PETER NEUGEBAUER

Enhancing HRP secretion in *Pichia pastoris*

Masterarbeit

zur Erlangung des akademischen Grades eines
Diplom-Ingenieurs
an der Naturwissenschaftlichen Fakultät der TU Graz

Dipl.-Ing. Christine Winkler
Ass.-Prof. Dipl.-Ing. Dr.techn. Harald Pichler
Univ.-Prof. Dipl.-Ing. Dr.techn. Helmut Schwab

Institut für Molekulare Biotechnologie
Technische Universität Graz

Danksagung

Ich möchte den folgenden Personen, die mich während dieser Arbeit und des gesamten Studiums unterstützt haben meinen herzlichsten Dank aussprechen:

- Meinem Betreuer Harald Pichler, der mich in seine Forschungsgruppe aufnahm und jederzeit für hilfreiche Diskussionen zur Verfügung stand. Er half mir den roten Faden nicht zu verlieren und gab mir damit die Möglichkeit diese Masterarbeit erfolgreich zu beenden.
- Herrn Prof. Helmut Schwab, der ebenfalls großen Anteil an der Ermöglichung dieser Arbeit hat und mir in vielen Gruppenseminaren mit Ratschlägen weiterhalf.
- Meiner Betreuerin Christine Winkler, die mir während der ersten Monate meiner Arbeit im Labor mit Rat und Tat zur Seite stand und mir später die Möglichkeit gab Erlerntes selbstständig anzuwenden, während sie trotz neun Stunden Zeitunterschieds dennoch den Überblick über meine Arbeit behielt.
- Meinen lieben Kolleginnen, zu denen sich nach etlichen gemeinsam im Labor und auch außerhalb der Universität verbrachten Stunden eine besondere Freundschaft entwickelt hat: Anita, mit der ich mein Leid teilen konnte wenn Mittwochs die Nudeln nicht rechtzeitig am Tisch standen, die mich in die Freuden den Skitouren-Gehens einweihte und mir sogar hin und wieder erlaubte einen von ihr reservierten Schikanenkolben zu benutzen. Melanie, die dafür sorgte, dass trotz Protesten der Vegetarier auch Speck in die Eierspeiß kam und die mir weiterhalf wenn diverse Computerprogramme nicht taten was sie sollten. Eva, die mich mit ihrer lieben und aufgeweckten Art dazu brachte die Namen aller Personen ihrer komplizierten Freundeskreise auswendig zu lernen und mich im Labor unter anderem mit der Herstellung ihrer berühmten, superkompetenten Zellen unterstütze. Ari, die mich mit den komplizierten Regeln des "Kubbens" vertraut machte und Tamara, die sich auch nicht aus der Ruhe bringen ließ wenn ich sie des Öfteren mit Fragen über den Verbleib diverser Laborutensilien löcherte.
- Meinen Eltern, meiner Schwester Katharina, meiner Großmutter und meiner lieben Elisabeth, die mir halfen nie das Ziel aus den Augen zu verlieren und hinter mir standen wenn die Motivation nachließ. Ihr seid mein Lebensmittelpunkt, habt immer an mich geglaubt und habt zugehört wenn ich etwas auf dem Herzen hatte.
- Und abschließend all den unzähligen lieben Freunden und Bekannten, die meine Studienzeit unvergesslich machten.

Abstract

Over the past years the methylotrophic yeast *Pichia pastoris* has become an important model organism in biotechnology. Although *Pichia pastoris* is an important eukaryotic system for production of heterologous proteins still little is known about molecular processes occuring during protein secretion into the culture supernatant. Horseradish peroxidase (HRP) and Alternative pig liver esterase (APLE) secreting strains were mutagenized by random integration of resistance cassettes into their genomes by Christine Winkler. Selected strains served as a starting point for the present work. Strains showing increased secretion behaviour, compared to a wild type control, were analyzed by genome walking experiments. Thereby, genes affected by the integration were identified and their influence on protein secretion was evaluated. Southern blot analysis was applied to check for multiple integrations that might have been overlooked in genome walking experiments. In order to verify the screening results directed target knock-outs of the respective genes were performed using knockout plasmids designed and constructed ad hoc. Target knock-out mutants were transformed with the model protein expression cassettes and were analyzed in activity assays comparing them to strains from the screening and wild type controls. Deep-well plate as well as shake flask assays confirmed the influence of affected genes on protein secretion behavior. Increases in HRP activity in the culture supernatant were found to be up to the two-fold compared to expression strains in the wild type background. Furthermore, evaluation of cell growth showed no influences on cell viability or altered growth behavior in any of the directed knock-out strains. Thus, numerous targets for enhancing recombinant protein (HRP) secretion from *P. pastoris* by cell engineering have been identified.

Zusammenfassung

Die methylotrophe Hefe *Pichia pastoris* hat sich im Laufe der letzten Jahre zu einem der bedeutensten Modellorganismen der Biotechnologie entwickelt. Obwohl sie bekannt ist für ihre positiven Eigenschaften wie Wachstum zu hohen Zelldichten und genetische Manipulationen verhältnismäßig einfach durchführbar sind, sowie ihre Produkte den GRAS-Status der FDA besitzen, ist noch immer wenig über die molekularen Hintergründe während der Proteinsekretion in das Medium bekannt. Als Ausgangspunkt für die vorliegende Arbeit standen von Christine Winkler konstruierte HRP- (Meerrettichperoxidase) und APLE- (alternative Schweineleberesterase) sekretierende *P. pastoris*-Stämme zur Verfügung. Durch an zufälligen Positionen im Genom integrierte Resistenz-Kassetten zeigten sich in einem Screening in deep-well Platten veränderte Aktivitäten des Modellproteins im Zellüberstand. Darunter befanden sich auch Mutanten, die im Vergleich zum Wildtyp stark erhöhte Sekretion aufwiesen und im folgenden näher charakterisiert wurden. Durch Genome-Walking Experimente konnte auf die durch Integration der Resistenz-Kassette betroffenen Gene geschlossen werden. Um den Einfluss besonders interessierender Gene auf die Sekretionscharakteristik zu verifizieren sollten entsprechende target-knock-out Stämme generiert und mit den korrespondierenden Wildtyp-Stämmen verglichen werden. Dabei konnte durch Zellkulturen in deep-well Platten und auch in Erlenmeyerkolben gezeigt werden, dass die Sekretion in allen getesteten Fällen erhöht, und gleichzeitig das Zellwachstum in keinem Fall beeinträchtigt war. Somit wurden mehrere Gene identifiziert, deren Manipulation zur erhöhten Sekretion von rekombinanter HRP führt.

Table of contents

1 Introduction 7

1.1 Secretion in *Pichia pastoris* 7

1.2 HRP 8

1.3 APLE 9

1.4 Integration of heterologous DNA into the genome of *P. pastoris* 10

2 Objectives 13

3 Materials and Methods 15

3.1 Strains, plasmids, primers 15

3.2 Instruments, reagents, media 17

3.3 Methods 20

3.3.1 General methods 20

3.3.1.1 General PCR conditions 20

3.3.1.2 Colony PCR 20

3.3.1.3 Gel elektrophoresis 21

3.3.1.4 Isolation of genomic DNA 21

3.3.1.5 Southern Blot 22

3.3.1.6 Electrotransformation in *E. coli* 24

3.3.1.7 Electrocompetent P. pastoris cells and Transformation 24

3.3.1.8 Integration of heterologous DNA at the AOX1 locus and screening for MutS phenotype and histidin prototrophy 26

3.3.2 Enzyme-activity measurements and strain characterization 27

3.3.2.1 Cell cultivation in DWPs and shake flasks for determination of secretion activity 27

3.3.2.2 HRP isolation from intracellular fractions 27

3.3.2.3 ABTS Assay 27

3.3.2.4 pNPA Assay 28

3.3.2.5 Calcofluor White Spotting 28

3.3.2.6 Growth curves 28

3.3.3 Target knock-outs in *P. pastoris* 29

3.3.3.1 Construction of Knock-out plasmids 33

3.3.3.2 Performing knock-outs in Pichia pastoris and screening for positive transformants 33

3.3.3.3 Removal of zeocin-marker 34
3.3.4 Genome Walking experiments 35
3.3.4.1 Template Blocking PCR 35
3.3.4.2 Genome Walker Kit Strategy 37
4 Results 38
4.1 Genome walking 38
4.2 Cell wall integrity tests 43
4.2.1 HRP activity in the supernatant and intracellular fractions 43
4.3 Establishment of HRP secreting strains as host strains for target knock-outs 44
4.4 Functions and properties of target knock-out genes 52
4.4.1 Kcs1p 52
4.4.2 Rim101p 53
4.4.3 Sgt2 54
4.4.4 Hypothetical protein of strain H8 55
4.5 Growth curves 55
4.6 HRP activity measurements in cell supernatant 57
4.6.1 Cell cultivation in DWPs 57
4.6.2 Cell cultivation in shake flasks 59
5 Discussion 61
5.1 Outlook 63
6 Supplementary material 64
6.1 Construction of knockout plasmids 64
List of figures 70
References 73

1 Introduction

1.1 Secretion in *Pichia pastoris*

The *P. pastoris* expression system has gained great acceptance as an outstanding host system for producing heterologous proteins. Major advantages of this yeast include the growth to high cell densities, little secretion of endogenous proteins, ease of genetic manipulations, no known human pathogenicity and GRAS status of *P. pastoris* products. Furthermore, its abilities regarding posttranslational modifications that include polypeptide folding, glycosylation, methylation, acylation, formation of disulfide bonds and targeting to subcellular compartments as well as to extracellular space make *P. pastoris* an important host for the production of recombinant proteins [1].
Due to the fact that if a protein of interest produced in *P. pastoris* is efficiently secreted it is almost the only protein found in the supernatant, downstream processing (purification) is strongly facilitated. Cell lysis and purification from cell extracts can thereby be circumvented.

One of *P. pastoris'* major benefits is the presence of a very strong and inducible promoter, the methanol induced alcoholoxidase promoter P_{AOX1}. It plays a key role in the methanol utilization pathway and is highly repressed on most carbon sources. Its power is shown by the fact that upon methanol induction the fraction of total protein that is composed of alcoholoxidase can rise up to 30% [2].

In cases where constitutive expression is beneficial another promoter, namely the *GAP* promoter derived from the *P. pastoris* glyceraldehydes-3-phosphate dehydrogenase gene, is available. It does not depend on methanol, nor is it necessary to shift the carbon source for induction purposes. For expression of proteins with potentially toxic effects on the cell the *GAP* promoter is not recommended.
Other available promoters are *FLD1*, *PEX8* and *YPT7*, which are not regularly used in the lab [2].

Several secretion signal sequences are known to work in *P. pastoris*, but the most important one is the *Saccharomyces cerevisiae* alpha mating factor secretion signal sequence (αMF) [1]. When placed in frame with the codons of the protein of interest it encodes a pre- and a pro-sequence of 19 and 66 aa-residues, respectively. During its processing the pre-signal is removed in the endoplasmic reticulum by a signal peptidase. The endopeptidase Kex2p cleaves off the pro-leader sequence in the Golgi. Ultimately Glu-Ala repeats are trimmed by

the Ste13 protein. Amino acids of the protein of interest adjacent to the αMF may have an influence on the efficiency of the processing process [3].

The vesicular transport hypothesis is commonly recognized as a mechanistic model for protein secretion in *P. pastoris.* In this model it is assumed that transport vesicles transfer cargo molecules between organelles of the secretory pathway. Important membrane-enclosed organelles include the endoplasmic reticulum (ER), the Golgi complex and secretory granules. In a tightly regulated process called “vesicle budding” vesicles bud from a donor compartment, selectively incorporating cargo, and fuse to “acceptor” organelles in which they unload secretory proteins. This series of several budding and fusing processes at consecutive transport steps is accompanied by retrieval of transport machinery components in order to permit organelle homeostasis by retrograde transport [4].
Figure 1 shows the compartements of the secretory, vacuolar and endocytic pathways including COPI and COPII vesicles as well as clathrin coated structures.

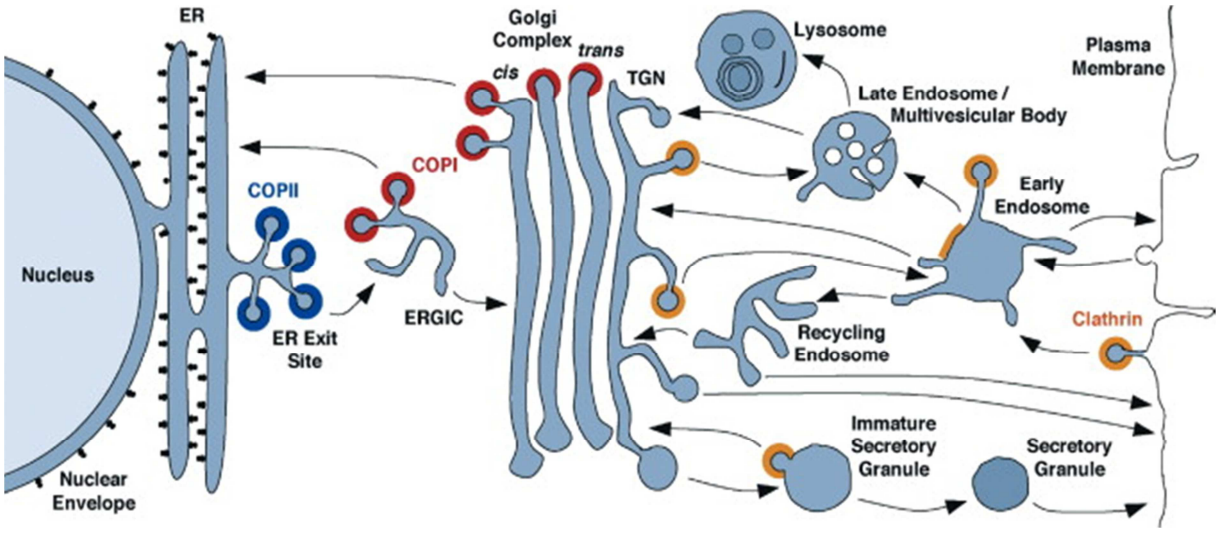


Figure 1: Intracellular transport pathways [4]

1.2 HRP

Horseradish peroxidase is an enzyme with a broad application range in biochemistry in general, more precisely in clinical diagnosis kits and immunoassays. In horseradish its major functions are the removal of H_2O_2 and the oxidation of toxic reductants. Moreover, it takes part in the biosynthesis and degradation of lignin and in response to environmental and oxidative stresses and pathogen attack [5]. Several isoenzymes exist, isoenzyme C being the most common and best characterized one regarding structure, catalytic ability and use

for industrial applications. It consists of a single peptide chain of 308 amino acids including 4 disulfide bridges and 1 salt bridge.

The motif Asn-X-Ser/Thr known to be readily available for N-glycosylation at the asparagine residue is found nine times in the protein sequence, of which 8 are occupied by carbohydrate groups. The glycosylation profile of the isoenzyme is heterogeneous, although a carbohydrate side chain consisting of 7 residues is found in 70-85% of the cases. Due to the reported heterogeneity the molecular weight varies between 40 and 45 kDa [5, 6].

Essential for both the catalytic activity as well as the stability of HRP are two calcium ions and a heme group (iron(III) protoporphyrin IX). A histidine side chain (His170) links the heme group to the protein by a coordinated bond. Aspartate, serine and threonine residues bind the two calcium ions [7].

1.3 APLE

Pig liver esterases (PLEs) isolated as crude mixtures of different isoenzymes have been used for more than 30 years as they have shown to be capable of metabolizing a wide range of different substrates. Although these inhomogeneous blends had found applications in kinetic resolutions, in desymmetrizations of prochiral compounds and in the synthesis of nucleosides assuring the production of pure PLE isoenzymes was aspired [8]. Attempts to separate mixtures of different PLEs were hampered by their similar purification properties [9]. Additionally, as an animal-derived product its use in pharmaceutical production processes was rejected. These problems could be overcome by producing isoenzymes like γ-PLE in different host systems, such as *Escherichia coli* and also *Pichia pastoris*. Cleavage of certain substrates could only be detected in crude PLE though, leading to the assumption that another PLE isoenzyme must be responsible for the relevant activity. Our laboratory collected more information by cloning and sequencing of several different PLE-isoforms and functionally expressing them in *P. pastoris* [10]. A new isoenzyme was discovered, henceforth known as alternative pig liver esterase (APLE) with a mass of 58 kDa and a size of 548 amino acids. It differs from γ-PLE in 21 amino acid residues. Furthermore, this study revealed that APLE is highly stable over a period of several months when stored at 4°C. It could clearly be demonstrated that recombinant APLE is retained in the periplasm of yeast, possibly due to its large size by the formation of trimers (~174 kDa) [10].

1.4 Integration of heterologous DNA into the genome of *P. pastoris*

DNA double strand breaks (DSBs) are considered to be most harmful injuries to DNA as they may ultimately lead to loss of chromosomes and/or cell death. Exogenous agents such as ionizing radiation, a range of different chemical compounds or free radicals can cause DSB as well as they may appear in various recombination events during the cell cycle [11, 12]. In order to maintain genome integrity and stability three major pathways exist in yeast for recombining DSBs: First an event called homologous recombination sequence homology is used to align and join DNA ends. Secondly when sequence homology is absent broken ends are drawn together and joined in a process known as non-homologous end joining (NHEJ). A third process called single-strand annealing uses generation and annealing of 3' single-stranded DNA [12, 13].

Homologous recombination (HR)

The exchange of DNA sequences between homologous DNA molecules is a process known as homologous recombination. It was first discovered in the prokaryotic organism *E. coli*, but its core mechanism was found to be similar in eukaryotes, however in a more complex fashion.

Homologous recombination can be divided into different basic steps, as suggested by Whyman et al. [14]. Upon DSB formation, the ends are processed by a specific nuclease, forming 3' single stranded tails. In a first step one of these tails transacts with a homologous intact piece of DNA forming a joint molecule and generating a D-loop structure. Both 3' ends can use their binding partner as template synthesizing new DNA in a process followed by ligation, thereby forming two four-stranded branched structures, also known as Holliday junctions. By branch migration these Holliday junctions can move freely alongside the DNA influencing the size of heteroduplexed DNA. In order to realize recombination these DNA transactions have to be performed in a very accurate and coordinated manner. During the last step structure-specific endonucleases (resolvases) cleave either the crossed or non-crossed strands of the Holliday junctions leaving either crossover or non-crossover products [14, 11].

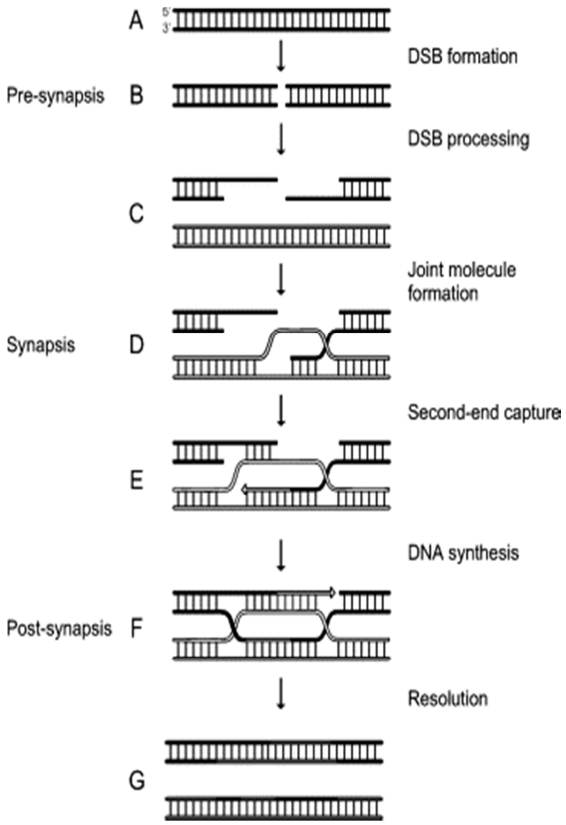


Figure 2: Sequence homology is used to restore genome integrity after double-strand breaks in a process called homologous recombination [14].

Non-homologous end joining (NHEJ)

During repair of DSBs by NHEJ, in a first step limited processing of the DNA-ends takes place, followed by joining and relegation requiring little or no sequence homology at the ends. This mechanism was first discovered in mammalian cells and since then was investigated thoroughly. In the meantime functional homologues were discovered in yeast suggesting that NHEJ is largely conserved in eukaryotes [15]. Broken DNA ends are recognized by Ku-proteins binding at DSBs and recruiting a complex of three proteins, namely Mre11, Rad50 and Xrs2. This complex was found to be forming a bridge and by the aid of a NHEJ-specific endonuclease and polymerase ends are trimmed to repair the lesion by a specific DNA-Ligase [15]. As a major drawback of this mechanism correct restoration of the original sequence is often hampered by loss or gain of nucleotides [13].

HR and NHEJ are used in integration events of heterologous DNA

Episomal plasmid systems are not ultimateley stable in *P. pastoris* rendering integration of foreign DNA into the host's genome necessary for industrial application. Integration of whole expression cassettes will depend on the mechanisms described above (HR and NHEJ).

Mostly, integration at a specific locus is desired, for example in the *HIS4* or the *AOX1* locus. This is achieved by homologous recombination making this process the most sought-after one. Several *P. pastoris* cloning vectors carry respective regions of homology in their backbone, such as pPIC9 and pGAPZ(α)A/B/C. After transformation of one of these (linearized) plasmids into *P. pastoris* in an event called single cross-over homologous regions of the plasmid and the genome align and the heterologous DNA gets inserted.

In an alternative situation called double cross-over expression vectors can be restricted in a way that the expression cassette and a marker gene are liberated. Two regions of homology, e.g. a 5' and 3' *AOX1* region, flanking this sequence of interest allow alignment leading to a gene replacement event in which the *AOX1* gene is deleted and replaced by the expression cassette. The absence of the *AOX1* gene leads to cells showing a MutS (methanol utilization slow, decelerated growth on methanol as sole carbon source) phenotype when grown on methanol as sole carbon source. The *AOX2* gene is transcriptionally weaker. The MutS property makes screening for positive transformants easy by plating them on methanol media and selecting for strains growing more slowly than the wild type [3, 16].

Depending on the length of the flanking homologous regions the frequency of gene replacement events in *P. pastoris* ranges from 0.1% (<500 bp) up to 30% in cases of extensive 1 kb regions of homology [17]. Success also seems to depend on the locus of integration, the *HIS4* locus and the *AOX1* locus being exceptionally easy to target. Nevertheless, homologous recombination is always accompanied by NHEJ limiting the efficiency of the generation of specific target knock-outs. Näätsaari et al. [18] showed that the Ku70p homologue in *Pichia pastoris*, a conserved DNA end-binding protein, is essential for NHEJ. Engineered *P. pastoris* strains lacking Ku70p have a strongly elevated efficiency in homologous recombination because of reduced occurrence of random integrations via NHEJ [18].

In this work, the use of Ku70p deficient strains was avoided although target knock-outs would potentially have been easier to establish. Possibilities of effects on protein expression and/or secretion have not been ruled out for ku70 deficient strains and would narrow the significance of the present study.

2 Objectives

The overall aim of this Master thesis project was to find and characterize possible target genes having an influence on expression and secretion of heterologous proteins in *P. pastoris*. Alternative Pig Liver Esterase (APLE) and the Horseradish Peroxidase (HRP) were chosen as model enzymes.

A *Pichia pastoris* GS115 *his4* strain was transformed with an expression cassette carrying the gene of the heterologous protein (HRP or APLE) fused with an alpha mating factor as a signal sequence for secretion. This cassette was introduced into the genome by homologous recombination at the *AOX1* locus leading to a MutS phenotype. The following strategy was pursued in order to create a library of clones carrying different mutations in their genome with a possible impact on protein secretion behavior: A DNA sequence containing a zeocin resistance gene but without any homologous region to the genome (see Figure 3) was transformed into the cells with the objective of obtaining a great number of independent mutants. In the best case the resistance cassette should integrate randomly.

As described above, this method is taking advantage of the process called “non homologous end joining” as no homologous regions to the host genome are present in the transformed part of DNA.

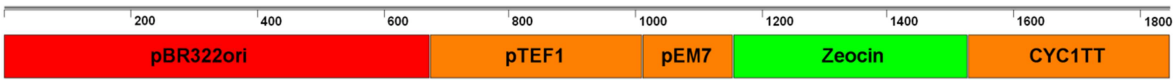


Figure 3: Integrated DNA sequence for random insertion into the genome of *P. pastoris*.
pBR322ori: replicon sequence, allows plasmid maintenance in *E. coli*
pTEF1: eukaryotic promoter region
pEM7: prokaryotic promoter region
Zeocin: Zeocin resistance gene
CYC1TT: transcriptional terminator

By doing an activity screening for the respective heterologous protein in the supernatant of cell cultures, mutants with an increased or decreased secretion behavior were identified.

The first goal was the identification of the locus of integration of the zeocin resistance cassette by genome walking experiments. In most of the cases, the gene most likely causing

the observed phenotyp could be determined. In some cases, though, the cassette did not hit any coding sequence but some untranslated region, which are most likely involved in transcriptional control of neighbouring genes.

After identifying genes supposedly affecting secretion efficiency, the ones not known yet to be of importance for protein secretion were selected for further research. This included in a first step a verification of the screening results by doing a directed target knock-out in the parenting strain. If targeted knock-out of the gene found in the screening leads to the same increase or decrease in secretion efficiency, the results of the screening and genome walking experiments will be confirmed.

In this thesis, the identification process of affected genes in 57 different mutants by genome walking experiments, the design and cloning of 18 target knock-out plasmids and 4 target gene knock-outs in *Pichia pastoris* including screening for positive mutants is described.

3 Materials and Methods

3.1 Strains, plasmids, primers

For cloning purposes and construction of diverse knock-out plasmids the *E. coli* TOP10 F' strain was used. The *P. pastoris* strain GS115 was used for random integrations of a zeocin-resistance cassette and following genome walking experiments. Knockouts of target genes were performed in the *P. pastoris* CBS7435 strain background.

Table 1: Basic strains used

Strain	Genotype	Source
E. coli TOP10 F'	F´{lacIq Tn10 (TetR)} mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL endA1 nupG	Invitrogen
P. pastoris GS115	his4 (Mut+ phenotype)	Invitrogen
P. pastoris CBS7435 Pp004	his4 (Mut+ phenotype)	Winkler C.M.

Table 2: Plasmids used For information about plasmid construction see section 3.3.3

#	Plasmid	Features	Source
1	pPIC9	pBR322ori, AOX1 promoter, MFα sig.seq., HIS4 marker, Amp. resistance cassette	Invitrogen
2	pGAPZalphaA	GAP promoter, MFalpha sig.seq, zeocin resistance marker	Invitrogen
3	pPpKC1	Flippase gene under *AOX1* promotor, 2 FRT regions, pUC ori, Zeocin resistance gene	Ahmad M.
4	pPpKC1 Dus1	3' and 5' homologous regions of Dus1 gene	this work
5	pPpKC1 Rim101	3' and 5' homologous regions of Rim101 gene	this work +Ahmad M.
6	pPpKC1 Sgt2	3' and 5' homologous regions of Sgt2 gene	this work
7	pPpKC1 hyp.prot. H7	3' and 5' homologous regions of hyp.prot. H7 gene	this work
8	pPpKC1 hyp.prot. H8	3' and 5' homologous regions of hyp.prot. H8 gene	this work
9	pPpKC1 Nam2	3' and 5' homologous regions of Nam2 gene	this work
10	pPpKC1 Gtr1	3' and 5' homologous regions of Gtr1 gene	this work
11	pPpKC1 Erd1	3' and 5' homologous regions of Erd1 gene	this work
12	pPpKC1 Cyc8	3' and 5' homologous regions of Cyc8 gene	this work +Ahmad M.
13	pPpKC1 Cat8	3' and 5' homologous regions of Cat8 gene	this work +Ahmad M.
14	pPpKC1 Mxr1	3' and 5' homologous regions of Mxr1 gene	this work +Ahmad M.
15	pPpKC1 Spt8	3' and 5' homologous regions of Spt8 gene	this work +Ahmad M.
16	pPpKC1 Ena1	3' and 5' homologous regions of Ena1 gene	this work +Ahmad M.
17	pPpKC1 Pas1	3' and 5' homologous regions of Pas1 gene	this work +Ahmad M.
18	pGaHSwaI	HRP expression cassette for transformation into AOX1 locus	Ahmad M.

Table 3: *P. pastoris* strains established for HRP and APLE secretion screening

#	Strain labeling	Relevant features	Source
1	Pp013 *P. pastoris* GS115 pPIC9-HRP	HRP protein with MFalpha sig.seq. under GAP promoter	Winkler C.M.
2	Pp015 *P. pastoris* GS115 pPIC9-APLE wt	APLE wild type protein with MFalpha	Winkler C.M.

		sig.seq. under GAP promotor	
3	*P. pastoris* CBS7435 ΔRim101	Rim101-knockout strain, dhis4	this work
4	*P. pastoris* CBS7435 ΔRim101 P_{GAP}-HRP	see #4 + integrated expression cassette	this work
5	*P. pastoris* CBS7435 ΔSgt2	Sgt2-knockout strain, dhis4	this work
6	*P. pastoris* CBS7435 ΔSgt2 P_{GAP} -HRP	see #6 + integrated expression cassette	this work
7	*P. pastoris* CBS7435 ΔKcs1	Hyp.prot.H7-knockout strain, dhis4	this work
8	*P. pastoris* CBS7435 ΔKcs1 P_{GAP} -HRP	see #8 + integrated expression cassette	this work
9	*P. pastoris* CBS7435 Δhypprot.H8	Hyp.prot.H8-knockout strain, dhis4	this work
10	*P. pastoris* CBS7435 Δhypprot.H8 P_{GAP} -HRP	see #10 + integrated expression cassette	this work
11	*P. pastoris* CBS7435 ΔNam2	Nam2-knockout strain, dhis4	this work

Table 4: HRP secreting MutS strains

#	Strain labeling	Relevant features	Source
1	CBS7435 P_{GAP} HRP A1	promoter: GAP, secretory protein: HRP	this work
2	CBS7435 dku70 P_{GAP} HRP C1	promoter: GAP, secretory protein: HRP, dku70	this work
3	CBS7435 P_{GAP} APLE E2	promoter: GAP, secretory protein: APLE	this work
4	CBS7435 dku70 P_{GAP} APLE G3	promoter: GAP, secretory protein: APLE, dku70	this work
5	CBS7435 P_{AOX1} HRP A1	promoter: AOX1, secretory protein: HRP	this work
6	CBS7435 dku70 P_{AOX1} HRP C3	promoter: AOX1, secretory protein: HRP, dku70	this work
7	CBS7435 P_{AOX1} APLE E2	promoter: AOX1, secretory protein: APLE	this work
8	CBS7435 dku70 P_{AOX1} APLE G4	promoter: AOX1, secretory protein: APLE, dku70	this work

Table 5: Primers used

Name	Sequence	Design/Source
3UTRDus1F	TCGGCCGATCAGGCCTCTCAATCTGGTTCAACATTGCCAG	this work
3UTRDus1R	GCCTTATTGGGGTAAATATTGTGGATTCCTAGCGGGG	this work
5UTRDus1F	CTAGGAATCCACAATATTTACCCCAATAAGGCAACGTGG	this work
5UTRDus1R	TCGGCCCTAGTGGCCTCGGCGACAAGCCTAGC	this work
Up5UTRDus1F	AACGTGTCGTCACCCGTAGA	this work
Down3UTRDus1R	GGGTCGCCTGGTGAATTAGC	this work
3UTRSgt2F	TCGGCCGATCAGGCCGGCGGTGATAAGAAGCCTTAAATTTATAATCTTTCT	this work
3UTRSgt2R	CCTGGAAGAGCATGAATATTATGTTCGTTAAGGTTAATTCGGTTTGTAGCT	this work
5UTRSgt2F	CCTTAACGAACATAATATTCATGCTCTTCCAGGAAACGTTACAAATAA	this work
5UTRSgt2R	TCGGCCCTAGTGGCCGTTTCCCTCCAGCTTGAAAGCTTC	this work
Up5UTRSgt2F	CTTGGAGACCAACTGCATAATATGGG	this work
Down3UTRSgt2R	CCGAACTCGTTTCTCAACTACAAGATC	this work
3UTRhypprotH8F	TCGGCCGATCAGGCCCGGAAGTGATACTAAATTTGAATATGGAAGGGC	this work
3UTRhypprotH8R	AATATCACTATGATCTTAAGTGAATTTAAATTATCTAGCTTTCCCAGGTACGCTC	this work
5UTRhypprotH8F	GGAAAGCTAGATAATTTAAATTCACTTAAGATCATAGTGATATTATAACTCAATCCTCC	this work
5UTRhypprotH8R	TCGGCCCTAGTGGCCCAACTCCTTCCAAGAATGATGTAAATACCTACC	this work
Up5UTRhypprotH8F	CCTTCGCAGTATATCTACCCAGGC	this work
Down3UTRhypprotH8R	CCCAGCTTGTCATCCTTGTCG	this work
3UTRhypprotH7F	TCGGCCGATCAGGCCCATAATACGCTGTATAATACATAATAATACAAAGAACTAGCCAT	this work
3UTRhypprotH7R	AGAAATCGTCCCGGGAAAGTCTCCCAGTTGACTAACTTTAC	this work
5UTRhypprotH7F	AGACTTTCCCGGGACGATTTCTCCGAGAACATCAA	this work
5UTRhypprotH7R	TCGGCCCTAGTGGCCTAGTAGCTCTGGATGACGCATCTCTAT	this work
Up5UTRhypprotH7F	GGGATCAGTTTACTGTATACCCAATCTTTGG	this work
Down3UTRhypprotH7R	CCTTATGCGCACTTGCTATCTCAAC	this work
3UTRNam2F	TCGGCCGATCAGGCCGAAGAAGGGCCTAATGGGAGTGTA	this work
3UTRNam2R	GTCTGCCAAGATTCAATTTAAATTCTGTTGAATTTTCTCCGAGTCTTGG	this work
5UTRNam2F	GAAAATTCAACAGAATTTAAATTGAATCTTGGCAGACAAAGATCAGAATC	this work
5UTRNam2R	TCGGCCCTAGTGGCCTATATGAAGCTTACAAACAACATTGCCGATAATG	this work
Up5UTRNam2F	GCCAATGGGAACAAACCCTTCG	this work
Down3UTRNam2R	CCCTAGTTTGAACGATACACCATGCC	this work
3UTRGtr1F	TCGGCCGATCAGGCCGAAAAATTTCGATATCCTTGTTAACCTCCTTACTTTC	this work
3UTRGtr1R	GTACTCGTCTATTTAAATACCATAGCTGCGAACTTATCTTCCTC	this work
5UTRGtr1F	GTTCGCAGCTATGGTATTTAAATAGACGAGTACGTCCTTGATATGGTC	this work
5UTRGtr1R	TCGGCCCTAGTGGCCCAAGGTGTCTAGTGTGTATAGAGCAGG	this work
Up5UTRGtr1F	GGCCATAACTTGTGGAATGACACAG	this work

Down3UTRGtr1R	GGAGAAGTGGTATTGACATCAGCTCC	this work
3UTRErd1F	TCGGCCGATCAGGCCATTTTCTCTATTCTATCTCATTCAGAATCAGTAATGGGTC	this work
3UTRErd1R	CCTCCAGTGAGAATTTAAATAATACATCACGTGACACTAGTTCTTCGTTAC	this work
5UTRErd1F	GTCACGTGATGTATTATTTAAATTCTCACTGGAGGTAAACGATACAAGC	this work
5UTRErd1R	TCGGCCCTAGTGGCCTCAGAGCTATTGCGTACTAAGCTAATGTTC	this work
Up5UTRErd1F	GCAACGAAGTAACTTAGAGACTGCTACC	this work
Down3UTRErd1R	GGGATGGTTCTAGTCAAAACGAAAAAATGC	this work
Ori_BamHIfw	GTGTGGATCCCCCGTAGAAAAGATCAAAGG	Winkler C.M.
Ori_BamHIrev	TTATGGATCCCATGTGAGCAAAAGGCCAG	Winkler C.M.
CYC1PstIrev	TCGCCTGCAGAGCTTGCAAATTAAAGCCTTCGAG	Winkler C.M.
GSPTEFa	TTCCAAACCTTTAGTACGGGT	Winkler C.M.
GSPTEFb	GCTGTGCTAGGGTGTTTTGAA	Winkler C.M.
GSPTEFc	TTTATTGCCTTTTTCGACGAAG	Winkler C.M.
GSPCYC1a	GAGTTAGACAACCTGAAGTCT	Winkler C.M.
GSPCYC1b	GTACAGACGCGTGTACGCATG	Winkler C.M.
GSPCYC1c	GTAATTAGTTATGTCACGCTTA	Winkler C.M.
CSF27	GACGCGTAATACGACTCACTATAGGGC	Winkler C.M.
CSR30	ATCTCCCTATAGTGAGTCGTATTACGCGTC	Winkler C.M.
CSF28	GACGCGTAATACGACTCACTATAGGGC	Winkler C.M.
CSR31	TACGCCCTATAGTGAGTCGTATTACGCGTC	Winkler C.M.
Genome Walker Adaptor	GTAATACGACTCACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGT	Winkler C.M.
AP1	GTAATACGACTCACTATAGGGC	Winkler C.M.
AP2	ACTATAGGGCACGCGTGG	Winkler C.M.
CP	ACGCGTAATACGACTCACTATAGGGCGTAC	Winkler C.M.
CPII	ACGCGTAATACGACTCACTATAGGGCGTAC	Winkler C.M.
pBR322seq	CTATGGAAAAACGCCAGCAAC	Winkler C.M.
GSPpUCa	GTAACAGGATTAGCAGAGCGA	Winkler C.M.
GSPpUCb	CCTTCGGAAAAAGAGTTGGTA	Winkler C.M.
P_{GAP}_cmw_fw	CATGTCATGAGATTATTGGAAACCA	Winkler C.M.
Adaptor	AGGGGCTGGT	Winkler C.M.
Adaptor_rev_TTAA	AATTACCAGCCCCT	Winkler C.M.
Adaptor_rev_TCGA	AGCTACCAGCCCCT	Ahmad M.
5'UTRAox1F	CTTTGATGCCTGAAATCCCAGCGCCTACAATGATGACA	Ahmad M
3'UTRAox1R	CCGTTCGGTATTAGAATTTGTGACTAACAGTGTTCTTAC	Ahmad M
Arg4TTF	GATCTCCTGAGACAAAGTTCACGGGTATCTAG	Ahmad M
AlphaFSSR	GGCAAAACAGCAACATCGAAATCCCCTTC	Ahmad M

3.2 Instruments, reagents, media

Table 6: Instruments and materials used

Application	Instrument	Manufacturer
Electrotransformation	MicroPulserTM	BIO-RAD, USA
Mixing	Vortex-Genie 2	Scientific Industries Inc, USA
OD600 measurements	BioPhotometer	Eppendorf, Germany
PCR	GeneAmp®PCR System 2700	Applied Biosystems, USA
Shaker	HT MiltronII	Infors AG, Swiss
Centrifuges	Centrifuge 5810	Eppendorf, Germany
	Centrifuge 5415R	Eppendorf, Germany
Absorption measurements	Nanodrop 2000c	Thermo Fisher Scientific Inc, USA
	SpectraMax Plus384	Molecular Devices, USA
	FLUOstar Omega	BMG Labtech, Germany
Thermomixer	Thermomixer comfort	Eppendorf, Germany
Agarose Electrophoresis instruments	Sub-cell® GT	Bio-Rad Laboratories GmbH, Vienna, Austria
pH-meter	InoLab pH720	WTW GmbH, Weilheim, Germany
Scale	TE Präzisionswaage TE1502S	Sartorius AG, Göttingen, Germany
Magnetic stirrer	Magnetic Stirrer Model L-71	Hartenstein, Würzburg, Germany
Certoclave	Certoclave	Certoclave Sterilizer GmbH, Traun, Austria
Blotting, filter assays	Filter paper	Whatman International Ltd., GB
Deep Well Plates	SCIENCEWARE 96 Deep-Well Plate	Bel-Art, Wayne, USA
Microtiter Plates	Micro-Plate, TC, sterile	Greiner Bio-One GmbH, Frickenhausen, Germany
Membrane for Southern Blots	Nitrocellulose Membran, 0,2 µm	Roth GmbH, Germany

Table 7: Reagents

Reagent	Supplier
ABTS	Boehringer Ingelheim Pharma GmbH & Co. KG, Germany
Agar	Bacto Laboratories Pty Ltd, Australia
Agarose	Biozyme, Germany
Ammonium acetate	Roth GmbH, Germany
Biotin	Roth GmbH, Germany
Citric Acid	Roth GmbH, Germany
Dimethylsulfoxid	Roth GmbH, Germany
dNTPs	Roth GmbH, Germany
DreamTaq Green DNA Polymerase	Thermo Fisher Scientific Inc, USA
EDTA	Roth GmbH, Germany
Ethanol	Roth GmbH, Germany
Gene Jet Plasmid Miniprep Kit	Thermo Fisher Scientific Inc, USA
Glucose	Roth GmbH, Germany
LB (Luria-Bertani)	Roth GmbH, Germany
Maleic Acid	Roth GmbH, Germany
Maxima Hot Start Green PCR Mastermix	Thermo Fisher Scientific Inc, USA
Methanol	Roth GmbH, Germany
N-Laurylsarcosine	Sigma-Aldrich, St.Louis, USA
NBT/BCIP	Merck, Germany
O'Generuler DNA ladder Mix	Thermo Fisher Scientific Inc, USA
Peptone	
Phusion® Hot Start High-Fidelity DNA Polymerase	Finnzymes, Finnland
Potassium dihydrogen phosphate	Roth GmbH, Germany
Potassium hydrogen phosphate trihydrate	Roth GmbH, Germany
Q5 Polymerase	New England Biolabs Inc., Ipswich, USA
Restriction enzymes	Thermo Fisher Scientific Inc, USA
SDS	Roth GmbH, Germany
Sodium chloride	Roth GmbH, Germany
Sodium hydroxide	Roth GmbH, Germany
T4 DNA Ligase	Thermo Fisher Scientific Inc, USA
TRIS	Roth GmbH, Germany
Tri-Na-citrate	Roth GmbH, Germany
Tween 20	Roth GmbH, Germany
Wizard SV Gel and PCR Clean Up System	Promega Corporation, USA
Yeast extract	Bacto Laboratories Pty Ltd, Australia
Yeast nitrogen base	Difco- Becton, USA
Zeocin	Cayla-InvivoGen, Toulouse, France
Klenow Fragment	Thermo Fisher Scientific Inc, USA
Calcofluor white	Sigma-Aldrich, St.Louis, USA

Table 8: Media and buffers

Medium/Buffer	Composition
10x ABTS stock	5.5 g/L
APLE Buffer	100 mM Tris-HCl, pH 7.0
BEDS	5.0% DMSO, 3.0% ethylene glycol, 1 M sorbitol, 10 mM bicine
Biotin (500x)	200 mg/L
BYPD	same as YPD + 200 mM phosphate buffer pH 6
Calcofluorwhite plates	YPD + 10µg/mL CW
DTT (1M)	154,25 g/L
HRP substrate solution	800 mL/L Na-acetate buffer, 200 mL/L ABTS (10x), 175 µL/L ddH_2O_2
LB-medium (+agar)	10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl (+20 g/L Agar)
MD-medium (Agar)	100 mL YNB (10x), 2 mL biotin (500x), 20 g/L glucose (20 g/L Agar)
MM-medium (Agar)	100 mL YNB (10x), 2 mL biotin (500x), 10 mL EtOH, (20 g/L Agar)
Na-acetate buffer	50 mM Na-acetate, pH 4.5
Phosphate buffer	1 M (K_2HPO_4 30 g/l, KH_2PO_4 118 g/l) pH 6 set with KOH
SOC-medium	20 g/L bacto tryptone, 0.58 g/L NaCl, 5 g/L bacto yeast extract, 2 g/L $MgCl_2$, 0.16 g/L KCl, 2.46 g/L $MgSO_4$, 3.46 g/L dextrose
Tris-HCl	100 mM Tris, pH 7, set with HCl
YPD (Agar)	10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose, (20 g/L Agar)

½ YPD	5 g/L yeast extract, 10 g/L peptone, 10 g/L glucose
YNB (10x)	134 g/L
Working Solutions for Southern Blots	
Blocking Stock Solution (10%=10x, 100 mL)	5g Blocking Reagent in Maleic Acid buffer (pH 7.5)
SDS solution (10% in 20 mL)	2 g SDS in H_2O
N-laurylsarcosine sodium salt (10%, 20 mL)	5.7 mL N-laurylsarcosine 35% in H_2O
Hybridization Solution (200 mL)	50 mL SSC 20x, 20 mL blocking stock solution, 2 mL N-laurylsarcosine, 0.4 mL SDS 10%, in H_2O
Washing Solution 1 (500 mL)	50 mL SSC 20x, 5 mL SDS 10%, in H_2O
Washing Solution 2 (500 mL)	2.5 mL SSC 20x, 5 mL SDS 10%, in H_2O
Maleic Acid Buffer (pH7.5, 1 L)	11.6 g Maleic Acid , 8.77 g NaCl+NaOH till pH 7.5, in H_2O
Washing buffer 3 (500 mL)	500 mL Maleic Acid Buffer, 1.5 mL Tween 20
Detection Buffer (pH 9.5, 100 mL)	1.21 g Tris, 0.58 g NaCl + HCl till pH 9.5, in H_2O
Color-substrate solution	200 µL NBT-BCIP stock solution, 10 mL detection buffer
TE-Buffer (pH 8.0, 1 L)	1.25 g Tris, 0.4 g EDTA +HCl till pH 8.0 in H_2O
HCl (250mM, 500 mL)	10 mL 37% HCl in H_2O
Denaturing Solution (500 mL)	10.0 g NaOH, 43.83 g NaCl, in H_2O
Neutralization Solution (1 L, pH 7.4)	121.14 g Tris, 87.66 g NaCl +HCl till pH 7.4, in H_2O
Transfer Buffer (20x SSC, pH 7.0, 2x2 L)	350.64 g NaCl, 176.46 g tri-Na-citrate*2 H_2O + Citric Acid till pH 7.0, in H_2O
Equilibration Buffer (500 mL)	250 mL transfer buffer (20x SSC) in H_2O

Table 9: Restriction enzymes used

Restriction enzyme	Recognition Sequence	Source
*Acc*65I	5'…G ↓ G T A C C…3' 3'…C C A T G ↑ G…5'	Thermo Fisher Scientific Inc, USA
*Bam*HI	5'…G ↓ G A T C C…3' 3'…C C T A G ↑ G…5'	Thermo Fisher Scientific Inc, USA
*Bcl*I	5'…T ↓ G A T C A…3' 3'…A C T A G ↑ T…5'	Thermo Fisher Scientific Inc, USA
*Bcu*I	5'…A ↓ C T A G T…3' 3'…T G A T C ↑ A…5'	Thermo Fisher Scientific Inc, USA
*Bgl*II	5'…A ↓ G A T C T…3' 3'…T C T A G ↑ A…5'	Thermo Fisher Scientific Inc, USA
*Box*I	5'…G A C N N ↓ N N G T C…3' 3'…C T G N N ↑ N N C A G…5'	Thermo Fisher Scientific Inc, USA
*Bsp*OI	5'…G ↓ C T A G C…3' 3'…C G A T C ↑ G…5'	Thermo Fisher Scientific Inc, USA
*Eco*RI	5'…G ↓ A A T T C…3' 3'…C T T A A ↑ G…5'	Thermo Fisher Scientific Inc, USA
*Eco*RV	5'…G ↓ A T A T C…3' 3'…C T A T A ↑ G…5'	Thermo Fisher Scientific Inc, USA
*Eco*147I	5'…A ↓ G G C C T…3' 3'…T C C G G ↑ A…5'	Thermo Fisher Scientific Inc, USA
*Hind*III	5'…A ↓ A G C T T…3' 3'…T T C G A ↑ A…5'	Thermo Fisher Scientific Inc, USA
*Pst*I	5'…C ↓ T G C A T…3' 3'…G A C G T ↑ A…5'	Thermo Fisher Scientific Inc, USA
*Sac*I	5'…G ↓ A G C T C…3' 3'…C T C G A ↑ G…5'	Thermo Fisher Scientific Inc, USA
*Sal*I	5'…G ↓ T C G A C…3' 3'…C A G C T ↑ G…5'	Thermo Fisher Scientific Inc, USA
*Sfi*I	5'…G G C C N N N N ↓ N G G C C C…3' 3'…C C G G N ↑ N N N N C C G G…5'	New England Biolabs Inc., Ipswich, USA

*Sma*I	5'…C C C ↓ G G G…3' 3'…G G G ↑ C C C…5'	Thermo Fisher Scientific Inc, USA
*Ssp*I	5'…A A T ↓ A T T…3' 3'…T T A ↑ T A A…5'	Thermo Fisher Scientific Inc, USA
*Swa*I	5'…A T T T ↓ A A A T…3' 3'…T A A A ↑ T T T A…5'	Thermo Fisher Scientific Inc, USA
*Xba*I	5'…T ↓ C T A G A…3' 3'…A G A T C ↑ T…5'	Thermo Fisher Scientific Inc, USA

3.3 Methods

3.3.1 General methods

3.3.1.1 General PCR conditions

PCRs were performed in either 25 µL or 50 µL of total volume. For routine PCRs the Hot Start Green PCR Master Mix was used, which already contains dNTPs, buffer and loading dye besides a Taq Polymerase. For cloning and engineering work a high fidelity polymerase such as Phusion or Q5 was used.

Table 10: Standard PCR conditions

Condition	Taq Polymerase	Phusion/Q5 Polymerase
Denaturation (initial)	95°C, 4 min	98°C, 30 s
Denaturation	95°C, 30 s	98°C, 10 s
Annealing	Temperature according to Tm of primers, 50 s	Temperature according to Tm of primers, 30 s
Extension	72°C, 1 min/kb	72°C, 25 s/kb
Extension (final)	72°C, 10 min	72°C, 10 min
Number of cycles	35	35

3.3.1.2 Colony PCR

Very little fresh cell material was picked from Agar plates with yellow pipette tips or toothpicks and resuspended in 10 µL of ddH_2O. For cell disruption the samples were heated to 95°C for 10 min and afterwards shock-frosted in liquid nitrogen to break up the cells. Finally, the samples were centrifuged at max. speed (16200xg) and 4°C for 10 min. One µL of supernatant was used for the subsequent PCR.

3.3.1.3 Gel electrophoresis

For standard control and preparative gels agarose was added to 1xTAE-buffer (~220 mL) to a final concentration of 1% and boiled for complete dissolution. Before pouring the gel 5 µL of ethidium bromide (10 mg/mL) were added. Different combs generating 10 to 30 slots were used according to requirements.

3.3.1.4 Isolation of genomic DNA

For the isolation of the genomic DNA an adapted protocol by Hoffman and Winston [19] was used.

Cell cultures were grown in shake flasks for high yield or in DWP for time efficiency.

Cell cultures in shake flasks:

Starter yeast culture

Five mL of YPD medium were inoculated with the respective *P. pastoris* strains in 50 mL Falcon. All samples and a sterile control were shaken overnight at 28°C and 150rpm.

Main culture

The main cultures were supposed to be inoculated to an OD of 1. Therefore, the starter cultures were grown to an OD of at least 7. YPD medium was added to an appropriate amount of starter culture to a final volume of 50 mL. During incubation on 28°C the cells grew to an OD between 3 and 10. The equivalent volume of 100 OD units (100 OD units = OD_{final} · X mL) were transferred into 50 mL tubes and the cells were spun down for 5 min at 1500 rpm. After decanting the supernatant the cells were resuspended in 0.5 mL sterile water, transferred into Eppendorf tubes and spun down again for 5 s at maximum speed.

Cell cultures in DWP:

Wells were filled with 500 µL of YPD each and were inoculated with *P. pastoris*. The DWP was shaken for 24 h at 320 rpm and 80% humidity. Subsequently, the cultures were pippetted into Eppendorf tubes and spun down for 5 min at 1500 rpm.

At this point the supernatant was decanted and 200 µL of Yeast lysis buffer, 200 µL phenol:chloroform:phenyalkohol (25:24:1) and 0.3 g of acid washed beads were added to the pellet. To obtain sufficient cells lysis the tubes were vortexed for at least 3 min and 400 µL of TE buffer were added eventually. By centrifugation the phases were separated and the aqueous phase was transferred into a new tube. By adding 1 mL of 100% ice-cold ethanol

the genomic DNA (gDNA) precipitated as white coils. Remaining RNA was digested by aspirating and dissolving the residual pellet in 400 µL TE buffer and 5 µL RNAse A (10 mg/mL). After incubation at 37°C for 2 h the gDNA was precipitated again with 10 µL of 4 M ammonium acetate and 1 mL of ice cold 100% ethanol. By centrifugation the pellet was fixed at the bottom of the tube and the supernatant was removed. For high purity of gDNA an additional ethanol washing step could be performed at this point by addition of ~1 mL 70% ethanol, vortexing for about 2 min plus centrifugation and aspiration. In a final step the gDNA was dissolved in 50 µL sterile water.

3.3.1.5 Southern Blot

Southern Blotting is a method to screen for a specific DNA sequence in a complex DNA pool. In this procedure DNA is isolated from each source and then digested with specific restriction enzymes. Twenty to thirty µg of digested gDNA are loaded onto an agarose gel and get separated according to size over night at 20 V. Then the DNA is transferred to a nylon membrane by assembling a device shown in Figure 4.

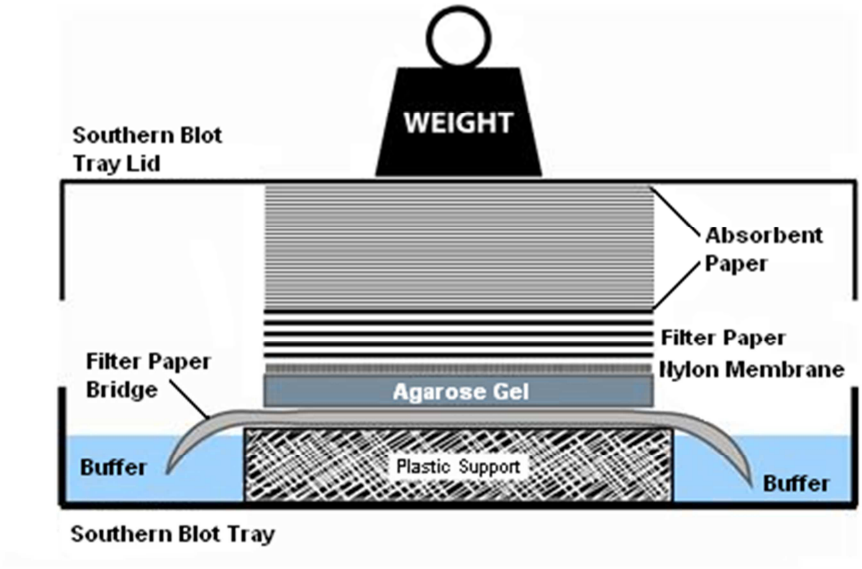


Figure 4: Assembly for Southern Blotting [20]
By capillary transfer the transfer buffer moves along a filter paper bridge and through the agarose gel, transferring DNA which adheres to the Nylon membrane.

A Southern Blot tray containing a plastic support slightly bigger than the agarose gel was filled with Transfer Buffer. Two basic layers of filter paper protruding into the buffer were overlaid by a sponge and another two layers of filter paper. On top of this the agarose gel was placed, carefully not to tear the sensitive structure. A nylon membrane covering the gel was flattened on it avoiding air blister formation in between. Three layers of filter paper and a

stack of absorbent paper form the top which was weighed down. By capillary transfer the buffer was sucked bottom-up and carries DNA molecules with it to be attached to the nylon membrane. After 15 to 20 h the blotting was accomplished and DNA was fixed to the membrane by baking it for 30 min at 120°C.

The following steps are described in more detail in the Instruction Manual by Roche Applied Science [21]. Generation of DIG-labeled probes was done by applying the method of random primed labeling [22] which is based on the hybridization of random oligonucleotides to the denatured DNA template. Random oligonucleotides as primers and a mixture of desoxyribonucleotides containing DIG-11-dUTP form probes according to a template and incorporating DIG-dUTP approximately every 20-25 nucleotides leading to optimal sensitivity in detection reactions.

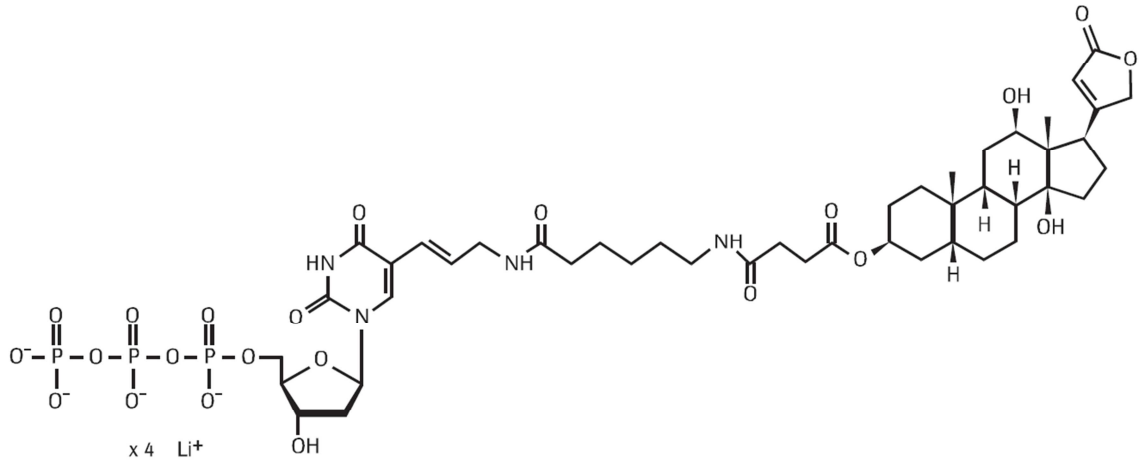


Figure 5: Chemical structure of digoxygenin (DIG)-11-dUTP

The labeled nucleic acid probe is added to the membrane and binds to complementary DNA segments. Excess probe is washed away in multiple washing steps to avoid unspecific results. The hybridized probes are immunodetected with anti-digoxigenin-antibodies and are then visualized with the colorimetric substrates NBT/BCIP.

In the present study Southern Blotting was applied on the one hand for detection of multiple insertions of HRP and APLE expression cassettes during the process of their transformation and integration into the genome of *Pichia pastoris*. To be able to deduce significant results from HRP/APLE activity assays it is of importance that multiple insertions can be ruled as the reason for increased activity.

Moreover, Southern blotting was employed to check for the number of integrated zeocin resistance cassettes in the strains investigated in genome walking experiments (see section

4.1). Genome walking may not yield extensive information about multiple integrations and therefore may lead to wrong assumptions concerning which gene affects secretion behavior. By employing Southern Blotting the number of integrations can be assessed and compared to the loci found in genome walking.

Searching for HRP/APLE expression cassettes was done by using probes binding either in the *GAP* promoter region or in the *HIS4* gene. Zeocin resistance cassettes were detected with a probe binding in the region of the eukaryotic promoter pTEF1.

3.3.1.6 Electrotransformation into E. coli

Fifty µL of electrocompetent *E. coli* TOP10F' cells were thawed on ice and gently mixed with 20-100 ng of heat-inactivated and desalted ligation mix containing plasmid DNA. The mixture was transferred to pre-cooled electroporation cuvettes and kept on ice for at least 5 min. For electroporation the program "EC2" was chosen (2.5 kV, 5-6 ms). Cells were regenerated immediately with 500 µL SOC-medium and kept at 37°C and 600 rpm for 45 min. For plating, the cells were spun down shortly, most of the supernatant was removed and the rest was plated on selective LB-medium.

3.3.1.7 Electrocompetent P. pastoris *cells and Transformation*

For preparation of electrocompetent *P. pastoris* cells a condensed protocol by Lin-Cereghino [23] was used. It combines the most efficient parts of electroporation and heat-shock transformation protocols.

A 5 mL ONC of the respective *P. pastoris* strain was grown and used to inoculate a 50 mL shake flask culture incubated at 28°C and 120 rpm. When the culture reached an OD of 0.8-1.0 cells were harvested by centrifugation for 5 min at 1600 rpm. After the removal of the supernatant the pellet was resuspended in filter-sterilized 9 mL BEDS-solution and 1 mL of 1 M DTT. The mixture was incubated for 5 min at 28°C and 100 rpm before cells were collected by centrifugation as above. After resuspension in 1 mL BEDS aliquots of 80 µL were prepared.

For transformation 800-1400 ng of linearized plasmid DNA were mixed with electrocompetent cells and incubated on ice for 5 min in electroporation cuvettes. Program "Pic" was chosen for pulsing the cells (1.5 kV). Afterwards 500 µL of 1 M sorbtiol were added immediately. After 30 min of incubation at 28°C without shaking, 500 µL of YPD were

added for another hour at 28°C. Finally, aliquots of 100 µL, 200 µL and the rest were plated on selective media.

Figure 6 shows the map of the plasmid carrying the HRP expression cassette for insertion into the genome of *P. pastoris* at the *AOX1* locus. Before transformation the plasmid was linearized by using two *Swa*I restriction sites.

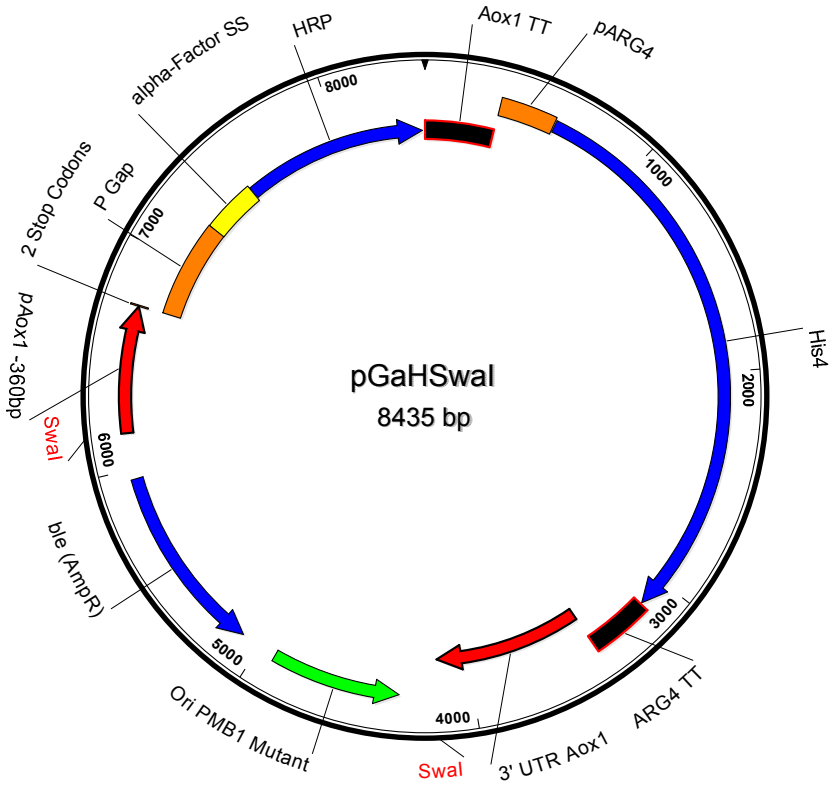


Figure 6: Plasmid pGaHSwaI carrying the HRP expression cassette for transformation into *AOX1* locus of *P. pastoris*.
After linearization via restriction enzyme *Swa*I the construct is flanked by two homologous regions (part of *AOX1* promoter region and 3'UTR AOX1 respectively). The HRP gene is fused to an α-factor signal sequence under the control of a GAP promoter. The *HIS4* gene complements *his4* auxotrophy in *P. pastoris* allowing easy screening on minimal medium.

3.3.1.8 Integration of heterologous DNA at the AOX1 locus and screening for MutS phenotype and histidin prototrophy

After integration of heterologous genes at the *AOX1* locus *P. pastoris* cells depend on their *AOX2* gene when grown on methanol medium deficient of any other carbon sources. As the *AOX2* promoter is comparatively weak, MutS cells grow much slower which is easily visible after 2-3 days.

Part of all expression constructs used in this thesis is a *HIS4* gene reinstalling histidin prototrophy in all dhis4 strains provinding another screening option.

After transformation of expression constructs into the *AOX1* locus of *P. pastoris* screening both for an intact *HIS4* gene as well as for MutS (methanol utilization slow) phenotype was performed allowing simple recognition of positive transformants.

Transformants were streaked out on MD plates and after three days clones were picked, transferred to DWPs filled with a small amount of YPD medium and pinned onto screening medium. Clones growing on MD medium (indicating histidin prototrophy) and showing MutS phenotype on MM medium (indicating integration of heterologous DNA at the *AOX1* locus) were considered to have the desired genotype.

Expression strains were checked by PCR. Therefore, primers were designed binding outside of the *AOX1* locus (5'UTRAox1F and 3'UTRAox1R) (see Figure 7). In wild type strains PCR yields fragments of 4270 bp. After integration of the HRP expression cassette the binding sites for these primers are quite far apart (6639 bp fragments), though. Therefore two additional primers were designed binding in α-factor signal sequence and the *ARG4* transcription terminator sequence of the integrated plasmid (see also Figure 6), both pointing outwards for amplification of two distinct fragments.

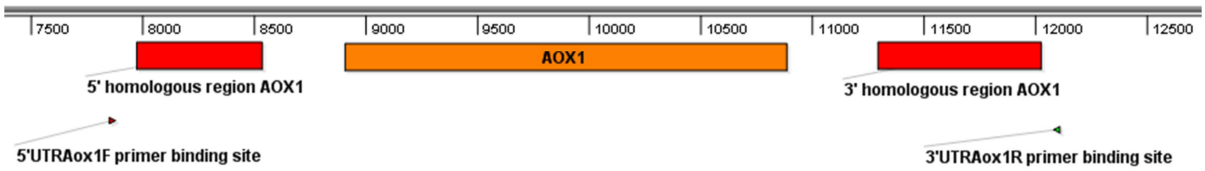


Figure 7: *AOX1* locus and primer binding sites

3.3.2 Enzyme activity measurements and strain characterization

3.3.2.1 Cell cultivation in DWPs and shake flasks for determination of secretion activity

HRP activity measurements in the supernatant were done with samples from two different cell cultivation conditions to examine influences on secretion behavior.

In a first fermentation method, starter cultures in DWPs with 250 µL of BYPD per well were inoculated with cells either from glycerol stocks or from plates and grown at 28°C and 320 rpm at 80% humidity. After 24 h 2 DWPs with 500 µL BYPD per well were inoculated by using a 96-well pin. After another 34-38 h OD was determined in microtiter plates and HRP activity was measured by ABTS assay.

Alternatively, 5 mL starter cultures in 50 mL tubes were grown over night at 28°C and were used to inoculate 50 mL main cultures in 300 mL baffled shake flasks (28°C, 120 rpm) to an OD of 0.1. Samples were taken at different time points between 20 h and 50 h of incubation time.

3.3.2.2 HRP isolation from intracellular fractions

Cells were pelleted by centrifugation for 10 min at 1500xg and room temperature. After decantating the supernatant cells were washed with 500 µL of ddH_2O and pelleted again. By adding 150 µL of 100 mM phosphate buffer, pH 6, to the pellet the cells were resuspended and lysed by addition of 100 µL of acid washed beads and vortexing. In order to minimize denaturation of HRP samples were alternately kept on ice and on a vortex for 30 s each. Together 10 cycles of vortexing and cooling were performed. To separate the intracellular fraction from cell debris samples were centrifuged at maximal speed (16200xg) and 4°C. The supernatant was collected and used for ABTS assays.

3.3.2.3 ABTS Assay

The ABTS assay was used to measure HRP activity in culture supernatants or intracellular fractions. The assay was performed in 96-well microtiter plates. Per plate 7.5 mL of substrate solution were prepared. Cell cultures were centrifuged at 1500g for 10 min. Seventy µL of Na-acetate buffer were provided per well using multichannel pipette and 15 µL of culture supernatant (mostly diluted 1:10 or 1:20 to avoid saturation of absorption curve) were added per well. Seventy µL of substrate solution were added per well right before putting the plate into the photometer as the reaction starts immediately. The absorbance change at 405 nm was measured over 5 min in intervals of 15 s and derived as milli absorption units per minute (mAu/min).

ABTS is an abbreviation for the IUPAC name 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid). This compound is widely used in biochemical investigations as it easily dissociates into radicals which can be detected spectrophotometric at 405 nm. In the HRP assay the peroxidase transfers electrons from hydrogen peroxide to ABTS leading to a change of color from transparent to green [24].

3.3.2.4 pNPA Assay

Esterases such as APLE catalyze the hydrolysis of esters into the corresponding acid and alcohol. In the presence of water APLE cleaves para-nitrophenyl acetate (pNPA) into acetic acid and p-nitrophenol (pNP), leading to a change of color as pNP is bright yellow with an absorption maximum at 405 nm.

Similar to the ABTS assay the pNPA assay was done in 96-well plates. 150 µL of Tris-HCl at pH 7 were provided per well and 10 µL of culture supernatant were added. Shortly before the measurement the APLE screening solution containing pNPA was prepared. Immediately after addition of 140 µL screeing solution per well the 96-well plate was put into the photometer and the increase of absorption was monitored for 20 min.

3.3.2.5 Calcofluor White Spotting

Calcofluor white is a negatively charged fluorescent compound that is not entering cells and stains chitin structures. Studies in *S. cerevisiae* have shown that calcofluor white can be used in screenings for all main types of cell wall mutants except for chitin defective strains [25]. If the same applies in *P. pastoris* cell wall defects could easily be identified. Rim101 knockouts in *S. cerevisiae* in turn have exhibited a relevant increase in resistance to this dye indicating a possible screening method for target knockouts of the homologue in *P. pastoris* [26].

Cell cultures were diluted to an OD of 0.5 in YPD and further dilutions thereof were prepared from $1:10^1$ to $1:10^6$. By using a pipette 10 µL of each dilution were applied on a YPD-plate containing 10 µg/mL of calcofluor white.

3.3.2.6 Growth curves

To detect possible growth defects in any of the knock-out strains growth curves were recorded for 300 mL shake flask cultures. ONCs with 5 mL of BYPD were prepared and inoculated with cells picked from agar plates. The next day the OD was measured in triplicate and 50 mL main cultures were inoculated to an OD of 0.1. For 58 h the OD of each flask culture was determined in short time intervals in triplicates. As the linear range of the photometer is confined samples were diluted up to 100-fold.

3.3.3 Target knock-outs in *P. pastoris*

Knock-outs of genes in *P. pastoris* were performed using knock-out plasmids as first proposed by Reuß et al. [27] and optimized for the use in *P. pastoris* by Näätsaari et al. [18]. A general map of the knock-out plasmids is shown in Figure 8. They contain the following features:

Table 11: Features of target knockout pPKC1 plasmids

Feature name	Function
pUCori	Origin of replication in *E. coli*
ZeoR	Zeocin resistance gene, complemented with both a prokaryotic (pEM72) and a eukaryotic (pARG4) promoter to confer resistance against zeocin to *E. coli* as well as to *P. pastoris*.
FRT	Flippase recognition target, sequence of 34bp recognized by flippase
FLP	Flippase gene under promoter, Intracellularly expressed Flippase recognizes FRT regions in the genome and is capable of cutting out DNA sequences in between (see also Figure 12)
3'UTR and 5'UTR	DNA sequences homologous to flanking regions of target genes in the genome, needed for homologous recombination events

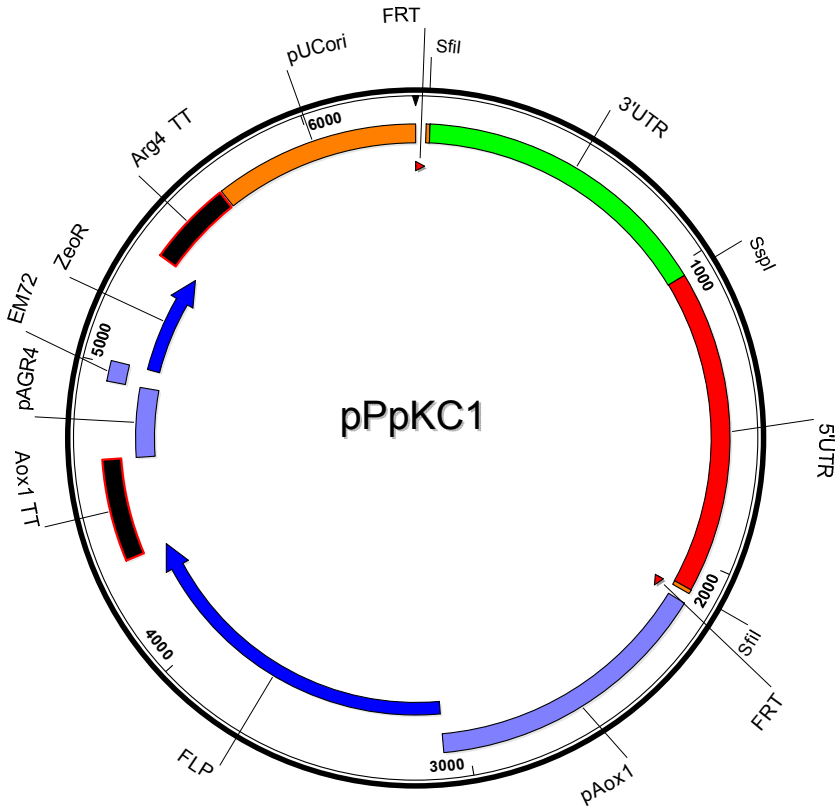


Figure 8: Map of the knock-out plasmid pPpKC1
For information about features of the plasmid please see Table 11.

For constructing different knock-out plasmids a plasmid backbone was ligated with an insert containing the 3' and 5' UTR regions of the target genes (Figure 8). Before cloning the inserts were amplified from the homologous regions of the *P. pastoris* CBS7435 genome and 3' and 5' UTRs were joined by overlap-extension PCR. By digestion with the restriction enzyme *Sfi*I the insert as well as the backbone were prepared for ligation.

3' and 5' UTR regions

The probability of a correct locus-integration in *P. pastoris* by homologous recombination is strongly dependent on the length of the homologous regions, besides other factors like accessibility of the part of genome. In the work presented here a length of about 1000 bp per

upstream and downstream region was chosen. During the process of a double crossover by homologous recombination a part of genomic DNA gets excised and released in the cell leading to a possibility of reintegration of a functional gene elsewhere. To minimize the chance of this event, homologous regions were designed to be partly inside of the target knock-out gene as shown in Figure 9.

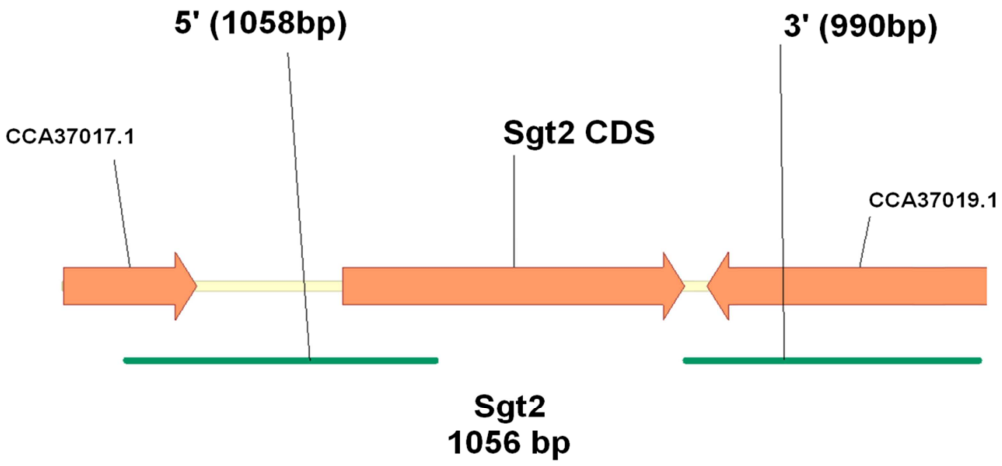


Figure 9: 3‘ and 5‘ homologue regions of knock-out target Sgt2 gene

Furthermore, homologous regions were designed with respect to the following specifications:

As the homologous regions subsequently were cloned into the knock-out plasmid backbone, the restriction sites used therefore must not be found in the sequences. Moreover, in the final plasmid a unique restriction site between the homologous regions is essential for linearization, which was important for their ends pointing towards each other (see Figure 10).

Primer design

Primers were designed as shown in Figure 10 for the 3'UTR of the knock-out target Sgt2.

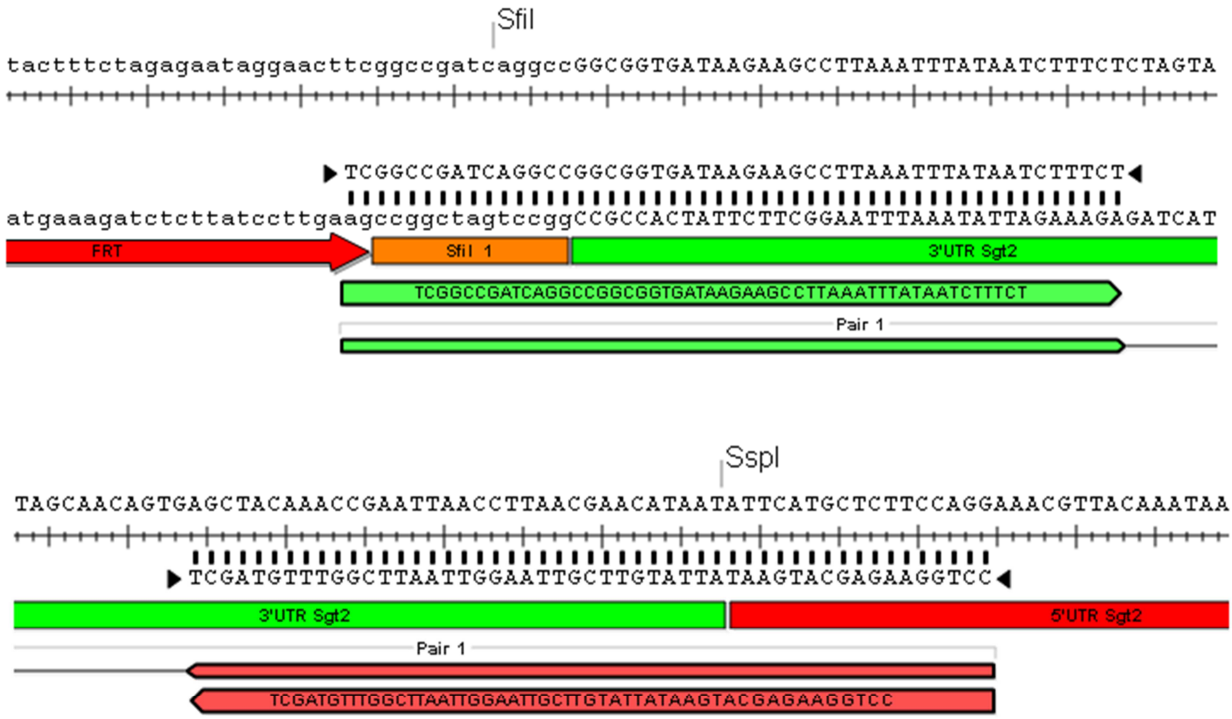


Figure 10: Primer design for amplification of the 3'UTR of *SGT2*. In addition to the 3' homologous region the forward primer carries a *Sfi*I restriction site plus 2 bases for proper recognition by the restriction enzyme later on, the reverse primer is attached to a sequence homologous to the 5'UTR to allow overlap-extension PCR. Mind the appearance of a unique restriction site (*Ssp*I) for linearizing the knockout plasmid before transformation.

Besides a region complementary to the sequence in the genome, primers were provided carrying the following features: For the 3'UTR the forward primer has an additional *Sfi*I restriction site to allow easy cloning into the vector backbone. The reverse primer is provided with a sequence complementary to the 5'UTR for an overlap-extension PCR to fuse the two homologous regions (Figure 10). The length of each part of the primers was chosen with respect to the Tm of the region active in the respective PCR. The same applied for the primers for amplification of the 5'UTR, with interchanged characteristics of forward and reverse primer.

The Tm of primers was calculated using the IDT Oligoanalyzer 3.1 software. A Tm of 58°C was aspired for every PCR.

3.3.3.1 Construction of Knock-out plasmids

Genomic DNA of *P. pastoris* CBS7435 dhis4 was isolated and used as a template for amplification of the 3' and 5'UTRs. After purification by preparative agarose gel equal amounts of the respective fragments were fused by overlap extension (oe-)PCR using the outer primers 3'UTR forward and 5'UTR reverse. Both the vector backbone and the oe-PCR products could now be digested with restriction enzyme *Sfi*I, resulting in compatible vector and insert fragments.

After ligation at a molar ratio of 3:1 (insert:vector) over night the constructs were transformed into *E. coli* TOP10F' cells and plated on LB-Zeo plates. Transformants were picked and colony-PCRs were performed using primers binding in the pUCori and the P_{AOX1} regions each pointing towards the insert. Plasmids with positive results in the colony-PCR were isolated and sequenced by LGC Genomics (Berlin, Germany).

3.3.3.2 Performing knock-outs in Pichia pastoris *and screening for positive transformants*

Knock-out plasmids were linearized and transformed into CBS7435 dhis4. As multiple integration of the cassettes into the genome should be avoided the amount of DNA for transformation was as low as 800 ng per transformation. After transformation cells were plated onto YPD+Zeo medium and incubated at 28°C for 3 days. Single colonies were picked with toothpicks carefully, resuspended in YPD medium in DWPs and pinned onto YPD omnitrays as well as YPD-Zeo omnitrays.

To search for clones with correct locus integration the following strategy was applied: Primers binding outside of the knock-out target were designed to make amplification of the region and sequencing possible. Because of the size of the integrated vector DNA (about 6.4 kb, depending on the size of the homologue regions), it was decided to do separate PCRs for each side as shown in Figure 11.

Figure 11: Situation in the genome after integration of a Sgt2-knockout plasmid in *P. pastoris*. In the flanking region upstream to the integrated sequence a short stretch of the target gene remained to avoid the release of a functional gene during excision. Green and red arrows indicate primer binding sites for screening purposes.

In a first step genomic DNA of 20 clones of each knock-out transformation was isolated and analyzed via PCR of the 5' region. Samples which gave a positive signal in the first PCR were examined in a second PCR with primers for the 3' region. Clones which have shown proper signals in both PCRs were selected for flipping out their zeocin marker.

3.3.3.3 Removal of zeocin-marker

The knock-out plasmids carry a feature, which facilitates its own removal from the genome after excision of the target DNA. A flippase recombinase system is utilized where a Flippase gene under control of the *AOX1* promoter allows excision to remove the integrated cassette. Sequences between the FRT (flippase recognition target) signals are cut out as shown in Figure 12 leaving only one FRT region and a *Sfi*I restriction site. This makes the knock-out as "clean" as possible as it results in very little additional alterations in the genome of the wild type. Moreover, cells lose their resistance against zeocin making this marker available for further genetic engineering of the strains. Last but not least, the removal of the integrated cassette makes sequencing over the whole region possible giving full information about the situation in the genome.

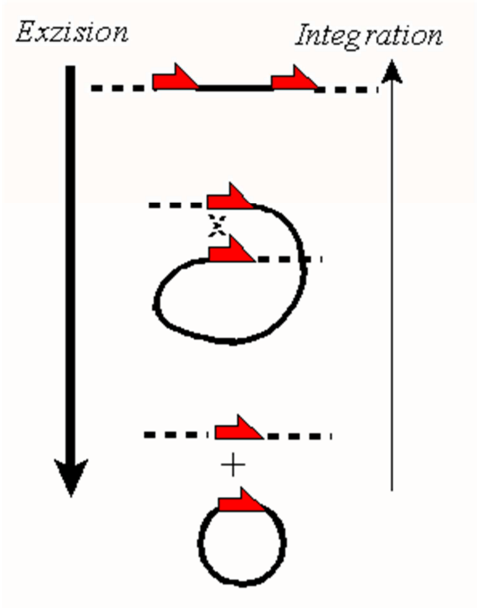


Figure 12: If recombination occurs between two homologous signal sequences (red arrows, FRT regions) the flanked DNA region gets excised as a circular molecule leaving one FRT region in the genome [28].

A colony carrying a knockout plasmid to be removed from the genome was used to inoculate a 50 mL ½ YPD culture and was kept in the shaker at 28°C for about 24 h. Thereafter the cells were induced by adding 0.8 mL methanol every 12 h. After 36 h a few µL of cell

suspension were streaked out on YPD plates. After two days single colonies were picked, resuspended in DWPs containing YPD and pinned onto YPD as well as YPD-Zeo plates. Cells growing on YPD but not on YPD-Zeo are showing the desired phenotype as the plasmid containing a zeocin resistance should now be cut out of the genome.

3.3.4 Genome Walking experiments

3.3.4.1 Template Blocking PCR

Various PCR-based methodologies have been developed to isolate the unknown genomic regions adjacent to a known region, which is referred to as “genome walking”. A promising new approach is the “Template-blocking PCR” [29]. Some of the common limitations of genome walking, like low specificity and efficiency, short walking distance and complex methodology can potentially be avoided by this method. In Figure 13 an overview is presented. Genomic DNA is digested with a restriction enzyme which does not have a recognition site in the known region of the DNA sequence. For blocking the 3’OH at this point a dideoxynucleotide is attached. Now a double stranded cassette with compatible sticky ends is generated and ligated to the ends of the digested DNA. Then the cassette-ligated genomic DNA fragments can be used as a template for the amplification of the target gene with a cassette primer (CP) and a gene specific primer (GSP). Nonspecific amplification by the CP is inhibited by the absence of a second primer binding site due to the ddNTP blocking of the 3’OH. Hence specific amplification should not be affected because target gene specific amplification is started from the target gene-specific primer (GSP) and nonspecific amplification should be completely inhibited. A subsequent, nested PCR with primers binding inside the fragment generated in the primary PCR to enhance fidelity is optional.

PCR products were isolated and sequenced by LGC Genomics. By BLASTing the results against the genome of *P. pastoris* GS115 the exact locus of integration was identified.

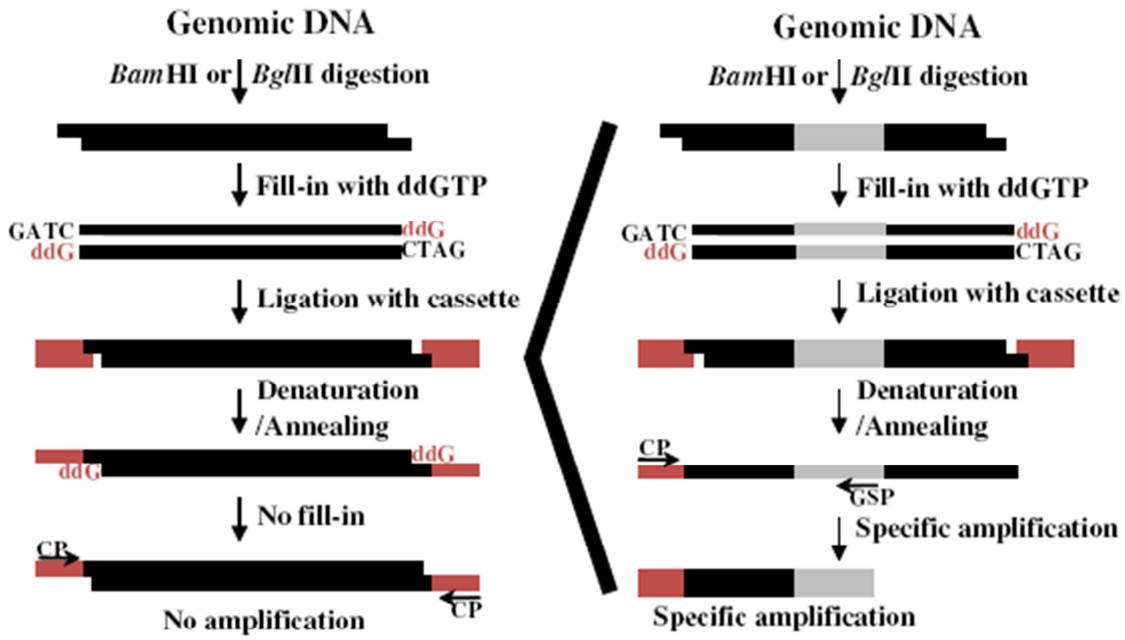


Figure 13: Overview of Template blocking PCR (adapted from Jung-Hoon et al. [29])

Restriction enzymes for digestion of genomic DNA

For the digestion of genomic DNA (gDNA) the restriction enzymes *Bgl*II and *Acc*65I, respectively, were chosen. The restriction enzyme *Bgl*II produces a ----CTAG overhang which leaves ----TAG after introduction of a dideoxyguanoside triphosphate (ddGTP). *Acc*65I leaves the overhang ----CATG which leaves ----ATG after introduction of ddGTP. The double-stranded cassettes were designed in respect to these overhangs.

Construction of double stranded cassette, the cassette primer and the gene-specific primer

For constructing the double stranded cassette complementary primers were designed according to the restriction enzyme used for digestion of genomic DNA. The two primers CSF27 and CSR30 as proposed in the literature [29] build up a cassette with an overhang complementary to the one-base-filled *Bgl*II digested gDNA-fragments, primers CSF28 and CSR31 match to one-base-filled *Acc*65I-fragments.

The cassette primer (CP/CPII) has extended GATC bases at the 3' end of CSF27/28. For the gene specific primer the sequence of the zeocin cassette was scanned for an appropriate binding site by using the program Primer3plus.

3.3.4.2 Genome Walker Kit Strategy

A quite similar approach is used by the Clontech Laboratories, Inc. in their commercial Genome Walker™ Universal Kit [30]. Here, not a dideoxynucleotid but an amine group at the 3' end of the adaptor-ligated genomic fragments blocks extension and consequently unspecific amplification (Figure 14). In contrast to the suggestions in the user manual not a blunt end digestion of genomic DNA but a digestion with restriction enzymes *Eco*RI and *Hind*III respectively leaving sticky ends was performed. Consequently the adaptor primers were designed accordingly. Primers AP1 and GSP1 were designed in a way similar to the primers in the template blocking PCR. Nested PCR was again optional by using primers AP2 and GSP2.

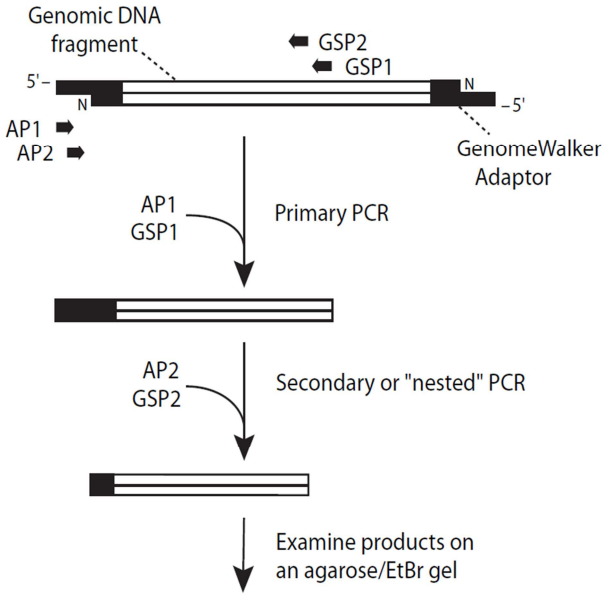


Figure 14: Overview of important steps in Genome Walker™ Kit Strategy. After digestion of genomic DNA fragments are ligated to a double stranded cassette carrying an amine group at the 3' end pointing outward (represented by 'N' in the figure). In the next step this amine group blocks extension and nonspecific amplification is prevented by the absence of a priming site. (Adapted from [30])

4 Results

4.1 Genome walking

In a screening of four thousand HRP- or APLE-secreting clones carrying a randomly integrated zeocin resistance cassette in their genome, 43 APLE secreting clones and 24 HRP-secreting clones were found to have significantly increased or decreased secretion behavior. In order to acquire information about factors influencing the secretion capacity of cells these clones became the matter of major interest. By genome walking experiments information about the location of the randomly integrated zeocin resistance cassette in the genome was accessible. By using two different approaches (template blocking PCR and Genome Walker^{TM} Kit Strategy) and 5 different restriction enzymes for digestion of gDNA (*Eco*RI, *Hind*III, *Bgl*II, *Acc*65I and *Bcl*I) 99 flanking regions of 67 integrated zeocin resistance cassettes in 60 different clones could be identified. In seven cases a double insertion was found, which did not give a direct hint as to which integration influenced secretion behavior of the clones. Seven clones were found to have an insertion in the HRP and APLE expression cassette, obviously impairing expression.

In Table 12 and 13 the genome walking results are listed including gene ID of the directly hit genes in *P. pastoris* GS115 and CBS7435 as well as protein BLAST suggestions from *Saccharomyces cerevisiae*. On the right side genes adjacent to the region hit by the integrated cassette are shown, which might have been influenced by insertion into thier promoter or operator sequences.

Table 12 and Table 13: Genome walking results of HRP secreting *P. pastoris* strains (Table 12) and APLE secreting P. pastoris strains (Table 13) showing increased or decreased secretion behavior compared to the wild type. On the left side (blue captions) genes directly hit by zeocin resistance cassette are shown, on the right side (greenish captions) possibly affected adjacent genes due to integrations into promoter or operator regions are indicated.

Identity	Activity in supernatant	GS115 Gene ID affected ORF	CBS7435 Gene ID	Protein BLAST suggestions	Gene ID maybe affected gene **1**	CBS7435 Gene ID	Protein BLAST suggestions
H1	increased	8197462 PAS_FragB_0019	CCA37611.1	Dus1p/Dus4p	PAS_FragB_0018 PAS_FragB_0020	no annotation CCA37612.1 (ATG 1200 bp away)	no ORF affected in CBS Vps70p
H2	increased	8197651 PAS_chr1-3_0184	CCA36351.1	Rim20p	PAS_chr1-3_0183	CCA36350.1 (ATG 500 bp away)	Cft1p
H3	increased	8197462 PAS_FragB_0019	as for H1	as for H1	PAS_FragB_0018 PAS_FragB_0020	as for H1	as for H1
H4	increased	8198181 PAS_chr2-1_0062	CCA38922.1	Rph1p (CBS annotation: hypothetical protein)	PAS_chr2-1_0061	CCA38923.1 (ATG 350 bp away)	Rad24p (CBS annotation: Cell division protease ftsH)
H5	increased	8199825 PAS_chr3_0625	CCA39536.1	Rim101p (CBS annotation: Zinc finger protein 347)			
H6	increased	8197329 PAS_chr1-4_0113	CCA37018.1	Sgt2p			
H7	increased	nothing annotated for GS115	CCA41142.1	hypothetical protein (Blast: Kcs1p? Inositol hexakisphosphate (IP6) and inositol heptakisphosphate (IP7) kinase			
H8	increased	8201292 PAS_chr4_0935 8200538 PAS_chr4_0012	CCA40244.1 CCA41154.1	no clear result Flo11p? no clear result Sed4p?	PAS_chr4_0013	CCA41153.1 (ATG 500 bp away)	beta-glucosidase
H9	increased	8197329 PAS_chr1-4_0113	as H6	as for H6			
H10	increased	8197651 PAS_chr1-3_0184	as H2	as for H2			
H11	decreased	8197243 PAS_chr1-1_0118	CCA36575.1	CBS annotation: Laminin subunit gamma-1	PAS_chr1-1_0117	CCA36574.1 (ATG 500 bp away)	Pmc1p (CBS annotation: putative membrane protein)
H12	decreased	expression cassette hit					
H13	decreased	expression cassette hit					
H14	decreased	8198479 PAS_chr2-1_0083	CCA38902.1	CBS annotation: UV radiation resistance-ass	PAS_chr2-1_0084	CCA38901.1 (ATG 400 bp away, intror	Lst7p (CBS annotation: folliculin)
H15	decreased	expression cassette hit					
H16	decreased	8198594 PAS_chr2-1_0212	CCA38772.1	Pmt1p	PAS_chr2-1_0211	CCA38773.1 (ATG 900 bp away)	Beclin-1-like protein
H17	decreased	expression cassette hit					
H18	decreased				8197753 PAS_chr1-4_0303	CCA37218.1 (ATG 20 bp away)	Hem3p (CBS annotation: hydroxymethylbilane synthase)
H21	increased	8199835 PAS_chr3_0635	CCA39524.1	Rim13p			
H22	increased	8197651 PAS_chr1-3_0184	CCA36351.1	Rim20p			
H23	decreased	8200026 PAS_chr3_1065	CCA39075.1	no clear result	no ORF annot. in GS115	CCA39074.1 (ATG 500 bp away)	Trm3p
H24	decreased	8197237 PAS_chr1-1_0111	CCA36569.1	Fra1p OR X-Pro dipeptidase	PAS_chr1-1_0110	PP7435_Chr1-0412 (300 bp away)	CBS annotation: mobile element
H27	decreased	expression cassette hit					
H45	decreased	8198692 PAS_chr2-2_0355 8201315 PAS_chr4_0022	CCA38070.1 CCA41145.1	Crc1p Ubp15p	PAS_chr2-2_0356	CCA38071.1 (ATG 800 bp away)	Tom22p

Identity	Activity in supernatant	GS115 Gene ID affected ORF	CBS7435 Gene ID	Protein BLAST suggestions	Gene ID maybe affected gene **1**	CBS7435 Gene ID	Protein BLAST suggestions
A1	increased	8197665 PAS_chr1-4_0534	CCA37453.1	Nam2p			
A2	increased	8200141_PAS_chr3_0316	CCA39851.1	Gtr1p	PAS_chr3_0317	CCA39850.1 (ATG 900 bp away)	Erg9p
A3	increased	8199199 PAS_chr2-2_0095 8198462 PAS_chr2-1_0066	CCA37798.1 CCA38919.1	Bem2p (annotated as Chitin synthase in CBS) Mon2p			
A4	increased	expression cassette hit					
A5	increased	8199120_PAS_chr2-2_0415	CCA38127.1	Apl6p			
A7	increased	expression cassette hit					
A8	increased	8197350 PAS_chr1-3_0038	CCA36203.1	Pam16p	PAS_chr1-3_0037	CCA36202.1 (ATG 450 bp away)	SET domain containing protein
A10	increased	8200059_PAS_chr3_1102	not annotated	Erd1p	PAS_chr3_1101	CCA39033.1 (ATG 250 bp away)	Mrpl11p
A11	increased	8197254_PAS_chr1-1_0131	CCA36593.1	Cyc8p	PAS_chr1-1_0131 PAS_chr2-1_0270	CCA36594.1 (ATG 800 bp away) CCA38714.1 (ATG 450 bp away)	Cyc8p Rtm1p
A13	increased	8198375_PAS_chr2-1_0757	CCA38204.1	Cat8p			
A15	increased				PAS_chr3_0826 PAS_chr3_0825	CCA39327.1 (ATG 150 bp away) CCA39328.1 (ATG 250 bp away)	Gpm1p Nsp1p
A16	increased				PAS_chr1-1_0470	CCA36532.1 (ATG 80 bp away)	Def1p
A17	increased	8201433_PAS_chr4_0487 8198724 PAS_chr2-2_0147	CCA40655.1 CCA37849.1	Mxr1p (methanol expression regulator I) CBS annotation: glycolipid transfer protein			
A18	increased	8197813_PAS_chr1-4_0365	CCA37281.1	Spt8p			
A20	increased	8201194 PAS_chr4_0107 8196722 PAS_chr1-1_0219	CCA41060.1 CCA36684.1	Rpl2bp	PAS_chr1_0535	CCA36685.1 (ATG 450 bp away)	Ybt1p
A23	decreased	8200886_PAS_chr4_0215	CCA40949.1	Bzz1p	PAS_chr4_0216	CCA40948.1 (ATG 600 bp away)	Ecm3p
A25	decreased				chr3 (only found in CBS) PAS_chr3_1218	CCA39388.1 (ATG 800 bp away)	Cse1p Doc1p
A26	decreased	8196921_PAS_chr1-4_0247	CCA37157.1	Smm1p	PAS_chr1-4_0246	CCA37156.1 (ATG 550 bp away)	Smf2p (CBS annotation: Smf1p)
A27	decreased	8200539 PAS_chr4_0013	CCA41153.1	beta-glucosidase	PAS_chr4_0014	CCA41152.1 (ATG 800 bp away)	no clear result
A29	decreased	8200580_PAS_chr4_0802	CCA40323.1	CBS annotation: Sodium/potassium/calcium exchanger 2			
A30	decreased		chr1 chr2 chr3 chr4	rRNA - 28S-ribosomal RNA		chr1 CCA37700 chr2 CCA38974.1 chr3 CCA40180.1 chr4 CCA41172.1	hypothetical protein, no clear Blast result
A31	decreased				PAS_chr3_0673	CCA39485.1 (ATG 350 bp away)	Slg1p
A32	decreased				PAS_chr1-3_0088 PAS_chr1-3_0089	CCA36254.1 (ATG 100 bp away) CCA40112.1 (ATG 500 bp away)	Trinucleotide repeat-containing gene 6B protein Chaperone protein dnaJ
A33	decreased	ORF only found in CBS, not directly hit	CCA40113.1 (ATG 70 bp away)	hypothetical protein; also Apl6p found as in A5	PAS_chr3_0069 PAS_chr3_0071	CCA40114.1 (ATG 600 bp away) CCA40112.1 (ATG 500 bp away)	Cat2p DNA-binding protein Rfx6p
A34	decreased				PAS_chr1-1_0428 PAS_chr1-1_0430	CCA36889.1 (ATG 130 bp away) CCA36890.1 (ATG 800 bp away)	Ena1p/Ena2p/Ena5p Ade4p
A36	decreased	PAS_chr3_1045	CCA39094.1	Pas1p aka Pex1p			
A37	decreased	PAS_chr1-4_0549	CCA37471.1	Mph1p			
A38	decreased	PAS_chr3_0886	CCA39258.1	no clear result	PAS_chr3_0887	CCA39257.1	Hpt1
A39	decreased	8198462 PAS_chr2-1_0066	CCA38919.1	Mon2p			
A40	decreased				PAS_chr2-1_0753 PAS_chr2-1_0752	CCA38210.1 CCA38211.1	Mns1p Str2p
A42	decreased	8200138 PAS_chr3_0312	CCA39855.1	Vps36	PAS_chr3_0311	CCA39856.1	no clear result
A44	decreased				PAS_chr1-4_0213	CCA37124.1 (ORF 200 bp away)	hypothetical protein
A45	decreased				PAS_chr2-1_0625	CCA38344.1 (ATG 100 bp away)	Vth1/Vth2/Pep1p

As can be seen in the list a few genes were found not only once but up to three times as for the pH-response regulator protein Rim20 (Table 12). Furthermore, not only Rim20p but several proteins of the Rim cascade could be identified making this pathway a promising target for further investigations. Striking though is the fact that none of the targets found in HRP-secreting strains appeared in the list for APLE-secreting strains suggesting that the influence of single genes on secretion rates depends on the properties of the secretory protein. This fact would make the aspired goal of identifying global enhancers of protein secretion invalid.

For further work targets that have shown to be increasing secretion upon deletion were of primary interest. Cases in which the gene responsible for the observed effect was clear were preferentially chosen, summing up to the following list of genes to validate in further experiments:

Table 14: Genes of primary interest found in genome walking experiments

Identity	CBS7435 gene ID	BLAST suggestion
H1	CCA37611.1	Dus1p
H5	CCA39536.1	Rim101p
H6	CCA37018.1	Sgt2p
H7	CCA41142.1	hypothetical protein
H8	CCA40244.1	hypothetical protein
A1	CCA37453.1	Nam2p
A2	CCA39851.1	Gtr1p
A10	not annotated	Erd1p
A11	CCA36593.1	Cyc8p
A13	CCA38204.1	Cat8p
A17	CCA40655.1	Mxr1p or Adr1p
A18	CCA37281.1	Spt8p
A34	CCA36889.1	Ena1p/Ena2p/Ena5p
A36	CCA39094.1	Pas1p aka Pex1p

As the number of interesting targets was quite extensive only selected ones were picked for further research in this thesis. It was decided to work on the target genes identified for improved secretion of HRP (see Table 14).

In order to confirm the number of Zeocin cassette integrations into the genome of strains H1-H8 Southern Blots employing a DIG-labeled zeocin resistance gene probe were done.

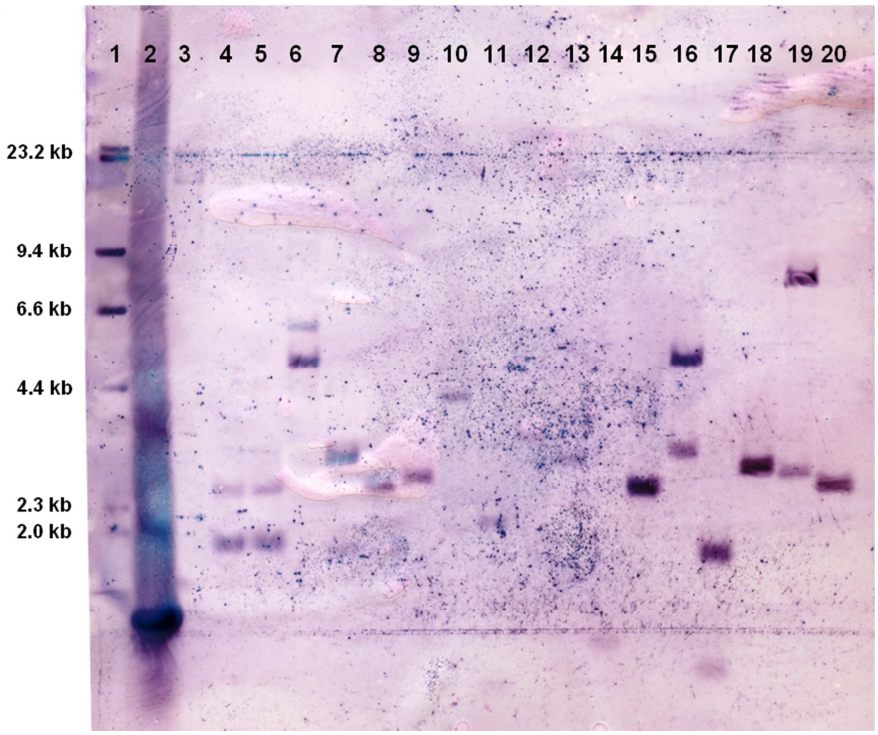


Figure 15: Southern Blotting of HRP secreting hits
Probe: Zeocin resistance gene DIG-labeled
Ladder: DNA Molecular Weight Marker II, DIG-labeled
Positive control: plasmid DNA containing one Zeocin resistance gene sequence
Restriction enzymes: *EcoR*V cuts in terminator of Zeocin reistance cassette (ca. 270 bp + 930 bp fragments), *Hind*III does not cut in cassette, therefore per gene copy two fragments expected if cut with *EcoR*V and one if cut with *Hind*III (see also Figure 3)

Table 15: Caption to numbers from Southern Blot Zeocin [Figure 15]
H1 – H9: Pp013 mutants, with Zeocin cassette integrated, found in screening to have increased HRP activity in supernatant

Lane	Sample/ integrated expression cassette	Restriction Enzyme
1	Ladder 18 µL	
2	Plasmid carrying Zeocin resistance gene	
3	H1	*Hind*III
4	H1	*EcoR*V
5	H1	*Hind*III+*EcoR*V
6	H5	*Hind*III
7	H5	*EcoR*V
8	H5	*Hind*III+*EcoR*V
9	H6	*Hind*III
10	H6	*EcoR*V
11	H6	*Hind*III+*EcoR*V
12	H9	*Hind*III
13	H9	*EcoR*V
14	H9	*Hind*III+*EcoR*V
15	H7	*Hind*III
16	H7	*EcoR*V
17	H7	*Hind*III+*EcoR*V
18	H8	HindIII
19	H8	*EcoR*V
20	H8	*Hind*III+*EcoR*V

HRP secreting clones of primary interest (Table 14) were analyzed by Southern Blot. In genome walking experiments strain H8 had been found to have two integrations in the genome. This could not be confirmed here (see lanes 18-20). Results for strains H1 and H7 clearly suggest single integrations. Contrary to expectations, the HindIII digested gDNA of strain H5 showed two bands but multiple insertions were not indicated in lanes 7 and 8. Lanes 9-14 are hard to read but seem to confirm single integrations. All in all one can say that results are difficult to interpret as different lanes of the same sample gave inconsistent results.

Blots for analyzing the remaining strains from Table 14 were performed showing bands only for the positive control. As the DIG-labeled probe was reused for the latter blots, probe concentration might have been rather low. It was abstained from any further zeocin resistance cassette probe blots due to lack of time.

4.2 Cell wall integrity tests

4.2.1 HRP activity in the supernatant and intracellular fractions

In order to investigate the consequences of mutating selected genes on protein secretion HRP activity was measured in cell extracts and culture supernatants of HRP expressing strains.

By comparing HRP activity in the supernatant to its activity inside the cells one might be able to draw conclusions about the fraction of secretory protein that is efficiently secreted (Figure 16).

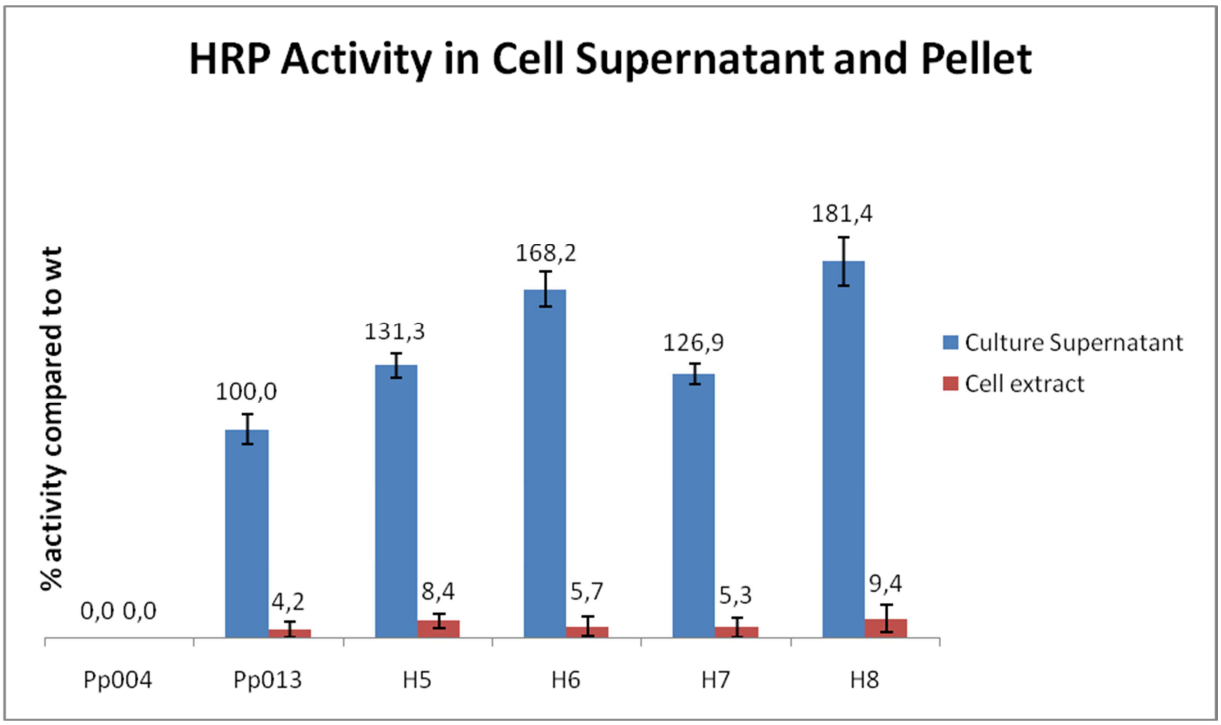


Figure 16: Results of HRP activity measurements in culture supernatant and cell extract. Every strain was investigated in triplicate and results represent means and STD. Pp004 represents the CBS7435 dhis4 strain without HRP expression cassette, Pp013 represents the HRP secreting GS115 strain with an expression cassette integrated at the AOX1 locus.

Screening hits H5-H8 showed increased HRP secretion into the culture supernatant consistent with the screening results (pers. communication Christine Winkler). The detection of differences in intracellular accumulation of HRP was hampered by several methodical limitations leading to doubts about reliability. Values for intracellular HRP activity scattered significantly, leading to high standard deviations (Figure 16). This may have been caused by loss of activity during glass bead lysis. It is unclear how much activity is lost due to HRP degradation/denaturation or incomplete and non-uniform cell disruption.

Furthermore, *Pichia pastoris* is known for its high level expression of catalase which catalyzes degradation of H_2O_2, leading to its depletion during the assay. On behalf of these limitations the strategy was not pursued in following experiments. It was suggested to assess intracellular HRP concentrations by methods such as Western Blotting.

4.3 Establishment of HRP secreting strains as host strains for target knock-outs

Based on genome walking results it was decided to knock-out genes identified to influence secretion behavior. Initially it was considered less laborious to first establish model protein expression cassettes in a CBS7435 strain (and additionally in a Ku70p deficient strain) in which the different target genes could be knocked-out. This led to the results presented in this section. The strains described here were not used for further research but moved into the instituts strain collection. This holds true for all strains except for one being useful as a wild type control in activity assays as described in section 4.6. The advantage of this approach is that the copy number of the integrated expression cassette can easily be checked upon by Southern blotting.

Later on it was decided to pursue another strategy: As it was desired to have the knock-out strains available also for other projects it appeared advantageous to establish the knock-outs in wild type strains before integrating expression cassettes. In particular, it was of interest for further studies if genes listed in Table 14 show the same influences on expression and secretion for different model proteins, i.e. APLE instead of HRP and vice versa. Furthermore, instead of using a constitutive promoter, e.g. P_{GAP}, it would be revealing to use an inducible promoter such as P_{AOX1}. These factors eventually were convincing, although the expression cassette has to be integrated in each knock-out strain separately, complicating the affirmation of single integration by Southern blotting.

Pursuing the first approach, the following expression constructs containing a *HIS4* marker gene were transformed into CBS7435 dhis4 wild type as well as CBS7435 dhis4 dku70 strains by Christine Winkler:

Transfomants were screened on minimal methanol plates for histidin prototrophy and MutS phenotype. In the following the characterization of these strains via Southern Blotting is described as strains with a single integration at the *AOX1* locus were required. The indication A1 – G4 refers to the position on a plate of glycerol stocks of the respective strains. Strains with the same promoter and the same letter of indication are carrying the same expression construct (i.e P_{AOX1} APLE dku70 G1-G4).

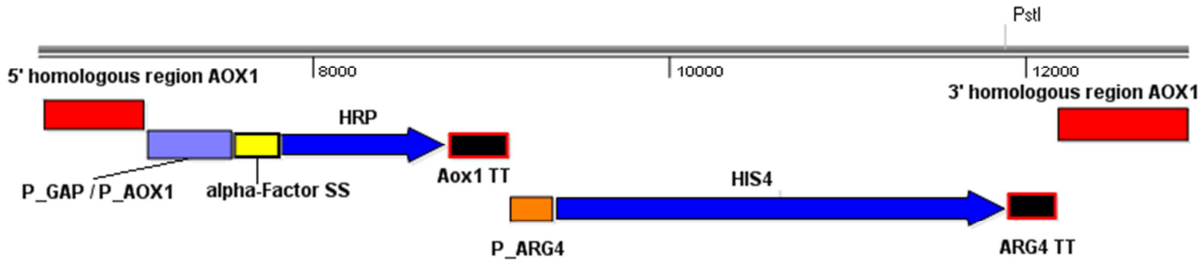


Figure 17: HRP expression cassette integrated at the AOX1 locus. P_{GAP}, P_{AOX1} and HIS4 probes of the following Southern blots bind to the respective sites in the figure.

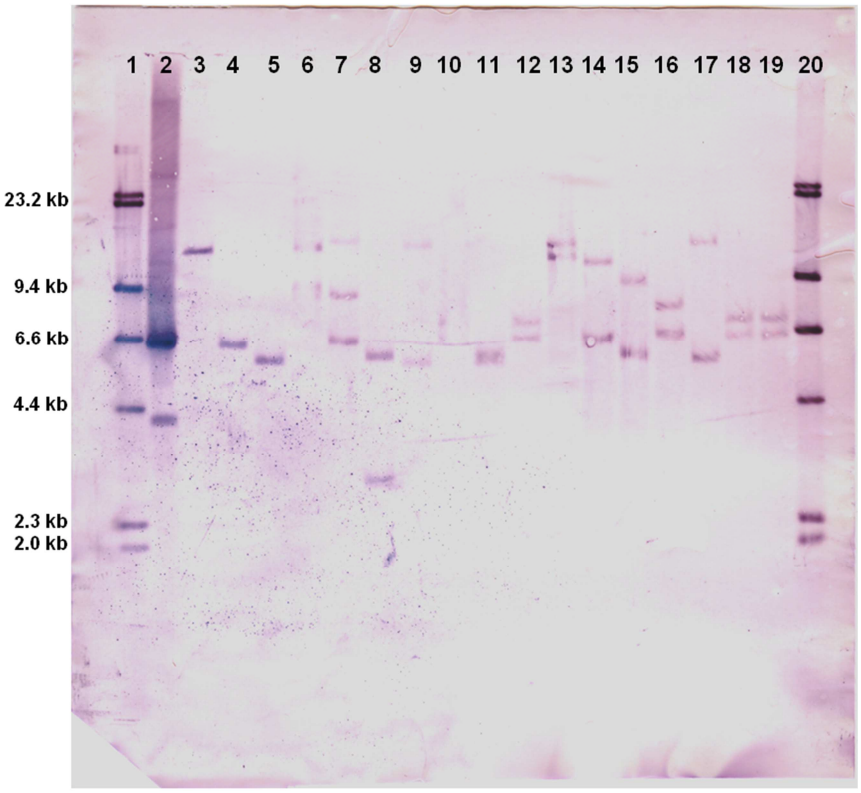


Figure 18: Southern Blot for P_{GAP}-driven expression of HRP and APLE #1
Probe: GAP promoter DIG-labeled
Ladder: DNA Molecular Weight Marker II, DIG-labeled
Positive control: plasmid DNA containing one GAP promoter sequence
Restriction enzymes: *Pst*I cuts at end of HIS4 (part of integrated cassette), *Hind*III does not cut in integrated cassette. None of the restriction enzymes cuts in P_{GAP} – sequence, therefore only one labeled band per copy in genome is expected. One copy of GAP promoter is present also in the wild type. Information about the positive control can be found in section 6.

Table 16: Caption to numbers from Southern Blot P_{GAP} #1 [Figure 18]

Lane	Sample/ integrated expression cassette	Restriction Enzyme
1	Ladder 18 µL	
2	Plasmid with GAP promoter	
3	CBS7435 d*his4* (control)	*Pst*I
4	CBS7435 d*his4* (control)	*Hind*III
5	CBS7435 d*his4* (control)	*Pst*I+*Hind*III
6	Pp013	*Pst*I
7	Pp013	*Hind*III
8	Pp013	*Pst*I+*Hind*III
9	P_{GAP} HRP A1	*Pst*I
10	P_{GAP} HRP A1	*Hind*III
11	P_{GAP} HRP A1	*Pst*I+*Hind*III
12	P_{GAP} HRP A2	*Hind*III
13	P_{GAP} APLE E1	*Pst*I
14	P_{GAP} APLE E1	*Hind*III
15	P_{GAP} APLE E1	*Pst*I+*Hind*III
16	P_{GAP} APLE E2	*Hind*III
17	P_{GAP} HRP Δku70 C1	*Pst*I
18	P_{GAP} HRP Δku70 C1	*Hind*III
19	P_{GAP} HRP Δku70 C2	*Hind*III
20	Ladder, 18 µL	

For Southern Blot P_{GAP} #1 the strains shown in Table 16 were applied. CBS7435 dhis4 is a wild type control as one copy of the GAP promoter is also present. Strains of the correct genotype carrying an expression cassette containing another GAP promoter should show two bands. This holds true for all restriction enzymes and combinations as no suiting restriction sites are found in the sequence in question.

Bands matching to the ones found in the wild type (lanes 3-5) were found for all samples. Lane 7 shows 3 bands indicating a possible double integration event. This is not affirmed in the lanes 6 and 8 rendering the situation unclear.

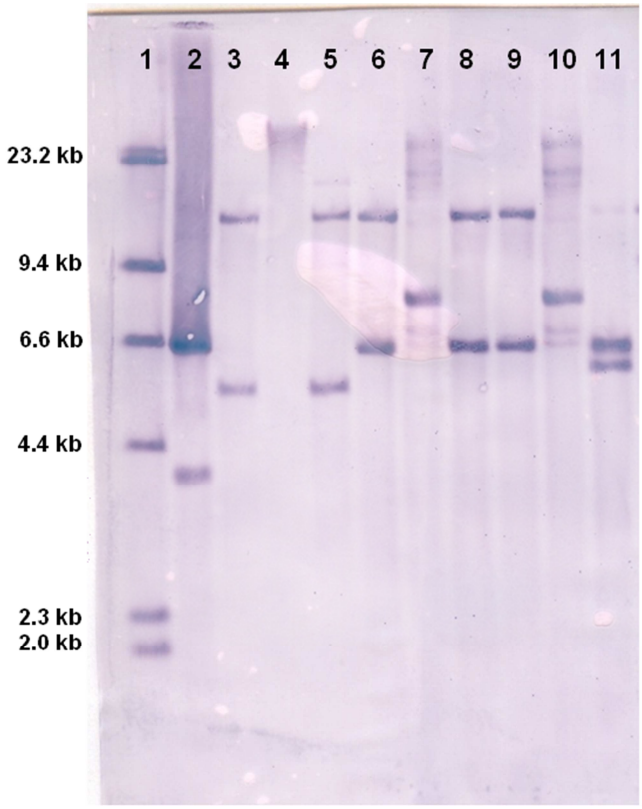


Figure 19: Southern Blot for P_{GAP}-driven expression of HRP and APLE #2
Probe: GAP promoter DIG-labeled
Ladder: DNA Molecular Weight Marker II, DIG-labeled
Positive control: Plasmid DNA containing one GAP promoter sequence
Restriction enzymes: *Pst*I cuts at end of HIS4 (part of integrated cassette), *Hind*III does not cut in integrated cassette. None of the restriction enzymes cuts in P_{GAP} – sequence, therefore only one labeled band per copy in genome is expected.
One copy of GAP promoter is present also in the wild type. Information about the positive control can be found in section 6.

Table 17: Caption to numbers from Southern Blot P_{GAP} #2 [Figure 19]
Pp015 represents an APLE secreting GS115 strain with an expression cassette integrated at the AOX1 locus.

Lane	Sample	Restriction Enzyme
1	Ladder 18 µL	
2	Plasmid with GAP promoter	
3	Pp015	*Pst*I
4	Pp015	*Hind*III
5	Pp015	*Pst*I+*Hind*III
6	P_{GAP} APLE Δku70 G1	*Pst*I
7	P_{GAP} APLE Δku70 G1	*Hind*III
8	P_{GAP} APLE Δku70 G1	*Pst*I+*Hind*III
9	P_{GAP} APLE Δku70 G2	*Pst*I
10	P_{GAP} APLE Δku70 G2	*Hind*III
11	P_{GAP} APLE Δku70 G2	*Pst*I+*Hind*III

On Southern Blot pGAP #2 remaining P_{GAP}-strains were applied. The results are not as distinct as the ones from blot #1. Likely for gDNA samples of strain Pp015 digestion by *Hind*III did not work out which would explain missing bands for lane 4 and the same pattern for the double digested gDNA of lane 5 as for the PstI digested sample of lane 3. The situation in lanes 7, 8, 10 and 11 remains unclear, again mainly because of incomplete digestion by *Hind*III.

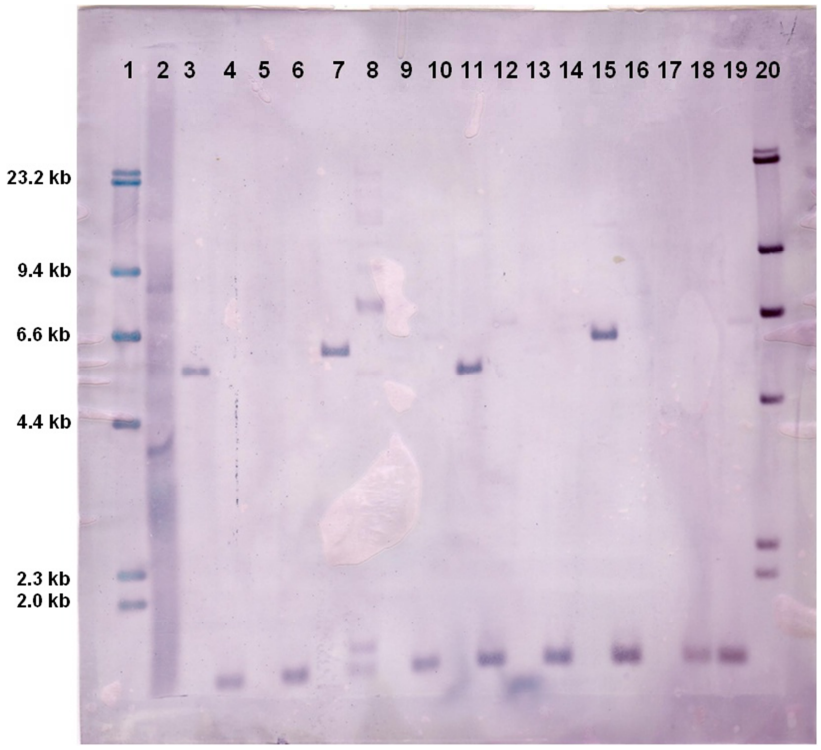


Figure 20: Southern Blot for P_{AOX1}-driven expression of HRP and APLE #1
Probe: AOX promoter DIG-labeled
Ladder: DNA Molecular Weight Marker II, DIG-labeled
Positive control: Plasmid DNA containing one AOX promoter sequence
Restriction enzymes: *Pst*I cuts at end of *HIS4* (part of integrated cassette), *Hind*III does not cut in integrated cassette. None of the restriction enzymes cuts in P_{AOX} – sequence, therefore only one labeled band per copy in genome is expected. Information about the positive control can be found in section 6.

Table 18: Caption to numbers from Southern Blot P_{AOX1} #1 [Figure 20]

Lane	Sample/ integrated expression cassette	Restriction Enzyme
1	Ladder 18 µL	
2	Plasmid with AOX promoter	
3	P_{AOX1} HRP A1	*Pst*I
4	P_{AOX1} HRP A1	*Hind*III
5	P_{AOX1} HRP A2	*Pst*I+*Hind*III
6	P_{AOX1} HRP A2	*Hind*III
7	P_{AOX1} APLE E1	*Pst*I
8	P_{AOX1} APLE E1	*Hind*III
9	P_{AOX1} APLE E1	*Pst*I+*Hind*III
10	P_{AOX1} APLE E2	*Hind*III
11	P_{AOX1} HRP Δku70 C1	*Pst*I
12	P_{AOX1} HRP Δku70C1	*Hind*III
13	P_{AOX1} HRP Δku70 C2	*Pst*I+*Hind*III
14	P_{AOX1} HRP Δku70 C2	*Hind*III
15	P_{AOX1} APLE Δku70 G1	*Pst*I
16	P_{AOX1} APLE Δku70 G1	*Hind*III
17	P_{AOX1} APLE Δku70 G1	*Pst*I+*Hind*III
18	P_{AOX1} APLE Δku70 G2	*Hind*III
19	CBS7435 his4	*Hind*III
20	Ladder, 18 µL	

Strains containing model enzymes HRP and APLE under the control of P_{AOX1} were applied on a gel for analysis via P_{AOX1}-probe. It can be seen that the agarose gel had run quite far and bands of small fragments were found only little from the very bottom. This applies in particular for *Pst*I digested samples. In the lanes for double digested DNA (lanes 5, 9, 13 and 17) no bands were found indicating migration of fragments out of the gel. As the *AOX1* promoter sequence of the wild type is replaced during integration of heterologous DNA at the *AOX1* locus only one band is expected in every case. Therefore the situation in lane 8 is unclear.

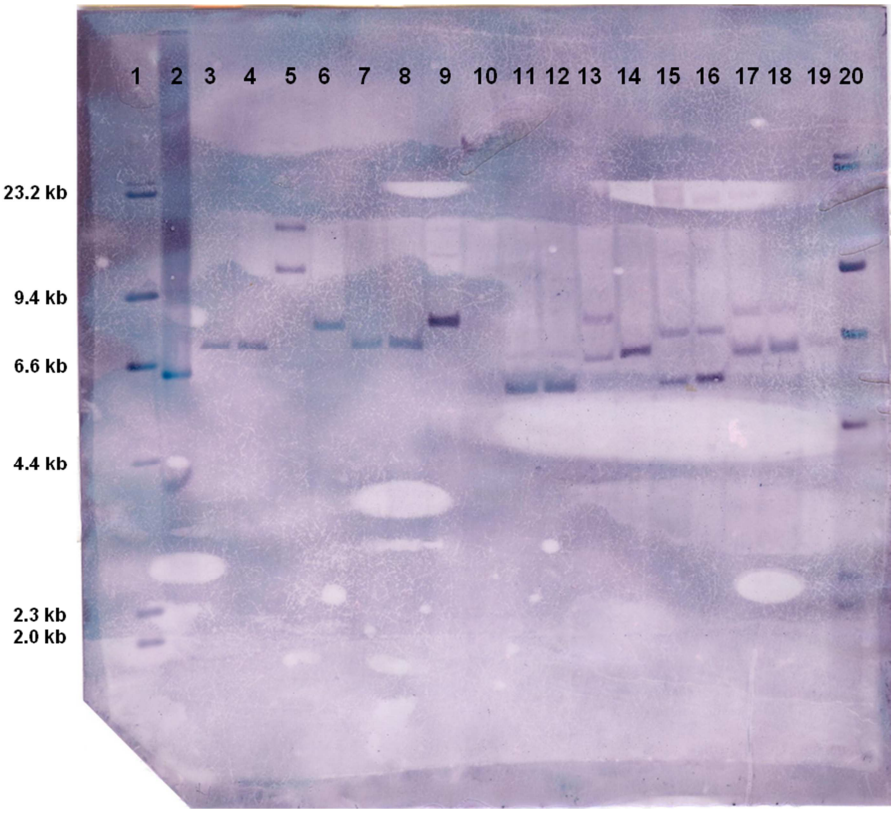


Figure 21: Southern Blot *HIS4* #1
Probe: HIS4 gene DIG-labeled Ladder: DNA Molecular Weight Marker II, DIG-labeled Positive control: Plasmid containing *HIS4* gene Restriction enzyme: *Hind*III does not cut in integrated cassette; Therefore only one labeled band per copy in genome expected. Information about the positive control can be found in section 6.

Table 19: Caption to numbers from Southern Blot *HIS4* #1 [Figure 21]

Lane	Sample/ integrated expression cassette	Restriction Enzyme
1	Ladder, 18 µL	
2	Plasmid pos. control	
3	P_{GAP} HRP A1	*Hind*III
4	P_{GAP} HRP A2	*Hind*III
5	P_{GAP} APLE E1	*Hind*III
6	P_{GAP} APLE E2	*Hind*III
7	P_{GAP} HRP dku70 C1	*Hind*III
8	P_{GAP} HRP dku70 C2	*Hind*III
9	P_{GAP} APLE dku70 G1	*Hind*III
10	P_{GAP} APLE dku70 G2	*Hind*III
11	P_{AOX1} HRP A1	*Hind*III
12	P_{AOX1} HRP A2	*Hind*III
13	P_{AOX1} APLE E1	*Hind*III
14	P_{AOX1} APLE E2	*Hind*III
15	P_{AOX1} HRP Δku70 C1	*Hind*III
16	P_{AOX1} HRP Δku70 C2	*Hind*III
17	P_{AOX1} APLE Δku70 G1	*Hind*III
18	P_{AOX1} APLE Δku70 G2	*Hind*III
19	CBS7435 his4	*Hind*III
20	Ladder, 18 µL	

Some of the Southern blots employing the P_{GAP} and P_{AOX1} probes did not give clear and unambiguous results. Thus, it was decided to detect integrations via a second probe binding to the HIS4 gene as it is part of the integrated cassette as well.

For this Southern blot all applied gDNA was digested with *Hind*III only. As the host strain CBS7435 is his4 deficient no bands can be observed in the respective lane. Strains with more than one band were likely to have a double integration and thus were discarded.

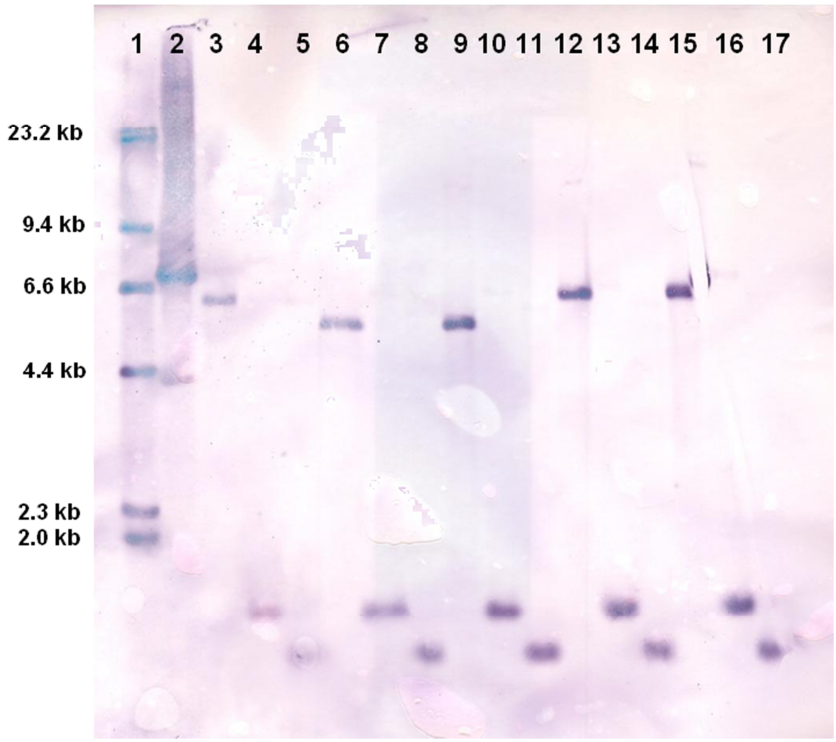


Figure 22: Southern Blot for P_{AOX1}-driven expression of HRP and APLE #2
Probe: AOX promoter DIG-labeled
Ladder: DNA Molecular Weight Marker II, DIG-labeled
Positive control: Plasmid DNA containing one AOX promoter sequence.
Restriction enzymes: *Pst*I cuts at end of HIS4 (part of integrated cassette), *Hind*III does not cut in integrated cassette. None of the restriction enzymes cuts in P_{AOX1} – sequence, therefore only one labeled band per copy in genome is expected. Information about the positive control can be found in section 6.

Table 20: Caption to numbers from Southern Blot P_{AOX1} #2 [Figure 22]

Lane	Sample/ integrated expression cassette	Restriction Enzyme
1	Ladder 18 µL	
2	Plasmid with AOX promoter	
3	P_{AOX1} APLE E2	*Pst*I
4	$P_{AOX\ 1}$ APLE E2	*Hind*III
5	P_{AOX1} APLE E2	*Pst*I+*Hind*III
6	P_{AOX1} HRP Δku70 C3	*Pst*I
7	P_{AOX1} HRP Δku70 C3	*Hind*III
8	P_{AOX1} HRP Δku70 C3	*Pst*I+*Hind*III
9	P_{AOX1} HRP Δku70 C4	*Pst*I
10	P_{AOX1} HRP Δku70 C4	*Hind*III
11	P_{AOX1} HRP Δku70 C4	*Pst*I+*Hind*III
12	P_{AOX1} APLE Δku70 G3	*Pst*I
13	P_{AOX1} APLE Δku70 G3	*Hind*III
14	P_{AOX1} APLE Δku70 G3	*Pst*I+*Hind*III
15	P_{AOX1} APLE Δku70 G4	*Pst*I
16	P_{AOX1} APLE Δku70 G4	*Hind*III
17	P_{AOX1} APLE Δku70 G4	*Pst*I+*Hind*III

As strains P_{AOX1} APLE E1, P_{AOX1} HRP dku70 C1+C2 and P_{AOX1} APLE dku70 G1+G2 showed two bands in the Southern blot HIS4 #1 (see Figure 21) they were discarded and other transformants with the same expression construct were applied on another gel and analyzed by an *AOX1* promoter probe (C3+C4, G3+G4) (Figure 22). In this case the gel electrophoresis was stopped in time so also the double digested DNA samples are visible. Every sample showed only one band indicating single integrations.

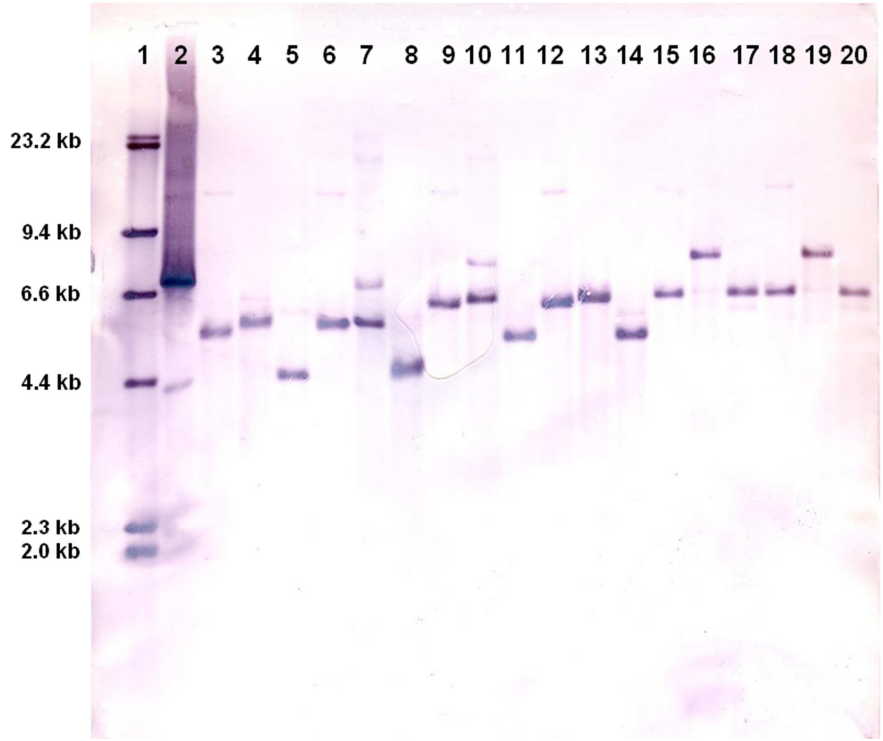


Figure 23: Southern Blot *HIS4* #2
Probe: *HIS4* gene DIG-labeled Ladder: DNA Molecular Weight Marker II, DIG-labeled Positive control: Plasmid containing *HIS4* gene Restriction enzyme: *Hind*III does not cut in integrated cassette; Therefore only one labeled band per copy in genome expected. Information about the positive control can be found in section 6.

Table 21: Caption to numbers from Southern Blot HIS4 #2 [Figure 23]

Lane	Sample/ integrated expression cassette	Restriction Enzyme
1	Ladder, 18 µL	
2	Plasmid pos. control	
3	P_{AOX1} HRP Δku70 C3	PstI
4	P_{AOX1} HRP Δku70 C3	HindIII
5	P_{AOX1} HRP Δku70 C3	PstI+HindIII
6	P_{AOX1} HRP Δku70 C4	HindIII
7	P_{AOX1} HRP Δku70 C4	PstI
8	P_{AOX1} HRP Δku70 C4	HindIII
9	P_{AOX1} APLE Δku70 G3	PstI+HindIII
10	P_{AOX1} APLE Δku70 G3	HindIII
11	P_{AOX1} APLE Δku70 G3	PstI
12	P_{AOX1} APLE Δku70 G4	HindIII
13	P_{AOX1} APLE Δku70 G4	PstI+HindIII
14	P_{AOX1} APLE Δku70 G4	HindIII
15	P_{GAP} APLE Δku70 G3	PstI
16	P_{GAP} APLE Δku70 G3	HindIII
17	P_{GAP} APLE Δku70 G3	PstI+HindIII
18	P_{GAP} APLE Δku70 G4	PstI
19	P_{GAP} APLE Δku70 G4	HindIII
20	P_{GAP} APLE Δku70 G4	PstI+HindIII

In the *HIS4* #2 blot we found single bands in all cases except for lanes 7 and 10. Faint bands above the main bands may be caused by little amounts of incompletely digested DNA. For all 8 different constructs transformants with single integrations were identified (see Table 22). Strain P_{GAP} HRP dhis4 A1 was used as wildtype control in activity assays (see section 0). The remaining strains were deposited in the strain collection of the IMBT.

Table 22: Selected strains

Strain	Analyzed on Southern Blot
CBS7435 P_{GAP} HRP A1	P_{GAP} #1, HIS4 #1
CBS7435 Δku70 P_{GAP} HRP C1	P_{GAP} #1, HIS4 #1
CBS7435 P_{GAP} APLE E2	P_{GAP} #1, HIS4 #1
CBS7435 Δku70 P_{GAP} APLE G3	HIS4 #2
CBS7435 P_{AOX1} HRP A1	P_{AOX1} #1, HIS4 #1
CBS7435 Δku70 P_{AOX1} HRP C3	P_{AOX1} #2, HIS4 #2
CBS7435 P_{AOX1} APLE E2	P_{AOX1} #2, HIS4 #1
CBS7435 Δku70 P_{AOX1} APLE G4	P_{AOX1} #2, HIS4 #2

4.4 Functions and properties of target knock-out genes

The results of the genome walking experiments elicited investigations on several genes concerning their role in protein secretion. For further validation four targets were selected as their deletion showed a significant increase in expression and/or secretion of the model protein HRP. The role of Sgt2p, Rim101p, Kcs1p and Flo11p in protein secretion pathways in *P. pastoris* have not been investigated yet. Here, information about function of *Saccharomyces cerevisiae* homologues is presented and their putative function in protein expression and secretion is discussed.

4.4.1 Kcs1p

By blasting the sequence of the affected gene in *P. pastoris* strain H7 against the genome of *S. cerevisiae* the highest similarity was found for the protein Kcs1p with a maximal identity of 80% at query coverage of 67%. Kcs1p is a multifunctional protein with a major role as a diphosphoinositol polyphosphate synthase [31]. Compared to the size of the *P. pastoris* homologue (595 amino acids) it is much larger in size (1050 amino acids) and has a molecular weight of ~120 kDa (compared to ~68 kDa in *P. pastoris)*.

In addition to a high overall sequence similarity a conserved consensus sequence of Ser-Leu-Leu found in homologues of higher eukaryotes as well as in Kcs1p is also present in the respective *P. pastoris* protein. It could be shown that they are part of the catalytic site essential for phosphate kinase activity [32]. Furthermore, Huang et al. [33] proposed two

groups of four heptad repeats of leucine residues to be conserved leucine zippers directing homo- and heterodimerizations of proteins and having putative DNA binding abilities. These are only loosely matched in the *P. pastoris* homologue.

Initial deletion studies of Kcs1p in *S. cerevisiae* led to 93% lower cellular steady-state levels of diphosphoinositol pentakisphosphate (PP-$InsP_5$) and bis-diphosphoinositol tetrakisphosphate ($(PP)_2$-$InsP_4$) compared to wild type cells [34]. Interestingly, effects on phosphate regulation were accompanied by fragmentation of vacuolar compartments, defective response to salt stress and fragility of the cells leading to elevated levels of alkaline phosphatase in the extracellular milieu at pH 9.8 but also at a less stressfull pH of 6.5 [35]. These findings may indicate defects in the maintenance of cell wall structure and cell wall stability. Observations were made demonstrating that the vacuolar H-ATPase is simultaneously significantly up-regulated. Its function is to keep the electrochemical gradient by sequestration of sodium into the vacuole, which links the defective response to salt stress to observed deformations of vacuolar compartments [34]. Another study by Luo et al. suggests that a mammalian diphosphoinositol polyphosphate synthase interacts with a guanine nucleotide exchange factor that regulates vesicle exocytosis. PP-$InsP_5$ and $(PP)_2InsP_4$ might be phosphate donors for protein phosphorylation and thereby supporting this theory.

4.4.2 Rim101p

The Rim101 protein is known to be part of the Rim-cascade comprising several other Rim proteins such as Rim8p, Rim9p, Rim13p and Rim20p. Rim101p acts as a transcriptional repressor by downregulation of several genes including two more transcriptional repressors. As a result the absence of Rim101p leads to both an increase and a decrease in the expression of certain genes [36]. Although its size of 628 amino acids is similar to the size of its putative homologue in *P. pastoris* (608 amino acids), significant amino acid identity is almost completely restricted to a conserved zinc finger region consisting of two cysteins and two histidins coordinating a zinc ion. In its native state Rim101p has to undergo proteolytic processing, as it gets activated by removal of approximately 100 residues from the C-terminus of the protein [37]. Deletions of Rim8, Rim9, Rim13 and Rim20 have shown to be preventing the proteolytic processing of Rim101p thereby blocking its activation [37]. The close interaction of Rim 13p, Rim20p and Rim101p, all found in HRP expressing mutants of

this thesis (see Table 12) is reviewed in [38], suggesting similar mechanisms leading to increased secretion of knock-out strains described in the present work.
Rim101 has functions as a positive acting regulator of sporulation, of adaption to alkaline pH, invasive growth, ion tolerance and acquisition and also influences cell-wall characteristics [38]. Castrejon et al. [36] assume a possible relationship between cell wall assembly in S. cerevisiae and the Rim101 pathway as some *rim* mutants show synthetic lethality with cell wall mutations and mutants having cell-wall defects showed reduced growth at alkaline pH [39].

4.4.3 Sgt2

The affected gene in strain H6 was found to be annotated in *P. pastoris* as a “Small glutamine-rich tetratricopeptide repeat-containing protein alpha” with a length of 351 amino acids. Blasting it against the proteome of *S. cerevisiae* led to a close homologue named Sgt2p with a sequence identity of 40% over 98% of sequence coverage. Both proteins carry three conserved tetratricopeptide repeat domains which can be involved in a variety of functions including protein-protein interactions [40].

Sgt2 is known to be an important part of a complex process for targeting tail-anchored (TA) proteins (proteins carrying a transmembrane domain at the C-terminus) to the ER membrane. In contrast to the well-known cotranslational process of transmembrane domain (TMD) recognition by the SRP, TA-proteins have to be targeted posttranslationally. This involves liberation of hydrophobic domains which are prone to aggregation if not shielded from the cytoplasm by specific chaperones. In eukaryotes this responsibility is accepted by the GET pathway which shuttles TA-proteins to the ER [41]. Schuldiner et al. [42] could demonstrate in several different strains, that deletions of GET pathway proteins lead to a wide array of phenotypes and many of the 55 predicted TA-proteins found in *S. cerevisiae* are found in mitochondria, peroxisomes and, importantly, throughout the secretory pathway. E.g. many of the SNARE proteins mediating fusion of secretory vesicles are TA-proteins making them important for enabling vesicular traffic. Although thereby a possible connection between Sgt2 and the secretory pathway can be presented, a clear theory about how deletion of this protein may influence secretion of heterologous proteins cannot be drawn.

4.4.4 Hypothetical protein of strain H8

The gene that was found to be disrupted in strain H8 is not annotated in *P. pastoris*. By blasting the gene product against the proteome of *S. cerevisiae* very little homology could be discovered. Small parts (query coverage 15%) were found to have similarity to a FLO11 domain containing GPI-anchored cell surface glycoprotein (flocculin) named Muc1p. Further speculations about the function of the hypothetical protein H8 were not reasonable.

In the following target knock-out strains are named according to their respective homologue in *S. cerevisiae*. That means the target knock-out strain of the gene found in strain H5 which shows homology to RIM101 is referred to as Δrim101. The same is done for H6 and its target knock-out counterpart Δsgt2, for H7 and Δkcs1, and for H8 and Δflo11.

4.5 Growth curves

The evaluation of cell growth gives information about cell viability and possible defects due to the knock-outs in the genome. In the laboratory, under favorable conditions, the interval for doubling of a *Pichia pastoris* population is about 2 h in YPD medium. This is true as long as there are no limiting factors such as supply of glucose and oxygen or accumulation of inhibitory metabolites.

Knock-outs in the genome of *P. pastoris* may lead to less robustness and a slower growth which would become apparent by a growth curve analysis as shown in Figure 24.

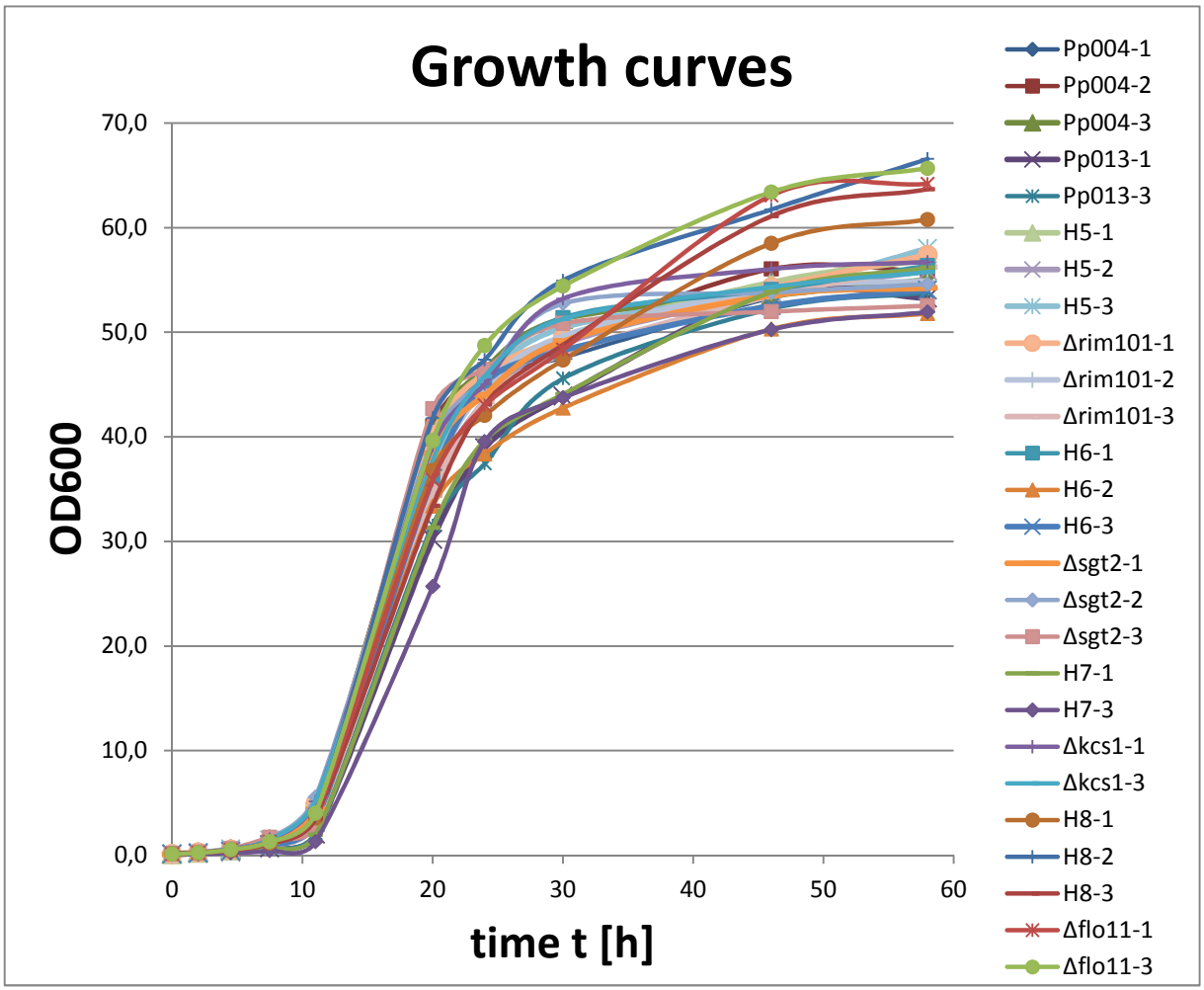


Figure 24: Growth curves of strains from the screening performed by Christine Winkler (H5-H8), the respective target-knock-out strains (Δrim101, Δsgt2, Δkcs1 and Δflo11) and the parent strains Pp004 (*P. pastoris* CBS7435 dhis4) and Pp013 (*P. pastoris* GS115 P_{GAP}-HRP).

Growth curves for strains from the original screening by Christine Winkler (H5-H8) as well as the respective target knock-out strains (Δrim101, Δsgt2, Δkcs1 and Δflo11) are compared to the parent strains Pp004 (*P. pastoris* CBS7435 dhis4) and Pp013 (*P. pastoris* GS115 P_{GAP}-HRP), every strain in duplicate or triplicate.

All mutated strains showed a lag time like the parental strains. In the exponential phase no significant deviations were observed. After about 30 h of incubation cells passed into the stationary phase and reached a maximal OD of ~55±5 OD-units at the end of the measurement after 58 h. Strains H8 and Δflo11 (target knock-out of the hypothetical protein hit in H8) deviated significantly, reaching a maximal OD of 60-66 OD-units.

4.6 HRP activity measurements in cell supernatant

The choice of cultivation vessel has a great influence on the conditions cells will find and therefore, growth and secretion characteristics may differ significantly. Its size and shape as well as the presence of baffles affect the oxygen supply and therefore, also cell growth. In order to investigate the secretion behavior of strains under two varying conditions activity measurements were not only performed in deep-well plates but also in shake flasks.

4.6.1 Cell cultivation in DWPs

Results of activity measurements in DWPs are shown in Figure 25. The target knock-out strains were compared to their counterparts obtained in screening by Christine Winkler on the one hand, as well as with the HRP-secreting wild type strain GS115 P_{GAP}-HRP and the strain CBS7435 P_{GAP}-HRP characterized via Southern blotting on the other hand. The results presented here stem from an assay with the following number of parallel determinations:

Table 23: Number of parallel determinations

Strain	Number of parallel determinations
GS115 dhis4 wt (negative control)	24
GS115 P_{GAP}-HRP	32
CBS7435 P_{GAP}-HRP	64
Strains containing randomly inserted zeocin resistance cassette (“H5”-“H8”)	32 each
Target knock-out strains	32 each

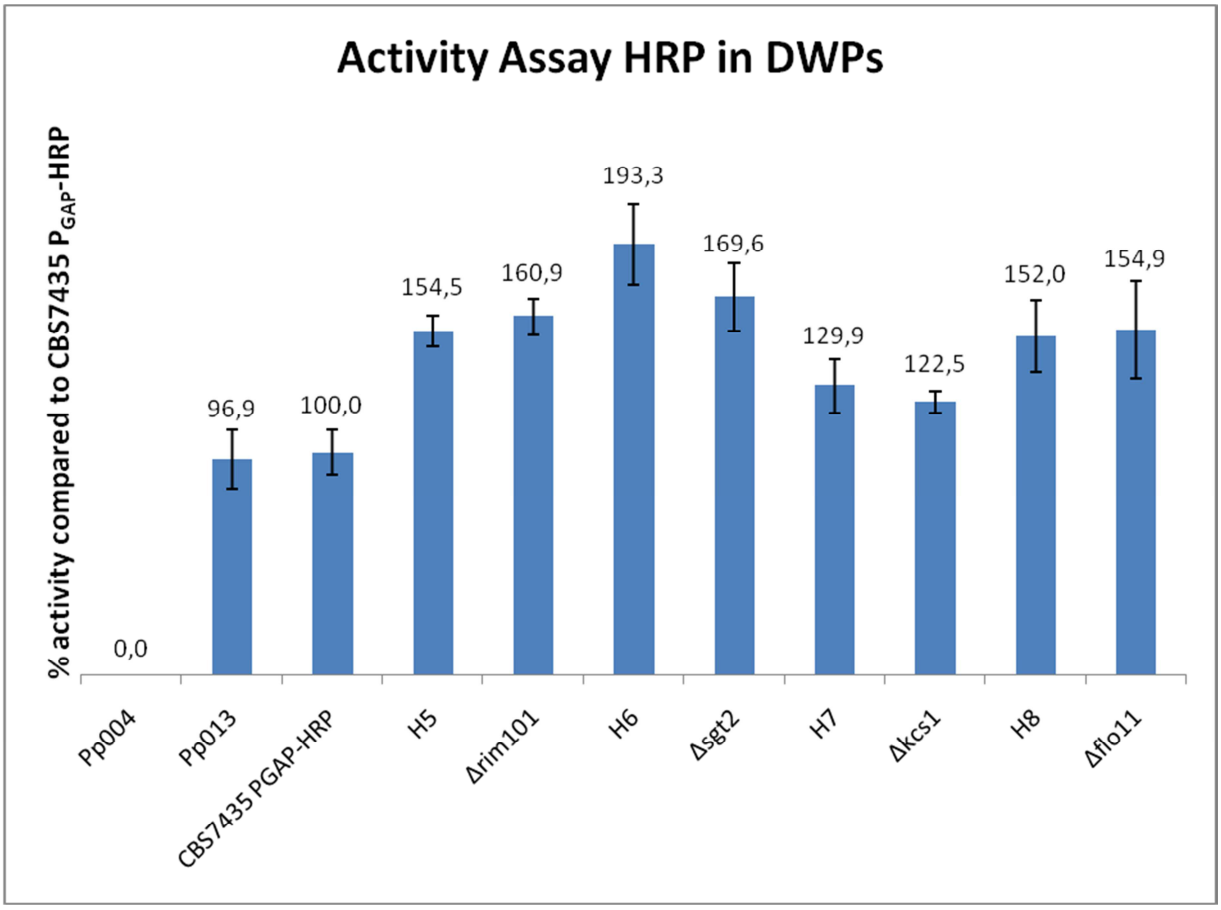


Figure 25: Results of the activity assay performed in deep-well plates. Labeling beneath the bars correlates with the following legend:

Pp004	**GS115 dhis4 wt (negative control)**
Pp013	**GS115 P_{GAP}-HRP**
CBS7435 P_{GAP}-HRP	
H5:	**Strain "H5" from screening**
Δrim101	**Strain with target knock-out of *RIM101* gene**
H6	**Strain "H6" from screening**
Δsgt2	**Strain with target knock-out of *SGT2* gene**
H7	**Strain "H7" from screening**
Δkcs1	**Strain with target knock-out of gene of hypothetical protein found in Strain H7**
H8	**Strain "H8" from screening**
Δflo11	**Strain with target knock-out of gene of hypothetical protein found in Strain H8**

In Figure 25 the activity in the supernatant of strain CBS7435 P_{GAP}-HRP was set to 100% and the results for the other strains were normalized for the same factor and to the OD of their cell suspension. Standard deviations for the measurements (marked as black lines) of the supernatants were moderate.

Results from the original screening could be confirmed for all four mutants and target knockouts, respectively. All knockouts led to a significant increase in HRP secretion behavior compared to the wild type. Furthermore, in all four target knock-outs the measurements led

to very similar results as their corresponding mutants carrying the zeocin cassette. This suggests that in every case the gene responsible for the mutant's enhanced secretion has been identified.

4.6.2 Cell cultivation in shake flasks

For shake flask analysis target knock-out strains were compared to strains from the original screening and the CBS7435 P_{GAP}-HRP wt strain described in Southern blotting. Every strain was investigated in triplicate and measurements were performed after 38 h of cultivation. Results were normalized on the wt which was set to 100% as well as on their OD600.

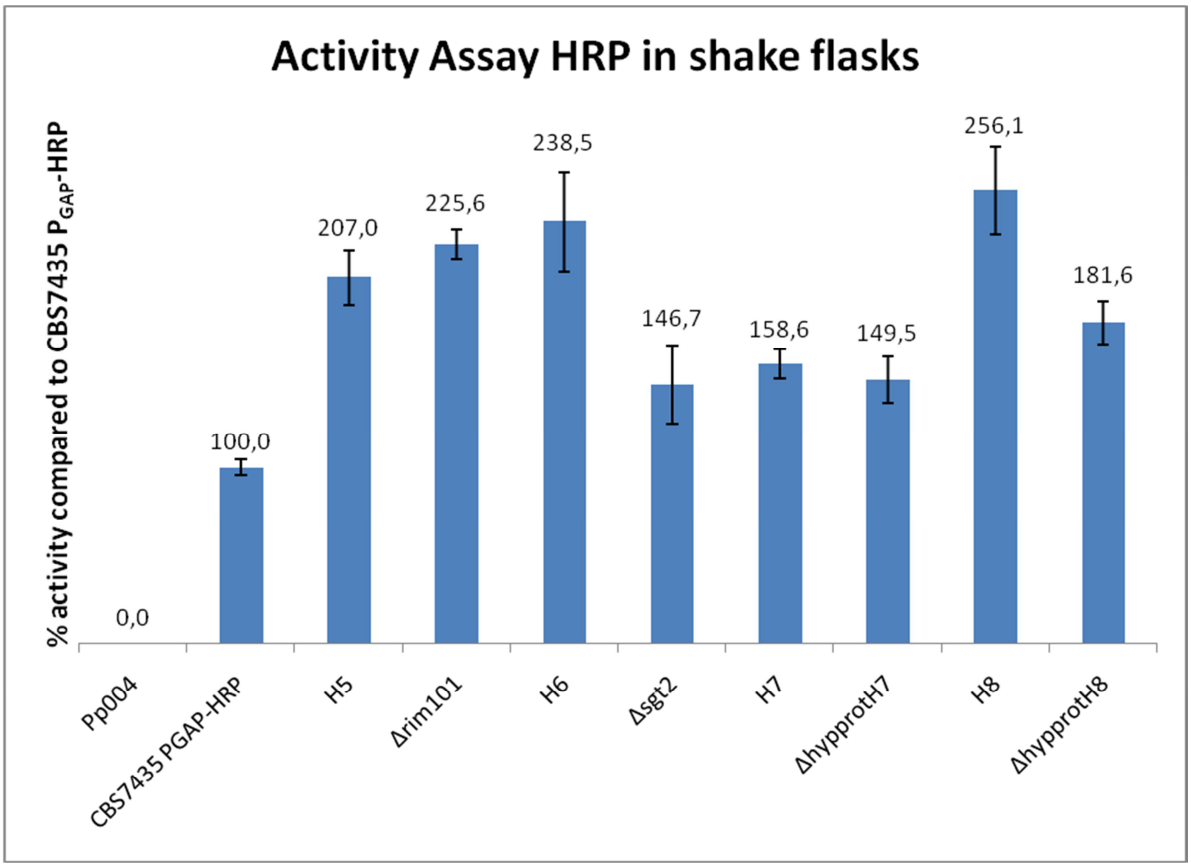


Figure 26: Results of the activity assay performed in shake flasks. Labeling beneath the bars correlates with the following legend:

Pp013	**GS115 P_{GAP} -HRP**
CBS7435 P_{GAP} -HRP	
H5:	**Strain "H5" from screening**
Δrim101	**Strain with target knock-out of RIM101 gene**
H6	**Strain "H6" from screening**
Δsgt2	**Strain with target knock-out of SGT2 gene**
H7	**Strain "H7" from screening**
Δkcs1	**Strain with target knock-out of gene of hypothetical protein found in Strain H7**
H8	**Strain "H8" from screening**
Δflo11	**Strain with target knock-out of gene of hypothetical protein found in Strain H8**

Two of four target knock-out strains showed the same elevation of activity in the supernatant as their respective strain from the original screening. For H5/Δrim101 and H7/Δkcs1 the results matched very well. Still a definite increase compared to the wild type can be found for dflo11 and dsgt2. The more than two-fold increase in HRP secretion observed in the strains H6 and H8 was not reached though in strains Δsgt2 and Δflo11 in shake flasks.

Contrary to the Southern Blots, genome walking experiments suggested two distinct loci of strain H8 being hit by the integration of the zeocin resistance cassette. Anyways, the elevated secretion of HRP found in H8 was confirmed in the target knock-out strain Δflo11, suggesting that we can assume that the right locus was knocked out.

5 Discussion

This thesis describes a process starting from screening of four thousand mutants obtained by random insertion of a zeocin resistance cassette into *P. pastoris*. Strains with altered secretion were chosen for a re-screening and thereafter a number of 43 APLE and 24 HRP secreting strains confirmed former results by clearly having an increase or decrease in secretion efficiency observed by activity measurements in the cell supernatant. By genome walking experiments the loci of integration of 60 clones were elucidated and validation of the results was decided to be done by target knockouts. Fourteen knockout plasmids were constructed, the screening for correct knockout mutants is time consuming though. Therefore, four genes were selected for further elaborations as they were regarded to be of major interest.

Secretion of proteins into the extracellular space is an enormously complex process whose components have not yet been discovered completely. Although the vesicular transport can be regarded as a well-known mechanism we have to assume that proteins not yet associated with secretion may have significant influence on its efficiency. By blasting the proteins found in strains presented in this thesis against the proteome of *S. cerevisiae* possible functional analogies were found. Still one has to be aware of the fact that even high sequence similarities do not necessarily imply identical function in the cell.

Nevertheless, some clues can be deduced from the short discussions presented above. Defects in the maintenance of cell wall stability may provide first indications on what mechanism triggers increased secretion in the strains analyzed. It is suggested that a porous and/or thinner cell wall is posing a barrier easier to be overcome. As an affected cell wall is likely leading to cells being sensitive against extracellular stresses industrial interest might be limited.

An important factor which has to be considered is the comparability of the strains. During construction of target knock-out mutants different undesirable genetic alterations are conceivable. Gene insertion into the genome via double crossover events leads to liberation of parts of genomic DNA with a potential of re-insertion elsewhere. In order to lower the chance of re-integration of functional target genes to be knocked-out in the present study the homologous regions in the knock-out plasmids were chosen not to cut out whole genes but only parts.

Moreover, multiple integrations of foreign DNA into the genome are undesired. This holds true not only for the knock-out plasmids but just as well for the expression cassette of the model protein. Clearly, differences of the copy number have major influence on the amount of protein expressed and, therefore, also on the activity in the supernatant, concealing the effect of the knock-out.

As presented earlier in this work not only HRP expressing mutants were analyzed by genome walking but also APLE expressing ones. A quite interesting aspect of the results is that a couple of genes were found twice or even more often in strains expressing the same secretory protein (e.g. proteins of the Rim pathway or Sgt2) but none of them appears in both lists for APLE and HRP secretion. Also there was no close relationship of gene products of HRP secreting strains compared to the APLE secreting ones implying that possibly no gene with a potential of being a global enhancer of secretion could be identified. Nevertheless one has to keep in mind the comparatively low coverage of genes identified in screening making a cross checking of gene targets and secretory proteins reasonable (see below).

In general one can say that even growth conditions for all strains can hardly ever be achieved. When inoculating the main cultures with aliquots from the starter cultures it is of importance that all the starter cultures are being in the same growth phase and aliquots are added carefully to reach equal OD in the main cultures. These limitations lose importance if more parallel determinations are performed as it is easily accessible in DWPs.

As an outstanding result of this work we have to regard the fact, that for all four characterized knock-outs growth defects in *P. pastoris* were absent. Growth curves confirmed, that not only during lag time after inoculation but also in the exponential phase the knock-out mutants showed identical growth behavior compared to the wild type. Along with this observation the highly increased secretion can be appreciated as an remarkable outcome of this thesis, that can be assumed to be of major biotechnological interest.

5.1 Outlook

Starting with the identification of over 60 different strains showing altered characteristics in protein secretion efficiency in the course of this work four were chosen to be validated by target knock-outs. On the one hand this leaves a great number of targets still to be subjected to the process described here. On the other hand results of this thesis encourage continuing the characterization of one or more genes discussed closely in section 4.4. Increases of measured activity in the supernatant of even more than 100% represent improvements hardly seen by classical genomics. As described, this was done using the constitutive promoter P_{GAP}, which in biotechnological applications is often replaced by inducible promoters such as the promoter of the alcoholoxidase variant 1, P_{AOX1}. This leaves open the interest in further validation of the results by cloning the respective promoter region in front of the model protein and repeating activity assays including a growth phase and an induction phase during fermentation.

A major issue still to be addressed is the influence on secretion of other secretory proteins besides HRP. As mentioned above no gene was found in both lists of HRP and APLE-secreting mutants (see Table 12 and 13). This might indicate that the four genes discussed here lead to better results only for the model protein at hand (HRP). By transforming an APLE expression cassette into the knockout strains instead of an HRP expression cassette this can easily be checked upon. When analyzing knock-outs found in APLE secreting strains the mentioned approach should be done vice versa. The model proteins used in this thesis allow easy activity screening in the supernatant based on simple and cheap reaction components. Other possible model proteins such as i.e. the human growth hormone hGH have to be detected by Western blotting, which is more laborious and secretion levels can not be assessed as accurately. Nevertheless, the use of further model proteins has to be considered in the search for a gene acting as a global secretion enhancer.

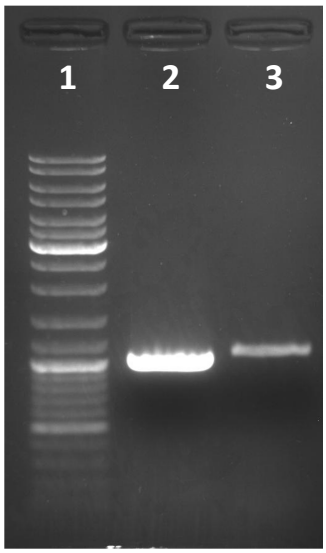
Additionally, by performing a target knock-out of a gene that has not shown to be influencing secretion, effects of genetic manipulations during the process of doing a target knock-out can be observed and potentially be eliminated.

Interestingly, growth conditions in DWP differ significantly from those in shake flasks, in the presented experiments, however, results are confirming each other. During up-scaling of biotechnological processes effects found in small fermentation vessels often cannot be repeated when it comes to the use of bigger volumes. Tests in fermenters are suggested as they would provide information about cell behavior in larger volumes and different conditions regarding e.g. oxygen supply.

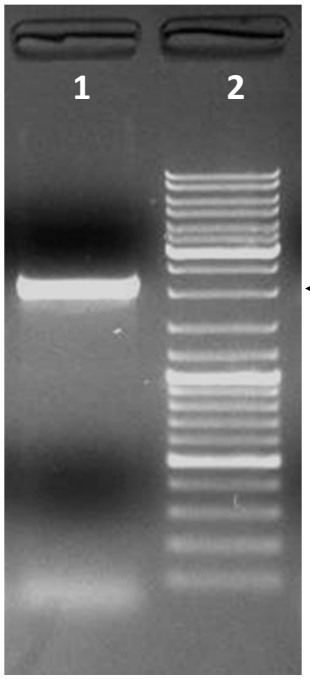
6 Supplementary material

6.1 Construction of knockout plasmids

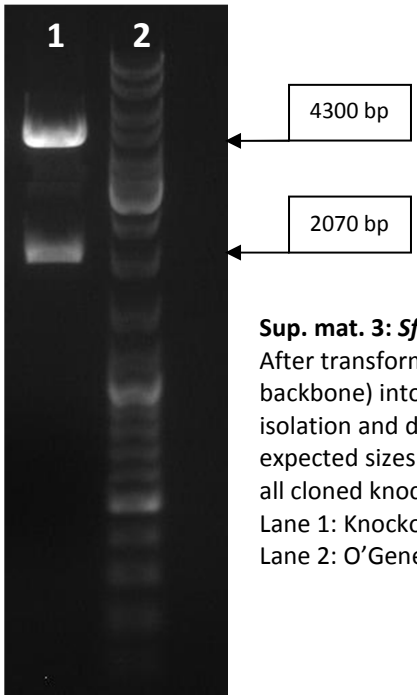
Gel images from the knockout procedure of the SGT2 homologue are taken as examples to document the process of construction of knockout plasmids and the target knockout in *P. pastoris*



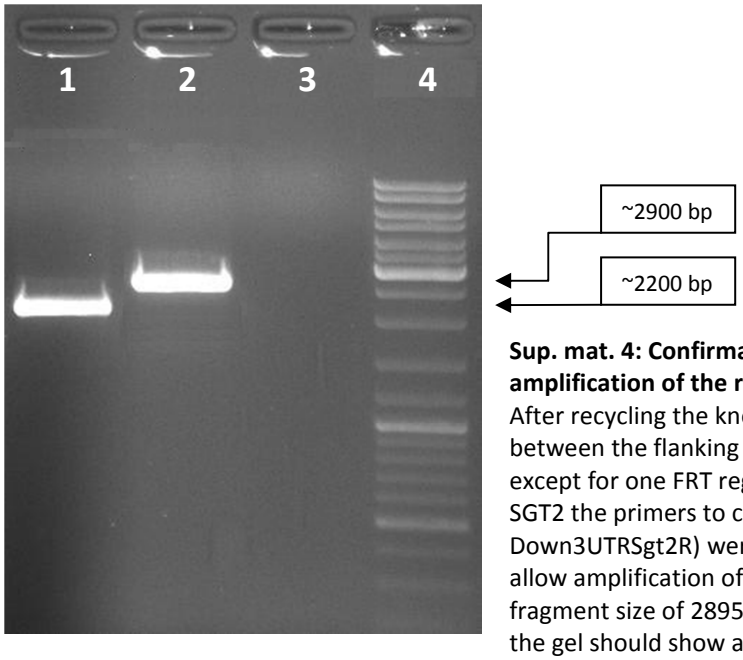
Sup. mat. 1: Amplification of 3'UTR and 5'UTR of knockout plasmid pPKC1 SGT2
By using primers 3UTRSgt2F + 3UTRSgt2R for amplification of the 3'UTR and 5UTRSgt2F + 5UTRSgt2R for the 5'UTR respectively the proper DNA sequences could be amplified from the genome of CBS7435 wt (Polymerase: Phusion)
Lane 1: O'GeneRuler DNA Ladder Mix, 10 µL
Lane 2: PCR product 3'UTR
Lane 3: PCR product 5'UTR



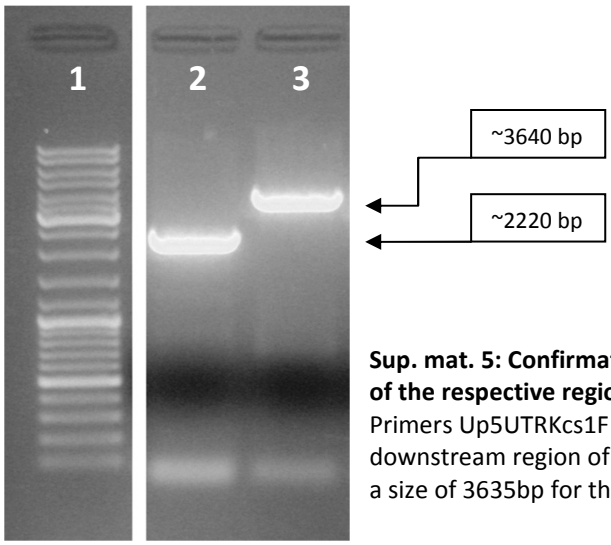
Sup. mat. 2: Overlap extension-PCR product of 3'UTR and 5'UTR of *SGT2*
As primers 3UTRSgt2R and 5UTRSgt2F carry complementary regions the flanking sequences of Sgt2 (see Sup. mat. 1) could be fused by using primers 3UTRSgt2F and 5UTRSgt2R.
Lane 1: oe-PCR product of Sgt2 flanking regions
Lane 2: O'GeneRuler DNA Ladder Mix, 10 µL
By digesting the oe-PCR product with the restriction enzyme *Sfi*I compatible ends to the pPKC1 backbone were generated.



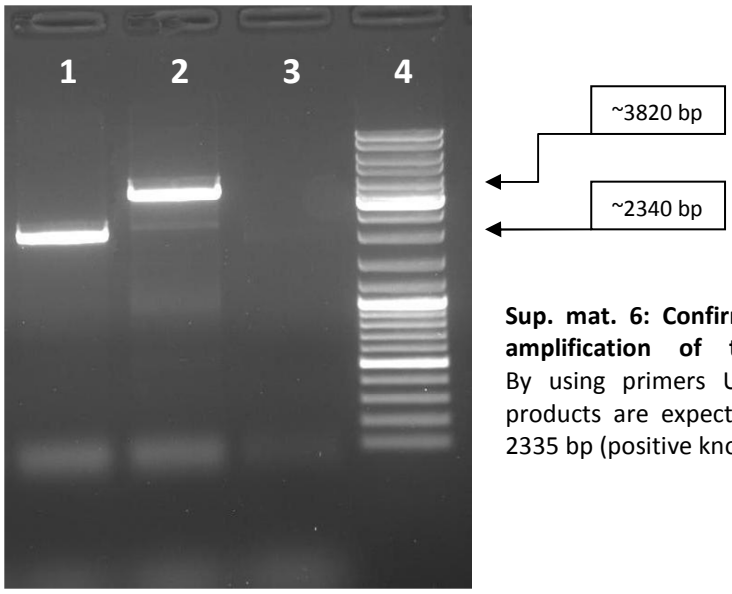
Sup. mat. 3: *Sfi*I-digestion of knockout plasmid pPKC1 *SGT2*
After transformation of the ligated fragments (oe-PCR product and backbone) into *E. coli* TOP10 cells. The correct plasmid was checked via isolation and digestion with *Sfi*I. The picture shows fragments with the expected sizes indicating successful cloning. For confirmation purposes all cloned knockout plasmids were sent for sequencing.
Lane 1: Knockout plasmid pPKC1 Sgt2, digested with *Sfi*I
Lane 2: O'GeneRuler DNA Ladder Mix, 10 µL



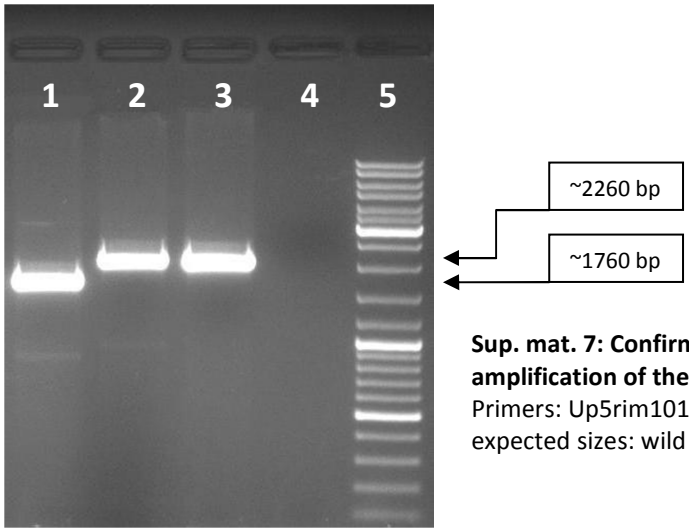
Sup. mat. 4: Confirmation of *sgt2* knockout in *P. pastoris* via amplification of the respective region in the genome.
After recycling the knockout plasmid, the DNA sequence between the flanking regions (3'UTR and 5'UTR) is removed except for one FRT region and a *Sfi*I restriction site. In case of SGT2 the primers to confirm the knockout (Up5UTRSgt2F and Down3UTRSgt2R) were binding in the genome in a way to allow amplification of the whole region. For the wild type a fragment size of 2895 bp is expected, for a positive knockout the gel should show a band at 2207 bp. The results in the picture confirm this.



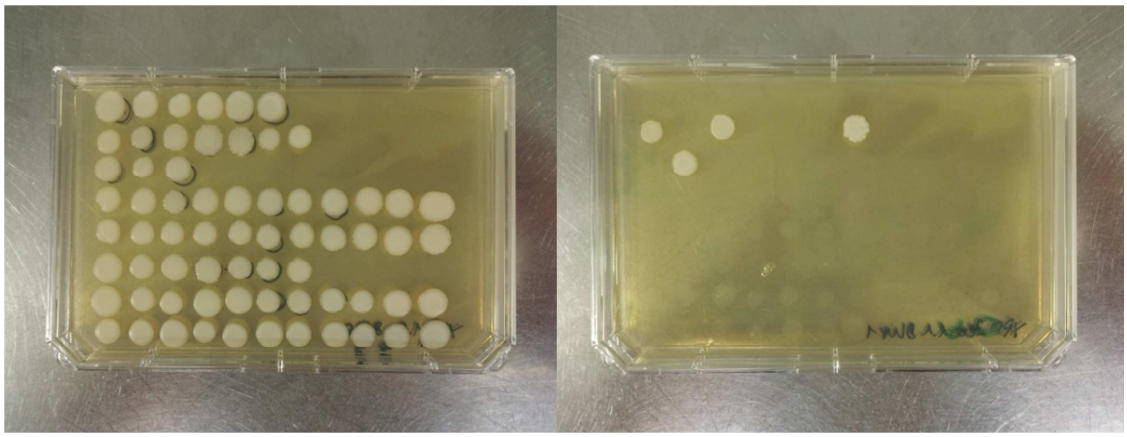
Sup. mat. 5: Confirmation of *kcs1* knockout in *P. pastoris* via amplification of the respective region in the genome
Primers Up5UTRKcs1F and Down3UTRKcs1R are binding in the upstream and downstream region of KCS1 gene. Products of the PCR are expected to have a size of 3635bp for the wild type and 2224bp in case of a positive knockout.



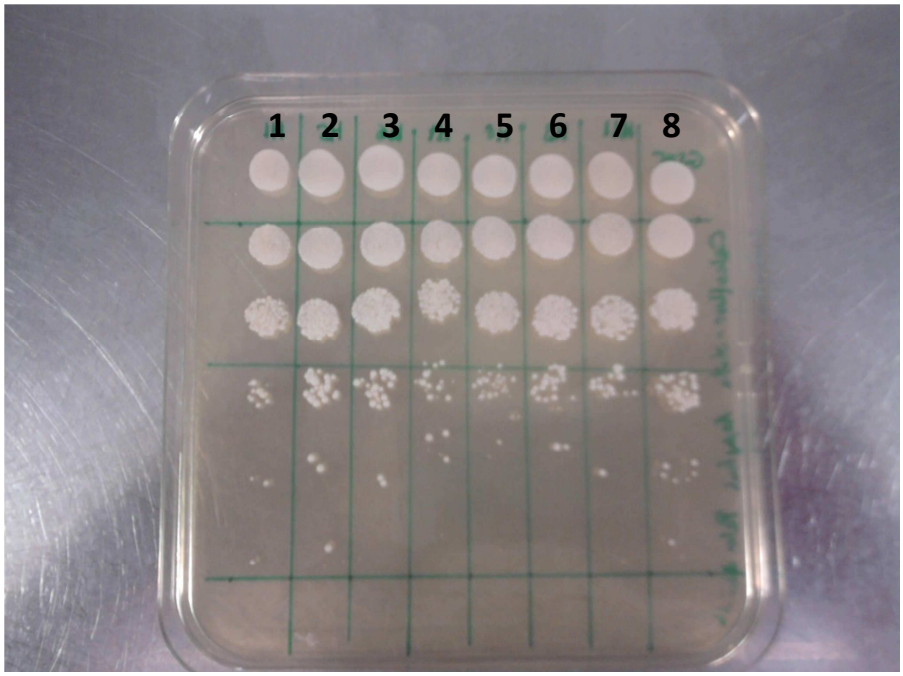
Sup. mat. 6: Confirmation of *flo11* knockout in P. pastoris via amplification of the respective region in the genome
By using primers Up5hypprotH8F and Down3hypprotH8R PCR products are expected to have sizes of 3822bp (wild type) and 2335 bp (positive knockout of *flo11*) respectively.



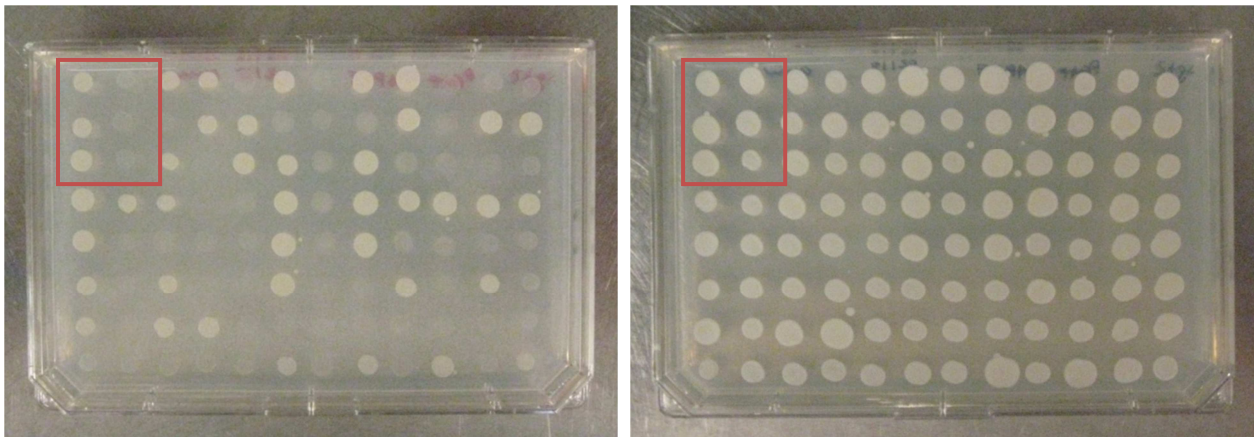
Sup. mat. 7: Confirmation of *rim101* knockout in *P. pastoris* via amplification of the respective region in the genome
Primers: Up5rim101F and Down3rim101R
expected sizes: wild type: 2260 bp; knockout: 1758 bp



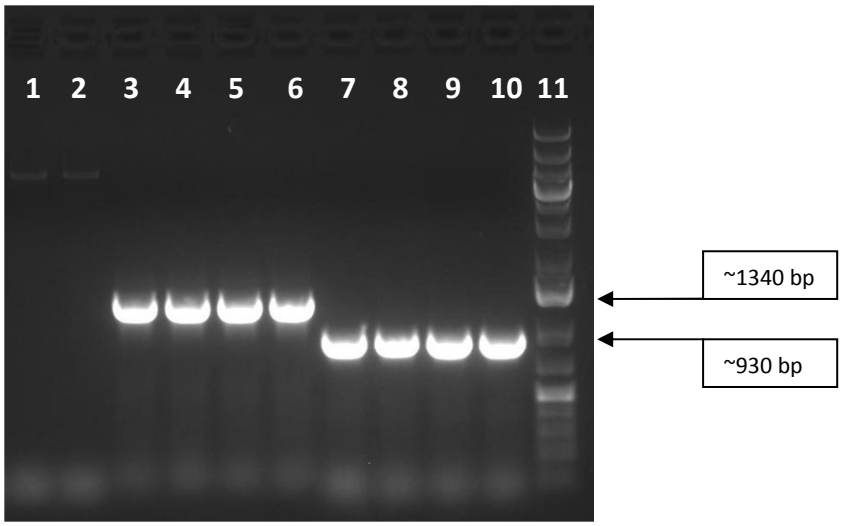
Sup. mat. 8: Removal of the Zeocin marker by induction of Flippase gene
After induction single colonies were picked and pinned onto a YPD plate (left) as well as onto a YPD-Zeo plate (right). About 95% of colonies were losing their zeocin resistance proving the effectiveness of the process. In following experiments protocol specifications were abbreviated (induction for ~36 h) leading to a percentage of 50-60% of positive cells.



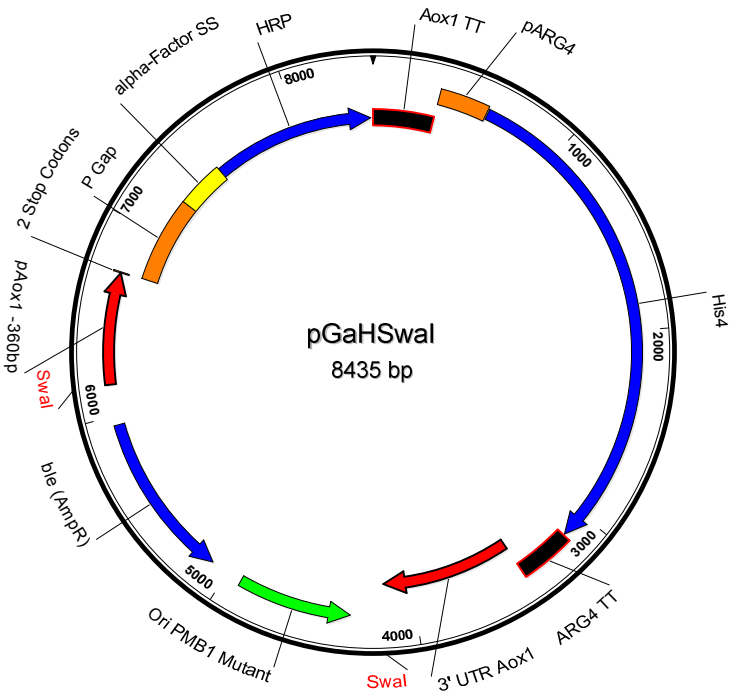
Sup. mat. 9: Calcofluor white Spotting of different HRP secreting mutants
S. cerevisiae cell wall mutants have shown to be sensitive against the fluorescent dye calcofluor white. Knockouts of several Rim genes (rim101, rim20) in turn lead to an increased resistance. Similar phenotypes did not appear in the respective *P. pastoris* strains as it can be seen here. Several different concentrations of calcofluor white were tested. Concentrations of more than 10 µg/mL have shown to inhibit growth completely. From left to right cell suspensions (OD=0.5) of the following strains were applied, including dilutions of 1:10 – $1:10^6$ in the rows beneath: Lane 1: H1; Lane 2: H5; Lane 3: H6; Lane 4: H7; Lane 5: H8; Lane 6: H9; Lane 7: H11; Lane 8: Pp004



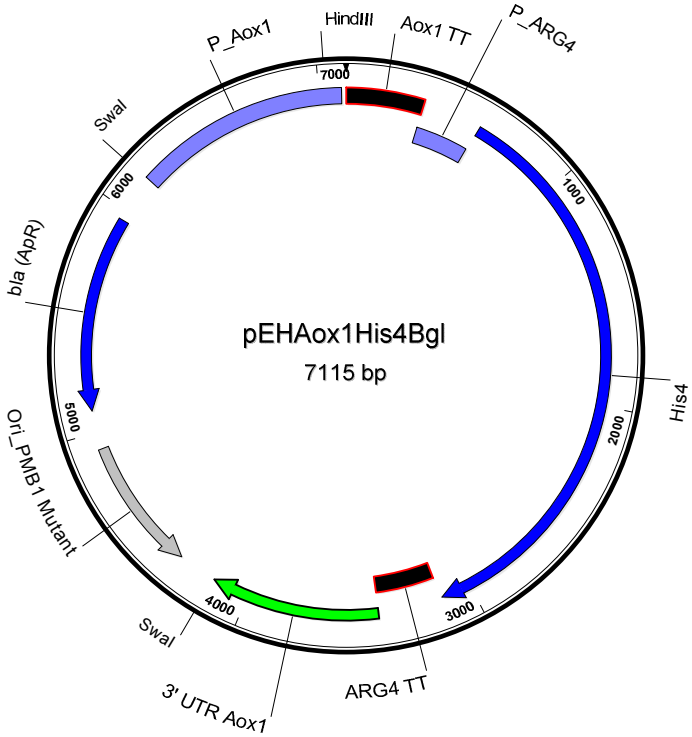
Sup. mat. 10: Screening for MutS phenotype
After transformation of the HRP expression cassette screening was done on minimal medium containing methanol (1%) as sole carbon source (left). The difference in growth of Mut+ and MutS strains is easily visible. The red rectangle indicates MutS and Mut+ control strains. On average more than 50% of transformants with histidin prototrophy show MutS phenotype,indicating a correct integration by homologous recombination (size of the flanking homologous regions: 570 bp and 740 bp)
In parallel the colonies were pinned onto minimal medium plates containing glucose (right).



Sup. mat. 11: Confirmation of correct integration of the HRP expression cassette at the *AOX1* locus
primer pairs: Lane 1+3-6: 5'UTRAox1F (binding in upstream region of *AOX1* locus)
AlphaFSSR (binding in the α-factor signal sequence)
expected fragment size: 1339bp
Lane 2+7-10: 3'UTRAox1R (binding in the downstream region of AOX1 locus)
Arg4TTF (binding in the Arg4 transcription terminator sequence)
expected fragment size: 930 bp
Lane 1 + 2: negative control, CBS7435 wt
Lane 3 + 7: CBS7435 Δrim101 P_{GAP}-HRP
Lane 4 + 8: CBS7435 Δsgt2 P_{GAP} -HRP
Lane 5 + 9: CBS7435 Δkcs1 P_{GAP} -HRP
Lane 6 + 10: CBS7435 Δflo11 P_{GAP} -HRP



Sup. mat. 12: As a positive control for Southern blots employing the P_{GAP} probe the plasmid pGaHSwaI was digested with the restriction enzyme *Swa*I, leading to a linear fragment of 6440 bp.



Sup. mat. 13: As a positive control for Southern blots employing the P_{AOX1} or *HIS4* probe the plasmid pEHAox1His4Bgl was digested with the restriction enzyme *Swa*I, leading to a linear fragment of 7115 bp.

List of figures

Figure 1: Intracellular transport pathways..... 8

Figure 2: Sequence homology is used to restore genome integrity after double-strand breaks in a process called Homologous recombination [14]..... 11

Figure 3: Integrated DNA sequence for random insertion into the genome of *P. pastoris*. pBR322ori: replicon sequence, allows plasmid maintenance in *E. coli* pTEF1: eukaryotic promoter region pEM7: prokaryotic promoter region Zeocin: Zeocin resistance gene CYC1TT: transcriptional terminator 13

Figure 4: Assembly for Southern Blotting [20] By capillary transfer the transfer buffer moves along a filter paper bridge and through the agarose gel, transferring DNA which adheres to the Nylon membrane. 22

Figure 5: Chemical structure of digoxygenin (DIG)-11-dUTP 23

Figure 6: Plasmid pGaHSwaI carrying the HRP expression cassette for transformation into *AOX1* locus of *P. pastoris*. After linearization via restriction enzyme *Swa*I the construct is flanked by two homologous regions (part of *AOX1* promoter region and 3'UTR AOX1 respectively). The HRP gene is fused to an α-factor signal sequence and under control of a GAP promoter. The *HIS4* gene complements *his4* auxotrophy in *P. pastoris* allowing easy screening on minimal medium. 25

Figure 7: AOX 1 locus and primer binding sites..... 26

Figure 8: Map of the knock-out plasmid pPpKC1 For information about features of the plasmid please see Table 11..... 30

Figure 9: 3' and 5' homologue regions of knock-out target Sgt2 gene..... 31

Figure 10: Primer design for amplification of the 3'UTR of *SGT2*. In addition to the 3' homologous region the forward primer carries a *Sfi*I restriction site plus 2 bases for proper recognition by the restriction enzyme later on, the reverse primer is attached to a sequence homologous to the 5'UTR to allow overlap-extension PCR. Mind the appearance of a unique restriction site (*Ssp*I) for linearizing the knockout plasmid before transformation. 32

Figure 11: Situation in the genome after integration of a Sgt2-knockout plasmid in *P. pastoris*. In the flanking region upstream to the integrated sequence a short stretch of the target gene remained to avoid the release of a functional gene during excision. Green and red arrows indicate primer binding sites for screening purposes..... 33

Figure 12: If recombination occurs between two homologous signal sequences (red arrows, FRT regions) the flanked DNA region gets excised as a circular molecule leaving one FRT region and (in this case) a *Sfi*I restriction site in the genome. 34

Figure 13: Overview of Template blocking PCR (adapted from Jung-Hoon et al. [29]) 36

Figure 14: Overview of important steps in Genome Walker™ Kit Strategy. After digestion of genomic DNA fragments are ligated to a double stranded cassette carrying an amine group at the 3' end pointing outward (represented by 'N' in the figure). In the next step this amine group blocks extension and nonspecific amplification is prevented by the absence of a priming site. (Adapted from [30]) 37

Figure 15: Southern Blotting of HRP secreting hits Probe: Zeocin resistance gene DIG-labeled Ladder: DNA Molecular Weight Marker II, DIG-labeled Positive control: plasmid DNA containing one Zeocin resistance gene sequence Restriction enzymes: *EcoR*V cuts in terminator of Zeocin reistance cassette (ca. 270 bp + 930 bp fragments), *Hind*III does not cut in cassette, therefore per gene copy two fragments expected if cut with *EcoR*V and one if cut with *Hind*III (see also Figure 3)..... 42

Figure 16: Results of HRP activity measurements in culture supernatant and cell extract. Every strain was investigated in triplicate and results represent means and STD. Pp004 represents the CBS7435 dhis4 strain without HRP expression cassette, Pp013 represents the HRP secreting GS115 strain with an expression cassette integrated at the AOX1 locus..... 43

Figure 17: HRP expression cassette integrated at the AOX1 locus. P_{GAP}, P_{AOX1} and HIS4 probes of the following Southern blots bind to the respective sites in the figure..... 45

Figure 18: Southern Blot for P_{GAP}-driven expression of HRP and APLE #1 Probe: GAP promoter DIG-labeled Ladder: DNA Molecular Weight Marker II, DIG-labeled Positive control: plasmid DNA containing one GAP promoter sequence Restriction enzymes: *Pst*I cuts at end of HIS4 (part of integrated cassette), *Hind*III does not cut in integrated cassette. None of the restriction enzymes cuts in P_{GAP} – sequence, therefore only one labeled band per copy in genome is expected. One copy of GAP promoter is present also in the wild type. 46

Figure 19: Southern Blot for P_{GAP}-driven expression of HRP and APLE #2 Probe: GAP promoter DIG-labeled Ladder: DNA Molecular Weight Marker II, DIG-labeled Positive control: Plasmid DNA containing one GAP promoter sequence Restriction enzymes: *Pst*I cuts at end of HIS4 (part of integrated cassette), *Hind*III does not cut in integrated cassette. None of the restriction enzymes cuts in P_{GAP} – sequence, therefore only one labeled band per copy in genome is expected. One copy of GAP promoter is present also in the wild type. 47

Figure 20: Southern Blot for P_{AOX1}-driven expression of HRP and APLE #1 Probe: AOX promoter DIG-labeled Ladder: DNA Molecular Weight Marker II, DIG-labeled Positive control: Plasmid DNA containing one AOX promoter sequence Restriction enzymes: *Pst*I cuts at end of *HIS4* (part of integrated cassette), *Hind*III does not cut in integrated cassette. None of the restriction enzymes cuts in P_{AOX} – sequence, therefore only one labeled band per copy in genome is expected. 48

Figure 21: Southern Blot *HIS4* #1 Probe: HIS4 gene DIG-labeled Ladder: DNA Molecular Weight Marker II, DIG-labeled Positive control: Plasmid containing *HIS4* gene Restriction enzyme: *Hind*III does not cut in integrated cassette; Therefore only one labeled band per copy in genome expected 49

Figure 22: Southern Blot for P_{AOX1}-driven expression of HRP and APLE #2 Probe: AOX promoter DIG-labeled Ladder: DNA Molecular Weight Marker II, DIG-labeled Positive control: Plasmid DNA containing one AOX promoter sequence. Restriction enzymes: *Pst*I cuts at end of HIS4 (part of integrated cassette), *Hind*III does not cut in integrated cassette. None of the restriction enzymes cuts in P_{AOX} – sequence, therefore only one labeled band per copy in genome is expected. 50

Figure 23: Southern Blot HIS4 #2 Probe: HIS4 gene DIG-labeled Ladder: DNA Molecular Weight Marker II, DIG-labeled Positive control: Plasmid containing *HIS4* gene Restriction enzyme: *Hind*III does not cut in integrated cassette; Therefore only one labeled band per copy in genome expected 51

Figure 24: Growth curves of strains from the screening performed by Christine Winkler (H5-H8), the respective target-knock-out strains (Δrim101, Δsgt2, Δkcs1 and Δflo11) and the parent strains Pp004 (*P. pastoris* CBS7435 dhis4) and Pp013 (*P.pastoris* GS115 P_{GAP}-HRP). 56

Figure 25: Results of the activity assay performed in deep-well plates..... 58

Figure 26: Results of the activity assay performed in shake flasks..... 59

References

1. Cregg JM, Tolstorukov I, Kusari A et al. (2009) Expression in the Yeast Pichia pastoris. Methods in Enzymology **463**: 170–189

2. Li P, Anumanthan A, Gao X et al. (2007) Expression of Recombinant Proteins in Pichia Pastoris. Applied Biochemical Biotechnology **142**: 105–124

3. Cereghino JL, Cregg JM (2000) Heterologous protein expression in the methylotrophic yeast Pichia pastoris. FEMS Microboiology Reviews **24**: 45–66

4. Bonifacino JS, Glick BS (2004) The Mechanism of Vesicle Budding and Fusion. Cellular Molecular Biology **116**: 153–166

5. Veitch NC (2004) Horseradish peroxidase: a modern view of a classic enzmye. Phytochemistry **65**: 249–259

6. Welinder KG (1976) Covalent Structure of the Glycoprotein Horseradish peroxidase. FEBS Letters **72**: 19–23

7. Veitch NC, Smith AT (2000) Horseradish peroxidase. Advances in Inorganic Chemistry **51**: 107–162

8. Dominguez de Maria P, Garcia-Burgos C, Bargeman G et al. (2007) Pig liver esterase (PLE) as biocatalyst in organic synthesis: From nature to cloning and to practical applications. ChemInform **10**: 1439–1452

9. Heymann E, Junge W (1979) Characterization of the isoenzymes of pig-liver esterase. European Journal of Biochemistry **95**: 509–513

10. Hermann M, Kietzmann MU, Ivancic M et al. (2008) Alternative pig liver esterase (APLE) – Cloning, identification and functional expression in Pichia pastoris of a. versatile new biocatalyst. Journal of Biotechnology **133**: 301–310

11. Dudáš A, Chovanec M (2004) DNA double-strand break repair by homologous recombination. Mutation Research **566**: 131–167

12. Ivanov EL, Sugawara Neal, Haber James E. (1996) Genetic Requirements for the Single-Strand Annealing Pathway of Double-Strand Break Repair in Saccharomyces cerevisiae. Genetics **142**: 693–704

13. Pastink A, Eeken JC, Lohman PH (2001) Genomic integrity and the repair of double-strand DNA breaks. Mutation Research **480**: 37–50

14. Wyman C, Ristic D, Kanaar R (2004) Homologous recombination-mediated double-strand break repair. DNA repair **3**: 827–833

15. Hefferin ML, Tomkinson AE (2005) Mechanism of DNA double-strand break repair by non-homologous end joining. DNA repair **4**: 639–648

16. Balamurugan V, Reddy G, Suryanarayana (2007) Pichia pastoris: A notable heterologous expression system for the production of foreign proteins—Vaccines. Indian Journal of Biotechnology **6**: 175–186

17. Higgins DR, Cregg JM (1998) Pichia Protocols. Methods in Molecular Biology, vol 103. Humana Press Inc., New Jersey

18. Näätsaari L, Mistelberger B, Ruth C et al. (2012) Deletion of the Pichia pastoris KU70 Homologue Facilitates Platform Strain Generation for Gene Expression and Synthetic Biology. Gene expression and Synthetic Biology **7**: 1–13

19. Hoffmann C, Winston F (1987) A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of Escherichia coli. Gene expression and Synthetic Biology **57**: 267–272

20. G-Biosciences Southern Blot Analysis. www.GBiosciences.com

21. Roche Applied Science (2004) DIG DNA Labeling and Detection Kit. Instruction Manual

22. Höltke H, Ankenbauer W, Rein K et al. (1995) The Digoxigenin (DIG) System for non-radioactive labeling and detection of nucleic acids - An overview. Cellular Molecular Biology **41**: 883–905

23. Lin-Cereghino J, Xiong W (2005) Condensed protocol for competent cell preparation and Transformation of pichia pastoris. BioTechniques

24. Gallati H (1979) Horseradish peroxidase: a study of the kinetics and the determination of optimal reaction conditions, using hydrogen peroxide and 2, 2'-azinobis 3-ethylbenzthiazoline-6-sulfonic acid (ABTS) as substrates. Journal of clinical chemistry and clinical biochemistry **17**

25. Ram AF, Wolters A, Hoopen RT et al. (1994) A New Approach for Isolating Cell Wall Mutants in Saccharomyces cerevisiae by Screening for Hypersensitivity to Calcofluor White. Yeast **10**: 1019–1030

26. Groot P de, Ruiz C, Vazquez de Aldana C et al. (2001) A genomic approach for the identification and classification of genes involved in cell wall formation and its regulation in Saccharomyces cerevisiae. Comparative and functional Genomics **2**: 124–142

27. Reuß O, Vik A, Kolter R et al. (2004) The SAT1 flipper, an optimized tool for gene disruption in Candida albicans. Gene expression and Synthetic Biology **341**: 119–127

28. Bode J (2013) Flippase. http://de.wikipedia.org/wiki/Rekombinase. Accessed 08 May 2013

29. Jung-Hoon B, Jung-Hoon S (2010) Template Blocking PCR: An Advanced PCR Technique for Genome Walking. Analytical Chemistry **398**: 112–116

30. Clontech (2007) Genome Walker Universal Kit User Manual, Mountain View, CA

31. Saiardi A, Erdjument-Bromage H, Snowman AM et al. (1999) Synthesis of diphosphoinositol pentakisphosphate by a newly identified family of higher inositol polyphosphate kinases. Current Biology **9**: 1323–1326

32. Saiardi A, Nagata E, Luo H et al. (2001) Identification and characterization of a novel inositol hexakisphosphate kinase. Journal of Biological Chemistry **276**: 39179–39185

33. Huang KN, Symington LS (1995) Suppressors of a Saccharomyces cerevisiae pkc1 mutation identify alleles of the phosphatase gene PTC1 and of a novel gene encoding a putative basic leucine zipper protein. Genetics **141**: 1275–1285

34. Dubois E, Scherens B, Vierendeels F et al. (2002) In Saccharomyces cerevisiae, the Inositol Polyphosphate Kinase Activity of Kcs1p Is Required for Resistance to Salt Stress, Cell Wall Integrity, and Vacuolar Morphogenesis. Journal of Biological Chemistry **277**: 23755–23763

35. Paravicini G, Cooper M, Friedli L et al. (1992) The osmotic integrity of the yeast cell requires a functional PKC1 gene product. Molecular and cellular biology **11**: 4896–4905

36. Castrejon F, Gomez A, Sanz M et al. (2006) The RIM101 pathway contributes to yeast cell wall assembly and its function becomes essential in the absence of mitogen-activated protein kinase Slt2p. Eukaryotic Cell **5**: 507–517

37. Weishi L, Mitchell AP (1997) Proteolytic activation of Rim1p, a positive regulator of yeast sporulation and invasive growth. Genetics **145**: 63–73

38. Peñalva M, Tilburn J, Bignell E et al. (2008) Ambient pH gene regulation in fungi: making connections. Trends in Microbiology **16**: 291–300

39. Serrano R, Bernal D, Simon E et al. (2004) Copper and iron are the limiting factors for growth of the yeast Saccharomyces cerevisiae in an alkaline environment. Journal of Biological Chemistry **279**: 19698–19704

40. Marchler-Bauer A, Lu S, Anderson J et al. (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acid Research **39**: 225–229

41. Chartron JW, Clemons WM, Suloway C (2012) The complex process of GETting tail-anchored membrane proteins to the ER. Current Opinion in Structural Biology **22**: 217–224

42. Schuldiner M, Metz J, Schmid V et al. (2008) The GET complex mediates insertion of tail-anchored proteins into the ER membrane. Cellular Molecular Biology **134**: 634–645