MARKUS KOLMBAUER

Construction of protease-deficient strains of the methylotrophic yeast *P. pastoris* for heterologous protein expression

Masterarbeit

zur Erlangung des akademischen Grades eines
Master of Science (MSc)
an der naturwissenschaftlichen Fakultät der TU Graz

Mudassar Ahmad
Univ.-Prof. Dipl.-Ing. Dr. techn. Helmut Schwab

Institut für Molekulare Biotechnologie
Technische Universität Graz

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen / Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtliche und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Graz, am ……. ……..

DANKSAGUNG

Da die Fertigstellung meiner Masterarbeit letztendlich auch den Abschluss meines Studiums bedeutet, möchte ich die Gelegenheit nutzen, um einigen Personen meinen herzlichsten Dank auszusprechen.

Zunächst danke ich meiner Familie und speziell meinen Eltern dafür, dass sie mir mein Studium vor allem finanziell ermöglicht haben und mir währenddessen den nötigen Rückhalt gaben, um nie das Ziel aus den Augen zu verlieren. Weiters möchte ich mich bei meinen Studienkollegen, Freunden und speziell bei meiner lieben Katrin einerseits für die moralische Unterstützung und andererseits für die unvergessliche Studienzeit bedanken.

Besonders danke ich Prof. Helmut Schwab, der es mir ermöglichte, meine Masterarbeit erfolgreich durchzuführen und bei den vielen Gruppenseminaren mit nützlichen Ratschlägen weiterhalf.

Ganz großer Dank gebührt natürlich auch Mudassar Ahmad für die ständige Betreuung während der Masterarbeit. Er stand mir mit Rat und Tat bei meiner Arbeit im Labor zur Seite und beantwortete mit viel Geduld all meine Fragen gern auch mehrmals. Vor allem aber brachte er mir bei, in größeren Maßstäben zu denken und zu arbeiten. Auch meinen lieben Laborkolleginnen und -kollegen möchte ich an dieser Stelle für die schöne Zeit im Labor danken.

ABSTRACT

The methylotrophic yeast *Pichia pastoris* has developed into a highly successful system for heterologous protein production during the past 15 years, as it allows expressing foreign proteins at milligram-to-gram quantities. A great advantage of *P. pastoris* is its suitability for protein secretion, because of the relatively low levels of endogenous secreted proteins and the subsequently simplified recovery and purification steps. A common problem occurring during heterologous protein expression is degradation of secreted recombinant target proteins by proteases and one strategy to reduce proteolysis is the use of protease-deficient strains. As an initial solution to this problem a set of protease-deficient strains was constructed by respective deletion of 27 *P. pastoris* protease genes. In a first step gene specific knockout cassettes were generated, which carry a Zeocin^{TM} resistance gene flanked by regions homologous to the 5´UTRs and 3´UTRs of the selected proteases. In this strategy *P. pastoris* CBS7435 strain was transformed with the respective linear knockout cassette, where replacement of the target ORF occurred by homologous recombination. After PCR-based screening for site-specific integration, the knockout cassette was removed by a flippase recombinase system in positive clones. The removal and subsequent deletion of the major part of the ORF sequence was verified by PCR and sequencing. Analysis of cell health and growth did with one exception not reveal any obvious aberrations for the generated protease-deficient *P. pastoris* strains compared to wild type. The established set of protease-deficient *P. pastoris* strains can be useful not only for practical application in protein production with regard to multiple gene deletions, but also for further examinations of proteolytic processes.

KURZFASSUNG

Die methylotrophe Hefe *Pichia pastoris* hat sich in den letzten 15 Jahren zu einem sehr erfolgreichen System für die Herstellung von rekombinanten Proteinen entwickelt. Da *P. pastoris* relativ geringe Mengen an endogenen Proteinen sekretiert und somit die nachfolgenden Reinigungsschritte wesentlich erleichtert werden, ist Sekretion des rekombinanten Zielproteins die bevorzugte Wahl. Aber ein generelles Problem, das während der Proteinexpression auftritt, stellt dabei die Degradierung des sekretierten Zielproteins durch proteolytische Enzyme dar und eine Möglichkeit dies zu reduzieren ist die Verwendung von Stämmen, die eine Deletion in den entsprechenden Protease-Genen tragen. Als eine anfängliche Lösung für dieses Problem wurde ein Set an *P. pastoris* Stämmen generiert, in denen 27 Protease-Gene einzeln deletiert wurden. Die erste Aufgabe bestand darin, Gen-spezifische Knockout-Kassetten herzustellen, die ein Zeocin-Resistenzgen flankiert von DNA-Abschnitten, die homolog zu den 5´UTR und 3´UTR Regionen der gewählten Protease sind, tragen. In weiterer Folge wurde der Wildtypstamm CBS7435 mit den linearen Knockout-Kassetten transformiert, wobei das Ziel-Gen durch homologe Rekombination ersetzt wurde. Nachdem die ortsspezifische Integration mittels PCR nachgewiesen worden war, erfolgte die Entfernung der Knockout-Kassetten mithilfe eines Flippase-Rekombinase-Systems in positiven Klonen. Die daraus resultierende Deletion der gesamten ORF-Sequenz wurde mittels PCR und Sequenzierung verifiziert. Die Analyse des Zellwachstums und der generellen Lebensfähigkeit der generierten Knockout-Stämme zeigte mit einer Ausnahme keine offensichtlichen Beeinträchtigungen verglichen mit dem Wildtypstamm. Das etablierte Set an *P. pastoris* Stämmen kann nicht nur für die praktische Anwendung bei der Proteinherstellung in Hinblick auf multiple Gen-Deletionen nützlich sein, sondern auch für weitere Untersuchungen von proteolytischen Prozessen.

TABLE OF CONTENTS

1 INTRODUCTION

1.1 The expression system *Pichia pastoris*

1.1.1 General features of *Pichia pastoris*

Pichia pastoris is a methylotrophic yeast and therefore able to utilize methanol as a sole source of carbon and energy. As a eukaryotic microorganism, *P. pastoris* has developed into a highly successful system for heterologous protein production during the past 15 years. One of several reasons for the increasing popularity of this particular expression system is the ability of *P. pastoris* to produce foreign proteins at milligram-to-gram quantities. Compared with mammalian cells, it does not require a complex growth medium or special culture conditions and the techniques needed for the molecular genetic manipulation are quite simple and similar to those of *Saccharomyces cerevisiae*, e.g. gene targeting, gene replacement, DNA-mediated transformation and cloning by functional complementation. A major advantage of *P. pastoris* over bacterial expression systems is that it can perform many of the eukaryotic post-translational modifications, such as O- and N-linked glycosylation, disulfide bond formation, folding, certain types of lipid addition and proteolytic processing. Furthermore, it is very useful that the *P. pastoris* expression system is available as a commercially available kit [1][2].

Foreign proteins can be produced either intracellularly or extracellularly in *P. pastoris* [1]. There is a great advantage of using this yeast for protein secretion, even at high molecular weights, because of the relatively low levels of endogenous secreted proteins and the subsequent simplified recovery and purification steps [3].

The most widely and successfully used promoter for heterologous protein expression is the *AOX1* promoter from the alcohol oxidase 1 (*AOX1*) gene [4]. It is a very strong and tightly regulated promoter whereas methanol is essential to the induction of high levels of transcription. However, there are circumstances in which the *AOX1* promoter might not be suited and therefore alternative promoters such as the *GAP*, *FLD1*, *PEX8* and *YPT1* promoters are used as well [2].

To express any foreign gene in *P. pastoris* three fundamental steps are required: 1) insertion of the target gene into an expression vector; 2) integration of the expression vector into the host genome; and 3) screening for possible strains expressing the foreign protein. There are several auxotrophic markers and antibiotic resistance genes available as selection markers for *P. pastoris* and all expression vectors have been designed as *Escherichia coli*/*P. pastoris*

shuttle vectors, containing an origin of replication for plasmid maintenance in *E. coli* and markers functional in one or both organisms [2].

1.1.2 Methanol metabolism

Methylotrophic organisms, which appear to be ubiquitous in nature, have the ability to use reduced C_1-compounds as the sole source of carbon and energy. Prokaryotic methylotrophs can utilize a variety of C_1-compounds such as methane, methanol and methylamine, while eukaryotic methylotrophs are only able to use methanol as a carbon source and therefore a diverse range of metabolic pathways for assimilating and dissimilating C_1-compounds exists [5]. The group of eukaryotic methylotrophic organisms is limited to four genera: *Hansenula, Candida, Turolopsis* and *Pichia* [6].

The conceptual basis for the *P. pastoris* expression system stems from the observation that some of the enzymes required for methanol utilization are present at substantial levels only when cells are grown on methanol [7]. During growth on this substrate, several key enzymes, e.g. alcohol oxidase, catalase, formaldehyde dehydrogenase and dihydroxyacetone synthase, are present in high amounts and peroxisomes proliferate. The regulation of the synthesis of these enzymes occurs at the transcriptional level and methanol assimilation depends on a carbon-source-dependent repression/derepression/induction mechanism. It is strongly repressed by multicarbon sources such as glucose and glycerol, but highly induced by methanol [8].

The first step in the methanol utilization pathway is the oxidation of the relatively toxic compound methanol, which results in the formation of two other very reactive compounds, formaldehyde and hydrogen peroxide (Figure 1). This reaction is catalyzed by the enzyme alcohol oxidase (*AOX*), which is sequestered within the peroxisome along with catalase. The latter enzyme degrades hydrogen peroxide to oxygen and water. A third peroxisomal enzyme of the methanol metabolism, dihydroxyacetone synthase, catalyzes the condensation of formaldehyde with xylulose 5-monophosphate to glyceraldehyde 3-phosphate and dihydroxyacetone. These two products leave the peroxisome and enter a cytoplasmic pathway that regenerates xylulose 5-monophosphate and, for every three cycles, one net molecule of glyceraldehyde 3-phosphate. A portion of formaldehyde generated by *AOX* leaves the peroxisome and is further oxidized to formate and carbon dioxide by two cytoplasmic dehydrogenases. These reactions are source of energy for cells growing on methanol [2][9].

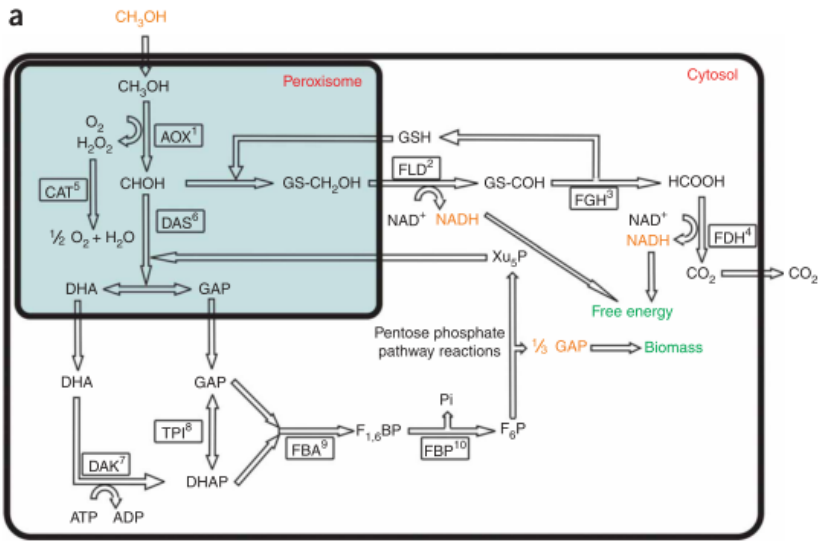


Figure 1: Methanol utilization pathway in *Pichia pastoris*: AOX: alcohol oxidase; FLD: formaldehyde dehydrogenase; FGH: S-formylglutathione hydrolase; FDH: formate dehydrogenase; CAT: catalase; DAS: dihydroxyacetone synthase; DAK: dihydroxyacetone kinase; TPI: triosephosphate isomerase; FBA: fructose-1,6-bisphosphate aldolase; FBP: fructose-1,6-bisphosphatase; DHA: dihydroxyacetone; GAP: glyceraldehyde 3-phosphate; DHAP: dihydroxyacetone phosphate; $F_{1,6}BP$: fructose-1,6-bisphosphate; F_6P: fructose-6-phosphate; P_i: phosphate; Xu_5P: xylulose-5-phosphate; GSH: glutathione (taken from [10]**).**

1.1.3 Commonly used promoters

The first enzyme of the methanol utilization pathway is alcohol oxidase, which catalyzes the oxidation of methanol to formaldehyde. In *P. pastoris* there are two genes that code for alcohol oxidase enzyme: alcohol oxidase 1 (*AOX1*) and alcohol oxidase 2 (*AOX2*) [6]. Grown on methanol as sole carbon source *AOX1* can represent up to 30% of total soluble protein in extracts of *P. pastoris*, what shows the enormous strength of the *AOX1* promoter (*pAOX1*). The second alcohol oxidase *AOX2* is under the control of a much weaker promoter (*pAOX2*) and therefore accounts for just 15% of the overall *AOX* activity in the cell [11]. There are three phenotypes in *P. pastoris* with regard to methanol utilization. The wild type or methanol utilization plus phenotype (Mut^+), and those resulting from deletions in the *AOX1* gene (methanol utilization slow Mut^S) or both *AOX* genes (methanol utilization minus Mut^-) [6][12].

Expression of *AOX1* is controlled at the level of transcription. In cells grown on methanol, ~5% of poly(A)$^+$ RNA is from *AOX1*, whereas in cells grown on most other carbon sources, *AOX1* message is undetectable. The *AOX1* gene seems to be regulated by two mechanisms: a repression/derepression mechanism plus an induction mechanism. Unlike *GAL1* regulation of *Saccharomyces cerevisiae*, the absence of a repressing carbon source such as glucose or glycerol in the medium does not result in substantial transcription of *AOX1*. To induce high levels of transcription the presence of methanol is essential [2]. The reasons of *pAOX1´*s predominant use for heterologous protein expression are, besides being

the first well-characterized promoter of this yeast, the strength and tight regulation by carbon sources [13].

There might be circumstances in which the *pAOX1* is not suitable. For example, methanol is a potential fire hazard and the use of methanol for gene expression may not be adequate for the production of food products. Therefore other promoters that are not induced by methanol like *GAP*, *FLD1*, *PEX8* and *YPT1* promoters can be used [2].

The glyceraldehyde 3-phosphate dehydrogenase promoter (*pGAP*) has been successfully used for constitutive expression on glucose for several heterologous proteins. Despite its comparable expression level to that observed with *pAOX1*, its use is not recommended, if the expressed protein is toxic to the cell [2][6]. Formaldehyde dehydrogenase (*FLD1*) is a key enzyme in the degradation of the formaldehyde received from the assimilation of methylated amines as nitrogen source. The *FLD1* promoter is independently inducible either by methanol or methylated amines, used as sole carbon or nitrogen source [14]. The strength of this promoter is similar to that of the *AOX1* promoter. In case *pAOX1* or *pGAP* driven expression is too strong, the weaker *PEX8* and *YPT1* promoters are sometimes used. *PEX8* gene encodes for a peroxisomal matrix protein and its promoter is induced by methanol or oleate. *YPT1* codes for a GTPase involved in secretion [15] and its promoter provides a low but constitutive level of expression in media containing glucose, methanol or mannitol as carbon sources [3].

1.1.4 Glycosylation in *Pichia pastoris*

P. pastoris is able to add both O- and N-linked carbohydrate moieties to secreted proteins. A major problem for the use of yeast-secreted glycoproteins as therapeutic products is the difference in the number and type of sugar units added by humans compared to *P. pastoris*. Introduced intravenously into mammals these glycoproteins can cause antigenic response and they are rapidly cleared from the bloodstream [16].

N-linked glycosylation starts in the endoplasmic reticulum with the transfer of an oligosaccharide unit, $Glc_3Man_9GlcNAc_2$ (Glc = glucose; GlcNAc = N-acetylglucosamine; Man = mannose), to asparagine at the recognition sequence Asn-X-Ser/Thr. Subsequently this oligosaccharide is trimmed to $Man_8GlcNAc_2$ and so far the glycosylation patterns of lower and higher eukaryotes are the same. In *Saccharomyces cerevisiae* N-linked core units are elongated in the Golgi through the addition of mannose outer chains, which have a typical length of 50-150 mannose residues. This condition is called hyperglycosylation. In mammals the Golgi forces a series of trimming and addition reactions leading to oligosaccharides composed of $Man_{5-6}GlcNAc_2$ (high-mannose type), a mixture of several different sugars (complex type), or a combination of both (hybrid type) [2]. In *P. pastoris* the outer

oligosaccharide chain of secreted proteins is mostly unaltered and consists of $Man_{8-9}GlcNAc_2$ and therefore lacks hyperglycosylation [1].

In yeasts, such as *P. pastoris*, O-linked saccharides are added to the hydroxyl groups of serine and threonine. These are composed solely of mannose, whereas higher eukaryotes have a more varied sugar composition in these oligosaccharides, including N-acetylgalactosamine, galactose and sialic acid [1][2].

1.1.5 Secretion in *Pichia pastoris*

P. pastoris has the ability to produce proteins both intracellularly and secreted into the culture medium. In most cases the latter option is preferred, because of the relatively low levels of endogenous secreted proteins of this yeast and the subsequent simplified recovery and purification steps. The presence of a signal sequence is required to target the protein to the secretory pathway [3]. Therefore foreign genes are cloned in *P. pastoris* vectors to align them in the correct reading frame with either the native secretion signal for the protein of interest, the *P. pastoris* acid phosphatase (*PHO1*) signal or the *Saccharomyces cerevisiae* α-factor prepro-peptide. However, the most widely and successfully used secretion signal is the *S. cerevisiae* α-factor prepro-signal (Figure 2). In many cases it is even better than the leader sequence of the native heterologous protein [1].

This signal sequence is composed of a 19-amino acid signal (pre) sequence followed by a 66-residue (pro) sequence. The pro sequence contains three consensus N-linked glycosylation sites and a dibasic Kex2 endopeptidase processing site. The first step in processing of this signal sequence is the removal of the pre signal by the signal peptidase in the ER followed by the Kex2 cleavage between Arg-Lys of the pro leader sequence. This is followed by trimming Glu-Ala repeats by the Ste13 protein. The cleavage efficiency of Kex2 as well as Ste13 can be influenced by amino acids adjacent to the α-factor prepro-peptide [2].

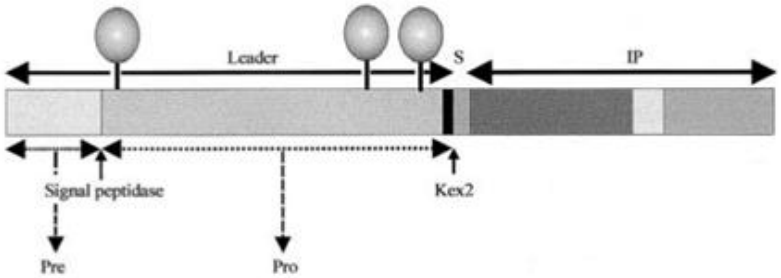


Figure 2: Schematic representation of the *S. cerevisiae* α-factor prepro signal (Leader) fused to an insulin precursor (IP) via a spacer peptide (S). Sites for attaching N-linked carbohydrate chains, for processing by the signal peptidase and for processing by the Kex2 endoprotease are indicated (taken from [17]).

1.1.6 Proteolysis

1.1.6.1 General features and functions of proteolytic enzymes

Proteolytic enzymes play an important role in the control and execution of many cellular events both intracellularly and extracellularly. Intracellular proteases are involved in following processes: 1) proteolytic processing of proproteins to generate completely active proteins; 2) removal of signal peptides after translocation of proteins through membranes; 3) inactivation of short-lived regulatory proteins; and 4) degradation of unwanted proteins derived from misfolding and mutations. Extracellular proteases are secreted to the medium and trim proteins for supplementation of peptides and amino acids for nutrition [18].

According to the MEROPS database proteases are classified on three criteria: 1) the chemical mechanism of catalysis; 2) the details of the reaction catalyzed and 3) the molecular structure and homology. The first type of classification depends on the respective chemical moiety that acts as a nucleophile at the catalytic centre. Thus proteolytic enzymes can be described as of serine (S), cysteine (C), threonine (T), aspartic (A), glutamic (G), asparagine (N), metallo catalytic (M), unknown (U) and mixed (M) type [19][20].

The second classification is based on the reaction proteases catalyze, which is for all of them the hydrolysis of a peptide bond. Endopeptidases hydrolyze internal, alpha-peptide bonds in a polypeptide chain, but they act away from the N- or C-terminus. In contrast exopeptidases require a free N-terminal amino group, C-terminal carboxyl group or both and are further divided into aminopeptidases, carboxypeptidases, dipeptidyl-peptidases, peptidyl-dipeptidases, tripeptidyl-peptidases and dipeptidases [19][20].

The classification of peptidases by molecular structure and homology is based on the availability of data for amino acid sequence and three-dimensional structures and is therefore the newest of the three methods. Each peptidase protein is assigned to a peptidase species. They are further grouped into families and the families into clans. Each protein has an ID of the form *S.01.001* that acts like an accession number [19][21].

Most of the protein degradation in the cell is executed by the proteasome and by the vacuolar pathway. Degradation by the ATP-dependent proteasome pathway occurs when proteins that are assigned for degradation are marked with the 8 kDa protein ubiquitin. In the vacuolar pathway proteins are transported to the vacuole and degraded by vacuolar proteases [18][22].

1.1.6.2 Proteolytic degradation of secreted recombinant proteins

Generally, secreted recombinant proteins can be degraded in the culture medium by extracellular proteases, cell-bound proteases and/or by intracellular proteases from lysed

cells. The extracellular proteases of *P. pastoris* are not well investigated and this yeast exhibits only low levels of secreted endogenous proteins [18][23].

Proteolysis can cause several problems in the production of recombinant proteins: 1) decreased product yield in case the protein is degraded; 2) absence of biological activity in case the product is truncated; and 3) contamination of the product by degradation intermediates in downstream processing due to their similar physicochemical and/or affinity characteristics [1].

1.1.6.3 Strategies to control proteolysis in *Pichia pastoris*

P. pastoris is capable of growing to very high cell densities up to 130 g l^{-1} dry cell weight [2]. Under these conditions *P. pastoris* cultures contain significant proteolytic enzyme activities that can degrade recombinant proteins. Therefore, the yields of recombinant proteins can be increased by reducing the degradation rate. Several strategies are used to overcome proteolytic degradation [18].

Cultivation-level strategies

P. pastoris is able of growing across a relatively broad pH range (3.0 – 7.0) without showing any significant effects on growth. Thus, this range allows considerable freedom in adjusting the pH to one that is not optimal for a problem protease providing that the protease that is degrading the protein is known [1].

The product stability can further be enhanced by addition of amino acid-rich supplements, such as peptone and casamino acids, to the culture medium. These supplements possibly act as alternative and competing substrates for one or more problem proteases and can also repress protease induction caused by nitrogen limitation [1][24].

Lower cultivation temperature and different methanol concentration can also influence the yields of recombinant proteins. Laccase activity could be increased by lowering the process temperature from 30°C to 20°C and by changing the methanol concentration from 1.0% to 0.5%. The mechanism behind the temperature effect may be due to poor stability, release of more proteases from dead cells, and folding problems at higher temperature [25].

Adding specific protease inhibitors to the culture medium may also be an option, but on an industrial scale, use of protease inhibitors could be too cost-intensive [1].

Protein-level strategies

A linker between the domains of a fusion protein could be particularly sensitive to degradation, if it contains an amino acid sequence recognized by native proteases. The respective amino acid sequence can be deleted if it is not essential for the function of the

protein [1]. Gustavsson et al. were able to design stable linker peptides for a cellulose-binding domain lipase fusion protein that decreased proteolysis [26].

Cell-level strategies

The use of several protease-deficient *P. pastoris* strains has been demonstrated to reduce the degradation of some foreign proteins. This is particularly distinguishable in fermenter cultures, because high cell density combined with lysis of a small percentage of cells results in a relatively high concentration of vacuolar proteases. Therefore, host strains with a disruption in these genes are available: SMD1163 (*his4 pep4 prb1*), SMD1165 *(his4 prb1*) and SMD1168 (*his4 pep4*) [2].

The *PEP4* gene codes for proteinase A, a vacuolar aspartic protease responsible for the activation of vacuolar proteases, such as carboxypeptidase Y (*PRC1*) and proteinase B (*PRB1*). Before its processing and activation by proteinase A, proteinase B has only about half the activity of the processed enzyme. Thus, *pep4* mutants exhibit a substantial decreased or even eliminated activity of proteinase A and carboxypeptidase Y, and a partial reduction of proteinase B activity. Whereas in *prb1* mutants only proteinase B activity is eliminated, *pep4 prb1* mutants show a significant reduction or elimination of all three of these protease activities [18].

Several more *P. pastoris* strains, which have a disruption in one or more protease genes, have been developed. Disruption of the *KEX1* gene, encoding a carboxypeptidase that can cleave carboxy-terminal lysines and arginines, has shown to be successful in expressing full-length murine and human endostatin [27] as well as hirudin [28]. Werten et al. were able to reduce proteolysis of secreted gelatin in a *P. pastoris kex2* disruptant, whereupon they reported an increase of generation time of the knockout strain by ~50% relative to the wild type [29].

The endoprotease Yps1 is a glycophosphatidylinositol (GPI)-anchored protein that is localized to the cell membrane and partially to other different subcellular locations such as Golgi apparatus and vacuole. Disruption of this gene reduced degradation of HSA-AX15(R13K) [30] and, combined with a *pep4* deletion, it decreased proteolysis of HSA/PTH as well [31]. Genome sequence annotation of *P. pastoris* revealed six more putative GPI-anchored aspartic proteases that belong to the yapsin family: *YPS2*, *YPS3*, *YPS7*, *MKC7*, *YPS´* and *YPS´´*. All of them have recently been knocked out, but only the *yps1* disruptant showed a decreasing effect on degradation of recombinant protein [31]. Unexpectedly *yps1*, *yps2*, *yps3* and *mkc7* knockout strains grew normally in the presence of the cell wall disturbing reagents calcofluor white and congo red compared to the wild type and *yps7* disruptant even showed increased resistance to these chemicals. There was not observed any negative change in cell wall integrity as it has been reported in *S. cerevisiae* [32].

Recently a secreted endogenous protease has been identified that belongs to serine type proteases and is encoded by the *SUB2* gene [33].

However, a total of 260 ORFs of *P. pastoris* were predicted to have secretion signals [4]. Thus, these proteins are somehow involved in the secretory pathway and could act as proteases. It is obvious that only a small percentage of putative proteases have been studied and that there is a need to put more effort into identifying and characterizing new proteolytic enzymes to overcome the major problem of degradation of heterologous expressed proteins.

1.2 α2, 6-Sialyltransferase ST6Gal I (Δ27ST6)

Protein glycosylation is a complex and frequent post-translational modification in eukaryotes. For many therapeutic proteins the attachment of the proper glycan structure is important for their folding, biological activity and pharmacological efficacy. Eukaryotic expression systems, such as yeast, insect and non-human mammalian cell lines are available that can perform protein glycosylation, but they insert non-native glycosylation patterns. As a result decreased biological potency and immunogenicity are observed [34].

Sialic acids exist in a variety of mammalian glycoproteins and glycolipids [35]. The presence of sialic acid moieties attached to the terminal site of its glycoprotein constitutes one of the most important examples of the influence of glycan structures on the biological activity of therapeutic proteins [36]. It has been shown that the plasma half-life of recombinant human erythropoietin (EPO) was reduced from more than 5 hours to 2 min, when terminal sialic acid moieties have been removed [37].

One promising possibility to produce fully sialylated proteins is to isolate the desired recombinant protein and attach the sialic acid moiety *in vitro* by using purified sialyltransferases and activated sialic acid donors. Therefore high yields of sialyltransferases are needed, but studies have shown that the yields of soluble and active mammalian sialyltransferases obtained from expression in bacteria or lower eukaryotes are very low [34].

The enzyme used in this work to express in protease-deficient *P. pastoris* strains is a truncated version of the human α2, 6-sialyltransferase ST6Gal I (EC 2.4.99.1) that sialylates O-linked glycoproteins [34]. The complete native protein consists of 406 aa with a molecular weight of ~46 kDa, including a typical hydrophobic signal anchor sequence (aa 11 to 25) and two potential N-linked glycosylation sites [38].

The first 27 amino acid residues, containing the anchor sequence, have been removed from the respective sialyltransferase. At the N-terminal end a FLAG tag and a secretion signal (*S. cerevisiae* α-factor prepro peptide or levanase secretion signal) have been fused to enable immunodetection as well as secretion. The protein was provided by Doris Ribitsch (ACIB GmbH, Petersgasse 14, Graz) and is further called Δ27ST6.

2 OBJECTIVES

The ultimate goal of this Master´s Thesis project was the development of protease-deficient *Pichia pastoris* host strains useful for effective production of protease-sensitive heterologous proteins.

Although the *Pichia* system has been successfully used to produce a variety of different heterologous proteins from bacterial (*Escherichia coli* phytase [39]), fungal (*Aspergillus oryzae* tannase [40]), viral (measules virus nucleoprotein MeN9 [41]) and human (anti-HBs Fab fragment [42]) origins, degradation of heterologous expressed proteins in *P. pastoris* is still a major problem. However, only few *P. pastoris* proteases have been documented and partially studied, but a total of 260 ORFs were predicted to have secretion signals [4].

Because of the presence of such a large number of putative proteases in the yeast *P. pastoris*, only 35 were selected as gene disruption targets by my supervisor Univ.-Prof. Dipl.-Ing. Dr. techn. Helmut Schwab. In Table 1 each protease is included with its gene name, gene ID, gene product, protease family and if it has a signal peptide (SP). Corresponding *S. cerevisiae* homologues for each protease are also noted with their values for identities and positives. 12 proteases have already been named and studied partially, whereas for the remaining ones the arbitrary abbreviation *kpx* was used to name them.

In a first step gene-specific knockout cassettes were generated for all 35 proteases. These carry a Zeocin^{TM} resistance gene and DNA sequences homologous to the flanking regions of the selected protease genes (5´UTR and 3´UTR). *P. pastoris* CBS7435 strain was transformed with the linearized knockout cassette, where replacement of the target ORF occurred by homologous recombination. After PCR-based screening for site-specific integration, the knockout cassette was removed by a flippase recombinase system. The removal and subsequent deletion of the full-ORF sequence was verified by PCR and sequencing.

At this point I have to mention that the data presented in this work arose from collaboration of my supervisor Mudassar Ahmad and me.

Table 1: *Pichia pastoris* (putative) proteases selected for gene deletion.

#	Gene	Gene ID CBS 7435	Gene ID GS 115	Gene product	Protease family	SP	Homologues		
							S. cerevisiae	Identities %	Positives %
1	*sub2*	299150304	254567147	Subtilisin 2 protease	S8	YES	*PRB1*	47	63
2	*pep4*	328352648	254572446	Proteinase A	A1	YES	*PEP4*	69	82
3	*prb1*	328350290	254565530	Proteinase B	S8	YES	*PRB1*	57	70
4	*prc1*	328350528	254566035	Carboxypeptidase Y	S10	YES	*PRC1*	60	72
5	*yps1*	38146742	254573826	Yapsin 1	A1	NO	*YPS1*	48	63
6	*yps2*	328353469	254570819	Yapsin	A1	YES	*YPS1*	30	49
7	*yps3*	328353465	254570827	Aspartic protease	A1	YES	*YPS1*	32	45
8	*yps7*	328353374	254571021	Cathepsin D	A1	YES	*YPS7*	28	50
9	*kex1*	3747109	238030858	Carboxypeptidase Kex1	S10	YES	*KEX1*	43	59
10	*kex2*	38146738	254568178	Kex2 proprotein convertase	S8	YES	*KEX2*	44	61
11	*prtP*	328352048	254567898	PI-type proteinase		YES	*FLO10*	27	39
12	*ctse*	328350442	254565855	Yapsin	A1	NO	*MKC7 (YPS2)*	29	44
13	*kpx1*	328354397	254573342	Immunglobulin A1 protease		YES	*ENT2*	44	51
14	*kpx2*	328351136	254567279	Aminopeptidase Y	M28	YES	*APE3*	49	63
15	*kpx4*	328352885	254571979	Aspartic protease	A1	YES	*MKC7 (YPS2)*	27	45
16	*kpx5*	328351070	254567147	Proteinase B	S8	YES	*PRB1*	47	63
17	*kpx6*	328350760	254566511	Proteinase B	S8	YES	YCR045C	46	61
18	*kpx7*	328354158	254573826	Aspartic protease	A1	YES	*YPS1*	48	63
19	*kpx8*	328353613	254570525	Aspartic protease	A1	YES	*YPS1*	29	49
20	*kpx9*	328350256	254565461	Aminopeptidase Y	M38	YES	*YDR415C*	45	63
21	*kpx10*	328352415	254567898	C5a peptidase		YES	*FLO10*	27	42
22	*kpx11*	328351136	254567279	Aminopeptidase Y	M28	YES	*APE3*	49	63
23	*kpx12*	328354499	254573130	Endoprotease	C13	YES	*GPI8*	78	89
24	*kpx13*	328353128	254571503	Carboxypeptidase Y	S10	YES	*YBR139W*	58	73
25	*kpx14*	328352413	254567902	Carboxypeptidase Y	S10	YES	*KEX1*	43	59
26	*kpx15*	328352277	254568178	Kex2 proprotein convertase	S8	YES	*KEX2*	44	61
27	*kpx17*	328353349	254571075	Metalloprotease	M14	YES	*ECM14*	54	74
28	*kpx18*	328352415	254567898	C5a peptidase		YES	*FLO10*	27	42
29	*kpx19*	328352413	254567902	Carboxypeptidase Y	S10	YES	*KEX1*	43	59
30	*kpx20*	328353755	254570259	PI-type proteinase		YES	*CWP1*	24	40
31	*kpx21*	328353754	254570261	C5a peptidase		YES	*FLO10*	31	42
32	*kpx22*	328353613	254570525	Aspartic protease	A1	YES	*YPS1*	29	48
33	*kpx23*	328353349	254571075	Metalloprotease	M14	YES	*ECM14*	54	74
34	*kpx24*	328350485	254565945	Aspartic protease	A	NO	*SRT1*	38	63
35	*kpx25*	328352792	254572167	Metalloprotease	M	NO	*YBR074W*	39	57

To show an application for the generated protease-deficient *P.* pastoris strains, above mentioned sialyltransferase Δ27ST6 was chosen to be expressed in those strains. Pretests were performed by a different work group from ACIB GmbH under supervision of Doris Ribitsch. When they expressed the sialyltransferase Δ27ST6 in a *P. pastoris* KM71H strain, high-quantity degradation of the secreted Δ27ST6 was observed especially after 120 h of fermentation, as shown in Figure 3. Adding of a protease inhibitor tablet, which generally eliminates most proteolytic activity, seemed to reduce proteolysis of the respective protein significantly.

Due to very high costs of protease inhibitors and therefore use of them for large scale fermentation is not a well suited option, our generated protease-deficient strains were destined for expression of Δ27ST6. The protease-deficient strains should be analyzed,

whether the deletion of certain proteases reduces proteolysis and leads to higher amounts of entire and active enzyme.

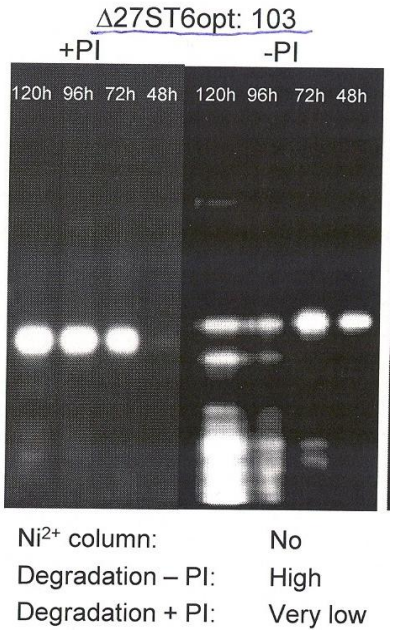


Figure 3: Expression of Δ27ST6 in a *P. pastoris* KM71H strain with and without protease inhibitors (PI). In contrast to high-quantity degradation of the protein especially after 120h of fermentation time without PI, proteolysis seems to be reduced significantly by adding protease inhibitors. Immunoblot detection was performed using a primary antibody against the protein (anti-hST6 antibody). (Source: Doris Ribitsch, ACIB Gmbh, Graz)

3 MATERIALS AND METHODS

3.1 Strains, plasmids, primers

The *E. coli* TOP10 F´ strain was used for cloning purposes and construction of knockout-plasmids. Knockouts of target genes were performed in the *P. pastoris* CBS7435 strain and *P. pastoris* CBS7435 *Δhis4* strain.

Table 2: Basic strains used

Strain	Genotype	Source
E. coli TOP10 F´	F´{lacIq Tn10 (TetR)} mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL endA1 nupG	Invitrogen
P. pastoris CBS7435	Wild type	
P. pastoris CBS7435 *Δhis4*	*Δhis4* (Mut+ phenotype)	Ahmad M.

Table 3: Plasmids used

#	Plasmid	Features	Source
1	pPpKC1	Flippase gene under *AOX1* promoter, 2 FRT regions, pUC ori, $Zeocin^R$	Ahmad M.
2	pPpKC1HIS4	Flippase gene under *AOX1* promoter, 2 FRT regions, pUC ori, Amp^R, *HIS4* marker	Ahmad M.
3	pAαZSwaIFlagD27ST6	Plasmid for expression of Δ27ST6, α-factor prepro peptide, Amp^R, $Zeocin^R$, integrates at *AOX1* locus	Ahmad M.
4	pAZSwaILevSSFlagD27ST6	Plasmid for expression of Δ27ST6, levanase secretion signal, Amp^R, $Zeocin^R$, integrates at *AOX1* locus	Ahmad M.
5	pAαHSwaIFLAGD27ST6	Plasmid for expression of Δ27ST6, α-factor prepro peptide, Amp^R, *HIS4* marker, integrates at *AOX1* locus	Ahmad M.
6	pAHSwaILevSSFLAGD27ST6	Plasmid for expression of Δ27ST6, levanase secretion signal, Amp^R, *HIS4* marker, integrates at *AOX1* locus	Ahmad M.

Table 4: Primers used

Gene Name	#	Primer Name	Nucleotide Sequence (5´ → 3´)
sub2	1	3UTRsub2F	TCGGCCGATCAGGCCTGACTCATTGACCCCAGCTCAAC
	2	3UTRsub2R	GGGACTGACCCGGGTGAGGAAAACACTCATTGAAATTCCTG
	3	5UTRsub2F	CCTCACCCGGGTCAGTCCCAACTTGTTGG
	4	5UTRsub2R	TCGGCCCTAGTGGCCGATCCCTGTAATTTCAGCGATGGAG
	5	Up5UTRsub2F	ACGATTAAGGCAAATCTTCCGGTTC
	6	Down3UTRsub2R	GAAACAAATCAGTGACGGCGATGTC
kpx1	7	3UTRkpx1F	TCGGCCGATCAGGCCTCTGTAGGAGATGCCAATGTCATTG
	8	3UTRkpx1R	GCTGATCTTATCCCGGGAAGTAATTTTTTAAACATATTGATAAACAGAC
	9	5UTRkpx1F	AAATTACTTCCCGGGATAAGATCAGCAGGTATGAATG
	10	5UTRkpx1R	TCGGCCCTAGTGGCCCTACACCAAAGCCAGGTTGCCAAAC
	11	Up5UTRkpx1F	GTGACCCTATCTGGAAAGTCGAGAC
	12	Down3UTRkpx1R	CTTCACCAGGTCCAAAGTTGATGG
kpx2	13	3UTRkpx2F	CGGCCGATCAGGCCCACAGTGTCGCCACTTATGCAAAG
	14	3UTRkpx2R	CTTTCTCTACCCGGGAAGCACAGCAAAGTATCTATGC
	15	5UTRkpx2F	GTGCTTCCCGGGTAGAGAAAGTTCATCTTCGTTCG
	16	5UTRkpx2R	TCGGCCCTAGTGGCCGTAGCAACGAGTGGCAAATATTTCATAGC
	17	Up5UTRkpx2F	GCAAGACGGTTTAACTTGGCACTG
	18	Down3UTRkpx2R	TTCAGTGGAAGGTGGACCTTCCACAG
yps3	19	3UTRyps3F	TCGGCCGATCAGGCCTAGCGTTGGATTAGAGTTCAGAAATACCAC
	20	3UTRyps3R	GGTTTTACCCGGGTGTCGCTTAATAACTGTTGTGTGATG
	21	5UTRyps3F	GCGACACCCGGGTAAAACCTCGAAGATGACTTCATTC
	22	5UTRyps3R	TCGGCCCTAGTGGCCTCAACTTTAAGGAGAGCAAAAGCAGAAG
	23	Up5UTRyps3F	CTCTGTTCCCGCAGGTAGCTATG
	24	Down3UTRyps3R	TAAGGACCACATAGTTTAACCCCTG
kpx4	25	3UTRkpx4F	TCGGCCGATCAGGCCCAACTACGACCCAAGCATATCAGATG
	26	3UTRkpx4R	GAATTGCTACCCGGGCACCTGGATTGAATGCAAG
	27	5UTRkpx4F	AGGTGCCCGGGTAGCAATTCCGAGTTAACATAACTG
	28	5UTRkpx4R	TCGGCCCTAGTGGCCCAATGGCGACGATTCCAGCATAG
	29	Up5UTRkpx4F	ATGGAGTGGCCCGTGATTGAAATATTG
	30	Down3UTRkpx4R	GGTTCTTCCAGTATTAAACCTAACTTGACAGG
pep4	31	3UTRapr1F	TCGGCCGATCAGGCCCTCAGTTTATGACCTAGGCAAAGATGC
	32	3UTRapr1R	GATAAAGGTCCCCGGGACCTCGGTTGTAAGCGGTAATTC
	33	5UTRapr1F	CCGAGGTCCCGGGGACCTTTATCACGTTGAATCTAGTTG
	34	5UTRapr1R	TCGGCCCTAGTGGCCGCTTGTGTATCTTAGCAGAATGAACTTTGG
	35	Up5UTRapr1F	GAAAATAGTGTATCACTGCCAGCATC
	36	Down3UTRapr1R	CTCATCTATACCCCAGGACCAG
kpx5	37	3UTRkpx5F	TCGGCCGATCAGGCCCATTGACCCCAGCTCAACTAAAAGC
	38	3UTRkpx5R	GAGAAGGGCCCGGGTTTCAGAAAGCTACGGATCTAG
	39	5UTRkpx5F	CTGAAACCCGGGCCCTTCTCTGCAAACACGTTTTG
	40	5UTRkpx5R	TCGGCCCTAGTGGCCGCAATTGGATATTCAATTGGTCAAGAGATGG
	41	Up5UTRkpx5F	ACACAGGCAGTGCTATATTTGGC
	42	Down3UTRkpx5R	CAGTAAACGACAAGCTACTATCATGGAGC
kpx6	43	3UTRkpx6F	TCGGCCGATCAGGCCCTAGAAGACATTGAAGAGGACGAGG
	44	3UTRkpx6R	CTACAACAACCCGGGTATAATTTCAAATTATGGGCCCAAC

	45	5UTRkpx6F	GAAAT ATACCCGGGTTGTTGTAGACCCTACCATCAATG
	46	5UTRkpx6R	TCGGCCCTAGTGGCCTGTGAGGCTTCTTCAGACGAACAAAATATTTTG
	47	Up5UTRkpx6F	CATTGTGGATGCTTTAGTCGACTTCAAGC
	48	Down3UTRkpx6R	CTCCGTTGATATTTCCAAGATTGACAGTG
prb1	49	3UTRprb2F	TCGGCCGATCAGGCCCACTGTCACCATTAGCACCAAACTG
	50	3UTRprb2R	GCCTCTAATCCCGGGAAAGTTTAACTTCATACAGAATAACTTCATG
	51	5UTRprb2F	AAACTTTCCCGGGATTAGAGGCGGTTGAACTCTG
	52	5UTRprb2R	TCGGCCCTAGTGGCCGTTGCTTCCTCCGACGATACTG
	53	Up5UTRprb2F	GCAGTATCCTGCTCATCTTCCCGTAC
	54	Down3UTRprb2R	CATGAACGTGTTGAACTTGGACGCC
kpx7	55	3UTRkpx7F	TCGGCCGATCAGGCCGGTACGCTCGCTATCAAAGCAAC
	56	3UTRkpx7R	CCAGACGCCCGGGTTTTCGTTTATCGGTATTACCGGAG
	57	5UTRkpx7F	CGAAAACCCGGGCGTCTGGTTGTTTGTATTAGCAC
	58	5UTRkpx7R	TCGGCCCTAGTGGCCGGTCCCCTTCAGCTACCTTTCTC
	59	Up5UTRkpx7F	CGTCGGCATTATCTGGTAGATCCG
	60	Down3UTRkpx7R	GTATCAAATTGACCATTATCCCTTTTCACC
kpx8	61	3UTRkpx8F	TCGGCCGATCAGGCCCTTCAGGTGCATCTTCTGCTACTCAAAATG
	62	3UTRkpx8R	CAAATATGTCCCGGGCATAAGTATCAATGTACTTCTCAATG
	63	5UTRkpx8F	GATACTTATGCCCGGGACATATTTGCCCTCTGATTGCAC
	64	5UTRkpx8R	TCGGCCCTAGTGGCCGCAACTTGTTAGCCTTGAAAGGCGATTG
	65	Up5UTRkpx8F	GCCTCTGACAGAGCGTTGACCTG
	66	Down3UTRkpx8R	CAAAGATCTTGGTGGCTTCGTCC
yps2	67	3UTRyps2F	TCGGCCGATCAGGCCCTGAGTGCAAGTAGAATTAAGCTGCTAG
	68	3UTRyps2R	CCAATAACCCGGGTCTGAACATCCTGATTGAAAGC
	69	5UTRyps2F	GTTCAGACCCGGGTTATTGGTGATCAAGGTTCCTTC
	70	5UTRyps2R	TCGGCCCTAGTGGCCGTCGGTGTCTCACATTAACACTAGTTC
	71	Up5UTRyps2F	AATTCATGATTCCGGAGTGCGTGTAATC
	72	Down3UTRyps2R	TTGATTGGCGTAGCTGGTGATGAC
yps7	73	3UTRyps7F	TCGGCCGATCAGGCCAATTAGCCGACGAAAAGCATATCAGAGAC
	74	3UTRyps7R	ATTCATGCCCGGGTTGAGAGGCATATCGAGAAG
	75	5UTRyps7F	CTCTCAACCCGGGCATGAATTTATTGGTGATTGCTTAAAG
	76	5UTRyps7R	TCGGCCCTAGTGGCCTCTTTACATCATTGGCCAGTCTGTTGAC
	77	Up5UTRyps7F	CAATTGACACCAAAGGACAGTTTAGACTC
	78	Down3UTRyps7R	GTGTGAAATGCGCTGATCGAACTG
kpx9	79	3UTRkpx9F	TCGGCCGATCAGGCCGGAGATTCCAAGGTACAATTGTCGC
	80	3UTRkpx9R	GATGGTGCCCGGGAGATTACGCGTCGAATCG
	81	5UTRkpx9F	GTAATCTCCCGGGCACCATCGTCCATAATGTC
	82	5UTRkpx9R	TCGGCCCTAGTGGCCTGAAAGTTTTGCGGTTAATCACGGTTTC
	83	Up5UTRkpx9F	TCGTATCGATCCAGCTATTTTGAACG
	84	Down3UTRkpx9R	GTCTTCCCCACTCGTATGACTTTAGG
kpx10	85	3UTRkpx10F	TCGGCCGATCAGGCCGTTATCGTTGCTGATTTGAACAGATGCTC
	86	3UTRkpx10R	GAAAAGACCCGGGCTTCGTAGGAGAGG
	87	5UTRkpx10F	TACGAAGCCCGGGTCTTTTCGACGTGGTTGATAAAG
	88	5UTRkpx10R	TCGGCCCTAGTGGCCGAATTCAAGCTGCTTCAAACAGCAC
	89	Up5UTRkpx10F	GTCTTGTTCCAGTGAACAACCAGC
	90	Down3UTRkpx10R	TGGGTAGGGTTGTATGGGAACG
kpx11	91	3UTRkpx11F	TCGGCCGATCAGGCCCAGTCAAACTTCAAGTACCACGGTC

	92	3UTRkpx11R	TCCTGCCCGGGAAGCACAGCAAAGTATCTATGC
	93	5UTRkpx11F	GCTTCCCGGGCAGGAGCATTTTTTTGTTTCAAG
	94	5UTRkpx11R	TCGGCCCTAGTGGCCAGGGTAGCAACGAGTGGCAAATATTTC
	95	Up5UTRkpx11F	CACTGATTGGAACTTGAGCCTCC
	96	Down3UTRkpx11R	TTCATTGAGCTGTTCATTGAGCTGTTC
prtP	97	3UTRprtPF	TCGGCCGATCAGGCCAAAGCAGATGCCAATCCTACCAG
	98	3UTRprtPR	TTAAAGACCCGGGCTGTTTTCACACTTGAGTCAG
	99	5UTRprtPF	AAAACAGCCCGGGTCTTTAAACTGTCCAATGGAAACC
	100	5UTRprtpR	TCGGCCCTAGTGGCCCGCTGCTGTGATTTTCTCAATTCTTCC
	101	Up5UTRprtPF	TATTCCTCGGAACAATCCTCTGTAAC
	102	Down3UTRprtPR	TCCTGACTACTACTCCTAACAGTAGAAAGG
kpx12	103	3UTRkpx12F	TCGGCCGATCAGGCCCTACTCTTGGGTTCACTTACACGC
	104	3UTRkpx12R	TAAACAAATCCCGGGAGTCCCTACTTCCGACAATATTG
	105	5UTRkpx12F	GTAGGGACTCCCGGGATTTGTTTAGCCATCCATCGAG
	106	5UTRkpx12R	TCGGCCCTAGTGGCCACGAGGCTGCAACACAACAAAGTATAC
	107	Up5UTRkpx12F	TGACTGATACTTACTTGAGTCTGCAAGC
	108	Down3UTRkpx12R	GTAGAAATTTCCGAAGAGCCTGCC
kpx13	109	3UTRkpx13F	TCGGCCGATCAGGCCATGGTTGTCAGATGGACACACG
	110	3UTRkpx13R	GACCTACTCCCGGGTATTCCTGGATACTTTGAAACCG
	111	5UTRkpx13F	CAGGAATACCCGGGAGTAGGTCTTCCCTGATACTC
	112	5UTRkpx13R	TCGGCCCTAGTGGCCAATGCCTGCGACAATGCTAGCAAAG
	113	Up5UTRkpx13F	GAGACAGTGGTTCAACTTATGGGC
	114	Down3UTRkpx13R	CATAGGTTTTCAGGTCAGGAAAGAAGG
kpx14	115	3UTRkpx14F	TCGGCCGATCAGGCCCTGACACCGCTACTAAAGGAAGC
	116	3UTRkpx14R	CCGACCCGGGTTGTCATTATTGGTGGTAAGGC
	117	5UTRkpx14F	TGACAACCCGGGTCGGCAGATGAGTCTTTG
	118	5UTRkpx14R	TCGGCCCTAGTGGCCACCTGGAATCCAACCTGGATCTTG
	119	Up5UTRkpx14F	TCATGGTAAAACCGTCGTCATTGC
	120	Down3UTRkpx14R	GGGTTCTCGGTCAATCCAAACG
kpx15	121	3UTRkpx15F	TCGGCCGATCAGGCCAAACGTGCGGCATTGTAAGAAAG
	122	3UTRkpx15R	GACTATCCCGGGCAATTTTTCTTGAAATATAGTTATGGAAGTG
	123	5UTRkpx15F	GAAAAATTGCCCGGGATAGTCGCCTTCCGGTTTCTG
	124	5UTRkpx15R	TCGGCCCTAGTGGCCCTCTCGGAAGTCTCCAAAGCC
	125	Up5UTRkpx15F	CCCATCCAGGACAAGTGCTG
	126	Down3UTRkpx15R	CGGACACATTCTCAAAGAGCAGC
prc1	127	3UTRprc1F	TCGGCCGATCAGGCCGGCATCTGCAAGGACAGACC
	128	3UTRprc1R	CACCTATCCCGGGAAAAGGCACATAAAGCAATCAATC
	129	5UTRprc1F	GCCTTTTCCCGGGATAGGTGATCCCTCAAAGAAGG
	130	5UTRprc1R	TCGGCCCTAGTGGCCGGCCCCATATGATCAGCCAG
	131	Up5UTRprc1F	CAAGTTCAAATGGCTCCATGGAGC
	132	Down3UTRprc1R	GCATTGAGGAAGTACATGGTCACG
kpx17	133	3UTRkpx17F	TCGGCCGATCAGGCCGGTCGTTGTTTCGCGCTC
	134	3UTRkpx17R	TACTGCCCGGGTGGCTGGTGGTAGTG
	135	5UTRkpx17F	GCCACCCGGGCAGTAAGGAGTCATCAAGAG
	136	5UTRkpx17R	TCGGCCCTAGTGGCCGTGTGCTATTGACAAGGTGGTCTTATAG
	137	Up5UTRkpx17F	GCTTATTCTCAGCACTGGATACACCTC
	138	Down3UTRkpx17R	GATTCGCCAAGATCCTGGCC

kpx18	139	3UTRkpx18F	TCGGCCGATCAGGCCCAGCACGAGCATCTCAGCTC
	140	3UTRkpx18R	GAACACTCCCGGGTCTTTTCGACGTGGTTGATAAAG
	141	5UTRkpx18F	GAAAAGACCCGGGAGTGTTCTCGTCCAGTC
	142	5UTRkpx18R	TCGGCCCTAGTGGCCCCAGATCCCAGGCCTTGC
	143	Up5UTRkpx18F	GCTGCAATCAGCTCCTCCAC
	144	Down3UTRkpx18R	CAGTGAACAACCAGCTTGTCCC
kpx19	145	3UTRkpx19F	TCGGCCGATCAGGCCTGTTACCACTACTCAAGCCACACAAG
	146	3UTRkpx19R	GACGTTCCCGGGTCTTTTCGACGTGGTTGATAAAG
	147	5UTRkpx19F	GAAAAGACCCGGGAACGTCTAGTGATGTGCTAC
	148	5UTRkpx19R	TCGGCCCTAGTGGCCCACGCAAACTTTGCAAATCCTGGAAAG
	149	Up5UTRkpx19F	GATTCCAGAAGGTCAACGTTGTTTCTAC
	150	Down3UTRkpx19R	AACAACCAGCTTGTCCCTGATTGAG
kpx20	151	3UTRkpx20F	TCGGCCGATCAGGCCCGGCAGCCTGTAGGATAATAAGAA G
	152	3UTRkpx20R	AATGCTCCCGGGAGTTTTCACGTGTTCTACGG
	153	5UTRkpx20F	TGAAAACTCCCGGGAGCATTCCAAAACATTGCG
	154	5UTRkpx20R	TCGGCCCTAGTGGCCCTTACACTGTGAGTGAGGCTCAAG
	155	Up5UTRkpx20F	CAGCACTATTTTGCTTGGCTTGC
	156	Down3UTRkpx20R	GTGTATTGATGCGCAGATACTCAGC
kpx21	157	3UTRkpx21F	TCGGCCGATCAGGCCTAAGGCTTGTTCGTAAGAAGGCAAAG
	158	3UTRkpx21R	TAACTGCCCGGGCAACATGGAAACACGTG
	159	5UTRkpx21F	ATGTTGCCCGGGCAGTTAGTCCTGTACTTGG
	160	5UTRkpx21R	TCGGCCCTAGTGGCCAGGACACTACGATAACAACTCATGTTAC
	161	Up5UTRkpx21F	TGCAGTATGTTACGCACTGCATTATATCG
	162	Down3UTRkpx21R	GATAAAATTCCACGGCTGACCGG
kpx22	163	3UTRkpx22F	TCGGCCGATCAGGCCGCGCACTTGCTAGTGCAATG
	164	3UTRkpx22R	TCTTTTGATCCCGGGCACACGGAAGCTG
	165	5UTRkpx22F	GTGTGCCCGGGATCAAAAGATCACCTTTTTCCAGTG
	166	5UTRkpx22R	TCGGCCCTAGTGGCCGCCTACTCCTCTACTTGGTCTAGC
	167	Up5UTRkpx22F	GTTCGGCTAAATCTTTAGGATTAGCACC
	168	Down3UTRkpx22R	CCGAAGAGTTGATGAGACCGGGAC
kpx23	169	3UTRkpx23F	TCGGCCGATCAGGCCGTGTGTGCTATTGACAAGGTGGTC
	170	3UTRkpx23R	CTTTAGCATCCCGGGCAGTAAGGAGTCATCAAGAG
	171	5UTRkpx23F	CTTACTGCCCGGGATGCTAAAGTTGGCAACTGAG
	172	5UTRkpx23R	TCGGCCCTAGTGGCCGGAACAGAGGTTTATGCCTCAATTCAG
	173	Up5UTRkpx23F	AGTCTTCTCTTTGAAGACAACGGAGATTC
	174	Down3UTRkpx23R	ATTGCTTTTGCAAACCCGTCAG
kpx24	175	3UTRkpx24F	TCGGCCGATCAGGCCTTGTGGATCTACTGAAGGACTTGTGAG
	176	3UTRkpx24R	GACTAGGACCCGGGTCATAATCGCATCTACTGTTGTG
	177	5UTRkpx24F	GATTATGACCCGGGTCCTAGTCTATGGACTAACAGC
	178	5UTRkpx24R	TCGGCCCTAGTGGCCCTGGCAGATGATGTGGATTCTGC
	179	Up5UTRkpx24F	CTTTGAAAGTGCACTACAACTGGTCC
	180	Down3UTRkpx24R	TTC GTA TCG TCC ACG ATA GGT AAA TGT C
kpx25	181	3UTRkpx25F	TCGGCCGATCAGGCCATTCCTGAAGGACTGCAAGTCTG
	182	3UTRkpx25R	CTTAGATCCCGGGAAGTTTAGAGGAAATCTGTCTCAAATAAG
	183	5UTRkpx25F	TCTAAACTTCCCGGGATCTAAGTTTGGAGATTCTGACC
	184	5UTRkpx25R	TCGGCCCTAGTGGCCGTTGGAATCGAATGGCTCTATGATTGG
	185	Up5UTRkpx25F	GACAGCCTCGATTATTTCTCTTTCTCTTCG

	186	Down3UTRkpx25R	AAGGCGTACCAAGAGCCTTTAGC
yps1	187	3UTRyps1F	TCGGCCGATCAGGCCTTCAGCTACCTTTCTCTCTGTTTGG
	188	3UTRyps1R	CTGGCCCGGGCGTCTGGTTGTTTGTATTAGC
	189	5UTRyps1F	ACGCCCGGGCCAGGGACCTAATTATGACATC
	190	5UTRyps1R	TCGGCCCTAGTGGCCCGCTGAAGTCCAACTGTTGAACG
	191	Up5UTRyps1F	CGAACCTAATCAATGACGGTTACGAG
	192	Down3UTRyps1R	TCGGCATTATCTGGTAGATCCGG
ctse	193	3UTRctseF	TCGGCCGATCAGGCCGTGTTACGTTGGCAGTTTGACTAAGG
	194	3UTRctseR	AGAAGTACCCGGGCGAACATGAACATATTGGCTG
	195	5UTRctseF	ATGTTCGCCCGGGTACTTCTCTGTTCACTTTGGGTCTTATTC
	196	5UTRctseR	TCGGCCCTAGTGGCCTGAGTGTTGGTCCCTGCATTATTG
	197	Up5UTRctseF	TTGATAAGCGGCTACCAAGTCAGAC
	198	Down3UTRctseR	AACATTGACCCTTGAGTTGTTACTCGG
kex1	199	3UTRkex1F	TCGGCCGATCAGGCCTGGACGACCTGGAATCCAAC
	200	3UTRkex1R	TGACAACCCGGGTCGGCAGATGAGTCTTTG
	201	5UTRkex1F	CCGACCCGGGTTGTCATTATTGGTGGTAAGGC
	202	5UTRkex1R	TCGGCCCTAGTGGCCTCGATTGGAGAGGCTGACACC
	203	Up5UTRkex1F	AGTGAAGAGAATTCACGAGTACAAGAGAC
	204	Down3UTRkex1R	GGTAAAACCGTCGTCATTGCTATTGC
kex2	205	3UTRkex2F	TCGGCCGATCAGGCCCTCTCGGAAGTCTCCAAAGCC
	206	3UTRkex2R	TGAAGGCCCGGGATAGTCGCCTTCCGGTTTC
	207	5UTRkex2F	GACTATCCCGGGCCTTCATCATCGGAGTC
	208	5UTRkex2R	TCGGCCCTAGTGGCCGATCTCATCCCAGCCGATGAC
	209	Up5UTRkex2F	CCTACATCAAATAAATCCGCCTGCG
	210	Down3UTRkex2R	CTCCGCATATAGTACCCATCCAGG
sub2	251	nCDSsub2F	GCACACTCGCTTTTGATACCATCTC
	252	nCDSsub2R	ATCCGAGTCATCAAGTACATCCTTGG
kpx1	253	nCDSkpx1F	AGGGCTTTGGTGTTAGCTGACTC
	254	nCDSkpx1R	CCAATAACACCATGGCAACTACAGC
kpx2	255	nCDSkpx2F	ATGAAATATTTGCCACTCGTTGCTACC
	256	nCDSkpx2R	GTCATCTTGAATGGGCTAGGCTC
yps3	257	nCDSyps3F	GCTTATCCAAACTTCTGCTTTTGCTCTC
	258	nCDSyps3R	GACATCGAGGAAGGCAAGTAATCAG
kpx4	259	nCDSkpx4F	GTGTTGTTACTGTCTACAGGCTATGC
	260	nCDSkpx4R	AGTTTTGCGGCTACCAGCATTTG
pep4	261	nCDSpep4F	CTCTCTACTCTAGGTATTGGTGCTGAAG
	262	nCDSpep4R	ACCTACTGCATCTTTGCCTAGGTC
kpx5	263	nCDSkpx5F	CACTCGCTTTTGATACCATCTCTTGAC
	264	nCDSkpx5R	GATCCGAGTCATCAAGTACATCCTTGG
kpx6	265	nCDSkpx6F	CCATACCGCTCCTAATTTTGCTTCTG
	266	nCDSkpx6R	TTCATCCCAATAATCCTCGTCCTCTTC
prb1	267	nCDSprb1F	AAACTCTTGGGCCAAGTTTTCAACAG
	268	nCDSprb1R	GATTGGCTATCTTATCTGCCATAGCAT
kpx7	269	nCDSkpx7F	TACCAAACAGAGAGAAAGGTAGCTGAAG
	270	nCDSkpx7R	TTTGATAGCGAGCGTACCGGTTTC
kpx8	271	nCDSkpx8F	CCAATCGCCTTTCAAGGCTAACAAG

	272	nCDSkpx8R	ACATGGCGTTGCACTTAAAGATGC
yps2	273	nCDSyps2F	GAACTAGTGTTAATGTGAGACACCGAC
	274	nCDSyps2R	CCTAGCAGCTTAATTCTACTTGCACTC
yps7	275	nCDSyps7F	ACAGTCAACAGACTGGCCAATGATG
	276	nCDSyps7R	GGTTTTGTCTCTGATATGCTTTTCGTCG
kpx9	277	nCDSkpx9F	CTGATACCTCCGCTAGAATCGAATTTGC
	278	nCDSkpx9R	CTAGTTCAACTTGTCTCTGCGACAATTG
kpx10	279	nCDSkpx10F	TAGATCCAGAGTTCTCGTTCCCAG
	280	nCDSkpx10R	CGGAGCATCTGTTCAAATCAGCAAC
kpx11	281	nCDSkpx11F	AATATTTGCCACTCGTTGCTACCCTG
	282	nCDSkpx11R	TGGACCGTGGTACTTGAAGTTTGAC
prtP	283	nCDSprtPF	CCAGTGGAAGAATTGAGAAAATCACAGC
	284	nCDSprtPR	AAGGCCTGGTAGGATTGGCATC
kpx12	285	nCDSkpx12F	CTTCTAGGTATACTTTGTTGTGTTGCAGC
	286	nCDSkpx12R	TTATTTAAGCGTGTAAGTGAACCCAAGAG
kpx13	287	nCDSkpx13F	CTTCTATCTTTGCTAGCATTGTCGCAG
	288	nCDSkpx13R	AGAAATCACCGTGTGTCCATCTGAC
kpx14	289	nCDSkpx14F	ATCGGCGTTTGCACCCTTATCTAG
	290	nCDSkpx14R	TACTGCTTCCTTTAGTAGCGGTGTC
kpx15	291	nCDSkpx15F	TATTTGCCAGCACTTCGCTTAGC
	292	nCDSkpx15R	AATGCCGCACGTTTGGGATG
prc1	293	nCDSprc1F	ATGAGAATTCTCTGGCTGATCATATGGG
	294	nCDSprc1R	TCCTAAAGCTATTGGTCTGTCCTTGC
kpx17	295	nCDSkpx17F	CTCTATAAGACCACCTTGTCAATAGCACAC
	296	nCDSkpx17R	GTCTTACTCCTTTCGGTATTCTGCTCC
kpx18	297	nCDSkpx18F	TGCGTCTATCCTAACGCTTTGCAAG
	298	nCDSkpx18R	TAAGAGGCAGAACTGTGAATAGTCCAG
kpx19	299	nCDSkpx19F	TCAGCATTTGCTACCCAACTGAGAG
	300	nCDSkpx19R	AGTTCTCGTTCCCAGCTGTAGTG
kpx20	301	nCDSkpx20F	AGATTTGGAACCCAAGTTGACTTGAG
	302	nCDSkpx20R	CAACACTTCTTATTATCCTACAGGCTGC
kpx21	303	nCDSkpx21F	CGTGGTTTGAGTAACATGAGTTGTTATCG
	304	nCDSkpx21R	ATTGCTGTTCTATTGTCAGCATTTGCTG
kpx22	305	nCDSkpx22F	AAGTTCGTGAAGTAGAGCCAGATGTG
	306	nCDSkpx22R	TTGGTATTGACAGCCCTCAGCATTG
kpx23	307	nCDSkpx23F	GTACTGAATTGAGGCATAAACCTCTGTTC
	308	nCDSkpx23R	CTATAAGACCACCTTGTCAATAGCACAC
kpx24	309	nCDSkpx24F	GATCGCTTGATAATAGCTCCATGTAAGC
	310	nCDSkpx24R	GCTCACAAGTCCTTCAGTAGATCCA
kpx25	311	nCDSkpx25F	TCCGTCAATTACGCTAACGGCATC
	312	nCDSkpx25R	AAGTCCCAAAGAGCAGAACGCTG
yps1	313	nCDSyps1F	AAAACGTTGTTGGCGTTCAACAGTTG
	314	nCDSyps1R	ATACTATACACACGCCGAGAATAACTACC
ctse	315	nCDSctseF	AGGCAATAATGCAGGGACCAACAC
	316	nCDSctseR	AAGGTACAAGTGTTTCCACCCTTAGTC
kex1	317	nCDSkex1F	TACTGCTTCCTTTAGTAGCGGTGTC

	318	nCDSkex1R	TGAGAAGTTTGCTCATCACACTATTGTCC
kex2	319	nCDSkex2F	CGGAATCCAAGGTGAATTGACTATTGG
	320	nCDSkex2R	CGCTTAGCATGCTGGATCTTAATTGG
pep4	221	NUp5UTRpep4F	CATACGAGCGAATTGAAGAAGCAGATG
	222	NDown3UTRpep4R	TTGAGCCATACTTCCAACTATCTAACCG
kpx5	225	NUp5UTRkpx5F	GTGAACTCGATTGGAATATCGCTGG
	226	Down3UTRkpx5R	CAGTGACGGCGATGTCATTAATTACC
kpx7	233	NUp5UTRkpx7F	AAATTCATCGTAGCAGATGGGACAGG
	234	NDown3UTRkpx7R	CGACAAATACAAAGAGCCTTCCAACG
		PucSeqF	CTTTTTACGGTTCCTGGCCTTTTGC
		PAox1SeqR	GGTTTCATTCAACCTTTCGTCTTTGGATG

3.2 Instruments, reagents, media

Table 5: Instruments and materials used

Application	Instrument	Manufacturer
Electrotransformation	MicroPulserTM	BIO-RAD, USA
Mixing	Vortex-Genie 2	Scientific Industries Inc, USA
OD_{600} measurements	BioPhotometer	Eppendorf, Germany
PCR	GeneAmp®PCR System 2700	Applied Biosystems, USA
Shaker	HT MiltronII	Infors AG, Swiss
Absorption measurement	Nanodrop 2000c	Thermo Fisher Scientific Inc, USA
Centrifuges	Centrifuge 5810R	Eppendorf, Germany
	Centrifuge 5415R	Eppendorf, 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
Deep Well Plates	SCIENCEWARE 96 Deep-Well Plate	Bel-Art, Wayne, USA
SDS-PAGE, Western Blot	XCell SureLock™ Mini-Cell Electrophoresis System	Life Technologies Corporation
	Ø NuPAGE® Novex 4-12 % BisTris Gel 1.0 mm, 15 well	Life Technologies Corporation
Nitrocellulose membrane for Western blotting	Roti®-NC	Roth Gmbh, Germany
Blotting	Filter paper	Whatman International Ltd., GB

Table 6: Reagents

Reagent	Supplier
Agar	Bacto Laboratories Pty Ltd., Australia
Agarose	Biozyme, Germany
Albumin Fraction V	Roth GmbH, Germany
Ammonium acetate	Roth GmbH, Germany
Biotin	Roth GmbH, Germany
Dimethylsulfoxid	Roth GmbH, Germany
dNTPs	Roth GmbH, Germany
Dream Taq 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
Glycerol	Roth GmbH, Germany
Glycine	Roth GmbH, Germany
LB (Luria-Bertani)	Roth GmbH, Germany
Methanol	Roth GmbH, Germany
O´Gene Ruler DNA Ladder Mix	Thermo Fisher Scientific Inc., USA
Peptone	
Phusion® Hot Start High-Fidelity DNA Polymerase	Finnzymes, Finnland
Potassium phosphate	Roth GmbH, Germany
Restriction enzymes	Thermo Fisher Scientific Inc., USA
SDS	Roth GmbH, Germany
Sodium chloride	Roth GmbH, Germany
Sodium hydroxide	Roth GmbH, Germany
Sorbitol	Roth GmbH, Germany
Super Signal® West Pico Chemiluminescence Substrate	Thermo Fisher Scientific Inc., USA
T4 DNA-Ligase	Thermo Fisher Scientific Inc., USA
TRIS	Roth GmbH, Germany
Tween®20	Roth GmbH, Germany
Triton X-100	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	Thermo Fisher Scientific Inc, USA
Anti-alpha2, 6-Sialyltransferase (C) Rabbit IgG Antibody (ST6Gall)	IBL International, Germany
Anti-Rabbit IgG-HRP Antibody	Sigma Aldrich
Anti-FLAG Antibody	Sigma Aldrich
Anti-Mouse IgG-HRP Antibody	Sigma Aldrich
Congo red	Sigma Aldrich
Calcofluor white	Sigma Aldrich

Table 7: Media and buffers

Medium / Buffer	Composition
BEDS	5% DMSO, 3% ethylene glycol, 1 M sorbitol, 10 mM bicine
Biotin (500x)	200 mg/l
BMGY	10 g/l yeast extract, 20 g/l peptone, 100 ml YNB (10x), 100 ml potassium phosphate (1 M, pH 6), 100 ml glycerol (10%), 4 ml biotin
DTT (1 M)	154.25 g/l
Glucose (10x)	200 g/l
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), 2ml biotin (500x), 20 g/l glucose, (20 g/l agar)
MM-medium (agar)	100 ml YNB (10x), 2 ml biotin (500x), 10 ml methanol, (20 g/l agar)
MOPS buffer (10x)	104.5 g/l MOPS, 60.6 g/l Tris, 10 g/l SDS, 3g/l EDTA
Phosphate buffer	1 M (K_2HPO_4 30 g/l, KH_2PO_4 118 g/l), pH 6 set with KOH
Tris-HCl	100 mM Tris, pH7 set with HCl
YPD (agar)	10 g/l yeast extract, 20 g/l peptone, 100 ml glucose (10x), (20 g/l agar)
YNB (10x)	134 g/l
Yeast lysis buffer	20 ml/l Triton X-100, 100 ml/l SDS (10%), 20 ml/l 5 NaCl (5M), 2 ml/l EDTA (0.5M), 10 ml Tris (1M, pH 8), dH_2O to 1l
Working Solutions for Western Blots	
TBS buffer (10x)	12.12 g/l Tris, 87.6 g/l NaCl, pH 7.5 set with HCl
TBS buffer (1x)	100 ml 10x stock, 900 ml dH_2O
TBS (10x, for Tween/Triton)	24.24 g/l Tris, 292 g/l NaCl, pH 7.5 set with HCl
TBS (1x, for Tween/Triton)	100 ml 10x stock, 900 ml dH_2O, 500 µl/l Tween®20, 2 ml/l Triton X-100
Blocking buffer	3 g/l Albumin Fraction V / 100 ml TBS buffer
Primary Antibody anti-hST6Gal I	1 µg/ml in ddH_2O
Secondary Antibody Anti-Rabbit IgG-HRP	1:1000 in blocking buffer
Primary Antibody anti-FLAG	
Secondary Antibody Anti-Mouse IgG-HRP	1:5000 in blocking buffer
Stripping buffer	620 ml/l Tris/HCl (100 mM, ph 7), 200 ml SDS (10%), 4.5 ml β-mercaptoethanol, dH_2O ad 1000 ml

Table 8: Restriction enzymes used

Restriction enzyme	Recognition sequence	Source
SfiI	5´…GGCCNNNN$^{\downarrow}$NGGCC…3´ 3´…CCGGN$_{\uparrow}$NNNNCCGG…5´	New England Biolabs Inc., Ipswich, USA
SmaI	5´…CCC$^{\downarrow}$GGG…3´ 3´…GGG$_{\uparrow}$CCC…5´	Thermo Fisher Scientific Inc., USA
SwaI	5´…ATTT$^{\downarrow}$AAAT…3´ 3´…TAAA$_{\uparrow}$TTTA…5´	Thermo Fisher Scientific Inc., USA

3.3 Methods

3.3.6 General methods

3.3.6.3 PCR

PCRs were performed using the Dream Taq DNA Polymerase, whereas for cloning and engineering work the Phusion® Hot Start High-Fidelity DNA Polymerase was utilized (see Table 9 and Table 10).

Colony PCRs were arranged using the Dream Taq DNA Polymerase mix, but instead of adding 1 µl of template DNA, very little of cell material from a single colony was transferred to the PCR mix using a sterile toothpick.

Table 9: Standard PCR Mix for Phusion and Dream Taq DNA Polymerase

Phusion DNA Polymerase	1x	Dream Taq DNA Polymerase	1x
HF 5x Buffer	10 µl	Go Taq Buffer 5x	1.5 µl
2 mM dNTPs	5 µl	2 mM dNTPs	2.5 µl
Phusion DNA Polymerase	0.3 µl	Dream Taq DNA Polymerase	0.15 µl
Primer forward	2.5 µl	Primer forward	1.25 µl
Primer reverse	2.5 µl	Primer reverse	1.25 µl
Template DNA (~20 ng/µl)	1 µl	Template DNA (~20 ng/µl)	1 µl
dH_2O	add to 50 µl	dH_2O	add to 25 µl

Table 10: Standard PCR conditions

Condition	Dream Taq DNA Polymerase	Phusion DNA Polymerase
Denaturation (initial)	95°C, 4 min	98°C, 30 s
Denaturation	95°C, 10 s	98°C, 10 s
Annealing	58°C, 30 s	58°C, 20 s
Extension	72°C, 1 min/kb	72°C, 15 s/kb
Extension (final)	72°C, 10 min	72°C, 10 min
Number of cycles	35	30

3.3.6.4 Gel electrophoresis

For control and preparative gels agarose was added to 1xTAE buffer (~230 ml) to a final concentration of 1% and boiled for complete dissolution. 5 µl of ehtidium bromide (10 mg/ml) were added before pouring the gel. Different combs generating 10 to 30 slots were used according to requirements.

3.3.6.5 Electro transformation into *E. coli*

80 µl of electrocompetent *E. coli* TOP10 F´ 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 cooled electroporation cuvettes and incubated on ice for at least 5 min.

For electroporation the program “Eco2” was chosen (2.5 kV, 5-6 ms). *E. coli* cells were regenerated with 1 ml LB medium and kept at 37°C and 600 rpm for 2 h. 200 µl of the culture were plated on selective LB agar. Additional the remaining cells were spun down shortly, most of the supernatant was removed and the rest was plated on selective LB agar.

3.3.6.6 Isolation of *P. pastoris* genomic DNA

For the isolation of genomic DNA an adapted protocol by Hofmann and Winston [43] was used.

Wells of a Deep Well Plate (DWP) were filled with 600 µl of YPD medium each and were inoculated with a single colony of the respective *P. pastoris* strain. The DWP was shaken for 24h at 28°C, 320 rpm and 80% humidity. Subsequently, the cultures were pipette into Eppendorf tubes and spun down for 2 min at maximum speed (14 000 rpm).

At this point the supernatant was decanted and in the following order 150 µl of Yeast lysis buffer, 150 µl of phenol: chloroform: isoamyl alcohol (25:24:1) and 0.3 g of acid washed glass beads were added to the pellet. For cell lysis the tubes were vortexed for at least 8 min and 150 µl of TE buffer were added. By spinning for 10 min at maximum speed in a centrifuge the phases were separated and the aqueous phase was transferred to a new tube. After adding 1 ml of 100% ice-cold ethanol the genomic DNA (gDNA) precipitated as white coils during incubation at -20°C for 30 min at least. By centrifugation at 4°C the pellet was fixed at the bottom of the tube and the supernatant was removed. The pellet was air-dried at 60°C and in a final step the gDNA was dissolved in 200 µl of sterile water.

3.3.6.7 Electro competent *P. pastoris* cells and Transformation

For preparation of electro competent *P. pastoris* cells a condensed protocol by Lin-Cereghino [44] was used.

A 50 ml overnight culture (ONC) of the respective *P. pastoris* strain in YPD was started inoculating it with a single colony. Next day the main culture was started with an OD_{600}=0.2 in 50 ml YPD and grown for approximately 4 h to an OD_{600}=0.8-1.0. After centrifuging the culture at 1600 rpm for 10 min at room temperature and discarding the supernatant, the cell pellet was resuspended in 9 ml BEDS solution and 1 ml 1M DTT. The resuspended cell pellet was warmed in hands and shaked lightely for 5 min. The suspension was centrifuged again at 1600 rpm for 10 min at room temperature, the supernatant was discarded and the cell pellet resuspended in 1 ml BEDS solution.

At this point 800-1000 ng of linearized plasmid DNA were mixed with 100 µl of electrocompetent *P. pastoris* cells, transferred into a pre-cooled electroporation cuvette and incubated on ice for 2-3 min. For transformation the program “Pic” was chosen for pulsing the

cells (1.5 kV) and subsequently 500 µl of 1M sorbitol and 500 µl of YPD media were added. When using a *Δhis* strain no YPD media was used, but only 1 ml of 1M sorbitol. After recovering at 28°C and 110 rpm for 2 h aliquots of 200 µl were plated on selective agar plates (YPD with 100 µg/ml zeocin or MD-plates for a *Δhis* strain).

3.3.6.8 Analysis of cell growth

Growth curves and measurement of cellular biomass produced during culturing of protease-deficient *P. pastoris* strains was performed by growing strains in 250 ml baffled shake flasks containing 50 ml YPD media at 28°C for 72 h. Cell growth was measured by OD_{600} in triplicate. The first 12 h OD_{600} was measured every 2 h and then after 24, 48 and 72 h.

After 24, 48 and 72 h of growth, 1 ml of each culture was removed and cells were pelleted by centrifugation at maximum speed (14 000 rpm) for 5 min. Cell pellets were washed once in dH_2O to remove medium components and dried at 120°C for 24 h. The dry cell weight (g l^{-1}) was calculated.

Protease-deficient *P. pastoris* strains were grown overnight in liquid YPD medium at 28°C and equal amounts of cells (approximately OD_{600} of 1) were further diluted in H_2O in from 10^{-1} – 10^{-3}. Aliquots of each dilution were applied on YPD agar supplemented with zeocin (10 µg/ml) and cell wall perturbing compounds SDS (0.025%), calcofluor white (10 µg/ml) and congo red (30 µg/ml) and cultured at 28°C for 2-3 days.

3.3.6.9 Expression of Δ27ST6

To express and secrete the sialyltransferase Δ27ST6GalI the linearized expression plasmid pAαZSwaIFlagD27ST6 or pAZSwaILevSSFlagD27 (provided by Ahmad M., see also Figure 4, Table 11) was transformed into protease-deficient *P. pastoris* strains, where they integrated at the *AOX1* locus. The only difference between these two plasmids is the N-terminal α-factor prepro signal peptide (AlphaF SP) or the levanase secretion signal (LevSS). Because the α-factor prepro signal peptide is processed by Kex2 protease, the plasmid with LevSS fused N-terminal to Δ27ST6 was included to be able to test *kex2* deletion strain as well.

After Mut^S-screening (see also Suppl. Figure 34) appropriate clones were used to inoculate 200 ml of BMGY media in 2 l baffled shake flasks. After growing them to an OD_{600}~100 at 28°C and 110 rpm, induction was started with 0.5% methanol for 120 h, in doing so, shaking at 28°C and 110 rpm was continued. 1 ml of culture was removed after 12, 24, 48, 72, 96 and 120 h. Cells were pelleted by centrifugation at maximum speed (14 000 rpm) for 5 min at 4°C. The supernatant was transferred into a new tube and both, supernatant as well as cell pellet, were stored at -20°C for further analysis.

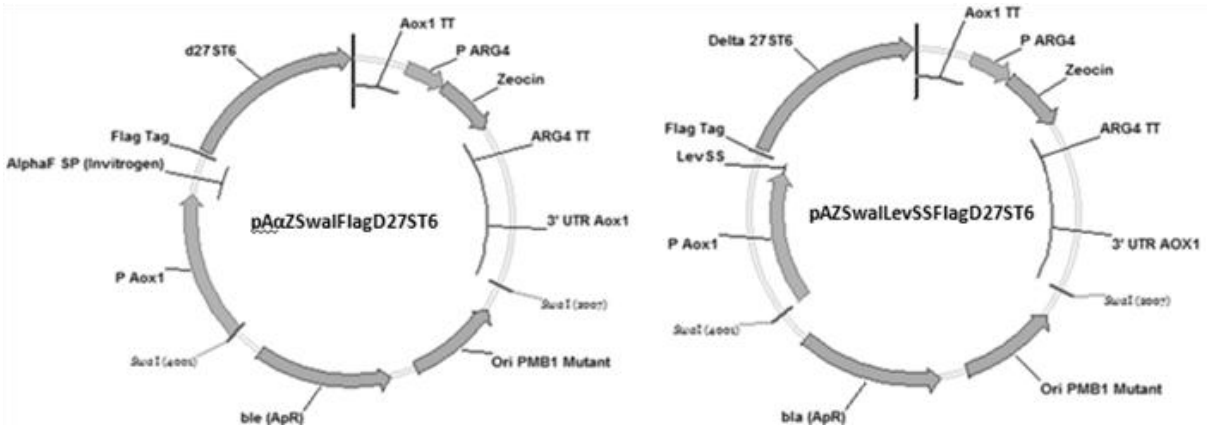


Figure 4: Maps of the plasmids pAαZSwaIFlagD27ST6 and pAZSwaILevSSFlagD27ST6 for expression of Δ27ST6 in protease-deficient *P. pastoris* strains. The *SwaI*-linearized plasmids integrate at the *AOX1* locus after transformation and thus, protein expression can be induced by adding methanol.

Table 11: Features of plasmids pAαZSwaIFlagD27ST6 and pAZSwaILevSSFlagD27ST6 for expression of Δ27ST6.

Feature name	Function
Ori PMB 1 Mutant	Origin of replication in *E. coli*
Ble (ApR)	Ampicillin resistance gene to confer resistance against ampicillin to *E. coli*
Zeocin	Zeocin resistance gene, complemented with an eukaryotic promoter (pARG4) to confer resistance against zeocin to *P. pastoris*
PAOX1 and 3´UTR AOX1	DNA sequences homologous to flanking regions of *AOX1* locus, needed for homologous recombination event
SwaI	Restriction site for linearizing the expression plasmid
Alpha SP	*S. cerevisiae* α-factor prepro signal peptide to secrete protein
LevSS	Levanase secretion signal to secrete protein (needed for *kex2* knockout)
d27ST6	α2, 6-Sialyltransferase ST6 with the first 27 residues removed

3.3.6.10 Western blotting

Western blotting was used to assess the quality of the Δ27ST6 protein secreted from various protease-deficient *P. pastoris* strains.

For SDS-PAGE an electrophoresis system from Life Technologies Corporation with 4-12% BisTris gels was used. Three different protocols for sample preparation were used. 1) 7 µl of 4xLDS sample buffer and 1 µl of 1M DTT were added to 20 µl of supernatant to a final volume of 28 µl. Samples were denatured at 99°C for 10 min and 15 µl were loaded onto the gel. 2) 100 µl of 100% TCA were added to 900 µl of supernatant. After incubating at 4°C for at least 2 h, the samples were centrifuged at maximum speed (14 000 rpm) at 4°C for 15 min. They were washed twice with ice-cold dH_2O and centrifuged at maximum speed at 4°C for 15 min. The concentrated protein was resuspended in 19.5 µl dH_2O, 7.5 µl 4x LDS sample buffer and 3 µl reducing agent were added to a final volume of 30 µl. The samples were denatured at 99°C for 10 min and 10 µl were loaded onto the gel. 3) Methanol-chloroform precipitation: 480 µl MeOH and 160 µl CHCl3 are added to 200 µl protein sample

and mixed by vortexing. After adding 640 µl ddH_2O the sample is vortexed and spinned for 5 minutes at 14 000 rpm at 4°C. The upper layer is sucked off, 300 µl MeOH is added, the sample is vortexed and spinned 30 minutes, 14 000 rpm at 4°C. Supernatant is sucked off completely and the air dryed pellet is resuspended in SDS sample buffer. The gels were run at least for 1 h 40 min with 125V, 120 mA and 25 W.

Subsequently the proteins were transferred to a nitrocellulose membrane using the XCell II™ Blot Module from Life Technologies Corporation with the specifications of the NuPageBlot program (25V, 160 mA, 17 W, 1 h)

The membrane was washed twice with 20 ml TBS buffer for 7 min each at room temperature with shaking and incubated for 1 h in blocking buffer at room temperature or at 4°C overnight. Thereafter it was washed twice with TBS-Tween/Triton buffer for 7 min each at room temperature and once with TBS buffer for 7 min at room temperature. After incubating with the primary antibody (Anti-α2, 6-Sialyltransferase (C) Rabbit IgG Antibody) for 1 h at room temperature the same wash steps with TBS-Tween/Triton buffer and TBS buffer as mentioned before were performed, before incubation with the secondary antibody (Anti-Rabbit IgG 1:1000, HRP-conjugated). The membrane was washed twice with TBS-Tween/Triton buffer for 7 min each and incubated with 7.5 ml of Super Signal West Pico Substrate Working Solution for 3 min. Signal detection was done with G: BOX.

Stripping and reprobing membranes

The complete removal of primary and secondary antibodies from membranes is possible following the method outlined here. The membrane was submerged in stripping buffer and incubated at 60°C for 30 min with occasional agitation. After washing the membrane twice with TBS-Tween/Triton buffer for 10 min each at room temperature, it was blocked by immersing in blocking buffer for 1 hour at room temperature. The immunodetection was performed as described above.

3.3.7 Construction of protease-deficient *P. pastoris* strains

For constructing different knockout cassettes a plasmid backbone (pPpKC1, see Table 12 for features) was ligated with an insert containing the 3´UTR and 5´UTR of the respective target gene. Before cloning the inserts were amplified from the homologous regions of the *P. pastoris* CBS7435 genome and 3´UTR and 5´UTR were joined by overlap-extension PCR. The insert as well as the backbone were prepared for ligation by digestion with the restriction enzyme *SfiI*.

3.3.7.3 Construction of knockout plasmids

The probability of proper locus integration in *P. pastoris* by homologous recombination is strongly dependent on the length of the homologous regions. In this work a length of ~1 kb 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 in a functional form elsewhere in the genome. Thus, homologous regions were designed to be partially inside (~200 bp) of the target gene to minimize the chance of this event.

A restriction site between the 3´UTR and 5´UTR region is needed for linearizing and therefore the *SmaI* restriction site CCCGGG was chosen. As shown in Figure 5 it was searched for CCC and GGG ~1kb downstream and upstream of the target ORF, so that after joining of 3´UTR and 5´UTR the *SmaI* restriction site is obtained. The primers for amplifying the 3´UTR and 5´UTR were designed in a way that the outer primers (P1: 5UTRgeneF, P4: 3UTRgeneR) contain the sequence for binding ~ 1 kb upstream or downstream of the target ORF, plus the CCC or GGG for *SmaI* restriction site, plus an overhang of ~10 bp for subsequent overlap extension PCR. The inner primers (P2: 5UTRgeneR, P3: 3UTRgeneF) bind ~200 bp inside of the target gene with an attached *SfiI*1 or *SfiI*2 restriction site. As shown in Table 8 the *SfiI* restriction site has five variable nucleotides and therefore cloning efficiency can be increased by changing these nucleotides and generating two different sites: *SfiI*1 and *SfiI*2. In general all primers were designed to have a T_m of the region active in the respective PCR of 58°C by using the IDT Oligoanalyzer 3.1 software and the Vector NTI software.

For amplification of the 3´UTRs and 5´UTRs genomic DNA of *P. pastoris* CBS7435 was used as a template. After purification by preparative agarose gel equal amounts of the respective fragments were fused by overlap-extension PCR using the outer primers P1 (5UTRgeneF) and P4 (3UTRgeneR). The plasmid backbones as well as the PCR products were digested with *SfiI*.

After ligation overnight the constructs were transformed into *E. coli* TOP10F´cells and plated onto LB_{Zeo} plates (LB with 25 µg/ml zeocin). Transformants were screened via colony PCRs using primers binding in the pUCori (PucSeqF) and the *pAOX1* (pAOX1SeqR) regions, each pointing towards the insert. Plasmids showing positive results in the colony PCR were isolated and sequenced by LGC Genomics (Berlin, Germany).

Isolated knockout plasmids were restricted with *SmaI* to obtain the linearized “Gene Disruption Fragment” as illustrated in Figure 6.

Table 12: Features of knockout plasmid pPpKC1.

Feature name	Function
pUCori	Origin of replication in *E. coli*
Zeocin	Zeocin resistance gene complemented with both a prokaryotic (pEM72) and a eukaryotic (pARG4) promoter to confer resistance against zeocin to *E. coli* and *P. pastoris*
FRT	Flippase recognition target. Sequence of 34 bp recognized by flippase.
FLP	Flippase gene under *AOX1* promoter. Intracellularly expressed flippase recognizes FRT regions and is capable of cutting out DNA sequences in between.
3´UTR and 5´UTR	DNA sequences homologous to flanking regions of target genes in the genome, needed for homologous recombination.

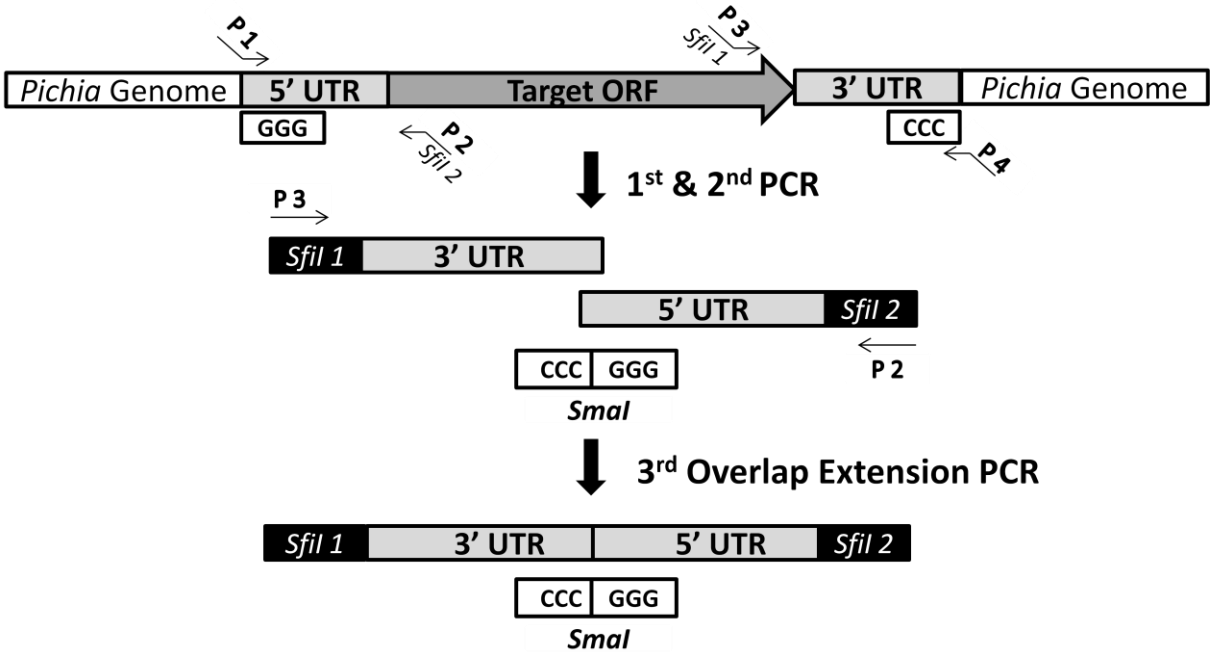


Figure 5: Schematic representation of generating the 3´UTR-5´UTR insert for cloning into pPpKC1. By doing the first two PCRs the 3´UTR and 5´UTR of the respective gene are amplified separately for subsequent third overlap-extension PCR. The resulting fragment contains the 3´UTR joined to the 5´UTR with a *SmaI* restriction site in between and flanked by two different *SfiI* restriction sites.

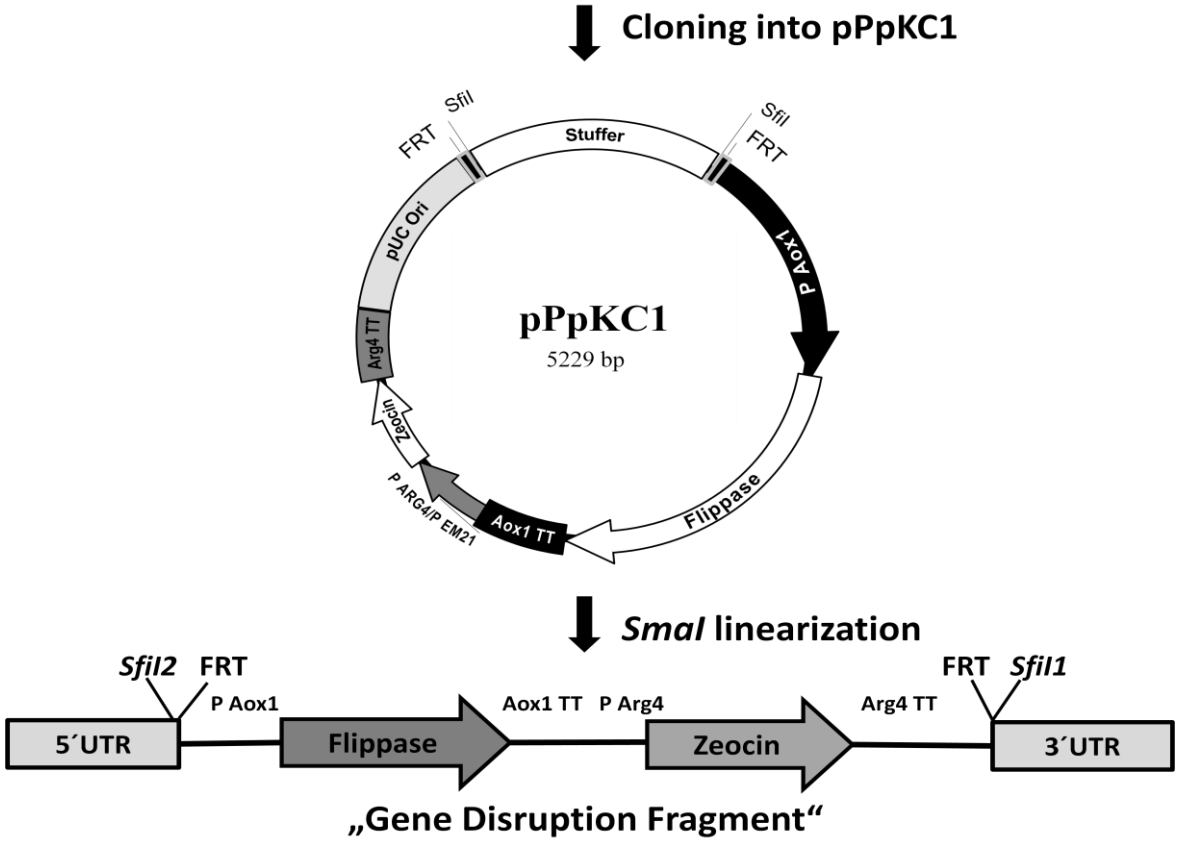


Figure 6: Schematic representation of cloning of the 3´UTR-5´UTR fragment into pPpKC1 for subsequent linearization with *SmaI*. The *SfiI*-restricted 3´UTR-5´UTR insert is cloned into the *SfiI*-cut pPpKC1 and after confirmation the knockout plasmid is linearized with *SmaI* for transformation into *P. pastoris* CBS 7435.

3.3.7.4 Generating knockouts in *P. pastoris*

Knockout plasmids were linearized with the restriction enzyme *SmaI* and ~800 ng DNA per transformation were used for transformation into *P. pastoris* CBS 7435. After transformation 200 µl of the cells were plated onto YPD_{Zeo} plates (YPD with 100 µg/ml zeocin) and incubated at 28°C for 3 days. Single colonies were picked with toothpicks and resuspended in YPD medium in DWP. Before isolating gDNA they were stamped onto YPD plates and YPD_{Zeo} plates.

To verify site-specific integration two independent PCRs were done to amplify the 5´UTR and 3´UTR regions. As shown in Figure 7 the outer primers P5 (Up5UTRgeneF) and P6 (Down3UTRgeneR) bind ~100 bp outside of the 5´UTR and 3´UTR regions, whereas the inner primers PAox1SeqR and PucSeqF bind inside the knockout plasmid. A PCR product is obtained only, if integration has occurred at the right locus.

In a first step gDNA of 24 clones of each knockout transformation was isolated and screened for the 5´UTR region. Clones, which gave positive results in the first PCR, were examined in a second PCR by using primers for the 3´UTR region. Clones with proper signals in both, 5´UTR as well as 3´UTR region, were selected for marker recycling.

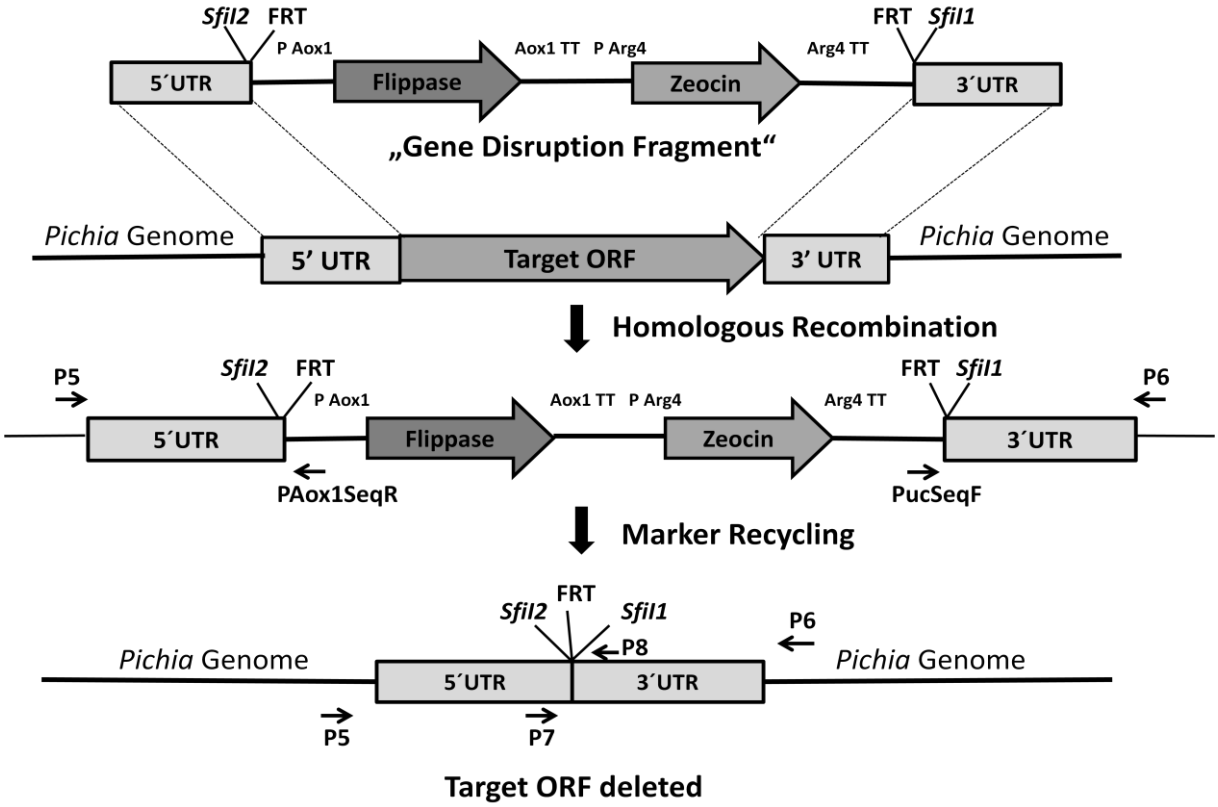


Figure 7: Schematic representation of generating knockouts in *P. pastoris*. After transformation homologous recombination occurs and the target ORF is replaced by the linear knockout plasmid. To verify proper integration event two independent PCRs were done to amplify the 5´UTR and 3´UTR regions. The outer primers P5 (Up5UTRgeneF) and P6 (Down3UTRgeneR) bind outside of the target locus, whereas the inner primers pAox1SeqR and PucSeqFbind inside the knockout plasmid. Clones, which gave positive results for both 5´UTR and 3´UTR regions after PCRs, were selected for marker recycling. The removal of the integrated cassette was proven by performing PCRs with two different primer pairs. P5 and P6 generating a shorter fragment compared to the wild type when the target ORF has been deleted. P7 (nCDSgeneF) and P8 (nCDSgeneR) were designed to eliminate the possibility that the excised target DNA has integrated somewhere else in a functional form in the genome.

Removal of zeocin marker

For removing the zeocin marker a flippase recombinase system is used that enables the plasmid its own removal from the genome after excision of the target DNA. A flippase under the control of the *AOX1* promoter recognizes the FRT sites (flippase recombination target) and cuts out the sequences between them as shown in Figure 8, leaving one FRT region flanked by two *SfiI* restriction sites. This way the knockout is performed as “clean” as possible, because only very little alterations in the genome of the wild type occur. Another advantage of removing the integrated cassette is that the cells lose their zeocin resistance and are available for further genetic engineering.

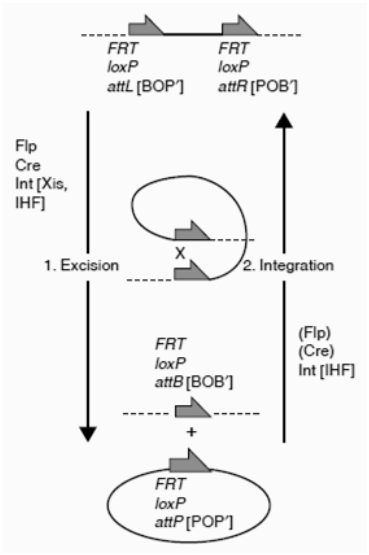


Figure 8: Mechanism of flipping out the sequence between two FRT sites by the flippase recombinase system (taken from [45]**).**

A colony with a proper integrated knockout cassette was used to inoculate a 50 ml ½ YPD culture and was shaken at 28°C and 110 rpm for 24 h. Cells were induced with 1% methanol for 3 days and subsequently a few µl of cell suspension were streaked out on YPD plates. After 2-3 days single colonies were picked, resuspended in DWP containing 300 µl YPD medium and stamped onto YPD plates and YPD_{zeo} plates. Cells growing on YPD, but not on YPD_{zeo} plates are showing the favored phenotype as they have lost the knockout cassette containing the zeocin resistance gene (see also Suppl. Figure 33).

Confirmation of deleted target ORF via PCR

To confirm the deleted target ORF of the respective gene a PCR with the outer primers P5 (Up5UTRgeneF) and P6 (Down3UTRgeneR) was performed by using gDNA as template as illustrated in Figure 7. If the ORF has been removed the resulting PCR fragment is shorter than the fragment obtained from the PCR with wild type gDNA. Additionally the whole region was sequenced by LGC Genomics (Berlin, Germany) using the same primers to receive full information about the situation in the genome.

But there is still a possibility that a part of excised and released genomic DNA reintegrates somewhere else in a fuctional form in the genome during the process of a double crossover by homologous recombination. To check for this event primers were designed that bind in the region between the start codon and the ~200 bp downstream area (P7: nCDSgeneF) and in the region between the stop codon and the ~200 bp upstream area (P8: nCDSgeneR) as indicated in Figure 7. This was possible because homologous regions were designed to be partially inside (~200 bp) of the target gene. By performing a PCR with this primer pair, a very short fragment of 100-300 bp is obtained, if the whole ORF is not present in the genome any more. Wild type gDNA as template leads to the full length ORF with the same primers

Additional notes

In some cases we were not able to knockout certain proteases. We assumed that either the respective protease is essential for viability of the cell or the zeocin as an antibiotic puts too much pressure on the cells. Especially the combination of both, potential lower viability due to a knockout and zeocin, might force this negative effect. To eliminate the problem with zeocin, we decided to change the selection marker to histidine (*HIS4*) for certain problem proteases. The whole knockout procedure remained the same as described above except of following modifications.

The 3´UTR-5´UTR insert was cut out from the original knockout plasmid with the zeocin resistance gene and ligated into the pPpKC1*HIS4* vector (Figure 9). The only differences of this plasmid compared to the pPpKC1 are the ampicillin resistance gene to confer resistance against ampicillin to *E. coli* and a *HIS4* marker for complementation in a *his4* deletion strain of *P. pastoris*. After transformation into electro competent *P. pastoris* cells, cells were plated onto MD plates. The procedure for confirmation via PCR and induction with methanol remained the same, but picked clones were stamped onto YPD and MD plates to find clones, which have lost their selection marker and are not able to grow on MD plates any more.

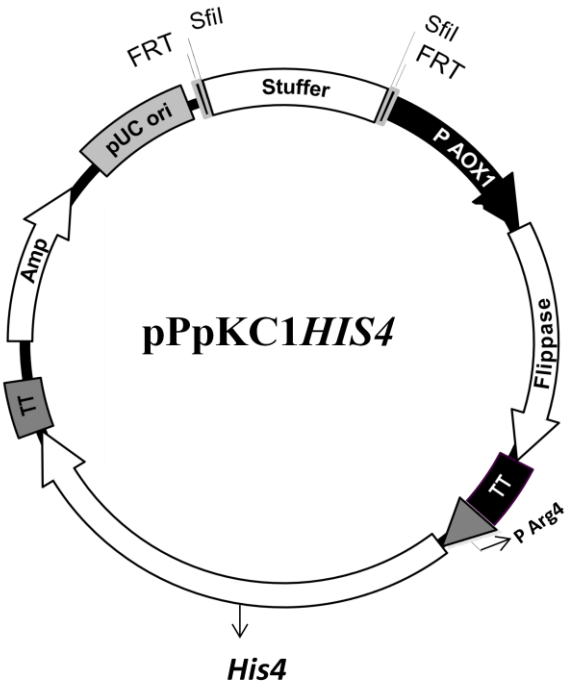


Figure 9: Map of the knockout plasmid pPpKC1*HIS4*. The only differences between pPpKC1*HIS4* and pPpKC1 are that pPpKC1*HIS4* has a *HIS4* marker for *P. pastoris* and an ampicillin resistance gene to confer resistance against ampicillin to *E. coli*.

4 RESULTS

4.1 Generated protease-deficient *P. pastoris* strains

As a result of the gene deletion and confirmation procedures, 27 out of the 35 selected protease genes were each successfully deleted in *P. pastoris*, as shown in Table 13. Beneath them are the already disrupted and partially studied ones *sub2, pep4, prb1, prc1, yps1, yps2, yps7, kex1 and kex2* and the putative proteases *kpx1 – kpx25*. Additionally a *pep4/prb1* double knockout strain was generated as well by transforming the *prb1* knockout cassette into *pep4* deletion strain followed by marker recycling as described in Material and methods.

We failed to obtain knockout strains of *yps3, kpx2, kpx6, kpx9, kpx11, kpx12, kpx13* and *kpx15*. Although the *yps3* gene has recently been knocked out [31], we decided to skip further attempts to generate this knockout, because to date there is no evidence that any of these knockouts have an influence on reducing degradation of expressed heterologous proteins.

Deletions of *pep4, pep4/prb1, kex2, kpx5* and *kpx7* were performed in CBS 7435 *Δhis4* strains, whereas all other knockouts were generated in CBS 7435 wild type strain.

Table 13: Protease-deficient *P. pastoris* strains generated in this project. Successfully deleted protease genes are marked green and each protease is listed with its gene name, gene ID, gene product, protease family, signal peptide (SP) and *S.* cerevisiae homologue. The strain names with their distinct ID numbers bracketed of the generated protease-deficient *P. pastoris* strains are listed in the right column.

#	Gene	Gene ID CBS 7435	Gene ID GS 115	Gene product	Protease family	SP	Homologues *S. cerevisiae*	Protease-deficient *P. pastoris* strain
1	*sub2*	299150304	254567147	Subtilisin 2 protease	S8	YES	*PRB1*	CBS 7435 Δ*sub2* (7)
2	*pep4*	328352648	254572446	Proteinase A	A1	YES	*PEP4*	CBS 7435 Δ*his4* Δ*pep4* (11)
3	*prb1*	328350290	254565530	Proteinase B	S8	YES	*PRB1*	CBS 7435 Δ*prb1* (b57_4)
4	*prc1*	328350528	254566035	Carboxypeptidase Y	S10	YES	*PRC1*	CBS 7435 Δ*prc1* (508)
5	*yps1*	38146742	254573826	Yapsin 1	A1	NO	*YPS1*	CBS 7435 Δ*yps1* (723)
6	*yps2*	328353469	254570819	Yapsin	A1	YES	*YPS1*	CBS 7435 Δ*yps2* (176)
7	*yps3*	328353465	254570827	Aspartic protease	A1	YES	*YPS1*	
8	*yps7*	328353374	254571021	Cathepsin D	A1	YES	*YPS7*	CBS 7435 Δ*yps7* (916)
9	*kex1*	3747109	238030858	Carboxypeptidase Kex1	S10	YES	*KEX1*	CBS 7435 Δ*kex1* (951)
10	*kex2*	38146738	254568178	Kex2 proprotein convertase	S8	YES	*KEX2*	CBS 7435 Δ*his4* Δ*kex2* (17b)
11	*prtP*	328352048	254567898	PI-type proteinase		YES	*FLO10*	CBS 7435 Δ*prtP* (435)
12	*ctse*	328350442	254565855	Yapsin	A1	NO	*MKC7 (YPS2)*	CBS 7435 Δ*ctse* (409)
13	*kpx1*	328354397	254573342	Immunglobulin A1 protease		YES	*ENT2*	CBS 7435 Δ*kpx1* (28)
14	*kpx2*	328351136	254567279	Aminopeptidase Y	M28	YES	*APE3*	
15	*kpx4*	328352885	254571979	Aspartic protease	A1	YES	*MKC7 (YPS2)*	CBS 7435 Δ*kpx4* (868)
16	*kpx5*	328351070	254567147	Proteinase B	S8	YES	*PRB1*	CBS 7435 Δ*his4* Δ*kpx5* (1)
17	*kpx6*	328350760	254566511	Proteinase B	S8	YES	YCR045C	
18	*kpx7*	328354158	254573826	Aspartic protease	A1	YES	*YPS1*	CBS 7435 Δ*his4* Δ*kpx7* (14)
19	*kpx8*	328353613	254570525	Aspartic protease	A1	YES	*YPS1*	CBS 7435 Δ*kpx8* (152)
20	*kpx9*	328350256	254565461	Aminopeptidase Y	M38	YES	*YDR415C*	
21	*kpx10*	328352415	254567898	C5a peptidase		YES	*FLO10*	CBS 7435 Δ*kpx10* (818)
22	*kpx11*	328351136	254567279	Aminopeptidase Y	M28	YES	*APE3*	
23	*kpx12*	328354499	254573130	Endoprotease	C13	YES	*GPI8*	

24	*kpx13*	328353128	254571503	Carboxypeptidase Y	S10	YES	*YBR139W*	
25	*kpx14*	328352413	254567902	Carboxypeptidase Y	S10	YES	*KEX1*	CBS 7435 Δ*kpx14* (342)
26	*kpx15*	328352277	254568178	Kex2 proprotein convertase	S8	YES	*KEX2*	
27	*kpx17*	328353349	254571075	Metalloprotease	M14	YES	*ECM14*	CBS 7435 Δ*kpx17* (542)
28	*kpx18*	328352415	254567898	C5a peptidase		YES	*FLO10*	CBS 7435 Δ*kpx18* (553)
29	*kpx19*	328352413	254567902	Carboxypeptidase Y	S10	YES	*KEX1*	CBS 7435 Δ*kpx19* (385)
30	*kpx20*	328353755	254570259	PI-type proteinase		YES	*CWP1*	CBS 7435 Δ*kpx20* (577)
31	*kpx21*	328353754	254570261	C5a peptidase		YES	*FLO10*	CBS 7435 Δ*kpx21* (616)
32	*kpx22*	328353613	254570525	Aspartic protease	A1	YES	*YPS1*	CBS 7435 Δ*kpx22* (625)
33	*kpx23*	328353349	254571075	Metalloprotease	M14	YES	*ECM14*	CBS 7435 Δ*kpx23* (657)
34	*kpx24*	328350485	254565945	Aspartic protease	A	NO	*SRT1*	CBS 7435 Δ*kpx24* (680)
35	*kpx25*	328352792	254572167	Metalloprotease	M	NO	*YBR074W*	CBS 7435 Δ*kpx25* (701)
36	*pep4* *prb1*	328352648 328350290	254572446 254565530	Proteinase B Carboxypeptidase Y	S8 S10	YES YES	*PRB1* *PRC1*	CBS 7435 *Δhis4 Δpep4 Δprb1* (6)

4.2 Confirmation of deleted target ORF

The knockout cassettes for all 35 (putative) protease genes were constructed during my project lab (MOL 737 Projektlabor Molekulare Mikrobiologie *Construction of knockout cassettes for generating protease-deficient* P. pastoris *strains*) and the procedure is shown by using the results for one protease in Supplementary information (6.1 Construction of knockout plasmids).

To verify site-specific integration of the knockout cassette, clones of each knockout transformation were screened for the 5´UTR region. Clones, which gave positive results in the first PCR, were examined for the 3´UTR region. Clones with proper signals in both, 5´UTR as well as 3´UTR region, were selected for marker recycling. The results of amplifying the 5´UTR and 3´UTR regions of each of the 27 generated knockout strains are shown in Supplementary information (6.2 Confirmation of site-specific integration of knockout cassette).

After recycling the knockout plasmid (see also Suppl. Figure 33), the DNA sequence between the flanking regions (5´UTR and 3´UTR) is removed except of one FRT site flanked by two *SfiI* restriction sites. For confirmation of the deleted target ORF of the respective protease, the particular region in the genome was amplified using two different primer pairs.

Primer pair I (Up5UTRgeneF, Down3UTRgeneR) binds in the upstream and downstream region of the respective gene, as shown in Figure 7. The PCR product of a successful protease knockout exhibits a shorter fragment compared to the wild type, e. g. *sub2* deletion leads to a fragment with a size of 2311 bp, whereas wild type is expected to have a size of 3391 bp (Figure 10, Table 14).

Primer pair II (nCDSgeneF, nCDSgeneR) binds inside the ORF of the respective gene 100 – 200 bp downstream of the start codon or rather upstream of the stop codon (Figure 7). This strategy was selected to rule out the potential event that excised and released gDNA reintegrates somewhere else in a functional form in the genome during the process of

homologous recombination. The PCR product obtained from a successful protease deletion is a very short fragment of 100 – 300 bp, whereas wild type shows the full ORF size, e. g. *sub2* deletion leads to a fragment with a size of 237 bp, whereas wild type is expected to have a size of 1317 bp (Figure 10, Table 14).

Additionally the whole region was sequenced by LGC Genomics (Berlin, Germany) using primer pair I to receive full information about the situation in the genome.

The results of the 27 successfully generated protease-deficient *P. pastoris* strains with the additional generated *pep4/prb1* double knockout strain are shown below (Figure 10 – Figure 38). The obtained agarose gel pictures after the two separate PCRs to confirm the successful deletions are explained by using the example of *sub2* deletion in Figure 10. Gel pictures of all protease knockouts are labeled the same way. The expected fragment lengths of the knockout and wild type strains obtained from PCRs using primer pair I and primer pair II are shown in Table 14.

Table 14: Expected fragment lengths of the knockout and wild type strains obtained from PCRs using primer pair I (Up5UTRgeneF & Down3UTRgeneR; PCR I) and primer pair II (nCDSgeneF, nCDSgeneR; PCR II). wt: wild type, Δ: protease knockout.

No.	Gene	PCR I		PCR II	
		Δ [bp]	wt [bp]	Δ [bp]	wt [bp]
1	***sub2***	2311	3391	237	1317
2	***pep4***	2141	3162	152	1172
3	***prb1***	2372	3754	237	1619
4	***prc1***	2350	3709	124	1483
5	***yps1***	2069	3604	148	1683
6	***yps2***	2390	3818	116	1544
8	***yps7***	2235	3822	126	1713
9	***kex1***	2216	3787	178	1749
10	***kex2***	2153	4167	198	2212
11	***prtP***	2453	7739	120	5406
12	***ctse***	2478	3912	132	1566
13	***kpx1***	2154	2894	206	946
15	***kpx4***	2250	3831	195	1776
16	***kpx5***	2210	3347	178	1315

No.	Gene	PCR I		PCR II	
		Δ [bp]	wt [bp]	Δ [bp]	wt [bp]
18	***kpx7***	2229	3788	182	1741
19	***kpx8***	2550	4068	175	1693
21	***kpx10***	2092	6536	169	4613
25	***kpx14***	2239	3829	221	1811
27	***kpx17***	1948	3558	173	1783
28	***kpx18***	2218	6983	196	4961
29	***kpx19***	2252	6726	321	4795
30	***kpx20***	2129	3695	134	1700
31	***kpx21***	2125	6233	167	4275
32	***kpx22***	1974	3343	251	1620
33	***kpx23***	1907	3354	120	1567
34	***kpx24***	1909	2601	137	829
35	***kpx25***	2247	4860	232	2845

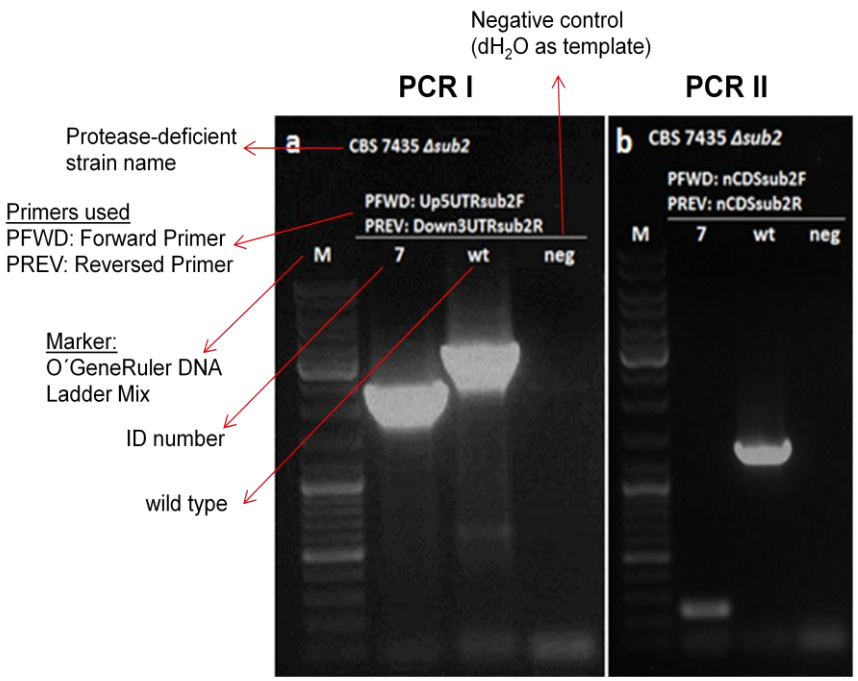


Figure 10: Agarose gel pictures for confirmation of *sub2* knockout. After recycling the knockout plasmid the DNA sequence between the flanking regions (5´UTR & 3´UTR) is removed except of one FRT site flanked by two *SfiI* restriction sites. For confirmation of deleted target ORF the particular region in the genome was amplified using two different primer pairs resulting in shorter fragments compared to the wild type that proves a successful knockout.
Gel pictures of all protease knockouts are labeled the same way.

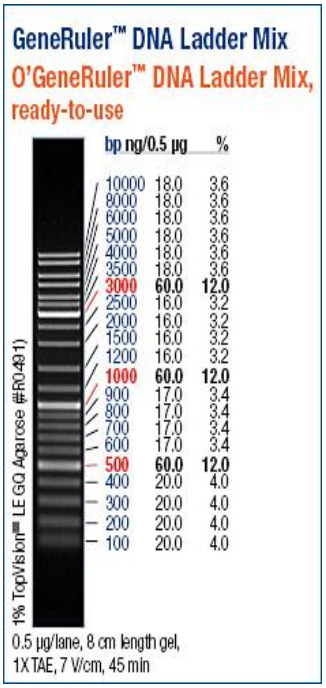


Figure 11: O´GeneRuler DNA Ladder Mix was used unless otherwise noted (taken from [46]**).**

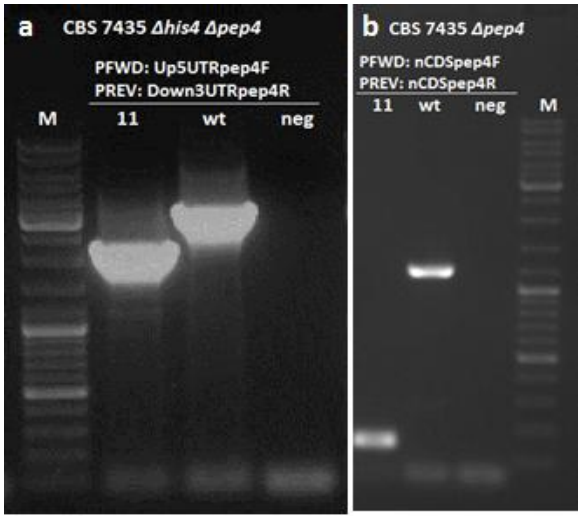


Figure 12: CBS 7435 *Δhis4 Δpep4*

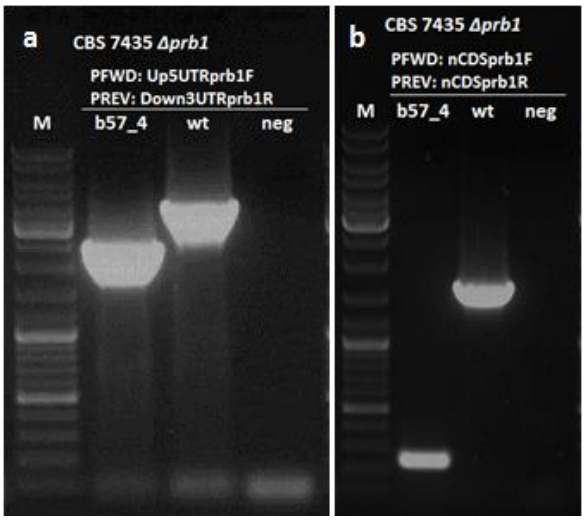


Figure 13: CBS 7435 *Δprb1*

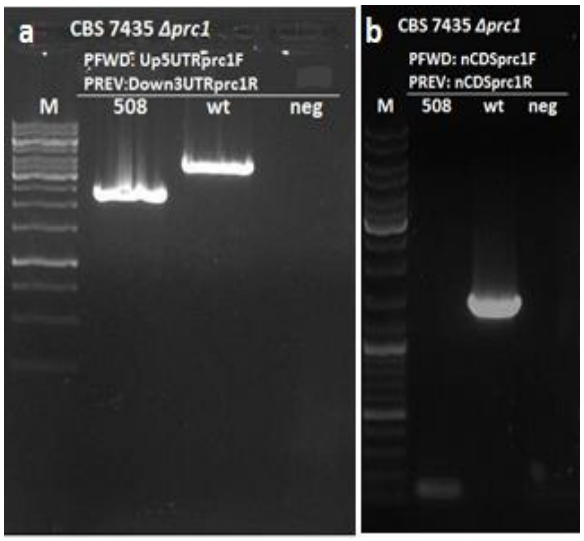


Figure 14: CBS 7435 *Δprc1*

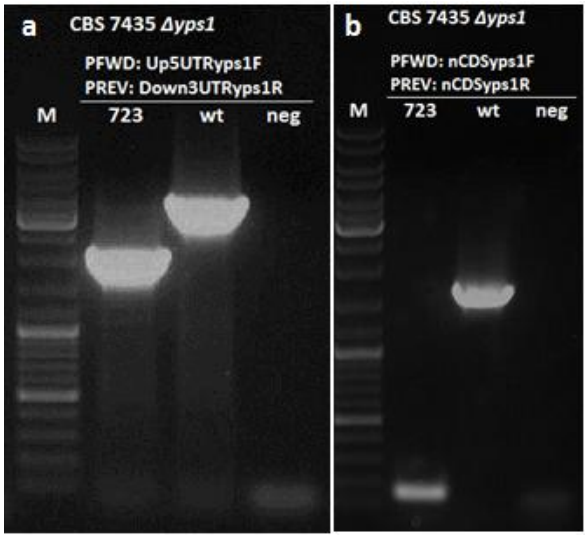


Figure 15: CBS 7435 *Δyps1*

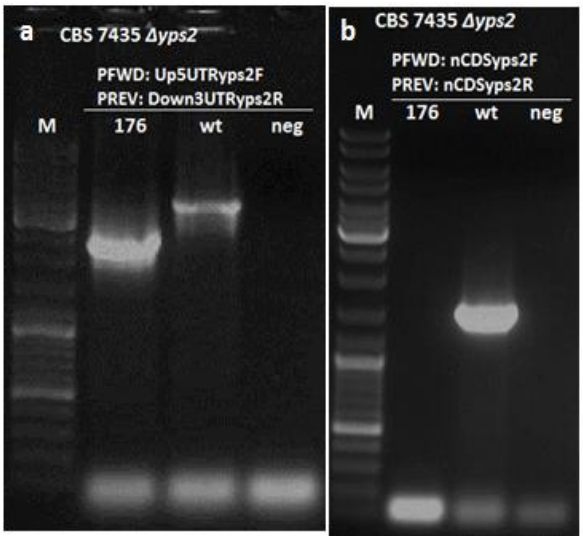


Figure 16: CBS 7435 *Δyps2*

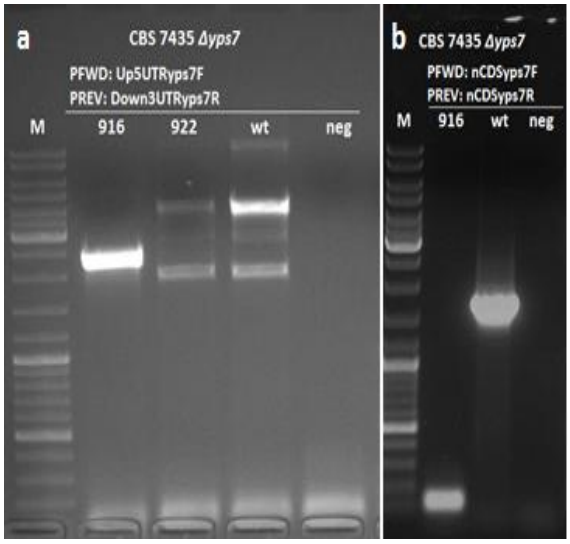


Figure 17: CBS 7435 *Δyps7*

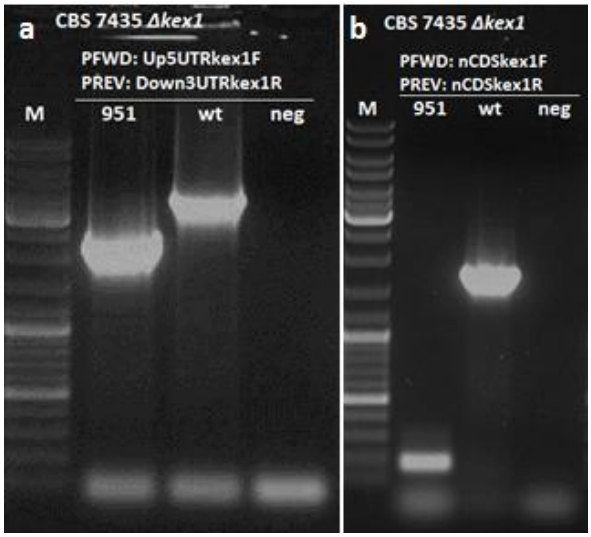


Figure 18: CBS 7435 *Δkex1*

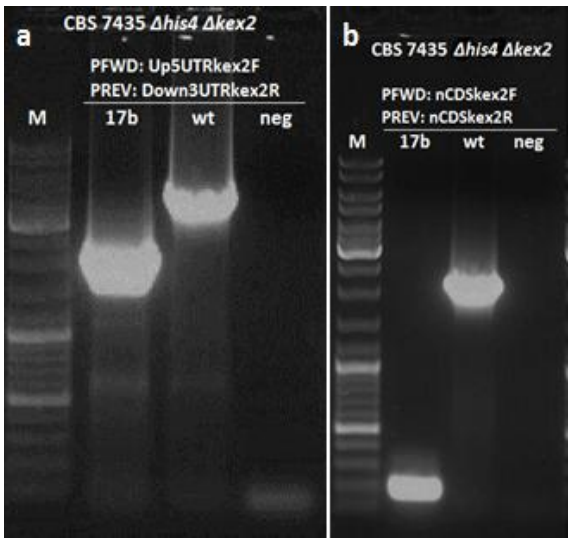


Figure 19: CBS 7435 *Δhis4 Δkex2*

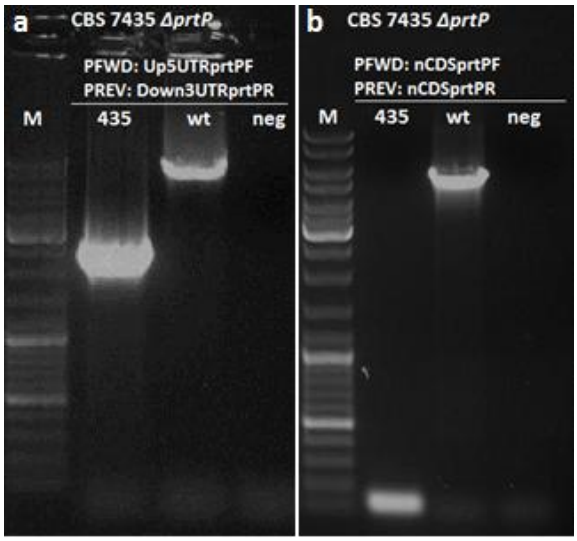


Figure 20: CBS 7435 *ΔprtP*

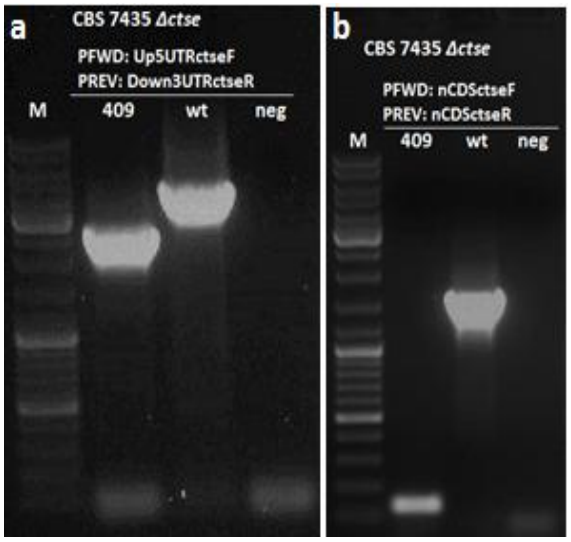


Figure 21: CBS 7435 *Δctse*

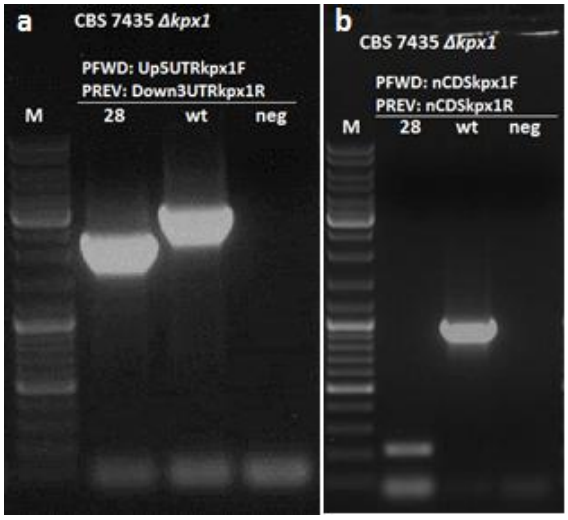


Figure 22: CBS 7435 *Δkpx1*

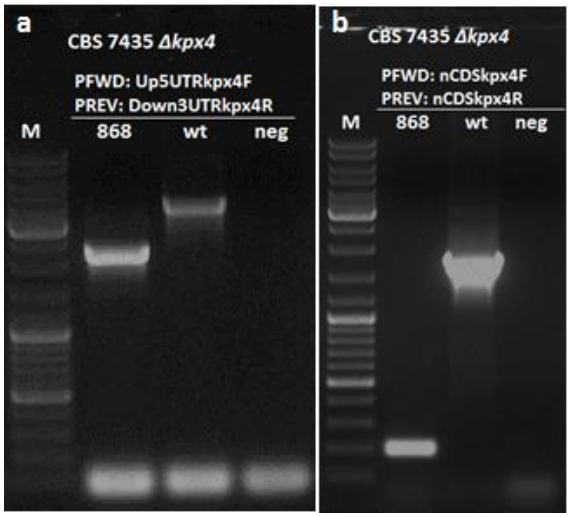


Figure 23: CBS 7435 *Δkpx4*

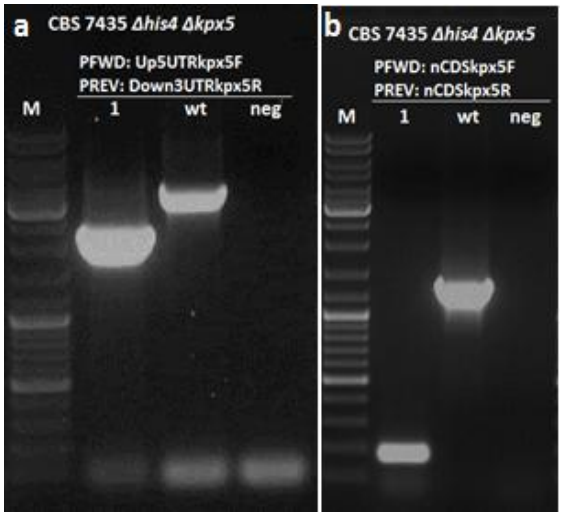


Figure 24: CBS 7435 *Δhis4 Δkpx5*

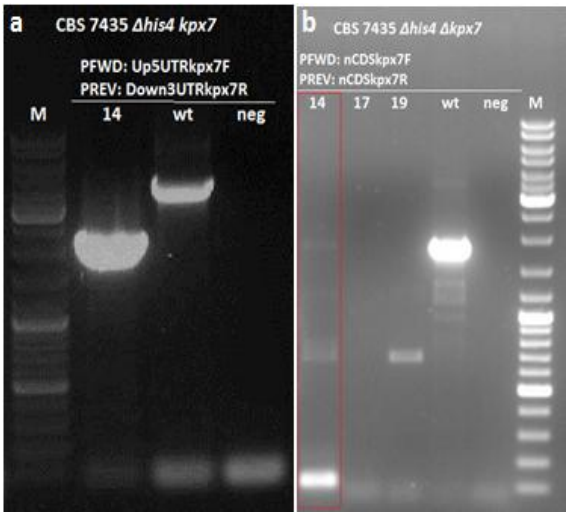


Figure 25: CBS 7435 *Δhis4 Δkpx7*

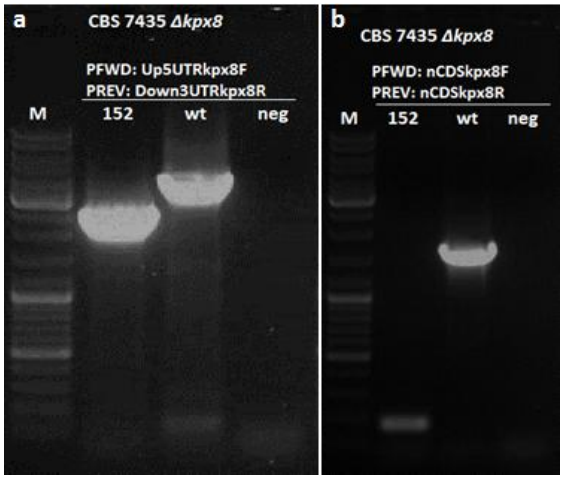


Figure 26: CBS 7435 *Δkpx8*

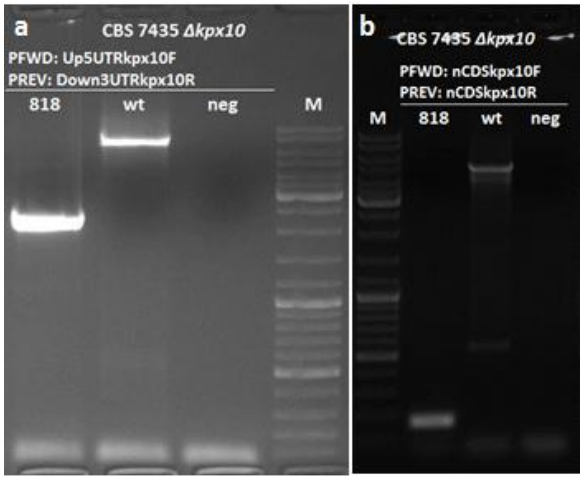


Figure 27: CBS 7435 *Δkpx10*

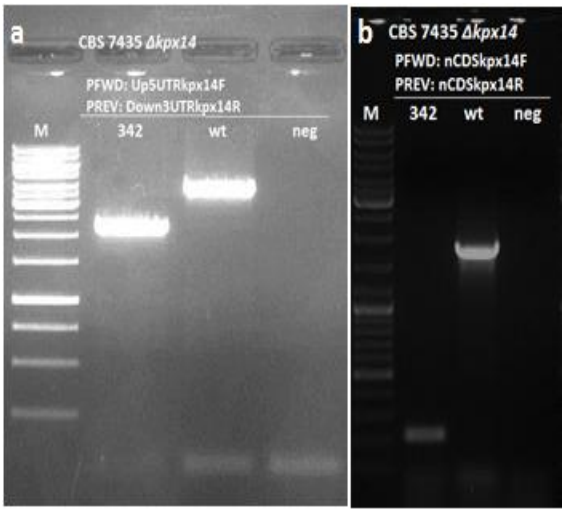


Figure 28: CBS 7435 *Δkpx14*

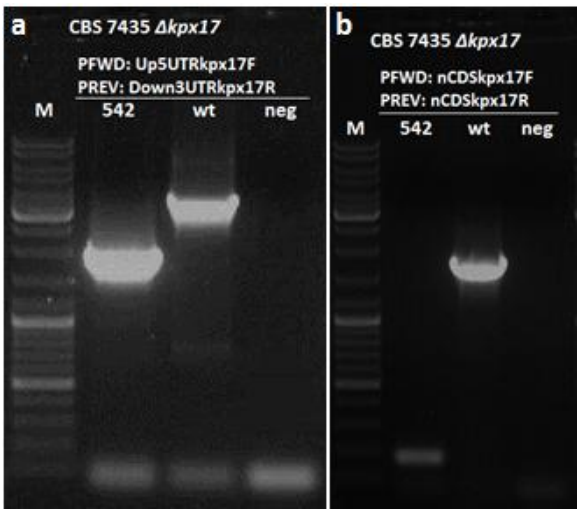


Figure 29: CBS 7435 *Δkpx17*

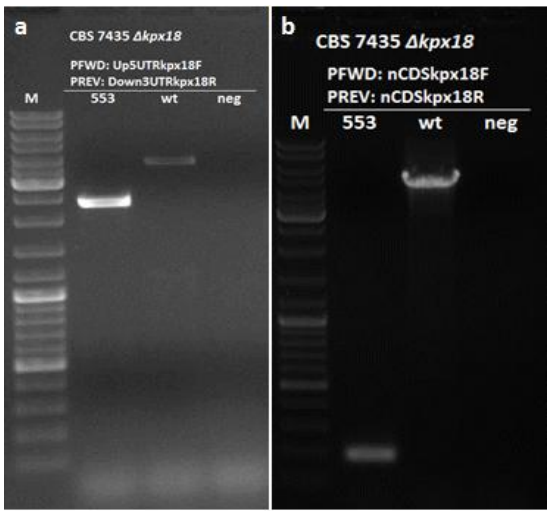


Figure 30: CBS 7435 *Δkpx18*

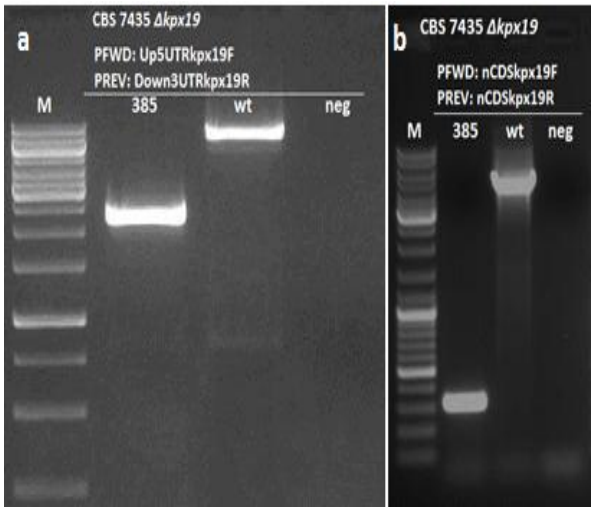


Figure 31: CBS 7435 *Δkpx19*

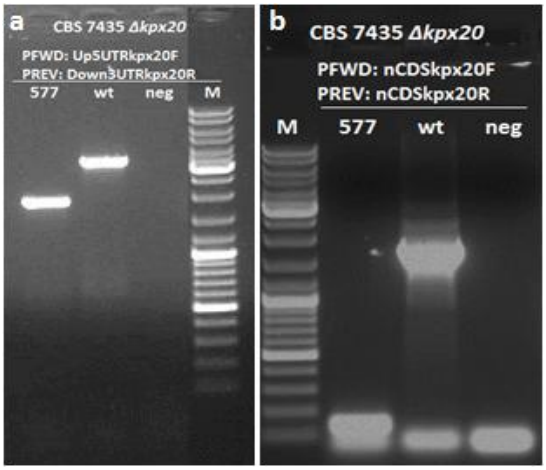


Figure 32: CBS 7435 *Δkpx20*

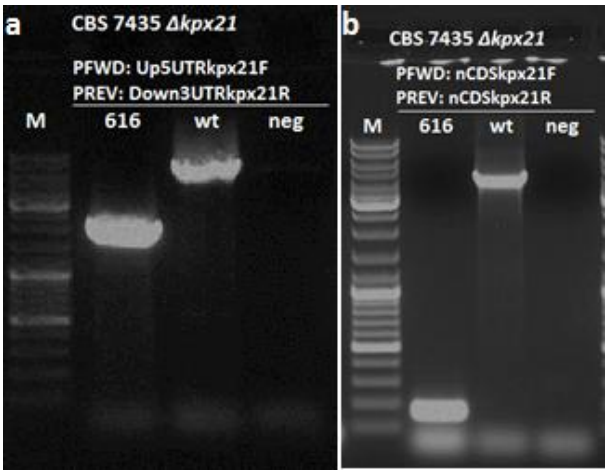


Figure 33: CBS 7435 *Δkpx21*

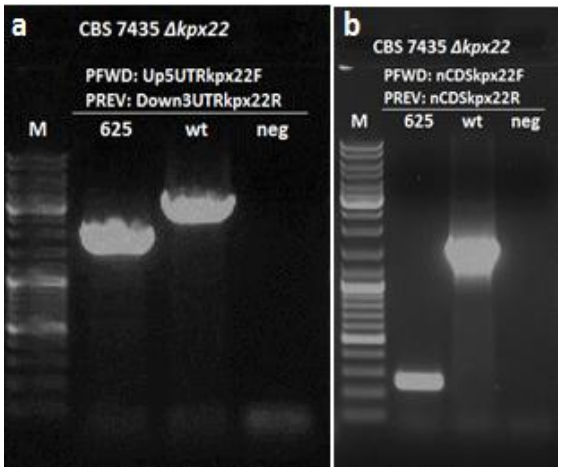


Figure 34: *CBS 7435 Δkpx22*

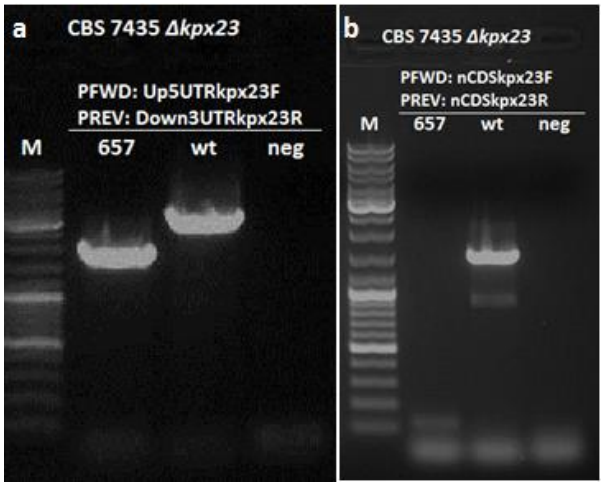


Figure 35: CBS 7435 *Δkpx23*

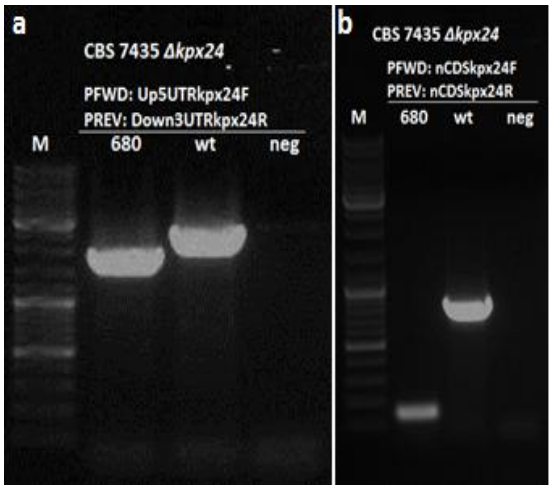


Figure 36: CBS 7435 *Δkpx24*

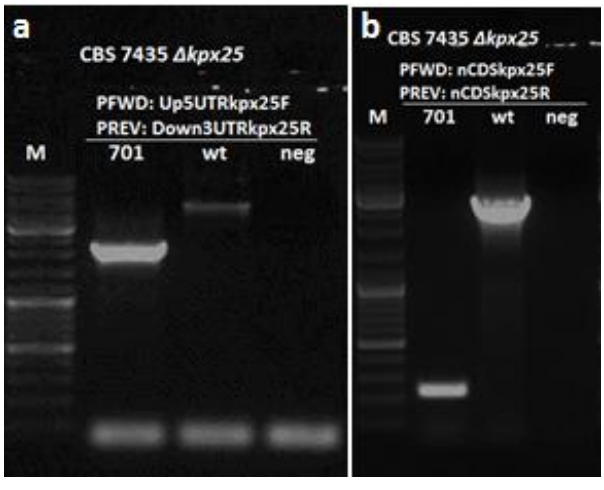


Figure 37: CBS 7435 *Δkpx25*

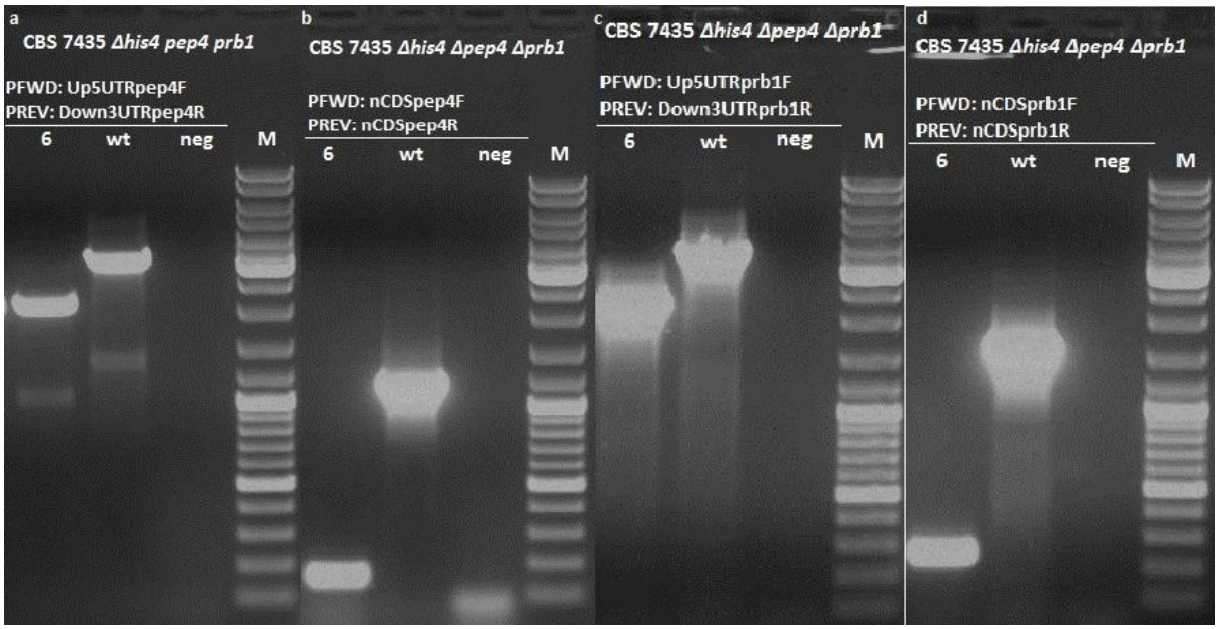


Figure 38: CBS 7435 *Δhis4 Δpep4 Δprb1*. Double knockout was performed by transforming *prb1* knockout cassette into *pep4* deletion strain followed by marker recycling as described in Material and methods.

4.3 Growth characteristics of protease-deficient *P. pastoris* strains

The overall health of a genetically modified yeast strain can significantly affect the yield of both secreted and intracellular heterologous proteins [47]. We therefore evaluated cell growth and health to receive information about cell viability and possible defects due to the knockouts in the genome.

4.3.1 Growth curves and dry cell biomass

A *Pichia pastoris* population possesses an interval for doubling of ~2 h in YPD medium under favorable conditions in the laboratory, as long as there are no limiting factors such as supply of glucose and oxygen.

The ability of the generated protease-deficient *P. pastoris* strains to grow effectively in liquid YPD medium was determined by analysis of growth kinetics and measurement of dry cell biomass produced over 72 h.

Growth curves for protease-deficient strains are compared to the CBS 7435 wild type strain as shown in Figure 39 – Figure 41. All knockout strains showed a lag time comparable to the wild type strain and in the exponential phase no significant aberrations were observed. After about 30 h of incubation cells passed into the stationary phase and reached a maximal OD_{600} of about 60±10 OD_{600}-units at the end of the measurement after 72 h.

Cereghino and Cregg [2] reported lower viability and lower specific growth rates for protease-deficient strains with respect to *PEP4*. We did not find any evidence to support this view. We also could not observe a ~50% increased generation time for *kex2* knockout strain as Werten

et al. [29] reported it for *Pichia pastoris*, nor did we notice obvious aberrant morphologies of cells grown in YPD as reported for *S. cerevisiae kex2* deletion strain [48]. In practical terms, all strains showed good viability. But what we observed is that overnight cultures started with a single colony of the *kex2*-deficient strain grew very slowly at the beginning, but as soon as an OD_{600} of ~0.2 is reached, growth is almost the same compared to wild type, as shown in Figure 39.

Measurement of dry cell biomass produced over 24, 48 and 72 h is illustrated in Figure 42. No aberrant cell growth or cell viability of the protease-deficient strains compared to the CBS 7435 wild type strain (wt) could be observed and after 72 h of growth a dry cell weight of ~12±5 g/l was reached by all knockout strains as well as by the wild type strain. The partially relatively high differences between the dry cell weights after 24, 48 and 72 h might be explained by the standard deviations.

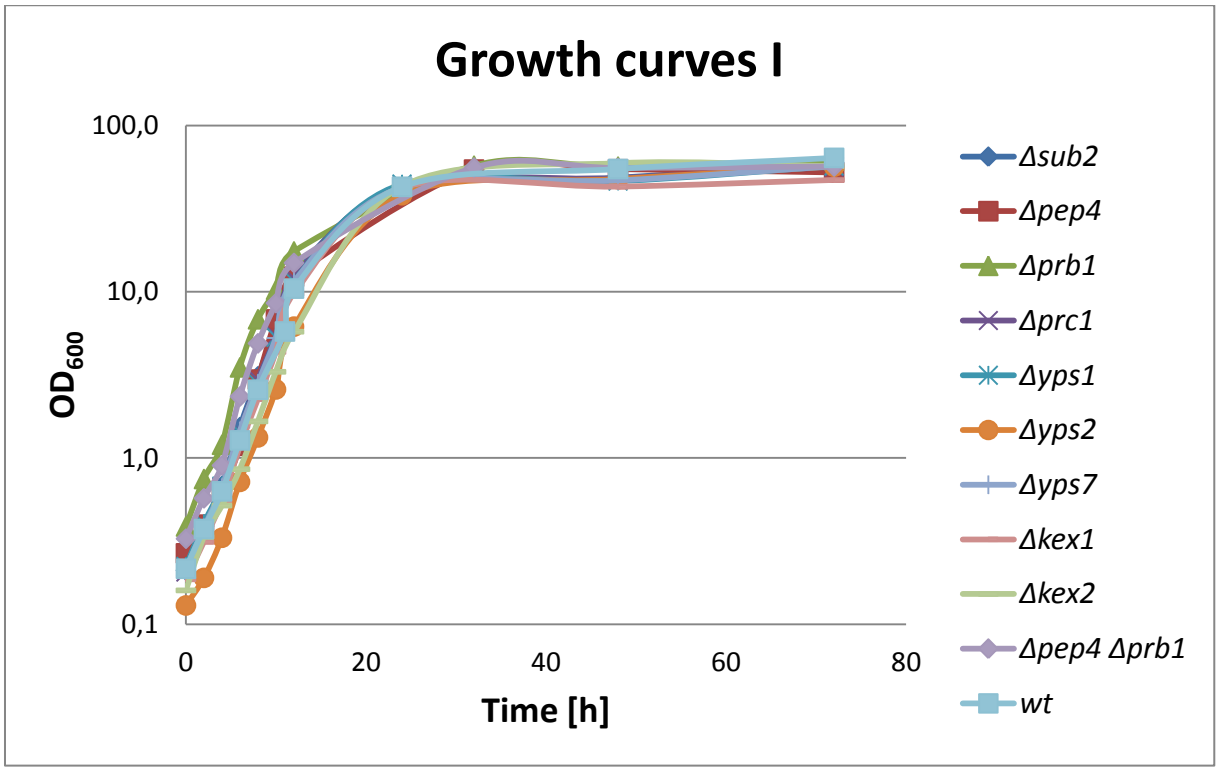


Figure 39: Growth curves of *sub2*, *pep4*, *prb1*, *prc1*, *yps1*, *yps2*, *yps7*, *kex1*, *kex2* and *pep4/prb1* knockout strains compared to the CBS 7435 wild type strain (wt).

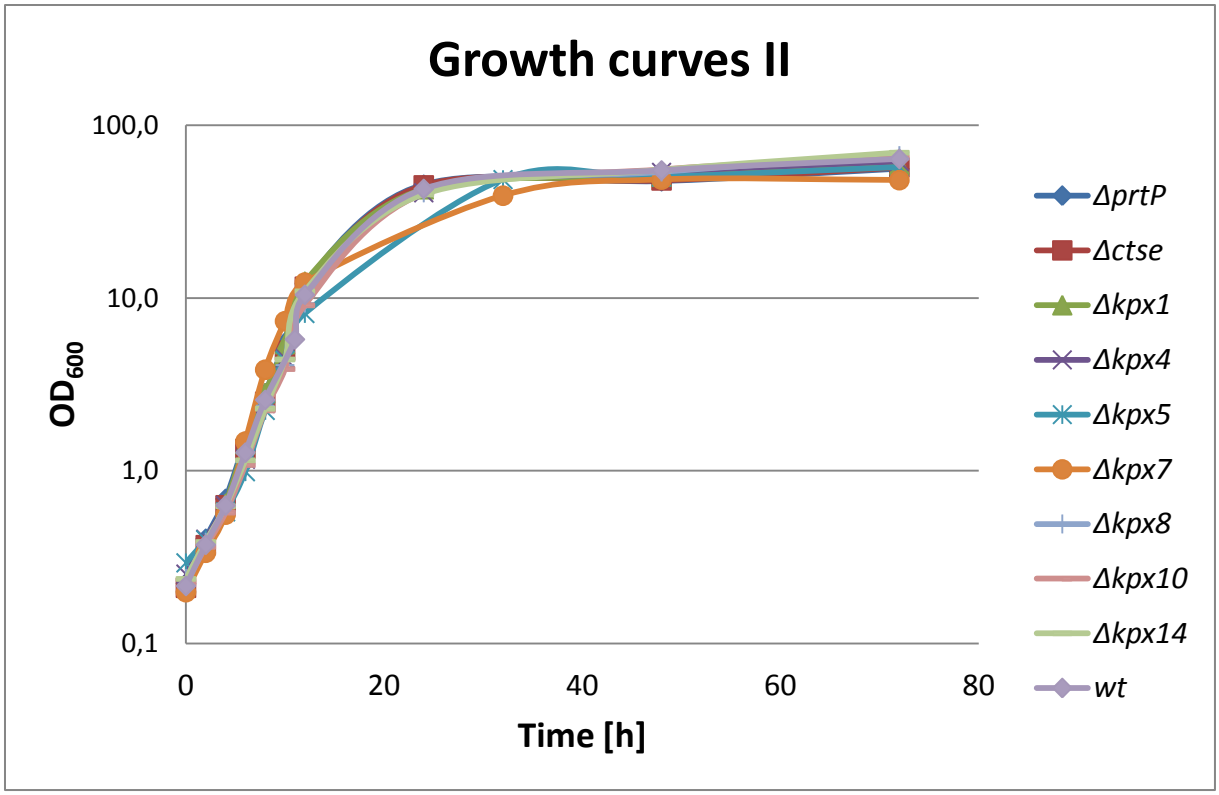


Figure 40: Growth curves of *prtP*, *ctse*, *kpx1*, *kpx4*, *kpx5*, *kpx7*, *kpx8*, *kpx10* and *kpx14* knockout strains compared to the CBS 7435 wild type strain (wt).

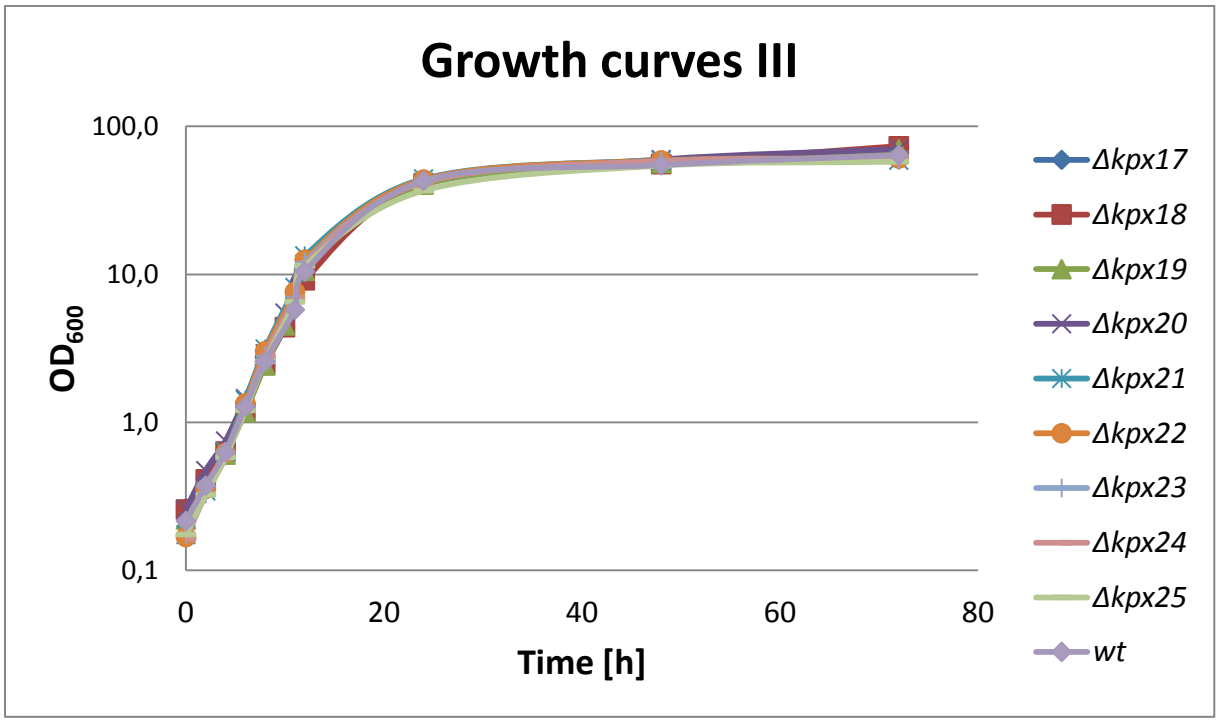


Figure 41: Growth curves of *kpx17*, *kpx18*, *kpx19*, *kpx20*, *kpx21*, *kpx22*, *kpx23*, *kpx24* and *kpx25* knockout strains compared to the CBS 7435 wild type strain (wt).

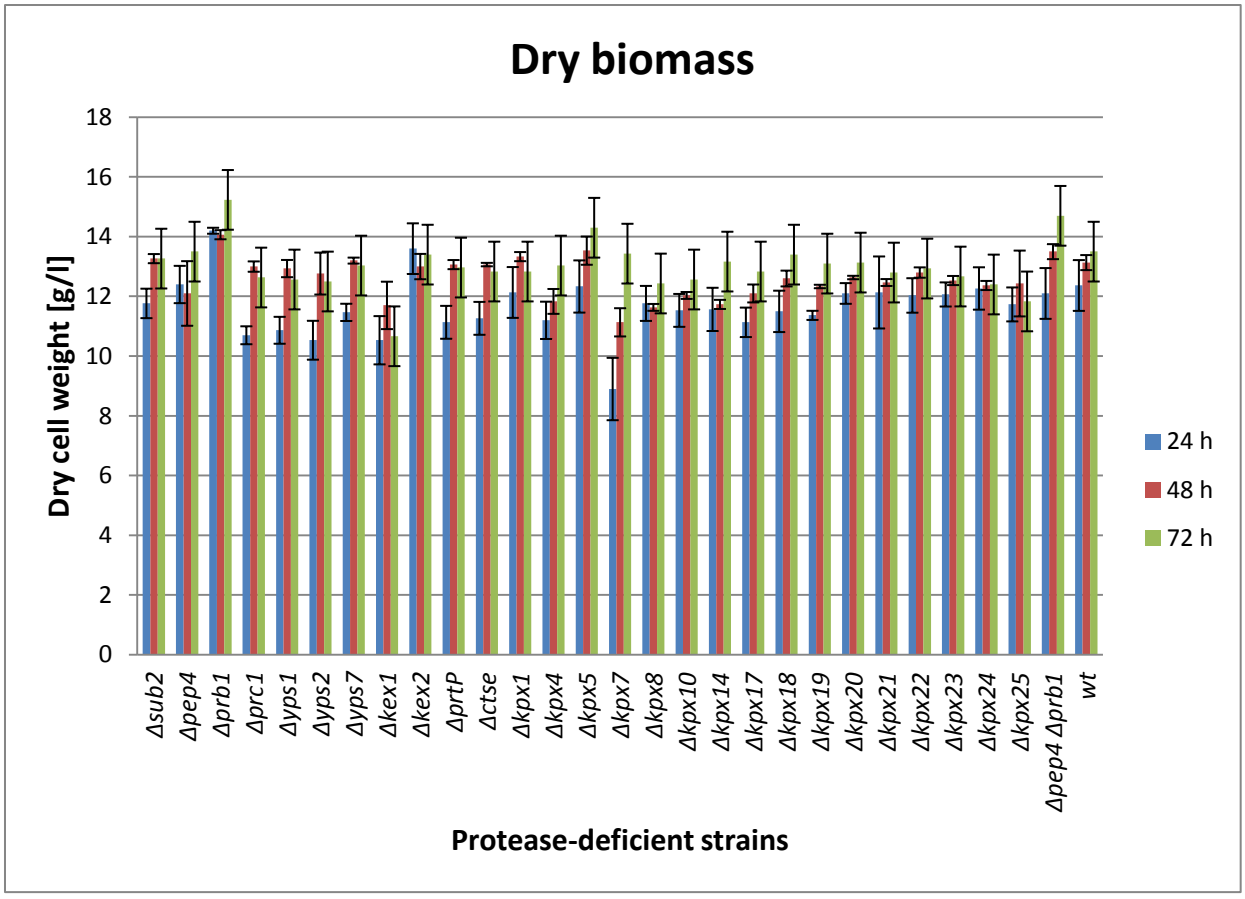


Figure 42: Dry biomass produced by each protease-deficient *P. pastoris* strain grown in liquid YPD medium after 24, 48 and 72 h. Each protease-deficient strain analyzed is indicated by its deleted protease name. wt = CBS 7435 wild type strain.

4.3.2 Growth on different media

To roughly determine any cell wall defects resulting from loss of protease function, cells of each deletion strain were grown on YPD agar supplemented with Zeocin™ (10 µg/ml) and the cell wall perturbing compounds SDS (0.025%), calcofluor white (CW, 10 µg/ml) and congo red (CR, 30 µg/ml) and cultured at 28°C for 2-3 days.

As the results show in Figure 43, every supplement (Zeocin™, SDS, CW and CR) affects growth of the protease-deficient strains as well as of CBS7435 wild type strain compared to the growth of the same strains on simple YPD agar. Unexpectedly, dilutions (10^{-1}-10^{-3}) of the respective cultures, all protease-deficients strains as well as wild type strain, do not show any growth on YPD supplemented with Zeocin™ and CW. The same is true for the growth in presence of CR in most cases, but interestingly 1:10 diluted cultures of some strains, such as *sub2, yps1, kpx7, kpx14, kpx17, kpx18, kpx24 and kpx25* knockout strains, show light growth. Compared to the wild type strain, this could lead to the assumption that the loss of the respective protease has a positive effect on cell wall integrity. But due to the fact that this

positive effect is only observed with CR and not with CW and SDS, this conclusion of increased cell wall resistance against CR cannot be seen as definitely consistent.

In contrast, *kex2* deletion strain exhibits decreased growth even on YPD agar compared to the wild type and all other protease-deficient strains. In presence of Zeocin™ this specific strain shows almost no growth and in presence of the cell wall perturbing reagent SDS completely growth arrest can be observed. It seems that growth on CW and CR is also diminished. Altogether, we can deduce from the obtained results in Figure 43 that only *kex2* deletion has certainly an impact on overall cell viability and leads to increased susceptibility to SDS and Zeocin™ and probably to CW and CR.

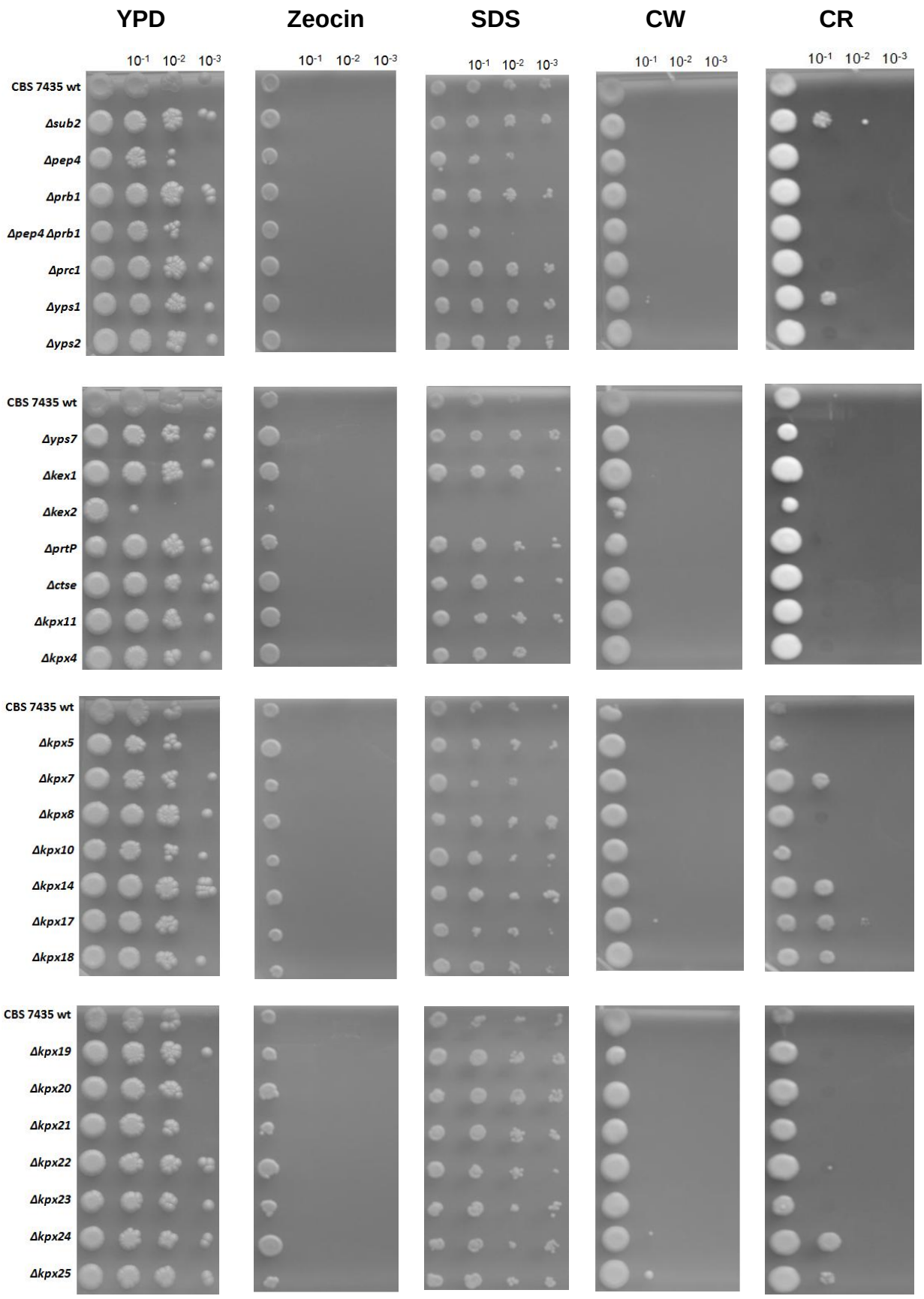


Figure 43: Growth characteristics of protease-deficient *P. pastoris* strains compared to CBS 7435 wt treated with Zeocin™ (10 µg/ml) and the cell wall perturbing compounds SDS (0.025%), calcofluor white (CW, 10 µg/ml) and congo red (CR, 30 µg/ml). Aliquots of each strain (approximately OD_{600} of 1) and dilution (10^{-1} – 10^{-3}) were stamped onto appropriate media and cultured at 28°C for 2-3 days.

4.4 Western blot analysis of secreted Δ27ST6

The enzyme used in this work for expression in protease-deficient *P. pastoris* strains is a truncated version of the human α2, 6-sialyltransferase ST6Gal I. The first 27 amino acid residues, containing the anchor sequence, have been removed and a N-terminal FLAG tag, a C-terminal HIS tag and a secretion signal (*S. cerevisiae* α-factor prepro peptide or levanase secretion signal) have been fused to enable immunodetection as well as secretion.

Expression of the sialyltransferase α-Δ27ST6 (α-factor prepro signal) in a *P. pastoris* KM71H strain led to high-quantity degradation of the secreted protein especially after 120 h of fermentation, as shown and described in Objectives (Figure 3). Based on these data received from Doris Ribitsch (ACIB GmbH, Graz) the protease-deficient *P. pastoris* strains generated during this project were analyzed, whether the deletion of certain proteases reduces proteolysis and leads to higher amounts of entire and active enzyme.

Surprisingly, western blot analysis revealed that secreted α-Δ27ST6 does not show any detectable degradation after 120 h of 0.5% methanol induction in BMGY medium, when expressed in *P. pastoris* CBS 7435 wild type strain, as shown in Figure 44 (lane 1 & lane 15). On the contrary, we could observe obviously increased degradation of α-Δ27ST6 expressed in some protease-deficient strains (Figure 44, lane 3-14 & lane 17-29), what constitutes a conflicting observation.

Expression of Δ27ST6 in *kex2* knockout strain needs further attention and explanations for several reasons. On the one hand the marker of the expression plasmid had to be changed to histidine (*HIS4*), because it is very sensitive to zeocin as mentioned above, and on the other hand Δ27ST6 was expressed with α-factor prepro signal (α-Δ27ST6) as well as with levanase secretion signal (LevSS-Δ27ST6) to illustrate the difference in protein processing. α-factor signal peptide has a Kex2 cleavage site and therefore α-Δ27ST6 should not be processed properly in absence of this protease. As shown in Figure 45 α-Δ27ST6 is visible as a ~25 kDa band, whereas LevSS-Δ27ST6 exhibits the same ~40 kDa band as apparent in all the other knockout strains. What exactly happens with α-Δ27ST6 during processing in *kex2* knockout strain is unknown and was not further examined.

However, a protein with a molecular weight of ~46 kDa is expected, but the bands visible on the western blots show a size of only ~40 kDa. The reason for that is the already known N-terminal cleavage of this protein, why immunodetection of the N-terminal Flag tag does not work as well. It is unidentified yet, which protease is involved in cleavage or if several proteases degrade the N-terminal part of Δ27ST6.

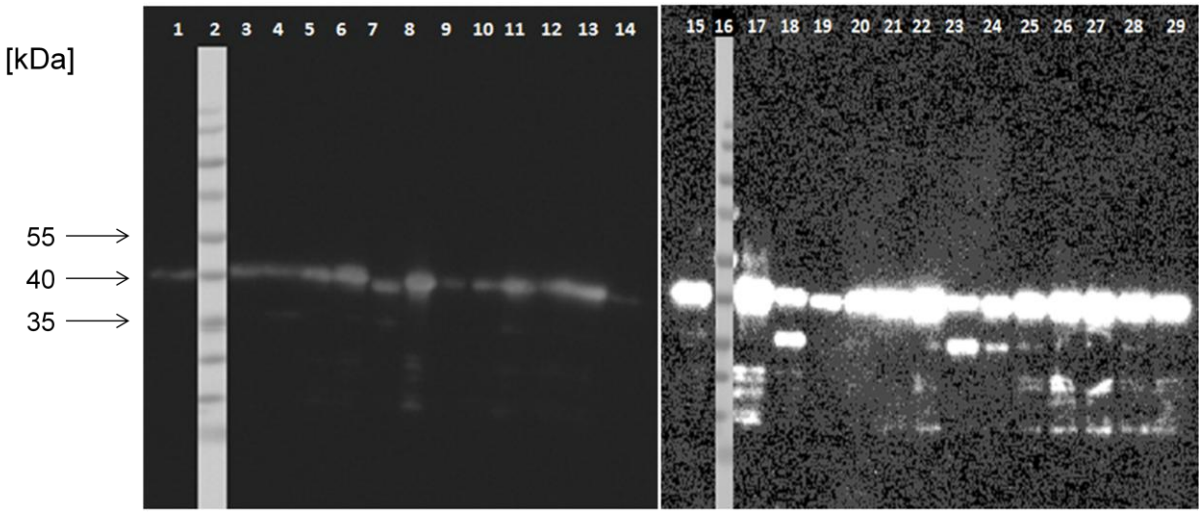


Figure 44: Western blot analysis of secreted Δ27ST6 from protease-deficient *P. pastoris* strains. After 120h of 0.5% methanol induction in BMGY medium, 15 µl of culture supernatant of each knockout strain were loaded onto a NuPage gel and subsequently blotted onto a nitrocellulose membrane. α-Δ27ST6 was detected using an anti-hST6 antibody. Secreted α-Δ27ST6 does not show any degradation when expressed in *P. pastoris* CBS 7435 wild type strain (1, 15). In contrast, degradation seems to be increased in some protease-deficient strains. No detection with anti-FLAG antibody. For explanation of sample numbers see Table 15.

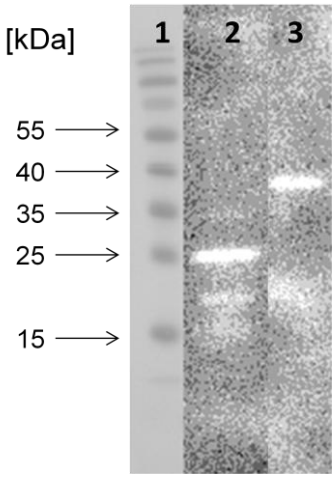


Figure 45: Western blot analysis of secreted Δ27ST6 from *kex2* deletion strain. After 120h of 0.5% methanol induction in BMGY medium, 200 µl of culture supernatant of *kex2* knockout strain were concentrated using methanol-chloroform precipitation and loaded onto a NuPage gel and subsequently blotted onto a nitrocellulose membrane.
1: PageRuler Prestained Protein Ladder
2: α-Δ27ST6 (α-prepro peptide)
3: LevSS-Δ27ST6 (Levanase secretion signal)

Table 15: Samples analyzed by western blot (Figure 44). Each protease-deficient strain analyzed for Δ27ST6 secretion is indicated by its deleted protease name. wt: wild type. Protein Ladder: PageRuler Prestained Protein Ladder (Figure 46).

#	*P. pastoris* strain	#	*P. pastoris* strain
1	wt	15	wt
2	Protein Ladder	16	Protein Ladder
3	*Δsub2*	17	*Δkpx7*
4	*Δpep4*	18	*Δkpx8*
5	*Δprb1*	19	*Δkpx10*
6	*Δyps1*	20	*Δkpx14*
7	*Δyps2*	21	*Δkpx17*
8	*Δyps7*	22	*Δkpx18*
9	*Δkex1*	23	*Δkpx19*
10	*ΔprtP*	24	*Δkpx20*
11	*Δctse*	25	*Δkpx21*
12	*Δkpx1*	26	*Δkpx22*
13	*Δkpx4*	27	*Δkpx23*
14	*Δkpx5*	28	*Δkpx24*
		29	*Δkpx25*

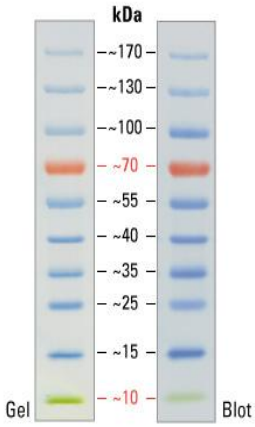


Figure 46: PageRuler Prestained Protein Ladder (taken from [50]**).**

5 DISCUSSION

The goal of this Master´s Thesis project was the development of protease-deficient *Pichia pastoris* host strains useful for production of protease-sensitive heterologous proteins. As a result of the gene deletion and confirmation procedures, 27 out of the 35 selected protease genes were each successfully deleted in *P. pastoris* and a *pep4/prb1* double knockout strain was generated as well (Table 13).

In some cases we were not able to knockout certain proteases. We assumed that either the respective protease is essential for viability of the cell or Zeocin™ as a very effective antibiotic, that shows strong toxicity against bacteria as well as fungi (including yeast) by binding DNA and cleaving it [51], puts too much pressure on the cells. Especially the combination of both, potential lower viability due to a knockout and Zeocin™ present in the medium, might force the effect that cells with a potential successful knockout die, although they possess a Zeocin™ resistance gene. To eliminate the problem with this antibiotic, we decided to change the selection marker to histidine (*HIS4*) for certain problem proteases and were finally able to generate five more protease-deficient strains in CBS7435 *Δhis4* strain *(Δpep4, Δpep4/Δprb1, Δkex2, Δkpx5* and *Δkpx7)*, so that at the end the most important protease genes, which´s deletions had already been reported by different workgroups (*Δpep4* [2][31]*, Δprb1* [2]*, Δpep4/Δprb1* [2]*, Δyps1* [30][32]*, Δyps7* [32]*, Δkex1* [27][28] *and Δkex2* [29]) were successfully knocked out.

Especially the *kex2*-deficient *P. pastoris* strain turned out to be very sensitive to Zeocin™ as well as to the cell wall perturbing compound SDS. This specific knockout strain shows decreased growth even on YPD agar compared to CBS7435 wild type strain as well as to all other protease-deficient strains. The strong susceptibility to even very low concentrations of Zeocin™ (10 µg/ml) was expected for *kex2* knockout strain, as on the one hand generating this particular knockout was the most difficult one and possible only with *HIS4* as selection marker. On the other hand further transformations (i.e. to integrate an expression plasmid in *kex2* deficient strain´s genome) never worked with *Zeocin^R* as selection marker. The observation that growth of *kex2* knockout strain is significantly decreased or even arrested in presence of the cell wall perturbing compounds SDS, CW and CR (Figure 43) might be due to the fundamental role of *Kex2* in processing proteins, which are possibly involved in cell wall composition.

Although a ~50% increased generation time for *kex2* knockout strain could not be observed as Werten et al. [29] reported it for *P. pastoris*, the viability of this protease-deficient strain is definitely influenced in a certain manner. We observed that overnight cultures started with a single colony of the *kex2*-deficient strain grew very slowly at the beginning, but as soon as an OD_{600} of ~0.2 is reached, growth is almost the same compared to wild type. The two *S. cerevisiae* GPI-anchored yapsins *Yps1* and *Yps2* [53] as well as *P. pastoris Yps1* [29] can

cleave the α-prepro signal and therefore are able to compensate partially for the loss of *Kex2*. This would explain why *kex2* knockout strain does not show obvious aberrant growth defects in YPD, especially after reaching a certain OD_{600}. Under stress conditions, such as transformation combined with additional pressure caused by Zeocin™ or other harmful substances, compensation by yapsins might not be sufficient and cells undergo cell death.

Concerning the deletions of GPI-anchored yapsins, which are thought to affect cell wall integrity, we neither could determine increased resistance of *yps7* knockout strain to SDS, CW and CR [32], nor were we able to assess higher susceptibility of *yps1*, *yps2* and *yps7* deletion strains to CW and CR [49]. For us cell viability and cell wall integrity was indistinguishable from CBS7435 wild type strain.

We are not sure yet, if deletions of *kpx2, kpx6, kpx9, kpx11, kpx12, kpx13* and *kpx15* are lethal to the cell and therefore generating knockout strains was not successful. Further attempts to delete these genes as well as *yps3*, which is not an essential gene definitely [32], were skipped, because to date there is no evidence that any of these knockouts have any influence on reducing degradation of expressed heterologous proteins.

Generally, the overall purpose of generating protease-deficient strains is to reduce proteolysis of heterologous expressed proteins and to achieve higher yields of intact and active product. We therefore investigated our set of knockout strains for reduced degradation of the sialyltransferase Δ27ST6 and found out that this particular protein is not degraded in the way, as Doris Ribitsch´s work group (ACIB GmbH, Graz) observed it with *P. pastoris* KM71H strain. We performed three independent fermentations with all knockout strains including wild type, but the results, showing high-quantity degradation in Figure 3, could never be produced, although the same conditions were chosen. CBS7435 wild type strain never showed visible degradation products in contrast to the protease-deficient strains (Figure 44, Figure 45), what is a conflicting observation anyway. Fact is that Δ27ST6 is somehow cleaved at the N-terminal site, but the protease-deficient strains tested did not prevent this cleavage event. We assumed that there might be involved several proteases at different stages of protein processing, so that a combination of multiple knockouts could be helpful in further experiments.

One advantage for analysis of Δ27ST6 degradation would also be to use an enzyme activity assay, because this method usually allows determining the enzyme activity differences between diverse strains relatively easy. But due to the fact that we did not have access to a sialyltransferase activity assay as described in [34], western blot analysis was used.

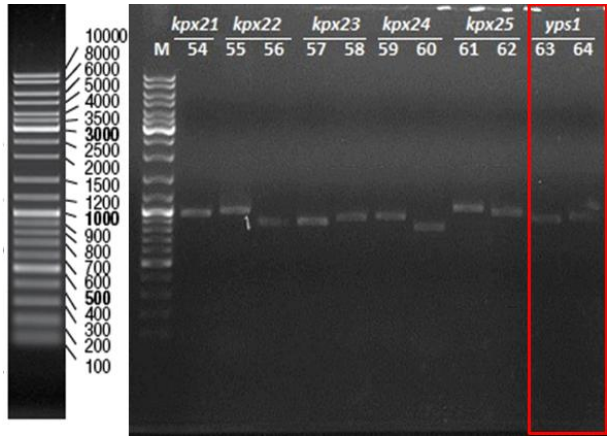
Summing up, during this Master´s Thesis project a set of protease-deficient *P. pastoris* strains has been generated, that can serve as a basis for further investigations. Although it was not possible to overcome the degradation problem with the sialyltransferase Δ27ST6,

the established set can be useful not only for practical application in protein production with regard to multiple gene deletions, but also for further examinations of proteolytic processes.

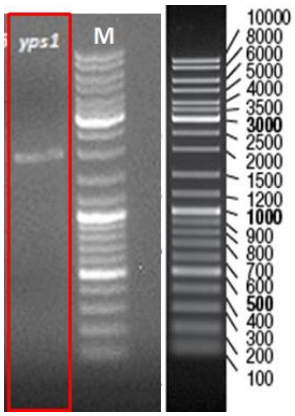
6 SUPPLEMENTARY INFORMATION

6.1 Construction of knockout plasmids

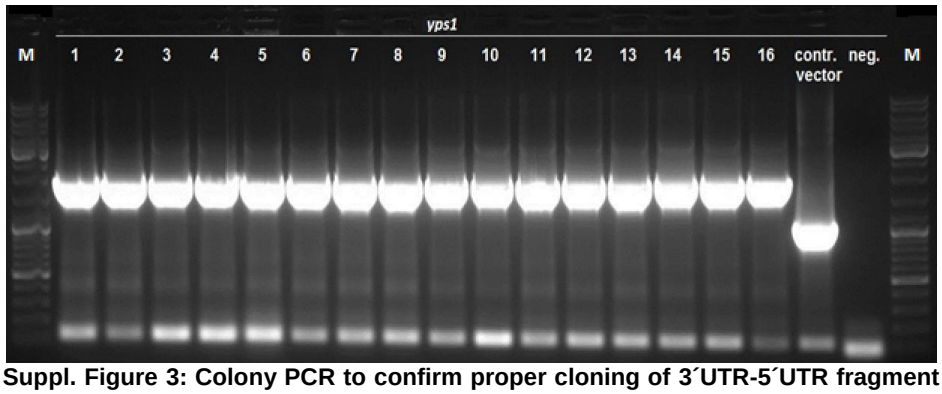
The procedure of constructing all 35 knockout cassettes for performing gene deletions in *P. pastoris* is shown by using the agarose gel pictures of *yps1*.



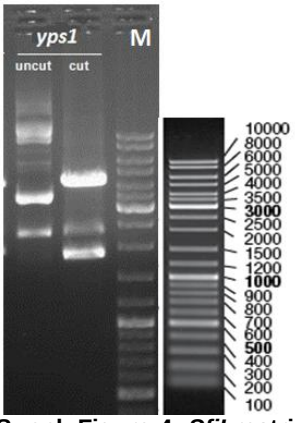
Suppl. Figure 1: Amplification of 3´UTR and 5´UTR of knockout plasmid pPpKC1yps1. By using the primers 3UTRyps1F & 3UTRyps1R for amplification of the 3´UTR and 5UTRyps1F & 5UTRyps1R for amplification of the 5´UTR proper DNA sequences could be generated from the *P. pastoris* CBS7435 genome.
M: O´GeneRuler DNA Ladder Mix.
63: PCR product 3´UTR (928 bp)
64: PCR product 5´UTR (964 bp)



Suppl. Figure 2: Overlap extension PCR product of 3´UTR and 5´UTR of *yps1*. The 3´UTR and 5´UTR PCR products were fused via overlap extension PCR by using the primers 3UTRyps1F & 5UTRyps1R, which carry complementary regions of the flanking sequences of *YPS1*.
M: O´GeneRuler DNA Ladder Mix
yps1: oe-extension PCR product 3´UTR-5´UTR (1877 bp)
The oe-PCR product was restricted with the restriction enzyme *SfiI* to make the ends compatible to the pPpKC1 backbone.



Suppl. Figure 3: Colony PCR to confirm proper cloning of 3´UTR-5´UTR fragment into pPpKC1 backbone after transformation into electro competent *E. coli* TOP10 F´ cells. Primers PucSeqF and PAox1SeqR, which bind on the pPpKC1 backbone and each pointing towards the insert, were used for verification.
M: O´GeneRuler DNA Ladder Mix
1-16: 3´UTR-5´UTR fragment of positive transformants (2022 bp)
contr. vector: pPpKC1 (stuffer ~1 kb)
neg.: negative control (dH_2O as template)



Suppl. Figure 4: *SfiI*-restriction of pPpKC1yps1. The correct plasmid was isolated and checked via restriction analysis with *SfiI*. The image shows an uncut and cut plasmid, whereas the cut plasmid exhibits a fragment of the expected size 1877 bp (lowest band). For confirmation purposes all cloned knockout plasmids were sent for sequencing.
M: O´GeneRuler DNA Ladder Mix

6.2 Confirmation of site-specific integration of knockout cassette

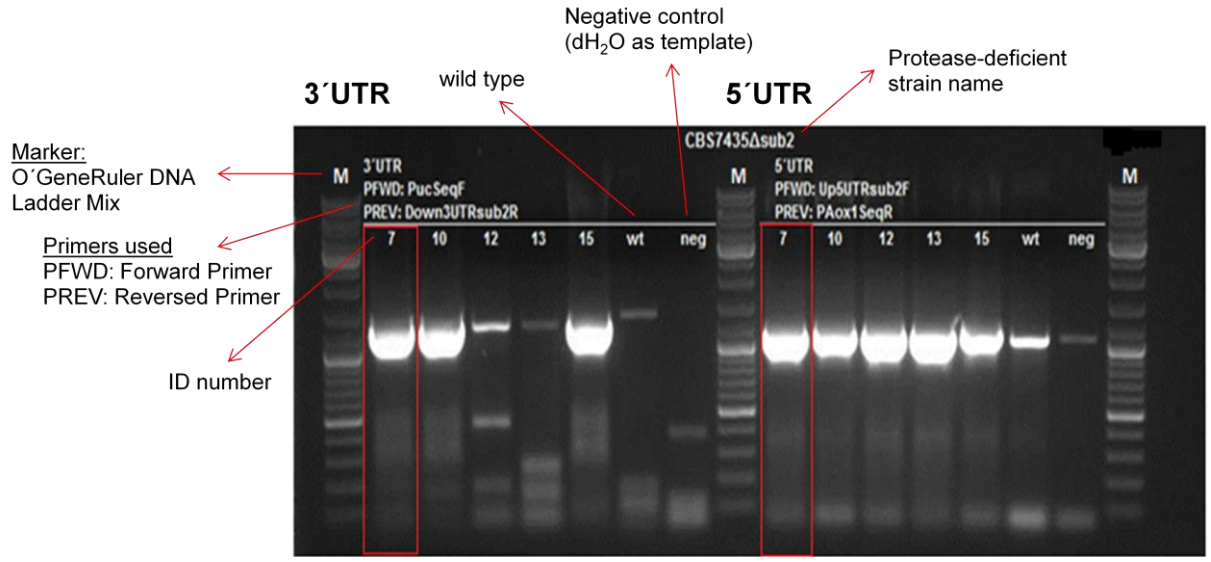
To verify site-specific integration of the knockout cassette, clones of each knockout transformation were screened for the 5´UTR and 3´UTR region. For amplifying the 5´UTR region the primers Up5UTRgeneF & PAox1SeqR were used and for amplifying the 3´UTR region the primers PucSeqF & Down3UTRgeneR were used (see also Figure 7).

The results of the 27 successfully disrupted proteases of *P. pastoris* are shown below (Suppl. Figure 5 -Suppl. Figure 32). The obtained agarose gel pictures after the two separate PCRs are explained by using the example of *sub2* disruption in Suppl. Figure 5. Positive clones encircled red were further used for marker recycling. The expected fragment lengths of the 5´UTR and 3´UTR regions obtained from the PCRs are listed in Suppl. Table 1.

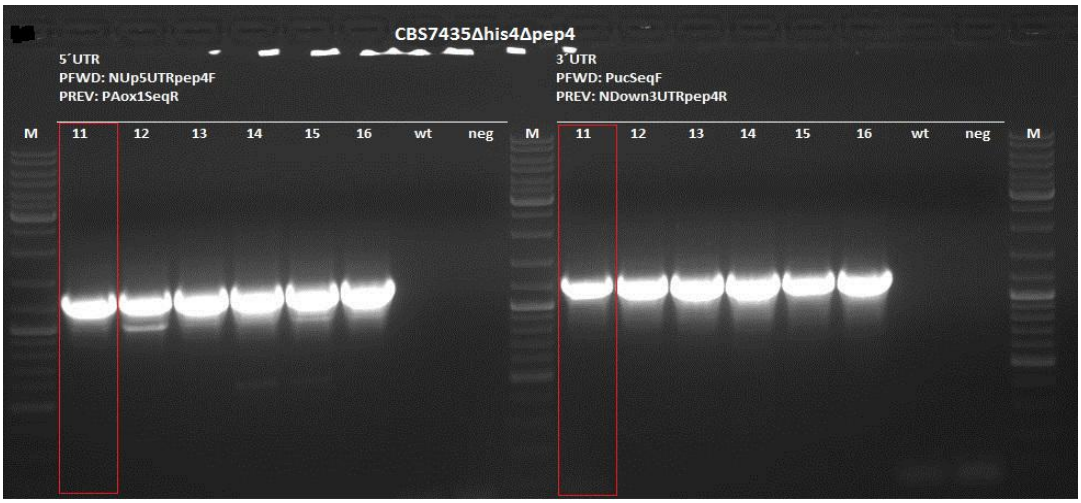
Suppl. Table 1: Expected fragment lengths of 5´UTR and 3´UTR regions after site-specific integration of the knockout cassette.

Knockout cassette integrated			
No.	**Gene**	**5´UTR [bp]**	**3´UTR [bp]**
		Up5UTRgeneF **PAox1SeqR**	**PucSeqF** **Down3UTRgeneR**
1	***sub2***	1130	1296
2	***pep4***	1287	1233
3	***prb1***	1191	1296
4	***prc1***	1289	1176
5	***yps1***	1107	1077
6	***yps2***	1234	1271
8	***yps7***	1056	1294
9	***kex1***	1241	1090
10	***kex2***	1114	1154
11	***prtP***	1084	1484
12	***ctse***	1311	1282
13	***kpx1***	1130	1140
15	***kpx4***	1193	1172
16	***kpx5***	1448	1293

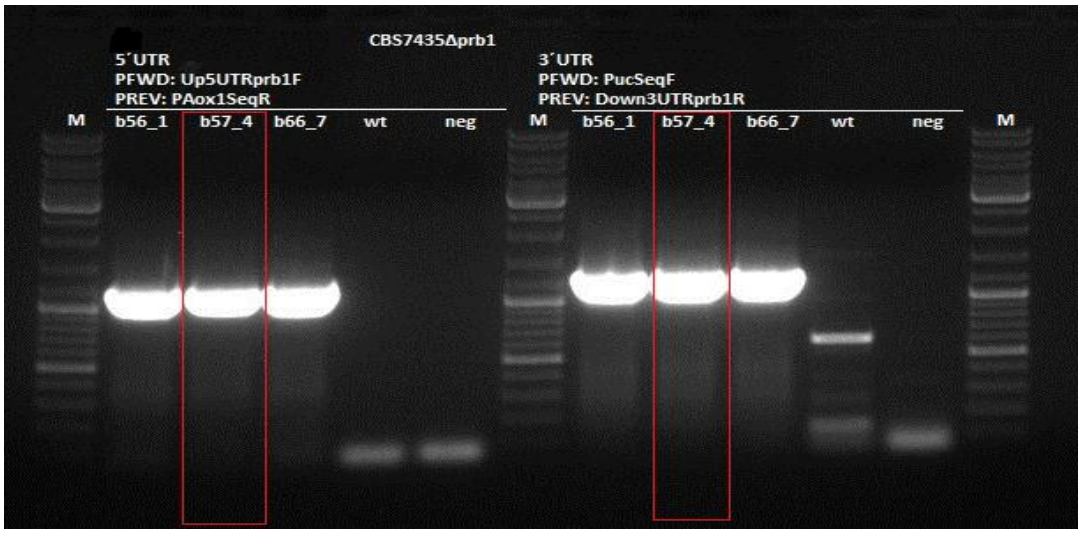
Knockout cassette integrated			
No.	**Gene**	**5´UTR [bp]**	**3´UTR [bp]**
		Up5UTRgeneF **PAox1SeqR**	**PucSeqF** **Down3UTRgeneR**
18	***kpx7***	1285	1382
19	***kpx8***	1203	1462
21	***kpx10***	1106	1101
25	***kpx14***	1081	1273
27	***kpx17***	1021	1042
28	***kpx18***	1248	1085
29	***kpx19***	1221	1146
30	***kpx20***	1114	1130
31	***kpx21***	1131	1109
32	***kpx22***	964	1125
33	***kpx23***	1031	991
34	***kpx24***	932	1092
35	***kpx25***	1183	1179



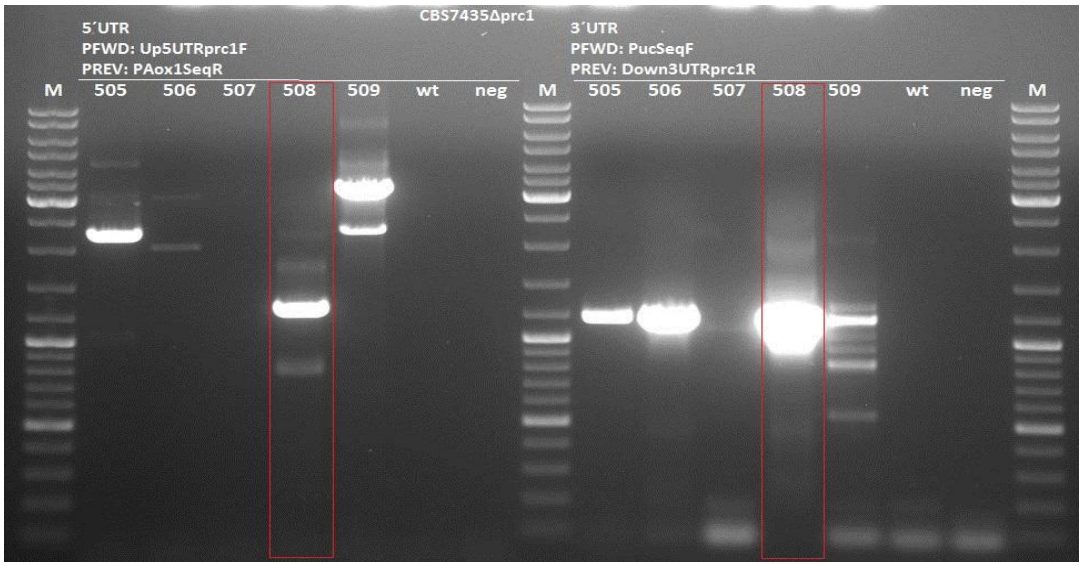
Suppl. Figure 5: Agarose gel picture for confirmation of site-specific integration of *sub2* knockout cassette in *P. pastoris* CBS 7435 genome. The clone framed red exhibited proper site-specific integration and was further used for marker recycling. Gel pictures for all 27 protease knockouts are labeled the same way.



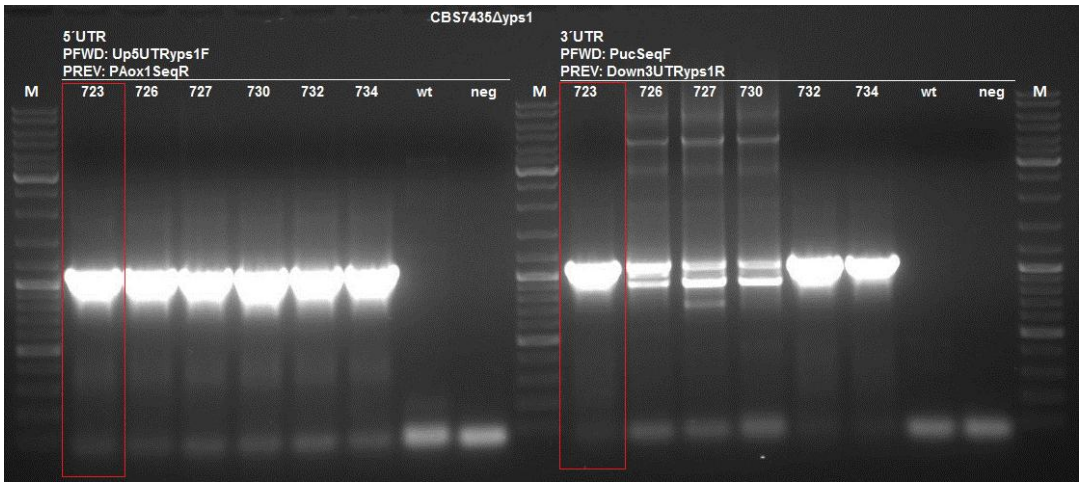
Suppl. Figure 6: CBS 7435 *Δhis4 Δpep4*



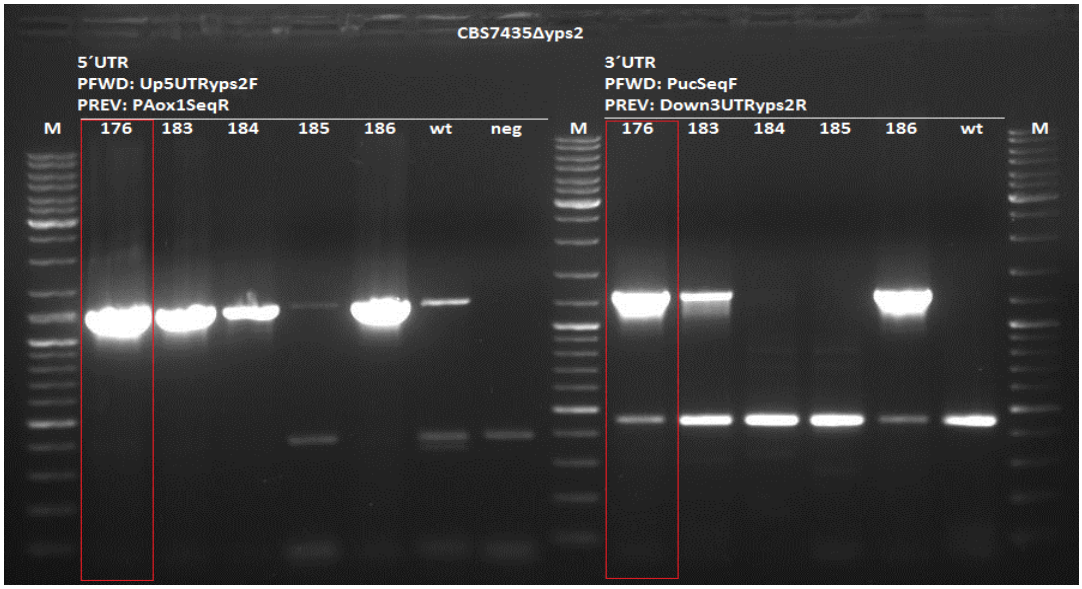
Suppl. Figure 7: CBS 7435 *Δprb1*



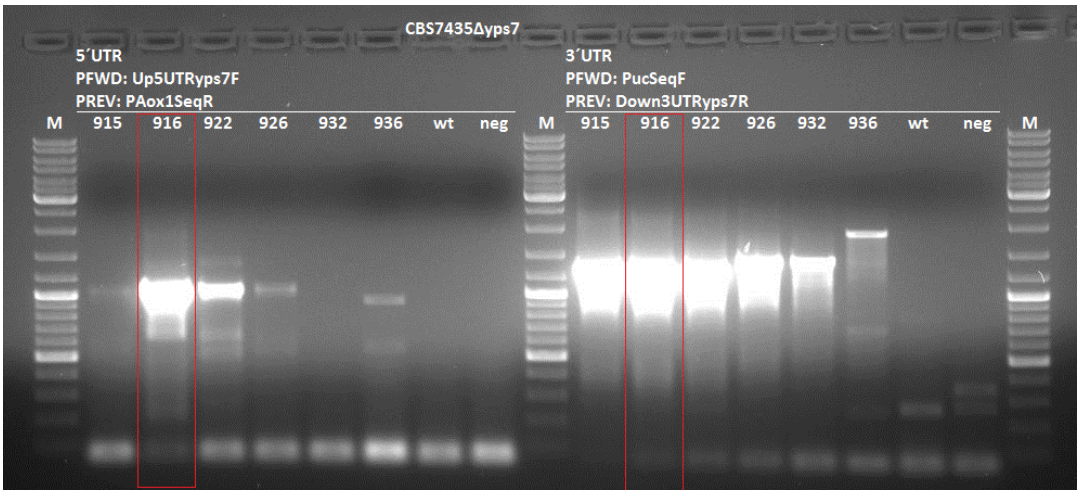
Suppl. Figure 8: CBS 7435 *Δprc1*



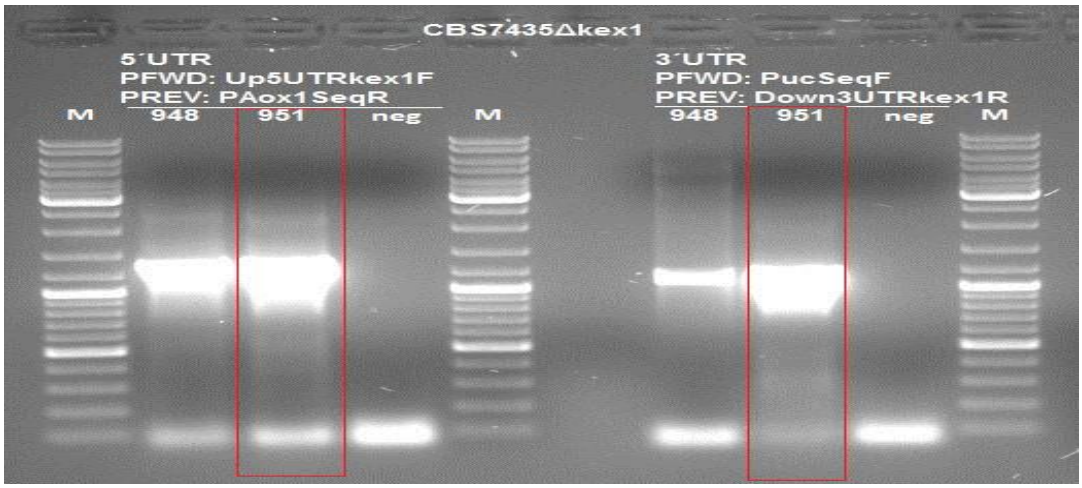
Suppl. Figure 9: CBS 7435 *Δyps1*



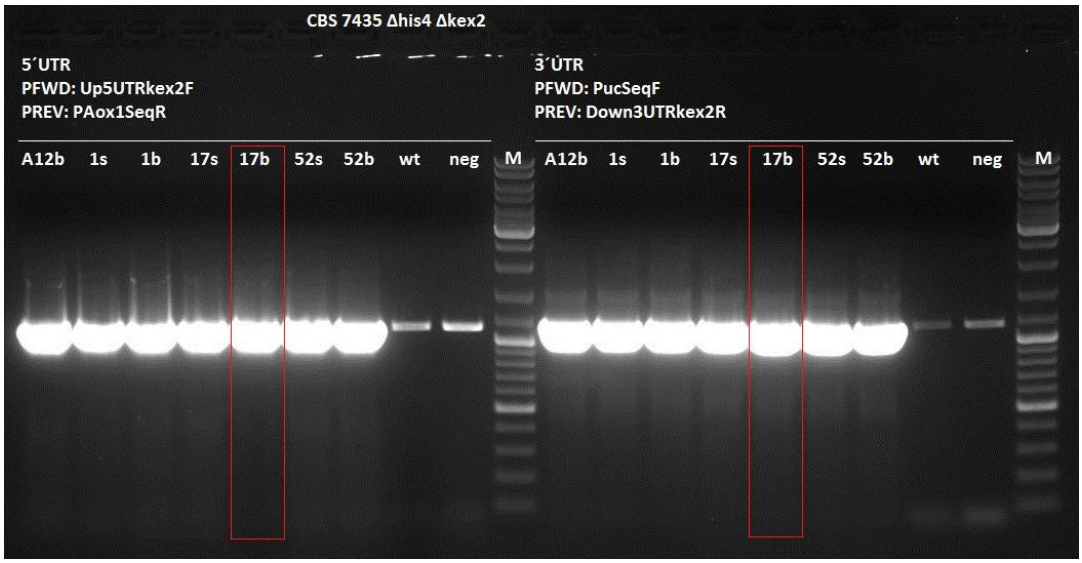
Suppl. Figure 10: CBS 7435 *Δyps2*



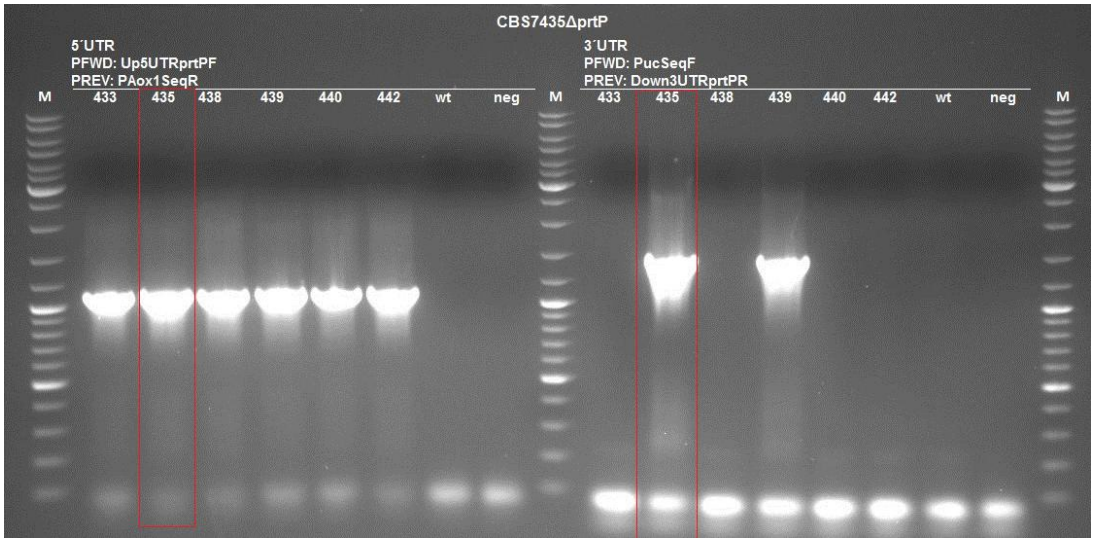
Suppl. Figure 11: CBS 7435 *Δyps7*



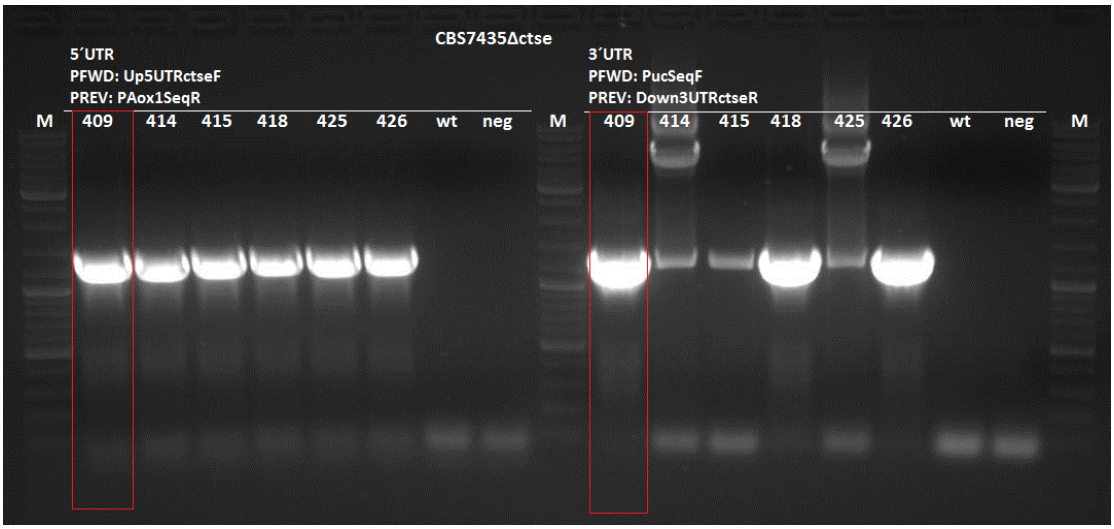
Suppl. Figure 12: CBS 7435 *Δkex1*



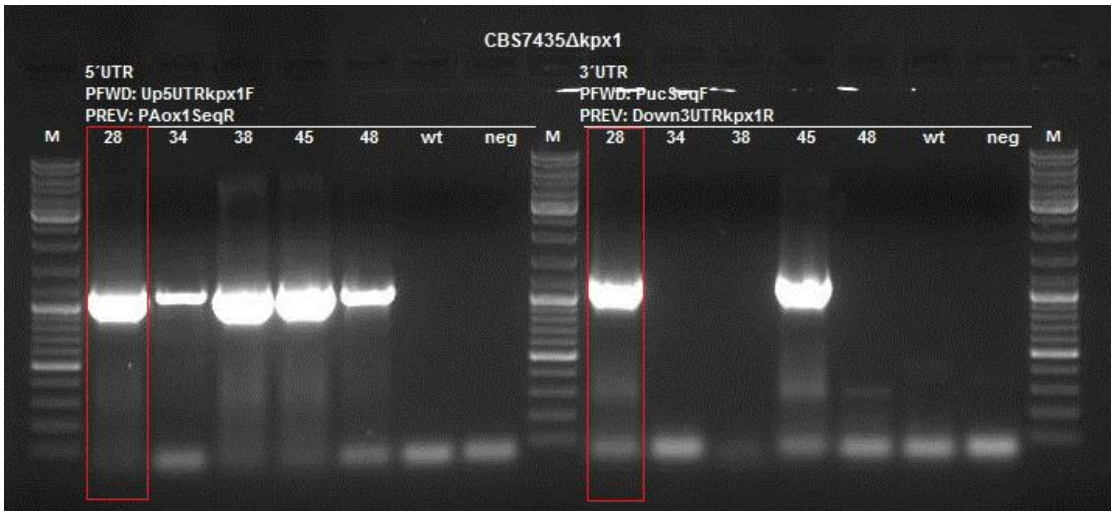
Suppl. Figure 13: CBS 7435 *Δhis4 Δkex2*



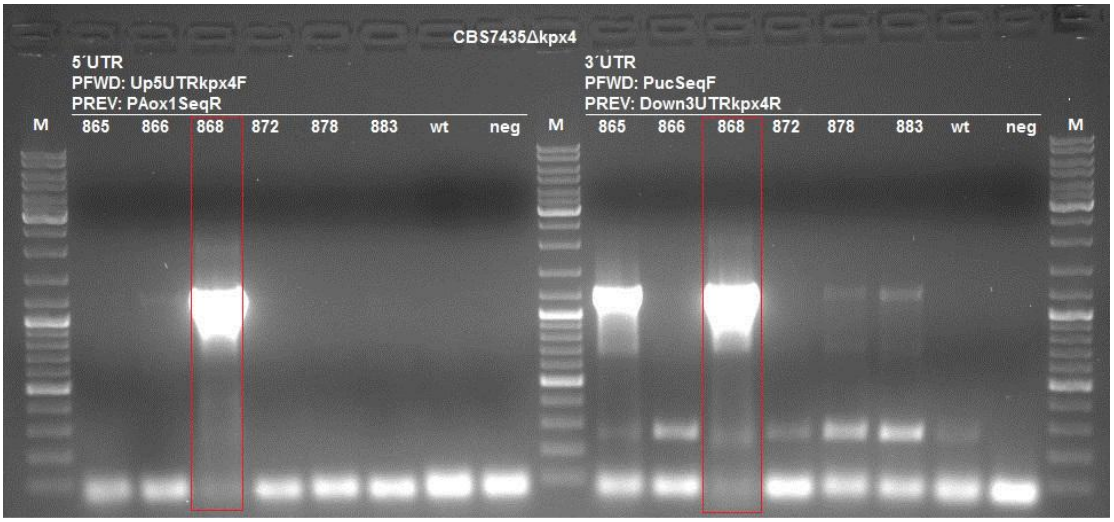
Suppl. Figure 14: CBS 7435 *ΔprtP*



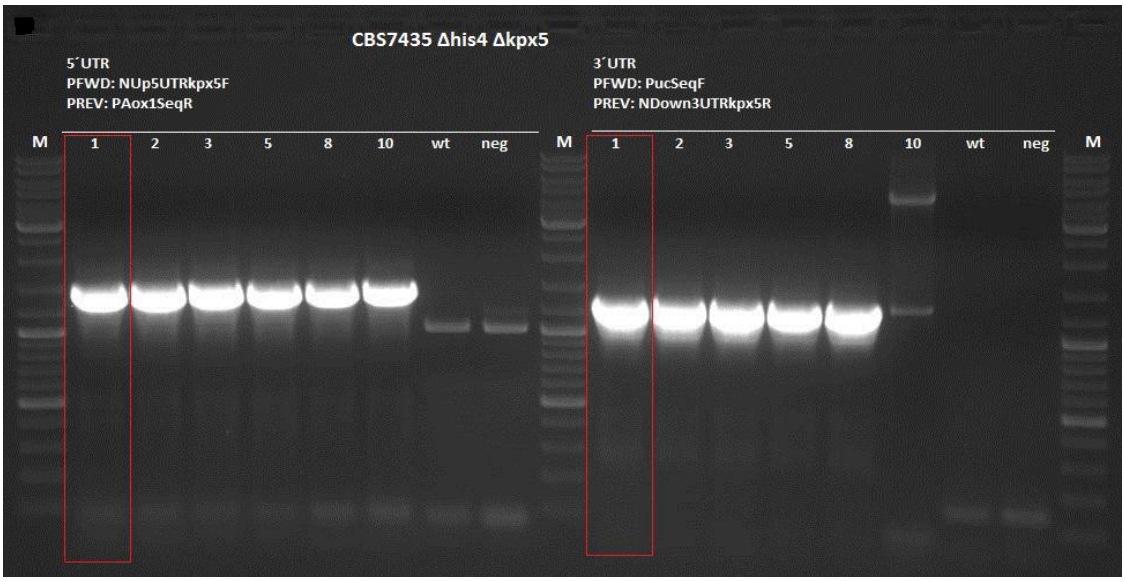
Suppl. Figure 15: CBS 7435 *Δctse*



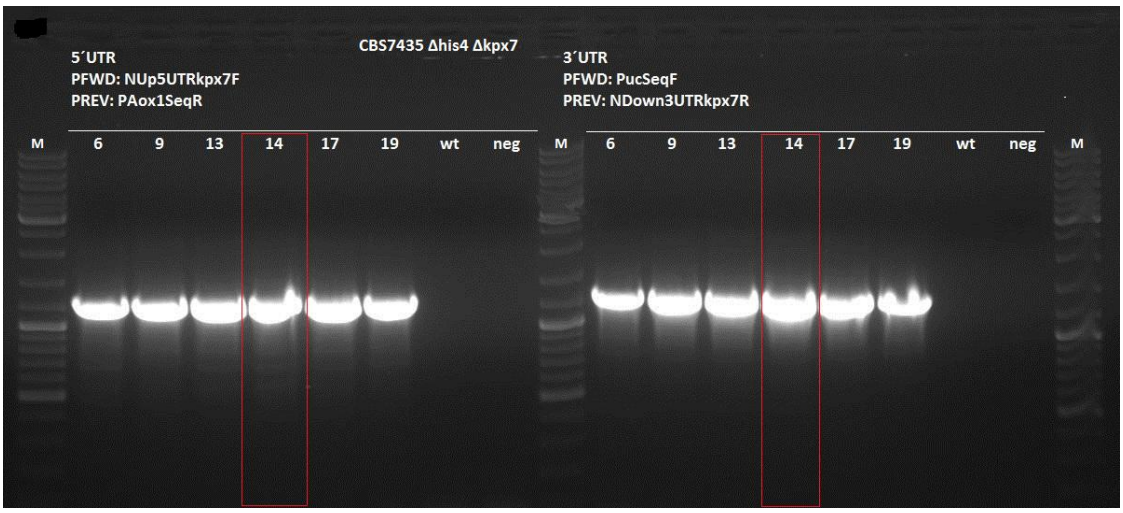
Suppl. Figure 16: CBS 7435 *Δkpx1*



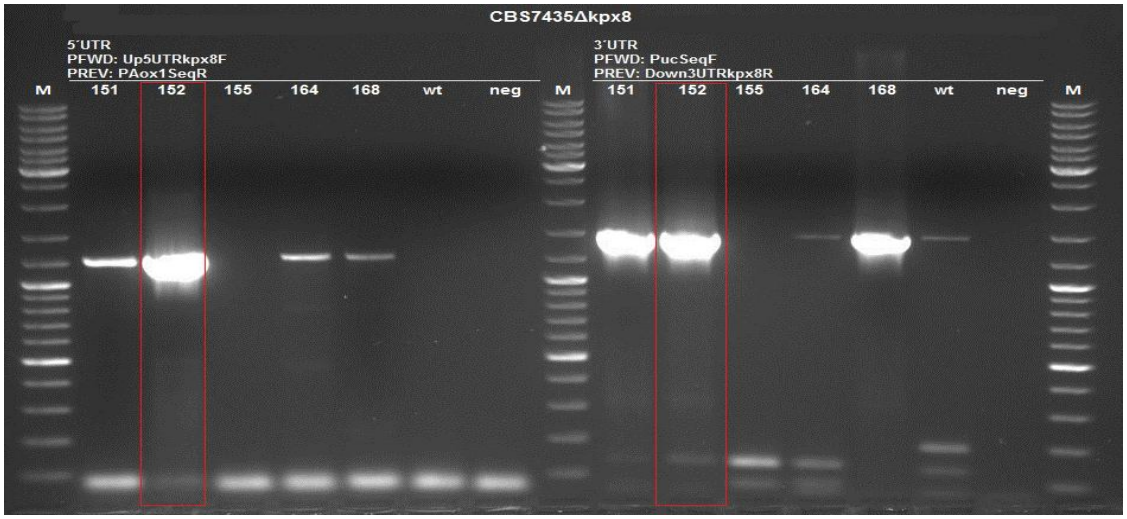
Suppl. Figure 17: CBS 7435 *Δkpx4*



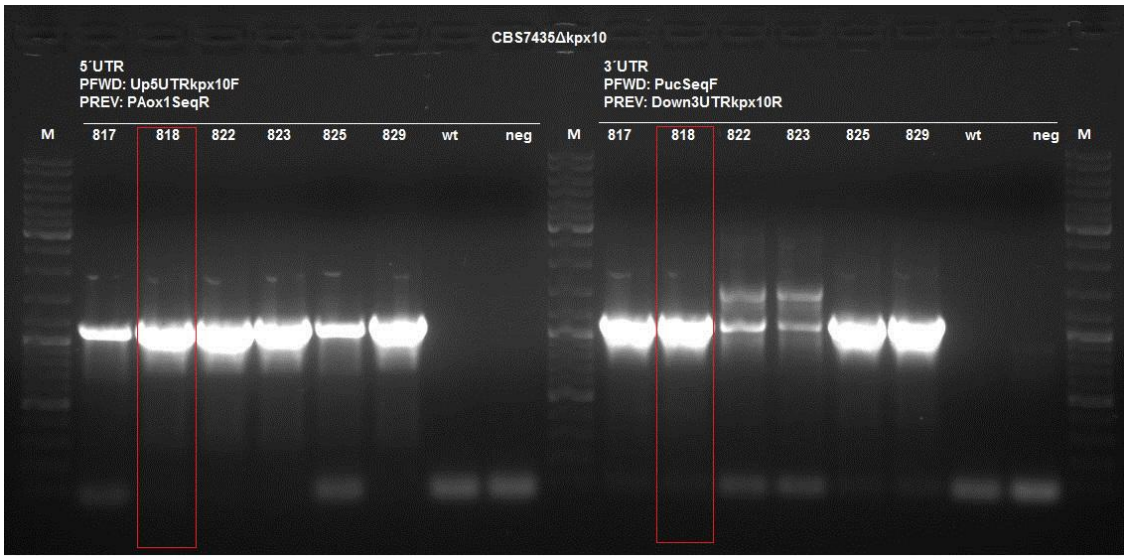
Suppl. Figure 18: CBS 7435 *Δhis4 Δkpx5*



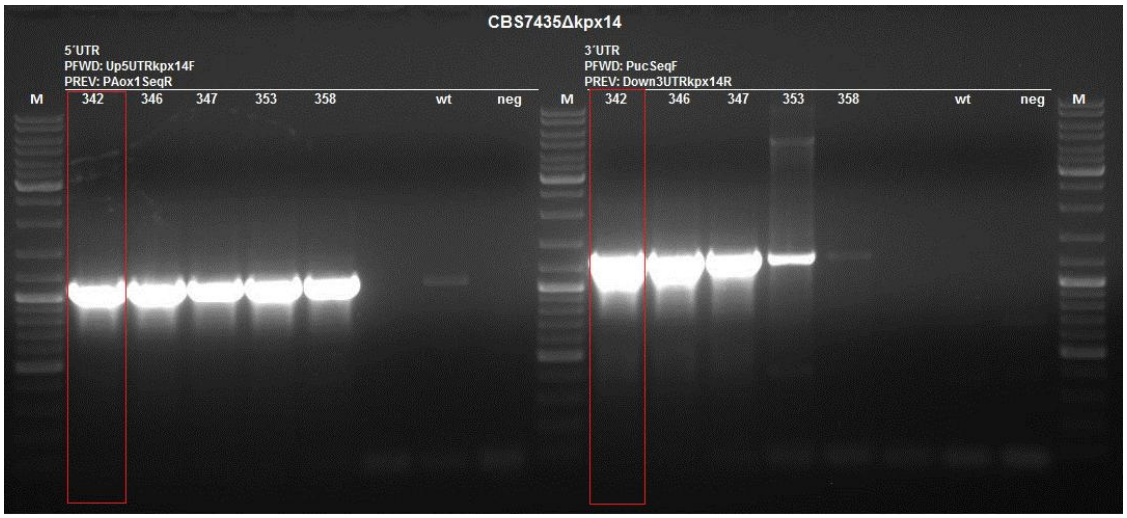
Suppl. Figure 19: CBS 7435 *Δhis4 Δkpx7*



Suppl. Figure 20: CBS 7435 *Δkpx8*

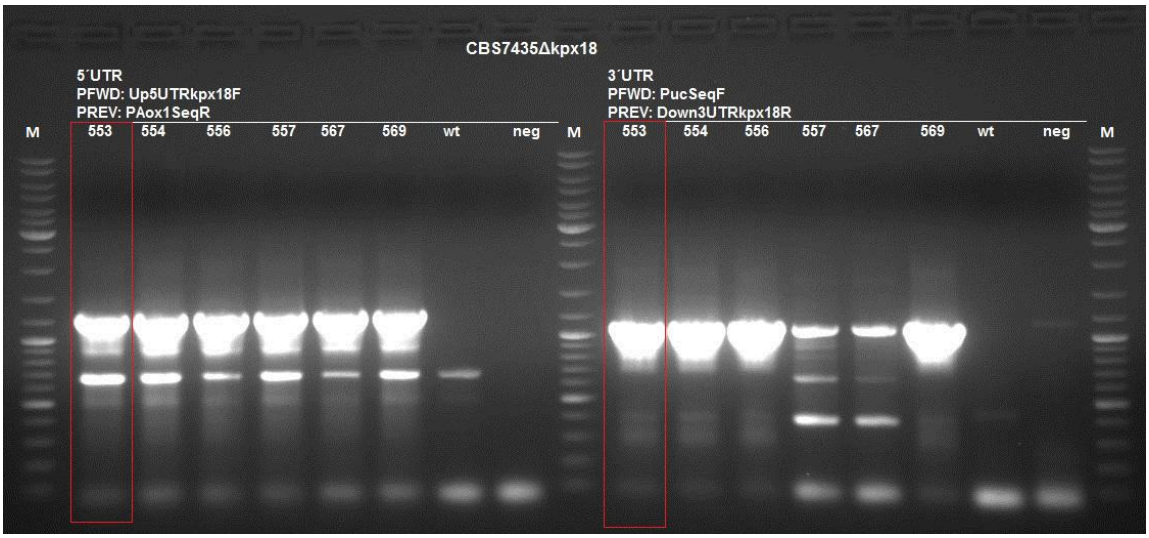


Suppl. Figure 21: CBS 7435 *Δkpx10*

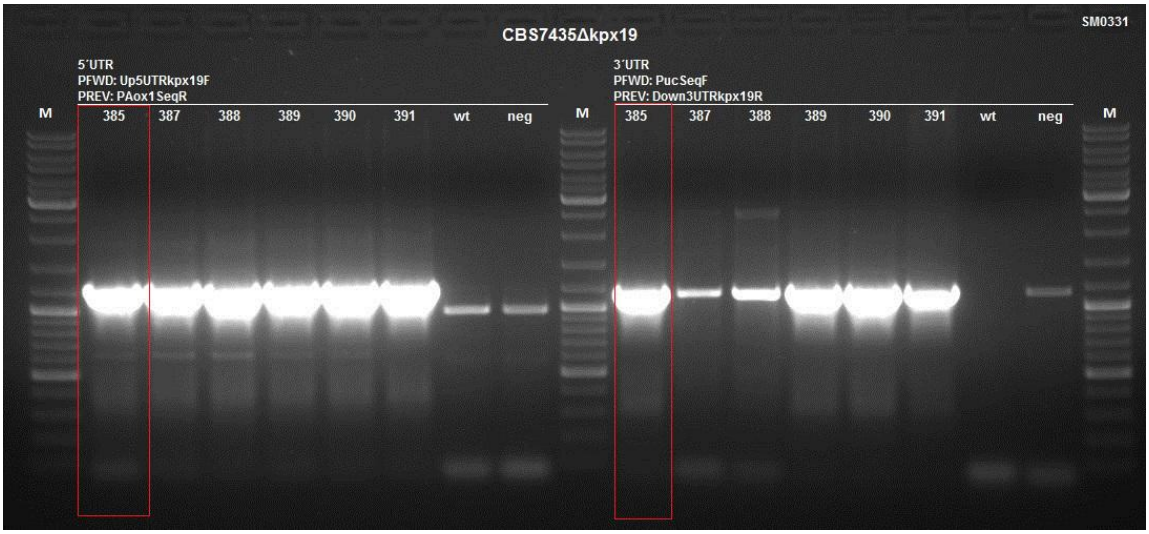


Suppl. Figure 22: CBS 7435 *Δkpx14*

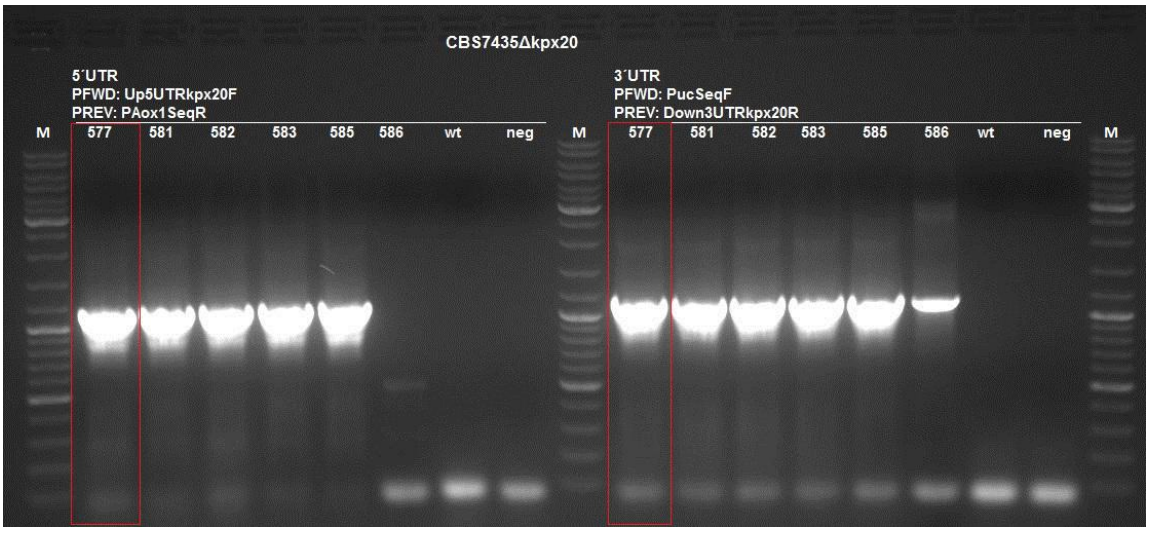
Suppl. Figure 23: CBS 7435 *Δkpx17*



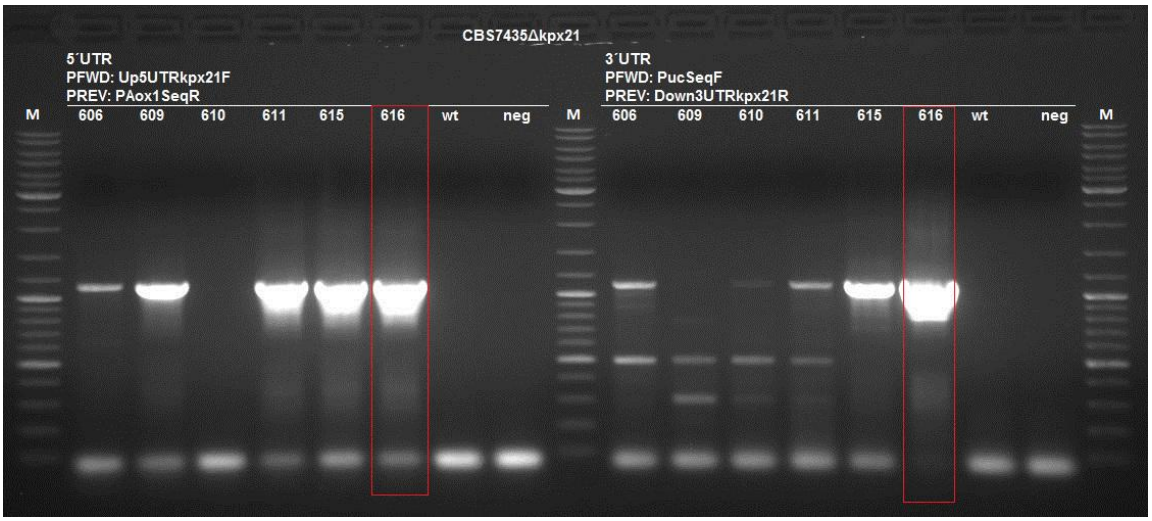
Suppl. Figure 24: CBS 7435 *Δkpx18*



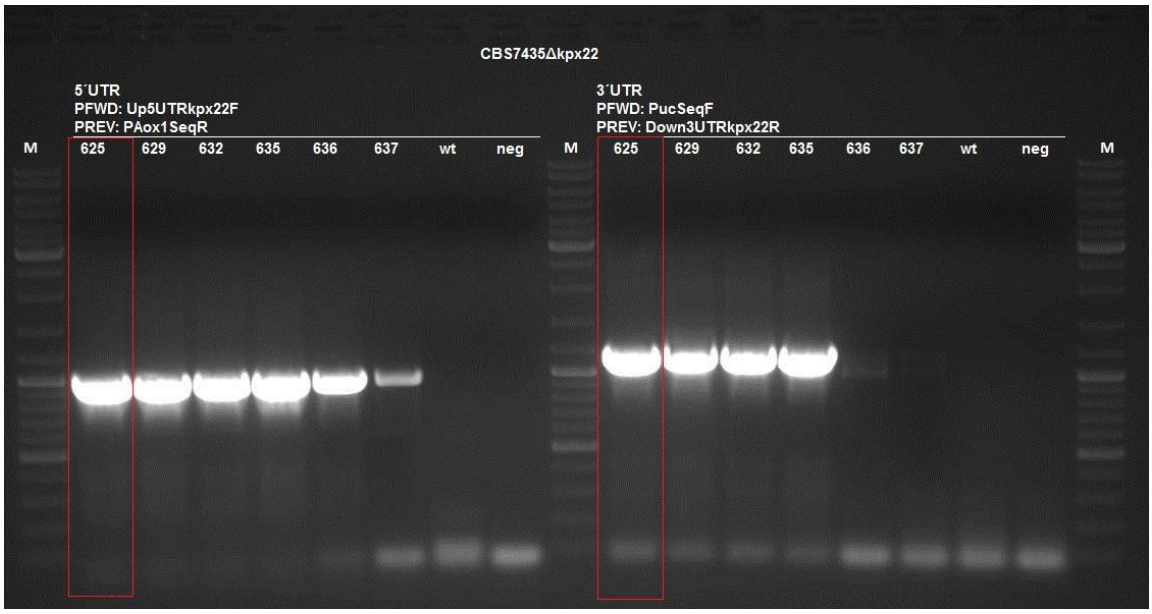
Suppl. Figure 25: CBS 7435 *Δkpx19*



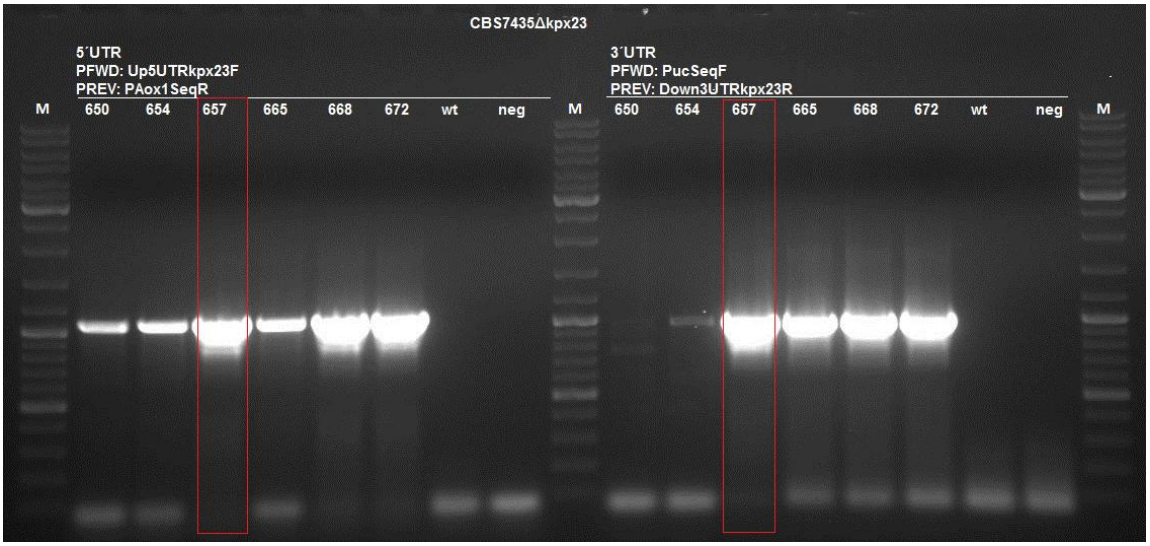
Suppl. Figure 26: CBS 7435 *Δkpx20*



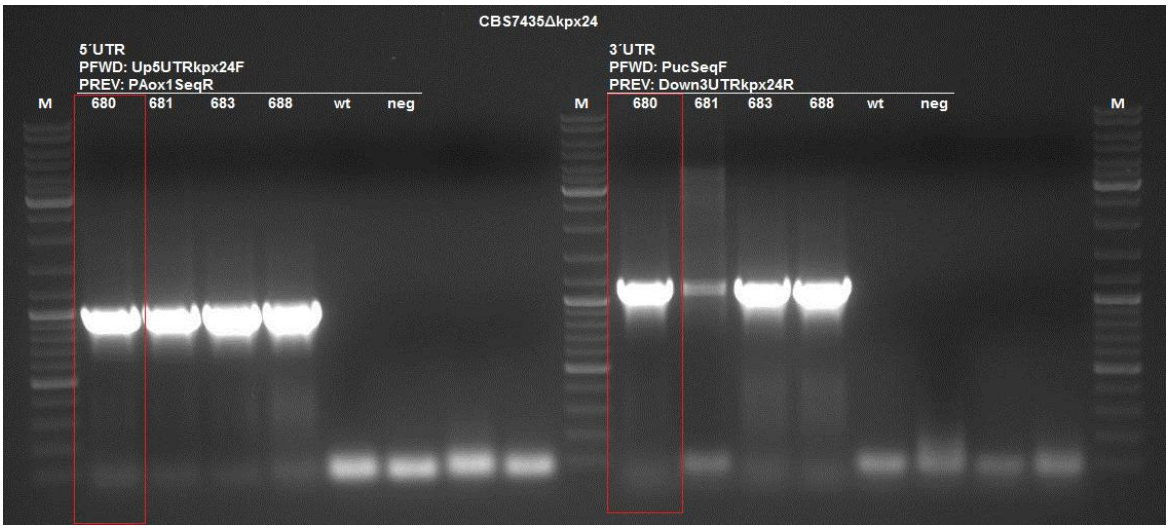
Suppl. Figure 27: CBS 7435 *Δkpx21*



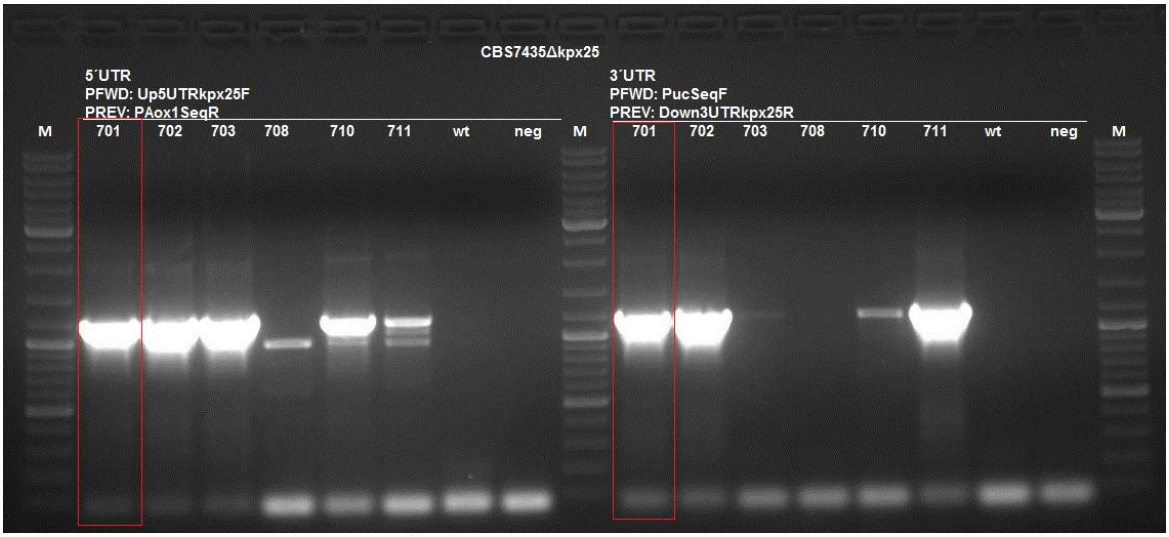
Suppl. Figure 28: CBS 7435 *Δkpx22*



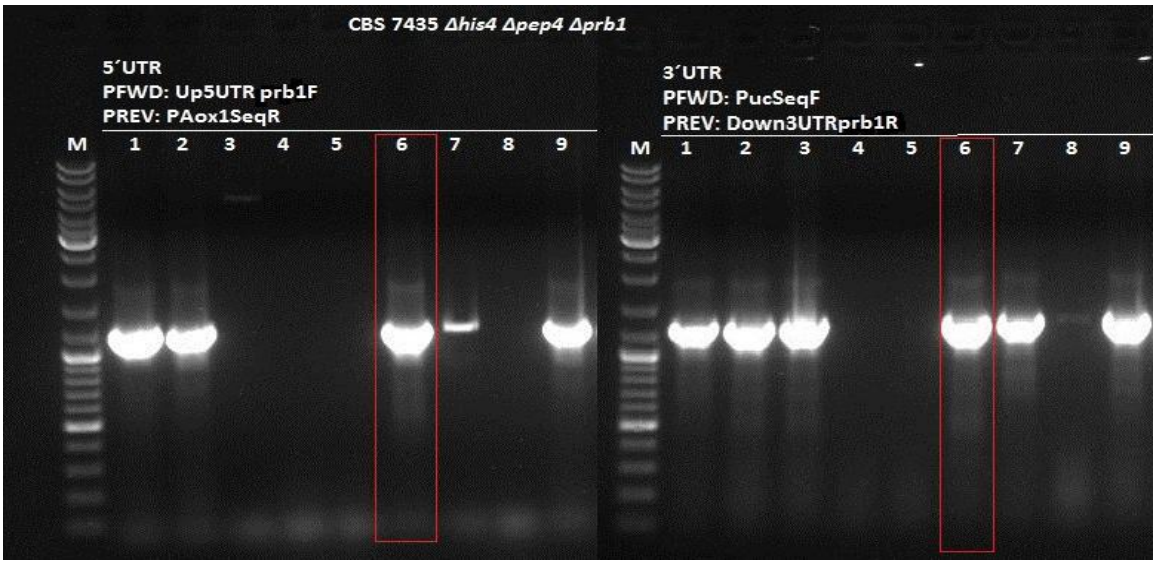
Suppl. Figure 29: CBS 7435 *Δkpx23*



Suppl. Figure 30: CBS 7435 *Δkpx24*

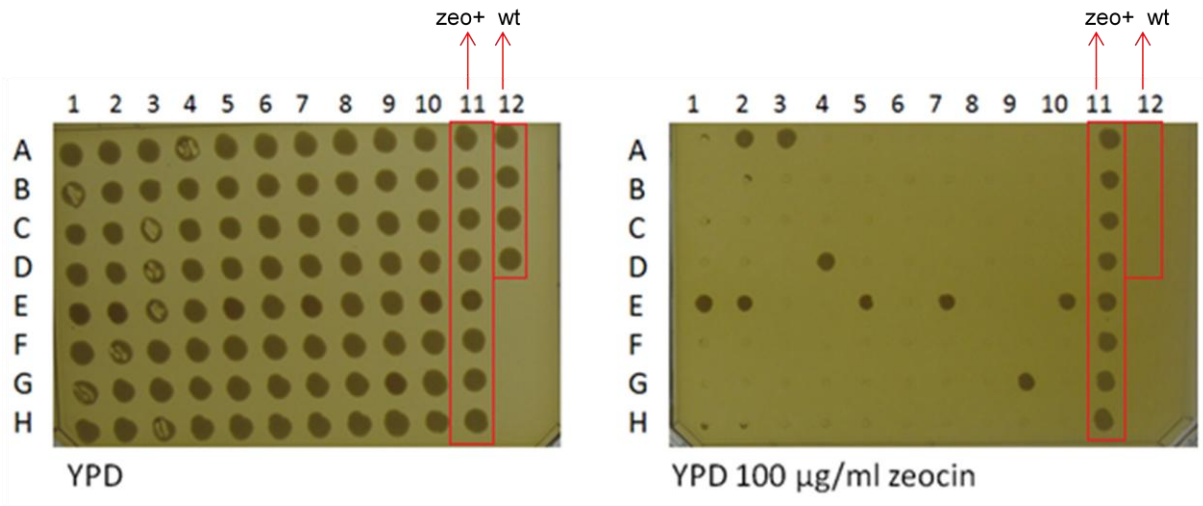


Suppl. Figure 31: CBS 7435 *Δkpx25*



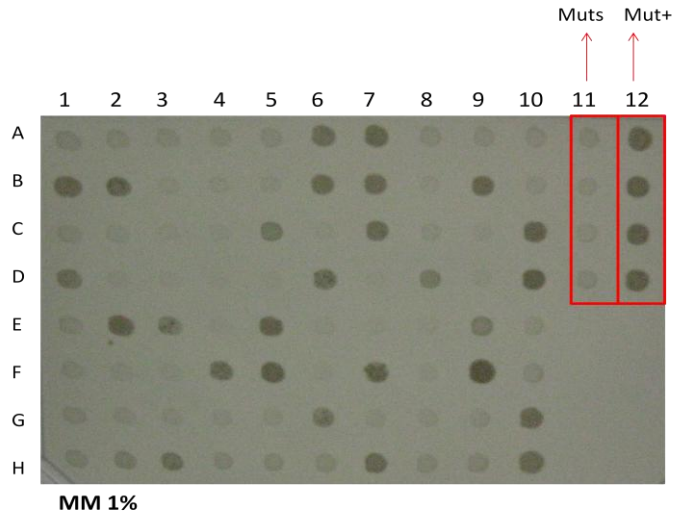
Suppl. Figure 32: CBS 7435 *Δhis4 Δpep4 Δprb1*. Double knockout was performed by transforming *prb1* knockout cassette into *pep4* deletion strain.

6.3 Removal of zeocin marker by induction of flippase gene



Suppl. Figure 33: Removal of zeocin marker by induction of flippase gene. After induction with 1% methanol for three days single colonies were picked and stamped onto a YPD plate (left) and onto a YPD plate supplemented with 100 µg/ml zeocin (right). About 90% of the colonies were losing their zeocin resistance that shows the effectiveness of this system.
zeo+: control strain with zeocin marker still inside
wt: wild type

6.4 Screening for Mut^S phenotype



Suppl. Figure 34: Screening for Mut^S phenotype. After transformation of Δ27ST6 expression cassette screening was done on minimal methanol medium (MM 1%). The difference in growth of Mut$^+$ and Mut^S strains is easily visible. The red rectangles indicate Mut^S and Mut$^+$ control strains.

6.5 Strains deposited in the culture collection

Suppl. Table 2: *Pichia pastoris* strains generated during this work and deposited in the culture collection.

Gene deleted	CC number	Abbreviation	Abbreviation VNTI file	Protease-deficient strain
sub2	6668	CBS7435Δsub2	sub2_locusAI	CBS 7435 Δ*sub2*
pep4	6911	CBS7435Δpep4	pep4_locus_AI	CBS 7435 Δ*his4* Δ*pep4*
prb1	6912	CBS7435Δprb1	prb1_locus_AI	CBS 7435 Δ*prb1*
prc1	6676	CBS7435Δprc1	prc1_locus_AI	CBS 7435 Δ*prc1*
yps1	6686	CBS7435Δyps1	yps1_locus_AI	CBS 7435 Δ*yps1*
yps2	6671	CBS7435Δctse1	ctse1_locus_AI	CBS 7435 Δ*yps2*
yps7	6907	CBS7435Δctsd	ctsd_locus_AI	CBS 7435 Δ*yps7*
kex1	6909	CBS7435Δkex1	kex1_locus_AI	CBS 7435 Δ*kex1*
kex2	6910	CBS7435Δkex2	kex2_locus_AI	CBS 7435 Δ*his4* Δ*kex2*
prtP	6673	CBS7435ΔprtP	prtP_locus_AI	CBS 7435 Δ*prtP*
ctse	6687	CBS7435Δctse2	ctse2_locus_AI	CBS 7435 Δ*ctse*
kpx1	6669	CBS7435Δkpx1	kpx1_locus_AI	CBS 7435 Δ*kpx1*
kpx4	6906	CBS7435Δkpx4	kpx4_locus_AI	CBS 7435 Δ*kpx4*
kpx5	6913	CBS7435Δkpx5	kpx5_locus_AI	CBS 7435 Δ*his4* Δ*kpx5*
kpx7	6914	CBS7435Δkpx7	kpx7_locus_AI	CBS 7435 Δ*his4* Δ*kpx7*
kpx8	6670	CBS7435Δkpx8	kpx8_locus_AI	CBS 7435 Δ*kpx8*
kpx10	6908	CBS7435Δkpx10	kpx10_locus_AI	CBS 7435 Δ*kpx10*
kpx14	6675	CBS7435Δkpx14	kpx14_locus_AI	CBS 7435 Δ*kpx14*
kpx17	6677	CBS7435Δkpx17	kpx17_locus_AI	CBS 7435 Δ*kpx17*
kpx18	6678	CBS7435Δkpx18	kpx18_locus_AI	CBS 7435 Δ*kpx18*
kpx19	6679	CBS7435Δkpx19	kpx19_locus_AI	CBS 7435 Δ*kpx19*
kpx20	6680	CBS7435Δkpx20	kpx20_locus_AI	CBS 7435 Δ*kpx20*
kpx21	6681	CBS7435Δkpx21	kpx21_locus_AI	CBS 7435 Δ*kpx21*
kpx22	6682	CBS7435Δkpx22	kpx22_locus_AI	CBS 7435 Δ*kpx22*
kpx23	6683	CBS7435Δkpx23	kpx23_locus_AI	CBS 7435 Δ*kpx23*
kpx24	6684	CBS7435Δkpx24	kpx24_locus_AI	CBS 7435 Δ*kpx24*
kpx25	6685	CBS7435Δkpx25	kpx25_locus_AI	CBS 7435 Δ*kpx25*
pep4, prb1	7013	CBS7435Δpep4Δprb1	pep4_locus_AI prb1_locus_AI	CBS 7435 Δ*his4* Δ*pep4* Δ*prb1*

Suppl. Table 3: *E. coli* strains generated during this work and deposited in the culture collection.

Gene	CC number	Knockout plasmid	Abbreviation VNTI file	Organism
sub2	6768	pPpKC1_sub2	pPpKC1_sub2, sub2_locus	*E. coli* TOP10 F´
kpx1	6769	pPpKC1_kpx1	pPpKC1_kpx1, kpx1_locus	*E. coli* TOP10 F´
kpx2	6770	pPpKC1_kpx2	pPpKC1_kpx2, kpx2_locus	*E. coli* TOP10 F´
yps3	6771	pPpKC1_kpx3	pPpKC1_kpx3, kpx3_locus	*E. coli* TOP10 F´
kpx4	6772	pPpKC1_kpx4	pPpKC1_kpx4, kpx4_locus	*E. coli* TOP10 F´
pep4	6773	pPpKC1_pep4	pPpKC1_pep4, pep4_locus	*E. coli* TOP10 F´
kpx5	6774	pPpKC1_kpx5	pPpKC1_kpx5, kpx5_locus	*E. coli* TOP10 F´
kpx6	6775	pPpKC1_kpx6	pPpKC1_kpx6, kpx6_locus	*E. coli* TOP10 F´
prb1	6776	pPpKC1_prb1	pPpKC1_prb1, prb1_locus	*E. coli* TOP10 F´
kpx7	6777	pPpKC1_kpx7	pPpKC1_kpx7, kpx7_locus	*E. coli* TOP10 F´
kpx8	6778	pPpKC1_kpx8	pPpKC1_kpx8, kpx8_locus	*E. coli* TOP10 F´
yps2	6779	pPpKC1_ctse1	pPpKC1_ctse1, ctse1_locus	*E. coli* TOP10 F´
yps7	6782	pPpKC1_ctsd	pPpKC1_ctsd, ctsd_locus	*E. coli* TOP10 F´
kpx9	6783	pPpKC1_kpx9	pPpKC1_kpx9, kpx9_locus	*E. coli* TOP10 F´
kpx10	6784	pPpKC1_kpx10	pPpKC1_kpx10, kpx10_locus	*E. coli* TOP10 F´
kpx11	6785	pPpKC1_kpx11	pPpKC1_kpx11, kpx11_locus	*E. coli* TOP10 F´
prtP	6786	pPpKC1_prtP	pPpKC1_prtP, prtP_locus	*E. coli* TOP10 F´
kpx12	6787	pPpKC1_kpx12	pPpKC1_kpx12, kpx12_locus	*E. coli* TOP10 F´
kpx13	6788	pPpKC1_kpx13	pPpKC1_kpx13, kpx13_locus	*E. coli* TOP10 F´
kpx14	6789	pPpKC1_kpx14	pPpKC1_kpx14, kpx14_locus	*E. coli* TOP10 F´
kpx15	6790	pPpKC1_kpx15	pPpKC1_kpx15, kpx15_locus	*E. coli* TOP10 F´
prc1	6791	pPpKC1_prc1	pPpKC1_prc1, prc1_locus	*E. coli* TOP10 F´
kpx17	6792	pPpKC1_kpx17	pPpKC1_kpx17, kpx17_locus	*E. coli* TOP10 F´
kpx18	6793	pPpKC1_kpx18	pPpKC1_kpx18, kpx18_locus	*E. coli* TOP10 F´
kpx19	6794	pPpKC1_kpx19	pPpKC1_kpx19, kpx19_locus	*E. coli* TOP10 F´
kpx20	6795	pPpKC1_kpx20	pPpKC1_kpx20, kpx20_locus	*E. coli* TOP10 F´
kpx21	6796	pPpKC1_kpx21	pPpKC1_kpx21, kpx21_locus	*E. coli* TOP10 F´
kpx22	6825	pPpKC1_kpx22	pPpKC1_kpx22, kpx22_locus	*E. coli* TOP10 F´
kpx23	6826	pPpKC1_kpx23	pPpKC1_kpx23, kpx23_locus	*E. coli* TOP10 F´
kpx24	6827	pPpKC1_kpx24	pPpKC1_kpx24, kpx24_locus	*E. coli* TOP10 F´
kpx25	6832	pPpKC1_kpx25	pPpKC1_kpx25, kpx25_locus	*E. coli* TOP10 F´
yps1	6828	pPpKC1_yps1	pPpKC1_yps1, yps1_locus	*E. coli* TOP10 F´
ctse	6829	pPpKC1_ctse2	pPpKC1_ctse2, ctse2_locus	*E. coli* TOP10 F´
kex1	6830	pPpKC1_kex1	pPpKC1_kex1, kex1_locus	*E. coli* TOP10 F´
kex2	6831	pPpKC1_kex2	pPpKC1_kex2, kex2_locus	*E. coli* TOP10 F´

REFERENCES

[1] S. Macauley-Patrick, M. L. Fazenda, B. McNeil, and L. M. Harvey, “Heterologous protein production using the Pichia pastoris expression system.,” *Yeast*, vol. 22, no. 4, pp. 249–70, Mar. 2005.

[2] J. L. Cereghino and J. M. Cregg, “Heterologous protein expression in the methylotrophic yeast Pichia pastoris.,” *FEMS Microbiol. Rev.*, vol. 24, no. 1, pp. 45–66, 2000.

[3] A. Q. Pedro, C. Maia, F. Sousa, J. A. Queiroz, and L. A. Passarinha, “Pichia pastoris : A Recombinant Microfactory for Antibodies and Human Membrane Proteins,” *J. Microbiol. Biotechnol.*, vol. 23, pp. 587–601, 2013.

[4] C.-J. Huang, L. M. Damasceno, K. a Anderson, S. Zhang, L. J. Old, and C. a Batt, “A proteomic analysis of the Pichia pastoris secretome in methanol-induced cultures.,” *Appl. Microbiol. Biotechnol.*, vol. 90, no. 1, pp. 235–47, Apr. 2011.

[5] H. Yurimoto, M. Oku, and Y. Sakai, “Yeast methylotrophy: metabolism, gene regulation and peroxisome homeostasis.,” *Int. J. Microbiol.*, vol. 2011, p. 101298, Jan. 2011.

[6] O. Cos, R. Ramón, J. L. Montesinos, and F. Valero, “Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast Pichia pastoris under different promoters: a review.,” *Microb. Cell Fact.*, vol. 5, p. 17, Jan. 2006.

[7] J. Jordà, P. Jouhten, E. Cámara, H. Maaheimo, J. Albiol, and P. Ferrer, “Metabolic flux profiling of recombinant protein secreting Pichia pastoris growing on glucose:methanol mixtures.,” *Microb. Cell Fact.*, vol. 11, p. 57, Jan. 2012.

[8] A. Solà, P. Jouhten, H. Maaheimo, F. Sánchez-Ferrando, T. Szyperski, and P. Ferrer, “Metabolic flux profiling of Pichia pastoris grown on glycerol/methanol mixtures in chemostat cultures at low and high dilution rates.,” *Microbiology*, vol. 153, no. Pt 1, pp. 281–90, Jan. 2007.

[9] I. J. van der Klei, H. Yurimoto, Y. Sakai, and M. Veenhuis, “The significance of peroxisomes in methanol metabolism in methylotrophic yeast.,” *Biochim. Biophys. Acta*, vol. 1763, no. 12, pp. 1453–62, Dec. 2006.

[10] K. De Schutter, Y.-C. Lin, P. Tiels, A. Van Hecke, S. Glinka, J. Weber-Lehmann, P. Rouzé, Y. Van De Peer, and N. Callewaert, “Genome sequence of the recombinant protein production host Pichia pastoris.,” *Nat. Biotechnol.*, vol. 27, no. 6, pp. 561–6, Jun. 2009.

[11] F. W. Krainer, C. Dietzsch, T. Hajek, C. Herwig, O. Spadiut, and A. Glieder, “Recombinant protein expression in Pichia pastoris strains with an engineered methanol utilization pathway,” *Microb. Cell Fact.*, vol. 11, no. 1, p. 22, 2012.

[12] V. Balamurugan, G. R. Reddy, and V. V. S. Suryanarayana, “Pichia pastoris : A notable heterologous expression system for the production of foreign proteins — Vaccines,” *Indian J. Biotechnol.*, vol. 6, no. April, pp. 175–186, 2007.

[13] F. S. Hartner, C. Ruth, D. Langenegger, S. N. Johnson, P. Hyka, G. P. Lin-cereghino, J. Lin-cereghino, K. Kovar, J. M. Cregg, and A. Glieder, “Promoter library designed for fine-tuned gene expression in Pichia pastoris,” *Nucleic Acid Res.*, vol. 36, no. 12, pp. 1–15, 2008.

[14] M. Bollok, D. Resina, F. Valero, and P. Ferrer, “Recent Patents on the Pichia Pastoris Expression System : Expanding the Toolbox for Recombinant Protein Production,” *Recent Pat. Biotechnol.*, vol. 3, pp. 192–201, 2009.

[15] T. Vogl and A. Glieder, “Regulation of Pichia pastoris promoters and its consequences for protein production,” *N. Biotechnol.*, vol. 30, no. 4, pp. 385–404, 2013.

[16] G. P. L. Cereghino, J. L. Cereghino, C. Ilgen, and J. M. Cregg, “Production of recombinant proteins in fermenter cultures of the yeast Pichia pastoris,” *Curr. Opin. Biotechnol.*, vol. 13, pp. 329–332, 2002.

[17] T. Kjeldsen, “Yeast secretory expression of insulin precursors.,” *Appl. Microbiol. Biotechnol.*, vol. 54, no. 3, pp. 277–86, Sep. 2000.

[18] M. Jahic, “Process Techniques for Production of Recombinant Proteins with Pichia pastoris,” Kungl Tekniska Högskolan, 2003.

[19] “MEROPS - the Peptidase Database.” [Online]. Available: http://merops.sanger.ac.uk/about/classification.shtml. [Accessed: 06-Jun-2013].

[20] N. D. Rawlings, A. J. Barrett, and A. Bateman, “MEROPS : the peptidase database,” *Nucleic Acid Res.*, vol. 38, no. November 2009, pp. 227–233, 2010.

[21] N. D. Rawlings and A. J. Barrett, “Evolutionary families of peptidases,” *Biochem. J.*, vol. 218, pp. 205–218, 1993.

[22] K. M. Sakamoto, “Ubiquitin-dependent proteolysis: its role in human diseases and the design of therapeutic strategies.,” *Mol. Genet. Metab.*, vol. 77, no. 1–2, pp. 44–56, 2002.

[23] O. Paper, “Proteolytic stability of recombinant human serum albumin secreted in the yeast Saccharomyces cerevisiae,” *Appl Microbiol Biotechnol*, vol. 53, pp. 575–582, 2000.

[24] M. W. T. Werten and T. J. V. A. N. D. E. N. Bosch, “High-yield Secretion of Recombinant Gelatins by Pichia pastoris,” *YEAST*, vol. 15, no. February, pp. 1087–1096, 1999.

[25] F. Hong, N. Q. Meinander, and L. J. Jo, “Fermentation Strategies for Improved Heterologous Expression of Laccase in Pichia pastoris,” *Inc. Biotechnol Bioeng*, vol. 79, pp. 438–449, 2002.

[26] S. Denman, M. Gustavsson, J. Lehtio, T. T. Teeri, K. Hult, and M. Martinelle, “Stable linker peptides for a cellulose-binding domain – lipase fusion protein expressed in Pichia pastoris Fusion proteins composed of a cellulose-binding domain,” *Protein Eng.*, vol. 14, no. 9, pp. 711–715, 2001.

[27] T. Boehm, S. Pirie-Shepherd, L. B. Trinh, J. Shiloach, and J. Folkman, “Disruption of the KEX1 gene in Pichia pastoris allows expression of full-length murine and human endostatin.,” *Yeast Chichester Engl.*, vol. 15, no. 7, pp. 563–572, 1999.

[28] Z. Ni, X. Zhou, X. Sun, Y. Wang, and Y. Zhang, “Decrease of hirudin degradation by deleting the KEX1 gene in recombinant Pichia pastoris,” *Yeast*, vol. 25, no. October 2007, pp. 1–8, 2008.

[29] M. W. T. Werten and F. A. De Wolf, “Reduced Proteolysis of Secreted Gelatin and Yps1-Mediated ␣ -Factor Leader Processing in a Pichia pastoris kex2 Disruptant,” *Appl. Environ. Microbiol.*, vol. 71, no. 5, pp. 2310–2317, 2005.

[30] X. Q. Yao, H. L. Zhao, C. Xue, W. Zhang, X. H. Xiong, Z. W. Wang, X. Y. Li, and Z. M. Liu, “Degradation of HSA-AX15(R13K) when expressed in Pichia pastoris can be reduced via the disruption of YPS1 gene in this yeast.,” *J. Biotechnol.*, vol. 139, no. 2, pp. 131–6, Jan. 2009.

[31] M. Wu, Q. Shen, Y. Yang, S. Zhang, W. Qu, J. Chen, H. Sun, and S. Chen, “Disruption of YPS1 and PEP4 genes reduces proteolytic degradation of secreted HSA/PTH in Pichia pastoris GS115.,” *J. Ind. Microbiol. Biotechnol.*, Mar. 2013.

[32] B. Guan, J. Lei, S. Su, F. Chen, Z. Duan, Y. Chen, X. Gong, H. Li, and J. Jin, “Absence of Yps7p, a putative glycosylphosphatidylinositol-linked aspartyl protease in Pichia pastoris, results in aberrant cell wall composition and increased osmotic stress resistance.,” *FEMS Yeast Res.*, vol. 12, no. 8, pp. 969–79, Dec. 2012.

[33] K. Salamin, D. Sriranganadane, B. Léchenne, O. Jousson, and M. Monod, “Secretion of an endogenous subtilisin by Pichia pastoris strains GS115 and KM71.,” *Appl. Environ. Microbiol.*, vol. 76, no. 13, pp. 4269–76, Jul. 2010.

[34] G. Skretas, S. Carroll, S. Defrees, M. F. Schwartz, K. F. Johnson, and G. Georgiou, “Expression of active human sialyltransferase ST6GalNAcI in Escherichia coli,” *Microb. Cell Fact.*, vol. 11, pp. 1–11, 2009.

[35] and I. T. Takeshi Yamamoto, Hideki Nagae, Yasuhiro Kajihara, “Mass Production of Bacterial alpha2,6-Sialyltransferase and Enzymatic Syntheses of Sialyloligosaccharides,” *Biosci. Biotechnol. Biochem.*, vol. 62, no. 2, pp. 210–214, 1998.

[36] A. Varki, “Sialic acids in human health and disease,” *Trends Mol Med*, vol. 14, no. 8, pp. 351–360, 2008.

[37] S. Erbayraktar, G. Grasso, A. Sfacteria, Q. Xie, T. Coleman, M. Kreilgaard, L. Torup, T. Sager, Z. Erbayraktar, N. Gokmen, O. Yilmaz, P. Ghezzi, P. Villa, M. Fratelli, S. Casagrande, M. Leist, L. Helboe, J. Gerwein, S. Christensen, M. A. Geist, L. Ø. Pedersen, C. Cerami-Hand, J.-P. Wuerth, A. Cerami, and M. Brines, “Asialoerythropoietin is a nonerythropoietic cytokine with broad neuroprotective activity in vivo.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 11, pp. 6741–6, May 2003.

[38] U. Grundmann, C. Nerlich, T. Rein, and G. Zettimeissl, “Complete cDNA sequence encoding human j-galactoside,” *Nucleic Acid Res.*, vol. 18, no. 3, p. 2188, 1990.

[39] C. Chen, P. Wu, C. Huang, and K. Cheng, “A Pichia pastoris fermentation strategy for enhancing the heterologous expression of an Escherichia coli phytase,” *Enzyme Microb. Technol.*, vol. 35, pp. 315–320, 2004.

[40] X. Yu and Y. Li, “Expression of Aspergillus oryzae Tannase in Pichia pastoris and Its Application in the Synthesis of Propyl Gallate in Organic Solvent,” *Food Technol. Biotechnol.*, vol. 46, no. 1, pp. 80–85, 2008.

[41] R. Slibinskas, D. Samuel, A. Gedvilaite, J. Staniulis, and K. Sasnauskas, “Synthesis of the measles virus nucleoprotein in yeast Pichia pastoris and Saccharomyces cerevisiae,” *J. Biotechnol.*, vol. 107, pp. 115–124, 2004.

[42] D. Ning, X. Junjianl, W. Xunzhang, S. Kuanyuan, R. Guirong, R. Xiangrong, C. Wenyin, and Z. Qing, “Expression , - HBs Fab Purification , Fragment and Characterization of Humanized Anti with Hepatitis Purification and Character Analysis of Recombinant Fab Fragment-The,” *J. Biocem.*, vol. 134, pp. 813–817, 2003.

[43] C. S. Hoffman and F. Winston, “A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for trausforma-,” *Gene*, vol. 57, pp. 267–272, 1987.

[44] J. Lin-cereghino, W. W. Wong, S. Xiong, W. Giang, T. Linda, J. Vu, S. D. Johnson, and G. P. Lin-cereghino, “Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast Pichia pastoris,” *Biotechniques*, vol. 38, no. 1, pp. 4–6, 2005.

[45] A. Baer and J. Bode, “Coping with kinetic and thermodynamic barriers : RMCE , an efficient strategy for the targeted integration of transgenes,” *Curr. Opin. Biotechnol.*, vol. 12, pp. 473–480, 2001.

[46] “Thermo Scientific Life Science Research.” [Online]. Available: http://www.thermoscientificbio.com/nucleic-acid-electrophoresis/ogeneruler-dna-ladder-mix-ready-to-use-100-10000-bp/. [Accessed: 01-Aug-2013].

[47] M. B. Ganatra, S. Vainauskas, J. M. Hong, T. E. Taylor, J.-P. M. Denson, D. Esposito, J. D. Read, H. Schmeisser, K. C. Zoon, J. L. Hartley, and C. H. Taron, “A set of aspartyl protease-deficient strains for improved expression of heterologous proteins in Kluyveromyces lactis.,” *FEMS Yeast Res.*, vol. 11, no. 2, pp. 168–78, Mar. 2011.

[48] H. Komano and R. S. Fuller, “Shared functions in vivo of a glycosyl-phosphatidylinositol-linked aspartyl protease, Mkc7, and the proprotein processing protease Kex2 in yeast.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 92, no. 23, pp. 10752–6, Nov. 1995.

[49] D. J. Krysan, E. L. Ting, C. Abeijon, L. Kroos, and R. S. Fuller, “Yapsins Are a Family of Aspartyl Proteases Required for Cell Wall Integrity in Saccharomyces cerevisiae,” *Eukaryot. Cell*, vol. 4, no. 8, pp. 1364–1374, 2005.

[50] “Thermo Scientific Pierce Protein Biology Products.” [Online]. Available: http://www.piercenet.com/browse.cfm?fldID=717EAB22-C50E-319F-D227-C1EB41C4343C. [Accessed: 01-Aug-2013].

[51] Invitrogen by Life Technologies, “Zeocin ™ Selection Reagent,” *User Guide*, 2012. [Online]. Available: http://tools.invitrogen.com/content/sfs/manuals/zeocin_man.pdf. [Accessed: 02-Aug-2013].

[52] D. Julius, A. Brake, L. Blair, R. Kunisawa, and J. Thorner, “Isolation of the Putative Structural Gene for the Endopeptidase Required for Processing of Yeast Prepro-a-Factor,” vol. 37, no. July, pp. 1075–1089, 1984.

[53] A. Albrecht, A. Felk, I. Pichova, J. R. Naglik, M. Schaller, P. de Groot, D. Maccallum, F. C. Odds, W. Schäfer, F. Klis, M. Monod, and B. Hube, “Glycosylphosphatidylinositol-anchored proteases of Candida albicans target proteins necessary for both cellular processes and host-pathogen interactions.,” *J. Biol. Chem.*, vol. 281, no. 2, pp. 688–94, Jan. 2006.