the samples were loaded on a preparative agarose gel and the DNA bands were cut out and purified.

**Table 14: Composition of the mastermix, added for the amplification of the whole constructs for secretion of moxY and citrine with the CBH2 signal peptide.**

	CBH2 constructs	
fw primer	P11148: CBH2_SpeI_fw	4 µL
rev primer for CBH2 moxY oePCR 1	P11149: moxY_NotI_rev	4 µL
rev primer for CBH2 citrine oePCR 2	P1115:1 Citrine_NotI_rev	4 µL
	Phusion polymerase	0.5 µL
	dNTP 10 mM	1 µL
	HF buffer 5 x	10 µL
	d H <sub>2</sub> O	30.5 µL

### Secretion of MBP fusion proteins with the $\alpha$ -factor signal sequence

To test the influence of the MBP-tag on the protein secretion, constructs were made with the  $\alpha$ -factor signal peptide, followed by the His-tag, MBP and moxY or citrine. Therefore, the primer P12061 [see Table 5], was designed to amplify the N-terminal part of the fusion protein with MBP. This primer contained a *XhoI* restriction site for cloning in the pPpT4\_ $\alpha$ \_S vector followed by the Lys-Arg-Glu-Ala-Glu-Ala sequence for processing of the signal peptide by Kex2 and Ste 13. Subsequently, the MBP fragment was linked to moxY or citrine via oePCR. The detailed compositions and temperature programs of the PCR reactions are listed in chapter 4.3.2 [see Table 16 to Table 25], where the preparation of all MBP constructs is described.

The strains of both secretion constructs (with moxY and citrine) were added to the strain collection. So far,  $\alpha$  His MBP moxY was not used for further experiments whereas the secretion of  $\alpha$  His MBP citrine in *P. pastoris* was measured spectrophotometrically and compared to the secretion of citrine without the MBP-tag.

### Molecular cloning:

The insert PCR products for the  $\alpha$ -factor signal sequence and the CBH2 constructs were cut with restriction enzymes in the same way as described for the vectors [see Table 8]. Before preparing the ligation reactions, the correct sizes of the vectors and inserts were checked via agarose gel electrophoresis. The inserts for secretion with the  $\alpha$ -factor signal sequence were cut with *XhoI* and *NotI* and ligated with the cut and dephosphorylated T4 $\alpha$ S vector, whereas the oePCR products with CBH2 were cut with *SpeI* and *NotI* and ligated with the cut and dephosphorylated T4 Smi vector (see 3.2.5.7).

The ligated plasmid DNA was transformed into *E. coli* Top10 cells as described in 3.2.5.10. Several clones were chosen, struck out on LB - zeocin plates and their DNA was isolated with the Miniprep Kit. Control cuts were performed to ensure that the insert was incorporated

into the vector. Therefore 1  $\mu\text{L}$  DNA of the *moxY* constructs was cut with 0.5  $\mu\text{L}$  *EcoRI* FastDigest using 10x Green FastDigest Buffer for 20 min at 37 °C. For the citrine constructs additionally 0.5  $\mu\text{L}$  *BamHI* FastDigest were added. Via agarose gel electrophoresis the clones containing the vector with the insert could be identified. To ensure that the whole sequence was correct a positive clone of each construct was sent to LGC Genomics for sequencing (using the 5 AOX primer provided by LGC Genomics and primer P08445 as well as P10120 for the *moxY* constructs). Clones with correct sequences could be found for all constructs, a silent mutation at an Arg (G→T) in the citrine sequence was identified, but this mutation was already present in the template and did not have any effects. The positive *E. coli* clones were struck out on LB - zeocin plates. Their plasmid DNA was isolated and linearized using 3  $\mu\text{L}$  FastDigest *SmlI* (*Swal*) restriction enzyme and 5.5  $\mu\text{L}$  of 10x FastDigest buffer (white). After incubation at 37 °C for 3 h an agarose gel electrophoresis was performed to confirm the successful linearization and the DNA was purified via a column. The concentration was determined and the linearized and purified DNA was finally ready for transformation to *P. pastoris* cells as described in 3.2.5.12. The clones grown on the YPD-zeocin agar plates were subsequently cultivated in 96-deep well plates (see chapter 3.2.5.13) for screening of enzyme secretion.

#### Screening:

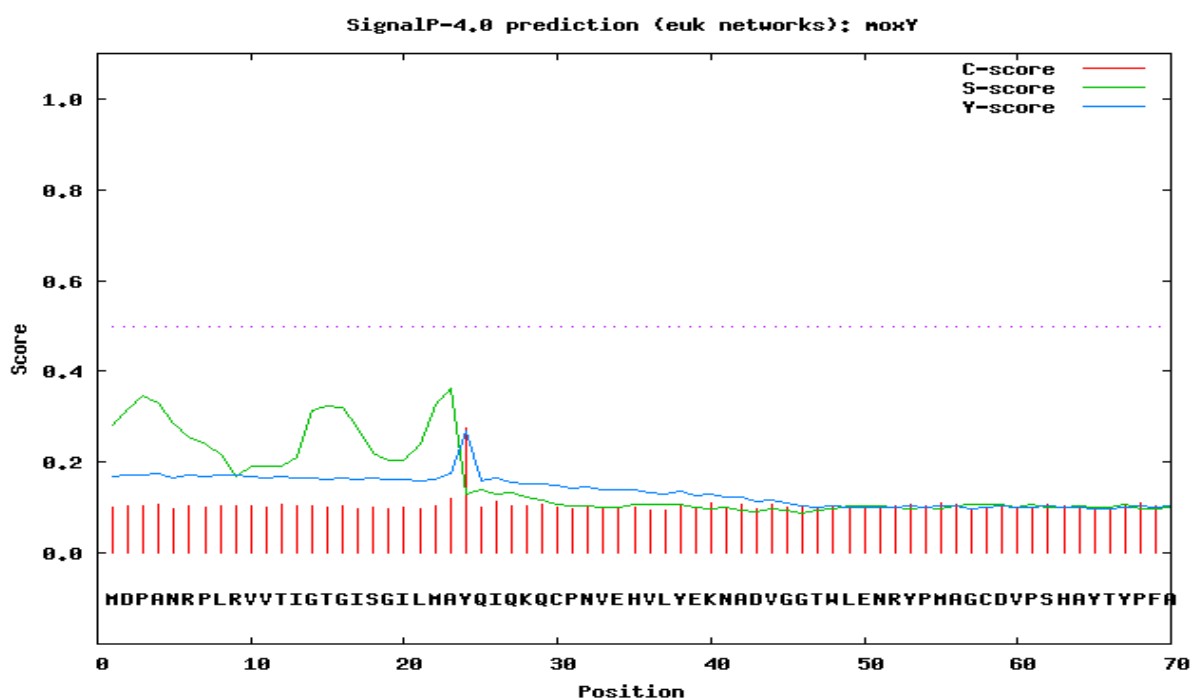
Screening for the secretion of citrine: Five DWPs were cultivated of each construct and the  $\text{OD}_{600}$  as well as the fluorescence of the cell suspensions was measured spectrophotometrically as described in chapter 3.2.5.17. Therefore, 10  $\mu\text{L}$  of the cell suspensions were diluted 1:10 with 90  $\mu\text{L}$   $\text{dH}_2\text{O}$ . Subsequently, the plates were centrifuged at full speed (3220 x g) for 10 min and 100  $\mu\text{L}$  of the supernatants were transferred in micro titer plates to determine the fluorescence of the supernatants.

Screening for the secretion of *moxY*: The deep well plates were centrifuged at full speed for 5 min and 300  $\mu\text{L}$  of the supernatant of randomly chosen clones were used for protein analysis via SDS-PAGE and determination of the protein concentration using the Bradford assay (see chapter 3.2.5.15). To ensure that the amount of protein, which was loaded on precast SDS gels was high enough, protein precipitation with TCA solution was performed as described in chapter 3.2.5.16.

### 3.4. Results and discussion

#### 3.4.1. *In silico* predictions of the localization of moxY

So far, the exact localization of moxY is not known. In case that the enzyme is membrane bound, this might explain difficulties in earlier purification trials. *In silico* studies with software programs and web tools could not make any clear predictions concerning the localization of moxY. No secretion signal was detected in the sequence for example with the webtool SignalP-4.0 [see Figure 26]. Thereby, the S-score represented the signal peptide prediction for every amino acid with high scores for those residues that are part of the signal peptide. A high C-score indicated the position of a cleavage site. For Y-max, these two scores were combined, indicating a possible cleavage site between the residues 23 and 24 but with very low values. The D-score was used for discrimination between secretory and non-secretory proteins, with low scores for non-secretory proteins, as it was the case for moxY (81).



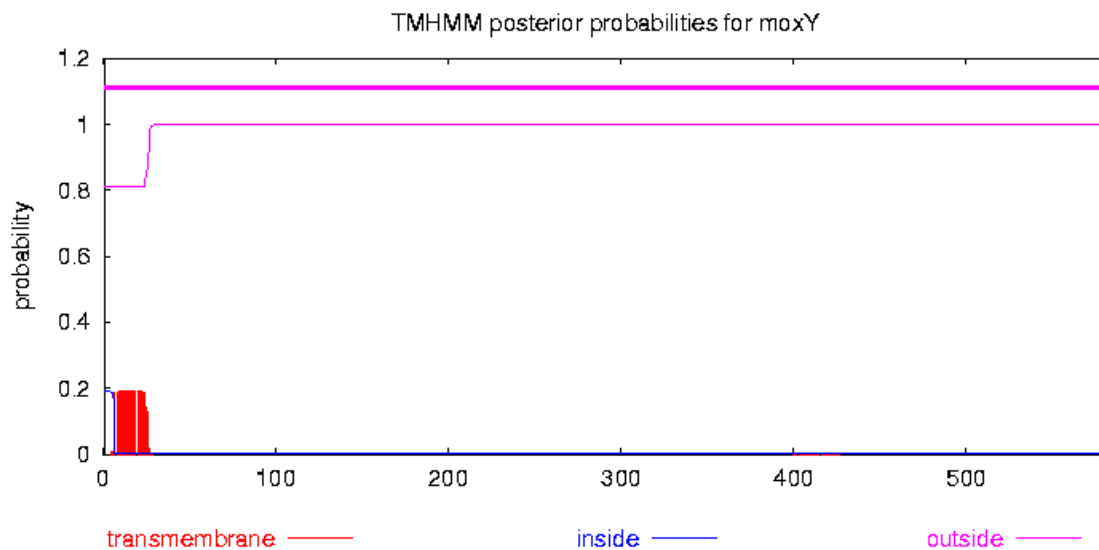
Measure	Position	Value	Cutoff	signal peptide?
max. C	24	0.276		
max. Y	24	0.268		
max. S	23	0.362		
mean S	1-23	0.262		
D	1-23	0.265	0.450	NO

Name=moxY SP='NO' D=0.265 D-cutoff=0.450 Networks=SignalP-noTM  
 citation: SignalP 4.0: discriminating signal peptides from transmembrane regions Petersen TN., Brunak S., von Heijne G. & Nielsen H. Nature Methods, 8:785

**Figure 26: *In silico* signal peptide prediction for moxY using SignalP-4.0 (81).**

These results were confirmed by the software PrediSi (<http://www.predisi.de>), which calculated a possible cleavage site at position 23 as well as no secretion of the enzyme.

In order to receive indications for possible transmembrane regions within the protein, the webtool TMHMM Server v. 2.0 was used [see Figure 27]. The prediction resulted in no transmembrane helix (TMH) within the sequence of *moxY* and only a very low number of 3.7 amino acids in a TMH within the first 60 residues. TMHs at the N-terminus of the protein would indicate a possible signal peptide, but in the case of *moxY* the results suggested that *moxY* contained no membrane helices (82).



```
# moxY Length: 581
# moxY Number of predicted TMHs: 0
# moxY Exp number of AAs in TMHs: 3.72702
# moxY Exp number, first 60 AAs: 3.71268
```

**Figure 27: Prediction of transmembrane helices in proteins with TMHMM Server, v 2.0, indicating that no transmembrane helix was found in *moxY* (82).**

The webtool PSORT II (<http://psort.hgc.jp/form2.html>) for yeast was used for predictions concerning the subcellular localization of *moxY*, based on the k-nearest neighbor (k-NN) algorithm (Horton and Nakai, 1997). Thereby no targeting signal was found.

Results of the k-NN Prediction, k =9/23: >> prediction is cyt (k=23)

- 69.6 %: cytoplasmic
- 13.0 %: mitochondrial
- 8.7 %: nuclear
- 4.3 %: vesicles of secretory system
- 4.3 %: cytoskeletal

Analysis with PSORT II suggested that *moxY* is most likely found in the cytoplasm (69.6 %), but it has to be taken into account that these results are predictions which depend on the applied algorithm and are not fully reliable. However, in general the results of the *in silico* studies indicated that *moxY* has no signal or targeting sequence.

### 3.4.2. Fluorescence microscopy study

Microscopy studies were applied, using the fluorescent protein citrine for detection, in order to experimentally investigate the localization of *moxY* and to check whether the enzyme is membrane associated or not. Different strains were analyzed, including *moxY* constructs with citrine once on the N- and once on the C-terminal end and a *P.p. mut<sup>S</sup>* strain as negative control. Strains which express only citrine without any other fusion protein served as comparison samples. Citrine was chosen as it is a well expressed, simple and frequently used cytosolic protein. Its distribution in the cytoplasm can be easily detected by fluorescence microscopy and can be compared to the pictures of the *moxY* - citrine fusion constructs. In case that *moxY* is a cytosolic protein, an image with similar fluorescence signals would be expected. However, if *moxY* is membrane bound or localized in specific cell compartments, different pictures were anticipated, with fluorescence signals concentrated along membranes or at certain cell organelles.

Fluorescence microscopy showed that citrine, which is present in multiple copies in the pPIC vector, was distributed over the whole cell [see Figure 28]. The same results were obtained when the B1 vector was used for the expression of citrine [see Figure 29]. Comparison between the fluorescence pictures and the phase contrast images showed, that more or less every cell was expressing citrine. As expected, the cells of the *P.p. mut<sup>S</sup>* strain did not show any fluorescence (picture not shown).

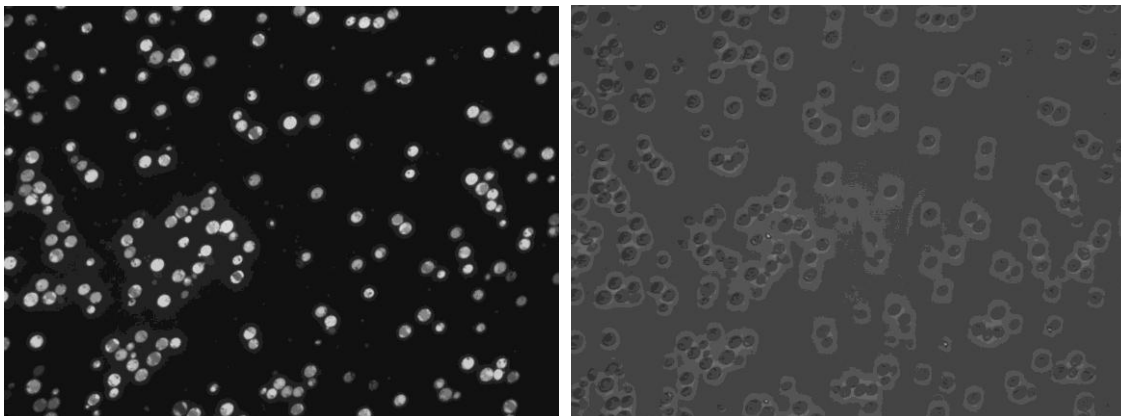
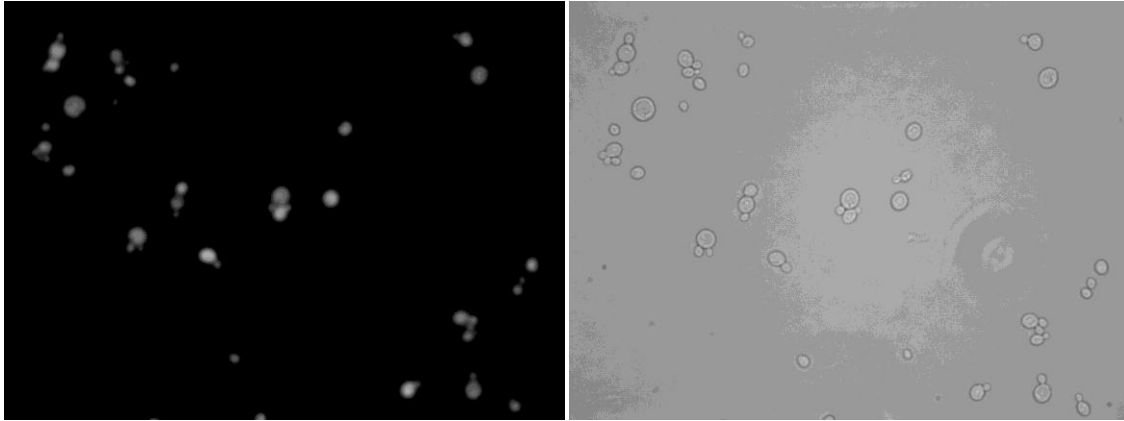


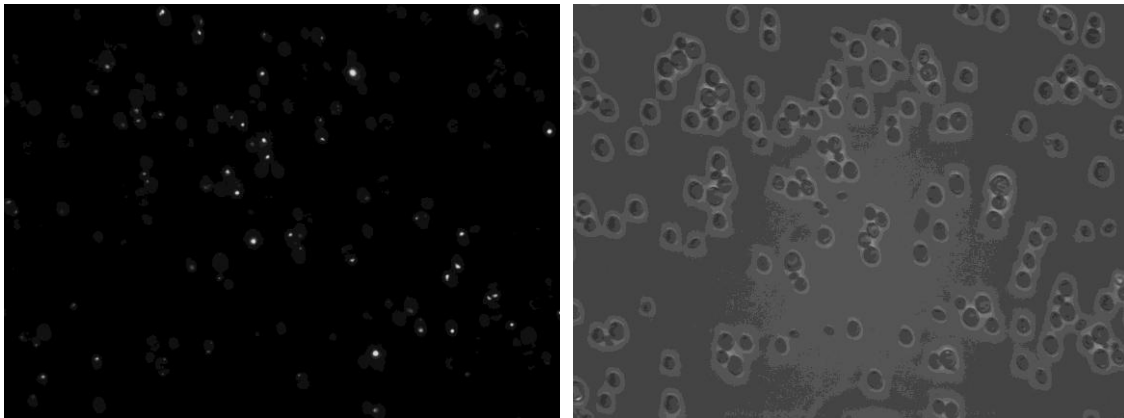
Figure 28: Fluorescence and phase contrast image of *P.p. pPic ZB\_citrine (3300) m.c.*



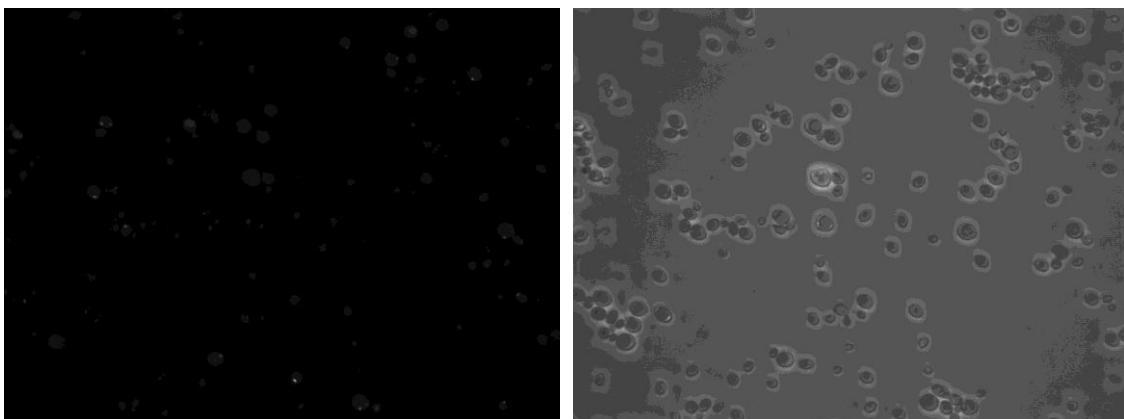


**Figure 29: Fluorescence and phase contrast image of *P.p. mutS B1 citrine*.**

The *moxY* long construct from *Aspergillus flavus* with citrine on the N-terminal end seemed to be accumulated at certain cell compartments [see Figure 30]. Very bright spots could be identified in the apart from that only faintly fluorescent cells. When citrine was fused to the C-terminal end, the whole cell was slightly shining and some spots of accumulated protein were detected again, but with a lower intensity [see Figure 31]. Interestingly, the fluorescence signals are mainly detected for connected cells and may correlate with the germination process.

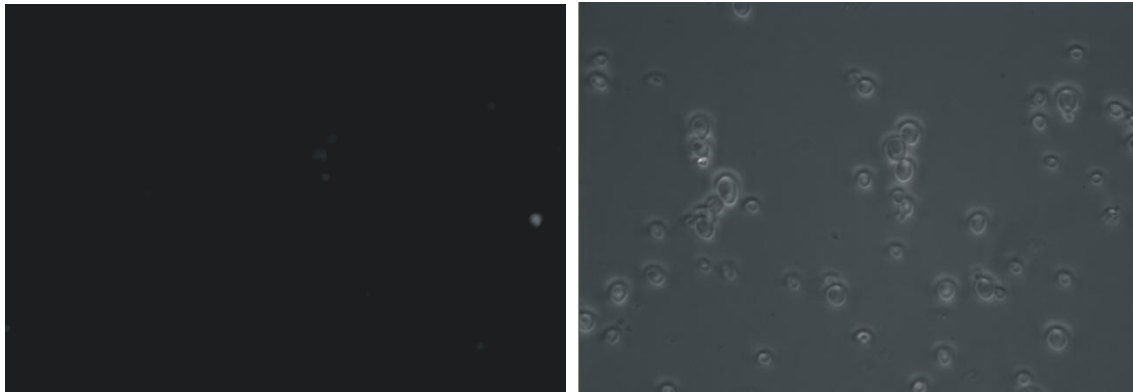


**Figure 30: Fluorescence and phase contrast image of *P.p. mutS B1\_his\_citrine\_moxYlong***

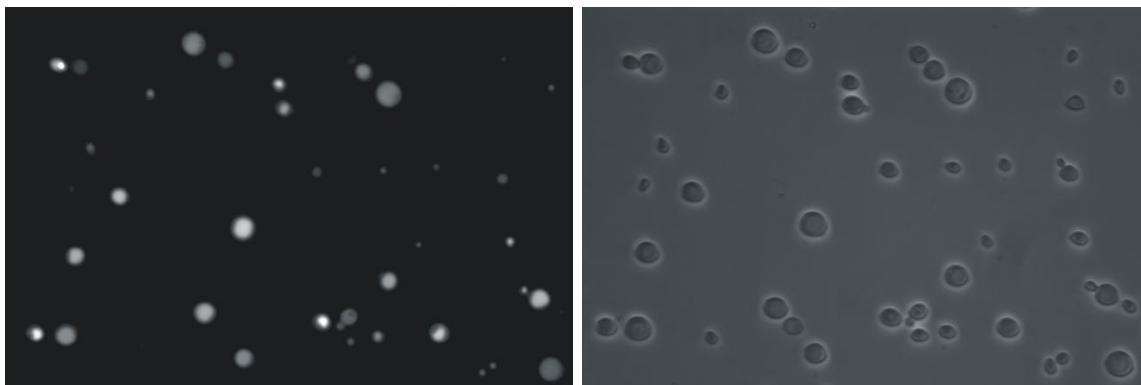


**Figure 31: Fluorescence and phase contrast image of *P.p. mutS B1\_moxYlong\_citrine\_his***

Additionally, two other moxY strains were analyzed. moxY short from *Asperillus flavus* showed only a very weak fluorescence of the whole cells [see Figure 32], whereas a much stronger signal could be seen at the moxY strain from Italy from *Aspergillus parasiticus* [see Figure 33]. Here mainly the whole cell was shining but sometimes there were again accumulations of the protein within the cells with higher intensities. Comparison between the fluorescence and phase contrast pictures revealed that the ratio of cells which express moxY citrine was usually very high; a low ratio was only found for the strain moxY short citrine.



**Figure 32: Fluorescence and phase contrast image of *P.p. mutS B1\_moxYshort\_citrine\_strep***



**Figure 33: Fluorescence and phase contrast image of *P.p. mutS moxY\_IT\_citrine\_his***

For moxY long a stronger fluorescence signal was measured compared to moxY short, indicating a favored expression, although a direct comparison was not possible since the copy numbers were not determined. An even more intense fluorescence was detected for moxY IT from *A. parasiticus*, which suggested that this moxY strain could be a promising option for future experiments. The fluorescence signal was distributed uniformly across the whole cell and indicated that the protein is located in the soluble, cytosolic fraction.

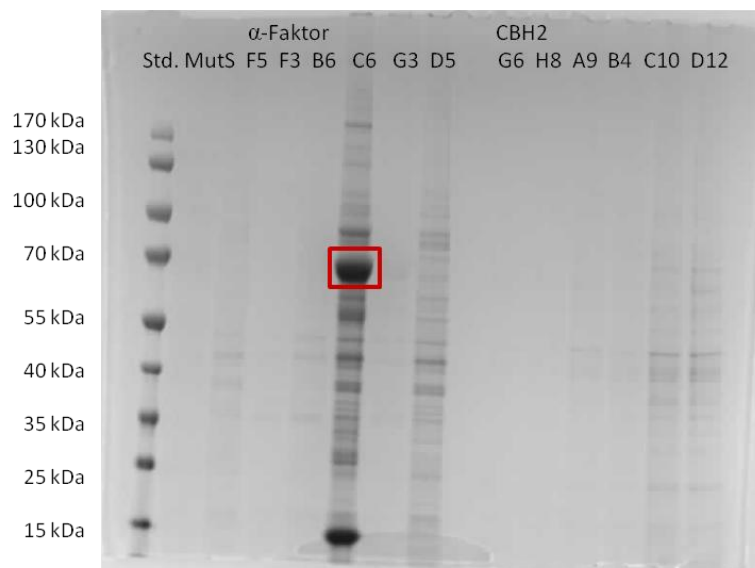
moxY long, which was used for all following experiments, showed accumulations of the protein, especially when citrine was fused to the N-terminal end of moxY. This showed that changing the position of the fusion protein effected the expression so that testing both

possibilities may be worthwhile. From the pictures it was hard to tell the exact localization of the accumulations, whether they were situated in certain cell organelles or protein aggregations. Perhaps analysis with a higher resolved microscope can clarify, if the protein is membrane bound or not since this will influence the planning of subsequent experiments and the choice of the purification strategy.

### 3.4.3. Secretion study

#### 3.4.3.1. Screening of the secretion of moxY

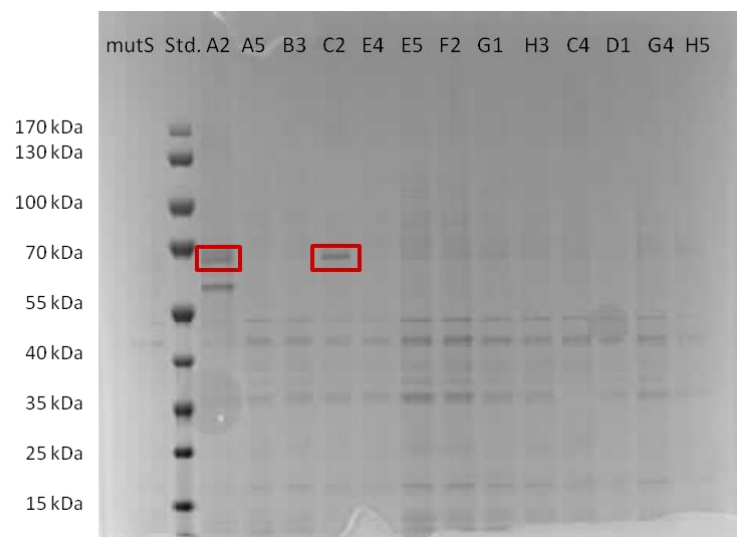
After cultivation of the moxY constructs with two different secretion signals, the protein concentrations in the supernatants were determined. Low protein concentrations were obtained, which indicated a low level of secretion and consequently TCA precipitation of the protein was necessary prior to SDS-PAGE analysis.



**Figure 34: SDS gel of the secretion of moxY with the  $\alpha$ -factor signal sequence and the CBH2 signal sequence. (Lane 1: Protein ladder Prestained PageRuler, lane 2: MutS strain as negative control, lane 3-8: the supernatant of clones with the  $\alpha$ -factor signal sequence for the secretion of moxY (66 kDa) and their position on the DWP, lane 9-14: the supernatant of clones with the CBH2 signal sequence for the secretion of moxY and their position on the DWP)**

The SDS gels revealed, that moxY, which was expected at a size of ~66 kDa, could not be efficiently secreted - neither with the  $\alpha$ -factor signal sequence nor with the CBH2 signal peptide [see Figure 34]. *P.p* mutS was applied as negative control but no bands at all could be detected, which may be due to difficulties at the TCA precipitation or application of the sample. Although the amount of proteins secreted by *Pichia* is usually very low, some slight bands were expected.

The first trials showed one clone of *moxY* with the  $\alpha$ -factor signal sequence that seemed to express big amounts of *moxY* (see lane C6 in Figure 34). However, MS analysis led to the result that this protein band represented albumin and not *moxY*. So far, the extracellular albumin expression of this *Pichia* strain in such high amounts could not be explained, whereby one possible option could be a contamination. An albumin homolog in *Pichia pastoris* could not be found in the NCBI database (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Some of the strains (D5, C10, D12) showed low secretion levels of various proteins but no secretion of *moxY* was detected. Consequently, more clones were investigated. The absorption values, which were measured for the determination of the protein concentration with the Biorad assay, were only slightly above the blank values and not further taken into account, so analysis was done only via SDS-PAGE.



**Figure 35: SDS gel of the secretion of *moxY* with  $\alpha$  - factor signal sequence. (Lane 1: MutS strain as negative control, lane 2: Protein ladder Prestained PageRuler, lane 3-15: the supernatant of clones with the  $\alpha$ -factor signal sequence for the secretion of *moxY* (66 kDa) and their position on the DWP)**

Two clones [A2 and C2 in Figure 35] were found when the  $\alpha$ -factor signal sequence was tested as secretion signal that seemed to express extracellular *moxY*, but the amounts were only small. What's more, these bands could also represent albumin, as it was the case for clone C6 in Figure 34. MS-analysis would be necessary to exclude this possibility and to confirm that these protein bands can indeed be attributed to *moxY*. In the case of strain A2 the proteins with molecular weights below 55 kDa that all other strains had in common could not be seen. Instead, a band at ~ 60 kDa was found, which may result from some kind of degradation of *moxY* or incorrect processing during secretion. For *P.p. mutS* only very faint protein bands were obtained, confirming the low amount of secretion of endogenous proteins which make *P. pastoris* a desirable host system for the expression of extracellular proteins.



**Figure 36: SDS gel of the secretion of moxY with CBH2. (Lane 1-13: the supernatant of clones with the CBH2 signal sequence for the secretion of moxY (66 kDa) and their position on the DWP, lane 14: MutS strain as negative control, lane 15: Protein ladder Prestained PageRuler)**

Also with the CBH2 signal peptide moxY could not be successfully secreted [see Figure 36]. For strain F4 a slight band that may correspond to moxY was detected, but in general the protein content of the samples was identical to that of *P.p. mutS*, indicating that no secretion of moxY was possible. However, although this may apply for moxY there are other examples, where proteins were efficiently secreted in *P. pastoris* with the CBH2 signal sequence. In recent experiments of A. Mellitzer *et al.* (ACIB Graz), new record levels of secretion could be reached in *P. pastoris* with TrCBH2 from *Trichoderma reesei*. Thereby, the optimization of the gene and the application of the synthetic promotor P(De2) led to protein concentrations of more than 15 g/L of secreted cellulase in fed-batch cultivations (77), (75).

Due to the results of the SDS gels, the secretion experiments were stopped at that point and the focus was laid on the MBP-tag purification strategy.

#### 3.4.3.2. Screening of the secretion of citrine

##### Secretion of citrine with the $\alpha$ -factor signal sequence

The secretion of citrine was easily detected by measuring the fluorescence of the protein directly in the cell suspension as well as in the supernatants after centrifugation. The obtained values were normalized to the OD<sub>600</sub> and depicted in diagrams [see Figure 37 to Figure 43]. Citrine single copy strains were added to the deep well plates as control samples, since these strains are known to successfully produce intracellular citrine. Their purpose was

having a reference for the expression yield and to make sure, that secretion of *moxY* is solely enabled by the added signal sequences.

It turned out that the extracellular expression of citrine was improved when the  $\alpha$ -factor signal peptide was used, since higher values of fluorescence were obtained with the  $\alpha$ -factor secretion signal compared to the signal of CBH2 [see Figure 38 and Figure 42]. In both cases, the fluorescence decreased significantly when the supernatant was analyzed instead of the medium with the whole cells [see Figure 37 and Figure 41]. This indicated that although citrine could be secreted successfully, a considerable amount still remained intracellular. Possible reasons therefore may be an overloading of the secretion machinery of the cells or limitations during signal sequence processing.

The intracellular expression of citrine s.c. strains was in the same range as the best expressing clones with a secretion signal. High deviations within the measured fluorescence/OD values were found for the citrine s.c. control strains. This may have been caused by slightly different cultivation conditions; uneven evaporation rates could for example lead to smaller volumes in the wells at the edge of the deep well plates and consequently higher methanol concentrations upon induction. When the supernatant was analyzed, as expected no fluorescence was observed neither in the negative controls *P.p. mutS* nor the citrine s.c. strains.

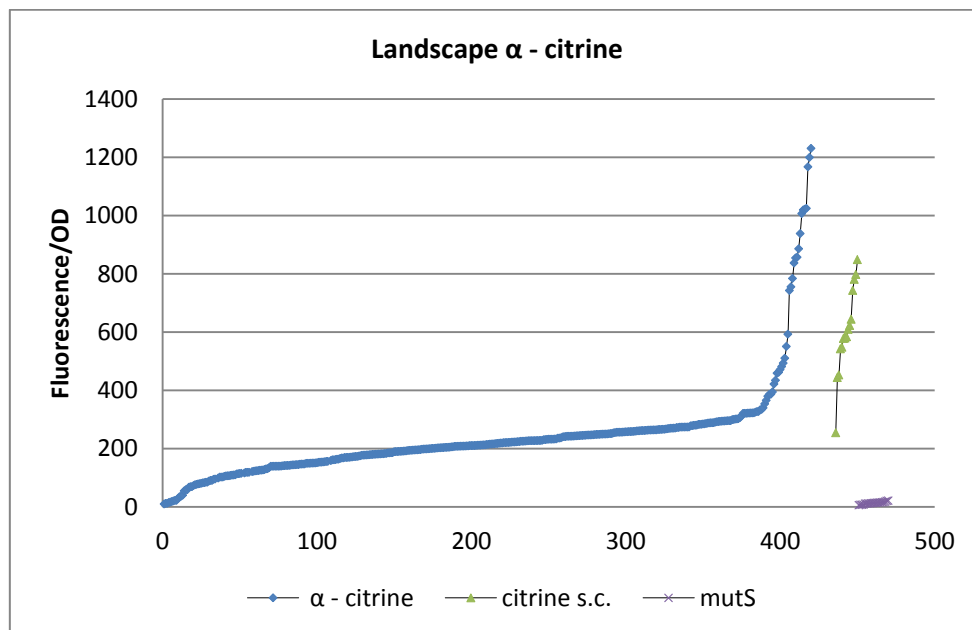
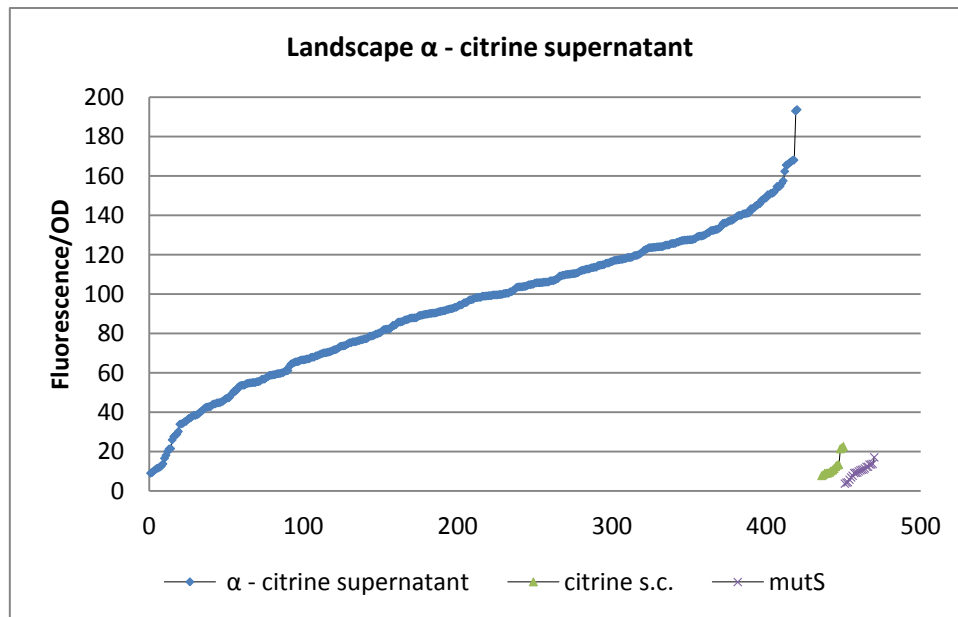


Figure 37: Fluorescence measurement of *P. pastoris*  $\alpha$ - factor signal sequence citrine clones in the culture including whole cells.



**Figure 38: Fluorescence measurement of the secretion of citrine with the  $\alpha$ -factor signal sequence in the supernatant.**

#### Secretion of the MBP citrine fusion protein with the $\alpha$ -factor signal sequence

In recent experiments of Z. Li *et al.*, the secretion of some cargo proteins could be enhanced in *P. pastoris* by attaching an MBP-tag, even though cleavage between MBP and the fusion protein occurred during secretion (83). In order to test whether an improvement in the extracellular expression can be achieved also in the present experiments, the construct for the secretion of MBP citrine with the  $\alpha$ -factor signal sequence was examined. It turned out that the landscapes of the fluorescence of  $\alpha$ -factor signal sequence citrine with and without MBP tag in the cell suspensions are similar, although the construct with the MBP tag showed a decreased secretion in the supernatant [see Figure 39 and Figure 40]. This showed that the expression of the citrine fusion protein was not affected by the MBP-tag, whereas secretion was possible but less efficient compared to citrine without MBP-tag.

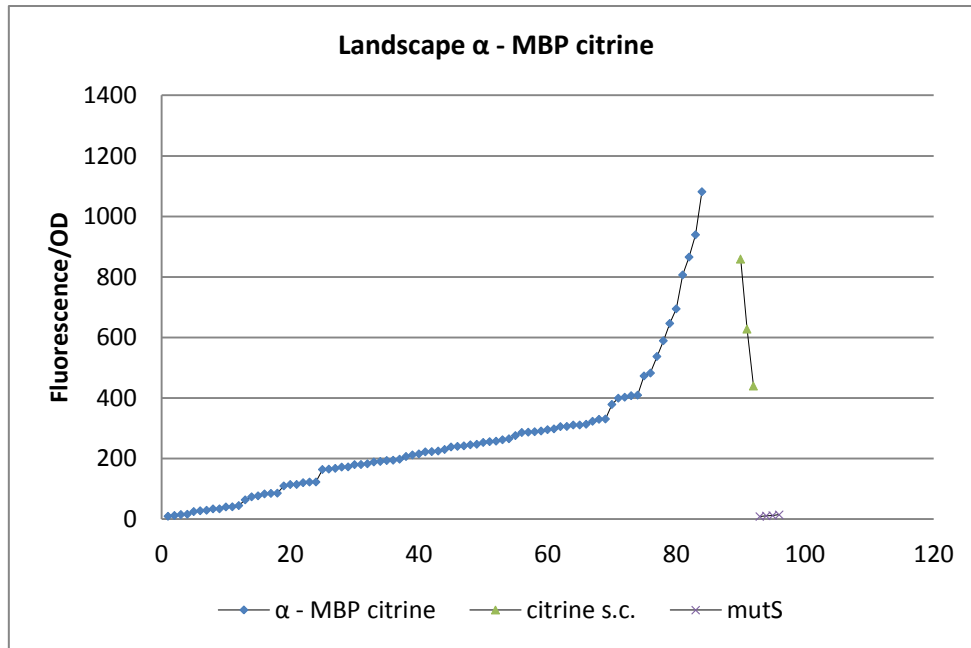


Figure 39: Fluorescence measurement of the *P. pastoris*  $\alpha$ - factor signal sequence His MBP citrine construct with MBP tag in the culture including whole cells.

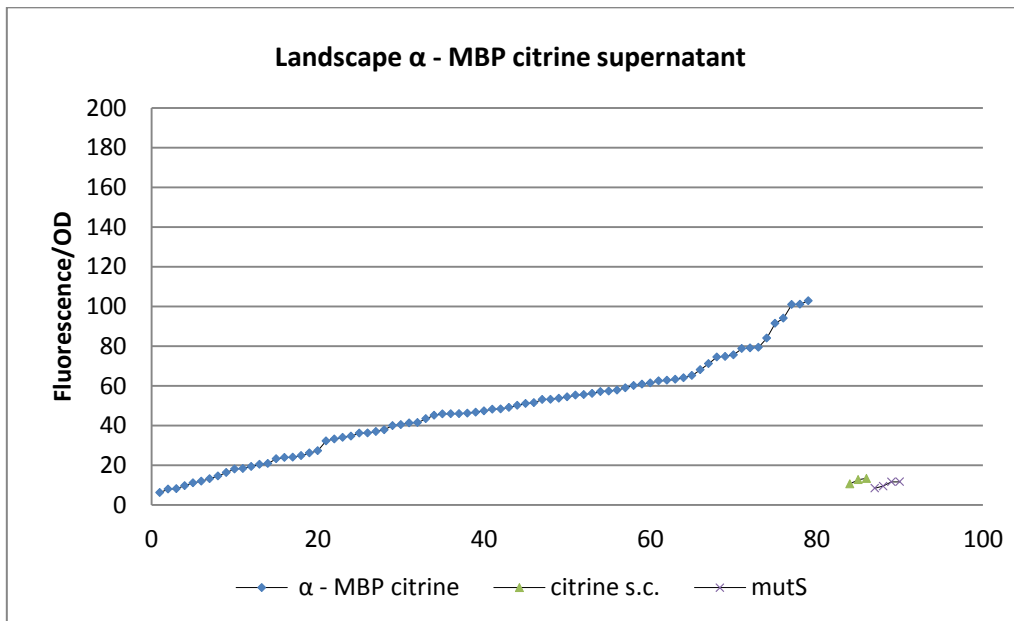


Figure 40: Fluorescence measurement of the secretion of the  $\alpha$ - factor signal sequence His MBP citrine construct with MBP-tag in the supernatant.

#### Secretion of citrine with the CBH2 signal peptide

The fluorescence and thus the expression levels of the CBH2 citrine strains in the cell suspensions were significantly lower than those of the strains with the  $\alpha$ -factor signal peptide



[see Figure 41]. The same applied for the secretion of the protein [see Figure 42], which seemed to be less efficient. In contrast to *moxY*, citrine could be secreted to a certain degree confirming the proper functioning of the CBH2 signal peptide. Nevertheless, the high secretion levels of the *TrCBH2* cellulase with its native signal sequence (77) could not be reached.

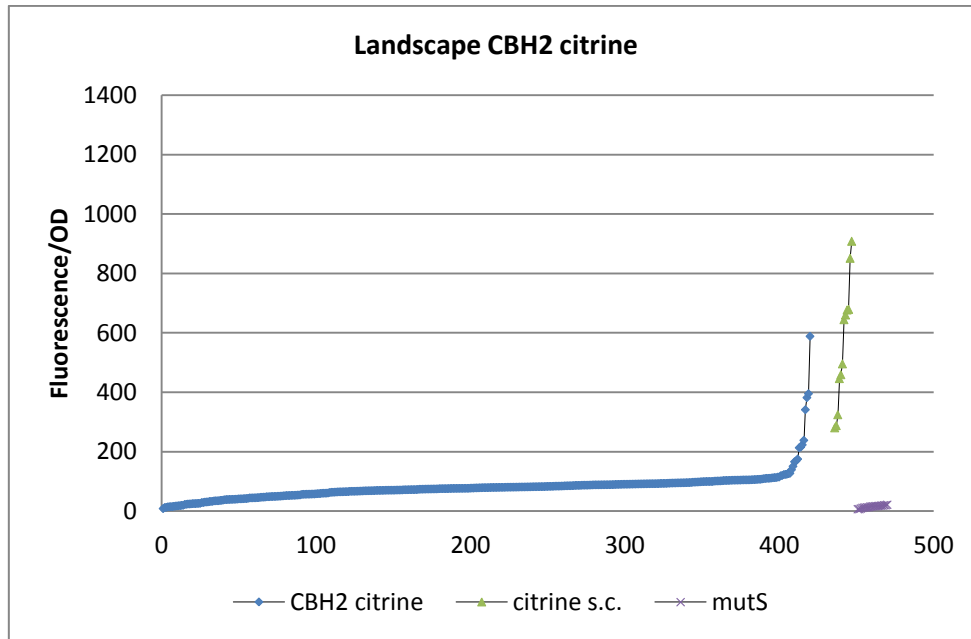


Figure 41: Fluorescence measurement of *P. pastoris* citrine constructs with the CBH2 signal sequence in the culture including whole cells.

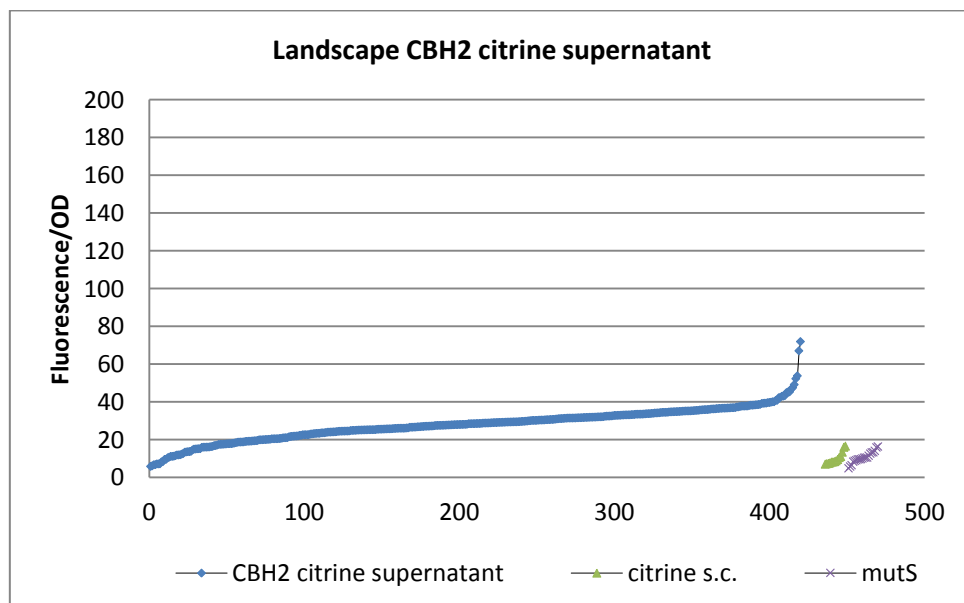
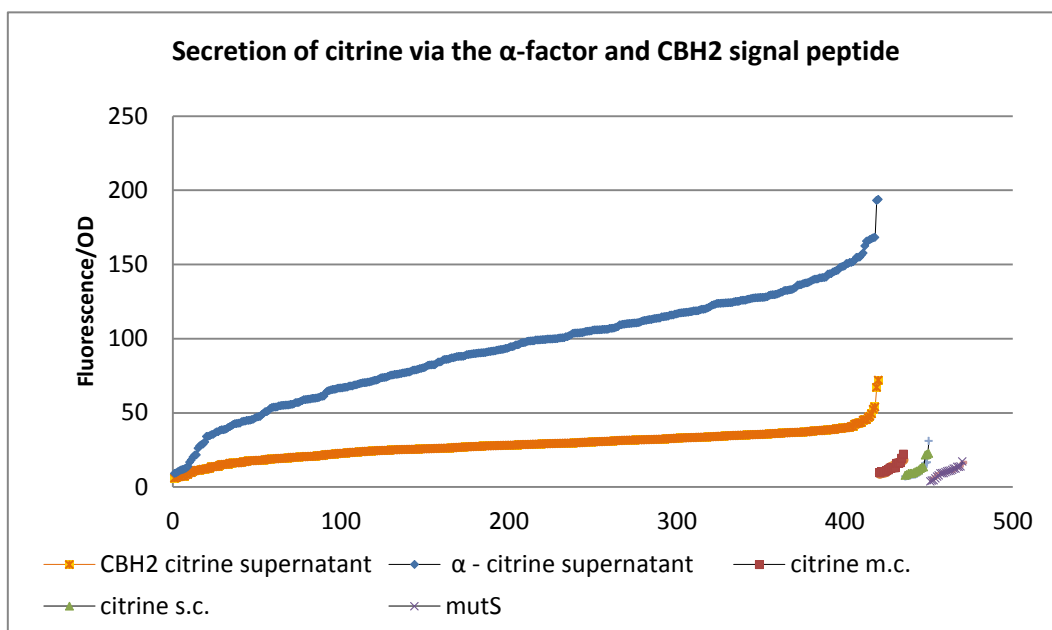


Figure 42: Fluorescence measurement of the secretion of citrine with the CBH2 signal sequence in the supernatant.

### 3.5. Conclusion

Fluorescence microscopy studies of moxY long, N-terminally fused to citrine, revealed that although the whole cells were slightly shining, there were protein accumulations or aggregations within the cells, which may cause insolubility problems during further working steps. However, in former experiments the successful expression of moxY with citrine could be proven via SDS gels of the cytosolic fraction, which indicated that the protein is soluble. Additionally, no targeting or signal sequences were detected in the course of *in silico* studies.

Alternatively, the protein expression of other strains like for example moxY IT from *Aspergillus parasiticus* could be tested, since here a stronger fluorescence signal, distributed over the whole cells, was detected. Changing the position of moxY may have consequences as well, given that less protein accumulations were found when the C-terminally fused moxY citrine constructs were analyzed. The exact localization of moxY still remains unclear, but it should be possible to get more informative pictures from a higher resolving microscope.



**Figure 43: Comparison of the secretion of citrine with the  $\alpha$ -factor signal sequence (blue) and the CBH2 signal sequence (orange). Citrine m.c. (red) and citrine s.c. strains (green) were used as positive controls and the empty *P.p. mut<sup>S</sup>* strain (violet) was used as a negative control.**

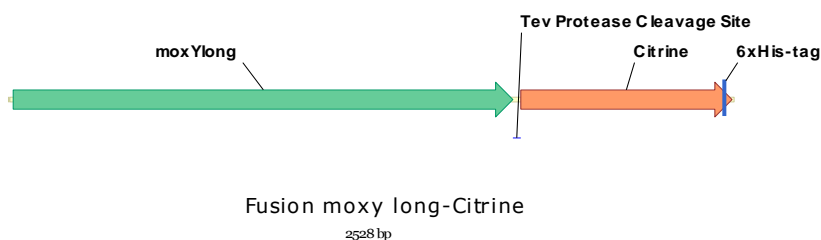
The secretion experiments demonstrated that the  $\alpha$ -factor signal sequence is better suited for the extracellular expression of citrine compared to the CBH2 signal peptide [see Figure 43]. moxY, on the other hand, could not be efficiently secreted neither with the  $\alpha$  - factor signal sequence nor with the CBH2 signal peptide. As a result, efforts were now focused on the intracellular expression of moxY together with a MBP-tag for purification of the protein via affinity chromatography.

## 4. Chapter II: Expression and purification of moxY

### 4.1. Background

The Baeyer - Villiger monooxygenase moxY is involved in the biosynthesis of aflatoxins by some *Aspergillus* species but, until now the mechanism and structure of the enzyme are unknown.

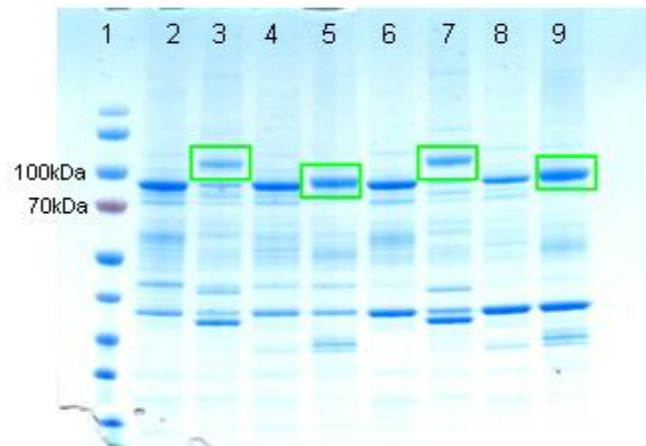
So far, there have been problems concerning the expression and purification of this monooxygenase. For first trials the host *E. coli* was used, which is a well established bacterial expression system and comparatively cheap and easy to handle. However, expression in *E. coli* was not possible and hindered by problems like the formation of inclusion bodies (unpublished results of the research group of Mattevi, Pavia). Given that moxY is a eukaryotic enzyme, its expression may be improved by using a fungal host system. Consequently, an expression construct was made for the yeast *Pichia pastoris*, including moxY long, followed by a TEV protease cleavage site, the reporter tag citrine and a His-tag for purification [see Figure 44].



**Figure 44: The first expression construct including moxY long with a TEV protease cleavage site, citrine as reporter tag and a His-tag.**

moxY could be successfully expressed in *P. pastoris* and SDS-PAGE analysis after His-tag purification of the cytosolic fraction indicated that moxY seemed to be available in the soluble fraction [see Figure 45, unpublished results from Martina Geier].

The identity of the protein could be proven by investigation of the amino acid sequence with MS-analysis. However, the expression of the His-tag was not confirmed and could also not be detected via dot blot analysis. It seemed that the His-tag was not expressed or not accessible since the fusion protein did not bind to the column for affinity chromatography but was found in the flow through. Earlier experiments showed, that using the StrepII-tag resulted in no binding either (unpublished results).



**Figure 45: SDS-PAGE gel of the His- tagged moxY variants after protein purification. Cytosolic fractions of recombinant *Pichia* strains were used. (Lane 1: PageRuler Prestained Protein Ladder, lane 2+6: eluate 1+2 of the negative control, the empty P.p. mut<sup>s</sup>, lane 3+7: eluate 1+2 of moxY<sub>long</sub> from *A. flavus*, lane 4+8: eluate 1+2 of moxY<sub>short</sub> from *A. flavus*, lane 5+9: eluate 1+2 of moxY<sub>IT</sub> from *A. parasiticus*) The green marked bands represent the moxY citrine fusions which were confirmed by MS analysis.**

There were various options how the problems with the purification of moxY could be handled, like changing or modifying the tag or trying to secrete the protein. For the present work, the MBP-tag (maltose binding protein), which seemed to be a good tag for protein folding, was chosen. Other possibilities would have been a SUMO-tag (small ubiquitin like modifier) or a choline binding domain, which allowed an easy and cheap purification of well secreted fusion proteins in other experiments (84).

There are several reasons, why MBP is commonly used in *E. coli*. It can improve the expression and folding properties as well as the solubility and stability of the enzyme (85). Furthermore, in combination with a signal sequence the secretion of the protein can be enhanced. Apart from that, the MBP-tag allows an easy purification via an amylose column.

Previous studies by Zhiguo Li *et al.* (83) showed that fusing MBP to the protein of interest can also enhance expression and secretion in *Pichia pastoris*. However, the main part of the secreted fusion protein seemed to be cleaved between MBP and the cargo protein. Dot blot analysis showed that intracellular this cleavage occurred only to a minor extent. Although different linkers as well as constructs with no linker at all were tested, it always resulted in the split off of MBP. Since no particular protease cleavage site could be detected, the cleavage seemed to be triggered by a certain 3D structure near the C-terminal end of MBP (83). In the case of moxY, maintaining the MBP-tag was crucial for its purification via affinity chromatography. Since moxY was expressed intracellularly, cleavage of the MBP-tag was expected only to a minor degree.

The first step of the present experiments was to optimize the MBP gene for *Pichia pastoris*. Different constructs were created where MBP was fused to the gene encoding for moxY either N- or C-terminally. The original protein sequence of MBP terminated with the amino acids DAQT [see appendix, chapter 10], whereas the construct which was optimized for *Pichia* contained a 23 amino acid long linker with a site for TEV protease cleavage which can be used to separate MBP and moxY after purification (86). Nevertheless, in the first instance the aim was to check whether expression and purification were generally possible and therefore a construct with no linker at all was designed, too. The next steps, after preparing the constructs, cloning them into *E. coli* and subsequent transformation into *P. pastoris* cells, consisted in the up-scaling of the cultivation of the cells and the expression and purification of moxY.

## 4.2. Materials and methods

Here the instruments and devices for the purification, fermentation and crystallization of the proteins are described. For a detailed list of all other materials, chemicals and methods used for the cloning, construct preparation and small scale cultivation of the strains see chapter 3.2.

### 4.2.1. Instruments

All chemicals were purchased from Sigma Aldrich, Merck, BD or Fluka and used without any further purification.

For the fermentation a fermenter from New Brunswick scientific CO., Inc (Enfield, CT, USA) was used.

All FPLC systems and columns were purchased from GE Healthcare (Uppsala, Sweden).

The crystallization plates were performed by an Oryx8 Protein Crystallization Robot from Douglas Instrument (Douglas House, UK).

### 4.2.2. Media and recipes

**Table 15: Media, buffers and solutions for the fermentation, cell breakage and purification moxY.**

Fermentation:		
YPD agar	yeast extract	1 % w/v
	peptone	2 % w/v
	dextrose	2 % w/v
	agar	1,5 % w/v
MGY media	YNB	1,34 % w/v
	glycerol	1 % w/v

	biotin	4 x 10 <sup>-5</sup> % w/v
basal salts medium, V= 1L	phosphoric acid 85 %	26,7 mL
	calcium sulfate	0,93 g
	potassium sulfate	18,2 g
	magnesium sulfate heptahydrate	14,9 g
	potassium hydroxide	4,13 g
	glycerol	32 mL
	tap water	filled up to 1 L
PTM <sub>1</sub> trace salts solutions, V = 100 mL	cupric sulfate pentahydrate	0,6 g
	sodium iodide	8 mg
	manganese sulfate hydrate	0,3 g
	sodium molybdate dihydrate	20 mg
	boric acid	2 mg
	cobalt chloride	50 mg
	zinc chloride	2 g
	ferrous sulfate heptahydrate	6,5 g
	biotin	200 mg
	sulfuric acid	500 µL
<b>Cell disruption:</b>		
breaking buffer, pH 7,9	KPi	50 mM
	EDTA	1 mM
	DTT	2 mM
	glycerol	5 % w/v
	PMSF	2 mM
	Roche complete protease inhibitor	50 x
<b>Dot blot:</b>		
10 x TBS buffer, pH 7,5	Tris	0,25 M
	NaCl	1,5 M
1 x TBST	10 x TBS buffer	10 %
	Tween <sup>20</sup>	0,003 %
TBST milk	dried whey	5 g (5 %)
	1 x TBST	filled up to 100 mL
<b>MBP - tag purification pretests:</b>		
binding buffer, pH 7,3	NaP <sub>i</sub>	20 mM
washing buffer, pH 7,3	binding buffer	
	NaCl	0,2 M
elution buffer, pH 7,3	washing buffer	
	maltose	10 mM
<b>His -tag purification pretests:</b>		
binding buffer, pH7,4	NaP <sub>i</sub> buffer	20 mM
	NaCl	500 mM
	imidazole	20 mM
washing buffer, pH7,4	NaP <sub>i</sub> buffer	20 mM
	NaCl	500 mM
	imidazole	50 mM
elution buffer, pH7,4	NaP <sub>i</sub> buffer	20 mM
	NaCl	500 mM
	imidazole	500 mM
<b>Protein purification:</b>		
washing buffer A, pH 7,4	NaP <sub>i</sub>	50mM
	NaCl	200 mM
	EDTA	1 mM
elution buffer A, pH 7,4	washing buffer A	
	maltose	10 mM
dialysis buffer, pH 8	Tris/Cl	25 mM
	NaCl	20 mM
	glycerol	5 % w/v
elution buffer B, pH 8	dialysis buffer	
	NaCl	1 M

The buffers and media were prepared using deionized water and sterilized either by filtration or by autoclaving.

### 4.3. Experimental

#### 4.3.1. Optimization of the MBP gene for *P. pastoris*

For the gene optimization the sequence of MBP including the linker and the TEV cleavage site was used. MBP was optimized for *Pichia pastoris* using the Gene Designer from DNA 2.0 Inc. (Menlo Park, CA, USA) and the *Pichia pastoris* high methanol codon usage table. The threshold of the software program was set to 10 %. Various parameter were considered during the optimization like a GC-content of 30 - 70 % without high peaks [see Figure 47], only few AT-repeats, not many rarely used codons after each other and not too energy rich RNA secondary structures (higher than -15 kcal/mol). Commonly used restriction sites for standard molecular cloning as well as certain motifs for polyadenylation or pre-termination were avoided within the gene sequence [see Table 28 and Table 29 in the appendix, chapter 10]. The gene named MBDPichiaopt was ordered at DNA 2.0 in the vector pUC57.

#### 4.3.2. Preparation of MBP moxY constructs:

The aim was to prepare the following constructs:

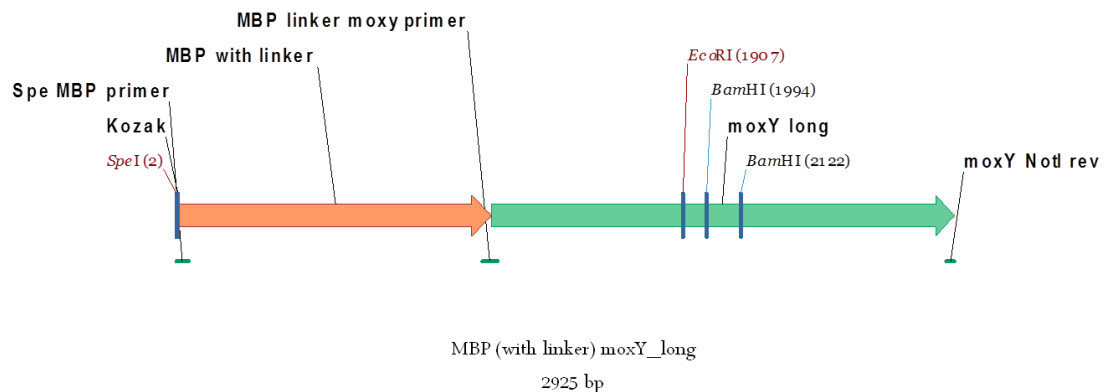
- pPp\_B1 MBP (with linker) moxY
- pPp\_B1 MBP (no linker) moxY
- pPp\_B1 moxY MBP (no linker)
- pPp\_B1 His-tag MBP (with linker) moxY
- pPpT4\_alpha\_S  $\alpha$  - factor signal sequence His-tag MBP (with linker) moxY

In order to confirm the successful expression of MBP fusion proteins the same constructs were prepared with citrine instead of moxY. Citrine, which is known as an easily detectable, relatively simple and usually well expressed fluorescent protein, served as a control sample to test whether the expression of MBP fusion proteins works in general. The constructs with the  $\alpha$ -factor signal sequence, the His-tag, MBP and moxY or citrine were prepared for the secretion study in chapter 3.

Design of the primers: For the N-terminal fusion of MBP to moxY [see Figure 46], the primers P11155 to P11162 [see Table 5] were designed. The restriction sites *SpeI* and *NotI* were added to allow cloning into the pPp\_B1 plasmid. Therefore, primers with a *SpeI* restriction site, a Kozak sequence and the start codon ATG followed by the first bases of MBP or the His-tag plus MBP were needed (P11162 and P11161). Additionally, forward and reverse

overlapping primers were prepared that connect the N-terminal end of *moxY* and the C-terminal end of MBP, once with and once without the linker sequence for TEV-cleavage.

For the C-terminal fusion of MBP to *moxY*, a forward primer with a *SpeI* cleavage site, a Kozak sequence and ATG followed by the first bases of *moxY* and a reverse primer with a *NotI* restriction site that bound to the C-terminal end of MBP without the linker were made. In addition, overlapping primers for the connection of both fusion partners were designed (P11153 to P11156).



**Figure 46: The construct of *moxY* long and the N-terminal fused MBP with linker. The primers designed for the preparation and cloning of the construct are shown as small green lines, important restriction sites are marked.**

Additionally, the primers P11178 to P11181 [see Table 5] were designed to allow the sequencing of the MBP constructs. For the preparation of the constructs with citrine the same setup was applied for the design of the primers (P11206 to P11211).

PCR reactions:

The optimized MBP gene in the pUC57 vector was diluted to a concentration of 10 ng/μL and used for transformation to *E. coli* Top10 cells as well as for the PCR reactions to produce the various *moxY* and citrine constructs. The components, primers, templates and temperature cycling programs that were used for the PCR reactions are listed in Table 16 to Table 20:

**Table 16: General composition of the PCR reactions for preparation of the MBP constructs.**

PCR for MBP constructs	
DNA template	1 μL
fw primer	2 μL
rev primer	2 μL
Phusion polymerase	0.5 μL
dNTP 10 mM	1 μL
HF buffer 5 x	10 μL
d H <sub>2</sub> O	33.5 μL
total volume	50 μL



**Table 17: The templates and primers used for the PCR reactions to prepare the constructs MBP (with linker) moxY, MBP (no linker) moxY and moxY MBP (no linker).**

reaction nr.	MBP moxY					
	MBP (with linker) moxY		MBP (no linker) moxY		moxY MBP (no linker)	
	1	2	3	4	5	6
template	pUC57_MBP_Pichiaopt [10 ng/μL]	B1_moxYlong_citrine [10 ng/μL]	pUC57_MBP_Pichiaopt [10 ng/μL]	B1_moxYlong_citrine [10 ng/μL]	B1_moxYlong_citrine [10 ng/μL]	pUC57_MBP_Pichiaopt [10 ng/μL]
fw primer	P11162: <i>SpeI</i> _MBP_fw	P11160: MBP_linker_moxY_fw	P11162: <i>SpeI</i> _MBP_fw	P11158: MBP_no_linker_moxY_fw	P11154: <i>SpeI</i> _moxY_fw	P11156: moxY_MBP_connection_fw
rev primer	P11159: MBP_linker_moxY_rev	P11149: moxY_NotI_rev	P11157: MBP_no_linker_moxY_rev	P11149: moxY_NotI_rev	P11155: moxY_MBP_connection_rev	P11153: MBP_no_linker_NotI_rev

**Table 18: The templates and primers used for the PCR reactions to prepare the constructs His MBP (with linker) moxY, α His MBP (with linker) moxY, MBP (with linker) citrine and MBP (no linker) citrine.**

reaction nr.	MBP moxY		MBP citrine			
	His MBP (with linker) moxY	α His MBP (with linker) moxY	MBP (with linker) citrine		MBP (no linker) citrine	
	7	8	9	10	11	12
template	pUC57_MBP_Pichiaopt [10 ng/μL]	pUC57_MBP_Pichiaopt [10 ng/μL]	pUC57_MBP_Pichiaopt [10 ng/μL]	B1_moxYlong_citrine [10 ng/μL]	pUC57_MBP_Pichiaopt [10 ng/μL]	B1_moxYlong_citrine [10 ng/μL]
fw primer	P11161: <i>SpeI</i> _HIS_MBP_fw	P12061: <i>XhoI</i> _alpha_His_MBP_fw	P11162: <i>SpeI</i> _MBP_fw	P11206: MBP_linker_Citrine_fw	P11162: <i>SpeI</i> _MBP_fw	P11208: MBP_no_linker_Citrine_fw
rev primer	P11159: MBP_linker_moxY_rev	P11159: MBP_linker_moxY_rev	P11207: MBP_linker_Citrine_rev	P11151: Citrine_NotI_rev	P11209: MBP_no_linker_Citrine_rev	P11151: Citrine_NotI_rev

**Table 19: The templates and primers used for the PCR reactions to prepare the constructs citrine MBP (no linker) and α His MBP (with linker) citrine.**

reaction nr.	MBP citrine		
	citrine MBP (no linker)		α His MBP (with linker) citrine
	13	14	15
DNA template	B1_moxYlong_citrine [10 ng/μL]	pUC57_MBP_Pichiaopt [10 ng/μL]	pUC57_MBP_Pichiaopt [10 ng/μL]
fw primer	P08511: <i>EcoRI</i> _fw	P11210: Citrin_MBP_connection_fw	P12061: <i>XhoI</i> _alpha_His_MBP_fw
rev primer	P11211: Citrin_MBP_connection_rev	P11153: MBP_no_linker_NotI_rev	P11207: MBP_linker_Citrine_rev

**Table 20: Temperature program for the PCR reactions to prepare the constructs with MBP.**

	temperature	time	} 30 cycles
initial denaturation	98 °C	30 sec	
denaturation	98 °C	10 sec	
annealing	58 °C	20 sec	
extension	72 °C	30 sec (moxY 1.7 kbp)	
		15 sec (citrine 700 bp)	
		20 sec (MBP 1.2 kbp)	
final extension	72 °C	7 min	
	4 °C	∞	

The PCR samples were loaded on a preparative agarose gel, the respective bands were cut out and the DNA was purified as described in 3.2.5.5 and used for oePCRs. The following oePCR reactions were prepared [see Table 21 to Table 23]:

**Table 21: General composition of the oePCRs to prepare the MBP moxY constructs.**

oePCR for MBP moxY constructs	
MBP fragment (1ng/μL)	4 μL
moxY fragment (1ng/μL)	6 μL
Phusion polymerase	0.5 μL
dNTP 10 mM	1 μL
HF buffer 5 x	10 μL
d H <sub>2</sub> O	28.5 μL
total volume	50 μL

**Table 22: General composition of the oePCRs to prepare the MBP citrine constructs.**

oePCR for MBP citrine constructs	
MBP fragment (1ng/μL)	6 μL
citrine fragment (1ng/μL)	3.5 μL
Phusion polymerase	0.5 μL
dNTP 10 mM	1 μL
HF buffer 5 x	10 μL
d H <sub>2</sub> O	28.5 μL
total volume	50 μL

**Table 23: The DNA-fragments which were connected via oePCR to prepare the MBP constructs. The numbers of the components refer to the numbers of the reactions in Table 17, Table 18 and Table 19.**

MBP (with linker) moxY	MBP (no linker) moxY	moxY MBP (no linker)	His MBP (with linker) moxY	α His MBP (with linker) moxY	MBP (with linker) citrine	MBP (no linker) citrine	citrine MBP (no linker)	α His MBP (with linker) citrine
1 MBP	3 MBP	5 moxY	7 His MBP	8 α His MBP	9 MBP	11 MBP	13 citrine	15 α His MBP
2 moxY	4 moxY	6 MBP	2 moxY	2 moxY	10 citrine	12 citrine	14 MBP	10 citrine

The numbers of the DNA fragments refer to the above mentioned PCR reaction numbers in Table 17, Table 18 and Table 19. The same temperature program as before was used with the only deviation that the extension time was extended to 45 sec for MBP moxY constructs (2.9 kb) and 30 sec for MBP citrine constructs (1.9 kb). After 15 cycles 50 μL of the following mastermix [see Table 24 and Table 25] were added for the amplification of the whole construct and the PCR was started for further 20 cycles.

**Table 24: General composition of the mastermix, added to the oePCRs for the amplification of the whole MBP constructs.**

Mastermix for oePCR for the MBP constructs	
Primer fw 5 μM	4 μL
Primer rev 5 μM	4 μL
Phusion polymerase	0.5 μL
dNTP 10 mM	1 μL
HF buffer 5 x	10 μL
d H <sub>2</sub> O	30.5 μL
total volume	50 μL

**Table 25: List of the forward and reverse primers added to the mastermix described in Table 24 for the amplification of the whole MBP constructs via oePCR.**

	MBP (with linker) moxY	MBP (no linker) moxY	moxY MBP (no linker)	His MBP (with linker) moxY	$\alpha$ His MBP (with linker) moxY
Primer fw	P11162: <i>SpeI</i> _MBP_fw	P11162: <i>SpeI</i> _MBP_fw	P11154: <i>Spe</i> _moxY_fw	P11161: <i>SpeI</i> _HIS_MBP_fw	P12061: <i>XhoI</i> _alpha_His_MBP_fw
Primer rev	P11149: moxY_ <i>NotI</i> _rev	P11149: moxY_ <i>NotI</i> _rev	P11153: MBP_no_linker_ <i>NotI</i> _rev	P11149: moxY_ <i>NotI</i> _rev	P11149: moxY_ <i>NotI</i> _rev
	MBP (with linker) citrine	MBP (no linker) citrine	citrine MBP (no linker)	$\alpha$ His MBP (with linker) citrine	
Primer fw	P11162: <i>SpeI</i> _MBP_fw	P11162: <i>SpeI</i> _MBP_fw	P08511: <i>EcoRI</i> _fw	P12061: <i>XhoI</i> _alpha_His_MBP_fw	
Primer rev	P11151: Citrine_ <i>NotI</i> _rev	P11151: Citrine_ <i>NotI</i> _rev	P11153: MBP_no_linker_ <i>NotI</i> _rev	P11151: Citrine_ <i>NotI</i> _rev	

After the oePCR and a preparative agarose gel, the DNA bands were cut out of and purified. The constructs as well as the B1 vector were cut with *SpeI* and *NotI* restriction enzymes and after dephosphorylation of the vector with SAP, the ligation with the inserts was performed as described for the preparation of the constructs for the secretion study (see chapter 3.3). For the citrine MBP construct the insert as well as the B1 vector were cut with *EcoRI* and *NotI*. The constructs with the  $\alpha$ -factor signal sequence and His-tag at the N-terminal end of MBP moxY and MBP citrine were cut with *XhoI* and *NotI* and ligated with the cut and dephosphorylated T4  $\alpha$  S vector (see 3.3.2.1).

The constructs were transformed into *E. coli* Top10 cells (see 3.2.5.10) and positive clones were identified via control cuts with the respective restriction enzymes followed by agarose gel electrophoresis and sequencing by LGC Genomics. They were struck out on LB zeocin plates and after isolation of the plasmid DNA, the DNA was linearized with *BglII* and transformed into *P. pastoris* cells as described in 3.2.5.12. The constructs  $\alpha$  His MBP (with linker) moxY and  $\alpha$  His MBP (with linker) citrine were linearized with *SmlI* (*Swal*) before transformation. For further experiments, like purification tests, the obtained *P. pastoris* clones were finally cultivated in deep well plates or 2 L shake flasks, as described in 3.2.5.13 and 3.2.5.14.

#### 4.3.3. Screening methods and purification pretests

After cell growth and induction of the protein expression the deep well plates were centrifuged at 4000 rpm for 5 min at room temperature and the supernatant was discarded. For the Y-Per protein extraction the protocol from ThermoScientific for yeast cells was used. Therefore, 200  $\mu$ L of Y-PER® Plus reagent were added to the cell pellets. The cells were resuspended by carefully vortexing the plates, before they were incubated at 24 °C and 500 rpm for 1 h. In order to improve the protein yield for *Pichia* cultures small modifications of the extraction procedure were tested, like increasing the temperature to 45 °C, as suggested by the manufacturer's protocol. For some trials, 20  $\mu$ L 1 M DTT were added together with the Y-

Per reagent and the resuspended cells were incubated at 45 °C and 500 rpm for 30 min. The plates were centrifuged again at 4000 rpm for 5 min and the supernatants were transferred in new, transparent micro titer plates.

The supernatants of several deep well plates were used to determine the protein concentration via the Bradford assay [see 3.2.5.15] and to calculate the amount necessary for TCA protein precipitation. 30 µL of the supernatants were precipitated with TCA [see 3.2.5.16] in order to load approximately 10 µg of protein on each slot of a SDS gel. After gel electrophoresis the bands for the MBP moxY constructs with the linker for TEV-cleavage were expected at a size of 108.8 kDa.

#### Dot blot for the screening of MBP-tagged proteins

A dot blot was applied to screen for high expression of the MBP moxY constructs with and without His-tag in a high-throughput format. After their cultivation in deep well plates, the protein extraction was done with Y-Per reagent and DTT at 45 °C and the supernatants were spotted on a membrane by pipetting and application of a vacuum. Therefore the Bio-Dot Apparatus from Bio-Rad and a nitrocellulose membrane (#RPN203D) from GE Healthcare were used. Before sample application, the membrane was moistened with 1 x TBST [see Table 15], placed on the sealing and fixed in the dot blot apparatus. 15 µL of the samples were applied eight times in a row and sucked through the membrane with vacuum. The membrane was rinsed and incubated in an omni tray at 4 °C over night with 20 mL 1 x TBST milk to block unspecific binding areas.

The membrane was then incubated for 1 h in 15 mL of the primary antibody while gently shaking the omni tray. Therefore, the MBP-tag Antibody, mAb, Mouse from GeneScript was used and diluted to a concentration of 0.5 µg/mL with TBST milk. To wash away residual antibodies, the membrane was incubated three times for 5 min in 20 mL 1 x TBST. Afterwards, the membrane was incubated with 15 mL secondary antibody for 1 h, using the Anti-Mouse goat IgG - Alkaline Phosphatase, which was diluted 1:30000 in TBST milk to a concentration of 1.4 µg/mL. Subsequently, the membrane was washed for 5 min with 20 mL 1 x TBST for three times. Finally, the membrane was wet with 5 mL BCIP/NBT solution (Calbiochem) and incubated for 15 min without light for the visualization of the protein bands. The color reaction was stopped with dH<sub>2</sub>O and the membrane was scanned. The clone with the darkest color, indicating the highest level of expression, was chosen, striked out on agar plates and used for cultivation.

## Dot blot for screening of proteins with His-tag

The same dot blot as described for MBP-tagged proteins was performed for screening of proteins with a His-tag. Therefore, the supernatants of the construct HIS MBP moxY after Y-Per extraction were spotted on a membrane. In contrast to the other samples, DTT was only added for the protein extraction of the samples in column 1 to 3 [see Figure 52]. For the dot blot the Anti-His-antibody (primary antibody, polyclonal, produced in rabbit, diluted 1:7500 in 1 x TBST milk) and Anti-Rabbit-IgG alkaline phosphatase antibody (secondary antibody, diluted 1:30000 in 1 x TBST milk) were used.

## MBP purification pretests:

One clone of each MBP moxY construct was chosen via SDS-PAGE analysis [see Figure 50] and cultivated in 2 L shake flasks [see 3.2.5.14] for small scale purification pretests. The following clones were cultivated:

- pPp\_B1 MBP (with linker) moxY E2
- pPp\_B1 MBP (no linker) moxY A4
- pPp\_B1 moxY MBP (no linker) A9
- pPp\_B1 His-tag MBP (with linker) moxY G12
- pPp\_B1 MBP (no linker) citrine
- *P.p.* CBS 7435 mutS (strain collection nr. 3445)

The cells were harvested by centrifugation at 3000 x g and 4 °C for 10 min. The pellets were resuspended in 100 mL ice-cold, autoclaved dH<sub>2</sub>O and centrifuged again. The breaking buffer was prepared according to M.D. Andersen *et al.* (2002) [see Table 15, but without complete protease inhibitor]. The pellets with 5 - 8 g cell wet weight were resuspended in 20 mL breaking buffer and the cell suspension was added to vials with glass beads. The cell breakage was done mechanically by putting the samples for 3 min in the Cell homogenizer MSK from P. Braun Biotech International GmbH. During the breakage, the cells were cooled by flashing the samples shortly with CO<sub>2</sub> gas all 30 seconds. The supernatant was transferred into centrifugation tubes and centrifuged at 10000 x g and 4 °C for 10 min. The pellets were discarded and the supernatants were ultracentrifuged for 1 h at 4 °C and 50000 rpm. Samples from the supernatants before as well as after ultracentrifugation were taken and used for small scale purification trials.

The protein concentrations in the supernatants were determined with the Bradford assay and concentrations of ~10 µg/µL were obtained. 100 µL amylose resin material was filled in 1.5 mL Eppendorf tubes, washed with 800 µL dH<sub>2</sub>O and equilibrated twice with 1 mL binding buffer [see Table 15]. 5 mg of protein, which corresponded to ~500 µL of supernatant, were added to the column material and incubated for at least 5 min. The amylose resin was

washed four times with 800  $\mu$ L of washing buffer before the protein was eluted by centrifugation of the samples after incubation for at least 5 min with 60  $\mu$ L elution buffer [see Table 15]. The various wash and elution fractions as well as the samples before and after loading on the column were applied on a SDS gel.

#### His - tag purification pretests

In order to test, whether the protein can be purified with the His-tag of the His MBP moxY construct, trials with His Spin Trap from GE Healthcare Bio-Sciences AB were made according to the manufacturer's recommendations. After removal of the storage solution, the column was equilibrated with 600  $\mu$ L binding buffer [see Table 15] and centrifuged for 30 sec at 100 x g. 450  $\mu$ L of the supernatants after ultracentrifugation of the constructs HIS MBP (with linker) moxY and MBP (without linker) moxY, which were cultivated in shake flasks, were applied on the column. The construct without His-tag served as negative control since in that case, no protein should be eluted. The column was washed with 600  $\mu$ L washing buffer before 200  $\mu$ L imidazole containing elution buffer were applied to elute the His-tagged proteins.

#### Glycerol stocks

Glycerol stocks with a final glycerol concentration of 20 % were prepared of all strains, shock frozen in liquid nitrogen and stored at - 80 °C. For *E. coli* strains 3.5 mL LB zeo medium were induced with fresh cells from an agar plate and incubated at 37°C at 110 rpm over night. In the morning 0.5 mL of 60 % glycerol were added to 1 mL ONC in a cryo tube. The *P. pastoris* cells were incubated at 28 °C at 110 rpm in 10 mL YPD zeo medium for 24 h. After addition of 5 mL 60 % glycerol, the ONCs were incubated for further 30 min before 1.5 mL were transferred in the labeled cryo tubes. Furthermore, all strains were added to the strain collection [see appendix, chapter 10].

#### 4.3.4. Large scale protein production

##### Growth and expression

*P. pastoris* cells from glycerol stocks were streaked on YPD agar plates containing 100 mg/L zeocin and grown for 24 h at 30 °C. A single colony was picked and precultured in 150 mL of MGY media containing 100 mg/L zeocin. The preinoculum was grown at 30°C, at 250 rpm for 24 h until an OD<sub>600</sub> of 2 - 6 was reached. 100 mL of PTM<sub>1</sub> trace salts solutions were prepared and filter sterilized. The preinoculum was added to 3 L of fermentation basal salts medium

enriched with 4.35 mL PTM<sub>1</sub> trace salts/L. For the fermentation a 5 L Bioflo 3000 Benchtop Fermenter was used. Before the inoculation, the temperature of the fermenter was set to 30 °C and the agitation to 500 rpm. The pH has been adjusted to 4.5 using ammonium hydroxide.

After the complete consumption of glycerol (after 24 - 36 h) the oxygen returned to the saturation and the glycerol fed batch phase was started. 50 % glycerol solution, containing 12 mL PTM<sub>1</sub> trace salts/L, was added at a rate of 12 mL/h/L. To prevent foaming, Antifoam was given to the media when needed. Furthermore, the oxygen content was constantly monitored and eventually added to keep its level dissolved in the broth above 30 %. After 16 h the glycerol feeding rate was increased to 18 mL/h/L for another 6 hours. After the glycerol feeding, the cell density, which was 80 - 100 g/L at the beginning, had doubled. The feeding was stopped for three hours before the methanol induction phase. For that purpose a 100% methanol solution, containing 12 mL PTM<sub>1</sub> trace salts/L, was added to the fermentation broth at a rate of 1 mL/h/L. This rate was increased to 1.5 mL/h/L after one night and 2 mL/h/L after another 10 h. Then the feeding was raised to 2.5 and finally to 3 mL/h/L every 24 h. During the fermentation, the oxygen consumption, pH, temperature, agitation speed and cell density were constantly monitored. After seven days, from 3 L of fermentation broth more than 500 g of cells could be obtained.

#### Cell harvest and breakage

The broth was centrifuged at 7000 x g at 10 °C for 15 min and the supernatant was discarded. 50 g of cells were resuspended in 150 mL of breaking buffer [see Table 15]. The leftover was shock frozen in liquid nitrogen and stored at -80 °C. Attempts were made to improve the protein recovery yield by addition of a detergent. Therefore, 0.2 % Triton<sup>TM</sup> X-100 was added and the concentration of protease inhibitor was doubled. The whole work was performed while keeping the cells and the protein solutions constantly on ice. An equal volume of zirconia beads was added to the cell suspension and put into the BeadBeater from Bio Spec Products Inc. (Bartlesville, OK, USA). The mechanical breaking was kept for 2 min before placing the cells on ice for 5 min, repeating this cycle 10 times. Afterwards, the cell suspension was recovered by filtration through a miracloth filter using a vacuum pump to remove the zirconia beads. The raw extract was centrifuged at 70000 x g for 45 min at 4 °C. The obtained supernatant was filtered through a 0.45 µm Millipore prefilter and loaded on a chromatographic column for purification.

#### 4.3.5. Protein Purification

##### MBP affinity chromatography

For the purification of moxY different chromatographic steps were applied, which were all performed at 4°C using the ÄKTA FPLC system of GE Healthcare Life Sciences (Uppsala, Sweden). First of all, a MBP affinity chromatography column was used to isolate the MBP-moxY fusion protein. Depending on the amount of cells used for the breakage and the amount of protein loaded, either a 1 mL or 5 mL MBPTrap HP column was used.

The MBP column (maximum pressure over the column: 0.3 MPa) was initially washed with water and equilibrated with washing buffer A [see Table 15]. The filtered supernatant was loaded on the column using a 150 mL superloop from GE Healthcare for the sample application. In order to ensure efficient binding of the MBP moxY fusion protein, the flow rate was set to 0.5 mL/min and loading was performed over night. Afterwards, the column was washed with washing buffer to remove unspecific proteins until the absorbance signal at 280 nm was not detected anymore. The MBP moxY fusion protein was eluted by applying the maltose-containing elution buffer A (100 %, without a gradient) and fractions with a volume of 1.5 mL were collected.

To monitor the purification steps, SDS-PAGE was used. The fractions of interest containing the MBP moxY fusion protein with a molecular weight of ~109 kDa were identified and pooled together.

##### UV-Vis Spectrum

The protein concentration and the ratio of flavin-bound protein were determined by UV-Vis absorbance spectroscopy, using a NanoDrop ND-100 from ThermoScientific (NanoDrop products, Wilmington, DE, USA). The absorbance at 280 nm was used for calculation of the protein concentration via the molar extinction coefficient of moxY ( $\epsilon = 140550 \text{ M}^{-1} \text{ cm}^{-1}$ ) and MBP moxY ( $\epsilon = 208390 \text{ M}^{-1} \text{ cm}^{-1}$ ). The molar extinction coefficients were determined with the ExPASy webtool on <http://web.expasy.org/protparam/>. The absorbance at 450 nm was used to determine the flavin content.

##### TEV cleavage and concentration

The MBP tag was specifically removed by TEV-protease digestion. Since UV-Vis absorbance spectra revealed that the protein was prone to lose FAD, this cofactor was added to the protein solution. The incubation of the protein with FAD and TEV-protease, which was added in a ratio of 1:40, was performed during dialysis at 4 °C over night, against dialysis buffer. In later attempts, 50 mM arginine and glutamic acid were added to the buffer in order to



improve the solubility of moxY. Various methods, like anion exchange chromatography, size exclusion chromatography or reiterative steps on a MBP affinity column, were used to facilitate the separation of MBP and moxY after the TEV cleavage.

Before the protein could be loaded to a column for further purification steps, it had to be concentrated to the appropriate volume. Therefore different membranes were tested, including cellulose membranes (Amicon® Ultra Centrifugal Filters from Millipore) and polyethersulfone membranes (Vivaspin 2® and Turbo 15® from Sartorius Stedium). The concentration steps were performed via ultrafiltration at 4 °C and 3000 x g.

#### Anion exchange chromatography

Anion exchange purification was performed to separate the two proteins by elution via a salt gradient using elution buffer B [see Table 15]. Therefore a Mono Q column was used and the flow rate, alarm pressure etc. were set according to the instructions in the manual of the manufacturer. The column was equilibrated with dialysis buffer and washed after loading of the sample. For protein elution, the ratio of the elution buffer was gradually increased. Next to the elution buffer containing 1 M NaCl which was used for the first purification of the fresh cells, also other Tris buffers with different concentrations of NaCl and KCl were tested. The elution of moxY occurred at a salt concentration of ~ 400 mM [see Figure 57].

#### Size exclusion chromatography

For subsequent attempts gel filtration was used to purify moxY, which is based on the separation of proteins according to their size. A Superdex 75 10/300 GL column was equilibrated with washing buffer A [see Table 15]. The bed volume of the column was 25 mL and its maximum pressure 1.8 MPa. The elution was carried with a flow rate of 0.2 mL and fractions with a volume of 0.3 mL were collected.

Moreover, trials were made using a second MBP trap column step to get rid of the cleaved MBP tag as well as the eventually uncleaved full length MBP moxY protein. moxY was expected to be found in the flow through, whereas the tagged species should remain bound to the column. A dialysis was done in advance to remove the maltose which was still present in the sample. Subsequently, removal of the TEV-protease was necessary in the last step, for example via size exclusion chromatography.

## Preliminary activity tests

Initial experiments were performed with the full MBP moxY fusion protein, to test whether some activity of the enzyme can be detected. Therefore, the enzyme (5  $\mu$ M) was incubated over night with a five-fold excess of flavin. The assay was done at 25 °C in 50 mM Tris/HCl buffer, pH 7.5, with a concentration of NADPH (or NADH) from 50 to 100  $\mu$ M. Acetone and phenylacetone (100-500  $\mu$ M) were used as artificial substrates, since the natural one is not available. For detection, the absorbance of NADPH (or NADH) was measured spectrophotometrically.

### 4.3.6. Crystallization

After the purification, the fractions were pooled, concentrated, supplied with FAD and used for first crystallization trials by vapor diffusion. Ammonium sulfate and PEG screenings from Nextal sold by Qiagen (Hilden, Germany) were performed by an Oryx8 Protein Crystallization Robot. These screenings provided a ready-to-use kit format where only the protein had to be added. 96 defined chemical solutions were available for each precipitant. The ammonium sulfate suite provided various buffers with different pH values and numerous salts which may be effective co-crystallizers. The PEG suite was used to test polyethylene glycol solutions with different molecular weights, concentrations and pH values. Half of the screening conditions contained various salts to investigate the effect of salt-PEG combinations on the protein ([www.qiagen.com](http://www.qiagen.com)).

For crystallization, a protein concentration of 2.9 mg/mL containing 1 eq FAD was used in 25 mM Tris/Cl buffer, pH 8.0, 20 mM NaCl and 5 % glycerol. The crystallization plates were stored at 4 °C during the first days and then put to 20 °C, where after 20 days the first results could be observed.

## 4.4. Results and discussion

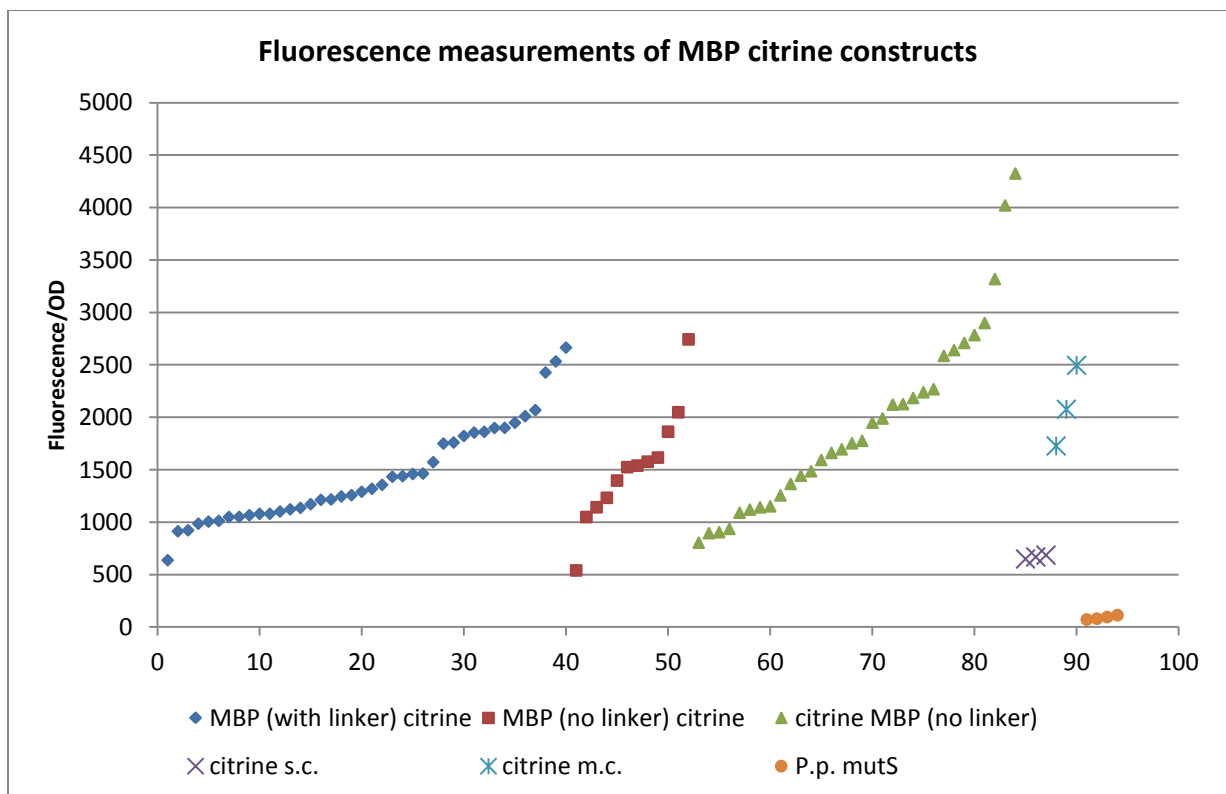
### 4.4.1. Optimization of the MBP gene for *P. pastoris*

The sequence of the MBP gene after optimization for *P. pastoris* is shown in the appendix in chapter 10. The GC content and GC peaks were determined via the webtool EMBOSS 6.3.1 [see Figure 47], giving a total GC percentage of 46.07 % (<http://mobyli.pasteur.fr/cgi-bin/portal.py?form=freak>).



#### 4.4.2. Screening of MBP citrine constructs

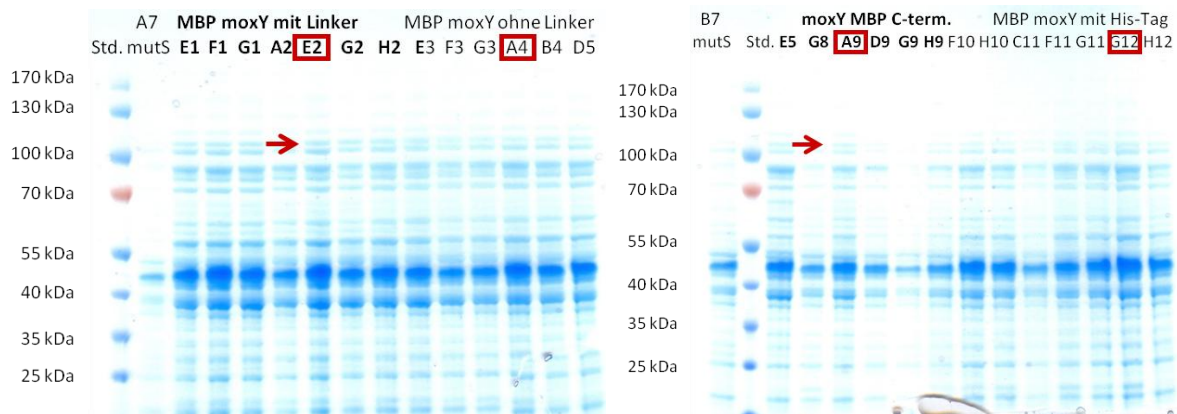
The aim of the MBP citrine constructs was to confirm that the expression of MBP fusion proteins in *P. pastoris* is generally possible and to investigate influences of the tag. Therefore, the fluorescence of the prepared constructs was determined with the spectrophotometer. The highest fluorescence/OD values were measured for the C-terminal construct citrine MBP [see Figure 49]. As expected, no fluorescence was measured for the wild type *P.p. mut<sup>S</sup>* strain. The signals of citrine m.c. were located in the range of the highest fluorescence signals obtained from the MBP citrine constructs and were clearly higher than those of citrine s.c.. For a direct comparison, the copy numbers of the MBP constructs could be determined by real time PCR. The variation between the signals of citrine s.c. was only small, whereas larger deviations were found for citrine m.c. which may result from fluctuations of the cultivation conditions. All in all, from these results it can be concluded, that the expression of citrine with MBP-tag was very efficient and that the expression of *moxY* with MBP should be possible, as well.



**Figure 49: Fluorescence measurements of the MBP citrine constructs and the control samples citrine m.c., citrine s.c. and *P.p. mutS*.**

#### 4.4.3. Screening of MBP moxY constructs and purification pretests

SDS-PAGE analysis was performed to investigate the expression of the different MBP moxY strains after protein extraction and precipitation [see Figure 50]. The bands corresponding to MBP moxY were expected at ~109 kDa, but they were very weak. The bands around 70 - 100 kDa, which were not found in the wild type, may represent some degradation products of the fusion protein. The protein bands of the negative control *P.p.* mutS were mainly limited to the region from 40 to 55 kDa, although the low protein levels may also be influenced by difficulties during the sample application. Differences in the total amounts of protein within the various strains could be due to different expression levels or fluctuations during the protein precipitation since a quantitative recovery of the very small protein pellets was difficult. Provided that the sample treatment was identical for all strains, the protein expression of the N-terminal constructs seemed to be higher than those of the C-terminal constructs. The four clones, which were selected for further cultivation based on their protein bands, were marked in red.



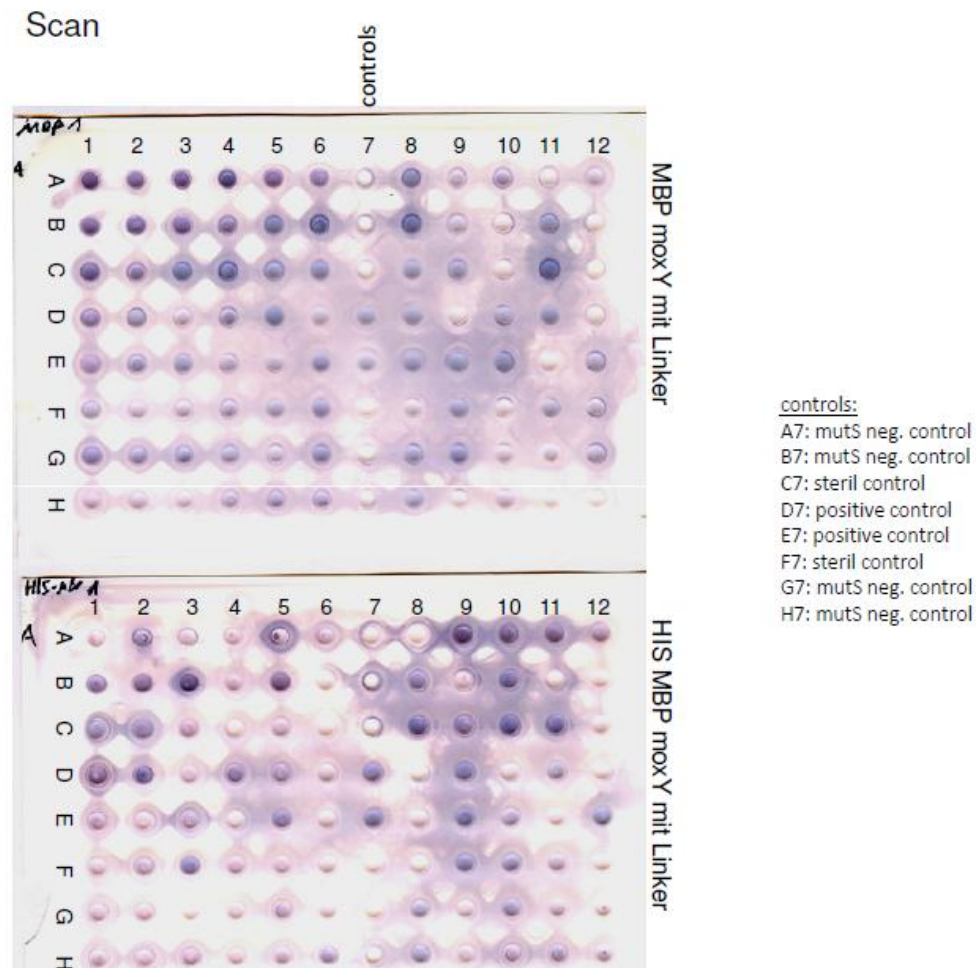
**Figure 50: SDS gels of the MBP-moxY constructs (109 kDa). Those clones that were selected for cultivation are marked in red.**

#### Dot blot analysis for MBP-tagged proteins

Screening for *P. pastoris* strains with a high expression level of moxY represented a major challenge since no enzymatic assay was available, which allowed an activity based screening. Furthermore, the selection by SDS-PAGE, which required Y-Per extraction and TCA precipitation of the protein, was very time-consuming and labor- as well as cost intensive. Therefore, its application was limited and a new screening system was required, which allowed the fast and reliable investigation of a large number of clones. As a result, dot blot analysis with MBP antibodies was tested, which represented a new tool for screening of the expression of MBP-tagged proteins in *P. pastoris*. Thereby, very promising results could be obtained.

After staining of the dot blot membrane, the clones with the highest protein expression could be easily identified by comparing their color intensity [see Figure 51]. As expected, no coloring was detected for the *P.p* mutS negative controls (see lane 7), whereas a dark violet staining was seen for some clones, indicating a high level of protein expression. The strain of MBP (with linker) moxY which was used for the first large-scale protein production served as positive control for direct comparison to other strains (see wells D7 and E7). Dot blot analysis revealed, that strains with a clearly more intense coloring could be found suggesting a higher amount of produced protein. Consequently, those clones were selected and used from then on.

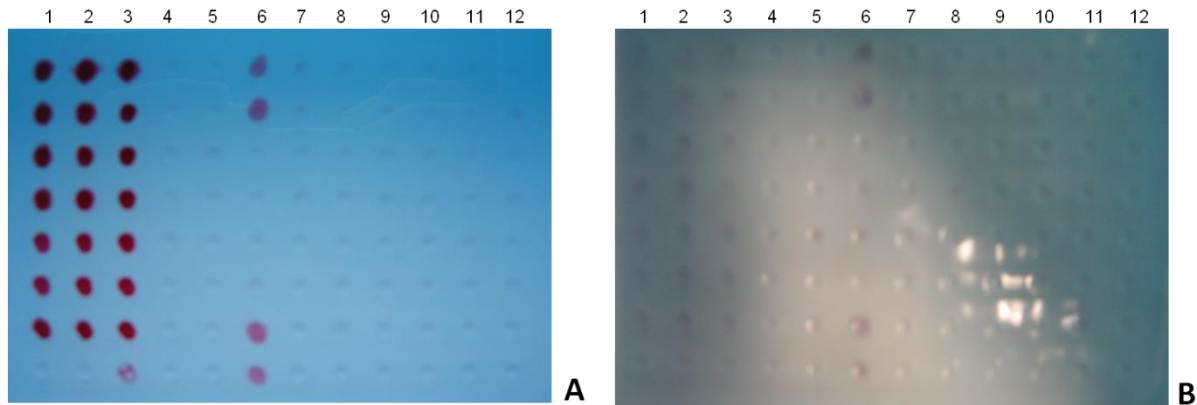
The results of this dot blot showed, that a new high throughput screening system for MBP-tagged proteins in *P. pastoris* could be established.



**Figure 51: Results of the dot blot with MBP-antibodies of the MBP moxY constructs with and without His-tag. Two clones with a very intense color were chosen for the MBP (with linker) moxY and the HIS MBP (with linker) moxY construct and stored as glycerol stocks, namely A4 and A9.**

## His - tag dot blot

The dot blot for screening for the expression of the His-tagged construct HIS MBP moxY, did not lead to as conclusive and clear results as the dot blot for MBP. Only a very faint coloring of the positive controls (the first two and the last two spots in column 6) and the samples of the first three lanes could be detected [see Figure 52 (B)].

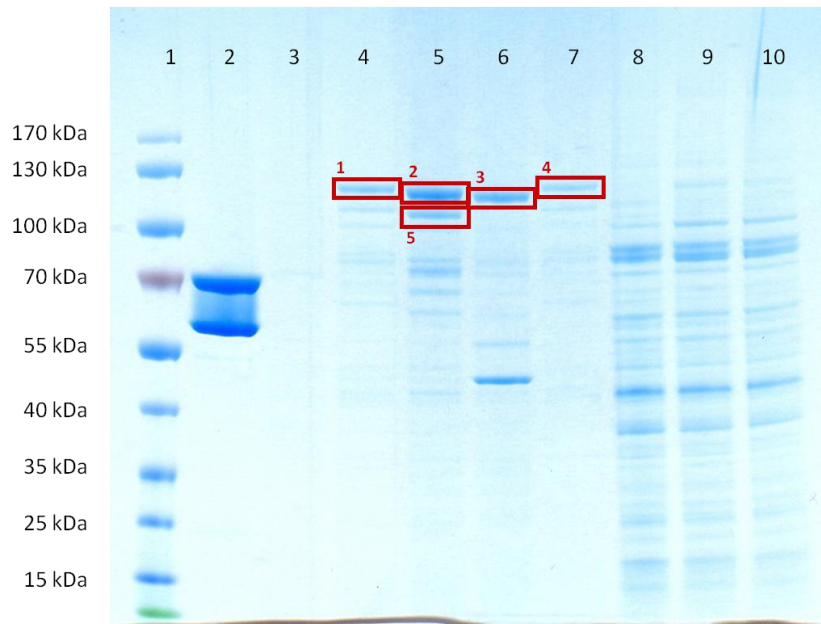


**Figure 52: Membranes of the His-tag dot blot. (A) The membrane after application of the sample and protein staining with Ponceau solution. (B) Dot blot membrane after final staining with BCIP/NBT solution.**

To ensure that the application of the protein was successful, the membrane was stained with Ponceau S staining solution for visualization of proteins [see Figure 52 (A)]. Thereby, a high protein concentration was detected for the samples of the first three columns, whereas no red coloring of the other samples could be seen, indicating that no protein was transferred to the membrane. In contrast to all others, the Y-Per protein extraction of the samples in the first three columns was performed at 45 °C and with addition of DTT. The much weaker signals in the last lane were traced back to application problems during pipetting. The increased protein yield after extraction with DTT at higher temperatures may be due to the ability of DTT to prevent miss-folding. Furthermore, reducing agents like DTT may act on disulfide bonds within the protein and can lead to depolymerization of polymeric compounds. Consequently, the protein solubility may be improved (88), (89). Instead of proceeding with this dot blot work was focused on the MBP-tag dot blot, which led to much better results.

## MBP purification pretests

Purification pretests with an amylose resin material were performed in Eppis to investigate whether elution of the MBP-tagged protein of interest was possible. The SDS gels of the respective purification fractions proved the successful elution of the MBP moxY fusion protein [see Figure 53].



**Figure 53: SDS gel of MBP purification pretests. (1: Prestained Protein Ladder, 2: eluate of MBP Citrin (67 kDa), 3: eluate of *P.p. mutS* (empty strain, negative control), 4: MBP moxY with linker sequence (109 kDa), 5: MBP moxY without linker (106 kDa), 6: moxY MBP C-terminal without linker (106 kDa), 7: MBP moxY with HIS tag and with Linker (110 kDa), 8: cell lysate of *P.p. mutS* (empty strain, negative control), 9: cell lysate of MBP moxY without linker before loading it on the column, 10: cell lysate of MBP moxY without linker after loading it on the column)**

At the gel the elution of the different MBP moxY constructs with an expected size of 106 -110 kDa is shown [see Figure 53]. MBP citrine (lane 2), which was cultivated for control reasons, is less complex and showed a much higher expression. For both MBP fusion proteins, with citrine as well as moxY, a version of the construct that is a few kDa lighter was found. There seemed to be some kind of degradation at one end of the protein. At the C-terminal construct in lane 6, a partial cleavage of MBP may have occurred, since an intense band at ~42 kDa, which could represent MBP, can be seen at the gel. However, in that case the second fusion partner moxY would be expected to be found in approximately the same amounts, but no clear band at ~66 kDa was detected. One possible reason therefore could be degradation of moxY. As expected, no proteins were eluted for the *P.p. mutS* strain (lane 3). The only difference in the protein distribution of the cell lysate of *P.p. mutS* (lane 8) and that of the prepared construct (lane 9) consisted in the band of the MBP moxY fusion protein (106 kDa). After loading the sample onto the column, this band got fainter, which indicated the successful but incomplete binding of the protein of interest to the column material (lane 10). Since the pretests were done in Eppis and with short incubation times, it is not astonishing that there was no perfect binding to the column. Exactly the same results were obtained when the cell lysates were ultracentrifuged to remove remaining parts of the membrane (gels not shown). This led to the assumption that moxY is not, or at least not fully, membrane bound. In that case, a different elution profile with a decreased yield of MBP moxY would



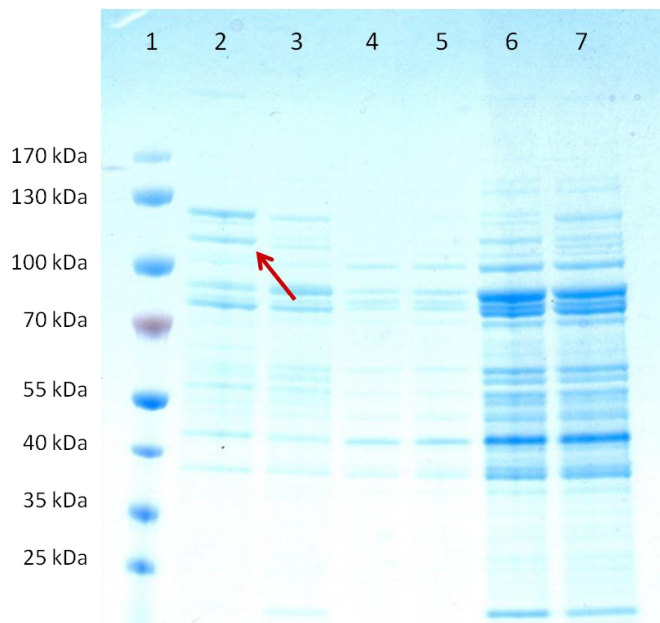
have been expected since after ultracentrifugation moxY should have been separated together with the membranes, if it is associated to them. The purification pretests also led to the same outcome when the frozen cell extracts were used.

MS analysis of the protein bands of all four constructs confirmed in general that the bands corresponded to the MBP moxY constructs, although not all amino acids were identified correctly (MS results are shown in the appendix in chapter 10, E). The second band at lane 5 was sent for MS analysis as well, but no significant differences in the sequence of the first and the second band, which may represent a degradation product, were identified. The coverage of the N-terminal MBP moxY constructs with linker was quite good, although the His-tag was not detected for the construct of lane 7. All in all, the MBP moxY constructs could be successfully eluted and so the up-scaling of the cultivation of the cells was started.

#### His - tag purification pretests

SDS-PAGE analysis after the His-tag purification pretests revealed that the protein of interest could not be obtained in a pure form [see Figure 54]. It was expected that only the protein His MBP moxY (~ 110 kDa) was eluted from the column but as it can be seen in lane 2, various other proteins were eluted as well. However, it seemed that His MBP moxY could be concentrated compared to the loading fraction in lane 4, where the band, that is thought to correspond to the moxY fusion protein, was not visible. Thereby it has to be considered that the loading fractions (lane 4 and 5) were diluted 1:20. In these samples the amount of MBP moxY, with or without His-tag, seemed to be too small to be detected.

Partly His MBP moxY seemed to be eluted already in the wash fraction (see lane 6), which may result from a weak binding to the column. The wash fraction of the MBP moxY construct without His-tag was applied on lane 7. As expected, no band was found at the corresponding height in the elution fraction of MBP moxY (lane 3) since it should not bind to the column without a His-tag.



**Figure 54: SDS gel after His-tag purification pretests, the red arrow indicated the band that should correspond to HIS MBP moxY. (Lane 1: Protein ladder Prestained PageRuler, 2: eluted fraction of HIS MBP moxY, 3: eluted fraction of MBP moxY, 4: loading fraction of HIS MBP moxY, 1:20 diluted, 5: loading fraction of MBP moxY, 1:20 diluted, 6: wash fraction of HIS MBP moxY, 7: wash fraction of MBP moxY)**

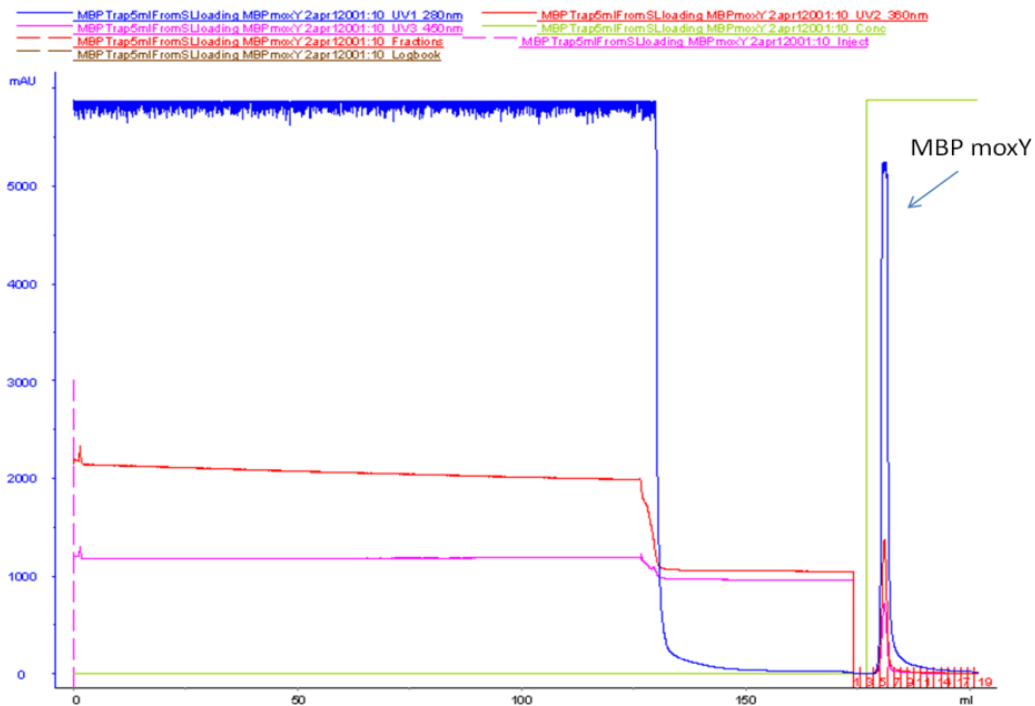
To sum up, in contrast to the MBP purification pretests, the His MBP moxY fusion protein could not be isolated in His-tag purification tests. Several other proteins were eluted from the column, possibly because unspecific binding occurred. Repeating the tests under modified conditions, for example longer binding times or more washing steps, may improve the yield of His MBP moxY. However, since the purification with the MBP-tag resulted in significantly better protein yields and a much higher degree of purity, the focus was laid on the purification strategy with MBP.

#### 4.4.4. Purification of moxY

##### MBP affinity chromatography

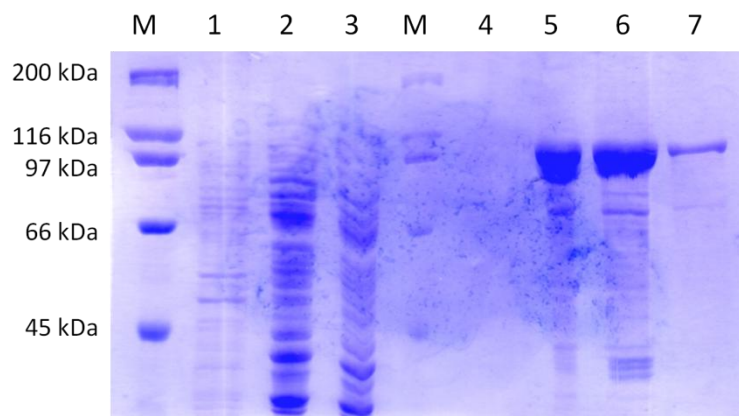
The first purification step consisted of a MBP affinity chromatography. Using the fresh cells directly after the fermentation, a large amount of the MBP moxY fusion protein could be eluted via applying the maltose containing elution buffer [see Figure 55]. The saturated signal of the UV-absorbance at 280 nm (blue line) was due to the large protein content in the flow through. After the loading, the column was washed to remove unspecific bound protein. Afterwards, the elution was started by applying the elution buffer (green line). A sharp and high protein peak of MBP moxY could be obtained. The increase of the absorbance both at

360 and 450 nm along with the 280 nm signal, indicated that the eluted protein was a flavoprotein.



**Figure 55: Chromatogram of the MBP affinity chromatography of MBP moxY. The UV absorbance of proteins at 280 nm is shown in blue, the flavin absorbance at 450 nm in pink and those at 360 nm in red, the concentration of the elution buffer is depicted in green and the fraction numbers are shown in red.**

SDS-PAGE confirmed that the eluted protein corresponded to MBP moxY (109 kDa), with a high protein content in the fractions of the top of the peak (see lane 5 and 6 in Figure 56) and nearly any protein at the beginning and at the end of the peak (lane 4 and 7). Furthermore, the elution fractions contained some other protein impurities or degradation products.



**Figure 56: SDS gel after MBP affinity column (M: marker, 1: pellet, 2: load, 3: flow through, 4 – 7: elution of MBP moxY (109 kDa), see fractions number 4-7 in Figure 55).**

The protein content in the pellet was very low (lane 1), which suggested that the cell breakage worked quite well and that most of the proteins could be extracted. The protein distributions of the loading fraction and the flow through (lane 2 and 3) were similar. It was expected that the protein band of MBP moxY will disappear in the flow through, since this was the only protein that should bind to the column. However, the proportion of MBP moxY compared to the other proteins seemed to be quite small so that no clear band of MBP moxY was detected neither in the loading fraction nor in the flow through.

### TEV cleavage and anion exchange chromatography

The fractions that contained the full construct MBP moxY with a size of 109 kDa were identified via SDS – PAGE [Figure 56], pooled and supplied with TEV – protease and FAD.

For the second purification step, the cleaved protein sample was loaded onto an anion exchange column. By applying a salt gradient (green line) the MBP tag and pure moxY could be separated [Figure 57]. MBP was eluted as a sharp and high peak at a salt concentration of ~ 200 mM, whereas the elution of moxY occurred at ~ 400 mM in the form of a broad, narrow and asymmetric peak (blue line). No UV-absorbance was detected at 360 and 450 nm for this peak which indicated that moxY did not retain the flavin cofactor.

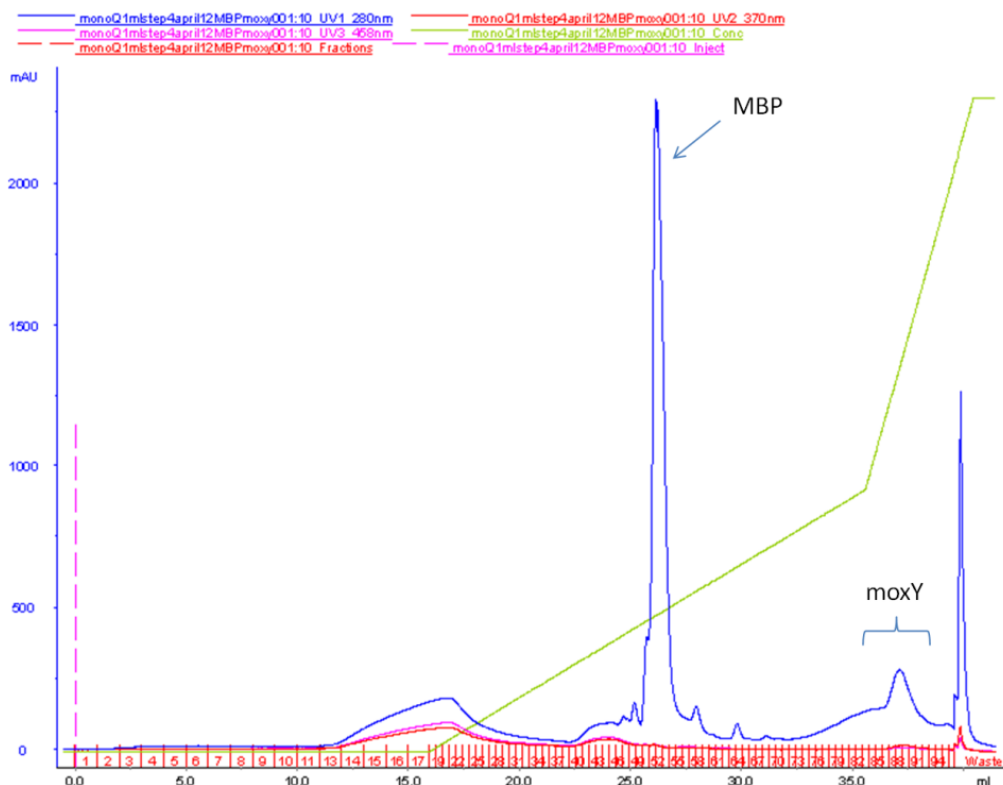
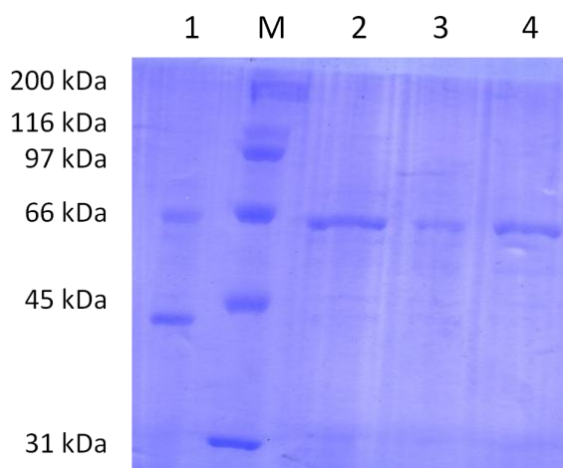


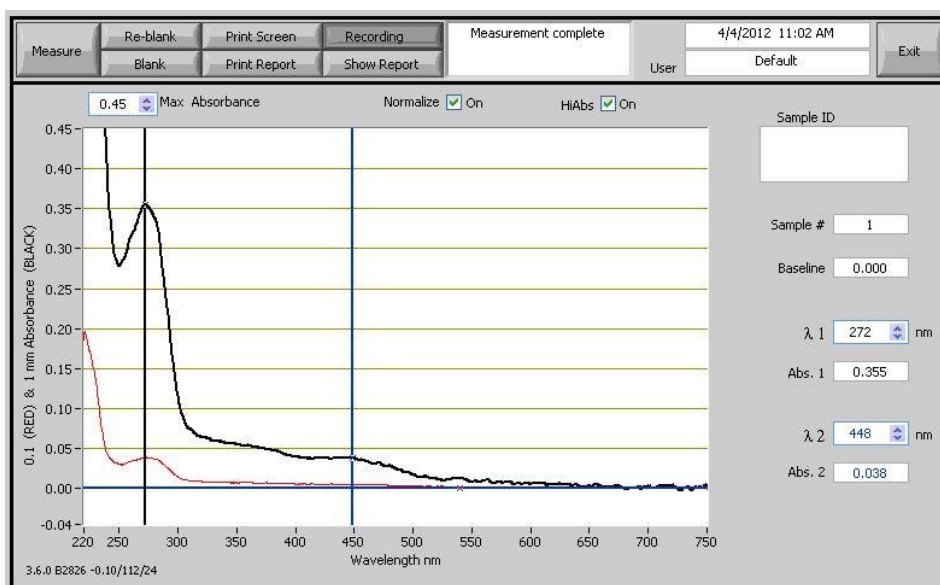
Figure 57: Chromatogram of the anion exchange purification of moxY. The UV absorbance of proteins at 280 nm is shown in blue, the flavin absorbance at 458 nm in pink and those at 370 nm in red, the concentration of the elution buffer is shown in green and the fraction numbers in red.

The successful TEV-cleavage and elution of moxY was confirmed via SDS-PAGE [Figure 58]. As expected, two bands were obtained after TEV-cleavage for moxY (66 kDa) and MBP (42 kDa), see lane 1. In case that the fusion protein was not cleaved completely, another band corresponding to MBPmoxY at 109 kDa would have been expected. The SDS-gel showed, that moxY could be eluted in a very pure form without other contaminants. The pure fractions of moxY (see lane 2 - 4) were combined and used for first crystallization trials.



**Figure 58: SDS gel after TEV cleavage and anion exchange column of moxY (M: marker, 1: MBP and moxY after TEV cleavage, 2 – 4: elution of moxY, fractions number 89 to 91 in Figure 57).**

The UV-Vis spectrum of the concentrated, pure moxY showed apart from the protein peak at 280 nm only very little flavin absorbance at 450 nm, indicating that the protein lost FAD [Figure 59]. Even when some FAD was added to the sample, for example during the TEV-cleavage, no FAD was retained at the end. Therefore, the protein was incubated with the flavin cofactor once more before crystallization trials were started.



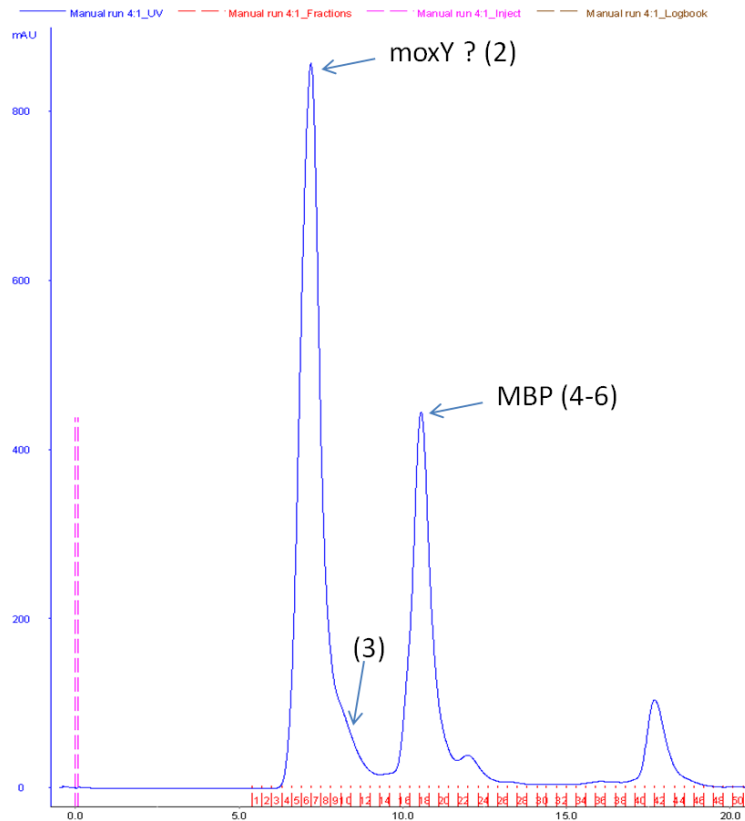
**Figure 59: UV-Vis spectrum of moxY with a protein concentration of 1,6 mg/mL.**

The loss of FAD may be caused by incorrect protein folding, which could impair the binding of FAD. Furthermore, the removal of the MBP-tag could affect the retention of the cofactor. The tag may also have an influence when moxY is still linked to MBP, since the flavin content was even directly after the MBP chromatography not very high. Especially because of the large size of the tag and the fact that moxY is only 1/3 bigger, it is possible that the tag itself disturb the correct folding of the protein.

One option would be to test constructs where MBP is fused to the opposite terminus to evaluate whether the binding of FAD is influenced. Alternatively, other affinity tags, like for example SUMO, may lead to a differently shaped protein, although it is not sure whether modifying the protein folding will indeed lead to a better retention of the FAD cofactor.

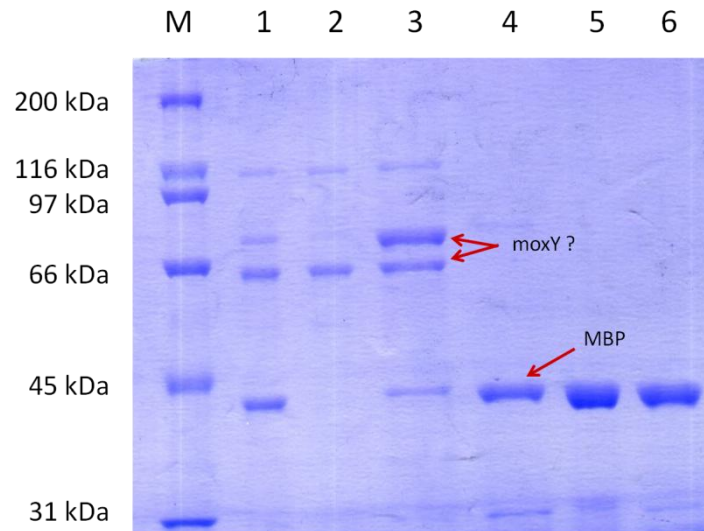
### Size exclusion chromatography

Instead of the anion exchange chromatography, a size exclusion chromatography could be carried out after the TEV-cleavage, as well. In that case, moxY and MBP were separated based on their different sizes. It was expected, that moxY is eluted first since it is larger than MBP and can therefore pass faster through the pores of the column material, followed by the MBP-tag and the TEV-protease, which has a size of only ~ 20 kDa. The chromatogram [see Figure 60] substantiated the elution of three peaks, though they had small shoulders. The amount of MBP and moxY was expected to be the same, but the peak of moxY was clearly higher. The third peak was expected to correspond to the TEV-protease but could not be seen on the SDS gel, as its size may have been too small.



**Figure 60: Chromatogram of the gel filtration of moxY after TEV-cleavage. The UV absorbance of proteins at 280 nm is shown in blue and the fraction numbers are given in red. The numbers of the peaks in the brackets correspond to their position on the SDS-gel in Figure 61.**

The SDS gel [see Figure 61] showed that the TEV-protease did not cleave the fusion protein completely (see lane 1). There was still a small amount of MBP moxY left and instead of one single band at 66 kDa for moxY, two bands which were slightly above and below the expected size, were obtained. The highest peak at the chromatogram (number 2) contained mainly the form of moxY with the slightly lower molecular weight. Interestingly, in the shoulder at the end of the peak the higher band dominated, although it was expected that this protein band is eluted earlier as it is larger. Additionally, small amounts of the uncleaved full length MBP moxY protein at 109 kDa and the MBP-tag itself were found in the first peak. MBP was eluted in the second peak (number 4-6).



**Figure 61: SDS gel of the gel filtration of moxY after TEV-cleavage. At this attempt 0.2 % Triton™ X-100 was added for the cell breaking. M: marker, 1: sample after TEV cleavage, 2: moxY in first elution peak, 3: various bands in the weight range of moxY in the shoulder of the first peak, 4 – 6: MBP in second elution peak.**

In most of the following purification trials the TEV-protease didn't cleave the construct completely. Perhaps the efficiency can be improved by adding more of the protease or by using a fresh enzyme, without refreezing it too often. What's more, instead of a single band of moxY most of the time two or three bands, which could not be separated, appeared in the same weight range as moxY. Later attempts showed the presence of a contaminant with a similar molecular weight, that could be recognized as the heat shock protein HSP70 from yeast (results from MS-analysis of an eluted fraction, data not shown). So far, the identity of each of the bands, which may be caused by some kind of degradation, cleavage or contamination, could not be resolved. Possibly, MS-analysis after cutting out the respective bands of a SDS gel can help in clarifying these results.

Another alternative to obtain moxY in its pure form was to perform a second MBP affinity chromatography with the aim to recover moxY in the flow through, whereas the MBP tag and uncleaved MBP moxY were bound to the column. However, so far the yields of moxY that could be isolated after a second MBP column were too small to be used for further experiments.



## Initial activity tests

Another interesting and important point would be to check, whether there is any enzyme activity or not. This was tested in preliminary experiments with the whole fusion protein, since the solubility and stability of moxY with the MBP-tag seemed to be higher. Yet no catalysis could be detected, suggesting that moxY is not active, although the still remaining MBP-tag and the large difference between the tested and the natural substrate might have influenced the results. Since moxY seemed to retain no FAD, the absence of activity may also be caused by impaired binding of the flavin cofactor.

## Challenges at the purification and possible solutions

The main challenge consisted in the substantial loss of protein that occurred especially during the TEV-cleavage, the second purification step and the concentration of moxY. Only few precipitations could be noticed, but it seemed that the protein got stuck to membranes, since after concentration nearly no protein could be recovered anymore. In spite of the same purification protocol, the amount of moxY, which was obtained after the purification of frozen cell pellets, was clearly below the amount obtained from the fresh cells. Various strategies were tested to improve the protein yield. The membrane material of the centrifugation tubes was changed and cellulose as well as polyethersulfone membranes were used, but without success.

Furthermore, the addition of glutamic acid and arginine was tested with the objective to increase the solubility and stability of the protein. Although this strategy worked in other experiments (90) it did not lead to significant improvements in the case of moxY.

Another strategy was to increase the concentration of protease inhibitors in the buffer. The amount of PMSF and complete protease inhibitor were raised and the detergent Triton™ X-100 was added to the breakage buffer. With these modifications the amount of protein recovered from the frozen pellets could indeed be enhanced.

In order to test, whether an improvement can be reached by using a different strain and whether freezing of the cells had any influence, a new fermentation was started. A better expressing clone was selected by dot blot analysis with MBP antibodies and fermented (therefore clone A4 in Figure 51 was chosen). Nevertheless, the results of the protein purification could not be improved. In that respect, technical problems with the oxygen tank during the fermentation need to be mentioned. As a result, the oxygen supply sank below the recommended value of 30 % during the first night. At that point the methanol induction was

not started yet and the cells continued to grow fast on glycerol, but it might have affected the metabolism. For more conclusive results, the fermentation will have to be repeated.

Another possible variation would be to test whether the protein solubility and therefore the final yield can be improved by using the construct of *moxY*, where the MBP-tag is fused to the C-terminal end. Using MBP as a C-terminal tag for purification already worked in *Saccharomyces cerevisiae*, but so far it has not been tested in *Pichia* (91). The present experiments demonstrated that the C-terminal *moxY* MBP construct could be expressed in *P. pastoris* and that first small scale cultivations and purification pretests were successful. Additionally, the construct where MBP is fused to the C-terminal end of citrine showed even higher fluorescence signals than the corresponding N-terminal construct [see Figure 49]. Therefore, it may be a promising option to test the large scale cultivation and purification of the C-terminal *moxY* MBP construct.

A further possibility would be to use a truncated version of MBP, where small parts of the C-terminus are removed. According to experiments by Zhiguo Li *et al.* (83), this region may cause certain structural features that could trigger the partial cleavage or degradation of the protein.

Using the construct His MBP (TEV-cleavage-site) *moxY* and a His-tagged TEV-protease should allow the purification of *moxY* after TEV-cleavage in a single step via His-tag nickel affinity chromatography. Only *moxY* is supposed to be found in the flow through, whereas the MBP-tag, the uncleaved fusion protein and the TEV-protease should bind to the column. Alternatively the His-tag could be fused directly to *moxY* in the following order: *moxY* His (TEV-cleavage-site) MBP. In that case, a nickel resin material could be added already during TEV-cleavage. Because of its His-tag *moxY* will stick to the resin and can afterwards be eluted with imidazole. Perhaps the stability of *moxY* can be improved when it is connected to the resin material.

Although the affinity tag could be changed again, this did not seem to be a very promising approach since after removal of the tag, the problem with the low solubility of *moxY* will most likely remain. Besides that, the MBP affinity chromatography worked quite fine and the MBP *moxY* fusion protein could be obtained in a relative high purity and yield. The selected host seemed to be a good choice as well, given that the cultivation of *P. pastoris* and the protein expression were successful whereas previously, difficulties with *E. coli* as host system were encountered.

Alternatively, completely new approaches for the production of *moxY* could be tested. One option would be to test homologous genes from other species, for example from *A.*

*parasiticus* instead of *A. flavus*, since this may influence the protein expression as well as various properties of the enzyme, like its solubility and folding.

#### 4.4.5. Crystallization

Various different results could be observed performing screening grids with ammonium sulfate and PEG. The crystallization with PEG as precipitant was based on the promotion of protein-protein interactions through molecular crowding, whereas ammonium sulfate caused the dehydration of the protein by competing for water molecules at its surface (for further information see manual of the manufacturer at [www.qiagen.com](http://www.qiagen.com)). Both crystallization screens showed multiple conditions that caused unregular protein precipitation. Small, irregular needles with white or light yellow color and phase separation (in some wells of the PEG screening) could also be found.

In condition C6 of the ammonium sulfate screening, containing 0.2 M KI and 2.2 M  $(\text{NH}_4)_2\text{SO}_4$ , a yellow, symmetrical, hexagonal shaped, very small crystal was found after three weeks of incubation at 20 °C [see Figure 62]. Unfortunately, it was not possible to fish the crystal due to technical limitations. However, these conditions could be promising starting points for further investigations.



**Figure 62:** Possible crystal resulting from the ammonium sulfate screening with 0.2 M KI and 2.2 M  $(\text{NH}_4)_2\text{SO}_4$ .

#### 4.5. Conclusion

The aim of this project was to express and purify the mostly uncharacterized Baeyer-Villiger monooxygenase *moxY* from *Aspergillus flavus*, in order to obtain the pure protein for structural resolution by crystallography. The first step was the successful intracellular expression of the MBP *moxY* fusion protein in *P. pastoris* which could be achieved in small scale in DWPs and shake flasks as well as in a 3 L fermenter.

In order to evaluate the expression level of the different strains, a high throughput screening system for MBP-tagged proteins in *P. pastoris* was needed. With the help of MBP-antibodies, a new dot blot method could be established, which allowed the fast and simple screening of *P. pastoris* clones in the 96-DWP format. In this way, strains for large scale cultivation could be selected.

The subsequent purification of *moxY* turned out to be a very difficult task. It was possible to obtain a pure fraction of *moxY* via affinity chromatography followed by an anion exchange chromatography. This sample could be used for first crystallization trials but the yield, especially after the second chromatographic step, was very low.

Several challenges, which have to be addressed in the future, were encountered when the purification of *moxY* was repeated. Apart from the significant loss of protein yield after TEV-cleavage, concentration and the second chromatographic step, the main concerns consisted in the low solubility and very poor FAD retention of the enzyme. Even after incubation with the cofactor, *moxY* did not retain the flavin cofactor, impairing therefore its catalytic activity.

The aim of future experiments will be to obtain crystals suitable for structural investigation and to perform activity tests. To reach these objectives, the above mentioned difficulties have to be approached in advance, to facilitate the purification of high amounts of *moxY* in a pure and active form.

## 5. Chapter III: Expression and purification of PAMO

### 5.1. Background

The phenylacetone monooxygenase PAMO from the thermophilic bacterium *Thermobifida fusca* is one of the best studied BVMOs. Next to many other oxidation reactions, it catalyzes the conversion of phenylacetone to phenylacetate. A major advantage of this BVMO is its exceptional thermal stability with an activity half life of 24 h at 52°C. PAMO is active over a broad range of pH values, showing the best results at pH 7 - 9 (30). Other attractive features next to the thermostability and robustness are that this enzyme can tolerate relative high concentrations of organic solvents (92) and catalyzes various enantioselective BV oxidations or sulfoxidations (32), (34). Additionally, extensive studies were performed to analyze the structure, kinetic aspects and catalytic mechanism of this enzyme and led to a broad knowledge about the characteristics of PAMO.

PAMO was expressed in *E. coli* cells and purified according to an already established purification protocol in order to prepare crystallization plates and to measure the redox activity and oxygen consumption of PAMO with modified cofactors and coenzymes.

### 5.2. Experimental

The materials and methods used for the cultivation, purification and crystallization of PAMO corresponded to those of moxY and are described in chapter 4.2. The media and buffers that were applied for the cultivation and purification of PAMO are listed in Table 26.

**Table 26: Media and recipes for the cultivation and purification of PAMO.**

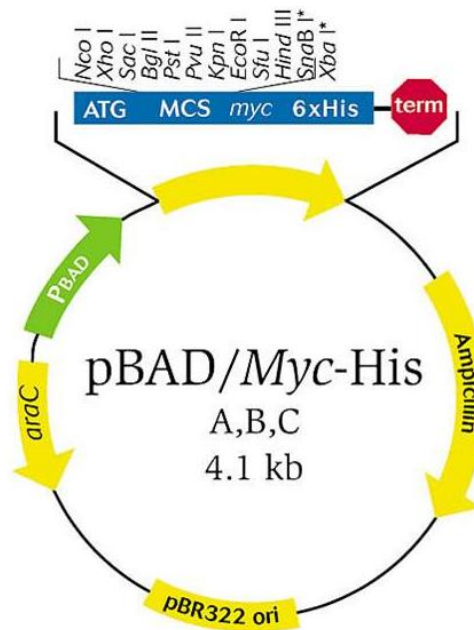
<b>Cultivation:</b>		
LB medium, V = 1 L	yeast extract	5 g/L
	NaCl	10 g/L
	trypton	10 g/L
TB medium , V = 1 L	yeast extract	24 g/L
	trypton	12 g/L
	glycerol	4 mL
	K <sub>2</sub> HPO <sub>4</sub> (1 M)	50 mL
	KH <sub>2</sub> PO <sub>4</sub> (1 M)	15 mL
<b>Protein purification:</b>		
breaking buffer, pH 7,5	Tris/Cl	25 mM
elution buffer, pH 7,5	breaking buffer	
	KCl	2 M

The buffers and media were prepared using deionized water and sterilized either by filtration or by autoclaving.

### 5.2.1. Cloning, cultivation and purification

#### Cloning

The gene encoding PAMO was already cloned into a modified pBAD/Myc-His vector [see Figure 63] from Invitrogen (where the His-tag codons have been silenced) and transformed into *E. coli* Top10 cells. The induction of the protein expression was elicited by the addition of L-arabinose.



\* Frame-dependent variations

**Figure 63:** The pBAD/Myc-His vector from Invitrogen for the expression of PAMO, the His-tag has been silenced (93).

For the transformation, 100 ng plasmid DNA was added to 50  $\mu$ L of chemical competent cells and incubated on ice for 30 min. After putting the sample at 42  $^{\circ}$ C for 25 sec, it was incubated on ice for another 2 min. 1 mL of LB medium [see Table 26] was added and the cells were incubated at 37  $^{\circ}$ C for 1 h. 100  $\mu$ L of this cell suspension were plated on LB agar plates, containing LB with 15 g/L agar and 50  $\mu$ g/mL ampicillin. The plates were incubated at 37  $^{\circ}$ C over night.

#### Growth and expression

Single colonies were used to inoculate 10 mL LB medium precultures containing 50  $\mu$ g/mL ampicillin and grown overnight at 37  $^{\circ}$ C. The precultures were then diluted 1:100 into 800 mL of TB medium [see Table 26]. After 3 h of cell growth at 37  $^{\circ}$ C at 180 rpm, an OD<sub>600</sub> of 0.6 was reached and the protein expression was induced by adding 0.02 % w/v L-arabinose along with a temperature shift to 30  $^{\circ}$ C. The cells were then harvested after 16 h.

## Cell harvest and breakage

The cell harvest was performed by centrifugation at 7000 x g for 10 min at 10 °C. 5 - 10 g of cells were resuspended in 40 mL of breaking buffer [see Table 26] before they were broken by ultrasonication using the settings described in Table 27.

**Table 27: Settings of ultrasonication.**

total time	2 min
pulse on	5 sec
pulse off	9 sec
amplitude	60 %

After the addition of FAD, the sample was subjected to thermo-precipitation. The broken cells were warmed up for 30 min at 50 °C to precipitate non-thermostable proteins and to simplify the purification of PAMO, before the sample was centrifuged at 70000 x g for 45 min. After centrifugation, the supernatant was filtered through a 0.45 µm filter. The clarified, filtered supernatant was subsequently applied on a column for the purification of the protein via two chromatographic steps.

## Protein purification

All chromatographic steps were carried out using the ÄKTA FPLC system from GE Healthcare. The first purification step consisted of an anion exchange chromatography, using a Q Sepharose HiPrep 16/10 QFF column from GE Healthcare. The maximum pressure over the packed bed was 0.15 MPa. After equilibration with breaking buffer [see Table 26], the filtered protein solution was loaded on the column with a flow rate of 2 mL/min. Since it was already known that the elution of PAMO occurred at ~ 250 mM KCl, the concentration of the elution buffer was set to 8 % (equal to 165 mM KCl) for elution of contaminants. The elution of the protein was facilitated by gradually increasing the ionic strength, therefore a gradient from 8 % to 30 % of elution buffer was applied and fractions with a volume of 1.5 mL were collected. The fractions containing the desired protein were identified via SDS-PAGE and combined. Additionally, an UV-Vis spectrum was recorded after each step.

Subsequently, a size exclusion chromatography was performed as a second purification step. The sample was concentrated via ultracentrifugation at 3000 x g to a volume of <2 mL, before it was loaded onto a Superdex 200 16/60 PrepGrade column with a maximum pressure of 0.3 MPa. The breaking buffer [see Table 26] was used for the equilibration as well as for the entire purification. The purification was performed as recommended in the manual of the manufacturer (Amersham Pharmacia Biotech Europe GmbH, Freiburg, Germany). After pooling the desired fractions, the protein was concentrated to >20 mg/mL for

crystallization trials. For the calculation of the protein concentration the absorbance at 441 nm of the UV-Vis spectra was divided by the molar extinction coefficient of PAMO ( $\epsilon_{441} = 12.4 \text{ mM}^{-1} \text{ cm}^{-1}$ ) and multiplied by the molecular weight of PAMO (62 kDa) to convert the molar concentration into mg/mL.

### 5.2.2. Crystallization

In previous experiments crystals of PAMO were obtained by vapor diffusion at 4 °C with a protein concentration of 18 mg/mL, 0.5  $\mu\text{M}$  FAD, 4 mM NADP<sup>+</sup>, 50 mM Tris/HCl, pH 7.5 and 40 % (w/v) PEG 4000, 100 mM MES/HCl, pH 6.5, 100 mM NaCl (34). To reproduce the crystals via the sitting drop method, a grid was performed around the already known crystallization conditions. Thereby, the protein, precipitant and salt concentrations were modified within a narrow range, using 14 - 20 mg/mL protein in 25 mM Tris/Cl pH 7.5, 34 - 40 % PEG 4000 and 50 - 150 mM NaCl and 100 mM MES buffer pH 6.5. The protein was heated up to 50 °C for 30 min and incubated with 2 - 4 mM NADP<sup>+</sup> and 0.5 - 5  $\mu\text{M}$  FAD. For the plates, 800  $\mu\text{L}$  of the respective mother solutions were filled in each well of the tray. 1  $\mu\text{L}$  protein sample was then added to 1  $\mu\text{L}$  of the mother solution on the provided bridge. The plates were carefully sealed with a plastic foil and stored at either 4 °C or 20 °C.

## 5.3. Results and discussion

### 5.3.1. Purification of PAMO

As a first step, an anion exchange chromatography was performed [Figure 64], which allowed the resolution of PAMO in an already reasonable pure form, separated from most of the other proteins. The first protein peak (blue line) represented the elution of contaminants with 8 % of elution buffer (~ 160 mM KCl, green line). Then, the gradient was raised to 30 % and PAMO was eluted at a salt concentration of 250 - 300 mM KCl (second chromatogram peak). The fractions of PAMO [see lane 6 + 7 in Figure 65] with a molecular weight of 62 kDa were identified via SDS-PAGE, pooled together and concentrated via ultracentrifugation.



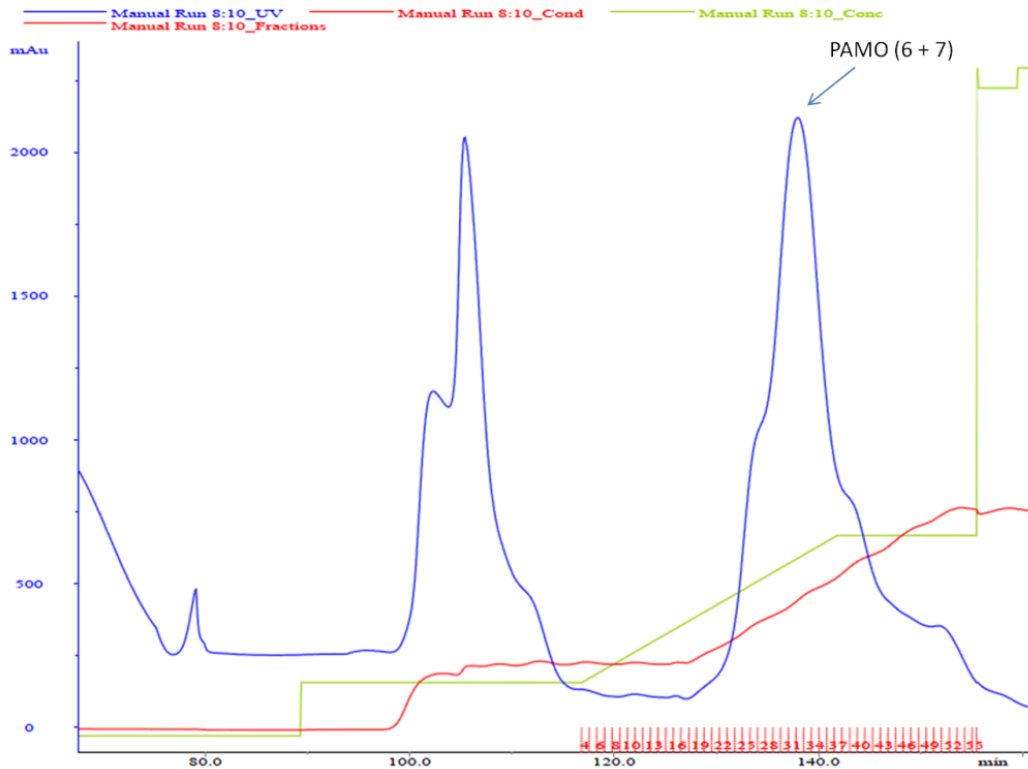


Figure 64: Chromatogram of the anion exchange purification of PAMO. The UV absorbance of proteins at 280 nm is shown in blue and the gradient of the elution buffer in green. The red line shows the increase in conductivity due to the higher salt concentrations. The red numbers represent the fraction numbers.

The SDS gel in Figure 65 showed the purification pattern: the raw extract was loaded on lane 1. It can be seen that the protein extraction worked well and that the main present protein consisted of PAMO (band with a size of ~62 kDa). In the second lane, the wash fraction was applied. As expected, various other protein contaminants were washed off while PAMO remained bound to the column. When the gradient was raised, PAMO could be eluted in a quite pure form (see lane 6 + 7), whereas other protein impurities were found in the shoulders on the left and right side of the peak (see lanes 3 - 5 and 8 - 9).

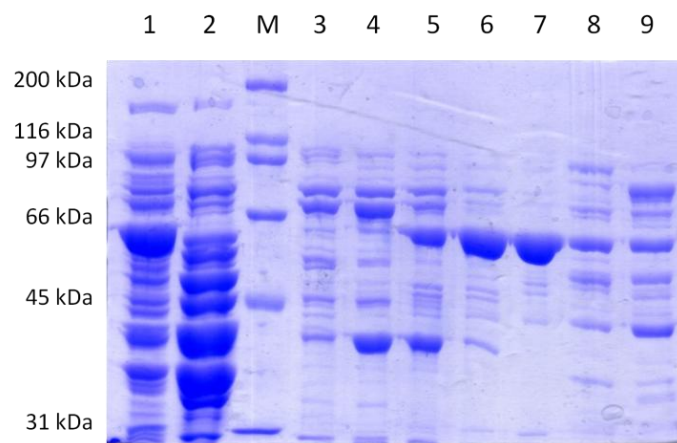


Figure 65: SDS gel of PAMO after anion exchange chromatography (1: load, 2: 8 % elution buffer, M: marker, 3 - 5: elution of impurities, 6 + 7: elution of PAMO, 8 + 9: elution of impurities).

To remove the remaining protein contaminants and undesirable aggregates, the sample was further purified via size exclusion chromatography [Figure 66]. Beside the small shoulders at the beginning, PAMO could be eluted in one clear peak (blue line) and was separated from the salt, as can be seen as an increase in the conductivity (red line).

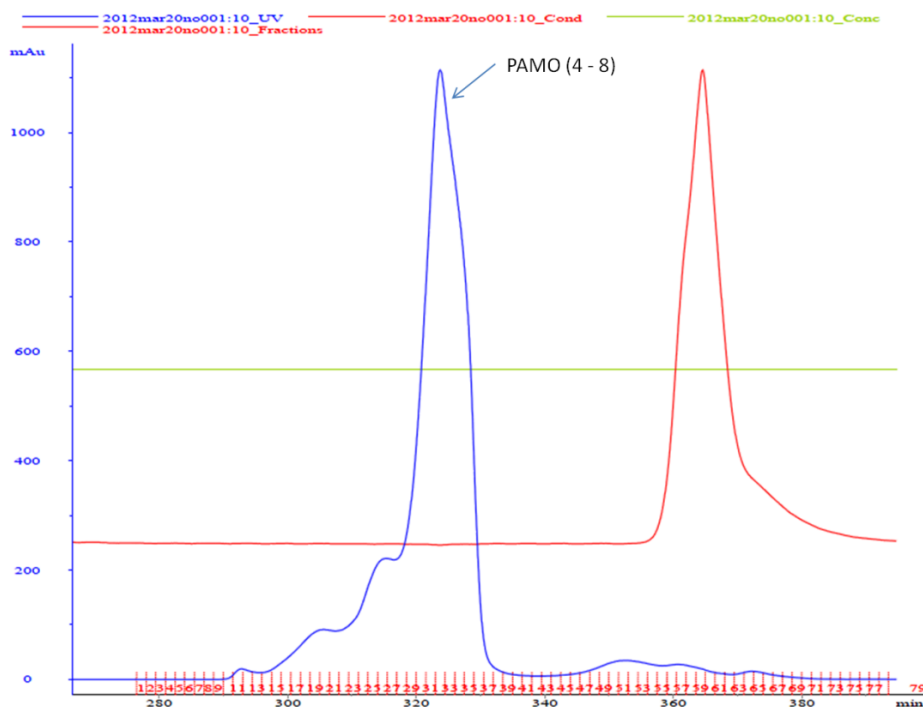


Figure 66: Chromatogram of size exclusion chromatography of PAMO. The UV absorbance of proteins at 280 nm is shown in blue. The red line shows the conductivity and the red numbers indicate the fraction numbers.

The SDS gel [see Figure 67] confirmed that other impurities (see lane 2 and 3) were found in the small shoulders on the left side of the peak, whereas the majority consisted of PAMO. The fraction that was loaded on the gel filtration column with PAMO as well as some other contaminants can be seen at lane 1.

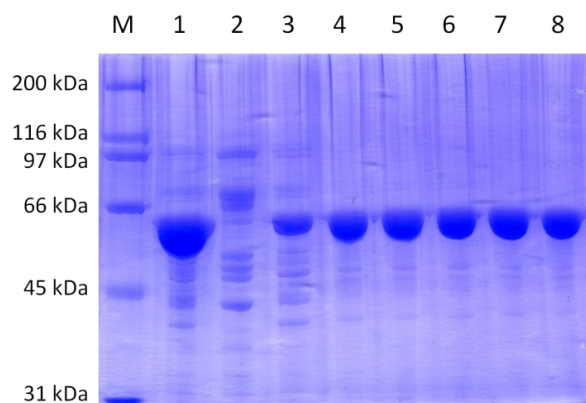
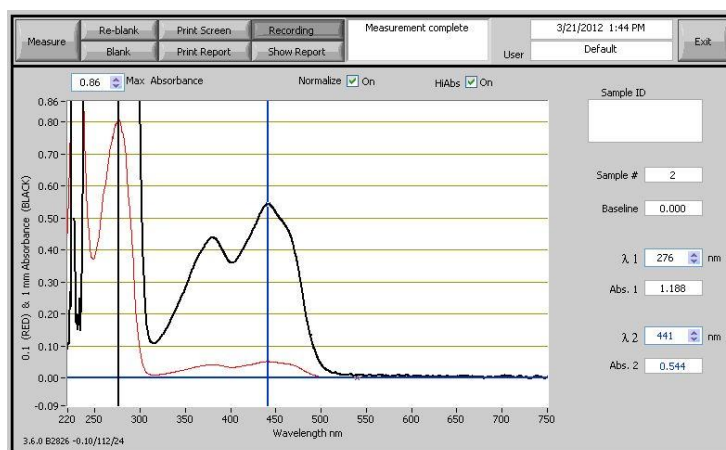


Figure 67: SDS gel of PAMO after size exclusion chromatography (M: marker, 1: load, 2 + 3: fractions of PAMO and other contaminants, 4 – 8: eluted fractions of pure PAMO).

Pure fractions of PAMO were collected (see lane 4 - 8) and could subsequently be used for the preparation of crystallization plates and other applications, like activity assays. Although the purification profiles were not always completely identical to the chromatograms shown, PAMO could be isolated in good yields and purities. The UV-Vis spectra of pure PAMO allowed the determination of the active holoprotein concentration [see Figure 68]. According to the literature, for globular monomeric flavoproteins the ratio of the absorbance at 280 nm/450 nm is 9.3 – 11.3. In the case of PAMO, often a higher ratio of ~15 was found, which suggested that some of the total protein lacks of FAD. However, this loss of FAD could be restored by the incubation with the flavin cofactor in slight excess.



**Figure 68: UV-Vis spectrum of pure PAMO holoprotein with a 280/450 ratio of 14,7.**

### 5.3.2. Crystallization

Starting from the already known conditions, different grids were tested, modifying the concentration of the enzyme, precipitant, salt and cofactor. Unfortunately, in most of the wells the protein precipitated or, instead of symmetrical, yellow crystals only protein aggregations, precipitations or salt crystals were obtained.

### 5.4. Conclusion

PAMO has become one of the best characterized BVMOs and a very popular biocatalyst especially because of its thermal stability and versatility. The expression in *E. coli* and the subsequent two-step purification are very well established and efficient. After an anion exchange chromatography and a size exclusion chromatography PAMO could be obtained in good yields and high purity and was used for crystallization trials. One of the main problems consisted in the reproducibility of crystals. Thus the crystallization trials with PAMO will be repeated in the future and various modified conditions will be tested.



The aim of the present work was to establish a method for the expression and purification of this monooxygenase, to obtain pure protein which can consequently be used for crystallography and structural biology studies. Therefore, the already known purification strategy of PAMO served as a reference protocol. Adjustments were done to suit it to the specific requirements of BVMO24.

## 6.2. Experimental

The same materials and methods were used for the growth, the protein purification and the crystallization trials of BVMO24 as described for *moxY* and PAMO in chapter 4.2 and 5.2.

### 6.2.1. Cultivation and purification

#### Cloning, growth and expression

Cloning, transformation, cultivation of the cells and protein expression were performed using the protocols already adopted for PAMO [see chapter 5.2.1]. TB-medium was chosen since larger cell pellets and less contaminants could be obtained compared with the cultivation with LB-medium.

#### Cell harvest and breakage

Cell harvest and cell breakage by ultrasonication were done as described for PAMO. The only modification consisted in the working temperature since BVMO24 was not as thermostable. Therefore, the whole procedure was done at 4 °C. In order to improve the protein recovery, 2 mM PMSF was added to the breaking buffer and the ultrasonication time was increased to 3 min.

#### Protein purification

The purification of BVMO24 was performed in two steps and based on the same concept as described for PAMO, at 4 °C. First of all, an anion exchange chromatography was conducted. After the cell breakage, the supernatant was filtered and loaded on a Q Sepharose HiPrep 16/10 QFF column with a flow rate of 1.5 mL/min (maximum pressure over the column: 0.15 MPa). The column was washed with 50 mL breaking buffer [see Table 26] before the concentration of the elution buffer was gradually increased to 50 % with a gradient length of 140 mL. Fractions with a volume of 1.5 mL were collected. Finally, the concentration of the elution buffer was increased to 100 % to remove residual proteins. Various attempts with different buffers, for example 20 mM  $KP_i$ , pH 7.5 and 25 mM HEPES,

pH 7.5 were performed in order to increase the yield of BVMO24 and to modify the crystallization conditions.

SDS-PAGE analysis allowed the identification of BVMO24 containing fractions (MW = 62.1 kDa), which were afterwards concentrated to a volume <1 mL for desalting. Therefore, a HiTrap Desalting column from GE Healthcare was used as described in the instructions of the manufacturer. Separation according to the size of the molecules was facilitated by a 5 mL HiTrap Desalting column with a Sephadex G-25 Superfine matrix.

Furthermore, an additional, more sensitive anion exchange chromatography with a Resource Q column was tried after the first step. This column should allow high resolutions and high flow rates. After equilibration of the column, loading of the sample and washing with breaking buffer, the concentration of the elution buffer [see Table 26] was gradually increased to elute the desired protein. The whole purification procedure was performed as described in the manual from GE Healthcare for Resource™ Q Ion Exchange columns.

The last step consisted in a size exclusion chromatography with a Superdex 200 16/60 PrepGrade column as described for PAMO. Once again the breaking buffer listed in Table 26 was used.

In later attempts, an additional Mono Q GL5/50 column from GE Healthcare with a bed volume of 1 mL was used after the gel filtration in order to remove remaining protein contaminants. This anion exchange column functioned according to the same principle as the Q Sepharose column but should provide a higher sensitivity. All settings were made as recommended by the manufacturer.

An UV-Vis spectrum was recorded and after SDS-PAGE analysis the protein was concentrated for crystallization trials (> 20 mg/mL). The molar extinction coefficient was calculated with the Expasy webtool <http://web.expasy.org/protparam/> and used for the determination of the protein concentration ( $\epsilon = 100.333 \text{ mM}^{-1} \text{ cm}^{-1}$ ).

### 6.2.2. Crystallization

With the purified protein BVMO24 various crystallization plates were produced by the robot, including PEG screens and ammonium sulfate screens from Nextal, and stored at 4 °C as well as at 20 °C. The plates were prepared with and without 1.5 - 2.5 mM NADP<sup>+</sup> using a protein concentration of 10 and 15 mg/mL in breaking buffer and 0.3 mM FAD.

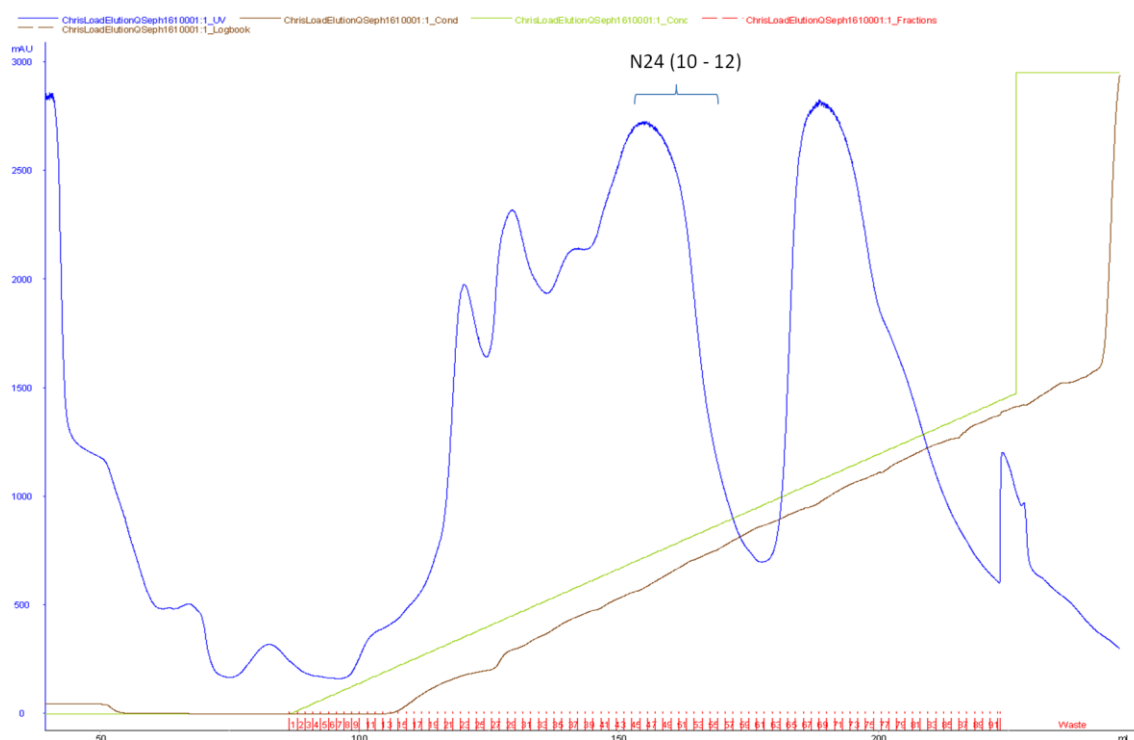
Additionally, Jena Start Screen #1 (from Jena Bioscience GmbH, Jena, Germany) and ammonium sulfate crystallization plates (from Nextal Biotech, sold by Qiagen) were done by hand using the sitting drop method and put at 4 °C. The Jena Start Screening # 1 allowed the

screening of a large variety of conditions with different concentrations and molecular weights of PEG, different co-precipitants, salts, buffers and pH-values. This kit provided ready to use solutions where only the protein, which was incubated with FAD and NADP<sup>+</sup> in advance, had to be added. Further plates were done by hand at 20 °C within the following conditions: 26 - 36 % PEG 2000, 50 - 300 mM ammonium sulfate and 8 - 14 mg/mL protein.

### 6.3. Results and discussion

#### 6.3.1. Purification

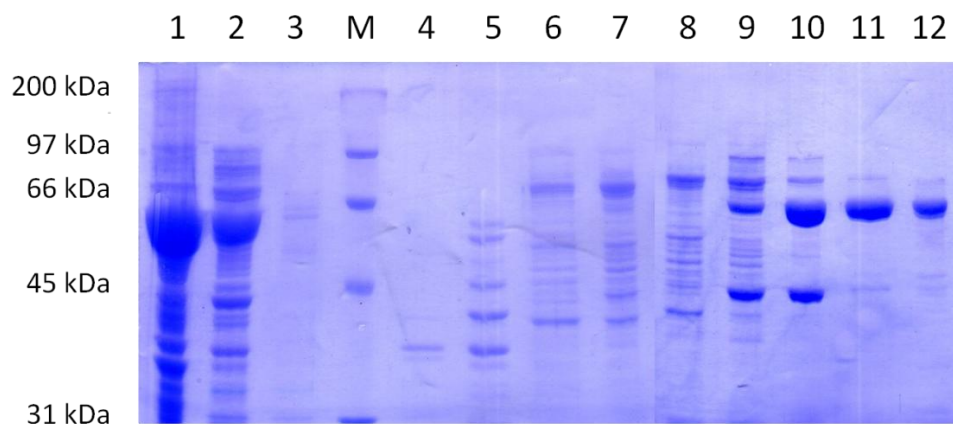
The chromatogram of the first purification step, an anion exchange chromatography with a QSepharose column, showed that several different proteins were eluted when the concentration of the elution buffer (green line) was gradually raised [Figure 70]. A better separation between the proteins (blue line) could be achieved by increasing the length of the gradient. Due to the ionic strength the conductivity (brown line) increased correspondingly to the concentration of the applied elution buffer.



**Figure 70: Chromatogram of the anion exchange purification of BVMO24. The UV absorbance of proteins at 280 nm is shown in blue and the gradient of the elution buffer in green. The brown line shows the increase in conductivity. The red numbers represent the fraction numbers. The numbers of the protein peak of BVMO24 correspond to the lanes on the SDS gel in Figure 71.**

Via SDS-PAGE the fractions containing BVMO24 with a molecular weight of 62 kDa were identified [see lane 10 - 12 in Figure 71], combined and concentrated to reach the appropriate volume for the subsequent gel filtration. The pellet was loaded on the first lane. It can be seen, that there was still a very large amount of BVMO24 left in the pellet, which indicated that the cell breakage needed to be improved in order to reach a better extraction. This may be achieved by prolonging the ultrasonication time. Alternatively, other ways of cell disruption could be tested. In the second lane of the gel the fraction that was loaded on the column can be seen, with BVMO24 being the most abundant protein. The flow through (see lane 3) seemed to contain hardly any proteins which indicated that the binding of the proteins to the column worked fine without overloading of the column.

The first peaks that were eluted when the gradient was started [see Figure 70] contained various other protein contaminants [see lane 4 - 9 in Figure 71]. From lane 10 on, the sample consisted mainly of BVMO24 but a contaminant with a molecular weight of ~44 kDa was found as well. Lane 11 and 12 contained almost exclusively BVMO24, in a quite good yield and purity.



**Figure 71: SDS gel of BVMO24 after anion exchange chromatography (1: pellet, 2: load, 3: flow through, M: marker, 4 - 9: elution of contaminants, 10: elution of BVMO24 plus contaminant at ~ 44 kDa, 11 + 12: elution of BVMO24).**

To further purify BVMO24 and to remove aggregates, a size exclusion chromatography was performed [see Figure 72]. The elution of BVMO24 did not occur as a single, sharp peak, which would be the ideal case, but with a shoulder at the beginning of the peak (blue line). The elution of the salt, which corresponded to the conductivity peak (brown line), partly overlapped with the elution of BVMO24.



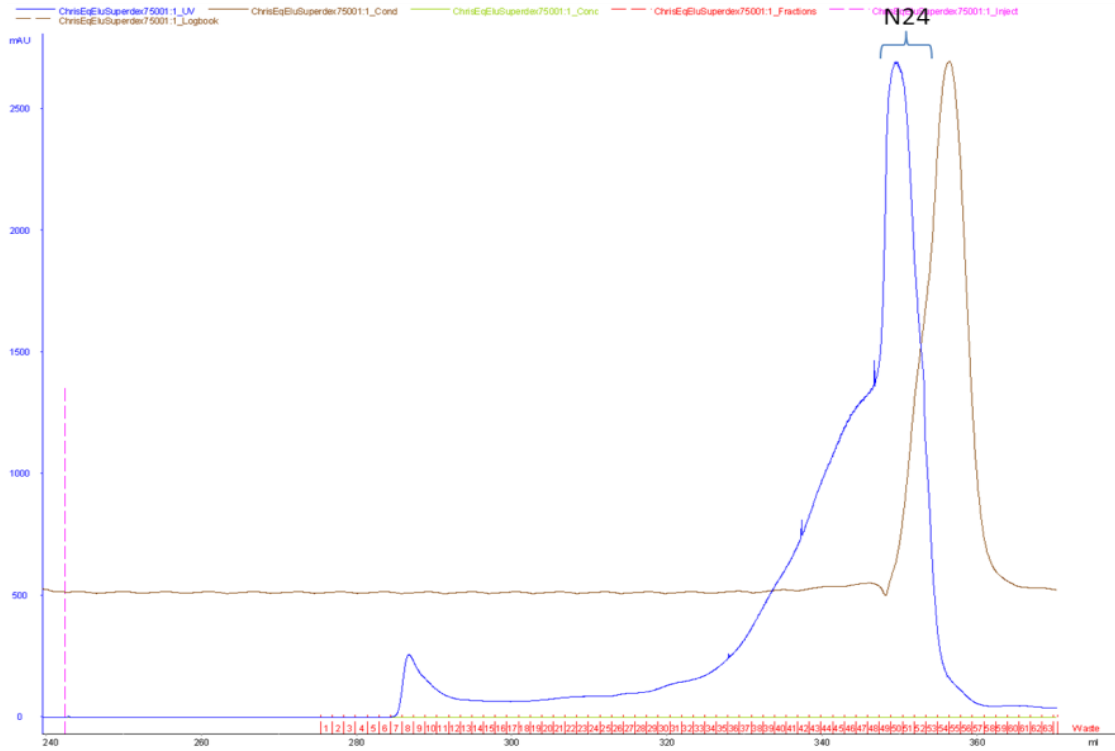


Figure 72: Chromatogram of the size exclusion chromatography of BVMO24. The UV absorbance of proteins at 280 nm is shown in blue and the conductivity, is depicted in brown. The red numbers represent the fraction numbers.

SDS-PAGE analysis showed that pure fractions of BVMO24 could be obtained (see lane 5 - 8 in Figure 73), which were used for crystallization trials. Nevertheless, many fractions contained also a lower molecular weight contaminant with ~44 kDa, which could hardly be separated. Especially the fractions of the first part of the peak contained still some other protein impurities (see lane 1 - 4).

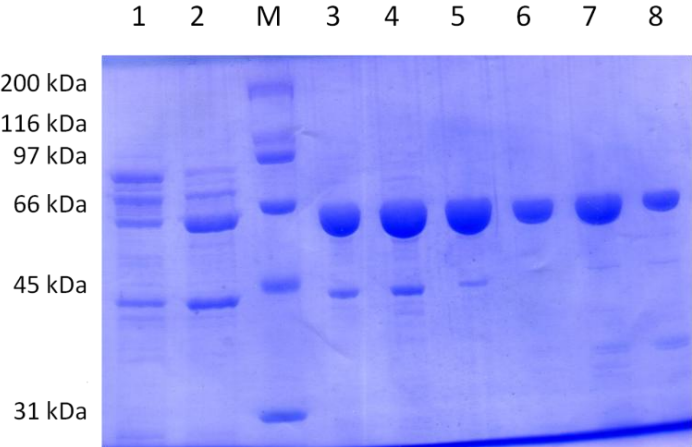
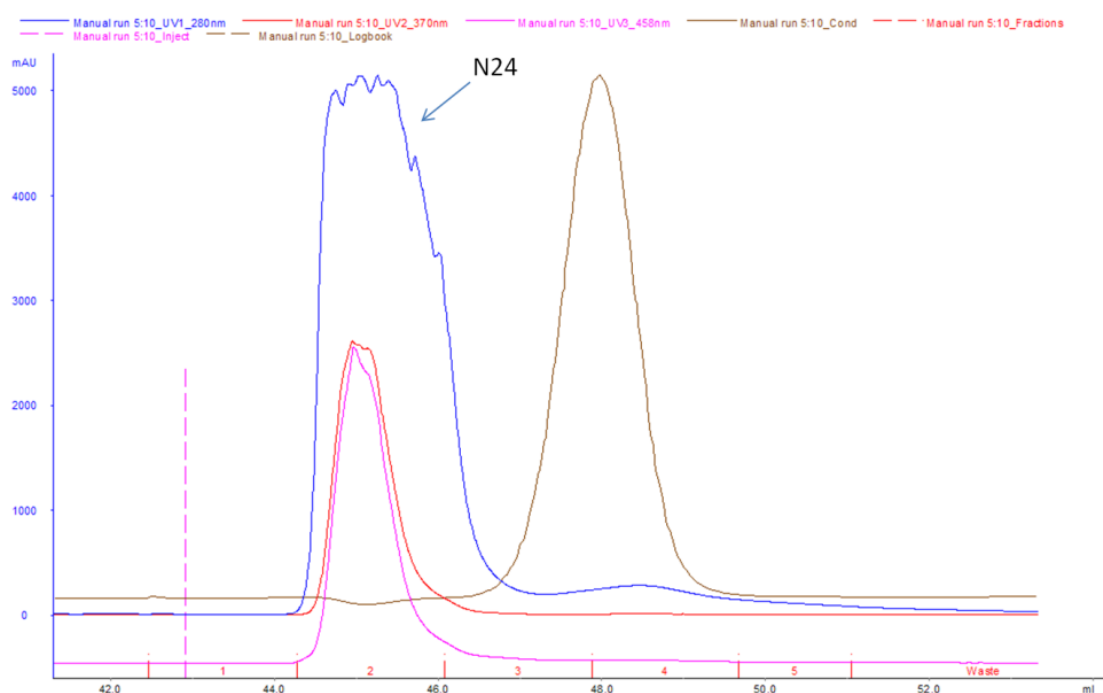


Figure 73: SDS gel of BVMO24 after the gel filtration (1 + 2: contaminants in the first part of the shoulder of the peak in Figure 72, M: marker, 3 + 4: shoulder of the peak, 5 – 8: elution of pure BVMO24).

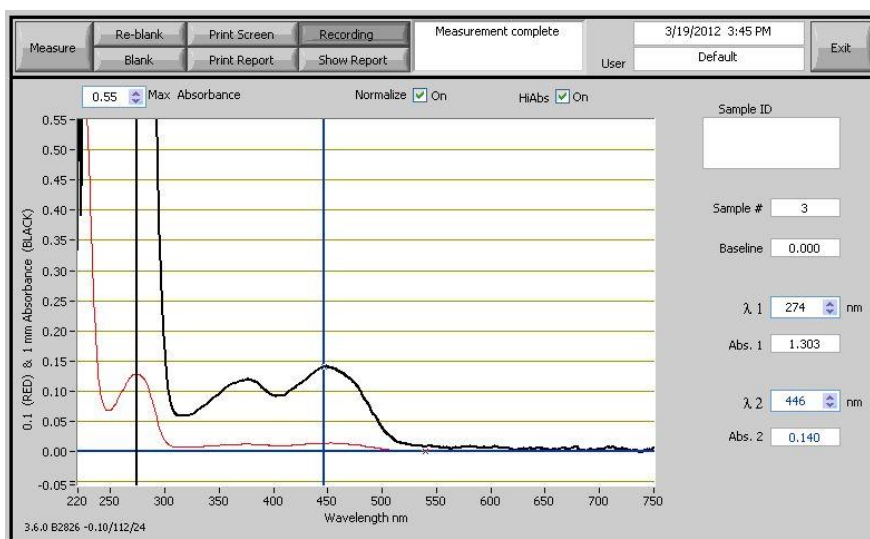
Additional attempts, that included a more sensitive anion exchange chromatography with a Resource Q or Mono Q column, were done. Therefore, the protein sample had to be desalted after the first purification step via a desalting column [Figure 74] or by dialysis. The chromatogram of the desalting showed the expected elution of the protein peak (blue line) as well as the subsequent elution of the salt (brown line). As expected, the flavin absorbance (pink and red lines) increased together with the protein peak (blue line of the absorbance at 280 nm), thus confirming that BVMO24 retained the flavin cofactor during purification.



**Figure 74:** Chromatogram of the desalting of BVMO24 with a HiTrap Desalting column. The UV absorbance of proteins at 280 nm is shown in blue, the flavin absorbance at 458 nm in pink and those at 370 nm in red, the conductivity is depicted in brown and the fraction numbers are shown in red.

The desalted protein sample was used for further chromatographic purifications but also with the Resource Q and Mono Q columns the contaminant at ~44 kDa could not be removed completely (chromatograms and gels are not shown).

UV-Vis spectra were recorded to determine the protein concentration and the flavin content in the samples. The measured absorbance ratio of 280 nm /450 nm of the pure, concentrated protein BVMO24 suggested a good retention of FAD [Figure 75].



**Figure 75: UV-Vis absorbance spectra of BVMO24 with a 280/450 absorbance ratio of 9.3.**

### 6.3.2. Crystallization

So far, no crystals could be obtained from the crystallization trials with BVMO24. However, various conditions could be identified in the PEG as well as ammonium sulfate screening plates, where the separation of two phases or the growth of small yellow, irregular shaped needles or aggregations suggested that they are promising starting points for the growth of analyzable crystals.

This may indicate that various crystalline forms might be possible with a different quality of nucleation. In future trials, several parameters like the concentration of the precipitant or salt, the pH or the temperature will have to be modified.

### 6.4. Conclusion

The successful expression of BVMO24 in *E. coli* and its purification via anion exchange and size exclusion chromatography could be achieved. Pure fractions of BVMO24 could be obtained and first crystallization screenings were started.

The main problem, which still has to be solved in prospective experiments, is that after elution many fractions contained apart from BVMO24 also a major contaminant with a lower molecular weight. To get rid of this contaminant, different other chromatographic steps or additives will have to be tested to further optimize the purification procedure. The lower thermal stability of BVMO24 compared to PAMO could be overcome by performing all working steps at 4 °C. In order to finally obtain crystals for elucidation of the three-

dimensional structure of this protein, many additional screenings and crystallization conditions will have to be tested in the future.

The interest in this protein can be explained by the results of previous studies which revealed the large variety of substrates it can accept, its high activity and good enantioselectivity. On top of that, BVMO24 shows preference for the opposite enantiomer than most other BVMOs. In contrast to the typical FxGxxxHxxxW(P/D) motif for BVMOs, it possesses a serine residue instead of the histidine in the core of this usually strictly conserved sequence and an asparagine at the end (36). As a result, this Baeyer-Villiger monooxygenase seems to be not only a very interesting and insightful enzyme but also a very versatile and valuable biocatalyst with a high potential.

## 7. References

1. **P.T. Anastas, J.C. Warner.** Green Chemistry: Theory and Practice. *Oxford Science Publications*. 1998.
2. **M. Poliakoff, J.M. Fitzpatrick, T.R. Farren, P.T. Anastas.** Green Chemistry: Science and Politics of Change. *Science*. 297, 2002, pp. 807-810.
3. **P. Tundo, P. Anastas, D. StC. Black, J. Breen, T. Collins, S. Memoli, J. Miyamoto, M. Polyakoff, W. Tumas.** Synthetic pathways and processes in green chemistry. Introductory overview. *Pure Appl. Chem*. 72, 2000, Vol. 7, pp. 1207–1228.
4. <http://www.rug.nl/oxygreen/index>. [Online] University of Groningen, Oxygreen Project. [Cited: 08 08, 2012.]
5. **D.E. Torres Pazmino, M. Winkler, A. Glieder, M.W. Fraaije.** Monooxygenases as biocatalysts: Classification, mechanistic aspects and biotechnological applications. *Journal of Biotechnology*. 146, 2010, pp. 9–24.
6. **A. Mattevi.** To be or not to be an oxidase challenging the oxygen reactivity of flavoenzymes. *Trends in Biochemical Sciences*. 31, 2006, Vol. 5, pp. 276-283.
7. **E. Gross, D.B. Kastner, C.A. Kaiser, D. Fass.** Structure of Ero1p, Source of Disulfide bonds for Oxidative Protein Folding in the Cell. *Cell*. 117, 2004, pp. 601–610.
8. **K. Faber.** *Biotransformations in Organic Chemistry, A Textbook, 6th Edition*. Berlin Heidelberg : Springer Verlag, 2011.
9. **M.D. Mihovilovic, B. Müller, P. Stanetty.** Monooxygenase Mediated Baeyer Villiger Oxidations. *Eur. J. Org. Chem*. 2002, pp. 3711-3730.
10. **J.H. Dawson, M. Sono.** Cytochrome P-450 and chloroperoxidase: thiolate-ligated heme enzymes. Spectroscopic determination of their active-site structures and mechanistic implications of thiolate ligation. *Chem. Rev*. 87, 1987, Vol. 5, pp. 1255–1276.
11. **W.J.H. van Berkel, N.M. Kamerbeek, M.W. Fraaije.** Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *Journal of Biotechnology*. 124, 2006, pp. 670–689.
12. **A. Baeyer, V. Villiger.** Einwirkung des Caro'schen Reagens auf Ketone. *Ber. Dtsch. Chem. Ges*. 1899, pp. 3625-3633.
13. **N.M. Kamerbeek, D.B. Janssen, W.J.H. van Berkel, M.W. Fraaije.** Baeyer Villiger Monooxygenases, an Emerging Family of Flavin Dependent Biocatalysts. *Adv. Synth. Catal.* , , 667. 345, 2003, pp. 667-678.
14. **G.J. ten Brink, I.W.C.E. Arends, R.A. Sheldon.** The Baeyer-Villiger Reaction: New Developments toward Greener Procedures. *Chem. Rev*. 104, 2004, pp. 4105-4123.
15. **G.E. Turfitt.** The Microbiological Degradation of Steroids. *Biochem. J*. 42, 1948, pp. 376-383.

16. **M.W. Fraaije, N.M. Kamerbeeka, W.J.H. van Berkelb, D.B. Janssen.** Identification of a Baeyer-Villiger monooxygenase sequence motif. *FEBS Letters*. 518, 2002, pp. 43-47.
17. **D.M.Ziegler.** Flavin-containing monooxygenases: enzymes adapted for multisubstrate specificity. *Trends Pharmacol. Sci.* 11, pp. 321-324.
18. **M. Stehr, H. Diekmann, L. Smau, O. Seth, S. Ghisla, M. Singh, P. Macheroux.** Studies with Lysine N6-Hydroxylase. Effect of a mutation in the assumed FAD binding site on coenzyme affinities and on lysine hydroxylating activity. *Biological Chemistry*. 380, 1999, pp. 47-54.
19. **D.E. Torres Pazmiño, M.W. Fraaije.** Discovery, Redesign and Applications of Baeyer-Villiger Monooxygenases. *Future Directions in Biocatalysis*. 2007, pp. 107-128.
20. **D.E. Torres Pazmiño, B.-J. Baas, D.B. Janssen, M.W. Fraaije.** Kinetic Mechanism of Phenylacetone Monooxygenase from *Thermobifida fusca*. *Biochemistry*. 47, 2008, pp. 4082–4093.
21. **D.E. Torres Pazmino, H.M. Dudek, M.W. Fraaije.** Baeyer–Villiger monooxygenases: recent advances and future challenges. *Current Opinion in Chemical Biology*. 14, 2010, pp. 138–144.
22. **V.Massey.** Activation of molecular oxygen by flavins and flavoproteins. *J. Biol. Chem.* 269, 1994, pp. 22459–22462.
23. **G. de Gonzalo, M.D. Mihovilovic, M.W. Fraaije.** Recent Developments in the Application of Baeyer–Villiger Monooxygenases as Biocatalysts. *ChemBioChem*. 11, 2010, pp. 2208 – 2231.
24. **J. Yu, P.-K. Chang, D. Bhatnagar, T. E. Cleveland.** Genes encoding cytochrome P450 and monooxygenase enzymes define one end of the aflatoxin pathway gene cluster in *Aspergillus parasiticus*. *Appl Microbiol Biotechnol*. 53, 2000, pp. 583-590.
25. **Y. Wen, H. Hatabayashi, H. Arai, H. K. Kitamoto, K. Yabe.** Function of the *cypX* and *moxY* Genes in Aflatoxin Biosynthesis in *Aspergillus parasiticus*. *Applied and Environmental Microbiology*. 71, 2005, Vol. 6, pp. 3192–3198.
26. **J. Yu, P.-K. Chang, K.C. Ehrlich, J.W. Cary, D. Bhatnagar, T.E. Cleveland, G.A. Payne, J.E. Linz, C.P. Woloshuk, J.W. Bennett.** Clustered Pathway Genes in Aflatoxin Biosynthesis. *Applied and Environmental Microbiology*. 70, 2004, Vol. 3, pp. 1253–1262.
27. **G.A. Payne, M.P. Brown.** Genetics and physiology of aflatoxin biosynthesis. *Annu. Rev. Phytopathol.* 36, 1998, pp. 329–362.
28. **F.S.Chu.** Mycotoxins: food contamination, mechanism, carcinogenic potential and preventive measures. *Mutation research*. 259, 1991, pp. 291-306.
29. **K. Yabe, N. Chihaya, S. Hamamatsu, E. Sakuno, T. Hamasaki, H. Nakajima, J.W. Bennett.** Enzymatic Conversion of Averufin to Hydroxyversicolorone and Elucidation of a Novel Metabolic Grid Involved in Aflatoxin Biosynthesis. *Applied and Environmental Microbiology*. 69, 2003, Vol. 1, pp. 66–73.
30. **M.W. Fraaije, J. Wu, D.P.H.M. Heuts, E.W. van Hellemond, J.H. Lutje Spelberg, D.B. Janssen.** Discovery of a thermostable Baeyer–Villiger monooxygenase by genome mining. *Appl Microbiol Biotechnol*. 66, 2005, pp. 393–400.

31. **C. Rodriguez, G. de Gonzalo, M.W. Fraaije, V. Gotor.** Enzymatic kinetic resolution of racemic ketones catalyzed by Baeyer–Villiger monooxygenases. *Tetrahedron: Asymmetry*. 18, 2007, pp. 1338–1344.
32. **G. de Gonzalo, D.E. Torres Pazmino, G. Ottolina, M.W. Fraaije, G. Carrea.** Oxidations catalyzed by phenylacetone monooxygenase from *Thermobifida fusca*. *Tetrahedron: Asymmetry*. 16, 2005, pp. 3077–3083.
33. **E. Malito, A. Alfieri, M.W. Fraaije, A. Mattevi.** Crystal structure of a Baeyer–Villiger monooxygenase. *PNAS*. 101, 2004, Vol. 36, pp. 13157–13162.
34. **R. Orru, H.M. Dudek, C. Martinoli, D.E. Torres Pazmino, A. Royant, M. Weik, M.W. Fraaije, A. Mattevi.** Snapshots of Enzymatic Baeyer-Villiger Catalysis; Oxygen activation and intermediate stabilization. *Journal of Biological Chemistry*. 286, 2011, Vol. 33, pp. 29284–29291.
35. **D.E. Torres Pazmiño, R. Snajdrova, D.V. Rial, M.D. Mihovilovic, M.W. Fraaije.** Altering the Substrate Specificity and Enantioselectivity of PAMO by Structure-inspired Enzyme Redesign. *Advanced Synthesis and Catalysis*. 349, 2007, pp. 1361-1368.
36. **A. Riebel, H.M. Dudek, G. de Gonzalo, P. Stepniak, L. Rychlewski, M.W. Fraaije.** Expanding the set of rhodococcal Baeyer–Villiger monooxygenases by high-throughput cloning, expression and substrate screening. *Appl Microbiol Biotechnol*. DOI 10.1007/s00253-011-3823-0, 2012.
37. **C. Szolkowy, L.D. Eltis, N.C. Bruce, G. Grogan.** Insights into sequence-activity relationships amongst Baeyer–Villiger monooxygenases as revealed by the intragenomic complement of enzymes from *Rhodococcus jostii* RHA1. *ChemBioChem*. 10, 2009, Vol. 7, pp. 1208–1217.
38. **K.N. Faber, W. Harder, G. Ab , M. Veenhuis.** Review: methylotrophic yeasts as factories for the production of foreign proteins. *Yeast*. 11, 1995, pp. 1331–1344.
39. **K. Ogata, H. Nishikawa, M. Ohsugi.** A yeast capable of utilizing methanol. *Agricultural and biological Chemistry*. 33, 1969, pp. 1519-1520.
40. **G.Wegner.** Emerging applications of the methylotrophic yeasts. *FEMS Microbiol. Rev.* 7, 1990, pp. 279-283.
41. **S. Macauley-Patrick, M.L. Fazenda, B. McNeil, L.M. Harvey.** Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*. 22, 2005, pp. 249-70.
42. **J.L. Cereghino, J.M. Cregg.** Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS microbiology reviews*. 24, 2000, pp. 45-66.
43. **J.M. Cregg, J.L. Cereghino, J. Shi, D.R. Higgins.** Recombinant Protein Expression in *Pichia pastoris*. *Molecular Biotechnology*. Molecular Biotechnology, 2000, pp. 23-52.
44. **G.Gellissen.** Heterologous protein production in methylotrophic yeast. *Appl Microbiol Biotechnol*. 54, 2000, pp. 741-750.
45. **R. Couderc, J. Baratti.** Oxidation of methanol by the yeast *Pichia pastoris*. Purification and properties of alcohol oxidase. *Agric. Biol. Chem.* 44, 1980, pp. 2279-2289.

46. **J.M. Cregg, K.R. Madden, K.J. Barringer, G.P. Thill, C.A. Stillman.** Functional Characterization of the Two Alcohol Oxidase Genes from the Yeast *Pichia Pastoris*. *Molecular and Cellular Biology*. 9, 1989, Vol. 3, pp. 1316-1323.
47. **M.Romanos.** Advances in the use of *Pichia pastoris* for high-level gene expression. *Current Opinion in Biotechnology*. 6, 1995, pp. 527-533.
48. **F.S. Hartner, A. Glieder.** Regulation of methanol utilisation pathway genes in yeasts. *Microbial Cell Factories*. 5, 2006, Vol. 39.
49. **M.A. Gleeson, P.E. Sudbery.** The methylotrophic yeasts. *Yeast*. 4, 1988, pp. 1-15.
50. **P. Koutz, G.R. Davis, C. Stillman, K. Barringer, J. Cregg, G. Thill.** Structural comparison of the *Pichia pastoris* alcohol oxidase genes. *Yeast*. 5, 1989, Vol. 3, pp. 167-177.
51. **W. Zhang, M. Inan, M.M. Meagher.** Fermentation Strategies for Recombinant Protein Expression in the Methylotrophic Yeast *Pichia pastoris*. *Biotechnol. Bioprocess Eng.* 5, 2000, pp. 275-287.
52. **A.A. Sibirny, V. I. Titorenko, M. V. Gonchar, V. M. Ubiyvovk, G. P. Ksheminskaya, O. P. Vitvitskaya.** Genetic control of methanol utilization in yeasts. *J. Basic. Microbiol.* 28, 1988, pp. 293-319.
53. **L.Näätäsaari, B. Mistlberger, C. Ruth, T. Hajek, F.S. Hartner.** Deletion of the *Pichia pastoris* KU70 Homologue Facilitates Platform Strain Generation for Gene Expression and Synthetic Biology. *PLoS ONE*. 7, 2012, Vol. 6, e39720.
54. **J.M. Cregg, K.J. Barringer, A.Y. Hessler, K.R. Madden.** *Pichia pastoris* as a Host System for Transformations. *Molecular and Cellular Biology*. 5, 1985, Vol. 12, pp. 3376-3385.
55. **J.F.Tschopp, P.F. Brust, J.M.Cregg, C.A.Stillman, T.R.Gingeras.** Expression of the lacZ gene from two methanol-regulated promoters in *Pichia pastoris*. *Nucleic Acids Research*. 15, 1987, Vol. 9, pp. 3859-3876.
56. **B.R. Glick, J.J. Pasternak, C.L. Patten.** *Molecular Biotechnology - Principles and Applications of Recombinant DNA*. 4. Washington, DC : ASM PRESS, 2010.
57. **J. Kurjan, I. Herskowitz.** Structure of a yeast pheromone gene (MF alpha): a putative alpha-factor precursor contains four tandem copies of mature alpha-factor. *Cell*. 30, 1982, pp. 933-943.
58. **A.J. Brake, J.P. Merryweather, D.G. Coit, U.A. Heberlein, F.R. Masiarz, G.T. Mullenbach, M.S. Urdea, P. Valenzuela, P.J. Barr.** Alpha-factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA*. 81, 1984, pp. 4642-4646.
59. **Invitrogen, Life Technologies Corporation.**  
[http://tools.invitrogen.com/content/sfs/manuals/pich\\_man.pdf](http://tools.invitrogen.com/content/sfs/manuals/pich_man.pdf). *Pichia Expression Kit, For Expression of Recombinant Proteins in Pichia pastoris*. [Online] [Zitat vom: 10. 09 2012.]
60. **T. Kjeldsen, A. Frost Pettersson, M. Hach.** Secretory expression and characterization of insulin in *Pichia pastoris*. *Biotechnol. Appl. Biochem.* 29, 1999, pp. 79–86.



61. **S. Kato, M. Ishibashi, D. Tasuda, H. Tokunaga, M. Tokunaga.** Efficient expression, purification and characterization of mouse salivary  $\alpha$ -amylase secreted from methylotrophic yeast, *Pichia pastoris*. *Yeast*. 18, 2001, pp. 643-655.
62. **R.J.M. Raemaekers, L. de Muro, J.A. Gatehouse, A.P. Fordham-Skelton.** Functional phytohaemagglutinin (PHA) and *Galanthus nivalis* agglutinin (GNA) expressed in *Pichia pastoris*: correct N-terminal processing and secretion of heterologous proteins expressed using the PHA-E signal peptide. *Eur. J. Biochem.* 65, 1999, pp. 394-403.
63. **M. Inan, D. Aryasomayajula, J. Sinha, M.M. Meagher.** Enhancement of Protein Secretion in *Pichia pastoris* by Overexpression of Protein Disulfide Isomerase. *Papers in Biochemical Engineering*. 5, 2005.
64. **H.F.Gilbert.** Protein disulfide isomerase. *Methods Enzymol.* 290, 1998, pp. 26–50.
65. **G.P.L. Cereghino, J.L. Cereghino, C. Ilgen, J.M. Cregg.** Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. *Current opinion in biotechnology*. 13, 2002, pp. 329-332.
66. **M. Thiry, D. Cingolani.** Optimizing scale-up fermentation processes. *TRENDS in Biotechnology*. 20, 2002, Vol. 3, pp. 103-105.
67. **M.M. Whittaker, J.W. Whittaker.** Expression of recombinant galactose oxidase by *Pichia pastoris*. *Protein Expr. Purif.* 20, 2000, pp. 105–111.
68. *Affinity Chromatography - Principles and Methods*. s.l. : Handbook from GE Healthcare. 18-1022-29.
69. *Ion Exchange Chromatography & Chromatofocusing - Principles and Methods*. s.l. : Handbook from Amersham Biosciences. 11-0004-21.
70. *Gel Filtration - Principles and Methods*. s.l. : Handbook from Amersham Biosciences. 18-1022-18.
71. **J.Drenth.** *Principles of Protein X-Ray Crystallography, Second Edition*. New York : Springer - Verlag, 1999.
72. **A.Messerschmidt.** *X-Ray Crystallography of Biomacromolecules, A practical Guide*. Weinheim, Germany : Wiley - VCH Verlag GmbH & Co. KGaA, 2007.
73. **A.Pennetta.** *Master's Thesis: Investigating the active site of a Baeyer - Villiger Monooxygenase*. Pavia : s.n., 2011.
74. <http://crystal.csiro.au/en/About/Vapour-Diffusion.aspx>. [Online] Collaborative Crystallisation Centre, CSIRO Materials Science and Engineering, Australia. [Cited: 08 31, 2012.]
75. **A. Mellitzer, C. Gustafsson, M. Welch, C. Ruth, R. Weis, A. Glieder.** High level protein expression in *Pichia pastoris* combining synthetic promoters and synthetic genes. [https://www.dna20.com/files/PDF/Poster\\_Pichia\\_2012.pdf](https://www.dna20.com/files/PDF/Poster_Pichia_2012.pdf), *Pichia 2012 Conference Poster*.
76. **UniProt Consortium.** <http://www.uniprot.org/uniprot/P07987>. *UniProtKB*. [Online] [Cited: 08 07, 2012.]

77. **A. Mellitzer, R. Weis, A. Glieder, K. Flicker.** Expression of lignocellulolytic enzymes in *Pichia pastoris*. [Hrsg.] <http://www.microbialcellfactories.com/content/11/1/61>. *Microbial Cell Factories*. 11, 2012, Bd. 61.
78. **Caprette, David R.** <http://www.ruf.rice.edu/~bioslabs/studies/sds-page/gellab2.html>. *Experimental Biosciences*. [Online] Rice University. [Cited: 08 12, 2012.]
79. **J.L. Cereghino, W.W. Wong, S. Xiong, W. Giang, L.T. Luong, J. Vu, S.S. Johnson, G.P. Lin-Cereghino.** Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast *Pichia pastoris*. *BioTechniques*. 38, 2005, pp. 44-48.
80. **R. Weis, R. Luiten, W. Skranc, H. Schwab, M. Wubbolts, A. Glieder.** Reliable high-throughput screening with *Pichia pastoris* by limiting yeast cell death phenomena. *FEMS yeast research* . 5, 2004, pp. 179-89.
81. <http://www.cbs.dtu.dk/services/SignalP/>. *Center for Biological Sequence Analysis, Technical University of Denmark*. [Online] [Cited: 08 16, 2012.]
82. <http://www.cbs.dtu.dk/services/TMHMM/>. *Center for Biological Sequence Analysis*. [Online] Technical University of Denmark. [Cited: 09 12, 2012.]
83. **Z. Li, W. Leung, A. Yon, J. Nguyen, V. C. Perez, J. Vu, W. Giang, L. T. Luong, T. Phan, K. A. Salazar, S. R. Gomez, C. Au, F. Xiang, D. W. Thomas, A. H. Franz, J. Lin-Cereghino, G. P. Lin-Cereghino.** Secretion and proteolysis of heterologous proteins fused to the *Escherichia coli* maltose binding protein in *Pichia pastoris*. *Protein Expression and Purification*. 72, 2010, pp. 113–124.
84. **J. Caubín, H. Martín, A. Roa, I. Cosano, M. Pozuelo, J. M. de la Fuente, J. M. Sánchez-Puelles, M. Molina, C. Nombela.** Choline-Binding Domain as a Novel Affinity Tag for Purification of Fusion Proteins Produced in *Pichia pastoris*. *Biotechnol Bioeng*. 74, 2001, pp. 164–171.
85. **J. McKinney, P. M. Knappskog, J. Pereira, T. Ekern, K. Toska, B. B. Kuitert, D. Levine, A. M. Gronenborn, A. Martinez, J. Haavik.** Expression and purification of human tryptophan hydroxylase from *Escherichia coli* and *Pichia pastoris*. *Protein Expression and Purification*. 33, 2004, pp. 185–194.
86. **R. B. Kapust, J. Tözsér, T. D. Copeland, D. S. Waugh.** The P1' specificity of tobacco etch virus protease. *Biochemical and Biophysical Research Communications*. 294, 2002, pp. 949–955.
87. [http://www.genebee.msu.su/services/rna2\\_reduced.html](http://www.genebee.msu.su/services/rna2_reduced.html). *GeneBee - Molecular Biology Server*. [Online] [Cited: 11 08, 2011.]
88. **C. Liu, H.L. Wang, Z.M. Cui, X.L. He, X.S. Wang, X.X. Zeng, H. Ma.** Optimization of extraction and isolation for 11S and 7S globulins of soybean seed storage protein. *Food Chem*. 102, 2007, S. 1310–1316.
89. **K. Deng, Y. Huang, Y. Hua.** Isolation of Glycinin (11S) from Lipid-Reduced Soybean Flour: Effect of Processing Conditions on Yields and Purity. *Molecules*. 17, 2012, pp. 2968-2979.
90. **A.P. Golovanov, G.M. Hautbergue, S.A. Wilson, L.-Y. Lian.** A Simple Method for Improving Protein Solubility and Long-Term Stability. *J. AM. CHEM. SOC*. 126, 2004, pp. 8933-8939.

91. **L. Hennig, E. Schäfer.** Protein Purification with C-Terminal Fusion of Maltose Binding Protein. *Protein Expression and Purification*. 14, 1998, pp. 367–370 .
92. **G. de Gonzalo, G. Ottolina, F. Zambianchi, M.W. Fraaije, G. Carrea.** Biocatalytic properties of Baeyer–Villiger monooxygenases in aqueous–organic media. *Journal of Molecular Catalysis B: Enzymatic*. 39, 2006, pp. 91–97.
93. <http://www.invitrogen.com/1/1/10731-pbad-myc-his-kit.html>. [Online] Life Technologies Corporation. [Cited: 08 09, 2012.]
94. [http://web.expasy.org/cgi-bin/compute\\_pi/pi\\_tool](http://web.expasy.org/cgi-bin/compute_pi/pi_tool). *Expasy Tools*. [Online] [Cited: 08 14, 2012.]
95. <http://bioinf.cs.ucl.ac.uk/psipred>. *UCL Department Of Computer Science*. [Online] [Cited: 08 23, 2012.]

## 8. Figures

Figure 1: The chemical structure of the reactive group, the isoalloxazine ring, of the flavin cofactors FAD and FMN (6). .....	2
Figure 2: Reaction mechanism of the Baeyer - Villiger oxidation (13). .....	4
Figure 3: The mechanism of BVMOs using flavin as cofactor for oxidations (13). .....	6
Figure 4: Mechanism of a BVMO (enzyme above) and a fused PTDH (below) for the regeneration of NADPH (23). .....	7
Figure 5: The aflatoxin pathway gene cluster with the old gene names on the right and the new names on the left side. ....	8
Figure 6: Alignment of various <i>moxY</i> protein sequences. ....	9
Figure 7: Alignment of the DNA sequences of <i>moxY</i> _long and <i>moxY</i> _short. The insertion of a T and C in the sequence of <i>moxY</i> _short is highlighted in blue. ....	10
Figure 8: Growth of an <i>Aspergillus</i> species on a peanut. ....	10
Figure 9: Parts of the metabolic pathway for the biosynthesis of aflatoxins in <i>Aspergillus</i> species, including the reaction catalyzed by <i>moxY</i> (25). ....	11
Figure 10: Kinetic resolution of racemic benzylketones, yielding optically active ( <i>S</i> )-benzylesters and ( <i>R</i> )-benzylketones. a) 50 mM Tris/HCl, PAMO, G6P/G6PDH (23). ....	13
Figure 11: The regiodivergent BV oxidation of substituted 1-indanone using a) HAPMO with 5% hexane as co-solvent or b) the PAMO mutant M446G with 5% methanol as co-solvent (23). ....	13
Figure 12: Crystal structure of PAMO in complex with NADP <sup>+</sup> (orange). ....	14
Figure 13: Detailed look at the active site of PAMO during the catalysis of a Baeyer Villiger reaction (34). ....	15
Figure 14: UV-Vis spectra of the native PAMO (1) and upon reduction with NADPH. ....	17
Figure 15: The methanol utilization pathway in <i>P. pastoris</i> (52). ....	21
Figure 16: The secretion pathway in eukaryotes via the endoplasmic reticulum and the Golgi apparatus (55). ....	23
Figure 17: The Kex2 and Ste13 cleavage sites for processing of the $\alpha$ - factor signal sequence (58). ....	24
Figure 18: Overview of various strategies available for protein purification (68). ....	26
Figure 19: Concept of affinity chromatography (67). ....	27
Figure 20: (A) The unit cell of a crystal with the unit cell parameters a, b, c and the angles $\alpha$ , $\beta$ , $\gamma$ . (B) A two - dimensional lattice with lattice planes in a distance d (here h = 2 and k = 1) (71), (72). ....	29
Figure 21: Sitting drop and hanging drop crystallization. The precipitant solution is shown in blue, the drop containing the protein and precipitant solutions is shown in pink (73). ....	30
Figure 22: Sequence annotation of the CBH2 gene of the exoglucanase 2 (74). ....	32
Figure 23: (A) Bands of the GeneRuler™ 1 kb DNA Ladder with the corresponding sizes given in basepairs (www.fermentas.com); (B) Bands of the PageRuler™ Prestained Protein Ladder #SM0671 with the corresponding protein sizes given in kDa (www.fermentas.com). ....	38
Figure 24: The vector pPpT4 with the $\alpha$ -factor signal sequence and the gene encoding for <i>moxY</i> long. ....	47
Figure 25: The construct for the secretion of <i>moxY</i> with the CBH2 signal sequence. The primers necessary for the preparation and cloning of the construct are indicated as arrows. ....	48
Figure 26: <i>In silico</i> signal peptide prediction for <i>moxY</i> using SignalP-4.0 (79). ....	53
Figure 27: Prediction of transmembrane helices in proteins with TMHMM Server, v 2.0, indicating that no transmembrane helix was found in <i>moxY</i> (80). ....	54
Figure 28: Fluorescence and phase contrast image of <i>P.p.</i> pPic ZB_citrine (3300) m.c. ....	55
Figure 29: Fluorescence and phase contrast image of <i>P.p.</i> mutS B1 citrine. ....	56
Figure 30: Fluorescence and phase contrast image of <i>P.p.</i> mutS B1_his_citrine_moxYlong. ....	56
Figure 31: Fluorescence and phase contrast image of <i>P.p.</i> mutS B1_moxYlong_citrine_his. ....	56
Figure 32: Fluorescence and phase contrast image of <i>P.p.</i> mutS B1_moxYshort_citrine_strep. ....	57
Figure 33: Fluorescence and phase contrast image of <i>P.p.</i> mutS <i>moxY</i> _IT_citrine_his. ....	57
Figure 34: SDS gel of the secretion of <i>moxY</i> with the $\alpha$ -factor signal sequence and the CBH2 signal sequence. ....	58

Figure 35: SDS gel of the secretion of moxY with $\alpha$ - factor signal sequence. ....	59
Figure 36: SDS gel of the secretion of moxY with CBH2. ....	60
Figure 37: Fluorescence measurement of <i>P. pastoris</i> $\alpha$ - factor signal sequence citrine clones in the culture including whole cells. ....	61
Figure 38: Fluorescence measurement of the secretion of citrine with the $\alpha$ - factor signal sequence in the supernatant. ....	62
Figure 39: Fluorescence measurement of the <i>P. pastoris</i> $\alpha$ - factor signal sequence His MBP citrine construct with MBP tag in the culture including whole cells. ....	63
Figure 40: Fluorescence measurement of the secretion of the $\alpha$ - factor signal sequence His MBP citrine construct with MBP-tag in the supernatant. ....	63
Figure 41: Fluorescence measurement of <i>P. pastoris</i> citrine constructs with the CBH2 signal sequence in the culture including whole cells. ....	64
Figure 42: Fluorescence measurement of the secretion of citrine with the CBH2 signal sequence in the supernatant. ....	64
Figure 43: Comparison of the secretion of citrine with the $\alpha$ -factor signal sequence (blue) and the CBH2 signal sequence (orange). ....	65
Figure 44: The first expression construct including moxY long with a TEV protease cleavage site, citrine as reporter tag and a His-tag. ....	66
Figure 45: SDS Page gel of the His- tagged moxY variants after protein purification. Cytosolic fractions of recombinant <i>Pichia</i> strains were used. ....	67
Figure 46: The construct of moxY long and the N-terminal fused MBP with linker. ....	71
Figure 47: GC content and GC peaks of the optimized MBP sequence. ....	82
Figure 48: The results of RNA secondary structure prediction of the optimized MBP sequence with GeneBee (86). Total free energy of the structure = -180.9 kcal/mol. ....	82
Figure 49: Fluorescence measurements of the MBP citrine constructs and the control samples citrine m.c., citrine s.c. and <i>P.p.</i> mutS. ....	83
Figure 50: SDS gels of the MBP-moxY constructs (109 kDa). Those clones that were selected for cultivation are marked in red. ....	84
Figure 51: Results of the dot blot with MBP-antibodies of the MBP moxY constructs with and without His-tag. ....	85
Figure 52: Membranes of the His-tag dot blot. (A) The membrane after application of the sample and protein staining with Ponceau solution. (B) Dot blot membrane after final staining with BCIP/NBT solution. ....	86
Figure 53: SDS gel of MBP purification pretests. ....	87
Figure 54: SDS gel after His-tag purification pretests, the red arrow indicated the band that should correspond to HIS MBP moxY. ....	89
Figure 55: Chromatogram of the MBP affinity chromatography of MBP moxY. ....	90
Figure 56: SDS gel after MBP affinity column. ....	90
Figure 57: Chromatogram of the anion exchange purification of moxY. ....	91
Figure 58: SDS gel after TEV cleavage and anion exchange column of moxY. ....	92
Figure 59: UV-Vis spectrum of moxY with a protein concentration of 1,6 mg/mL. ....	92
Figure 60: Chromatogram of the gel filtration of moxY after TEV-cleavage. ....	94
Figure 61: SDS gel of the gel filtration of moxY after TEV-cleavage. ....	95
Figure 62: Possible crystal resulting from the ammonium sulfate screening with 0.2 M KI and 2.2 M $(\text{NH}_4)_2\text{SO}_4$ . ....	98
Figure 63: The pBAD/Myc-His vector from Invitrogen for the expression of PAMO, the His-tag has been silenced (92). ....	101
Figure 64: Chromatogram of the anion exchange purification of PAMO. ....	104
Figure 65: SDS gel of PAMO after anion exchange chromatography. ....	104
Figure 66: Chromatogram of size exclusion chromatography of PAMO. ....	105
Figure 67: SDS gel of PAMO after size exclusion chromatography. ....	105
Figure 68: UV-Vis spectrum of pure PAMO holoprotein with a 280/450 ratio of 14,7. ....	106
Figure 69: Radial branching diagram comparing the residues of the active site of selected BVMOs. moxY, PAMO and BVMO24 are marked in red (36). ....	107

Figure 70: Chromatogram of the anion exchange purification of BVMO24..	110
Figure 71: SDS gel of BVMO24 after anion exchange chromatography	111
Figure 72: Chromatogram of the size exclusion chromatography of BVMO24.	112
Figure 73: SDS gel of BVMO24 after the gel filtration	112
Figure 74: Chromatogram of the desalting of BVMO24 with a HiTrap Desalting column.	113
Figure 75: UV-Vis absorbance spectra of BVMO24 with a 280/450 absorbance ratio of 9.3.	114
Figure 76: Results of the structure prediction of optimized MBP with the PSIPRED webtool. $\alpha$ - Helices are shown in pink and $\beta$ -strands are shown as yellow arrows (94).	133
Figure 77: pPp_B1_citrin_Spe/Not	137
Figure 78: pPpT4_alpha_S_moxY	137
Figure 79: pPpT4_alpha_S_citrin	138
Figure 80: pPpT4_Smi_CBH2_moxY	138
Figure 81: pPpT4_Smi_CBH2_citrin.	139
Figure 82: pPp_B1_MBP_moxY_with_Linkers	139
Figure 83: pPp_B1_MBP_moxY_no_Linkers	140
Figure 84: pPp_B1_moxY_MBP_no_Linkers	140
Figure 85: pPp_B1_HIS_MBP_moxY_with_Linkers	141
Figure 86: pPpT4_alpha_S_HIS_MBP_moxY_with_Linkers	141
Figure 87: pPp_B1_MBP_citrin_with_Linkers	142
Figure 88: pPp_B1_MBP_citrin_no_Linkers	142
Figure 89: pPp_B1_citrin_MBP_no_Linkers	143
Figure 90: pPpT4_alpha_S_HIS_MBP_citrin_with_Linkers	143

## 9. Tables

Table 1: Substrate conversion of various BVMOs from <i>R. jostii</i> (given as percentage %); the enantioselectivities are given as enantiomeric excess in % (36). .....	18
Table 2: Different <i>P. pastoris</i> strains and their genotype as well as phenotype. ....	22
Table 3: Composition of BMD 1 %, BMM 2 and BMM 10 media (V = 1 L). ....	36
Table 4: Composition of the dephosphorylation reaction mixture with SAP. ....	40
Table 5: A list of all primers including their number, name, sequence and properties according to the data sheets. ....	41
Table 6: PCR reaction mixture for the preparation of the B1 citrine construct. ....	45
Table 7: Temperature program for the PCR to prepare the B1 citrine construct. ....	46
Table 8: Composition of the reaction mixtures for the restriction cuts of the T4 $\alpha$ S and the T4 Smi vector. ....	49
Table 9: PCR reactions for the secretion constructs of moxY with the $\alpha$ -factor signal sequence and CBH2 signal peptide. ....	49
Table 10: PCR reactions for the secretion constructs of citrine with the $\alpha$ -factor signal sequence and CBH2 signal peptide. ....	49
Table 11: Temperature program for the PCR reactions to prepare the secretion constructs of moxY and citrine with the $\alpha$ -factor signal sequence and the CBH2 signal peptide. ....	50
Table 12: Reaction mixtures for the oePCRs to prepare the secretion constructs of moxY and citrine with the CBH2 signal peptide, using the products of the PCRs in Table 9 and Table 10. ....	50
Table 13: Temperature program for the oePCR reactions to prepare the secretion constructs of moxY and citrine with the CBH2 signal peptide. ....	50
Table 14: Composition of the mastermix, added for the amplification of the whole constructs for secretion of moxY and citrine with the CBH2 signal peptide. ....	51
Table 15: Media, buffers and solutions for the fermentation, cell break up and purification moxY. ....	68
Table 16: General composition of the PCR reactions for preparation of the MBP constructs. ....	71
Table 17: The templates and primers used for the PCR reactions to prepare the constructs MBP (with linker) moxY, MBP (no linker) moxY and moxY MBP (no linker). ....	72
Table 18: The templates and primers used for the PCR reactions to prepare the constructs His MBP (with linker) moxY, $\alpha$ His MBP (with linker) moxY, MBP (with linker) citrine and MBP (no linker) citrine. ....	72
Table 19: The templates and primers used for the PCR reactions to prepare the constructs citrine MBP (no linker) and $\alpha$ His MBP (with linker) citrine. ....	72
Table 20: Temperature program for the PCR reactions to prepare the constructs with MBP. ....	72
Table 21: General composition of the oePCRs to prepare the MBP moxY constructs. ....	73
Table 22: General composition of the oePCRs to prepare the MBP citrine constructs. ....	73
Table 23: The DNA-fragments which were connected via oePCR to prepare the MBP constructs. ....	73
Table 24: General composition of the mastermix, added to the oePCRs for the amplification of the whole MBP constructs. ....	73
Table 25: List of the forward and reverse primers added to the mastermix described in Table 24 for the amplification of the whole MBP constructs via oePCR. ....	74
Table 26: Media and recipes for the cultivation and purification of PAMO. ....	100
Table 27: Settings of ultrasonication. ....	102
Table 28: Restriction sites which were avoided within the optimized MBP gene sequence. ....	130
Table 29: Motifs which were avoided at the optimization of the MBP gene because of polyadenylation or pre - termination. ....	130
Table 30: Codon distribution of the optimized MBP gene using the high methanol codon usage table for <i>Pichia pastoris</i> . ....	131
Table 31: Strains for the strain collection with their host, strain collection number (CC Nr.) and vector. ....	144
Table 32: Cultivation conditions for the <i>E. coli</i> and <i>P. pastoris</i> strains of the strain collection. ....	145

## 10. Appendix

### A. Sequences of CBH2

#### CBH2 - protein sequence (76)

The sequence of exoglucanase 2 (length: 471 bp, molecular mass: 49,6 kDa, P07987) with the signal peptide at the beginning (blue), the CBM1 binding domain (green) and the linker (yellow) (76):

```
      10      20      30      40      50      60
MIVGILTTLA TLATLAASVP LEERQACSSV WGQCGQNWS GPTCCASGST CVYSNDYYSQ
      70      80      90     100     110     120
CIPGAASSSS STRAASTSR VSPTTSRSSS ATPPPGSTTT RVPPVGS GTA TYSGNPFVGV
      130     140     150     160     170     180
TPWANAYYAS EVSSLAIPSL TGAMATAAAA VAKVPSFMWL DTLDKTPLME QTLADIRTAN
      190     200     210     220     230     240
KNGGNYAGQF VVYDLPRDC AALASNGEYS IADGGVAKYK NYIDTIRQIV VEYSDIRTL
      250     260     270     280     290     300
VIEPDSLNL VTNLGTPKCA NAQSAYLECI NYAVTQLNLP NVAMYLDAGH AGWLGWPANQ
      310     320     330     340     350     360
DPAAQLFANV YKNASSPRAL RGLATNVANY NGWNITSPPS YTQGNVAVNE KLYIHAIGPL
      370     380     390     400     410     420
LANHGWSNAF FITDQGRSGK QPTGQQQWGD WCNVIGTGFG IRPSANTGDS LLDSFVWVKP
      430     440     450     460     470
GGECDGTSDS SAPRFD SHCA LPDALQPAPQ AGAWFQAYFV QLLTNANPSF L
```

#### CBH2 - DNA sequence

The first part of the CBH2 gene, which was used for the secretion of moxY and citrine is depicted in blue, containing the signal peptide at the beginning, the CBM1 binding domain and the linker.

```
>gi|121855|sp|P07987.1|GUX2_HYPJE RecName: Full=Exoglucanase 2; AltName:
Full=1,4-beta-cellobiohydrolase; AltName: Full=Exocellobiohydrolase II;
Short=CBHII; AltName: Full=Exoglucanase II; Flags: Precursor
```

```
atgattgtgggaatgaccacgctcgtacacctggctacgtggccgcatcagtagcttggaggaaagacagggatgctcctcagtaggggt
caatgtggaggacagaactggagtgacacctggtgtgctcaggatcaacatgcgtctattctaagtagtactactctcaatgtctgctggag
ctgcaagctctcatcgtccactcgtgctgtcaacaactcagctgttccctacgacctcgcgctctagttcagcaactcctcctcctggtcaac
tactaccggagtcaccaccgctcgggtcaggtagcgaacatattctggaatcctttgtggtgttacgccctgggccaatgcgtattatgcctcaga
agtatctcattagctataccaagcctgactggagctatggcaacagcagctgcagcagttgctaaggctcctcttcatgtggtggatacgttg
acaagacccttggatggagcagaccctggcagatccgtagctgtaacaaaaatggaggaactatgctggtcaattgtggttatgattgcc
tgacagagactgtgcagcttggtagtaatggagagtattcaatagcggatggcgggtgtgtaagtaacaaaactatattgacacgatcagac
aatcgttgggaatattcagatatacaggacttggtagttatcgaacctgattcactgtaattgtgtagtaatactggggaccctaagtgtcca
atgccaatcagcactgtagtgataaactatgctgtcacacagttgaactgcctaattgtgcatgtatctggacgcaggccacgctggtt
ggctaggatggcctgtaatacaggaccctgctgcacaactttcgttaattgtgtacaaaaatgcctcgtcccctagagcattgagaggactggct
actaatgtagcaactataatgggtggaatcacctcaccacctcttatacacaaggaatgctgtatacaacgaaaagctgtatattcagca
ataggccctggttagtaatacagtaggtcaaatgccttttactgactgaccaaggacgtagtggaagcaaccaacagggcaacaacaat
ggggtgattggtgtaattgattggtacaggattggtatccgtcattctgtaatacaggagacagtagtctcctggatagcttgttgggtcaagcctg
gaggagagtgtagcgggactgtagtcatctgcaccagctttagctcccattgtgcgctacctgatgccttcagccagctccacaagccggt
gcatggttcaagcctattctgtaaatgttgacaaatgcaaatccctcattttgtaa
```



## B. Sequences of citrine

### Citrine - protein sequence

MASKGEELFTGVVPILEVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPT  
LVTTFGYGLMCFARYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEEDTL  
VNRIELKGIDFKEDGNILGHKLEYNYNSHNVIITADKQKNGIKANFKIRHNIEDGSVQLA  
DHYQQNTPIGDGPVLLPDNHLYSYQSKLSKDPNEKRDMVLLLEFVTAAGITHGMDELYK

### Citrine - DNA sequence

atggctagcaaaggagaagaactttcactggagtgcccaattctgtgaattagatggatgtaatgggcacaaatttctgcagtggaga  
gggtgaaggtgatgctacatacggaaagcttaccctaaattattgactactggaaaactacgttccatggccaacactgtcactacttctg  
ggtatggttgatgcttctgctggtatccggatcatatgaaacggcatgacttttcaagagtgccatgccgaagggtatgtacaggaacgcact  
atatcttcaaagatgacgggaactacaagacgctgctgaagtcaagttgaaggtgataccctgtaatcgatcgagtaaaaggattgatt  
taaagaagatggaacattctcggacacaaactgagtacaactataactcacacaatgtatacatcacggcagacaaacaaaagaatgg  
aatcaaagtaactcaaaatcgccacaacattgaagatggttccgtcaactagcagaccattatcaacaaaatactccaattggcgatggc  
cctgtcctttaccagacaaccattacctgtcgtaccaatctaagcttgcgaagatcccaacgaaaagcgtgaccacatggtccttctgagttgt  
aactgctgctgggattacacatggcatggatgaattgtacaagtaa

## C. Sequences of moxY long

### moxY long - protein sequence

MDPANRPLRVVTIGTGISGILMAYQIQKQCPNVEHVLYEKNADVGGTWLENRYPMAGCDV  
PSHAYTYPFAPNPDWPRYFYSYAPDIWNYLDRVCKVFGRLRRYMFVHTEVVGCYWNEDRGEW  
TVRLRQHASGSEPREFEDHCHVLVHASGVFNPNQWPQIPGLHDFRQGRVLHTARWPDYQ  
ESQWKSDRVAVIGSGASSIQTVPGMQPTVKHLDVVFVRTGVWFGVLAGNTGSQTKEYSPT  
RDEFRRNPAALVAHAKAIEDQVNGMWGAFYTGSKGQAMGSAFFRQRTANLIKDERLREGL  
DPPFAFGCRRITPGDPYMEAIQKENVHVHFTPVVSCTEKGVVGGDVERQVDTVVCATGF  
DASYRPRFPIVGRDGMDLREKWKECPNSYLGLAVPEMPNFFTFIGPTWPIQNGSVIGPLQ  
AVSKYVVQWIKKAQENLRSFVPRQDRTDQFNDHVQEWVKHTVWKDNCRSCTFLLSTLTA  
EFTENHLGYKNNETGRVNAIWPGSSSLHYQQVIDQPRYEDFDIRS FHENPWACLGMGWTIQ  
DRKGPKEADVSPHLGLQEIDPKWVESIKNENGSNGVSDSNR

Theoretical pI/Mw: 6.36 / 66192.33 Da (94).

### moxY long - DNA sequence

atggaccggccaaccgcccgtgctgggtggatgacctcggcagggcatctcgggatactgatggcataccagatccagaagcaatgcc  
ctaattgctgagcagctttgtatgagaaaaacgcgacgtgggtgactctggttagaaaaccgttccccatggccggtgctgatgcccag  
ccatgcctacacacctacccttctccgaatccagactgccccgatactctcctatgcgccgatctggaattacctcgaccgggtatgcaa  
agcttctgctcgcggttacatggtgtttcacacggaggtcgtggctgctactggaacgaagaccgcgagaaatggactgtccgactcgcac  
agcacgctagcggcagtgagccccgagagttgaagaccattgtcacgtcttggttcacgctcgggggtattcaataatcctcagtgccgca  
aatccccgcctcatgaccggtccaaggcctgtgctcctataccgcacgatggcccagcactaccaagagtcaaatggaagagtgata  
gggtggcagtcattggctccggggcatcatcctcaaacggtgcccggcatgcagcccaggtgaagcatctcgatgcttctgctggcagcg  
gctgatggtcggagtgctagcgggcaacactggcagccagaccaaggaataactcgcccaccgagcgcgacgaattccggcgaatccag

cagcgctgtggcccacgcaaaggccatcgaggaccaggtcaacggcatgtgggggcttttacactggatccaaagggcaagcgatgg  
gatctgcggtttccgccagcgcaccgccaactgatcaaggatgagcggctacgagaagggctgacccgccccttgcggttggtcgtcgca  
tccccgggggatccctacatggaggcgattcagaaggagaacgtccatgtacactttacccccgtgtaagctgtacggagaagggcgtg  
gtggcggtgacggagtagagcgcagggtgcacacggtagtatgtccacgggcttcgatgatcctatcgccacgctccccattgtgggc  
cgagacggcatggactgaggagaagtggaaggagtgcccaattcgtatctaggcctggccgtgccgagatgcccaacttctcacattc  
attggccgacatggcctatccaaaatggcagtgctattgggcccgtgcaggccgtctcgaagtagctgggtgagtgataagaagggccag  
aatgagaatctccgtagcttgcgcccgcagggaccgcacggatcaattcaacgatcatgtccaggagtggtgagacacacgggtggaa  
agataactgccgaagctgtacgttcttctccacgctcacagctgaattactgagaaccatttaggttacaagaacaacgagacgggtcgg  
gtcaatgccatctggcccgggttcttgcctgactaccagcaagtcattgaccagcctcgtacgaagactcgacatccgctccttccatgagaat  
ccctggcctgtctggcagtggtggacaatccaggaccgcaaagggcccgaaggaagcagatgtcagtcgcatctgggcttacaggaga  
ttgaccgaaatggtgggaatcaataaaaatgagaatggttcaatggagtctgacagtaaccgtag

#### D. Sequences of MBP and MBP gene optimization

##### MBP - protein sequence *Pichia* optimized

MKIEEGKLVIIWINGDKGYNGLAEVGGKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPD I  
IFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNK  
DLLPNPPKTWEEI PALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGYDIK  
DVGVDNAGAKAGLTFVLVLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSK  
VNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL  
GAVALKSYEEELAKDPRIAATMENAQKGEIMPNI PQMSAFWYAVRTAVINAASGRQTVDE  
ALKDAQTLINGDGAGLEVL FQGPENLYFQA

Linker

TEV Cleavage site

##### MBP - DNA sequence *Pichia* optimized

ATGAAGATCGAGGAAGGTAAGTTAGTGATTTGGATCAACGGAGACAAAGGTTACAATGGTTTGGCTGAAGTTGGTAAGA  
AATTCGAAAAGGACTGGTATCAAGGTTACCGTCGAGCACCTGACAAGTTGGAGGAAAAGTTCCACAAGTTGCTGCA  
ACCGTGATGGACCAGATATTATCTTTGGGCCATGACAGATTCGGTGGTTACGCACAGTCCGGTCTTTGGCTGAGATT  
ACTCCAGATAAAGCATTCCAAGACAAGTTGTATCCTTTCACTGGGATGCTGTTCTGTTACAACGGTAAGCTGATTGCCTATC  
CTATTGCTGTGAAGCCTTATCTTTGATCTACAACAAGGACTTGTGGCAAATCCACCTAAGACCTGGGAGGAAATTCCTGC  
ACTTGACAAGGAGTTGAAAGCCAAGGGAAAGTCTGCTGTGATGTTCAACCTTCAAGAACCATACTTTACTTGGCCATTGAT  
TGCCGCTGACGGAGGTTATGCTTTCAAGTACGAGAACGGTAAATACGATATCAAGGACGTGGGTGTCGATAATGCTGGTG  
CTAAGGCTGGATTGACCTTCTTGTGACTTAATCAAGAACAAGCACATGAACGCTGACACTGACTACTCTATTGCTGAAGC  
TGCATTCAACAAAGGTGAAACCGCATGACTATTAACGGTCTTGGGCTTGGTCCAACATTGACACTTCTAAGGTCAACTA  
CGGTGTTACTGTCCTTCCAACCTTCAAGGGACAACCTTCTAAGCCATTCGTTGGTGTCTTATCAGCTGGAATCAATGCCGCT  
TCTCAAACAAAGAAGTTCGAAAGGAGTTCTGGAAAACACTTGTGCTGACCGACGAGGTTTTGGAGGCTGTTAACAAGGA  
CAAGCCTCTGGGAGCTGTTGATTGAAGTCATACGAGGAAGAGTTGGCTAAAGATCCAAGAATTGCCGCTACTATGGAGA  
ACGCACAAAAGGGAGAAATCATGCCAAACATCCCTCAAATGTCGCTTTCTGGTACGCCGTTAGAAGTCTGTGATTAACG  
CCGATCTGGTAGACAGACCGTCGACGAGGCTTTGAAGGATGCACAAACCTTGATCAACGGTGACGGTGCTGGACTTGAA  
GTTTTGTTCCAGGGTCTGAGAACCTTACTTTCAAGCT

**MBP (with linker) moxY construct:**

MKIEEGKLVIIWINGDKGYNGLAEVGGKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPD I  
 IFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNK  
 DLLPNPKTWEEIIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGGYDIK  
 DVGVDNAGAKAGLTFVLVDLIKKNHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSK  
 VNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL  
 GAVALKSYEEELAKDPRIAATMENAQKGEIMPNI PQMSAFWYAVRTAVINAASGRQTVDE  
 ALKDAQTLINGDGAGLEVLFGQGPENLYFQADPANRPLRVVTIGTGISGILMAYQIQKQCP  
 NVEHVLYEKNADVGGTWTLENRYPMAGCDVPSHAYTYPFAPNPDWPRYFSYAPDIWNYLDR  
 VCKVFGLLRRYMFVHTEVVGWYWNEDRGEWTVRLRQHASGSEPREFEDHCHVLVHASGVFN  
 NPQWPQIPGLHDFQGRVLTARWPDYQESQWKS DRVAVIGSGASSIQTVPGMQPTVKH  
 LDVVFVRTGVWFGVLAGNTGSQTKESPTERDEFRRNPAALVAHAKAIEDQVNGMWGAFYT  
 GSKGQAMGSAFFRQRTANLIKDERLREGLDPPFAFGCRRITPGDPYMEAIQKENVHVHFT  
 PVVSCTEKGVVGGDGVVERQVDTVVCATGFDASYPFRFPVGRDGM DLREKWKECPN SYLG  
 LAVPEMPNFFTFIGPTWPIQNGSVIGPLQAVSKYVVQWIKKAQENENLRSFVPRQDRTDQF  
 NDHVQEWVKHTVWKDNCRSCFTLLSTLTAEFTENHLGYKNNETGRVNAIWPGSSLHYQQV  
 IDQPRYEDFDIRSFHENPWACLGMGWTIQDRKGPKEADVSPHLGLQEIDPKWWESIKNEN  
 GSNVSDSNR

Calculation of the theoretical pI and MW using with the corresponding webtool on expasy.org (94):

Theoretical pI/Mw: 5.70 / 108830.74 Da (with linker for TEV-cleavage)

Theoretical pI/Mw: 5.84 / 106383.01 Da (without linker for TEV-cleavage)

**Table 28: Restriction sites which were avoided within the optimized MBP gene sequence.**

<i>Ascl</i>	<i>NcoI</i>	<i>SphI</i>
<i>AvrII</i>	<i>NotI</i>	<i>SwaI/SmlI</i>
<i>BamHI</i>	<i>PstI</i>	<i>XbaI</i>
<i>BglII</i>	<i>SacI</i>	<i>XhoI</i>
<i>EcoRI</i>	<i>SpeI</i>	

**Table 29: Motifs which were avoided in the optimization of the MBP gene because of polyadenylation or pre - termination.**

WWWWW	W = weak base (A,T) R = purine (A,G) Y = pyrimidine (T,C)
YGTGTTY	
TTAAGAAC	
TATGTT	
TAYRTA	

**Table 30: Codon distribution of the optimized MBP gene using the high methanol codon usage table for *Pichia pastoris*.**

TARGET				TARGET number	ACTUAL					
AA	Codon	%	%		%	aa	codon	number	/1000	fraction
A-Ala	GCT	54,00	54,67	25	56	ALA	GCT	26	66,66	0,56
	GCC	23,00	23,67	11	21	ALA	GCC	10	25,64	0,21
	GCA	21,00	21,67	10	21	ALA	GCA	10	25,64	0,21
	GCG	2,00			0	ALA	GCG	0	0,00	0
C-Cys	TGT	70,00	70,00	0	0	CYS	TGT	0	0,00	0
	TGC	30,00	30,00	0	0	CYS	TGC	0	0,00	0
D-Asp	GAC	66,50	66,50	17	68	ASP	GAC	17	43,58	0,68
	GAT	33,50	33,50	8	32	ASP	GAT	8	20,51	0,32
E-Glu	GAG	53,00	53,00	15	51	GLU	GAG	15	38,46	0,51
	GAA	47,00	47,00	14	48	GLU	GAA	14	35,89	0,48
F-Phe	TTC	69,00	69,00	12	76	PHE	TTC	13	33,33	0,76
	TTT	31,00	31,00	5	23	PHE	TTT	4	10,25	0,23
G-Gly	GGT	64,25	69,63	23	69	GLY	GGT	23	58,97	0,69
	GGA	25,25	30,63	10	30	GLY	GGA	10	25,64	0,3
	GGC	8,50		0	0	GLY	GGC	0	0,00	0
	GGG	2,25		0	0	GLY	GGG	0	0,00	0
H-His	CAC	80,00	80,00	2	66	HIS	CAC	2	5,12	0,66
	CAT	20,00	20,00	1	33	HIS	CAT	1	2,56	0,33
I-Ile	ATT	52,00	54,15	12	52	ILE	ATT	12	30,76	0,52
	ATC	43,70	45,85	11	47	ILE	ATC	11	28,20	0,47
	ATA	4,30		0	0	ILE	ATA	0	0,00	0
K-Lys	AAG	74,50	74,50	27	77	LYS	AAG	28	71,79	0,77
	AAA	25,50	25,50	9	22	LYS	AAA	8	20,51	0,22
L-Leu	TTG	43,70	46,25	16	50	LEU	TTG	17	43,58	0,5
	CTT	21,20	23,75	8	23	LEU	CTT	8	20,51	0,23
	CTG	13,50	16,05	5	14	LEU	CTG	5	12,82	0,14
	TTA	11,50	14,05	5	11	LEU	TTA	4	10,25	0,11
	CTC	7,70			0	LEU	CTC	0	0,00	0
	CTA	2,50		0	0	LEU	CTA	0	0,00	0
M-Met	ATG	100,00	100,00	7	100	MET	ATG	7	17,94	1
N-Asn	AAC	78,50	78,50	18	82	ASN	AAC	19	48,71	0,82
	AAT	21,50	21,50	5	17	ASN	AAT	4	10,25	0,17
P-Pro	CCA	52,00	55,00	12	54	PRO	CCA	12	30,76	0,54
	CCT	42,30	45,30	10	45	PRO	CCT	10	25,64	0,45

	CCC	6,00		0	0	PRO	CCC	0	0,00	0
	CCG	0,00			0	PRO	CCG	0	0,00	0
Q-Gln	CAA	66,50	66,50	7	72	GLN	CAA	8	20,51	0,72
	CAG	33,50	33,50	4	27	GLN	CAG	3	7,69	0,27
R-Arg	AGA	71,50	75,70	4	80	ARG	AGA	4	10,25	0,8
	CGT	20,20	24,40	1	20	ARG	CGT	1	2,56	0,2
	AGG	6,00		0	0	ARG	AGG	0	0,00	0
	CGC	0,00		0	0	ARG	CGG	0	0,00	0
	CGA	1,20		0	0	ARG	CGA	0	0,00	0
	CGG	1,20		0	0	ARG	CGC	0	0,00	0
S-Ser	TCT	48,20	52,50	6	32,5	SER	TCT	7	17,94	0,58
	TCC	27,20	31,50	4	25	SER	TCC	3	7,69	0,25
	TCA	11,30	15,60	2	22,5	SER	TCA	2	5,12	0,16
	AGT	4,30		0	20	SER	AGT	0	0,00	0
	TCG	4,30			0	SER	AGC	0	0,00	0
	AGC	4,30			0	SER	TCG	0	0,00	0
T-Thr	ACT	46,50	54,00	10	52	THR	ACT	10	25,64	0,52
	ACC	38,00	45,50	9	47	THR	ACC	9	23,07	0,47
	ACA	9,00		0	0	THR	ACA	0	0,00	0
	ACG	6,00		0	0	THR	ACG	0	0,00	0
V-Val	GTT	48,50	50,67	11	52	VAL	GTT	11	28,20	0,52
	GTC	29,80	31,97	7	33	VAL	GTC	7	17,94	0,33
	GTG	15,30	17,47	4	14	VAL	GTG	3	7,69	0,14
	GTA	6,50		0	0	VAL	GTA	0	0,00	0
W-Trp	TGG	100,00	100,00	8	100	TRP	TGG	8	20,51	1
Y-Tyr	TAC	78,00	78,00	12	81	TYR	TAC	13	33,33	0,81
	TAT	22,00	22,00	4	18	TYR	TAT	3	7,69	0,18
Stop	TAA	78,00	78,00							
	TAG	0,00								
	TGA	25,00	25,00							



## E. Results of MS - analysis of MBP moxY constructs

>Sample\_1\_MBP\_moxY\_with\_Linker

# Mascot Search Results

## Protein View

Match to: 00001|Sample\_1\_MBP\_moxY\_mit\_Linker Score: 11622

Found in search of D:\Data\Projects\Geier-Glieder3\sample-1.RAW

Nominal mass ( $M_r$ ): 109446; Calculated pI value: 5.70

NCBI BLAST search of [00001|Sample\\_1\\_MBP\\_moxY\\_mit\\_Linker](#) against nr Unformatted [sequence string](#) for pasting into other applications

Fixed modifications: Carbamidomethyl (C)

Variable modifications: Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: 69%

Matched peptides shown in **Red**

1	MKIEEGKLV	WINGDKGYNG	LAEVGKKFEK	DTGIKVTVEH	PDKLEEKFPQ
51	VAATGDGPD	IFWAHDRFGG	YAQSGLLAEI	TPDKAFQDKL	YPFTWDAVRY
101	NGKLIAYPIA	VEALSIIYNK	DLLPNPPKTW	EEIPALDKEL	KAKGKSALMF
151	NLQEPYFTWP	LIAADGGYAF	KYENGKYDIK	DVGVDNAGAK	AGLTFLLVDLI
201	KNKHMNADTD	YSIAEAAFNK	GETAMTINGP	WAWSNIDTSK	VNYGVTVLPT
251	FKGQPSKPFV	GVLSAGINAA	SPNKELAKEF	LENYLLTDEG	LEAVNKDKPL
301	GAVALKSYEE	ELAKDPRIAA	TMENAQKGEI	MPNIPQMSAF	WYAVRTAVIN
351	AASGRQTVDE	ALKDAQTLIN	GDGAGLEVLF	QGPENLYFQA	DPANRPLRVV
401	TIGTGISGIL	MAYQIQKQCP	NVEHVLYEKN	ADVGGTWLEN	RYPMAGCDVP
451	SHAYTYPFAP	NPDWPRYFSY	APDIWNYLDR	VCKVFGLRRY	MVFHTEVVGC
501	YWNEDRGEWT	VRLRQHASGS	EPREFEDHCH	VLVHASGVFN	NPQWPQIPGL
551	HDRFQGRVLH	TARWPDDYQE	SQWKSDRVAV	IGSGASSIQT	VPGMQPTVKH
601	LDVVFRTGVW	FGVLAGNTGS	QTKEYSPTER	DEFFRNPAAL	VAHAKAIEDQ
651	VNGMWGAFYT	GSKGQAMGSA	FFRQRTANLI	KDERLREGLD	PPFAFGCRRI
701	TPGDPYMEAI	QKENVHVHFT	PVVSCTEKGV	VGGDGVERQV	DTVVCATGFD
751	ASYPFRFPPIV	GRDGMDLREK	WKECPNSYLG	LAVPEMPNFF	TFIGPTWPIQ
801	NGSVIGPLQA	VSKYVVQWIK	KAQENLRSE	VPRQDRTDQE	NDHVQEWVKH
851	TVWKDNCRSC	TFLSLTLTAE	FTENHLGYKN	NETGRVNAIW	PGSSLHYQQV
901	IDQPRYEDFD	IRSFHENPWA	CLGMGWTIQD	RKGPKEADVS	PHLGLQEIDP
951	KWWESIKNEN	GSENGVSDSNR			

## >Sample\_2\_MBP\_moxY\_without\_Linker

### First band:

Match to: 00002|Sample\_2\_MBP\_moxY\_ohne\_Linker Score: 10066

Found in search of D:\Data\Projects\Geier-Glieder3\Sample-2.RAW

Nominal mass ( $M_r$ ): 107000; Calculated pI value: 5.84

NCBI BLAST search of [00002|Sample\\_2\\_MBP\\_moxY\\_ohne\\_Linker](#) against nr  
Unformatted [sequence string](#) for pasting into other applications

Fixed modifications: Carbamidomethyl (C)

Variable modifications: Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: 64%

Matched peptides shown in **Bold Red**

```
1 MKIEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ
51 VAATGDGPD IFWAHDRFGG YAQSGLLAEI TPKAFQDKL YPFTWDAVRY
101 NGKLIAYPIA VEALSLIYNK DLLPNPPKTW EEIPALDKEL KAKGKSALMF
151 NLQEPYFTWP LIAADGGYAF KYENGKYDIK DVGVDNAGAK AGLTFVLVDLI
201 KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK VNYGVTVLPT
251 FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL
301 GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN
351 AASGRQTVDE ALKDAQTDPA NRPLRVVTIG TGISGILMAY QIQKQCPNVE
401 HVLYEKNADV GGTWLENRYP MAGCDVPSHA YTYPFAPNPD WPRYFSYAPD
451 IWNYLDRVCK VFGLRRYMFV HTEVVGCYWN EDRGEWTVRL RQHASGSEPR
501 EFEDHCHVLV HASGVFNNPQ WPQIPGLHDR FQGRVLHTAR WPDDYQESQW
551 KSDRVAVIGS GASSIQTVPG MQPTVKHLDV FVRTGVWFGV LAGNTGSQTK
601 EYSPTERDEF RRNPAALVAH AKAIEDQVNG MWGAFYTGSK GOAMGSAFFR
651 QRTANLIKDE RLREGLDPPF AFGCRITPG DPYMEAIQKE NVHVHFTPVV
701 SCTEKGVVGG DGVERQVDTV VCATGFDAZY RPRFPVIVGRD GMDLREKWKE
751 CPNSYLGLAV PEMPNFFTFI GPTWPIQNGS VIGPLQAVSK YVVQWIKKAQ
801 NENLRSFVPR QDRTDQFNDH VQEWVKHTVW KDNCRSCTFL LSTLTAEFTE
851 NHLGKYKNNET GRVNAIWPGS SLHYQQVIDQ PRYEDFDIRS FHENPWACLG
901 MGWTIQDRKG PKEADVSPHL GLQEIDPKWW ESIKNENGSN GVSDSNR
```

### Second band below:

Match to: 00002|Sample\_2\_MBP\_moxY\_ohne\_Linker Score: 9417

Found in search of D:\Data\Projects\Geier-Glieder3\Sample-5.RAW

Nominal mass ( $M_r$ ): 107000; Calculated pI value: 5.84

NCBI BLAST search of [00002|Sample\\_2\\_MBP\\_moxY\\_ohne\\_Linker](#) against nr  
Unformatted [sequence string](#) for pasting into other applications

Fixed modifications: Carbamidomethyl (C)

Variable modifications: Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: 55%

Matched peptides shown in **Bold Red**

```
1 MKIEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ
51 VAATGDGPD IFWAHDRFGG YAQSGLLAEI TPKAFQDKL YPFTWDAVRY
101 NGKLIAYPIA VEALSLIYNK DLLPNPPKTW EEIPALDKEL KAKGKSALMF
151 NLQEPYFTWP LIAADGGYAF KYENGKYDIK DVGVDNAGAK AGLTFVLVDLI
201 KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK VNYGVTVLPT
251 FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL
301 GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN
351 AASGRQTVDE ALKDAQTDPA NRPLRVVTIG TGISGILMAY QIQKQCPNVE
401 HVLYEKNADV GGTWLENRYP MAGCDVPSHA YTYPFAPNPD WPRYFSYAPD
451 IWNYLDRVCK VFGLRRYMFV HTEVVGCYWN EDRGEWTVRL RQHASGSEPR
501 EFEDHCHVLV HASGVFNNPQ WPQIPGLHDR FQGRVLHTAR WPDDYQESQW
551 KSDRVAVIGS GASSIQTVPG MQPTVKHLDV FVRTGVWFGV LAGNTGSQTK
601 EYSPTERDEF RRNPAALVAH AKAIEDQVNG MWGAFYTGSK GOAMGSAFFR
651 QRTANLIKDE RLREGLDPPF AFGCRITPG DPYMEAIQKE NVHVHFTPVV
701 SCTEKGVVGG DGVERQVDTV VCATGFDAZY RPRFPVIVGRD GMDLREKWKE
751 CPNSYLGLAV PEMPNFFTFI GPTWPIQNGS VIGPLQAVSK YVVQWIKKAQ
801 NENLRSFVPR QDRTDQFNDH VQEWVKHTVW KDNCRSCTFL LSTLTAEFTE
851 NHLGKYKNNET GRVNAIWPGS SLHYQQVIDQ PRYEDFDIRS FHENPWACLG
901 MGWTIQDRKG PKEADVSPHL GLQEIDPKWW ESIKNENGSN GVSDSNR
```



## >Sample\_3\_moxY\_MBP\_C\_term\_without\_Linker

Match to: 00001|Sample\_1\_MBP\_moxY\_mit\_Linker Score: 9065

Found in search of D:\Data\Projects\Geier-Glieder3\Sample-3.RAW

Nominal mass ( $M_r$ ): 109446; Calculated pI value: 5.70

NCBI BLAST search of 00001|Sample\_1\_MBP\_moxY\_mit\_Linker against nr  
Unformatted [sequence string](#) for pasting into other applications

Fixed modifications: Carbamidomethyl (C)  
Variable modifications: Oxidation (M)  
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
Sequence Coverage: 63%

Matched peptides shown in **Bold Red**

```
1 MKIEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ
51 VAATGDGPDI IFWAHDRFGG YAQSGLLAEI TPDKAFQDKL YPFTWDAVRY
101 NGKLIAYPIA VEALSLIYNK DLLPNPPKTW EEIPALDKEL KAKGKSALMF
151 NLQEPYFTWP LIAADGGYAF KYENKYDIK DVGVDNAGAK AGLTFLVDLI
201 KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK VNYGVTVLPT
251 FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL
301 GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN
351 AASGRQTVDE ALKDAQTLIN GDGAGLELVF QGPNLYFQA DPANRPLRVV
401 TIGTGISGIL MAYQIQKQCP NVEHVLYEKN ADVGGTWLEN RYPMAGCDVP
451 SHAYTYPFAP NPDWPRYFSY APDIWNYLDR VCKVFGLRRY MVFHTEVVGC
501 YWNEDRGEWT VRLRQHASGS EPREFEDHCH VLVHASGVFN NPQWPQIPGL
551 HDRFQGRVLH TARWDDYQE SQWKSDRVAV IGSGASSIQT VPGMQPTVKH
601 LDVFVRTGVW FGVLAGNTGS QTKEYSPTER DEFRNPAAL VAHAKAIEDQ
651 VNGMWGAFYT GSKGQAMGSA FFRQRTANLI KDERLREGLD PPFAFGCRRI
701 TPGDPYMEAI QKENVHVHFT PVVSCTEKGV VGGDGVERQV DTVVCATGFD
751 ASYRPRFPIV GRDGMDLREK WKECPNSYLG LAVPEMPNFF TFIGPTWPIQ
801 NGSVIGPLQA VSKYVVQWIK KAQENLRSF VPRQDRTDQF NDHVQEWVKH
851 TVWKDNCRSC TFLSTLTAE FTENHLGYKN NETGRVNAIW PGSSLHYQQV
901 IDQPRYEDFD IRSFHENPWA CLGMGWTIQD RKGPKEADVS PHLGLQEIDP
951 KWWESIKNEN GSNGVSDSNR
```

## >Sample\_4\_HIS\_MBP\_moxY\_with\_Linker

Match to: 00001|Sample\_1\_MBP\_moxY\_mit\_Linker Score: 6074

Found in search of D:\Data\Projects\Geier-Glieder3\Sample-4.RAW

Nominal mass ( $M_r$ ): 109446; Calculated pI value: 5.70

NCBI BLAST search of 00001|Sample\_1\_MBP\_moxY\_mit\_Linker against nr  
Unformatted [sequence string](#) for pasting into other applications

Fixed modifications: Carbamidomethyl (C)  
Variable modifications: Oxidation (M)  
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
Sequence Coverage: 59%

Matched peptides shown in **Bold Red**

```
1 MKIEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ
51 VAATGDGPDI IFWAHDRFGG YAQSGLLAEI TPDKAFQDKL YPFTWDAVRY
101 NGKLIAYPIA VEALSLIYNK DLLPNPPKTW EEIPALDKEL KAKGKSALMF
151 NLQEPYFTWP LIAADGGYAF KYENKYDIK DVGVDNAGAK AGLTFLVDLI
201 KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK VNYGVTVLPT
251 FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL
301 GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN
351 AASGRQTVDE ALKDAQTLIN GDGAGLELVF QGPNLYFQA DPANRPLRVV
401 TIGTGISGIL MAYQIQKQCP NVEHVLYEKN ADVGGTWLEN RYPMAGCDVP
451 SHAYTYPFAP NPDWPRYFSY APDIWNYLDR VCKVFGLRRY MVFHTEVVGC
501 YWNEDRGEWT VRLRQHASGS EPREFEDHCH VLVHASGVFN NPQWPQIPGL
551 HDRFQGRVLH TARWDDYQE SQWKSDRVAV IGSGASSIQT VPGMQPTVKH
601 LDVFVRTGVW FGVLAGNTGS QTKEYSPTER DEFRNPAAL VAHAKAIEDQ
651 VNGMWGAFYT GSKGQAMGSA FFRQRTANLI KDERLREGLD PPFAFGCRRI
701 TPGDPYMEAI QKENVHVHFT PVVSCTEKGV VGGDGVERQV DTVVCATGFD
751 ASYRPRFPIV GRDGMDLREK WKECPNSYLG LAVPEMPNFF TFIGPTWPIQ
801 NGSVIGPLQA VSKYVVQWIK KAQENLRSF VPRQDRTDQF NDHVQEWVKH
851 TVWKDNCRSC TFLSTLTAE FTENHLGYKN NETGRVNAIW PGSSLHYQQV
901 IDQPRYEDFD IRSFHENPWA CLGMGWTIQD RKGPKEADVS PHLGLQEIDP
951 KWWESIKNEN GSNGVSDSNR
```

F. Vector charts:

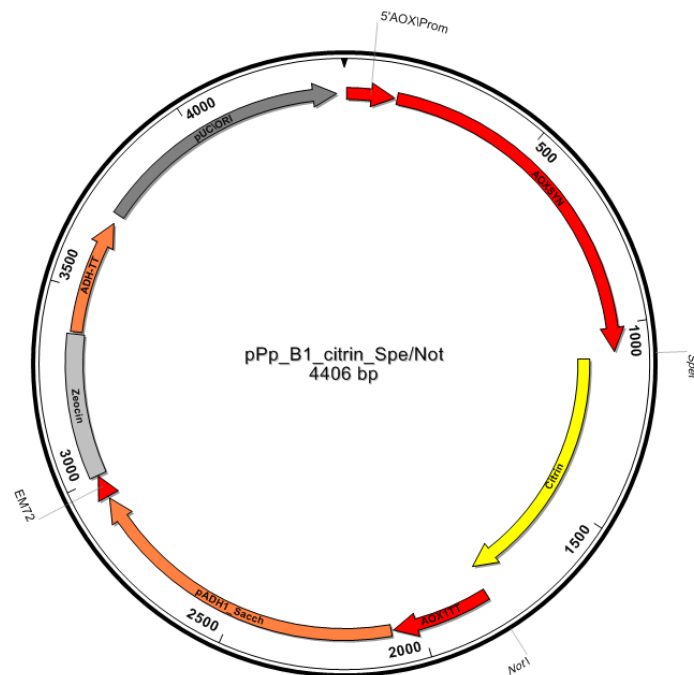


Figure 77: pPp\_B1\_citrin\_Spe/Not: pUC\_ORI: pUC origin of replication, 5'AOX\_Prom: 5' region of AOX1 promoter, AOXSYN: synthetic AOX1 promoter region, Citrin: citrine, AOX1TT: AOX1 transcription termination region, pADH1\_Sacch: eukaryotic promoter of yeast alcohol dehydrogenase, EM72: prokaryotic promoter, Zeocin: Zeocin™ resistance gene, ADH-TT: ADH transcription terminator.

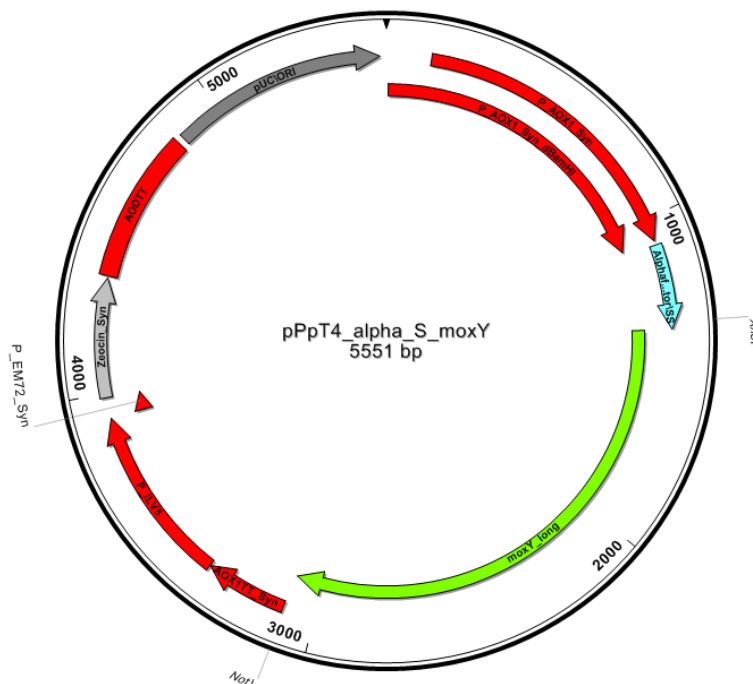


Figure 78: pPpT4\_alpha\_S\_moxY: pUC\_ORI: pUC origin of replication, P\_AOX1\_Syn\_dBamHI: synthetic AOX1 promoter upstream region, P\_AOX1\_Syn: synthetic AOX1 promoter region, Alphafactor\_SS:  $\alpha$ -factor signal peptide, moxY\_long: moxY\_long from *A. flavus*, AOX1TT\_syn: synthetic transcription termination region, P\_ILV5: eukaryotic promoter, P\_EM72\_Syn: synthetic prokaryotic promoter, Zeocin\_Syn: Zeocin™ resistance gene, AODTT: AOD transcription terminator.

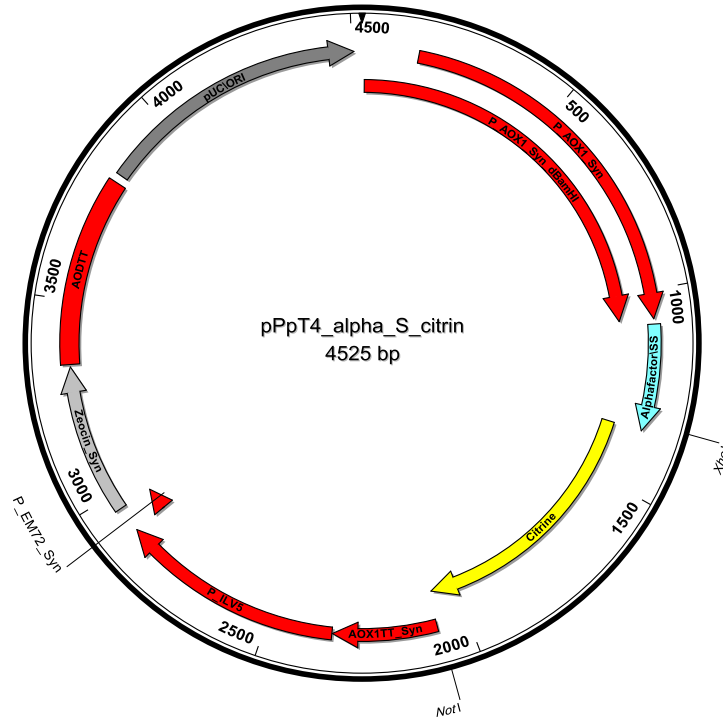


Figure 79: pPpT4\_alpha\_S\_citrin: pUC\_ORI: pUC origin of replication, P\_AOX1\_Syn\_dBamHI: synthetic AOX1 promoter upstream region, P\_AOX1\_Syn: synthetic AOX1 promoter region, Alphafactor\_SS:  $\alpha$ -factor signal peptide, Citrine: citrine, AOX1TT\_syn: synthetic transcription termination region, P\_ILV5: eukaryotic promoter, P\_EM72\_Syn: synthetic prokaryotic promoter, Zeocin\_Syn: Zeocin<sup>TM</sup> resistance gene, AODTT: AOD transcription terminator.

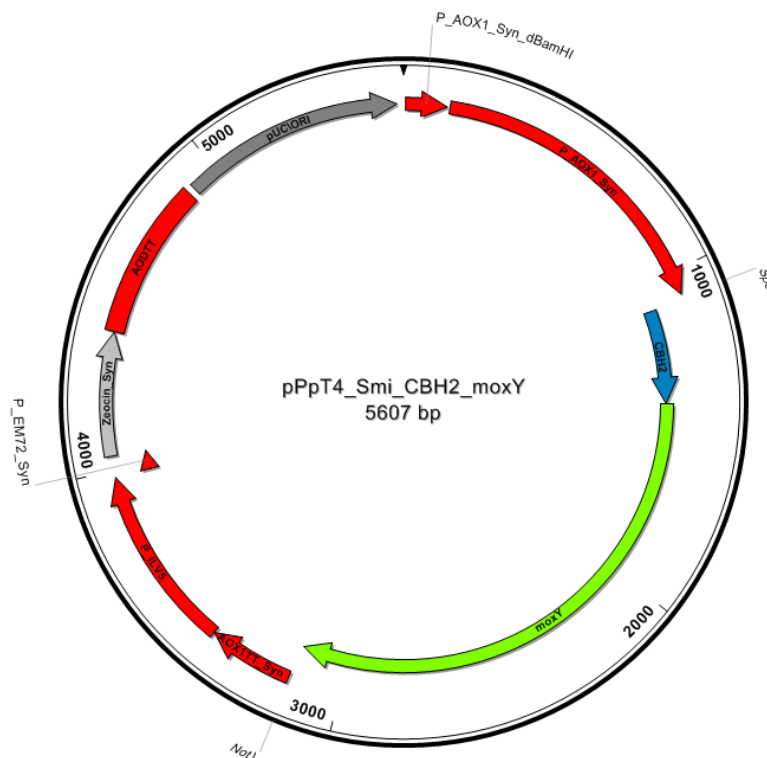


Figure 80: pPpT4\_Smi\_CBH2\_moxY: pUC\_ORI: pUC origin of replication, P\_AOX1\_Syn\_dBamHI: synthetic AOX1 promoter upstream region, P\_AOX1\_Syn: synthetic AOX1 promoter region, CBH2: signal peptide sequence of CBH2, moxY: moxY\_long from *A. flavus*, AOX1TT\_syn: synthetic transcription termination region, P\_ILV5: eukaryotic promoter, P\_EM72\_Syn: synthetic prokaryotic promoter, Zeocin\_Syn: Zeocin<sup>TM</sup> resistance gene, AODTT: AOD transcription terminator.

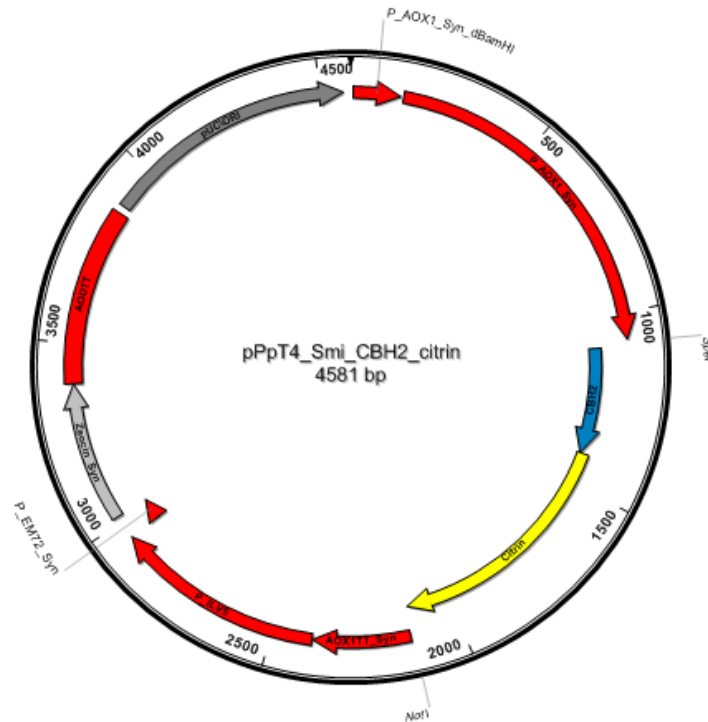


Figure 81: pPpT4\_Smi\_CBH2\_citrin: pUC\_ORI: pUC origin of replication, P\_AOX1\_Syn\_dBamHI: synthetic AOX1 promoter upstream region, P\_AOX1\_Syn: synthetic AOX1 promoter region, CBH2: signal peptide sequence of CBH2, Citrin: citrine, AOX1TT\_syn: synthetic transcription termination region, P\_ILV5: eukaryotic promoter, P\_EM72\_Syn: synthetic prokaryotic promoter, Zeocin\_Syn: Zeocin™ resistance gene, AODTT: AOD transcription terminator.

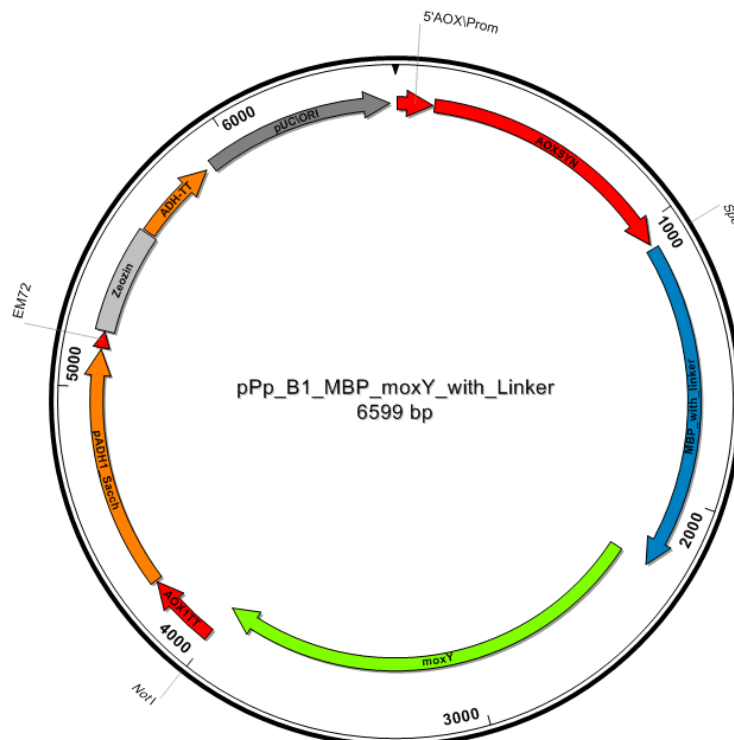


Figure 82: pPp\_B1\_MBP\_moxY\_with\_Linkers: pUC\_ORI: pUC origin of replication, 5'AOX1Prom: 5' region of AOX1 promoter, AOX1TT: synthetic AOX1 promoter region, MBP\_with linker: sequence optimized MBP with linker, moxY: moxY\_long from *A. flavus*, AOX1TT: AOX1 transcription termination region, pADH1\_Sacch: eukaryotic promoter of yeast alcohol dehydrogenase, EM72: prokaryotic promoter, Zeocin: Zeocin™ resistance gene, ADH-TT: ADH transcription terminator.

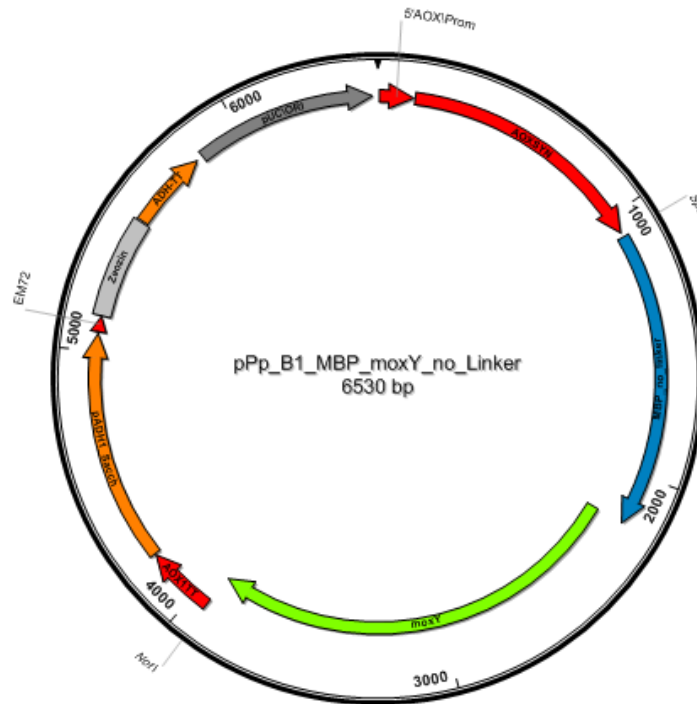


Figure 83: pPp\_B1\_MBP\_moxY\_no\_Linker: pUC\_ORI: pUC origin of replication, 5'AOX\_Prom: 5' region of AOX1 promoter, AOXSYN: synthetic AOX1 promoter region, MBP\_no linker: sequence optimized MBP without linker, moxY: moxY\_long from *A. flavus*, AOX1TT: AOX1 transcription termination region, pADH1\_Sacch: eukaryotic promoter of yeast alcohol dehydrogenase, EM72: prokaryotic promoter, Zeocin: Zeocin™ resistance gene, ADH-TT: ADH transcription terminator.

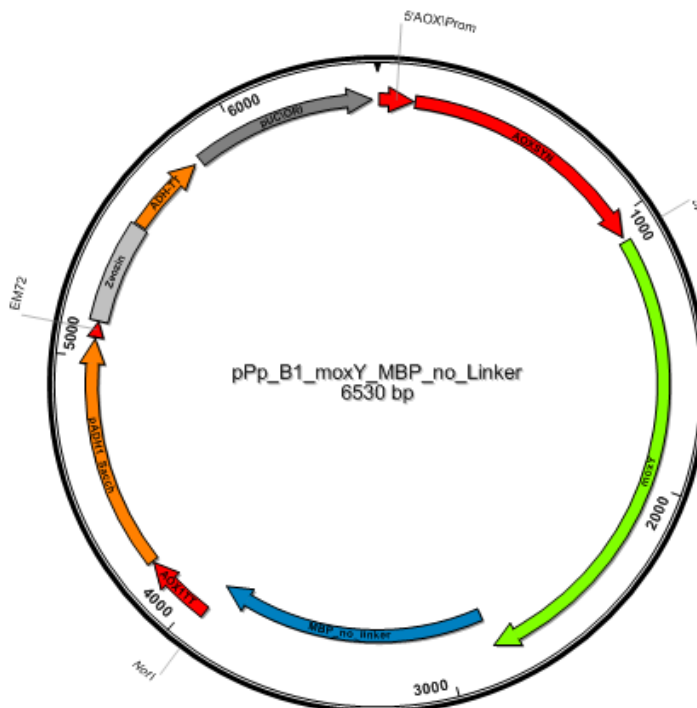


Figure 84: pPp\_B1\_moxY\_MBP\_no\_Linker: pUC\_ORI: pUC origin of replication, 5'AOX\_Prom: 5' region of AOX1 promoter, AOXSYN: synthetic AOX1 promoter region, moxY: moxY\_long from *A. flavus*, MBP\_no linker: sequence optimized MBP without linker, AOX1TT: AOX1 transcription termination region, pADH1\_Sacch: eukaryotic promoter of yeast alcohol dehydrogenase, EM72: prokaryotic promoter, Zeocin: Zeocin™ resistance gene, ADH-TT: ADH transcription terminator.

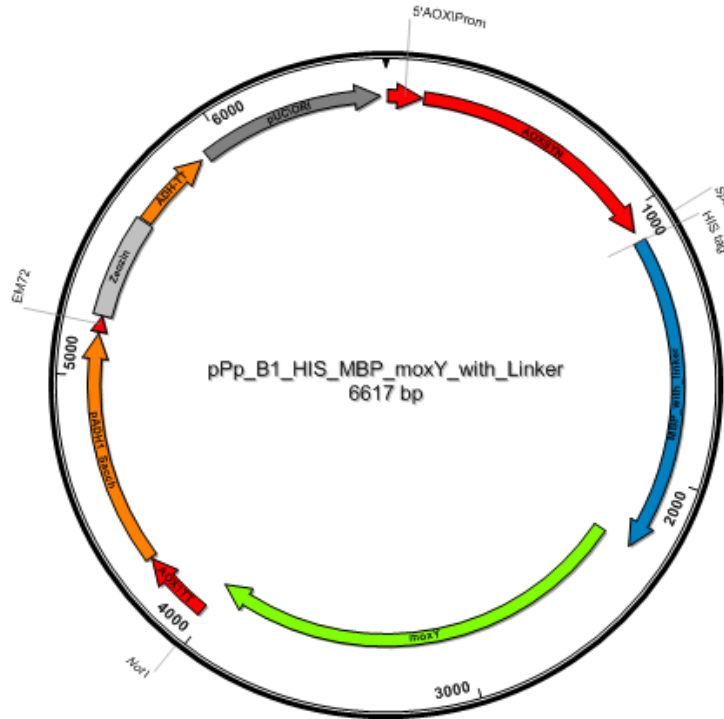


Figure 85: pPp\_B1\_HIS\_MBP\_moxY\_with\_Linker: pUC\_ORI: pUC origin of replication, 5'AOX\_Prom: 5' region of AOX1 promoter, AOXSYN: synthetic AOX1 promoter region, HIS tag: histidine - tag, MBP\_with linker: sequence optimized MBP with linker, moxY: moxY\_long from *A. flavus*, AOX1TT: AOX1 transcription termination region, pADH1\_Sacch: eukaryotic promoter of yeast alcohol dehydrogenase, EM72: prokaryotic promoter, Zeocin: Zeocin™ resistance gene, ADH-TT: ADH transcription terminator.

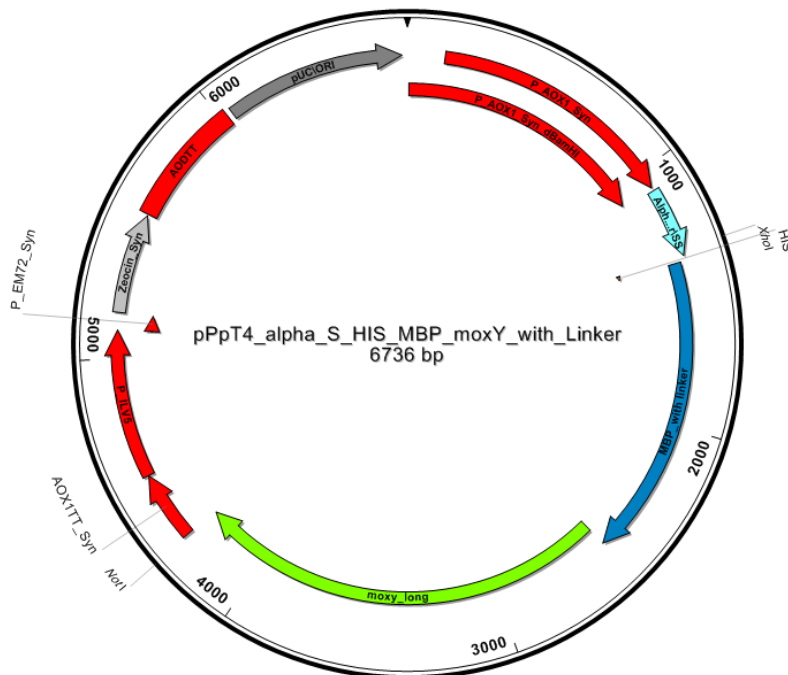


Figure 86: pPpT4\_alpha\_S\_HIS\_MBP\_moxY\_with\_Linker: pUC\_ORI: pUC origin of replication, P\_AOX1\_Syn\_dBamHI: synthetic AOX1 promoter upstream region, P\_AOX1\_Syn: synthetic AOX1 promoter region, Alphafactor\_SS:  $\alpha$  - factor signal peptide, HIS: histidine - tag, MBP\_with linker: sequence optimized MBP with linker, moxY\_long: moxY\_long from *A. flavus*, AOX1TT\_syn: synthetic transcription termination region, P\_ILV5: eukaryotic promoter, P\_EM72\_Syn: synthetic prokaryotic promoter, Zeocin\_Syn: Zeocin™ resistance gene, AODTT: AOD transcription terminator.

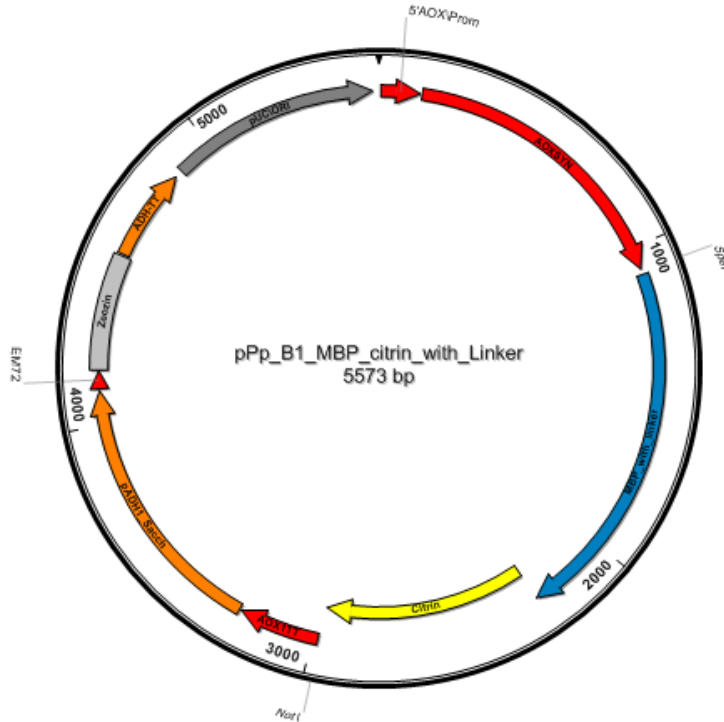


Figure 87: pPp\_B1\_MBP\_citrin\_with\_Linker: pUC\_ORI: pUC origin of replication, 5'AOX\_Prom: 5' region of AOX1 promoter, AOXSYN: synthetic AOX1 promoter region, MBP\_with linker: sequence optimized MBP with linker, Citrin: citrine, AOX1TT: AOX1 transcription termination region, pADH1\_Sacch: eukaryotic promoter of yeast alcohol dehydrogenase, EM72: prokaryotic promoter, Zeocin: Zeocin<sup>TM</sup> resistance gene, ADH-TT: ADH transcription terminator.

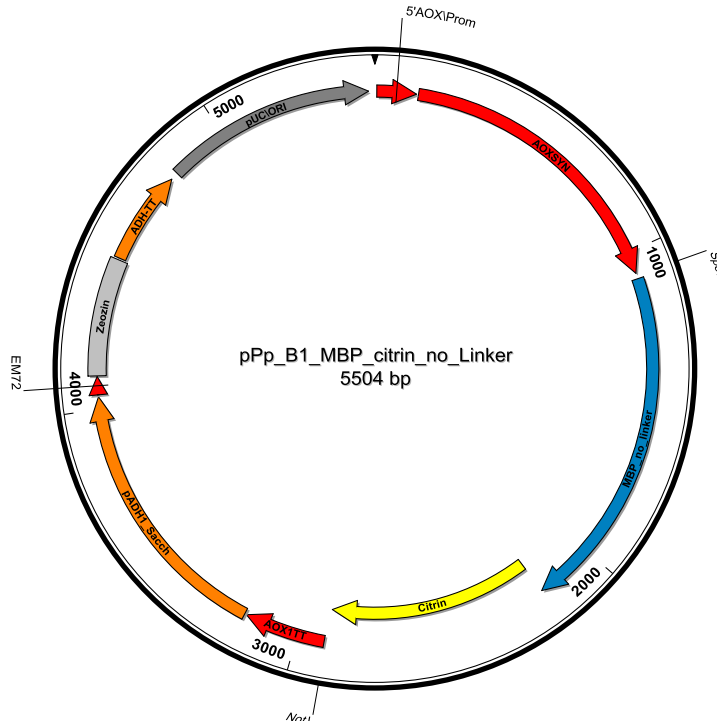


Figure 88: pPp\_B1\_MBP\_citrin\_no\_Linker: pUC\_ORI: pUC origin of replication, 5'AOX\_Prom: 5' region of AOX1 promoter, AOXSYN: synthetic AOX1 promoter region, MBP\_no linker: sequence optimized MBP without linker, Citrin: citrine, AOX1TT: AOX1 transcription termination region, pADH1\_Sacch: eukaryotic promoter of yeast alcohol dehydrogenase, EM72: prokaryotic promoter, Zeocin: Zeocin<sup>TM</sup> resistance gene, ADH-TT: ADH transcription terminator.

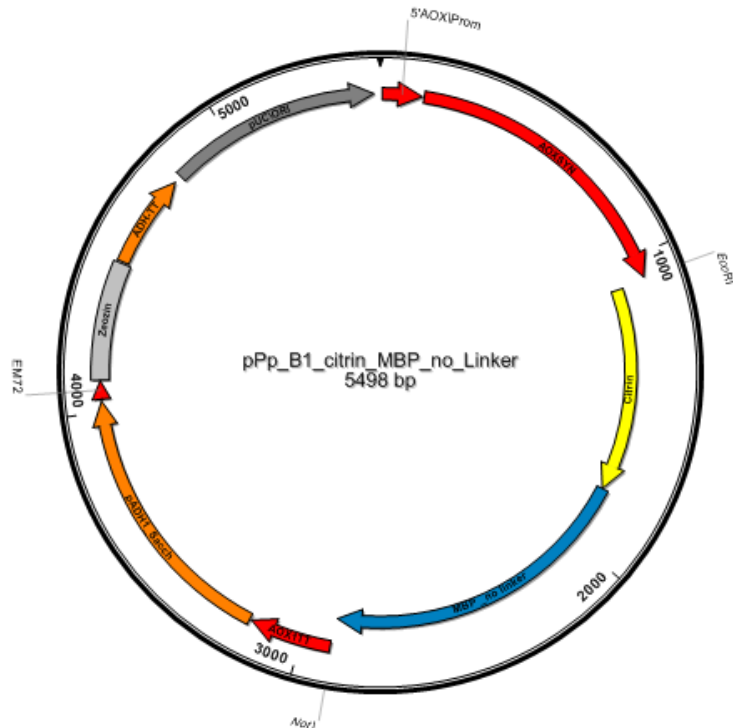


Figure 89: pPp\_B1\_citrin\_MBP\_no\_Linker: pUC\_ORI: pUC origin of replication, 5'AOX\_Prom: 5' region of AOX1 promoter, AOXSYN: synthetic AOX1 promoter region, Citrin: citrine, MBP\_no linker: sequence optimized MBP without linker, AOX1TT: AOX1 transcription termination region, pADH1\_Sacch: eukaryotic promoter of yeast alcohol dehydrogenase, EM72: prokaryotic promoter, Zeocin: Zeocin™ resistance gene, ADH-TT: ADH transcription terminator.

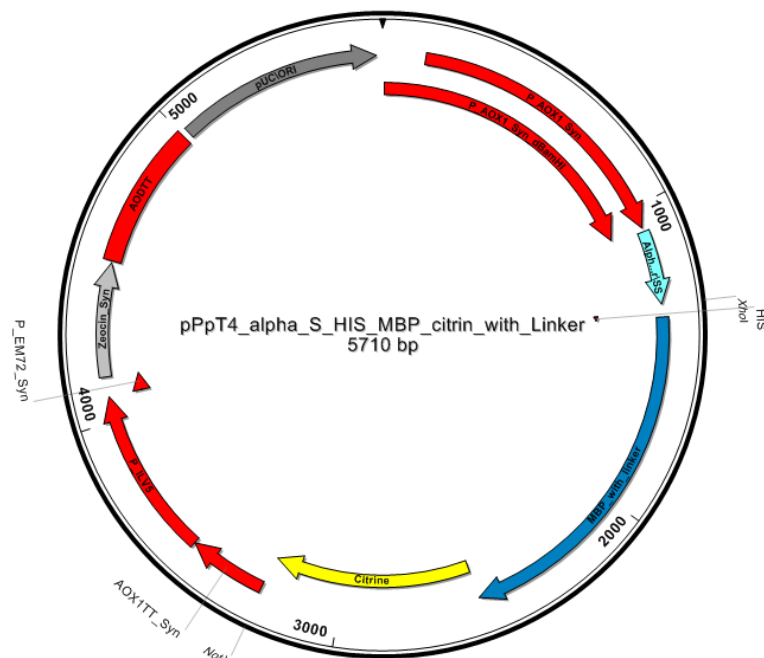


Figure 90: pPpT4\_alpha\_S\_HIS\_MBP\_citrin\_with\_Linker: pUC\_ORI: pUC origin of replication, P\_AOX1\_Syn\_dBamHI: synthetic AOX1 promoter upstream region, P\_AOX1\_Syn: synthetic AOX1 promoter region, Alphafactor\_SS:  $\alpha$  - factor signal peptide, HIS: histidine - tag, MBP\_with linker: sequence optimized MBP with linker, Citrine: citrine, AOX1TT\_syn: synthetic transcription termination region, P\_ILV5: eukaryotic promoter, P\_EM72\_Syn: synthetic prokaryotic promoter, Zeocin\_Syn: Zeocin™ resistance gene, AODTT: AOD transcription terminator.



## G. Strain collection

**Table 31: Strains for the strain collection with their host, strain collection number (CC Nr.) and vector.**

CC Nr.	Host strain	Sequence file (vector designation)	Vector (size/marker/cloning sites)	Description of construction
6539	<i>E. coli</i>	pPpT4_alpha_S_citrin.gb	4525 bp/Zeo-R/XhoI/NotI	secretion of citrin with alpha factor signal sequence
6540	<i>E. coli</i>	pPpT4_Smi_CBH2_citrin.sbd	4581bp/Zeo-R/SpeI/NotI	secretion of citrin with CBH2
6541	<i>E. coli</i>	pPpT4_alpha_S_moxY.gb	5551 bp/Zeo-R/XhoI/NotI	secretion of moxY with alpha factor signal sequence
6542	<i>E. coli</i>	pPpT4_Smi_CBH2_moxY.sbd	5607bp/Zeo-R/SpeI/NotI	secretion of moxY with CBH2
6543	<i>E. coli</i>	pPp_B1_citrin_Spe_Not.sbd	4406 bp/Zeo-R/SpeI/NotI	citrine cloned into B1 vector with SpeI and NotI
6544	<i>E. coli</i>	pPp_B1_MBP_moxY_with_Linkers.sbd	6599 bp/Zeo-R/SpeI/NotI	N-terminal fusion of MBP-tag with TEV cleavage - linker sequence to moxY
6545	<i>E. coli</i>	pPp_B1_MBP_moxY_no_Linkers.sbd	6530 bp/Zeo-R/SpeI/NotI	N-terminal fusion of MBP-tag without TEV cleavage - linker sequence to moxY
6546	<i>E. coli</i>	pPp_B1_moxY_MBP_no_Linkers.sbd	6530bp/Zeo-R/SpeI/NotI	C-terminal fusion of MBP-tag without TEV cleavage - linker sequence to moxY
6547	<i>E. coli</i>	pPp_B1_HIS_MBP_moxY_with_Linkers.sbd	6617 bp/Zeo-R/SpeI/NotI	N-terminal fusion of HIS and MBP-tag with TEV cleavage - linker sequence to moxY
6548	<i>E. coli</i>	pPp_B1_MBP_citrin_with_Linkers.sbd	5573 bp/Zeo-R/SpeI/NotI	N-terminal fusion of MBP-tag with TEV cleavage - linker sequence to citrin
6549	<i>E. coli</i>	pPp_B1_MBP_citrin_no_Linkers.sbd	5504 bp/Zeo-R/SpeI/NotI	N-terminal fusion of MBP-tag without TEV cleavage - linker sequence to citrin
6550	<i>E. coli</i>	pPp_B1_citrin_MBP_no_Linkers.sbd	5498 bp/Zeo-R/EcoRI/NotI	C-terminal fusion of MBP-tag without TEV cleavage - linker sequence to citrin
6551	<i>E. coli</i>	pPpT4_alpha_S_HIS_MBP_citrin_with_Linkers.gb	5710 bp/Zeo-R/XhoI/NotI	N-terminal fusion of HIS and MBP-tag with TEV cleavage - linker sequence to citrin plus the alpha factor signal sequence for secretion
6552	<i>E. coli</i>	pPpT4_alpha_S_HIS_MBP_moxY_with_Linkers.gb	6736 bp/Zeo-R/XhoI/NotI	N-terminal fusion of HIS and MBP-tag with TEV cleavage - linker sequence to moxY plus the alpha factor signal sequence for secretion
6553	<i>P. pastoris</i>	pPpT4_alpha_S_citrin.gb	4526 bp/Zeo-R/XhoI/NotI	secretion of citrin with alpha factor signal sequence
6554	<i>P. pastoris</i>	pPpT4_Smi_CBH2_citrin.sbd	4581 bp/Zeo-R/SpeI/NotI	secretion of citrin with CBH2
6555	<i>P. pastoris</i>	pPpT4_alpha_S_moxY.gb	5551 bp/Zeo-R/XhoI/NotI	secretion of moxY with alpha factor signal sequence, secreting clone C6 (albumin?)
6556	<i>P. pastoris</i>	pPp_B1_citrin_Spe_Not.sbd	4406 bp/Zeo-R/SpeI/NotI	citrine cloned into B1 vector with SpeI and NotI
6557	<i>P. pastoris</i>	pPp_B1_MBP_moxY_with_Linkers.sbd	6599 bp/Zeo-R/SpeI/NotI	N-terminal fusion of MBP-tag with TEV cleavage - linker sequence to moxY, used for first fermentation
6558	<i>P. pastoris</i>	pPp_B1_MBP_moxY_with_Linkers.sbd	6599 bp/Zeo-R/SpeI/NotI	N-terminal fusion of MBP-tag with TEV cleavage - linker sequence to moxY, higher expressing clone A4
6559	<i>P. pastoris</i>	pPp_B1_MBP_moxY_no_Linkers.sbd	6530 bp/Zeo-R/SpeI/NotI	N-terminal fusion of MBP-tag without TEV cleavage - linker sequence to moxY

6560	<i>P. pastoris</i>	pPp_B1_moxY_MBP_no_Linker.sbd	6530 bp/Zeo-R/Spel/NotI	C-terminal fusion of MBP-tag without TEV cleavage - linker sequence to moxY
6561	<i>P. pastoris</i>	pPp_B1_HIS_MBP_moxY_with_Linker.sbd	6617 bp/Zeo-R/Spel/NotI	N-terminal fusion of HIS and MBP-tag with TEV cleavage - linker sequence to moxY
6562	<i>P. pastoris</i>	pPp_B1_HIS_MBP_moxY_with_Linker.sbd	6617 bp/Zeo-R/Spel/NotI	N-terminal fusion of HIS and MBP-tag with TEV cleavage - linker sequence to moxY, higher expressing clone A9
6563	<i>P. pastoris</i>	pPp_B1_MBP_citrin_with_Linker.sbd	5573 bp/Zeo-R/Spel/NotI	N-terminal fusion of MBP-tag with TEV cleavage - linker sequence to citrin
6564	<i>P. pastoris</i>	pPp_B1_MBP_citrin_no_Linker.sbd	5504 bp/Zeo-R/Spel/NotI	N-terminal fusion of MBP-tag without TEV cleavage - linker sequence to citrin
6565	<i>P. pastoris</i>	pPp_B1_citrin_MBP_no_Linker.sbd	5498 bp/Zeo-R/EcoRI/NotI	C-terminal fusion of MBP-tag without TEV cleavage - linker sequence to citrin
6566	<i>P. pastoris</i>	pPpT4_alpha_S_HIS_MBP_citrin_with_Linker.gb	5710 bp/Zeo-R/XhoI/NotI	N-terminal fusion of HIS and MBP-tag with TEV cleavage - linker sequence to citrin plus the alpha factor signal sequence for secretion
6567	<i>P. pastoris</i>	pPpT4_alpha_S_HIS_MBP_moxY_with_Linker.gb	6736 bp/Zeo-R/XhoI/NotI	N-terminal fusion of HIS and MBP-tag with TEV cleavage - linker sequence to moxY plus the alpha factor signal sequence for secretion

**Table 32: Cultivation conditions for the *E. coli* and *P. pastoris* strains of the strain collection.**

	Vector	Culture medium	Temperature	Incubation
<i>E. coli</i>	<i>E. coli</i> - <i>P. pastoris</i> shuttle vector	LB with 25 µg/ml zeocin	37 °C	1 day
<i>P. pastoris</i>	<i>E. coli</i> - <i>P. pastoris</i> shuttle vector	YPD with 100 µg/ml zeocin	28 °C	2 days