EXPRESSION OF RECOMBINANT PROTEINS IN *PICHIA PASTORIS* AND THEIR PURIFICATION

Master Thesis in the Degree Program:
Pharmaceutical Biotechnology

HAW Hamburg
Faculty of Life Sciences

Leonhard Ruh

Supervised by:
Prof. Dr. Gesine Cornelissen
Prof. Dr. Juan Asenjo

Date of Submission: 15.02.2023

This thesis was carried out in collaboration with the Centre for Biotechnology and Bioengineering (CeBiB) in Santiago de Chile.

Statutory Declaration

I herewith formally declare that I have written the submitted dissertation independently. I did not use any outside support except for the quoted literature and other sources mentioned in the thesis. I clearly marked and separately listed all of the literature and all of the other sources which I employed when producing this academic work, either literally or in content. I am aware that the violation of this regulation will lead to failure of the thesis.

Leonhard Ruh

Acknowledgements

First of all, I would like to express my sincere gratitude to Sebastian Jesus Rodriguez Vega for guiding through this project. Further, I would like to thank my supervisors Prof. Dr. Juan Asenjo and Prof. Dr. Cornelissen for the thoughtful comments and recommendations on this dissertation. I am also thankful to the Centre for Biotechnology and Bioengineering Chile (CeBiB) and all its member's staff for the support inside and outside the lab. Last but not least, a big thank you to my family and the DAAD for supporting me financially during my time in Santiago de Chile.

Abstract

For multiple decades enzymes have been essential to industrial sectors such as health, food, agriculture, cosmetics, energy, and others. However, many naturally occurring enzymes cannot stay active under the harsh conditions needed in those processes. On the search for more stable enzymes, scientists started looking at microorganism living under extreme conditions. Chile is home to many almost uninhabitable environments hosting very-well adapted extremophiles. The aim of the project was to produce and purify two proteases (subtilisin and trypsin) and a glycosidase (xylanase) whose sequences all originate from bacteria isolated from the Chilean Antarctica or Atacama Desert. The sequences were inserted into the genome of *Pichia pastoris* KM71 via homologous recombination utilizing the vector pPIC9K. The vector generates His^{+} Mut^{S} Pichia strains. Though all sequences could be confirmed inside the yeast's genome, only trypsin was successfully expressed. Proper conditions for the expression xylanase could not be found while there was no time left to work on the expression of subtilisin. An attempt was made to purify trypsin by utilizing the propeptides' His-tag. However, the zymogen trypsinogen did not bind to the column while active trypsin was found in the flow-through. Since the cleaved propeptide did bind to the HisTrap column it is likely that the tertiary structure of trypsinogen occludes the histidine tag.

Table of Contents

1 Introduction ... 1

1.1 Extremophiles in Chile and their potential ... 1

1.2 Thermostable enzymes in industrial applications ... 2

1.3 Recombinant DNA techniques ... 3

1.4 Cultivation and expression ... 5

1.5 Protein purification ... 6

1.6 The project ... 7

1.6.1 *Pichia pastoris* KM71 ... 8

1.6.2 pPIC9K ... 8

1.6.3 Subtilisin T8 ... 9

1.6.4 Trypsin ... 10

1.6.5 Xylanase Xyl-L ... 12

2 Materials and Methodology ... 13

2.1 Materials ... 13

2.1.1 Stock solutions, buffer, and media recipes ... 13

2.1.2 Kits ... 16

2.1.3 Strains ... 16

2.2 Methodology ... 17

2.2.1 DNA quantification ... 17

2.2.2 Gel electrophoresis ... 17

2.2.3 Glycerol stocks ... 17

2.2.4 Preparation of competent *E. coli* DH5α cells 17

2.2.5 Transformation of competent *E. coli* DH5α cells 18

2.2.6 Linearizing plasmids with *SacI* 18

2.2.7 Phenol-Chloroform Extraction and Ethanol Precipitation 19

2.2.8 Preparation of *P. pastoris* cells for the transformation 19

2.2.9 Electroporation of *P. pastoris* 20

2.2.10 *In Vivo* Screening of Multiple Inserts 21

2.2.11 PCR Analysis of Transformants 21

2.2.12 Activity Assays 23

2.2.13 Cultivation and Expression.. 23

2.2.14 Scale up 24

2.2.15 SDS-PAGE, Coomassie blue and silver staining 24

2.2.16 Purification via FPLC 26

3 Results 27

3.1 Multiplying, purifying and digesting plasmid DNA 27

3.2 Transformation and screening of *P. pastoris* 28

3.3 Expression... 29

3.3.1 Trypsin 29

3.3.2 Xylanase 33

3.4 Scale up and purification of trypsin... 35

4 Discussion 40

4.1 Multiplying, purifying and digesting plasmid DNA 40

4.2 Transformation and screening of *P. pastoris* 40

4.3 Expression of trypsin 41

4.4 Expression of xylanase 44

4.5 Scale up and purification 45

5 Summary and outlook 47

6 Sources 49

7 Appendix 62

7.1 Chemicals 62

7.2 Primer 65

7.3 Plasmids 65

7.3.1 pPIC9K-subtilisin 65

7.3.2 pPIC9K-trypsin 66

7.3.3 pPIC9K-xylanase 66

7.4 Protein sequences 67

7.4.1 Subtilisin: 67

7.4.2 Trypsin 68

7.4.3 Xylanase-L 68

Abbreviations

bp	Base pairs
BLAST	Basic Local Alignment Search Tool
BMG	Buffered minimal glycerol medium
BMM	Buffered minimal methanol medium
BMGY	Buffered complex glycerol medium
BMMY	Buffered complex methanol medium
CBM4	Carbohydrate Binding Module Family 4
CRISPR	Clustered regularly interspaced short palindromic repeats
CUB	Codon usage bias
DNA	Deoxyribonucleic acid
DSB	Double-stranded break
dsRNA	Double stranded RNA
FPLC	Fast protein liquid chromatography
HCDF	High cell density fermentation
HDR	Homology directed repair
IMAC	Immobilized metal affinity chromatography
mADH	Mitochondrial alcohol dehydrogenase isozyme III
MPB	Maltose binding protein
mRNA	Messenger RNA
nt	Nucleotides
ORF	Original reading frame
PCR	Polymerase chain reaction
RNA	Ribonucleic Acid
UPR	Unfolded protein response

1 Introduction

1.1 Extremophiles in Chile and their potential

Extremophiles are microorganisms inhabiting environments that would be fatal for other forms of life. They can survive in water close to boiling or below the freezing point. However, “extreme” does not only refer to temperature. Besides thermo- and psychrophiles (high and low temperatures) there are also acido- and alkaliphiles (extreme pH), barophiles (high pressure) and halophiles (high salt concentrations).[1]

These organisms are interesting not only for their incredible adaptation capabilities but also because of their applications in the field of biotechnology. For example, thermophile and psychrophile organisms can produce enzymes which retain their function at very high or low temperatures. In a lot of industrial processes enzymes that work under harsh conditions are essential.[2,3]

Chile is home to a lot of extreme habitats. In the north of the country, the Atacama Desert is located. It is the driest and oldest desert on earth and hosts salt flats, geysers and geothermal springs which provide extreme conditions such as alkaline/acidic pH, water stress and strong UV radiation.[4] It’s microbial and genetic richness has been explored in the past two decades and novel, interesting metabolites have been discovered.[5] In 2011, a group of scientists isolated *Streptomyces* strains from Laguna de Chaxa, Salar de Atacama. The strains were screened and multiple new compounds showing antibacterial activity discovered (Chaxamycin A-D).[6] Another more recent discovery were two extracellular xylanases from halotolerant *Bacillus* bacteria. These glycosidases cleave xylan, one of the most common polysaccharides on planet earth. Compared to similar enzymes they showed great thermal stability and stayed active at a wide range of pH. The enzyme is specifically interesting for industrial processes, as

it has been used to improve the quality of flour, biofuel production and bleaching in the paper industry, among other examples.[7] But not only enzymes are relevant for the food industry. In 2020 several *Nostocaceae* strains were discovered in the Atacama Desert, whose phycobiliproteins can be used as natural colorants in dairy beverages.[8]

Chile also claims part of Antarctica, an environment that shows extreme conditions below and above the water surface. Great gradients in temperature, salinity and irradiation can be observed. Naturally, scientists started looking for extremophiles producing novel metabolites.[4] Especially cold-adapted enzymes can be interesting for industrial processes. In 2021, a group of Chilean researchers screened several hundred strains isolated from the Antarctic in the search for cold-active β-galactosidases. Many strains displayed enzymatic activity with *Rahnella inusitats* being the most promising candidate. Assays showed its β-galactosidase variant keeps up to 62% of its activity at 15°C. The enzyme is especially interesting for the dairy industry since it hydrolyzes lactose and can be used to fabricate lactose-free products.[9] Another example of a cold-adapted enzyme from the Antarctica is an alcohol-dehydrogenase discovered in 2006. This enzyme facilitates the interconversion between alcohols and aldehydes or ketones. The variant discovered was the first cold-adapted and thermostable of its type, showing activity at temperatures between 0-85°C.[10]

1.2 Thermostable enzymes in industrial applications

For years enzymes have been essential to industrial sectors such as health, food, agriculture, cosmetics, energy, and others. More than 75% of commercial enzymes are hydrolases used to degrade numerous substances. However, a lot of naturally occurring enzymes do not work under conditions needed for certain processes. Their lack of stability makes them unviable. On the search for novel, better suited enzymes, scientists developed multiple strategies. Traditional strategies include screening

extremophiles for enzyme variants or mimicking evolution on a laboratory scale. A more modern approach is using the help of computers to develop enzyme-engineering strategies. Advances in recombinant DNA techniques have made it possible to insert sequences of promising enzymes into organism which are established in fermentation processes. As a results, expression systems with high-yield expression of heterologous proteins can be used for production and make the use of enzymes in industrial processes profitable.[11] The numbers of the industrial enzymes market growth underline the importance of enzyme for industrial applications. According to Precedence Research the market was valued at 5.9 billion USD in 2020 and is expected to grow to more than 10 billion USD in 2030.[12]

Popular examples of a commercial enzymes that originate from thermophile organism are DNA polymerases (Taq, Pfu, and Vent). These more thermostable variants stay active at temperatures above the denaturation temperature of long DNA fragments. Therefore, they helped scientific research remarkably by making polymerase chain reaction (PCR) a standard technique in laboratories worldwide. Further examples of industrial enzymes that need to withstand high temperatures in biotechnological processes include cellulases, amylases, xylanases, lipases, proteases, and esterases.[11]

1.3 Recombinant DNA techniques

With the help of recombinant DNA techniques, it is possible to modify genetic material and improve characteristics of a living organisms or their products. Usually, one or even multiple genes are introduced into an organism together with regulatory elements such as promotors. In some cases, it is also desirable to block or decrease expression of endogenous genes.[13] Different DNA fragments are generated by enzymatic cleavage and later joined in a vector using ligases. This vector will then be introduced

into an organism who will incorporate the desired genes into its genome.[14] Nowadays multiple recombinant DNA technologies are available.[15]

One way to express a desired gene in microorganism is the use of circular or linearized plasmids. They are a common choice for this purpose because of their easy manipulation and regulated expression. However, they suffer from instability issues and to ensure the existence of plasmids in the host cell, it is necessary to use selective agents (e.g., antibiotics) which increase costs.[16] Other downsides include the limited number of expression vectors for yeast cells and the limited control of copy numbers in the host organism. An alternative is the integration the genes directly into the genome via homologous recombination. It may not be optimal for the overexpression of genes but ensures segregational stability and control over the number of copies.[17]

In homologous recombination exchanges of DNA sequences occur between DNA molecules in a region of shared homology. This means the sequences must not be identical for their whole entirety but rather a limited region. The event consists of three phases: pre-synapsis (the early phase), synapsis (homologous strand exchange) and post-synapsis (the late phase). The event can occur in prokaryotes or eukaryotes and is catalysed by specialized enzymes.[18] While integration of circular plasmids can occur, digesting it with restriction enzymes helps to increase the integration efficiency remarkably.[19] Some vectors even allow for multiple gene insertion events.[20]

A more recent development is the CRISPR-Cas9 system (clustered regularly interspaced short palindromic repeats – CRISPR associated) which can target and cut specific parts of DNA. It induces a double-stranded break (DSB) to the DNA and subsequently, the homology directed repair system (HDR) of the cell is abused to introduce genes. The mechanism compares the sequence of the molecule carrying the DSB to another DNA molecule which is intact. If they correspond, the cell will repair

the DNA according to the homologue's DNA fragment. If this fragment contains a novel gene, the mechanism will introduce the gene itself while repairing the DSB.[21,22]

1.4 *Pichia pastoris* cultivation and recombinant protein expression

Pichia pastoris is a yeast that has been used for the production of recombinant proteins for many years.[23] It is able to grow on numerous carbon sources with the most commonly used being glucose, methanol and glycerol.[24] The yeast allows for high-cell-density fermentation (HCDF). Thereby, it is possible to achieve high levels of cell mass and large quantities of proteins can be produced.[25]

Many of the available protocols for recombinant protein expression are based on the commercial Invitrogen expression kit.[26] The manual suggests using BMGY/BMMY (buffered complex glycerol or methanol medium) or BMG/BMM (buffered minimal glycerol or methanol medium) for the expression. These medias are buffered to provide a stable pH in case the proteins are being secreted. Furthermore, BMGY/BMMY contains yeast extract and peptone. These components allow for better growth/biomass accumulation and stabilize secreted proteins. Another important parameter is the aeration. The volume of the culture should not exceed 10-30% of the total flask volume. Temperatures during growth and expression may depend on the strain but should generally be 28-30°C with shaking conditions between 225 – 300 rpm.[20]

In recent years it became a common approach to move away from standard protocols and develop strategies which are specifically designed for certain genetic constructs and the available equipment.[26] An example is the expression strategy applied by AG Mattanovich at the Universität für Bodenkultur in Vienna. The group recently started to collaborate with the CeBiB and has been working with *Pichia* for many years.[27–29] Their

protocol alterations include growing the yeast on minimal media and the use of glucose limitation during induction.[30]

1.5 Protein purification

Producing and purifying proteins still entails high costs with downstream processing being a big proportion of it. To make a process profitable, it is essential to have profound knowledge about possible purification methods and their optimization.[31] All separation steps exploit differences in properties of molecules such as shape, charge, size, isoelectric point, charge distribution, density, hydrophobicity, ligand-binding affinity, metal binding or specific sequences or structures.[32] Firstly, proteins must be separated from the nonproteins of the broth. These first steps usually include methods with low-resolution but high capacity (e.g., two-phase partition systems or fractional precipitation). Secondly, the desired protein must be separated from other proteins present in the mixture while preserving its biological activity and chemical integrity. Later steps utilize methods with higher resolution but lower capacity such as chromatography.[33–35]

Proteins can whether accumulate inside the cells or be secreted to the outside. This will make a huge difference in the purification process. When working with intracellular proteins, cells must first be broken by homogenization. Established techniques include continuous-fill French press, sonication, or enzymatic digest. Naturally, the cell lysate will have a lot more impurities such as nucleic acids and cell debris that must be removed before chromatography steps. Either way, centrifugation is an essential technique that must be performed whether the protein is intracellular or secreted. It helps to remove unbroken cells, cell debris and ribosomal material and other particulates. If necessary, phase partitioning or membrane filtration can be performed subsequently.[36]

There are numerous chromatographic techniques available for protein purification though affinity chromatography is regarded as the most specific and effective. This technique exploits the ability of molecules to form reversible bonds with affinity ligands.[37] For the purification of recombinant proteins, immobilized metal affinity chromatography (IMAC) is commonly utilized.[38] The technique involves 4 steps. Firstly, metal ions are immobilized onto the matrix. Subsequently, the sample will be loaded onto the column. The targeted protein should interact with the metal ions and be retained. The third step consists of washing the column to remove unwanted components. Lastly, the targeted protein will be eluted.[39,40]

A popular example of the IMAC is purification with His6-tag. Six consecutive histidines are introduced in a cloning vector to the N- or C-terminus of a recombinant protein. Transition metal ions (e.g., Ni^{2+} or Co^{2+}) are immobilized on a resin matrix inside the column. The His-tags show high affinity to the metal-ion and bind to it. A substance called imidazole competes with the protein for binding and can be used to elute the protein from the column after washing steps have been performed. Usually, a low concentration of imidazole is part of the binding and wash buffer to interfere with weak interactions some undesired proteins may have with the column.[41] Other notable chromatography techniques in protein purification include ion exchange, hydrophobic interactions, and reverse-phase chromatography.[31]

1.6 The project

Aim of the project is to produce and purify two proteases and a glycosidase whose sequences all originate from bacteria isolated from the Chilean Antarctica or Atacama Desert.[42] These sequences will be inserted into the genome of *Pichia pastoris* KM71 via homologous recombination utilizing the vector pPIC9K. The expression is induced with methanol and the enzymes will be purified subsequently.

1.6.1 *Pichia pastoris* KM71

Pichia pastoris is a methylotrophic yeast that has been widely used to produce recombinant proteins. Unlike bacteria, eukaryotic expression systems have the advantage of post-translational modifications. The expression is usually induced via methanol and can be repressed in the presence of other carbon sources.[43] The first enzyme in the methanol utilization pathway is called alcohol oxidase (AOX) and is encoded by the two genes AOX1 and AOX2. AOX1 shows greater methanol-oxidizing activity though the protein-coding regions of the two genes have similarity greater than 90%. Strains with a disrupted AOX1 gene are still able to grow on methanol, however, significantly slower than the wild type.[44]

The strain *Pichia pastoris* KM71 that was used for this project. It cannot grow in the absence of histidine because of a mutation in the *HIS4*. This gene encodes for an enzyme catalysing multiple crucial steps in the biosynthesis of histidine.[45] Furthermore, the *ARG4* gene was used to disrupt *AOX1*. *ARG4* allows for the synthesis of arginine. Therefore, *P. pastoris* KM71 is a Mut^S (methanol utilization slow), Arg^+ and His^- strain.

1.6.2 pPIC9K

pPIC9K is an established vector for the integration of recombinant genes into *Pichia pastoris* KM71 and GS115. The plasmid is shown in figure 1.

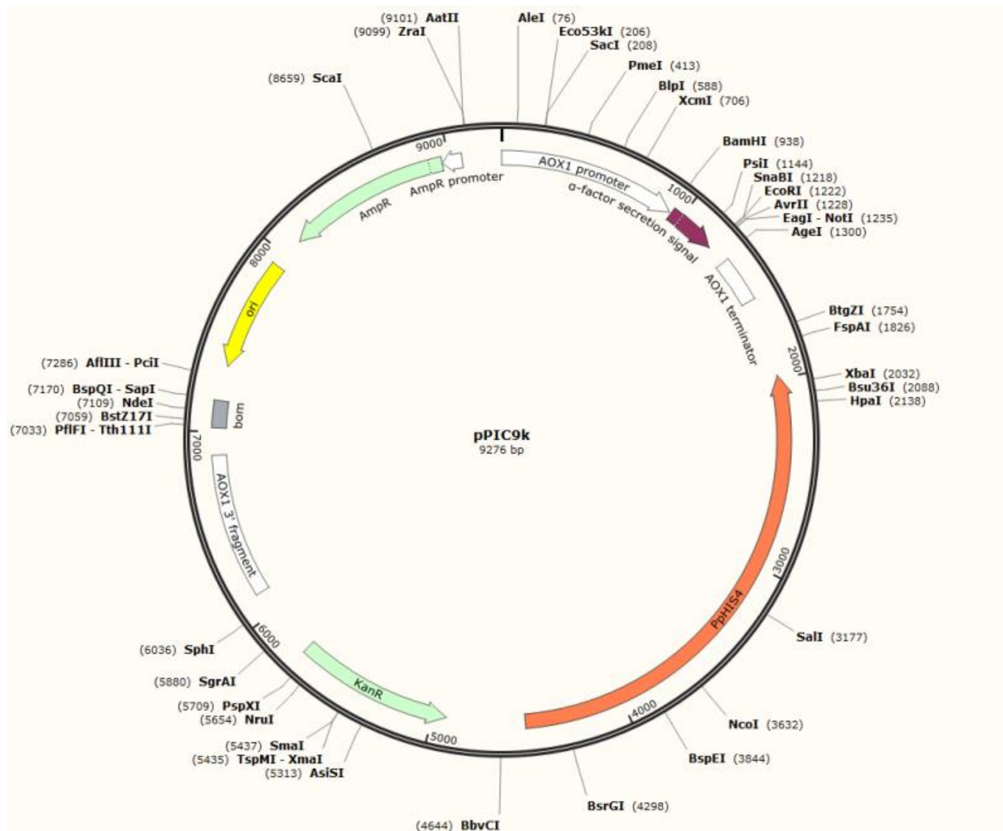


Figure 1: pPIC9K vector used for integration of genes in Pichia pastoris

pPIC9K allows for the integration of multiple inserts and can generate His$^+$ Mut$^+$ and His$^+$ Mut^S transformants. Four restriction sites (*Sna*B I, *Eco*R I, *Avr* II, *Not* I) can be utilized to insert genes at the desired location. The proteins will be secreted outside of the cell thanks to the α-factor secretion signal. Besides two antibiotic resistances (kanamycin and ampicillin), transformants can be screened via *HIS4* selection. This gene encodes for an enzyme crucial for the biosynthesis of histidine. Therefore, transformants will be able to grow on media lacking the amino acid. [45,46]

1.6.3 Subtilisin T8

Subtilisin T8 originates from bacteria isolated from the Chilean Antarctic and so far, had only been expressed in the native organism.

In the search for thermostable enzymes, several bacteria isolated from the Chilean Antarctic were cultivated in liquid media and their supernatant screened. A *Polaribacter*

sp. strain seemed to produce an interesting, cold-adapted enzyme with proteolytic activity. Using hydrophobic interaction chromatography (phenyl-sepharose) and SDS-PAGE it was possible to isolate two peptides which were further analyzed via mass spectrometry. The peptides corresponded to the same enzyme but in different states. Subsequently, primers were designed to identify the original reading frame (ORF) via genome walking.

The detected enzyme turned out be a type of subtilisin, a protease containing serine in its active site. The ORF is encoding for a large protein of 1100 amino acids. However, the active protein is much shorter than the coding region, so various proteolytic cleavages are necessary for the secretion and activation of the subtilisin. The plasmid construct utilized for this project contained a propeptide and the sequence of the active enzyme. The propeptide is needed to maintain subtilisin inactive during the secretion.

Multiple attempts have been made to achieve heterologous expression of the enzyme in *E. coli* but without any success. The protease had only been produced by the native strain so far and was further characterized. Greatest activity was observed at a pH between 7-8 and temperature about 45°C. Subtilisin T8 still retains 25% of its activity at 20°C what makes it specifically interesting for the use in detergents. [42]

In 2021 an alkaline variant of subtilisin was successfully expressed in *Pichia pastoris* SMD1168 and X33 utilizing the vector pPICZαC. The results suggested a second signal peptide could potentially enhance secretion levels of the protease but is not essential.[47]

1.6.4 Trypsin

The trypsin sequence originates from Micromonospora M12, isolated from the Atacama Desert. It was found making use of the Basic Local Alignment Search Tool

(BLAST) and a trypsin sequence found in *Streptomyces griseus*. Trypsin is another protease having serine in its active site and needed in various industrial applications such as leather bating, food processing, pharmaceuticals, clinical diagnoses, and biochemical testing applications.[48,49]

Sequences of the enzyme have been discovered in multiple sources, bacterial and mammalian organisms.[50] The inactive precursor of trypsin is called trypsinogen and can be activated by enteropeptidases or autoactivation.[51] Heterologous trypsin has been expressed in numerous systems already including *E. coli*[52], *transgenic plant cells*[53], *and mammalian cells*[54]. However, none of these expression systems were suitable for production on an industrial scale. Problems included low activity, safety, and operational difficulties.[55] Therefore, producing the enzyme in *Pichia pastoris* is considered a viable option.[50,56]

Previous studies have shown that highest trypsin expression could be achieved with an optimal number rather than a maximum number of copies inside the genome of *Pichia pastoris*. The data suggested the reason being accumulated enzyme causing an unfolded protein response (UPR) inside the yeast.[57] In 2016 a paper was published on how to improve the production of active, secreted trypsin in *Pichia pastoris* using the vector pPIC9K. They fused an artificial peptide to the N-terminus of the sequence, consisting of thioredoxin (Trxa), glycine-serine linker (GS*3), His_6-tag and the partial bovine trypsinogen pro-peptide (DDDDK). Thioredoxin functions as a chaperone while His_6-tags facilitate purification by having the ability to bind to metal ions in columns. The pro-peptide was kept so trypsin stays inactive during secretion. The production of active trypsin in fed-batch fermentation was improved significantly.[50]

The construct used in this project is similar to the one described above.

1.6.5 Xylanase Xyl-L

The sequence of xylanase xyl-L originates from a *Psychrobacter* strain isolated from Antarctica in 2008. The project aimed at finding cold-active lipases using degenerate primer. One of the amplifications turned out have high homology with xylanases of the GH10 family.[42,58]

The enzyme degrades xylan, one of the most common polysaccharides on earth. Xylan is a natural branched heteropolymer consisting of hexoses and pentoses, and a major component of plant cell walls. Xylanases already find wide applications in the food, biofuel, and paper industry.[7]

Xylanase xyl-L has already been produced in an *E. coli* strain at CeBiB. The enzyme showed optimal activity at pH 7 – 8 and retained 25% of its activity at 10°C. Despite being active at low temperatures, the enzyme was not very thermostable what would make its application in industrial processes challenging. Several attempts have been made to construct fusion proteins which help increasing thermostability while keeping the enzyme cold-active. With the information gained from these experiments, the gene used in this project was designed.[42] The maltose binding protein (MBP) was added to the N-terminus of xylanase. MBP was shown to increase folding and yield of recombinant proteins in yeast.[59] Furthermore, Carbohydrate Binding Module Family 4 (CBM4) was added to the C-terminus to improve binding to xylan.[60]

In 2013 a xylanase from *Aspergillus niger* was successfully expressed in *Pichia pastoris* GS115 using pPIC9K. Out of the multiple transformants obtained, one that showed highest level of resistance against geneticin also showed highest xylanase activity.[61] Furthermore, a thermostable variant of the enzyme was successfully expressed in *Pichia pastoris* GS 115 using a similar vector (pPIC9).[62]

2 Materials and Methodology

2.1 Materials

2.1.1 Stock solutions, buffer, and media recipes

If not further specified, recipes were taken from "Pichia Expression Kit" by Invitrogen™.[20] All percentages refer to g/100 mL.

BMGY medium contains:

Ingredient	Concentration
Biotin	$4*10^{-5}$%
Glycerol	1%
Peptone	2%
Potassium buffer, pH 6	100 mM
Yeast extract	1%
YNB	1.34%

BMMY medium contains:

Ingredient	Concentration
Biotin	$4*10^{-5}$%
Methanol	0.5%
Peptone	2%
Potassium buffer, pH 6	100 mM
Yeast extract	1%
YNB	1.34%

LB (agar) medium contains:

Ingredient	Concentration
Agar	2%
NaCl	1%
Tryptone	1%
Yeast Extract	0.5%

RDB (agar) medium contains:

Ingredient	Concentration
Agar	2%
L-glutamic acid	0.005%
L-isoleucine	0.005%
L-leucine	0.005%
L-lysine	0.005%
L-methionine	0.005%
Biotin	$4*10^{-5}$%
Dextrose	2%
Sorbitol	1M
YNB	1.34%

YPD (agar) medium contains:

Ingredient	Concentration
Agar	2%
Dextrose	2%
Peptone	2%
Yeast Extract	1%

For YPD plates containing geneticin the antibiotic is added before pouring the plates in respective concentrations.

Potassium buffer (1 M, pH 6) contains per litre:

Ingredient	Volume
1 M K_2HPO_4	132 mL
1 M KH_2PO_4	868 mL

Tris buffer (1.5/0.05 M, pH 7.4) contains:[63]

Ingredient	Concentration
Tris base	1.5/0.05 M

The pH was adjusted using HCl.

Binding buffer (pH 7.4) contains:[64]

Ingredient	Concentration
Sodium phosphate	20 mM
NaCl	0.5 M
Imidazole	20 mM

Elution Buffer (pH 7.4) contains:[64]

Ingredient	Concentration
Sodium phosphate	20 mM
NaCl	0.5 M
Imidazole	500 mM

SDS Running Buffer (10x) contains:[65]

Ingredient	Concentration
Tris base	0.2501 M
NaCl	1.924 M
Imidazole	0.03467 mM

Ampicillin and Geneticin

An ampicillin stock solution with a concentration of 25 mg/mL and a geneticin stock solution of 200 mg/mL were prepared. Both were sterilized with a 0.22 µm syringe filter.

2.1.2 Kits

To extract plasmids from bacteria the GeneJET Plasmid Miniprep Kit (ThermoFisher cat. no. 00338468) was used.

2.1.3 Strains

Escherichia coli DH5α was used to transform and multiply plasmids. It is one of the most used strains for this purpose.[66]

Pichia pastoris KM71 was used for the methanol-induced expression of the proteins. The strain is widely used for the expression of recombinant proteins.[67]

2.2 Methodology

2.2.1 DNA quantification

For measuring DNA concentrations, a MaestroNano by Maestrogen was used. It measured 3 wavelengths for RNA, DNA (A260), protein concentrations (A280) and other contaminants (A230) so it also provides information about the purity of the sample. For each measurement 2 µL were used.

2.2.2 Gel electrophoresis

For analyzing the size of DNA fragments gel electrophoresis was performed with 0.7% agarose gels with GelRed Nucleic Acid Stain 10000x in DMSO. The power supply EPS 3500 XL by Pharmacia Biotech was used, and the gel ran at 70 V for 40 minutes in TAE 1x buffer.

Afterwards the gel was analyzed with the Gel Doc™ EZ System by Bio-Rad.

2.2.3 Glycerol stocks

To prepare glycerol stocks, 500 µL of bacteria/yeast culture were added to 500 µL of glycerol 40 % and frozen at -80°C.

2.2.4 Preparation of competent *E. coli* DH5α cells

Competent cells are bacterial cells that can accept foreign DNA and plasmids from the environment.[68] They are needed to perform transformation. The cells were prepared using the $CaCl_2$ and $MgCl_2$ method adapted from Sambrook et al. (1989).[69]

A culture of *E. coli* DH5α was grown overnight (37°C in LB medium). On the next day 50 mL of fresh LB were inoculated with 1 mL (1/50) of the overnight culture and incubated at 37°C and 250 rpm until an OD_{600} of 0.2 – 0.3 was reached. Afterwards the culture was placed on ice for 10 min and then centrifuged at 3000 rpm and 4°C for 10 min. The supernatant was discarded, and the pellet resuspended in 12.5 mL (1/4) 0.1 M $MgCl_2$ before being put on ice for 5 min. Then the resuspended cells were

centrifuged again for 10 min at 4000 rpm at 4°C. The supernatant was removed again, and the cells got resuspended in 2.5 mL (1/20) 0.1 M $CaCl_2$ before being put on ice for 20 min. Afterwards the cells were centrifuged one last time at 4°C and 4000 rpm for 10 min. The pellet got resuspended in 1 mL (1/50) of 85% 0.1 M $CaCl_2$ solution plus 15% glycerol. The cells were aliquoted in 100 µL in sterile 0.6 mL Eppendorf tubes before being stored at -80°C until use.

2.2.5 Transformation of competent *E. coli* DH5α cells

Transformation is the uptake of DNA from the environment by a bacteria cell.[70] The procedure was adapted from Sambrook et al. (1989) and the Pichia Expression Kit by Invitrogen™ (2014).[20,69]

2.5 µL (625 ng) of the diluted genes were mixed with 100 µL of chemically competent *E. coli* DH5α cells and kept on ice for 30 minutes. Afterwards a heat shock was performed at 42 °C for 45 sec before being placed on ice again for 1 min. Then 900 µL of LB medium were added and the tube incubated in a shaker at 37°C for an hour. The bacteria were cultivated on LB plates (supplemented with 100 µg/mL ampicillin) and grown at 37°C.

2.2.6 Linearizing plasmids with *SacI*

The digestion was performed at 37°C for 1h followed by a heat inactivation at 65°C for 20 min. The protocol was taken from New England Biolabs' website and the digestion mixture is shown in table 1.[71]

Table 1: Digestion mixture *SacI*

Ingredient	Mass or Volume
DNA	20 µg
NEBuffer™ r1.1	50 µL
SacI	10 µL
Water	to 500 µL

2.2.7 Phenol-Chloroform Extraction and Ethanol Precipitation

The protocol was adapted from Thermo Fisher Scientific Inc.[72]

500 µL of phenol:chloroform:isoamyl alcohol (25:24:1) were mixed with 500 µL of the sample containing the DNA and vortexed for 20 sec. Afterwards, the mixture was centrifuged for 5 minutes at 4°C at 16000 x g. The upper, aqueous phase (~500 µL) was removed and transferred into a fresh tube.

After the phenol-chloroform extraction, 250 µL of 7.5M NH_4OAc and 1875 µL of 100% ethanol were added. The sample was stored at -80°C for 1.5h. Afterwards, the sample was centrifuged at 4°C for 30 minutes at 16000 x g and the supernatant discarded. 150 µL of 70% ethanol were added before centrifuging the sample at 4°C for 2 min. The supernatant was discarded, and the sample placed in a drying cabinet for 10 min. In a final step, the pellet was resuspended in TE buffer by pipetting up and down 30-40 times.

2.2.8 Preparation of *P. pastoris* cells for the transformation

The procedure was taken from "*Pichia* Expression Kit" by invitrogen™.[20]

A 5 mL culture *Pichia pastoris* GM71 was grown overnight in YPD media in a 50 mL falcon tube at 30°C and 200 rpm. The next day, an Erlenmeyer flask containing 500

mL of fresh YPD media was inoculated with 0.2 mL of the overnight culture and grown until an OD_{600} of 1.3 – 1.5 was reached (30°C, 200 rpm).

The cells were centrifuged four times at 1500 x g at 4°C for 5 min. After the first centrifugation step, the cells were resuspended in 500 mL ice cold, sterile water and after the second one in 250 mL. In the last two centrifugation steps, the cells were first resuspended in 20 mL, then in 1 mL of 1 M ice-cold sorbitol.

Transformation of cells was performed immediately after.

2.2.9 Electroporation of *P. pastoris*

The procedure was taken from "*Pichia* Expression Kit" by invitrogen™.[20]

80 µL of the cells were mixed with 10 µg of linearized DNA and transferred to a 0.2 cm electroporation cuvette which was put on ice. The cuvette containing the mixture was stored on ice for 5min before the electroporation was performed.

The electroporation device used was the Gene Pulser Xcell Electroporator by Bio-Rad Laboratories, Inc. The manufacturer suggested for electroporation of *Pichia pastoris* the following parameters:

Table 2: Parameters for electroporation of Pichia pastoris

Pulse type	**Capacitance (µF)**	**Resistance (Ω)**	**Volt**	**Cuvette (cm)**	**Cell vol (µL)**
Exponential decay	25	200	2000	0.2	40

Immediately after the electroporation ended, 1 mL of ice-cold 1 M Sorbitol was added to the cuvette and its content transferred to a sterile microcentrifugation tube.

200 µL of the mixture was plated onto RDB plated and incubated at 30°C until colonies could be seen.

2.2.10 *In Vivo* Screening of Multiple Inserts

The procedure was taken from "pPIC9K - A Pichia Vector for Multicopy Integration and Secreted Expression" by Invitrogen™.[46]

2 mL of sterile water was pipetted on the RDB plates containing transformants. By running a sterile spreader above the top layer of the agar, the colonies were removed from it. The cell suspensions were then transferred into a 50 mL conical centrifuge tube and briefly vortexed. By using a spectrophotometer, the cell density was determined. 1 OD_{600} correspond to 5 x 10^7 cells/mL.

10^5 cells were plated on YPD agar containing different concentrations of geneticin (0.25, 0.5, 0.75, 1.0, 1.5 and 2.0 mg/mL). The plates were incubated at 30°C and checked daily if colonies appeared.

2.2.11 PCR Analysis of Transformants

To check the transformants for inserts, colony PCR was performed. The procedure was adapted from Royle and Polizzi (2017).[73]

As pre-treatment, single colonies were resuspended in 10 µL 20 mM NaOH and incubated at 100°C for 15 min. 2 µL of the supernatant were then used as template for PCR.

The mix for the PCR can be seen in the table 3.

Table 3: Mix for colony PCR

Component	Concentration/Volume
Phusion HF Buffer	1X
dNTPs	10 mM
Forward primer	10 µM
Reverse primer	10 µM
DMSO	2% (v/v)
Phusion DNA polymerase	1 U
DNA template	2 µL
Total Volume	50 µL

The protocol used for the thermocycler is shown in the table 4.

Table 4: Thermocycler protocol for colony PCR

Temperature	Time
Initial denaturation	
95°C	7 min
Loop	
95°C	20 sec
53°C	30 sec
72°C	80 sec
Final elongation	
72°C	7 min

2.2.12 Activity Assays

2.2.12.1 Trypsin

The assay was adapted from Iversen and Jørgensen (1995). [74]

Tris buffer: 50 mM, pH 7.6

Azocasein: 6 g/L in Tris buffer

Trypsin: 0.05 g/L in Tris buffer

Trichloroacetic acid (TCA): 2 M

Positive Control: 200 µL azocasein (6 g/L) + 200 µL trypsin in Tris buffer

200 µL of the samples were mixed with 200 µL of 6 g/L azocasein and incubated at 37°C for 24h. The reaction was stopped by adding 800 µL of TCA. Afterwards the tubes were centrifugated for 10 min at 13000 rpm and the absorbance of the supernatant measured at 440 nm. Water was used as a blank.

2.2.12.2 Xylanase

The protocol was adapted from Sanhueza et al. (2018).[75]

A stock solution of 2% (w/v) xylan (Megazyme) was prepared in 50 mM Tris-HCl buffer, 4 mM $CaCl_2$ (pH 8). To start the assay, 40 µL of the stock solution and 5 µL of the sample were mixed with 35 µL of the same buffer. The mixture was incubated at 25°C for 30 min. To end the reaction, 80 µL of DNS were added and incubated at 95°C for 10 min. Afterwards the absorbance at 550 nm was measured.

2.2.13 Cultivation and Expression

The procedure was adapted from "*Pichia* Expression Kit" by invitrogen™.[20]

100 mL of BMGY media in a 2 L baffled flask were inoculated with a single colony. The culture was grown at 28°C and 250 rpm until an OD_{600} of 2 was reached.

Afterwards, the cells were harvested by centrifuging at 3000 x g for 5 minutes at room temperature. The supernatant was discarded, and the cells resuspended in 20 mL of BMMY media. The culture was grown at 28°C and 250 rpm in a 250 mL baffled flask covered with 2 layers of sterile gauze. Once a day a day, methanol was added to the culture to a final concentration of 0.5% (v/v).

2.2.14 Scale up

The procedure was adapted from "Standard Protocol: Screening for Pichia pastoris production strains using enzymatic glucose release in 24-deep-well-plates" by AG Mattanovich and "*Pichia* Expression Kit" by invitrogen™.[20,30]

10 mL of BMGY media in a 100 mL baffled flask were inoculated with a single colony. The culture was grown at 28°C and 250 rpm until an OD_{600} of 2 was reached.

Another 500 mL of BMGY media were prepared in a 2 L baffled flask and inoculated with the 10 mL culture. As soon as the culture reached the same OD_{600}, the cells were harvested by centrifuging at 3000 x g for 5 minutes at room temperature.

The supernatant was discarded, and the cells resuspended in 100 mL of BMMY media. The culture was grown at 28°C and 250 rpm in a 2 L baffled flask covered with 2 layers of sterile gauze. Twice a day, methanol was added to the culture to a final concentration of 0.5% (v/v).

2.2.15 SDS-PAGE, Coomassie blue and silver staining

The composition of the 12% polyacrylamide gel was adapted from the "Handcasting Polyacrylamide Gels Protocol" by Bio-Rad Laboratories, Inc.[76] It is shown in table 5.

Table 5: Composition of 12% polyacrylamide gel

12% Polyacrylamide gel		
Ingredient	Resolving gel	Stacking gel
30% Acrylamide/Bis-acrylamide	6 mL	1.98 mL
Buffer Tris (1.5M, pH 8.8)	3.75 mL	-
Buffer Tris (0.5M, pH 6.8)	-	3.78 mL
10% SDS	150 µL	150 µL
Water	5.03 mL	9 mL
TEMED	7.5 µL	15 µL
APS 10% (w/v)	75 µL	75 µL
Total volume	15 mL	15 mL

The gel was run at 200 V until the dye front has reached the bottom of the gel.

The chamber used was BioRad Mini-PROTEAN tetra cell.

The power supply used was BioRad PowerPac 1000.

The Coomassie blue staining protocol was adapted from protocolsonline.[77]

The gel was fixed in a solution containing 25% isopronanol and 10% acetic acid in water for 60 minutes. Afterwards, the gel was stained in a solution of 10% acetic acid and 60 mg/L Coomasssie Blue R-250. The staining was continued until bands appeared. The gel can be destained in 10% acetic acid.

The silver staining staining protocol was adapted from “Silver Staining Protocol” by ConductScience.[78]

The gel was fixed in solution that contained 40% ethanol, 10% acetic acid and 50% water for 1h. In the following washing step, the gel was placed in water for 1h. Subsequently, the gel was put in a solution containing 0.5% dichromate (DTT) for 30

min. Afterwards, the gel was equilibrated in a 0.1% silver nitrate solution for 30 min. This part of the silver staining procedure had to be done in darkness. In the final part, the gel was washed with water and revealed using a 250 mL solution containing 9 g sodium carbonate and 0.15 mL of formaldehyde (37%). To stop the reaction, 6 grams of acetic acid were given to the solution.

The gel was analyzed using the Gel Doc™ EZ System by Bio-Rad.

2.2.16 Purification via FPLC

The procedure was adapted from "HisTrap™ FF, 1 ml and 5 ml Instructions" by Laboratory Supplies Limited. [64] Buffers, sample and Millipore water were filtered and sonicated before being applied to the column.

The column used was HisTrap FF 1 mL. As a first step it was washed with 3 column volumes of Millipore water at a flow rate of 1 mL/min. The column was then equilibrated with 5 column volumes of binding buffer at the same flow rate. Afterwards the sample was applied using a pump and the column washed until the absorbance reached a steady value. The protein was eluted using a linear gradient of the elution buffer (8 column volumes). The samples were collected in a 96 deep well plate.

3 Results

3.1 Multiplying, purifying and digesting plasmid DNA

The pPIC9K plasmids containing the respective enzymes were transformed into *E. coli* DH5α according to the protocol. Additionally, the original pPIC9K plasmid was transformed as well to have a negative control in future experiments. After plating and growing them overnight, transformants could be found. Single colonies were picked, and glycerol stocks prepared.

To obtain plasmids in greater quantities, one colony of each *E. coli* transformant was picked and grown overnight in 15 mL LB media + ampicillin (100 µg/mL).

The following day, plasmids were extracted with the GeneJET Plasmid Miniprep Kit in 250 µL sterile water. Using nanodrop, the concentrations were determined. The results are shown in table 6:

Table 6: Plasmid concentrations after extracting

Plasmid	Concentration (ng/µL)	Mass (µg)
pPIC9K-subtilisin	180.64	45.16
pPIC9K-trypsin	115.48	28.87
pPIC9K-xylanase	145.72	36.43
pPIC9K	134.97	33.74

30 µg of each DNA sample was digested with *SacI* to eventually obtain His^{+} Mut^{S} transformants of *Pichia pastoris* KM71. Afterwards Phenol/Chloroform extraction was performed, and the DNA precipitated using ethanol. The DNA was resuspended in 15 µL water and analyzed using nanodrop. Results are shown in the table 7:

Table 7: DNA concentrations after digestion, extraction, and precipitation

Digested plasmid	Concentration (ng/µL)	Mass (µg)
pPIC9K-subtilisin	692.28	10.38
pPIC9K-trypsin	847.33	12.71
pPIC9K-xylanase	708.67	10.63
pPIC9K	739.46	11.09

3.2 Transformation and screening of *P. pastoris*

The *Pichia pastoris* cells were prepared and electroporation performed according to the protocol. Afterwards, the aliquots were grown on RDB plates at 30°C.

Two days later, colonies could be found. To screen for multiple inserts, colonies were picked and grown on YPD plates containing geneticin concentrations ranging from 0.25 – 2 mg/mL. After 3 days, pPIC9k-trypsin transformants appeared on plates containing 0.25 – 1.0 mg/mL. Colonies of the remaining transformants could only be found on plates containing 0.25 and 0.5 mg/mL. To confirm the insert, one colony for each respective insert was picked and colony PCR performed. The α-forward and AOX-reverse primer were used. The original plasmids sent by Integrated DNA Technologies, Inc. were utilized as positive control. The samples were analyzed by gel electrophoresis and the results are shown in figure 2.

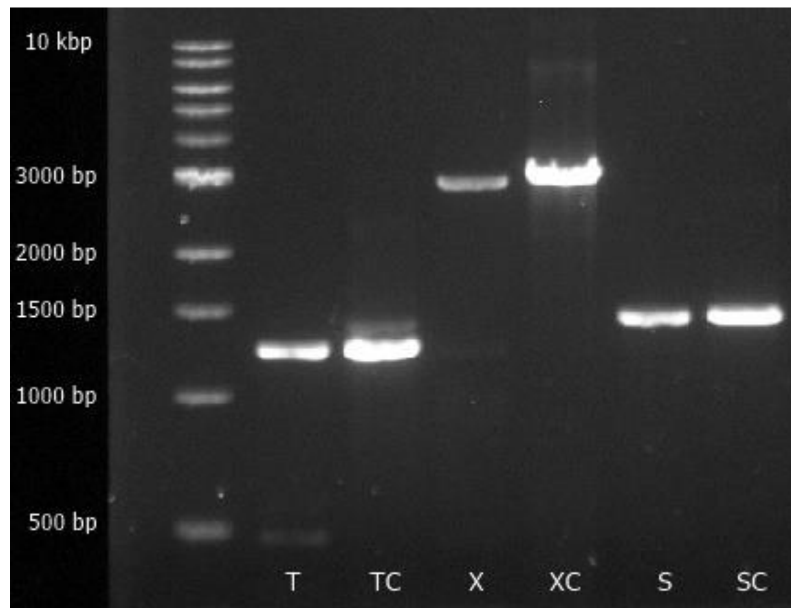


Figure 2: Gel electrophoresis of colony PCR samples (T: trypsin, X: xylanase, S: subtilisin, C: positive control)

The bands line up with their respective positive control and have the correct size (trypsin: 1246 bp, xylanase: 3001 bp, subtilisin: 1405 bp).

Glycerol stocks were prepared and the strains containing the respective insert called *Pichia pastoris* KM71-trypsin, KM71-subtilisin, KM71-xylanase and KM71-pPIC9k.

3.3 Expression

3.3.1 Trypsin

3.3.1.1 Screening transformants for trypsin expression

Multiple strains were tested for the expression of trypsin taken from plates containing different concentrations of geneticin (0.25 – 1 mg/mL).

The first expression was done according to the protocol by Invitrogen™. Methanol was added every 24h to a final concentration of 0.5% and the cultivation was performed at 28°C. A sample was taken every time the culture was fed, and an activity assay performed after all samples were taken.

Protease activity could be observed in all strains tested though it differed remarkably. In figure 3, the results of an activity assays of two trypsin strains are shown. *Pichia pastoris* KM71-Trypsin (1) showed a high resistance (1 mg/mL) while *P. pastoris* KM-71-Trypsin (2) showed low resistance (0.25 mg/mL) against geneticin. As a negative control *P. pastoris* KM71-pPIC9K was used.

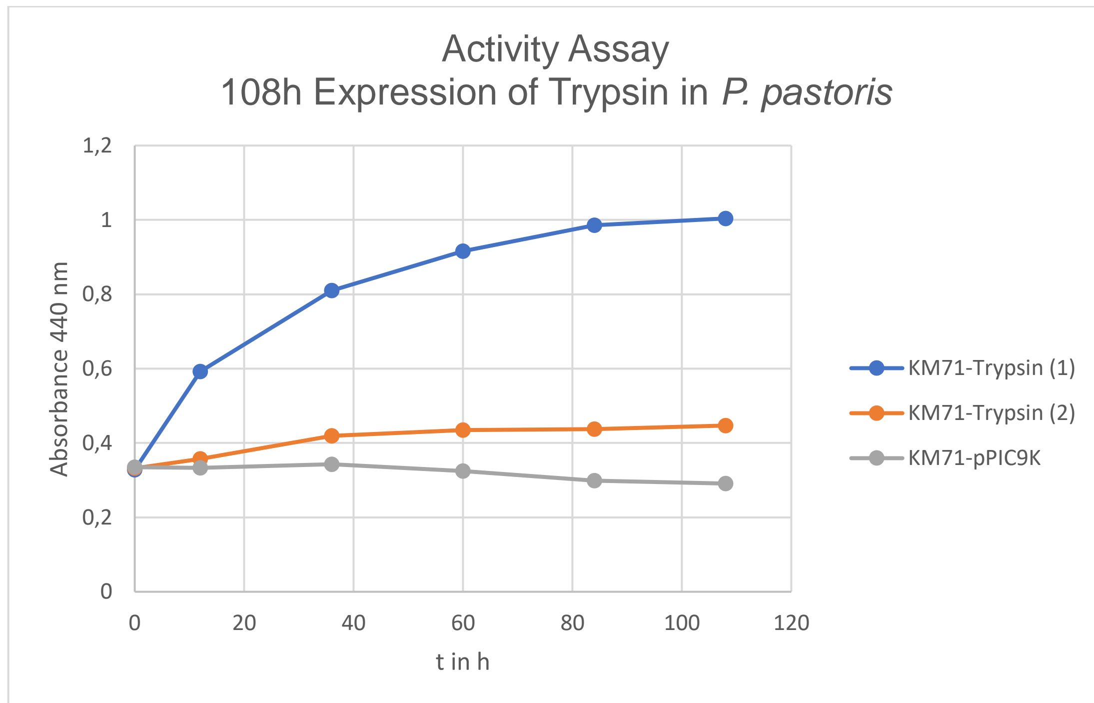


Figure 3: Trypsin expression levels of samples taken from different P. pastoris strains

The absorbance of the samples of *P. pastoris* KM71-Trypsin (1) jumped from 0.328 to 0.592 in the first 12h after expression. It continued to rise but with each sample the jump was smaller. The final value of 1.004 was reached after 108h.

The samples taken from *P. pastoris* KM71-Trypsin (2) showed a smaller jump in absorption after 12h (0.357). It continued to rise until after 60h the absorbance did not change a lot anymore (0.435 – 0.447).

The control strain started around the same value and was almost steady for the first 48h of cultivation (0.335 – 0.343). Afterwards the absorbance sank to a final value of 0.291 at 108h.

The supernatant of *P. pastoris* KM71-trypsin (1) was further analyzed via SDS-PAGE. *P. pastoris* KM71-pPIC9K was used as a negative control. The samples were denaturated and applied to a polyacrylamide gel. Afterwards, Coomassie blue staining was performed. The gel is shown in figure 4.

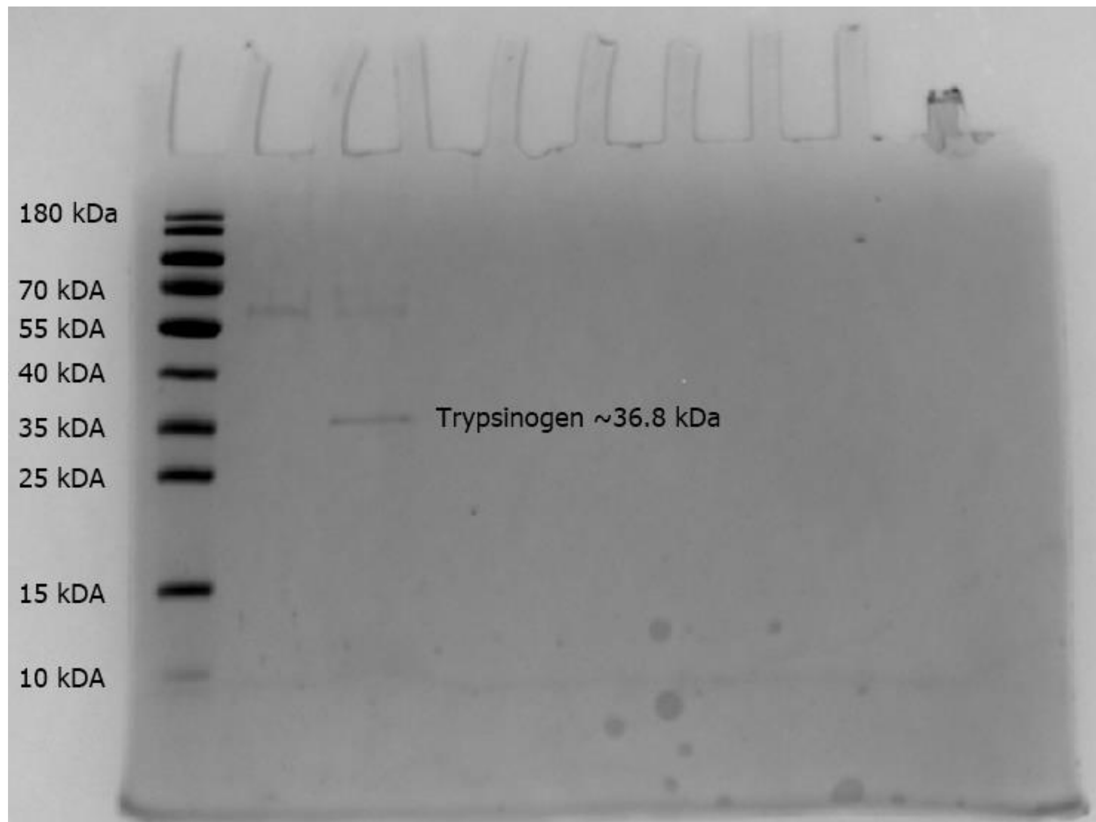


Figure 4: Coomassie blue staining of the transformant's supernatant

The culture of the trypsin transformant (right sample) showed a band that can be assigned to trypsinogen (36.8 kDa). Further bands of higher molecular weight can be seen, too. The negative control (left sample) only showed bands above the size of 55 kDa.

It was decided to analyse the gel with the more sensitive silver staining. The gel is shown in figure 5.

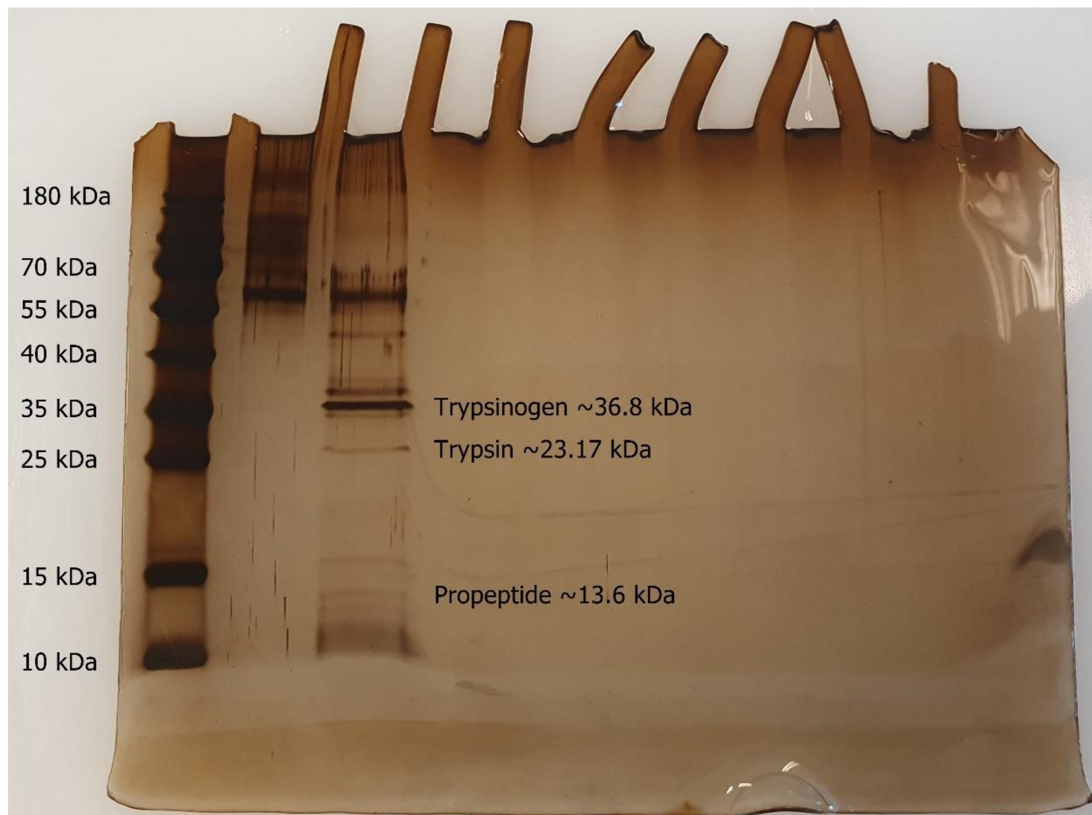


Figure 5: Silver staining of the transformant's supernatant

Silver staining revealed additional bands of higher and lower molecular weight. Two of them could be assigned to trypsin (23.17 kDa) and the propeptide (13.6 kDa). The negative control (left sample) only showed bands above the size of 55 kDa.

3.3.1.2 Enhancing trypsin expression

Different feeding strategies which were adapted from “Standard Protocol: Screening for Pichia pastoris production strains using enzymatic glucose release in 24-deep-well-plates” by AG Mattanovich were tested. Two cultures were fed once a day with methanol to a final concentration of 0.5% and 1%, respectively. The other two concentration were fed twice a day to the same final concentrations. The results of the following activity assay are shown in figure 6:

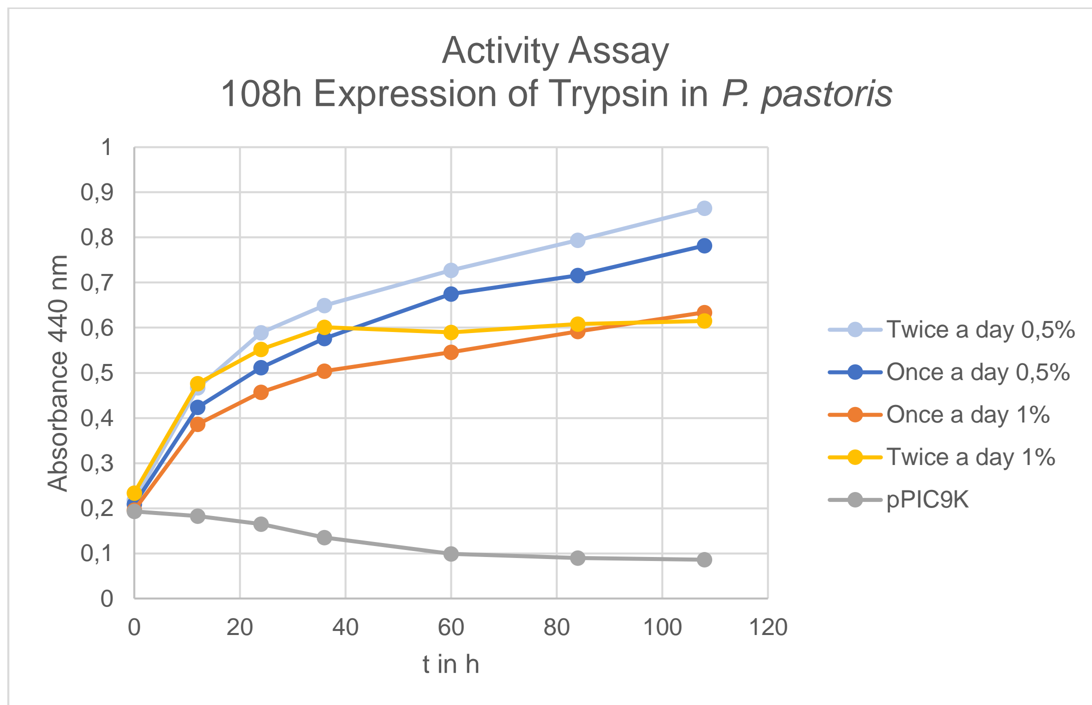


Figure 6: Trypsin expression levels of samples taken from cultures with different feeding strategies

Feeding twice a day to a final concentration of 0.5% resulted in highest expression of trypsin. The cultures of *P. pastoris* KM71-Trpysin (1) all started at similar 0h absorbance values (0.196 – 0.234) and showed a jump in the first 12h (0.386 – 0.476). Afterwards the absorbance continued to rise but less steep until 36h after induction (0.504 – 0.649). In the following hours, the activity of the culture that was fed twice a day to a final concentration of 1% methanol stagnated between 0.590 – 0.615 until the end of the cultivation. The remaining cultures all showed a slower but continuous rise in absorbance until the cultivation ended with maximum values between 0.634 – 0.865.

The absorbance of the negative control started at a similar value (0.193) but declined steadily until a value of 0.086 was reached after 108h.

3.3.2 Xylanase

The xylanase expression was done after the second trypsin expression so feeding strategy that showed best results was chosen (2-times a day, final concentration of 0.5%). Two transformant strains were cultivated and a *Pichia pastoris* strain (clone 6)

producing another xylanase variant was used as a positive control. The expression temperature was lowered (20°C) compared to the one used for trypsin (28°C) since the xylanase gene originates from bacteria from the Antarctic to create a more natural environment.[42]

Unfortunately, no xylanase activity could be measured in the supernatant of in *P. pastoris* KM71-xylanase. The positive control showed xylanase activity. Results of the activity assay are shown in figure 7.

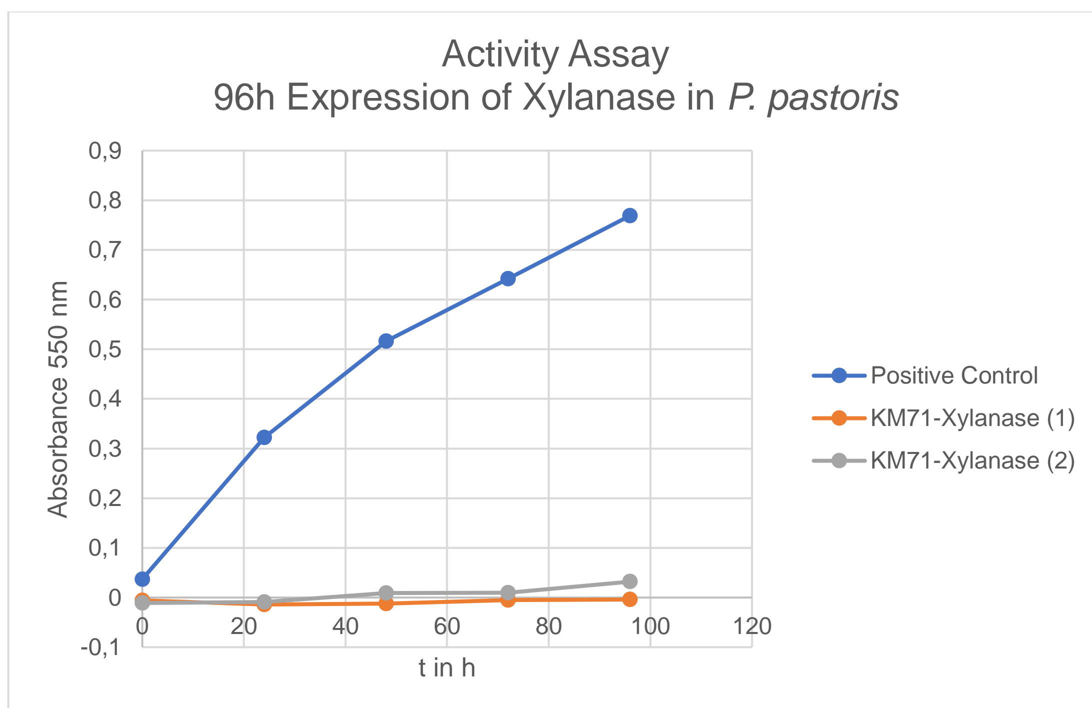


Figure 7: Xylanase expression levels of samples taken from different P. pastoris strains

The absorbance of the positive control started at a value of 0.037 and rose steadily to 1.022 after 168h. For the two transformants absorbance values between -0.014 and 0.032 were observed, indicating no enzyme activity.

The expression was repeated at 28°C and 15°C but no xylanase activity detected.

3.4 Scale up and purification of trypsin

The expression of trypsin was scaled up to a total volume of 100 mL. To continue working with the supernatant, the culture was centrifuged at 3000 x g for 10 min. The supernatant was tested for trypsin activity and the results are shown in figure 8. *Pichia pastoris* KM71-pPIC9K was used as a negative control.

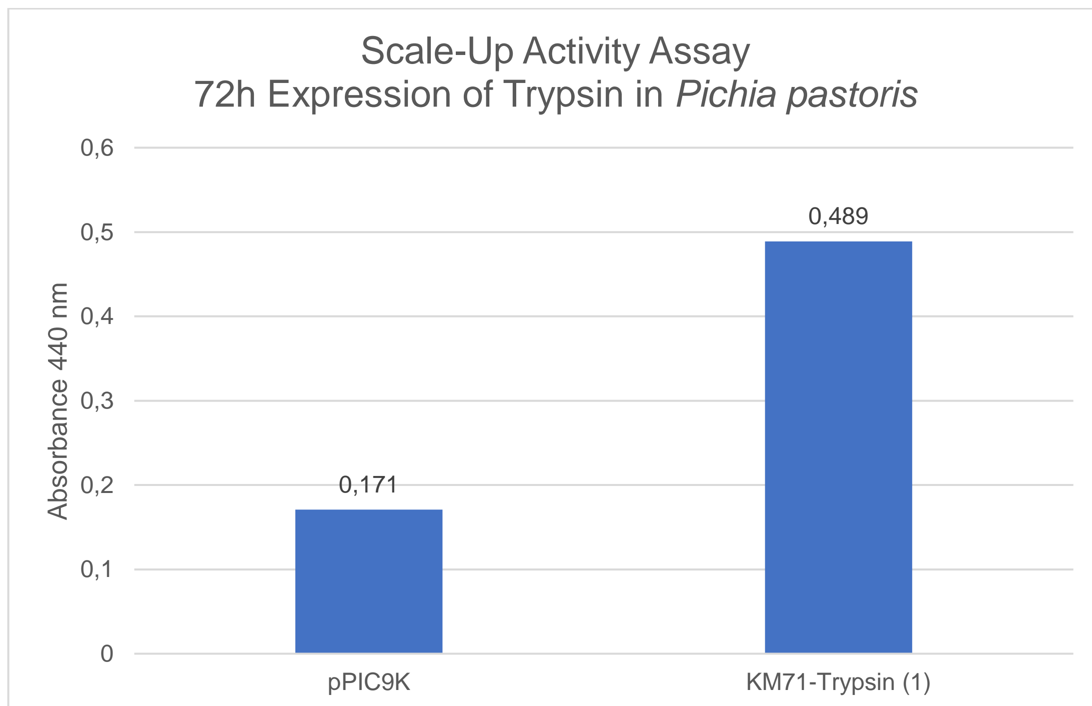


Figure 8: Trypsin expression levels after Scale-Up

The absorbance value (0.489) is not as high as the one observed after 60h – 82h when testing different feeding strategies (0.727 – 0.794; figure 4). Nevertheless, sufficiently high trypsin expression could be observed to continue with the purification.

After ensuring the expression of trypsin in the scale-up culture, purification with FPLC was performed. A HisTrap column was used because the propeptide of trypsin is histidine-tagged. 20 mL of the supernatant were kept for analysis.

The chromatogram of the purification is shown in figure 9. The blue line depicts the UV absorption at 280 nm in mAU (milli-absorbance unit) which provides information about

the protein content. The green line shows the concentration of the elution buffer which was added in a gradient.

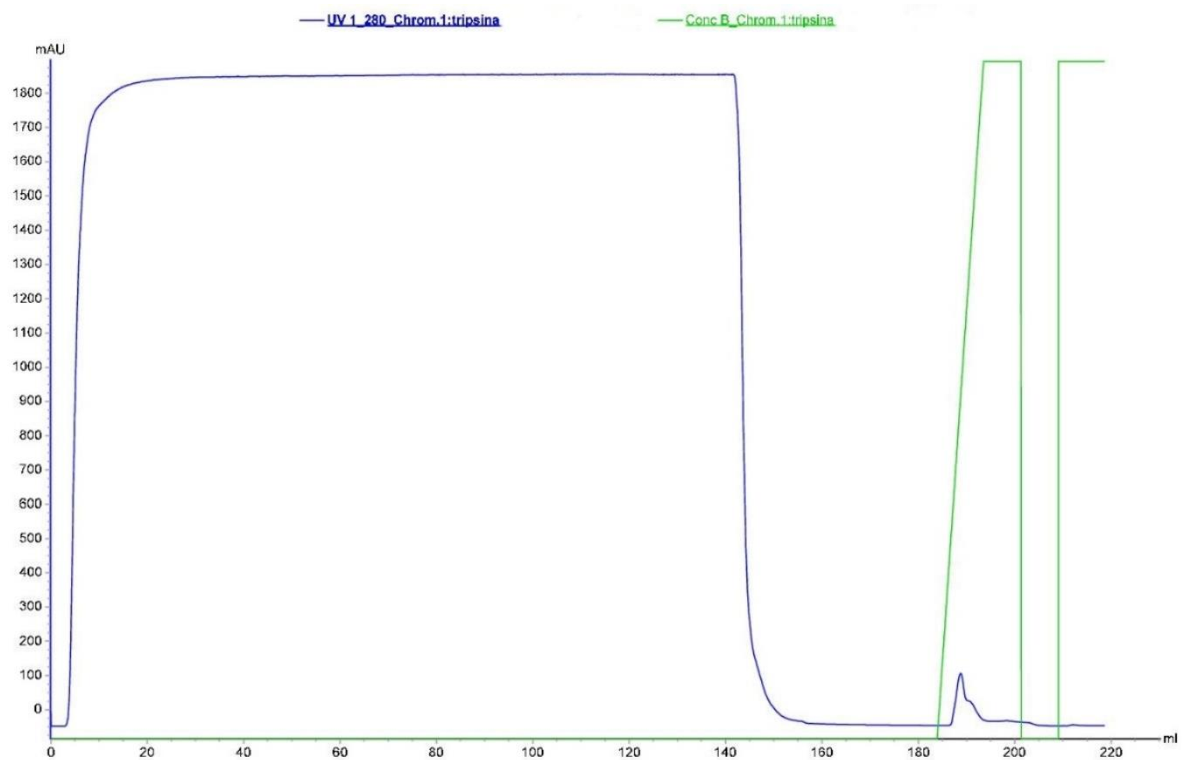


Figure 9: Chromatogram of trypsin purification

When the application of the sample was started (0 mL) the absorption jumped rapidly from -50 mAU to 1800 mAU. During the whole application process (0 - 140 mL) the value did not change. Shortly after the application was done, the absorption fell quickly back to a value of -50 mAU. At around 184 mL the elution was started. Soon after (~187 mL) a small peak could be observed that reached a maximum at 189 mL with a value of ~100 mAU. At around 194 mL the absorbance had reached its initial value of -50 mAU again.

Figure 10 depicts a more detailed image of the absorbance peak during the elution.

Figure 10: Zoom-in of the elution peak (trypsin purification)

The zoomed-in picture of the elution peak provides information about the fractions collected. The fractions that did show absorbance (4 – 10) were pooled to be tested for activity. 20 µL of each fraction were kept to be further analyzed on a polyacrylamid gel. For comparison, the flow-through, original sample a negative control were tested for activity, too. The results are shown in the figure 11.

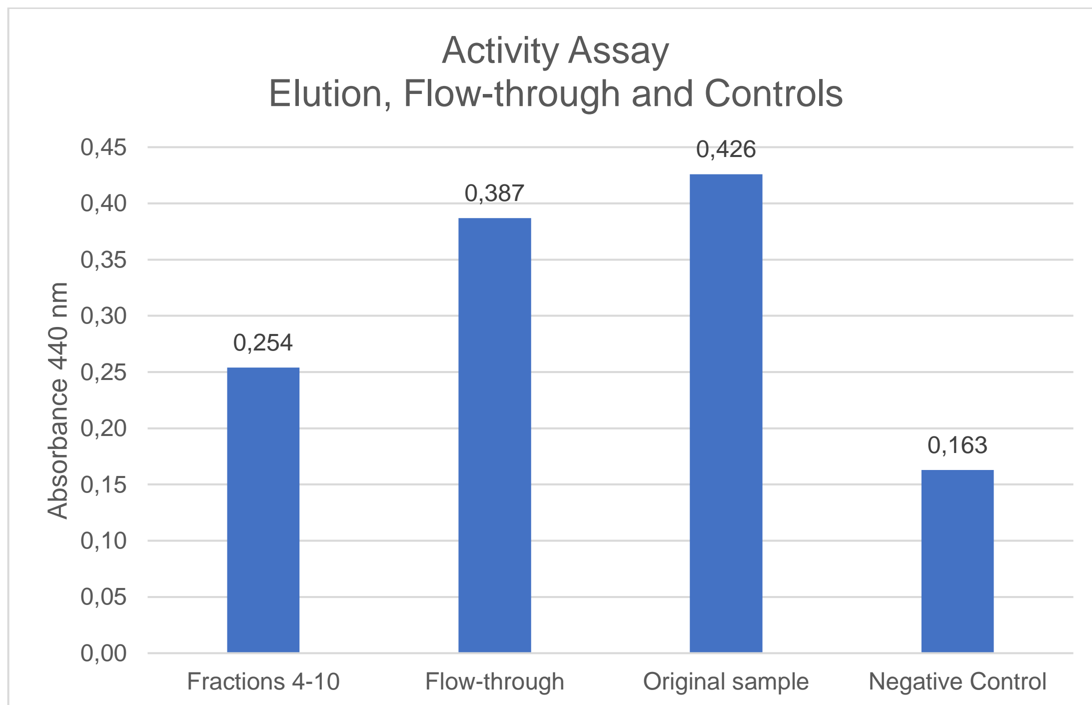


Figure 11: Activity assay of fractions 4-10, flow-through, original sample and negative control

Fractions 4 – 10 only showed very little absorbance (0.254) compared to the flow-through (0.387). The absorbance of the negative control (0.163) and original sample (0.426) were lower but similar to the ones measured before purification (0.171, 0.489; figure 7).

Fractions 1 – 13 were denaturated and applied to a polyacrylamide gel. Afterwards silver staining was performed. The result is shown in figure 12.

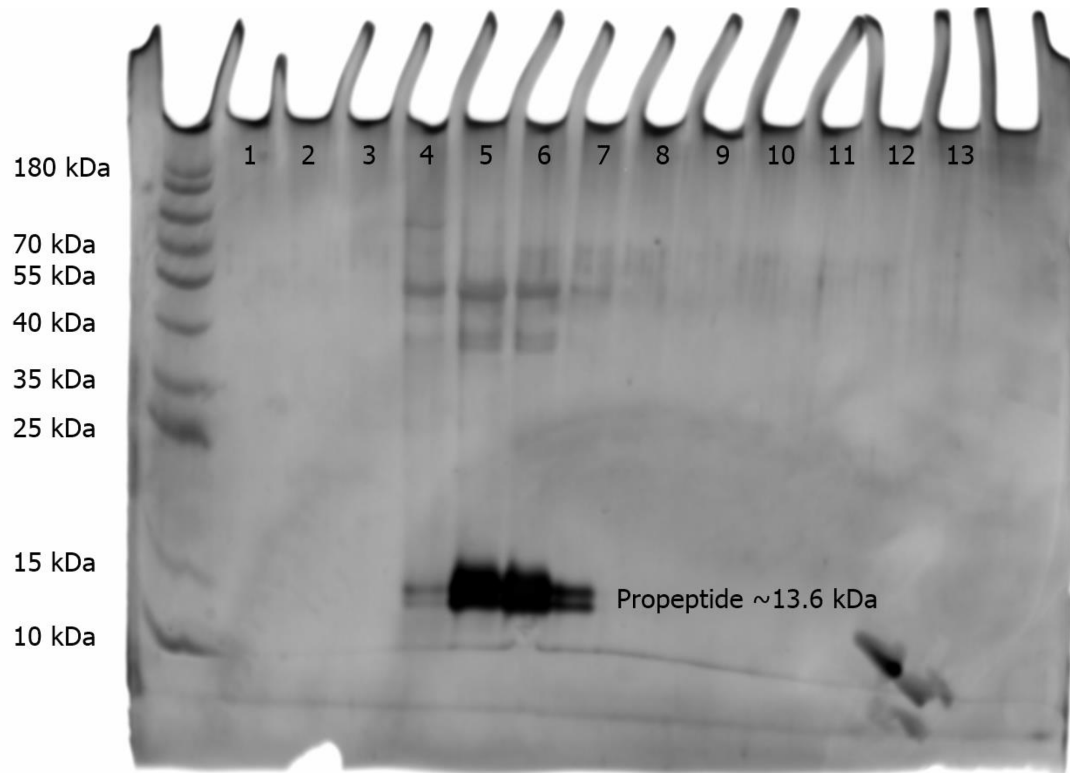


Figure 12: Polyacrylamide gel of fractions 1-13

The gel revealed that great amounts of the propeptide (13.6 kDa) had bound to the column while no band of trypsinogen (36.8 kDa) could be observed. A few further bands appeared after silver staining with molecular weights greater than 40 kDa.

4 Discussion

4.1 Multiplying, purifying and digesting plasmid DNA

The plasmids were amplified without any problems. All pPIC9K plasmids containing the sequence of a respective enzyme could be transformed into *E. coli* DH5α. The $CaCl_2$ and $MgCl_2$ method is one of the most common ways to prepare chemical competent *E. coli* DH5α cells and usually shows good transformation results.[79] The ampicillin resistance encoded in pPIC9K allowed screening for the colonies hosting it.

In the following plasmid extraction and DNA quantification concentrations ranging between 134.97 – 180.64 ng/µL were measured. The concentrations were all in the same magnitude and sufficiently high to continue to digest the plasmids.

After digesting the plasmids with *SacI,* the DNA was extracted and precipitated. Analysing the samples with nanodrop revealed only around a third (10.38 – 12.71 µg) of the originally used DNA (30 µg) was left. This seem like a huge loss, however, phenol/chloroform extraction followed by ethanol precipitation is prone to loss of DNA.[80] The protocol used for the transformation of *P. pastoris* suggested to use 5 – 20 µg of linearized DNA.[20] Therefore, enough DNA was recovered to perform electroporation of the yeast cells.

4.2 Transformation and screening of *P. pastoris*

The electroporation of *Pichia pastoris* was successful and His^+ transformants were found on the RDB plates. *Pichia pastoris* KM71 lacks the *HIS4* gene and is unable to grow on media lacking histidine.[20] The gene encodes for a trifunctional enzyme, catalysing 3 crucial steps in the biosynthesis of the amino acid.[45] The pPIC9K vector hosts the gene so after successful homologous recombination the transformants gain the ability to synthesize histidine themselves. However, screening via histidine does

not give information about possible multiple inserts. Hence, the transformants were pooled and screened on YPD plates containing geneticin with concentrations ranging between 0.25 – 2 mg/mL. The cells could not be screened using geneticin immediately after transformation because they need time to express sufficient amounts of the resistance factor.[46]

A single copy of the plasmid integrated into the genome confers resistance to 0.25 mg/mL geneticin. Thus, greater resistance indicates multiple inserts. Transformants hosting the trypsin sequence were found on plates with concentrations of 0.25 and 0.5 mg/mL. Therefore, it is likely trypsin transformants have at least 2 inserts. Transformants hosting the remaining sequences could only be found on plates containing 0.25 mg/mL, indicating single inserts. Statistically, 1-10% of transformants have multiple inserts.[46] Hence, it is not surprising only very few transformants were found on plates containing a concentration higher than 0.25 mg/mL geneticin.

Single colonies for each respective insert were picked. Via colony PCR the sequences of the enzymes were confirmed inside the genome of the *P. pastoris* strains created. All bands appearing on the gel corresponded to the expected sizes: trypsin (1246 bp), xylanase (3001 bp) and subtilisin (1405 bp).

4.3 Expression of trypsin

In the first part multiple strains showing different levels of geneticin resistance were tested. The azocasein assay that was performed afterwards cannot be used to determine trypsin-specific activity but protease activity in general.[81] However, the polyacrylamide gel (figure 4) which was prepared afterwards confirmed the expression of trypsinogen. Trypsinogen becomes the active form trypsin by cleaving the N-terminal propeptide.[82] The plasmid construct used allows for self-activation of the enzyme.[50] Since bands of trypsinogen, trypsin and the cleaved propeptide could be

seen on the polyacrylamide gel, the measured protease activity can be linked to trypsin.

The negative control of the very first activity assay (figure 3) showed a noticeably higher absorbance (0.335) compared to future assays (0.163 – 0.193; figure 5, 7, 10). Reasons for this could be the expression of an endogenous protease or the occurrence of mistakes when performing the assay. Either way, the result did not affect the subsequent experiments of the project.

The previous assumption of multiple integration events having happened in strains growing on plates containing higher geneticin concentrations seems highly probable. Higher gene copy numbers in *Pichia pastoris* can increase recombinant protein production remarkably[83] and the absorbance in the activity assay of *Pichia pastoris* KM71-Trypsin(1) was a lot higher compared to *Pichia pastoris* KM71-Trypsin(2). However, some cases of decreased protein production after multiple integration are known, including human trypsinogen. A study from 2004 showed, a copy number greater than 2 resulted in decreased trypsinogen expression when using the AOX promotor. It was concluded that accumulation of unfolded intracellular proteins causes an unfolded protein response (UPR) in *Pichia pastoris*.[57] UPR is a signalling network to restore functionality of the endoplasmatic reticulum. If the cell fails to adapt to the stress, apoptosis is induced.[84] Thus, an optimal number rather than maximum number of copies should try to be achieved.[85]

Since time spent on the thesis is limited, it was decided to first try to improve the expression by altering the feeding strategy. Another round of transformation and screening of multiple *P. pastoris* strains would be more time-consuming but is a valid option in case of too low expression levels.

To improve the expression of trypsin, methanol feeding strategies were adapted from "Mattanovich Lab: Metabolic and Cell Engineering" working group at the Universität für Bodenkultur in Vienna. The team has conducted research on *Pichia pastoris* for many years with numerous publications unravelling the mechanisms of recombinant protein production in yeast cells.[27–29] Since they are in close contact with the CeBiB center, they send us the protocol.

The results suggested the feeding strategy could be improved by feeding twice a day to a final concentration of 0.5% methanol. The difference in activity was not huge compared to the old feeding strategy (once a day, 0.5% methanol) but noticeable. Additionally, it was tested to feed the cultures to a final concentration of 1% methanol. Enzyme activity could be observed in both of these cultures though it was lower. Feeding the culture twice a day to final concentration of 1% methanol seemed to be too much to handle for the cells since the activity decreased after 60h. High methanol concentrations can lead to the accumulation of toxic by-products (formaldehyde and peroxide) which is probably the reason for the drop in activity.[86]

The protocol by Invitrogen™ provides other suggestions for the improvement of expression such as scale-up of the culture, expression in Mut^S strains and expressing the enzyme intracellular.[20] Further adaption from the Mattanovich Lab could be the use of a higher OD_{600} (8) and glucose limitation of the main culture before induction.[30] Other sources suggest the co-overexpression of enzymes of the methanol utilization pathway or proteins which improve folding.[87,88]

Since the timeframe for this thesis is limited, it was decided to first attempt to express xylanase and purify trypsin before trying to further improve the expression.

4.4 Expression of xylanase

Unfortunately, no xylanase activity could be observed after multiple attempts at expressing the enzyme. A common problem of the expression of cold-active enzymes mesophilic hosts is the formation of inactive aggregates.[89]

An approach to solve this problem could be the co-expression of proteins improving the cell's folding machinery such as chaperons. Studies have shown that in many popular host systems like *E. coli* and *P. pastoris* the secretion of active enzymes could be improved remarkably.[90,91] Furthermore, it is possible to refold proteins post expression. In this approach a denaturant (e.g., urea or 2-mercaptoethanol) is added and then slowly removed to refold the protein. However, the method can be very time-consuming and result in low yield.[89] A third approach is optimizing the gene codon according to the genetic preference of *P. pastoris*. Multiple codons encode the same amino acid. However, these synonymous codons are not found equally often in genomes. Therefore, every organism has a codon usage bias (CUB).[92] Multiple studies have shown the choice of codon can influence the translational efficiency.[93–95] However, CUB was considered in the design process of the vector.

Unfortunately, there was no time to determine if the enzyme was produced and aggregated or not produced at all. The focus was put on the purification of trypsin instead. However, a group of researchers at the Universität für Bodenkultur in Vienna successfully expressed the xylanase. They used the same temperature for the expression and detected xylanase activity after 48h so their positive results must come down to their *Pichia* strain and expression protocol. The plasmid construct used was similar, however, they expressed the enzyme in *Pichia pastoris* CBS7435. The transformations resulted in Mut^S strain, but the plasmid does not introduce the *His4*

gene. Furthermore, their expression protocol differed. The yeast was grown to a higher OD_{600} (8) on minimal media and glucose limitation was utilized during the induction.

At this point it is not possible to determine if the strain, their protocol, or the combination of both was the reason for their success. Correspondingly, work with the strain that was created as part of this thesis will not be continued.

4.5 Scale up and purification

The expression of trypsin was scaled up without any problems. The subsequent activity assay did show sufficiently high absorbance to link it to the expression of trypsin and no other proteases.

The following purification was done with a His-tag column. The elution peak was relatively small, and the activity assay revealed that most of the protease activity could be found in the flow through. The activity of the original sample was a bit lower than before the purification. This could be due to trypsin being prone to autolysis.[96] The polyacrylamide gel showed very intense bands for the propeptide in the eluted fractions 4-7. In the trypsin sequence utilized, the His-tag is part of the propeptide. Therefore, no band of active trypsin could be seen. Further bands that can be seen on the gel are most likely other native proteins expressed by *Pichia pastoris* which are rich in histidine. The band at around 40 kDa could potentially be trypsinogen (36.8 kDa). However, when comparing it to the gel after the first expression (figure 4), the trypsinogen band seems to be lower and closer to the 35 kDa marker band. Furthermore, the low protease activity suggests the absence of detectable amounts of trypsinogen in the elution fractions. A common contaminant when purifying His-tagged proteins from *Pichia pastoris* is mitochondrial alcohol dehydrogenase isozyme III (mADH) which has a similar molecular weight of ~40 kDa.[97,98] The band that can be seen around 40 kDa could therefore be mADH.

It can be concluded that the purification of trypsin utilizing the histidine tag is improper. After cultivation a considerable amount of trypsinogen already converted into active trypsin and propeptide what leaves trypsin in the flow through. The histidine tag being part of the propeptide was expected to be a possible problem in the purification process. But since the propeptide construct worked well in the expression of *Streptomyces griseus* trypsin in *Pichia pastoris*[50], it was decided to not alter the sequence for this project. Another problem is trypsinogen not binding to the column. A possible explanation could be that the tertiary structure of the protein occludes the histidine tag. This problem could be fixed by denaturing trypsin with urea or guanidium hydrochloride to make the histidine tag accessible.[99] The enzyme could after purification be refolded by dialyzing away the denaturants.[100] However, this would not fix the problem of active trypsin in the flow through. A better option would be the use of benzamidine-sepharose as stationary phase. The technique has been proven to be an efficient purification strategy for trypsin.[101] Benzamidine is an aromatic amidine which can inhibit serine proteases and bind them reversible.[102]

5 Summary and outlook

The aim of this work was to introduce the genes of 3 enzymes (subtilisin, trypsin and xylanase) in *Pichia pastoris* KM71. The sequences of these enzymes all derive from bacteria isolated from extreme environments in Chile, namely the Atacama Desert and Chilean Antarctica. For this project the vector pPIC9K was utilized which had showed great results in previous studies with *Pichia pastoris*.[50,103,104] It contains the α-secretion factor, AOX promotor, two antibiotic resistances (kanamycin and ampicillin) and multiple restriction sites. When the vector is used for transformation of *P. pastoris* KM71 it results in His^+ and Mut^S strains.

The transformation was performed via electroporation. Colonies were first screened on medium lacking histidine and afterwards on plates containing different concentrations of geneticin. The following colony PCR confirmed that all sequences were successfully introduced into 3 novel *Pichia pastoris* strains. The expression of trypsin was an immediate success. Multiple trypsin transformants were tested and strains that had showed a higher resistance against geneticin showed greater results in the following activity assay. Hence, greater resistance against the antibiotic implies the integration of multiple inserts and higher enzyme production. The expression was then tried to be improved by altering the methanol feeding strategy. The results suggest feeding twice a day to final concentration of 0.5% is the best of the ones tested though further improvements can be made. In the final part of the project the purification of trypsin was attempted. Unfortunately, only the cleaved propeptide bound to the column since it contains the His-Tag that was utilized. The zymogen trypsinogen should have bounded as well so it is likely the tertiary structure of trypsinogen occludes the His-tag. Multiple attempts to express xylanase failed and there was no time left to attempt the expression of subtilisin.

Work on the purification of trypsin will be continued. A good option would be the use of column containing benzamidine-sepharose. Benzamidine is an aromatic amidine which can inhibit and bind trypsin reversible.[102] A group of researchers from Vienna which are in close contact with the CeBiB managed to express xylanase in a different *Pichia pastoris* strain with a different expression protocol and plasmid construct. Therefore, work with the xylanase strain developed as part of this thesis will not be continued. However, their protocol will be used to improve the expression of trypsin and attempt to produce subtilisin.

6 Sources

1. Rampelotto, P. H. Extremophiles and Extreme Environments. *Life : Open Access Journal* **3**, 482 (2013).

2. Santos, H. & da Costa, M. S. Compatible solutes of organisms that live in hot saline environments. *Environ Microbiol* **4**, 501–509 (2002).

3. Turner, P., Mamo, G. & Karlsson, E. N. Potential and utilization of thermophiles and thermostable enzymes in biorefining. *Microb Cell Fact* **6**, 1–23 (2007).

4. Orellana, R. *et al.* Living at the Frontiers of Life: Extremophiles in Chile and Their Potential for Bioremediation. *Front Microbiol* **9**, 2309 (2018).

5. Gómez-Silva, B. & Batista-García, R. A. The Atacama Desert: A Biodiversity Hotspot and Not Just a Mineral-Rich Region. *Front Microbiol* **13**, 23 (2022).

6. Rateb, M. E. *et al.* Chaxamycins A–D, Bioactive Ansamycins from a Hyper-arid Desert Streptomyces sp. *J Nat Prod* **74**, 1491–1499 (2011).

7. Contreras, F., Amenabar, M. J. & Blamey, J. M. Purification and characterization of two thermostable xylanases from a halotolerant Bacillus sp. Asc6BA isolated from Salar de Ascotán, Atacama Desert. *Extremophiles* **25**, 51–59 (2021).

8. Galetovic, A. *et al.* Use of Phycobiliproteins from Atacama Cyanobacteria as Food Colorants in a Dairy Beverage Prototype. *Foods 2020, Vol. 9, Page 244* **9**, 244 (2020).

9. Núñez-Montero, K. *et al.* Antarctic Rahnella inusitata: A Producer of Cold-Stable β-Galactosidase Enzymes. *International Journal of Molecular Sciences 2021, Vol. 22, Page 4144* **22**, 4144 (2021).

10. Kazuoka, T., Oikawa, T., Muraoka, I., Kuroda, S. & Soda, K. A cold-active and thermostable alcohol dehydrogenase of a psychrotorelant from Antarctic seawater, Flavobacterium frigidimaris KUC-1. *Extremophiles* **11**, 257–267 (2007).

11. Rigoldi, F., Donini, S., Redaelli, A., Parisini, E. & Gautieri, A. Review: Engineering of thermostable enzymes for industrial applications. *APL Bioengineering* vol. 2 Preprint at https://doi.org/10.1063/1.4997367 (2018).

12. Precedence Research. Industrial Enzymes Market Size to Hit USD 11.02 Bn by 2030. https://www.precedenceresearch.com/industrial-enzymes-market (2021).

13. Bazan-Peregrino, M. *et al.* Combining virotherapy and angiotherapy for the treatment of breast cancer. *Cancer Gene Ther* **20**, 461–468 (2013).

14. Venter, M. Synthetic promoters: genetic control through cis engineering. *Trends Plant Sci* **12**, 118–124 (2007).

15. Khan, S. *et al.* Role of Recombinant DNA Technology to Improve Life. *Int J Genomics* **2016**, (2016).

16. Gu, P. *et al.* A rapid and reliable strategy for chromosomal integration of gene(s) with multiple copies. *Scientific Reports 2015 5:1* **5**, 1–9 (2015).

17. da Silva, N. A. & Srikrishnan, S. Introduction and expression of genes for metabolic engineering applications in Saccharomyces cerevisiae. *FEMS Yeast Res* **12**, 197–214 (2012).

18. Kuzminov, A. Homologous Recombination—Experimental Systems, Analysis and Significance. *EcoSal Plus* **4**, (2011).

19. Orr-Weaver, T. L., Szostak, J. W. & Rothstein, R. J. Yeast transformation: a model system for the study of recombination. *Proceedings of the National Academy of Sciences* **78**, 6354–6358 (1981).

20. invitrogen. *Pichia Expression Kit.* https://tools.thermofisher.com/content/sfs/manuals/pich_man.pdf (2014).

21. Pardo, B., Gómez-González, B. & Aguilera, A. DNA Repair in Mammalian Cells. *Cellular and Molecular Life Sciences* **66**, 1039–1056 (2009).

22. Gratz, S. J. *et al.* Highly Specific and Efficient CRISPR/Cas9-Catalyzed Homology-Directed Repair in Drosophila. *Genetics* **196**, 961–971 (2014).

23. Peña, D. A., Gasser, B., Zanghellini, J., Steiger, M. G. & Mattanovich, D. Metabolic engineering of Pichia pastoris. *Metab Eng* **50**, 2–15 (2018).

24. Inan, M. & Meagher, M. M. Non-repressing carbon sources for alcohol oxidase (AOX1) promoter of Pichia pastoris. *J Biosci Bioeng* **92**, 585–589 (2001).

25. Liu, W. C. *et al.* Fed-batch high-cell-density fermentation strategies for Pichia pastoris growth and production. *https://doi.org/10.1080/07388551.2018.1554620* **39**, 258–271 (2019).

26. Looser, V. *et al.* Cultivation strategies to enhance productivity of Pichia pastoris: A review. *Biotechnol Adv* **33**, 1177–1193 (2015).

27. Mattanovich, D. *et al.* Recombinant protein production in yeasts. *Methods Mol Biol* **824**, 329–358 (2012).

28. Peña, D. A., Gasser, B., Zanghellini, J., Steiger, M. G. & Mattanovich, D. Metabolic engineering of Pichia pastoris. *Metab Eng* **50**, 2–15 (2018).

29. Zahrl, R. J., Gasser, B., Mattanovich, D. & Ferrer, P. Detection and Elimination of Cellular Bottlenecks in Protein-Producing Yeasts. *Methods Mol Biol* **1923**, 75–95 (2019).

30. AG Mattanovich. *Standard Protocol: Screening for Pichia pastoris production strains using enzymatic glucose release in 24-deep-well-plates.* (2019).

31. Labrou, N. E. Protein purifi cation: An overview. *Methods in Molecular Biology* **1129**, 3–10 (2014).

32. Kallberg, K., Johansson, H. O. & Bulow, L. Multimodal chromatography: An efficient tool in downstream processing of proteins. *Biotechnol J* **7**, 1485–1495 (2012).

33. Chon, J. H. & Zarbis-Papastoitsis, G. Advances in the production and downstream processing of antibodies. *N Biotechnol* **28**, 458–463 (2011).

34. Rosa, P. A. J., Azevedo, A. M., Sommerfeld, S., Bäcker, W. & Aires-Barros, M. R. Aqueous two-phase extraction as a platform in the biomanufacturing industry: economical and environmental sustainability. *Biotechnol Adv* **29**, 559–567 (2011).

35. Kalyanpur, M. Downstream processing in the biotechnology industry. *Mol Biotechnol* **22**, 87–98 (2002).

36. Wingfield, P. T. Overview of the Purification of Recombinant Proteins Produced in Escherichia coli. *Curr Protoc Protein Sci* **30**, 6.1.1-6.1.37 (2002).

37. Roque, A. C. A., Silva, C. S. O. & Taipa, M. Â. Affinity-based methodologies and ligands for antibody purification: advances and perspectives. *J Chromatogr A* **1160**, 44–55 (2007).

38. Falke, J. J. & Corbin, J. A. Methods | Affinity Tags for Protein Purification. *Encyclopedia of Biological Chemistry: Third Edition* **3**, 701–705 (2013).

39. Corona, G., Skaper, S. D. & Callegaro, L. Purification of Different Biologically Active Forms of Mouse β-Nerve Growth Factor by Immobilized Metal Ion Affinity Chromatography. *Techniques in Protein Chemistry IV* 343–351 (1993) doi:10.1016/B978-0-12-058757-5.50042-1.

40. el Hajj, S. *et al.* Application in nutrition: mineral binding. *Biologically Active Peptides: From Basic Science to Applications for Human Health* 455–494 (2021) doi:10.1016/B978-0-12-821389-6.00016-9.

41. Hengen, P. N. Purification of His-Tag fusion proteins from Escherichia coli. *Trends Biochem Sci* **20**, 285–286 (1995).

42. Centro de Biotecnología y Bioingeniería. *CeBiB's enzyme portfolio: Cold-active enzymes isolated from the Antarctica*. (2022).

43. Chang, C. H., Hsiung, H. A., Hong, K. L. & Huang, C. T. Enhancing the efficiency of the Pichia pastoris AOX1 promoter via the synthetic positive feedback circuit of transcription factor Mxr1. *BMC Biotechnol* **18**, (2018).

44. Cregg, J. M., Madden, K. R., Barringer, K. J., Thill, G. P. & Stillman, C. A. Functional characterization of the two alcohol oxidase genes from the yeast Pichia pastoris. *Mol Cell Biol* **9**, 1316 (1989).

45. Crane, D. I. & Gould, S. J. The Pichia pastoris HIS4 gene: nucleotide sequence, creation of a non-reverting his4 deletion mutant, and development of HIS4-based replicating and integrating plasmids. *Curr Genet* **26**, 443–450 (1994).

46. invitrogen. *pPIC9K - A Pichia Vector for Multicopy Integration and Secreted Expression*. https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FLSG%2Fmanuals%2Fppic9k_man.pdf (2010).

47. Mechri, S. *et al.* Cloning and heterologous expression of subtilisin SAPN, a serine alkaline protease from Melghiribacillus thermohalophilus Nari2AT in Escherichia coli and Pichia pastoris. *Process Biochemistry* **105**, 27–41 (2021).

48. Trop, M. & Birk, Y. The specificity of proteinases from Streptomyces griseus (pronase). *Biochemical Journal* **116**, 19 (1970).

49. Olafson, R. W., Jurášek, L., Carpenter, M. R. & Smillie, L. B. Amino acid sequence of Streptomyces griseus trypsin. Cyanogen bromide fragments and complete sequence. *Biochemistry* **14**, 1168–1177 (1975).

50. Zhang, Y., Ling, Z., Du, G., Chen, J. & Kang, Z. Improved Production of Active Streptomyces griseus Trypsin with a Novel Auto-Catalyzed Strategy. *Scientific Reports 2016 6:1* **6**, 1–11 (2016).

51. Rawlings, N. D. & Salvesen, G. Handbook of Proteolytic Enzymes. *Handbook of Proteolytic Enzymes* **1–3**, (2013).

52. Vasquez, J. R., Evnin, L. B., Higaki, J. N. & Craik, C. S. An expression system for trypsin. *J Cell Biochem* **39**, 265–276 (1989).

53. Kim, N. S. *et al.* Production of functional recombinant bovine trypsin in transgenic rice cell suspension cultures. *Protein Expr Purif* **76**, 121–126 (2011).

54. Craik, C. S. *et al.* Redesigning trypsin: alteration of substrate specificity. *Science* **228**, 291–297 (1985).

55. Ling, Z. *et al.* Functional expression of trypsin from Streptomyces griseus by Pichia pastoris. *J Ind Microbiol Biotechnol* **39**, 1651–1662 (2012).

56. Zhao, M., Wu, F. & Xu, P. Development of a rapid high-efficiency scalable process for acetylated Sus scrofa cationic trypsin production from Escherichia coli inclusion bodies. *Protein Expr Purif* **116**, 120–126 (2015).

57. Hohenblum, H., Gasser, B., Maurer, M., Borth, N. & Mattanovich, D. Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant Pichia pastoris. *Biotechnol Bioeng* **85**, 367–375 (2004).

58. Parra, L. P. *et al.* Cloning and fusion expression of a cold-active lipase from marine Antarctic origin. *Enzyme Microb Technol* **42**, 371–377 (2008).

59. Dälken, B., Jabulowsky, R. A., Oberoi, P., Benhar, I. & Wels, W. S. Maltose-Binding Protein Enhances Secretion of Recombinant Human Granzyme B Accompanied by In Vivo Processing of a Precursor MBP Fusion Protein. *PLoS One* **5**, e14404 (2010).

60. Abou Hachem, M. *et al.* Carbohydrate-binding modules from a thermostable Rhodothermus marinus xylanase: cloning, expression and binding studies. *Biochemical Journal* **345**, 53 (2000).

61. Li, J. F. *et al.* Modified pPIC9K vector-mediated expression of a family 11 xylanase gene, Aoxyn11A, from Aspergillus oryzae in Pichia pastoris. *Ann Microbiol* **63**, 1109–1120 (2013).

62. Damaso, M. C. T. *et al.* Optimized Expression of a Thermostable Xylanase from Thermomyces lanuginosus in Pichia pastoris. *Appl Environ Microbiol* **69**, 6064–6072 (2003).

63. AAT Bioquest. Tris Buffer Preparation and Recipe. https://www.aatbio.com/resources/buffer-preparations-and-recipes/tris-buffer.

64. Scientific Laboratory Supplies Limited. HisTrap™ FF, 1 ml and 5 ml Instructions. https://www.scientificlabs.co.uk/handlers/libraryFiles.ashx?filename=Manuals_1_17525501_A.PDF (2014).

65. AAT Bioquest. SDS-PAGE SDS Running Buffer (10x) Preparation and Recipe. https://www.aatbio.com/resources/buffer-preparations-and-recipes/sds-page-10x-sds-running-buffer.

66. Kostylev, M., Otwell, A. E., Richardson, R. E. & Suzuki, Y. Cloning Should Be Simple: Escherichia coli DH5α-Mediated Assembly of Multiple DNA Fragments with Short End Homologies. *PLoS One* **10**, e0137466 (2015).

67. Salamin, K., Sriranganadane, D., Léchenne, B., Jousson, O. & Monod, M. Secretion of an endogenous subtilisin by Pichia pastoris strains GS115 and KM71. *Appl Environ Microbiol* **76**, 4269–4276 (2010).

68. Das, S. & Dash, H. R. Cloning and Transformation. in *Microbial Biotechnology- A Laboratory Manual for Bacterial Systems* 35–72 (Springer India, 2015). doi:10.1007/978-81-322-2095-4_2.

69. Sambrook, J., Fritsch, E. F. & Maniatis, T. Molecular cloning: a laboratory manual. *Molecular cloning: a laboratory manual.* (1989).

70. Johnston, C., Martin, B., Fichant, G., Polard, P. & Claverys, J. P. Bacterial transformation: Distribution, shared mechanisms and divergent control. *Nat Rev Microbiol* **12**, 181–196 (2014).

71. New England Biolabs. SacI. https://international.neb.com/products/r0156-saci#Product%20Information.

72. Thermo Fisher Scientific Inc. How to Use Phenol / Chloroform for DNA Purification | Thermo Fisher Scientific - CL. https://www.thermofisher.com/cl/es/home/references/protocols/nucleic-acid-purification-and-analysis/dna-extraction-protocols/phenol-chloroform-extraction.html.

73. Royle, K. E. & Polizzi, K. A streamlined cloning workflow minimising the time-to-strain pipeline for Pichia pastoris. *Sci Rep* **7**, (2017).

74. Iversen, S. L. & Jørgensen, M. H. Azocasein assay for alkaline protease in complex fermentation broth. *Biotechnology Techniques 1995 9:8* **9**, 573–576 (1995).

75. Sanhueza, C., Carvajal, G., Soto-Aguilar, J., Lienqueo, M. E. & Salazar, O. The effect of a lytic polysaccharide monooxygenase and a xylanase from Gloeophyllum trabeum on the enzymatic hydrolysis of lignocellulosic residues using a commercial cellulase. *Enzyme Microb Technol* **113**, 75–82 (2018).

76. Bio-Rad. Handcasting Polyacrylamide Gels Protocol. https://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_6201.pdf (2014).

77. Coomassie | Protocols Online. https://www.protocolsonline.com/protein-science/coomassie/.

78. Conduct Science. Silver Staining Protocol. https://conductscience.com/silver-staining-protocol/.

79. Chan, W., Verma, C. S., Lane, D. P. & Gan, S. K. A comparison and optimization of methods and factors affecting the transformation of Escherichia coli. *Biosci Rep* **33**, 931–937 (2013).

80. Mygind, T., Østergaard, L., Birkelund, S., Lindholt, J. S. & Christiansen, G. Evaluation of five DNA extraction methods for purification of DNA from atherosclerotic tissue and estimation of prevalence of Chlamydia pneumoniae in tissue from a Danish population undergoing vascular repair. *BMC Microbiol* **3**, 1–12 (2003).

81. Côlho, D. F. *et al.* Azocasein Substrate for Determination of Proteolytic Activity: Reexamining a Traditional Method Using Bromelain Samples. *Biomed Res Int* **2016**, (2016).

82. Briinger, A. T., Huber, R. & Karplus, M. Trypsinogen-Trypsin Transition: A Molecular Dynamics Study of Induced Conformational Change in the Activation Domain. *Biochemistry* **26**, 5153–5162 (1987).

83. Mansur, M. *et al.* Multiple gene copy number enhances insulin precursor secretion in the yeast Pichia pastoris. *Biotechnol Lett* **27**, 339–345 (2005).

84. Hetz, C. The unfolded protein response: controlling cell fate decisions under ER stress and beyond. *Nature Reviews Molecular Cell Biology 2012 13:2* **13**, 89–102 (2012).

85. Zhu, T. *et al.* Efficient generation of multi-copy strains for optimizing secretory expression of porcine insulin precursor in yeast Pichia pastoris. *J Appl Microbiol* **107**, 954–963 (2009).

86. Zhang, W., Bevins, M. A., Plantz, B. A., Smith, L. A. & Meagher, M. M. Modeling Pichia pastoris Growth on Methanol and Optimizing the Production of a

Recombinant Protein, the Heavy-Chain Fragment C of Botulinum Neurotoxin, Serotype A. *Biotechnol Bioeng* **70**, 1–8 (2000).

87. Damasceno, L. M., Huang, C. J. & Batt, C. A. Protein secretion in Pichia pastoris and advances in protein production. *Appl Microbiol Biotechnol* **93**, 31–39 (2012).

88. Krainer, F. W. *et al.* Recombinant protein expression in Pichia pastoris strains with an engineered methanol utilization pathway. *Microb Cell Fact* **11**, 1–14 (2012).

89. Söderberg, J. J., Grgic, M., Hjerde, E. & Haugen, P. Aliivibrio wodanis as a production host: Development of genetic tools for expression of cold-active enzymes. *Microb Cell Fact* **18**, 1–16 (2019).

90. Georgiou, G. & Valax, P. Expression of correctly folded proteins in Escherichia coli. *Curr Opin Biotechnol* **7**, 190–197 (1996).

91. He, H. *et al.* A Combinational Strategy for Effective Heterologous Production of Functional Human Lysozyme in Pichia pastoris. *Front Bioeng Biotechnol* **8**, 118 (2020).

92. Huang, Y. *et al.* Codon pair optimization (CPO): a software tool for synthetic gene design based on codon pair bias to improve the expression of recombinant proteins in Pichia pastoris. *Microb Cell Fact* **20**, 1–10 (2021).

93. Tats, A., Tenson, T. & Remm, M. Preferred and avoided codon pairs in three domains of life. *BMC Genom* **9**, 463 (2008).

94. Baeza, M., Alcaíno, J., Barahona, S., Sepúlveda, D. & Cifuentes, V. Codon usage and codon context bias in Xanthophyllomyces dendrorhous. *BMC Genom* **16**, 293 (2015).

95. Prabha, R., Singh, D. P., Sinha, S., Ahmad, K. & Rai, A. Genome-wide comparative analysis of codon usage bias and codon context patterns among cyanobacterial genomes. *Mar Genom* **32**, 31–39 (2017).

96. Smolin, D. *et al.* Accelerated trypsin autolysis by affinity polymer templates. *RSC Adv* **10**, 28711–28719 (2020).

97. Chen, Y., Li, Y., Liu, P., Sun, Q. & Liu, Z. Optimized expression in Pichia pastoris eliminates common protein contaminants from subsequent His-tag purification. *Biotechnol Lett* **36**, 711–718 (2014).

98. ADH3p Mitochondrial alcohol dehydrogenase isozyme III - Saccharomyces boulardii (nom. inval.) | UniProtKB | UniProt. https://www.uniprot.org/uniprotkb/A0A0L8VJQ7/entry.

99. Bornhorst, J. A. & Falke, J. J. [16] Purification of Proteins Using Polyhistidine Affinity Tags. *Methods Enzymol* **326**, 245 (2000).

100. Hermodson, M. Current protocols in protein science. *Proteins: Structure, Function, and Bioinformatics* **24**, 409–409 (1996).

101. Chamrád, I., Strouhal, O., Řehulka, P., Lenobel, R. & Šebela, M. Microscale affinity purification of trypsin reduces background peptides in matrix-assisted laser desorption/ionization mass spectrometry of protein digests. *J Proteomics* **74**, 948–957 (2011).

102. Tanizawa, K., Ishii, S. I., Hamaguchi, K. & Kanaoka, Y. Proteolytic EnzymesVI. Aromatic Amidines as Competitive Inhibitors of Trypsin. *The Journal of Biochemistry* **69**, 893–899 (1971).

103. Xie, T., Liu, Q., Xie, F., Liu, H. & Zhang, Y. Secretory Expression of Insulin Precursor in Pichia pastoris and Simple Procedure for Producing Recombinant Human Insulin. *http://dx.doi.org/10.1080/10826060802165147* **38**, 308–317 (2008).

104. Eissazadeh, S. *et al.* Production of recombinant human epidermal growth factor in Pichia pastoris. *Brazilian Journal of Microbiology* **48**, 286–293 (2017).

7 Appendix

7.1 Chemicals

Acrylamide/Bis-acrylamide	B-0100 Winkler
Acetic acid	131008.1612 PanReac
Agar, Bacto	2145430 Difco
APS 10% (w/v)	15523-012 Gibco-BRL
Azocasein	A2765-5G Sigma Aldrich
Biotin	B4501-1G Sigma Aldrich
$CaCl_2$	CA-0520 Winkler LTDA
Coomassie blue R-250	1.12553 Sigma Aldrich
Dithiothreitol (DTT)	BM0665 Winkler
DMSO	1.02931.1000 Merck
dNTPs	N0447S NEB
Formaldehyde (37%)	BM-0780 Winkler
GC Buffer 5x	B0519S NEB
Geneticin	SC-29064A Enco
Glucose	1.08337.1000 Merck
Glycerol	BM-0820 Winkler
Glycine	B0820 Winkler
HCl	00172 Loba Chemie

Hepes	H3375 Merck
Imidazole	1.04716.0250 Merck
Isopropanol	563935 Sigma Aldrich
KH_2PO_4	131509.1211 PanReac
K_2HPO_4	131512.1211 PanReac
KOH	1.05033.1000 Emsure
L-glutamic acid	49449 Sigma Aldrich
L-isoleucin	12752 Sigma Aldrich
L-leucine	4330 Calbiochem
L-lysine	440-100G Calbiochem
L-methionine	9265 Sigma Aldrich
Mannitol	105982 Merck
Methanol	131091.1612 PanReac
NEBuffer™ r1.1	B7201 NEB
$MgCl_2$	1.05833.1000 Merck
NaCl	1.06404.1000 Merck
NaOH	1.06498.1000 Merck
NdeI	00222478 NEB
Peptone	2053079 Gibco
Phusion HF buffer	F518L Thermo Scientific

Phusion Polymerase	F-530 M0530S NEB
SacI	R606A Promega
Silver nitrate	131459.1608 PanReac
Sodium carbonate	1.06392.1000 Merck
Sodium dodecyl sulfate (SDS)	428015 Calbiochem
Sorbitol	56755 Calbiochem
Taq DNA Ligase	M0208S NEB
TCA (Trichloroacetic acid)	131067.1611 Panreac
TEMED	A1148 PanReac
Tris base	BM 2000 Winkler
Trypsin	15400054 Gibco
Tryptone	211705 Difco
XbaI	00205486 Thermofisher
Xylan	P-XYLNBE-10G Megazyme
Yeast extract	212750 Difco
Yeast Nitrogen Base	239210 Difco
1 kb ladder	N3232S NEB
Protein ladder	26616 Thermo Scientific

7.2 Primer

Primers were ordered from Integrated DNA Technologies, Inc.

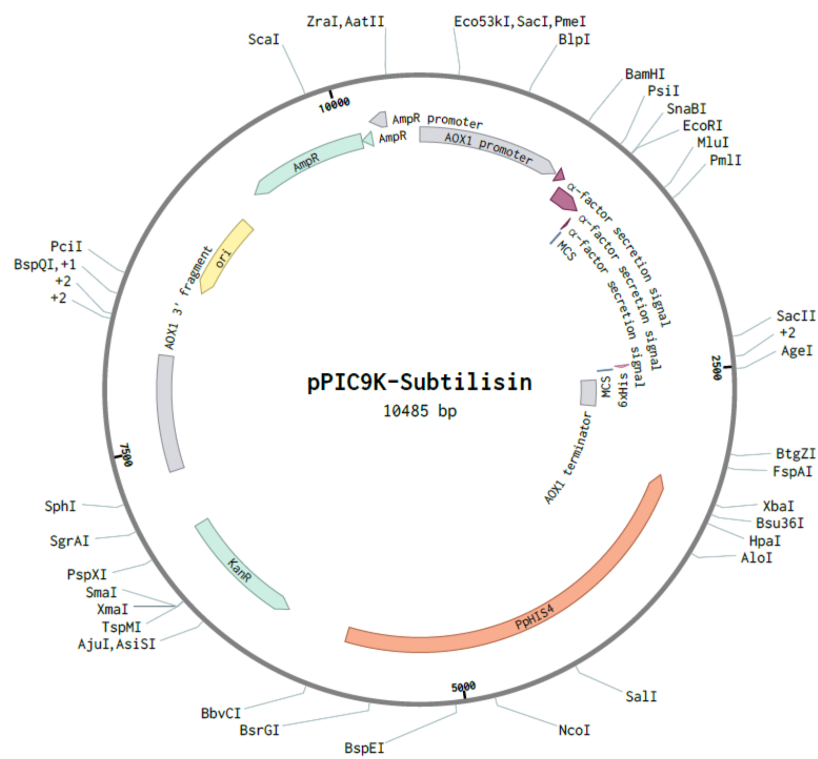
Table 8: List of used primers

Name	Sequence
α-forward	TACTATTGCCAGCATTGCTGC
AOX-reverse	GCAAATGGCATTCTGACATCC

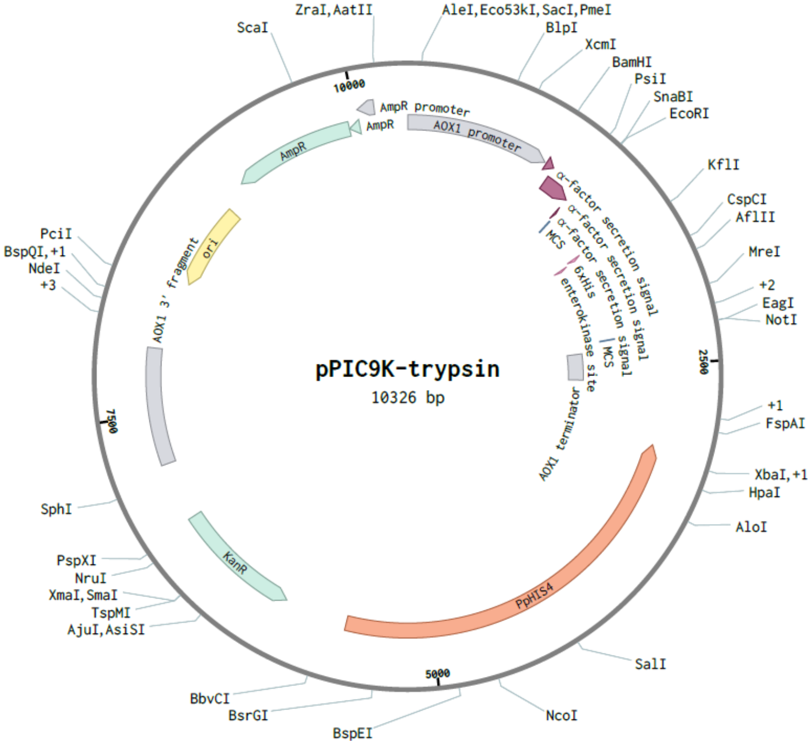
7.3 Plasmids

Plasmids were ordered from Integrated DNA Technologies, Inc.

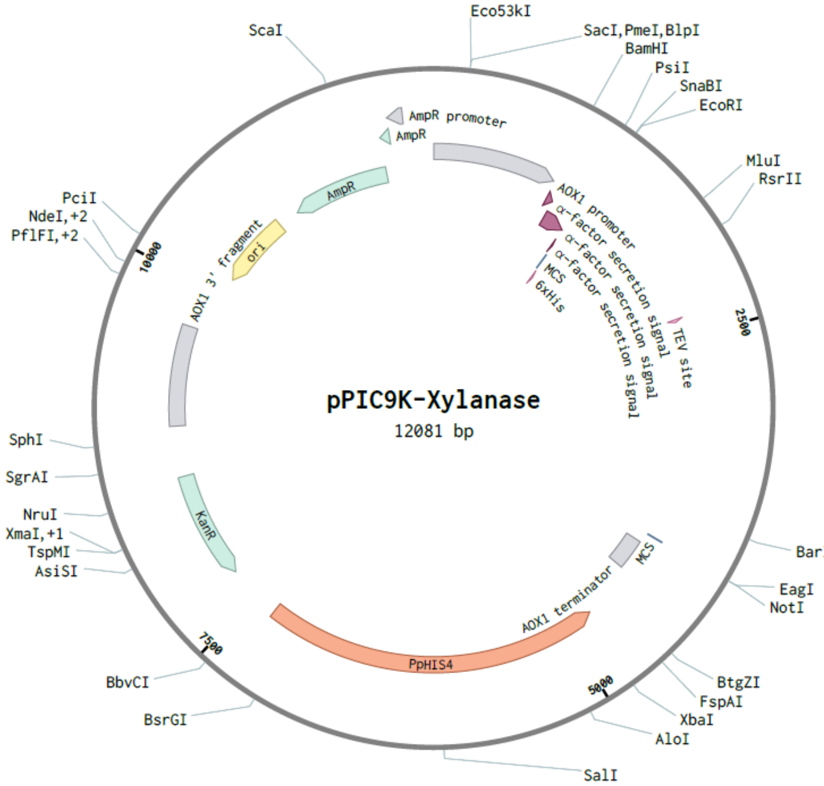
7.3.1 pPIC9K-subtilisin



7.3.2 pPIC9K-trypsin



7.3.3 pPIC9K-xylanase



7.4 Protein sequences

7.4.1 Subtilisin:

```
MKKRYINLLLTIGVFMISAFNMNAQKQEELTKISSKYNQEKLTTLKNDFKQKASLDKQNAITIAKSKGWKTRFTNKKGELLEIQK
VVNGKPIYYTTFNVAAAKSTRTNHLNNGGSLGLNLMGENMTAHVWDGGLARASHQEYDGAGGTNRFSIGDGTTALHYHSAH
VTGTIMASGVVANAKGMAPHASAVGYDWNNDTSEAINAASNGMLVSNHSYGFATRNAQGQPQLPDYYFGGYITDSRDWDNI
MFNAPNYLMVVAAGNDGNDNSANGAPLAGNSSYDKLSGHATAKNGMVVANANDANIDVNGNLLSVTINSSSSEGPTDDYRI
KPDITGNGTSVYSTYSSSNTAYNSTTGTSMASPNVAGTLLILQQHANNVRGSFIKASTLKGIALHTADDAGSNGPDAIFGWGL
MNAKRAAVAITQNGTESKIEELTLSSRQTYQITVDADGVNDLMASISWTDRAGTATTTANSSTAVLVNDLDIRVSKNGTTYTPW
RLTGVTTNGKGDNTVDPYERVDVANASGTYTITVTHKGSLTGGSQNYSLIVTGLAGTPVVCNATIPSNLTVDESGASTATVSW
NTVAGTSYDFRYRKTGTSTWTTSAVAGTSVSLTGLSTQTSYQTQVRSKCPNNSTSAYSSAVSFTTSDVQLNYCASNGNSVA
DEYISKVVLGGINNTTGASSSGYADYTSQSTSLTKGVSSTITITPTWAGASYNEGYAVFIDYNKDGDFTDNGETVWTKIASKTK
PVSGSFTVPTSATTGATRMRVVMQYNTVPAACGTYNYGETEDYTVNITGSSADTIAPTAPTNVSASAITQTTATLSWTASTDN
VGVAGYEIFSNGTSVGTVTATSANITGLTANTSYSYTIKANDAAENTSNSSNSVSFTTLGSTLVYCSSKGNRVTYEWIDYVSFG
GMTNTTAANAGYGDFTSKTATVSKGSDNQLIISAGFASTAYTEHWAVWIDFNQNGTFEESEKVTSGSSSSAANLTATISIPSSA
NTGQTRMRVSMKYNSAQTACETFSDGEVEDYTVNITNATANYTTFINTNSKNELGNESKAFDFTVYPNPVKGTVLNIHLNDAR
EVNFAITNMLGQTLKSGILTKQPIDVSTIKTGVYMLEITDGQKSVVKKFVRQ
```

Annotations:		
	1-26:	signal peptide
	27-96:	pro-peptide
	113-422:	peptidase S8/S53 domain
	558-646:	FN3 domain
	800-887:	FN3 domain

7.4.2 Trypsin

```
MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGE
VAATKVGALSKGQLKEFLDANLAGSGSGSHHHHHHDDDDKVVGGTRAAQGEFPFMVRLSMGCGGALYSSRLVLTAAHCVG
ATGTNTSITATLGTVDLQSSSRITVRSNYVYRAPGYNGSGRDWALIRLATPVTGLSTLKIANTTAYDSGTFTIAGWGAAREGGG
QQRYLLKATVPFVSDSTCNSYYGGEIIPAQEICAGYAAGGVDTCQGDSGGPMFRRDASNAWVQVGIVSWGNGCARPNYPG
VYTQVSYFASSIASAAASLGG
```

Annotations:	1-109:	thioredoxin TrxA
	110-115:	linker
	116-121:	His-tag
	122-126:	enterokinase recognition site
	127-351:	trypsin

7.4.3 Xylanase-L

```
MNKSIFRNTGLVTLVSLLMACGGNNKDTPVPEPIPEVVAPDTPEPEAPIAAEITNGGFEEDTAGQTTPVGNWVFRPTQESSAT
STIEVIESEEGVNTYQGTKAVEVNVNTLGDNPWGIEIAYEDLPITGGKNYEFSVWAKGEEGTSADFWIQTPAPDYGQLSLVKE
TLTGEWQKITLTAATAEADSLVRLAIHFSKEENINKSIYLDEFSGFILDDVPAQEIPDVQYSEVTAQSLKALAPNFNIGVAVPAGG
FGNSVIDRPEIKTIIEQHFNQLSAENIMKPTYLQPTQGEFFYDDSDELVNYAKDNSLTVHGHVFVWHSQIAPWMQSFQGDKAA
WITMMENHITQVATHFEEEGDNDTVVSWDVVNEAFMENGKYRGEKTTDDSADESVWFENIGAEFLPLAYKAARAADPDADL
YYNDYNLIWNADKLDAVIAMVNDFHNNGVPIDGIGFQSHISLNSPDISTIQAHLQKVVDIRPKIKVKITELDVRMNNEGGIPLTYL
TSERADEQKQYYYDIVKTYLETVPEDQRGGITIWGVIDEDSWLQNWPEPKTEWPLLFFNDFTAKPALQGFANALKELIEVVQP
APSSELLTNGDFEAGLDSWQARGSASITLESTQAHSGNNSALVQGRTETWNGLQKDVKGLFTADKTYNVSAWVKLSDDTST
VSPDIKLTLQIEHTSTEYLELTPVTTVAAGEWVQLSGTYTHSITTQESAALLYVESSELTADFYVDDVSVTLVE
```

Annotations:	1-25:	signal peptide
	53-194:	CBM22 (binding domain)
	233-575:	GH10 (xylanase domain)
	590-722:	CBM4 (binding domain)