EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Masterarbeit identisch.

Datum

Unterschrift

Diese Arbeit ist meinem Opa Elmar Koch gewidmet

"Bims ich etwa verrückt geworden?" fragte der Hutmacher traurig. "Ich fürchte, ja", sagte Alice, "Du bimst total durchgeknallt. Aber soll ich dir ein Geheimnis verraten? Das macht die Besten aus."

Nach Lewis Carroll: Alice im Wunderland

Danksagung

Mein größter Dank gilt meiner Familie, im speziellen meinen Eltern Sabine und Kurt Rieder und meiner Oma Mena Koch, die mich nicht nur finanziell, sondern auch mit vielen Ratschlägen in den nicht immer einfachen Zeiten während meines Studium unterstützt haben.

Ich möchte auch Prof. Anton Glieder danken, der mir nicht nur die Chance gegeben hat in seiner Gruppe meine Masterarbeit zu schreiben, sondern mich auch bei meinem weiteren Werdegang unterstützt hat.

Dr.techn. Christoph Reisinger gilt ein weiteres Dankeschön, als Hauptbetreuer meiner Arbeit war er immer für meine Fragen da und hat mir sehr viel Freiraum für meine eigenen Ideen gelassen, auch wenn diese nicht immer von Erfolg gekrönt waren.

Der wohl besten Arbeitsgruppe möchte ich für die lustige und lehrreiche Zeit danken und hoffe, dass sich unsere Wege auch in Zukunft wieder kreuzen werden.

Abstract

The filamentous fungus *Trichoderma reesei* is an important production host for homologous and heterologous protein production in modern enzyme industry. *T. reesei* is mainly used for the sake of its high efficiency for secretion, of about 100 g/l for homologous proteins and up to several grams for recombinant proteins, which can reportedly be reached[1].

For modern expression systems, coexpression of multiple genes can be indispensable as some proteins need simultaneous expression of redox partners, are composed of multiple subunits or chaperons are needed for proper protein folding. Such coexpression is easily possible by using bidirectional promoters, which frequently also give the added benefit of a smaller plasmid design compared to the classical two-promoter-system approach.

This work provides data on 8 different *T. reesei* strains, QM6a, QM9414, RutC30, NRRL15449, NRRL15500a, NRRL15500b, NRRL15500c and NRRL15502, including analyses of native protein secretion on different primary carbon sources and optimization of cultivation conditions for sporulation, a crucial factor for spore-forming filamentous fungi.

Furthermore, a new design and assembly of a vector system for recombinant protein production in *T. reesei*, using bidirectional promoters, was studied. Although a low specific targeting efficiency can be expected, the system was based on targeted integration of the expression cassette in the *cbh1* locus. In industry, this is the most frequently used gene locus for chromosomal integration in *T. reesei*. Furthermore, a new tool based on Excel spreadsheets and customized VBA codes was designed and successfully applied for the discovery of bidirectional promoters in the genome of *T. reesei.* To our knowledge, this was the first-time that bidirectional promoters were studied in *T. reesei*.

Finally, *T. reesei* was transformed and the co-expression of the two fluorescent reporter proteins eGFP and sTomato was analysed by microscopy and using fluorescence measurements with protein preparations of the transformed cells in order to prove the functionality of the new expression vectors in *T. reesei*.

[1] It must be noted that protein quantification is often done using colorimetric assays that have strong variations in specificity in dependence on the protein species that is measured. Therefore, reported absolute numbers for protein titers must be seen in such light.

Zusammenfassung

Der filamentöse Pilz *Trichoderma reesei* ist in der modernen Industrie einer der wichtigsten Organismen für die Produktion von homologen und heterologen Proteinen. Ein Hauptgrund für den Einsatz von *T. reesei* ist dessen hohe Effizienz bei der Sekretion von Proteinen. Bei homologen Proteinen werden bis zu 100 g/l sekretiert und bei rekombinanten Proteinen wurden Titer von mehreren Gramm pro Liter berichtet.

Ko-expression von unterschiedlichen Genen ist für modernere Expressiossysteme unabdingbar, da für manche Proteine eine Expression von Redox Partnern notwendig ist, die funktionellen Proteine aus unterschiedlichen Untereinheiten bestehen können oder die korrekte Proteinfaltung erst durch die simultane Expression von Chaperonen ermöglicht wird. Die Ko-expression mehrerer Proteinen kann durch den Einsatz von bidirektionalen Promotorsystemen einfach realisiert werden, welche, im Vergleich zum klassischen Ansatz mit zwei separaten Promotoren, häufig den zusätzlichen Vorteil eines kleineren Expressionskonstruktes zeigen.

Diese Arbeit enthält Daten zu acht *T. reesei* Stämmen QM6a, QM9414, RutC30, NRRL15449, NRRL15500a, NRRL15500b, NRRL15500c und NRRL15502, inklusive Analysen zur nativen Sekretion von Proteinen durch die verschiedenen Stämme bei Wachstum mit unterschiedlichen primären Kohlenstoffquellen, sowie zur Optimierung der Kultivierungsbedingungen, welche zu optimaler Sporenbildung, einem wichtigen Faktor für sporulierende Pilze, von *T. reesei* führen.

Ein weiterer Teil dieser Arbeit beschäftigt sich mit dem Design und der Assemblierung eines Plasmids für rekombinante Proteinexpression mittels bidirektionalen Promotoren in *T. reesei*. Obwohl eine niedrige spezifische Targetingrate erwartet wurde, wurde das System auf einer Strategie zur ortsspezifischen Integration der Expressionkassette in den *cbh1* Lokus, den am häufigsten verwendet Genlokus zur Integration in *T. reesei,* aufgebaut. Zusätzlich wurde, unseres Wissensstandes nach, zum ersten Mal ein System für die Erforschung bidirektionaler Promotersysteme im Genom von *T. reesei* entwickelt und erfolgreich angewendet.

Letztlich wurde *T. reesei* transformiert und die erfolgreiche Expression der beiden Fluoreszenzproteine eGFP und sTomato wurde sowohl mikroskopisch als auch durch Fluoreszenzbestimmung an Proteinpräparationen der transformierten Stämme analysiert, um so die Funktionalität des entwickelten Vektorsystems in *T. reesei* nachzuweisen.

Table of contents

1. Introduction

The introduction part is intended as part of a book published by WILEY-VCH and will focus on industrial production systems for enzymes.

1.1 Fungal production systems for industrial enzyme production

Proteins produced by different organisms under different cultivation conditions can be found everywhere in our modern life. This article focuses on, eukaryotic microbial production hosts, and explains special considerations for such hosts and the consequences for modern industrial enzyme production. Sales figures (Figure 1) for industrial enzymes show an estimated value of about €3.0 billon per year compared to about €100 billon in pharmaceutical enzyme production in the year 2015. [1] With a share of 48% of the total enzyme production market, Novozymes was the leading company in this field.(data adapted from Paloheimo et al. [2]) while the top 3 players in the field shared about 75% of the total market value.

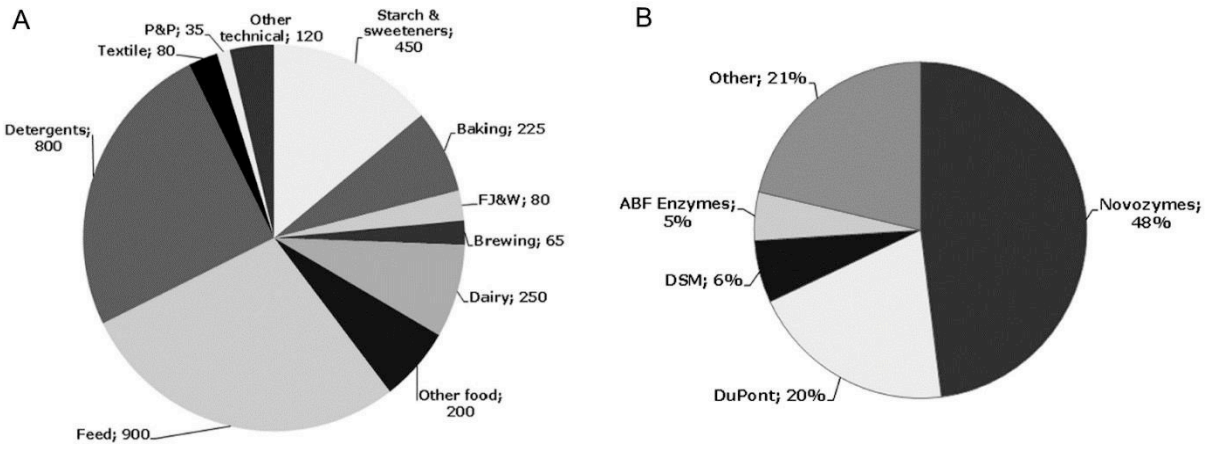


Figure 1 Industrial fields of enzyme application, their quantity based on sales(a) and the leading companies in this field (b). Around 50% of the produced enzymes are located to the sectors Detergents (Laundry detergents) and Feed. Based on the sales from 2015. 4 companies were the major players in the production of enzymes. The largest was Novozymes with almost 50% of the whole market. (data adapted from Paloheimo et al. [2])

1.2 Special considerations for working with eukaryotic expression systems

1.2.1 Choosing an expression host

Different expression hosts have different advantages or drawbacks for the production of individual target proteins. Major criterion deciding about the best suited system are total yield, space time yield, productivity, folding and post-translational processing efficiency, cost of media, downstream processing (DSP) or compatibility with a desired application in food,

feed or pharma applications – summarized as complexity of legal product approval.

In spite of the vast biological diversity just a few microorganisms are used as hosts for industrial protein expression. Table 1, adapted from Gomes et al. [3], shows advantages and disadvantages for industrial application of the most commonly used organisms . In addition, the application depends on target specific challenges and therefore no "master production host" is suitable for all applications.

Table 1 List of industrial organisms for protein production. Data were adapted from Gomes et al [3]

		Organism	Advantages	Disadvantages
Prokaryotic	Bacteria	*Bacillus subtilis*	GRAS, efficient secretion	secreted proteases, bad expression of GOI, low stability of plasmids, sporulation of WT strains
		Escherichia coli	easiest, fast and cheap system, well defined genetic and metabolic models, no glycosylation, different platform strains.	N-, and O-, glycosylation, codon usage, photogenic strains exist, no efficient secretion
		Pseudomonas	high cell densities	plasmid stability, photogenic strains exist
Eukaryotic	Yeast	*Pichia pastoris*	high cell density, easy scale up, high levels of functional proteins	hyper mannosylation of mammalian proteins
		Saccharomyces cerevisiae	well established, can be used for therapeutic application	hyper glycosylation of mammalian proteins
	Fungi	*Aspergillus niger*	GRAS, secretion potential, eukaryotic post translational modifications,	complex manipulation
		Myceliophthora thermophila	use of complex growth substrates	complex manipulation
		Trichoderma reesei	efficient secretion, eukaryotic post translational modifications	complex manipulation, secreted proteases

The optimal choice for a specific production hosts depends on the final application. The specific requirements for proteins produced for pharmaceutical applications are fundamentally different from those for proteins produced for consumer applications, food or feed purposes or biocatalysis for chemical production. In case of a pharmaceutical product it

is absolutely necessary to have flawless post-translational modifications, low error rates in translation and perfect folding of the protein. [4] On the other hand the price, specific productivity, cost of media, complexity in DSP and hurdles in commercial product approval will be effectors for choosing a production host for laundry enzymes

This article focuses on "White Biotechnology", thus biotechnological solutions for industry by using living cells or pure enzymes for chemical manufacturing producing less waste and using less energy [5]. Examples are biodegradable plastic from bacteria, [6]; the production of citric acid in *Aspergillus niger* [7] to modern biocatalytic applications for conversion of new substrates in *Pichia pastoris* (*Komagataella phaffii*) [8].

The following paragraphs focus on yeasts and fungi as expression hosts, as these are the most commonly used eukaryotic organisms in white biotechnology and very well complementary to frequently used bacterial hosts.

1.3 Comparison of cell structure and their influence on molecular biology

The molecular organisation of cells is a major issue which influences the potential use as "cell factories" for the production of enzymes as chemicals. Eukaryotic cells show many significant differences to the prokaryotic cells, which has consequences for their respective fields of application.

Cellular compartmental isolation to subcellular organelles with substrate permeability, redox conditions and different sets of specifically targeted enzymes are specific biological activities in eukaryotic hosts. For example, protein secretion involves specific targeting to the endoplasmic reticulum (ER) directed by specific signal peptides such as the CBH1 secretion signal of *T. reesei*. Specific posttranslational modifications such as N-, and O- glycosylation take place in this compartment and the quality of folding of disulphide bridges containing enzymes is important before the respective enzymes follow their path of further transport and modification through different Golgi compartments before fusing with the cell membrane and secretion to the culture supernatant. [9][10]

In prokaryotic systems, all processes connected to the DNA, such as replication and transcription, take place in the cytoplasm, while in eukaryotes these processes are located to the nucleus. As a consequence, in a prokaryotic system like *E. coli* the DNA introduced during the genetic engineering only needs to pass over the outer cell membranes to be transcribed and translated, which might be one of the reasons for lower transformation rates for eukaryotic hosts.

Even in lower eukaryotes such as *Trichoderma reesei* or *Pichia pastoris* as expression host, foreign DNA has to get through the cell wall but also through the pores of the nucleolus to reach the place for transcription while degrading enzymes attack such foreign DNA in the cytoplasm. The mRNA formed and processed within the nucleus then has to be exported outside of the nucleolus were translation and further translocation for example secretion to the extra cellular space takes place. [11], [12]

Such host specific differences also influence the design of specific expression vectors, complete expression cassettes and finally the complete expression strategy.

1.4 Differences in vector design for eukaryotic and prokaryotic hosts

In order to enable simple cloning and DNA amplification most expression constructs are made as plasmids which can be propagated by *E. coli.* In general plasmids can be summarized as circular or linear extrachromosomal replicons that can be used for the introduction, modification or removal of target genes [13]. It is favourable to have as small expression vectors as possible, due to the fact, that the metabolic load on the organisms increases with increasing size of the expression construct [14] and also plasmid instability increases with expanded size [15]. Probably one of the largest plasmids found in nature was from *Azospirillum brasilense* CBG497 with 1.59MBp [16], while typical natural plasmids of *E. coli* are 10-100 KBp and lab vectors are usually not bigger than 3-5 KBp. [17]

The exact design of an expression vector depends on the specific host of choice. Nevertheless, we can distinguish between autonomously replicating plasmid based expression system, also called episomal vectors [18], or integration vectors systems, which do not replicate autonomously but get integrated into the genome and are maintained during cell division due to chromosomal replication origins.

Plasmid based expression systems are more prominent for their use in prokaryotic organisms but they can also be used in lower eukaryotes, while their application is limited in higher eukaryotes because of the transient expression. [14]

Plasmid based systems for both types of organisms, prokaryotic and eukaryotic, have some parts in common, but also show some significant differences, which are mostly host related. For eukaryotes, normally shuttle plasmid with regulatory elements working in prokaryotes and eukaryotes as cloning is much easier and faster in *E. coli*. [19]

Schemes of typical vectors for plasmid based expression in prokaryotes and eukaryotes are shown in Figure 2 and Figure 3. In both systems, an origin of replication (ORI) for

independent replication of the plasmid within the cell, a selection marker for the use in prokaryotes and a multiple cloning site (MCS) for the integration of the gene of interest into the vectors is needed. For transcription during the cloning procedure in *E. coli* and in the expression host promoter (green) and terminator (red) regions are needed for selection marker as well was for the final target gene.[20]

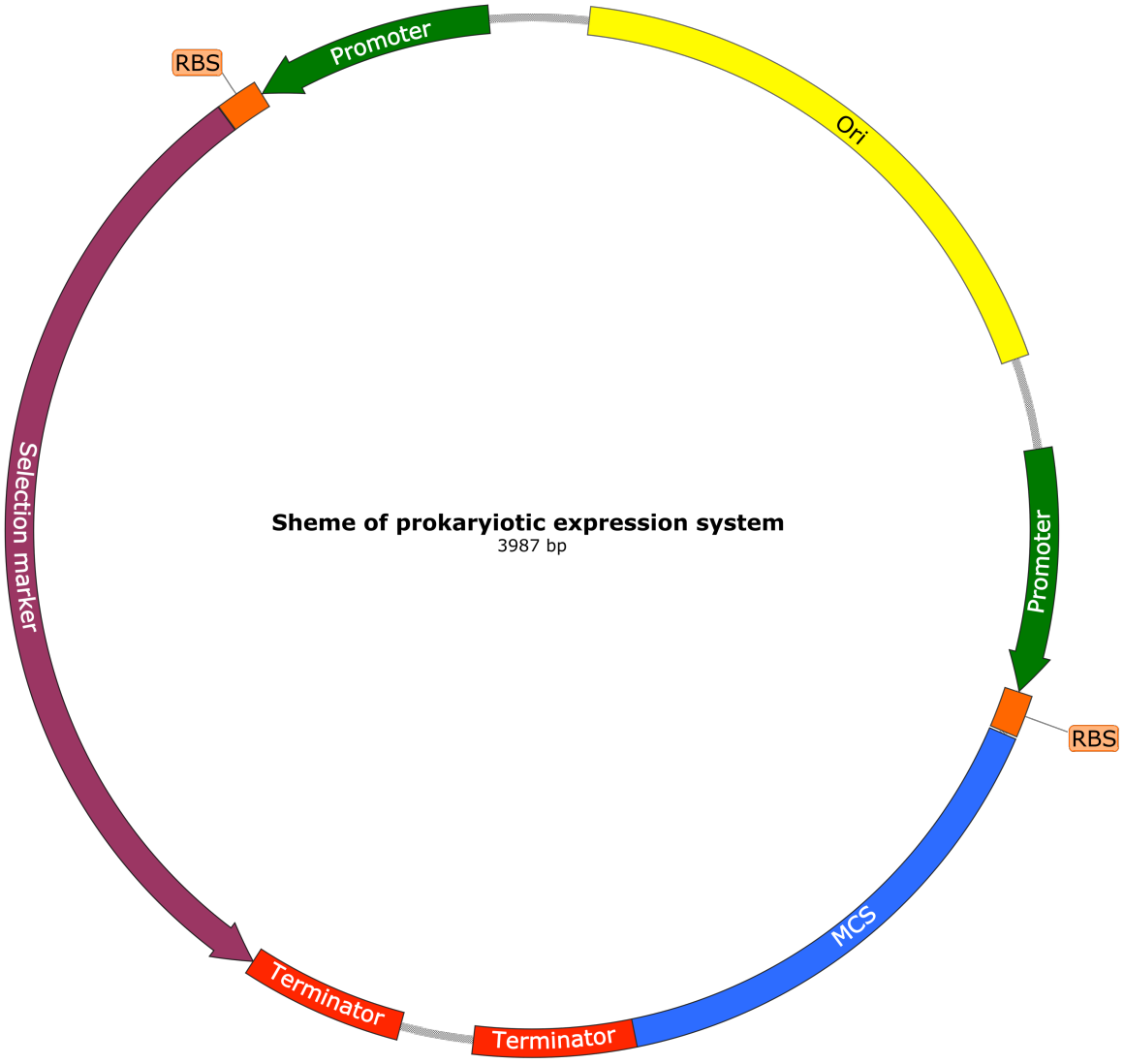


Figure 2 Scheme of typical prokaryotic expression vector. In general, three elements are absolutely necessary to have: the origin of replication (ORI) responsible for autonomous replication of the plasmid in the cellular environment. The selection marker for selection of transformed cells after transformation, including promoter and terminator. Multiple cloning site (MCS) the part of the plasmid were the gene of interest should be intergraded between another promoter and terminator for transcription. In most cases also a Shine- Dalgarno (SD) sequences as ribosome binding site (RBS) to enable the translation starting at an ATG start codon located next to it.

Next to the regulatory elements mentioned above episomal plasmids in eukaryotes need to have another selection marker for the selection in the final eukaryotic host including promoter

and terminator. (Figure 3) An alternative approach uses selection markers working in prokaryotic and eukaryotic organisms while the promoter for prokaryotic transcription is located within the 5'-UTR of the eukaryotic core promoter region. (Figure 4) The second element, which is indispensable, is a feature responsible for the replication of the plasmid similar to the ORI in prokaryotes. In eukaryotes this element is usually called autonomously replicating sequence (ARS). [21]

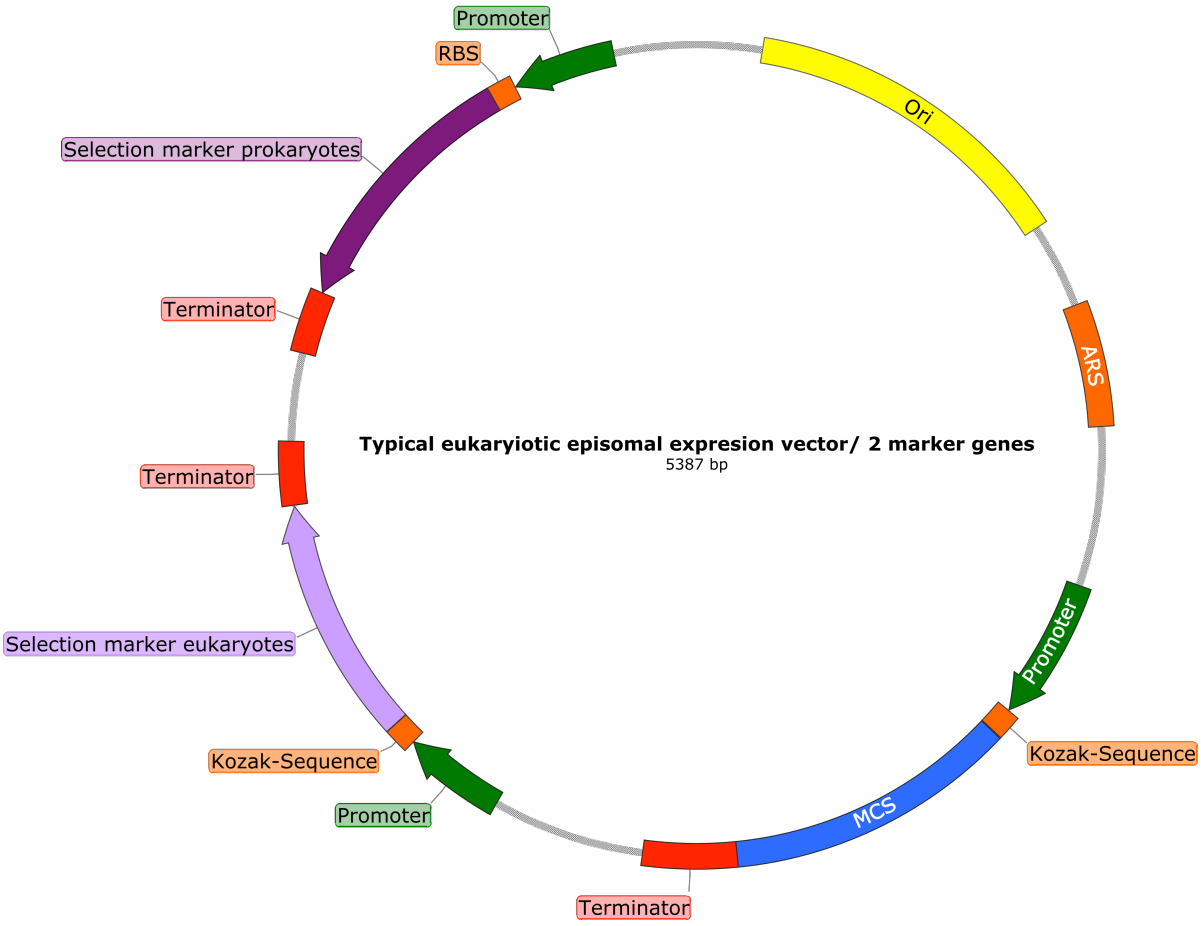


Figure 3 Scheme of eukaryotic plasmid based expression system. Essential elements include: the origin of replication (ORI) responsible for plasmid amplification in *E. coli*, autonomously replicating sequences (ARS) for independent plasmid replication in eukaryotic cells. The selection marker for selection of positive transformed cells after transformation of the eukaryotic expression host and its respective eukaryotic promoter and terminator sequences. Further the vector needs a selection marker for prokaryotic organisms for cloning procedures, including promoter and terminator element. A Multiple cloning site (MCS) to introduce the gene of interest (GOI) for expression is also part of the plasmid.

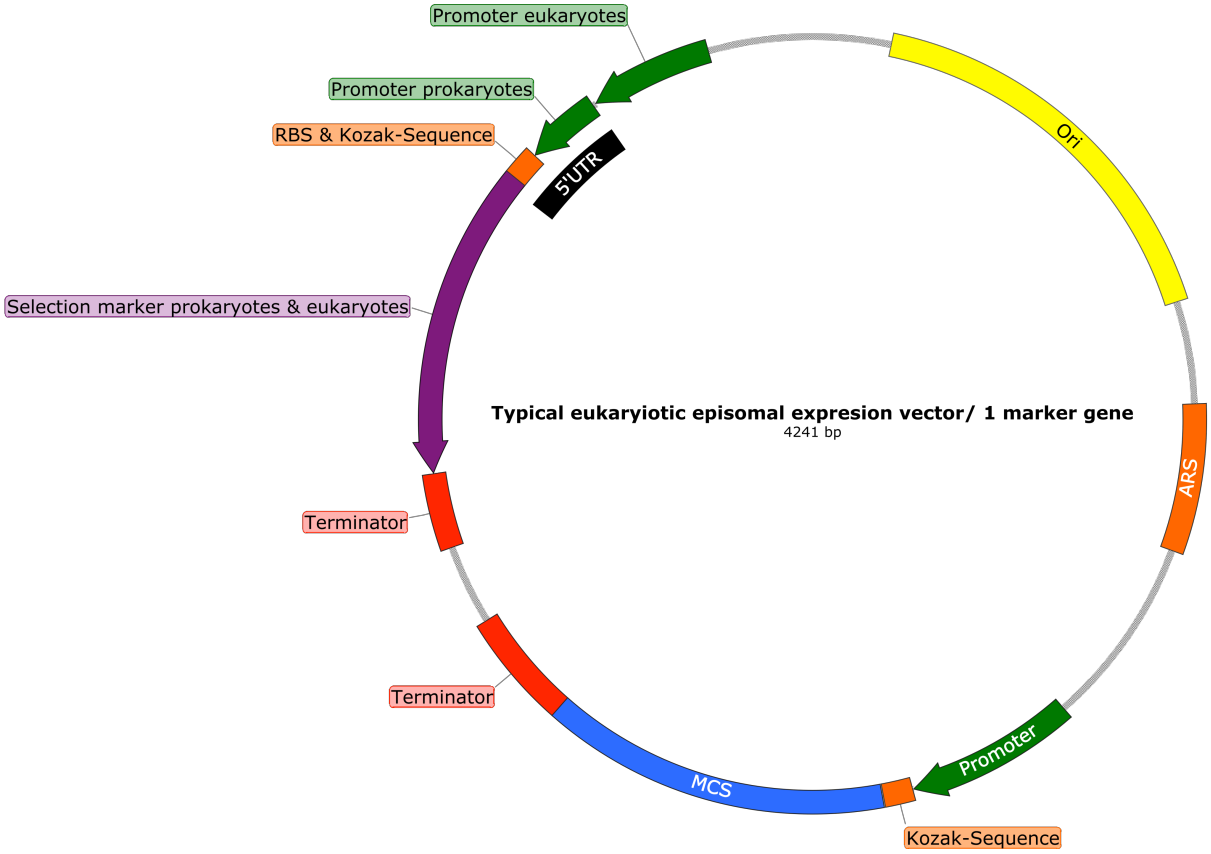


Figure 4 Scheme of eukaryotic plasmid based expression system. Essential elements include: the origin of replication (ORI) responsible for plasmid amplification in *E. coli*, autonomously replicating sequences (ARS) for independent plasmid replication in eukaryotic cells. The selection marker for selection of positive transformed cells after transformation of the eukaryotic expression host and its respective eukaryotic promoter and terminator sequences Selection marker for prokaryotic organisms for cloning procedures is in this special case the same as for the eukaryotic selection, including promoter and terminator element with the 5'UTR of the selection marker gene. Multiple cloning site (MCS) as part of the plasmid to introduce the gene of interest (GOI) for expression.

While plasmids are still frequently used for the protein expression in prokaryotic hosts and baker's yeast integration vectors are more common for many other yeasts (e.g. *Pichia pastoris*, *Hansenula polymorpha*) and standard for filamentous fungi. [22] [23]

One of the two mayor benefits of genomic integration is the high chance to obtain a stable production strain, in contrast to plasmid based autonomously replicating vectors. [24] The second reason why production strains with expression cassettes which are integrated to the genome into preferred for industrial applications, is that no further selection pressure (e.g. by antibiotics) is needed. [25], [26].

Expression strength also depends on the specific locus of the integration. [24] such site-specific effects lead to the strategy of using homologous flanking regions, (Figure 5) to direct the expression cassette to a defined and preferred locus. Such specific loci can show special

properties such as. inducible, or known for strong expression or a locus with neglectable epigenetic effects. An example for this strategy is the directed integration into the *AOX1* locus in the methylotrophic yeast *Komagataella phaffii* (*P. pastoris*). Expression from the *AOX1* promoter is strongly inducible by methanol and as an added benefit, disruption of the *AOX1* gene reduces biomass production from methanol to a minimum, therefore uncoupling cell growth from product production is possible [27] In addition less toxic hydrogen peroxide and heat are produced as by-products of the methanol metabolism. Site specific integration can be identified due to the slow growth on methanol as sole carbon source cause by the disruption or replacement of the gene open reading frame coding for the alcohol oxidase 1 (AOX1).

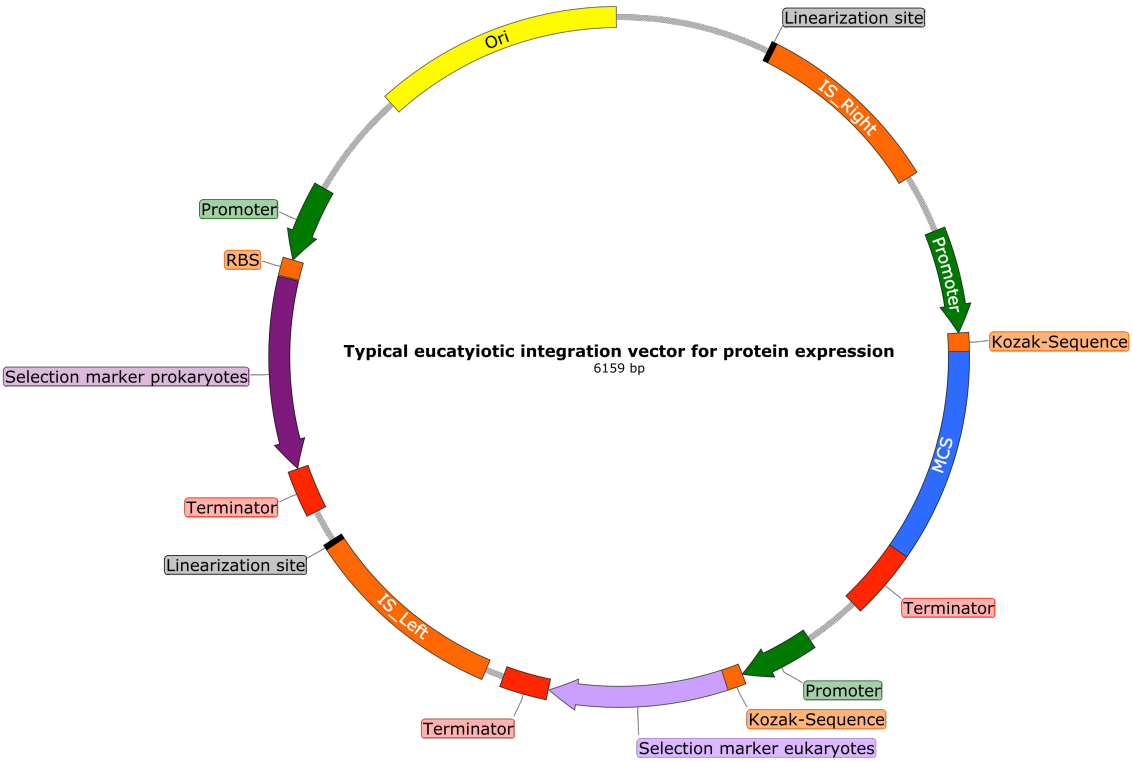


Figure 5 Scheme of eukaryotic plasmid based expression system. Elements which are absolutely necessary to have: the origin of replication (ORI) responsible for autonomous replication of the plasmid in *E. coli*. Integrations sites flanking the multiple cloning site and selection markers for eukaryotes for directed integration to the genome. The selection marker for selection of positive transformed cells after transformation to the eukaryotic expression host, including promoter and terminator. Selection marker for prokaryotic organisms for cloning procedure, including promoter and terminator elements Multiple cloning site (MCS) the part of the plasmid to introduce the gene of interest (GOI) for expression.

While many lab experiments use antibiotic resentence markers due to their efficiency and simplicity, markers relying on growth deficiency such as auxotrophies or lack of carbon source utilization are preferred for industrial applications. [28]

Table 2 Frequently used selection markers based on antibiotics for prokaryotic and eukaryotic organisms. Based on data from addgene [28]

	Substrate	resistance gene	Interaction	Concentration
Prokaryotic	Kanamycin	*nptII*	Binds 30S ribosomal subunit; causes miss-translation	50-100 µg/mL
	Ampicillin	*bla*	Inhibits cell wall synthesis	100-200 µg/mL
	Bleomycin (Zeocin)*	*ble*	Induces DNA breaks	5-100 µg/mL
	Carbenicillin	*bla*	Inhibits cell wall synthesis	100 µg/mL
	Chloramphenicol	*cat*	Binds 50S ribosomal subunit; inhibits peptidyl translocation	5-25 µg/mL (EtOH)
	Erythromycin	*erm*	Blocks 50S ribosomal subunit; inhibits aminoacyl translocation	50-100 µg/mL (EtOH)
	Spectinomycin	*aadA14*	Binds 30S ribosomal subunit; interrupts protein synthesis	7.5-50 µg/mL
	Streptomycin	*aadA14*	Inhibits initiation of protein synthesis	25-100 µg/mL
	Tetracycline	*tet*	Binds 30S ribosomal subunit; inhibits protein synthesis (elongation step)	10 µg/mL
Eukaryotic	Blasticidin	*bsd, bls, bsr*	Inhibits termination step of translation	2-10 µg/mL
	G418/Geneticin*	*neo*	Blocks polypeptide synthesis at 80S; inhibits chain elongation	100-800 µg/mL
	Hygromycin B*	*hygB*	Blocks polypeptide synthesis at 80S; inhibits chain elongation.	50-500 µg/mL
	Puromycin	*pac*	Inhibits protein synthesis; premature chain termination	1-10 µg/mL

In order to keep the expression construct small in size, antibiotics working in prokaryotes as well as in eukaryotes can be used. Examples are marked with an * Table 2. For expression in prokaryotes as well as in eukaryotes, dual promoters are useful. Therefore, the prokaryotic promoter is usually put to a non-functional and non-translated area of the eukaryotic core promoter or the 5' untranslated region of the transcript (5'-UTR).

Another possibility for selection is to use auxotrophic systems. This works by using strains lacking functionality of certain essential proteins or enzymes (auxotrophic strains). Subsequently, for selection and the complementation for such deficit, a functional copy of the corresponding gene is provided via the plasmid/linear fragment [29]. Auxotrophic systems are

mostly based on complementation of missing functionality in the synthesis of amino acids or nucleotides. Table 3 is based on data from addgene [28] and Nett et al. [30] lists frequently used autotrophic markers.

Table 3 Frequently used auxotrophic systems based on data from adgene [28] **and Nett et al.** [30]

Gene	Amino acid
HIS3	L-hisitidine
URA3	pyrimidine (uracil)
LYS2	L-lysine
LEU2	L-leucine
TRP1	L-tryptophan
MET15	L-methionine and overproduces hydrosulfide ions
ura4+	pyrimidine (uracil)
leu1+	L-leucine
ARG1	L-arginine
ARG2	L-arginine

Likewise, also the deficiencies in carbon source utilization can be used for the selection e.g. GUT1. The advantage in carbon source utilization deficiencies for selection lies in the simplicity of media preparation and strain cultivation as well as the possibility to apply complex media for growth, which is not the case for amino acid auxotrophic based selection. [31] Another example for such a system would be the TPI1- system developed from Novo for the production of Insulin. The *TPI1* gene which is coding for the triose-phosphate isomerase is deleted in the genome of the organism and by introducing a plasmid with the *POT1,* the corresponding gene from *Schizosaccharomyces pombe,* the pathway is completed. The clones generated by this method are very stable as cells which have no plasmid will hardly grow as a major enzyme for glycolysis is missing [32]

Another marker type are counter selection markers. Counter selection uses negative selection strategies meaning that the plasmid gets changed during cloning procedure leading to the survival of the cells. [33] An important counter selection marker for the use in yeast and fungi is amdS maker system. This marker was for example used for introducing several mutations to the genome of *S. cerevisiae* as marker recycling is possible. In the first round of selection the *amdS* gene is introduced to the cell which allows the use of acatamide as sole nitrogen source. After selection, the marker is recycled and the cells are counter selected on media with fluoroacetamide which is toxic for the cells when cleaved by the amidase. [34]

However, all these considerations only apply to expression strains employing episomal

plasmids. In case of stable integration of the expression cassettes into the genome no selection is necessary any more in the production process and scale up. Thus, the problematic of applying antibiotic resistance genes for selection is more or less reduced to an approval and public acceptance issue.

The strains with lacking synthesis of certain amino acids can also be used for labelling proteins by feeding labelled amino acids during cultivation procedure of the organism. [35]

1.5 Differences in regulation of expression in eukaryotes compared to prokaryotes

Promoters are special regulatory elements of the DNA, it is the part of the DNA where the RNA Polymerase binds to initiate transcription of the transcription start site. It was also observed, that these DNA elements show unusual structures and have low stability. [36]

Promoter regions from prokaryotic cells (Figure 6) show differences in structure compared to eukaryotic organisms (Figure 7)

For prokaryotic promoters, the sigma factor binding sites the -35 and -10 regions (green) are highly conserved and most important sequences. These conserved hexameric motifs are centred or near to the positions -35 and -10. The term -10 indicates the position ten base pairs upstream (left side in Figure 6) of the transcription start, which is indicated as +1. [36] Different consensus sequences and sigma factors are known for constitutive promoters compared to inducible promoters.

For translation of the GOI two different DNA motifs are needed. The ribosome recognizes the ribosome binding site, also called Shine- Dalgarno sequence and is binding with the 16s rRNA to this region. Against some assumption the RBS is not that strongly conserved within the whole kingdom of prokaryotes and is located 5-10 base pairs upstream of the start codon, the starting point of the translation. [37]

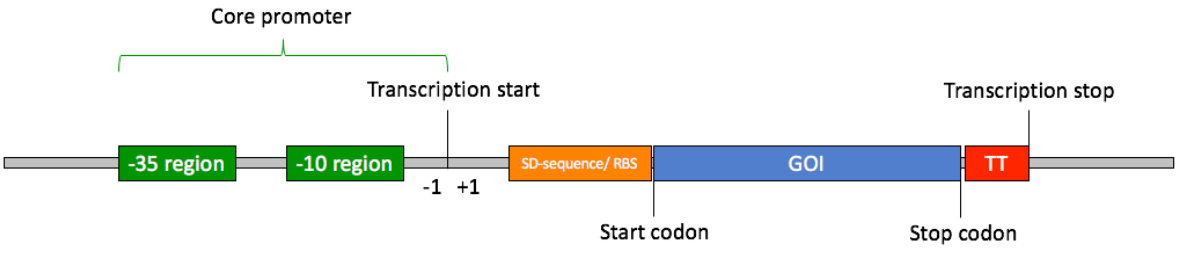


Figure 6 Scheme of prokaryotic promoter region. The green parts show the -35 and -10 region necessaries for binding the RNA- Polymerase located 10 base pairs upstream of the transcription start indicated as +1. The orange box indicates the ribosome binding site (Shine- Dalgarno sequence) 5-10 base pairs upstream of the start codon. Transcription is stopped by the terminator sequence.

Eukaryotic promoters are more complex in structure than prokaryotic ones (Figure 7). The promoter sequence can be divided into a core promoter and other regulatory elements an enhancer region, upstream of the core promoter (green). The most important element in the core promoter is the TATA-Box, located upstream of the transcription start. The exact location is dependent on the organism and promoter. This sequence is quite conserved, although some mismatches can occur and the whole core promoter region is usually AT rich. Eukaryotic systems also use different transcription factors, which induce the binding of the polymerase at the core promoter region.[38]

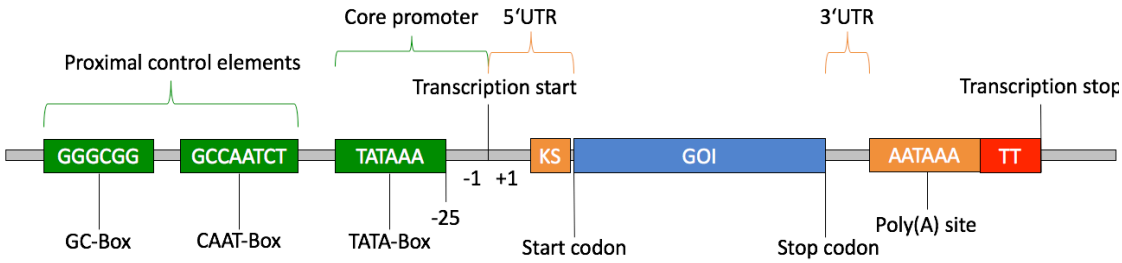


Figure 7 Scheme of eukaryotic promoter region. Elements belonging to the promoter are highlight in green. TATA-Box is the element which is recognized by the RNA-Polymerase. Position of the TATA-Box is dependent on the organism and the promoter. Other control elements like GC-Box and CAAT-Box are located upstream of the core promoter. Translation is introduced by the Kozak-Sequence (KS) is the element recognized by the ribosome. 5'UTR and 3'UTR are important factors for good expression. Before the transcription is stopped at the Terminator sequence a poly(a) site for post-translational modifications is needed.

Generally, translation is more tightly regulated in eukaryotic hosts. The Kozak-Sequence (KS) is the central element that facilitates the initiation of translation of the start codon AUG of the mRNA to start protein synthesis. The Kozak-Sequence is different from species to species and can be used for phylogenetic analysis, but generally the following base pairs are conserved amongst most species: -3A/G, -2AC and +5C although -3A/G is the most crucial position for good translation. (Figure 8) [39]

Figure 8 Detail of Kozak-Sequence showing the important positions (red) for good translation. In nature the bases -3A/G, -2AC and +5C are preferred.

Over the last few years it has further been proven, that the whole region not translated but transcribed, the so called 5'UTR, is important for good expression efficiency. [40] Kozak showed already in 1986 that hairpin structures within the 5'UTR are crucial for the expression level and that the expression level decreases with the strength of the hairpin. [41] A new approach of Weenink et al. [42] showed that it is possible predict the downregulation in expression by introducing hairpin structures to the 5'UTR.

1.5.1 Different types of promoters

Regardless the differences between the kingdoms, there exist constitutive and inducible promoters.

Constitutive promoters are active all the time, although they still show some regulation during different phases of growth. In *T. reesei* the promoter of the pyruvate decarboxylase promoter (P_{PDC}) is used in industrial applications for the production of xylanase II. [43], [44] Another example would be the Glyceraldehydes-3-phosphate dehydrogenase promoter (P_{GAP}) in *P. pastoris* which leads to an high space time yields (STY) in case the product causes no physical damage to the host. [35]

Inducible promoters can be divided into positive and negatively controlled ones. P_{AOX1} is tightly regulated promoter for application with the host *P. pastoris*. The wildtype promoter can be induced by methanol, but mutant variants can be used without the toxic and flammable substance.[27] Similarly the methanol oxidase promoter P_{MOX} of *H. polymorpha* is not repressed by glycerol and can be induced by simple depression while it is typically repressed by glucose. [45]

P_{CBH1} is an example for an inducible promoter for *T. reesei,* which is induced by the presence of complex sugars and repressed when glucose is the sole carbon source [46]. The promoters mentioned are used in industrial applications.

Besides the well-known monodirectional promoters (MDPs) also bidirectional promoters (BDPs) or dual promoters can be found in nature. While MDPs only transcribe in one direction BDPs transcribe in both direction. This application brings an enormous advantage, as we have to coexpress in many applications. Applications would be, for example, coexpression of an selection marker with the GOI, or the simultaneous coexpression of a chaperon or a redox partner [47].

A well-known example of such dual promoters of *Saccharomyces cerevisiae* is the strong

inducible Gal1/Gal10 promoter. [48]

1.6 Industrial enzyme production

Multiple integration of expression cassettes is one of the most frequently used strategies to increase expression yields. [49]

However similar effects can often be obtained by codon optimization (more efficient translation), or stronger promoters which lead to high abundance of mRNA like multiple expression cassettes. In general, a combination of the strategies lead to the most efficient production strains where a maximum of folded and processed active enzyme should be obtained while keeping the gene copy number at a minimum in view of potential genetic inhibitions due to the homologous recombination and cassette loss which might cause troubles during long time cultivation and large scale in industry. [50], [51]

Another strategy for enhancing protein production is the development of industrial production strains. The *T. reesei* strain QM6a discovered during the Second World War on the Solomon Island [52] is the basis for the most popular production strains QM9414 and RutC30 (Figure 9) which are used due to their high capacity of cellulase secretion. [53], [54]

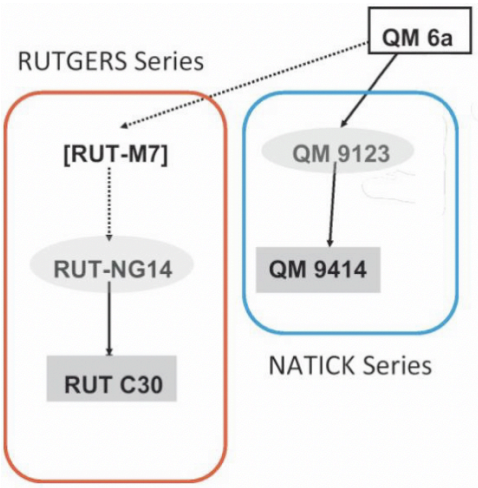


Figure 9 Genealogical tree of *T. reesei*. *T. reesei* QM6a as origin for modern production strains QM9414 and RutC30 made in several rounds of random mutagenesis. Adopted from Kubicek et al.[55]

Similarly, most industrial production strains of *P. pastoris* were derived from *K. phaffii* NRRL- Y11430. [56] Frequently used derivatives of this WT strain are the GS115 and the X-33 strain form Thermo Fischer Scientific or another platform described more recently by Näätsaari et al. [31] and Sturmerberger et al. [57]

Strain engineering was done mainly by using by UV or chemical substrates for random

mutagenesis which is a stable and easy method but causes a lot of screening effort. [58], [54]

More recent approaches use directed methods such as CRISPR/Cas9 to knock out specific genes e.g. proteases in order to enhance protein production, eliminate side products or to avoid protein degradation due to protease effects. [59] Screening of the knockout strains can for example be done by using compounds that become a toxic (suicide substrates) when used as substrate of a certain enzyme. By destroying the gene by a knockout only the organism having this knockout will survive on cultivation on a media containing the beforehand non-toxic substrate. [60]

1.7 Enzyme production in industrial applications

Although the exact vector/host combination and the scale of cultivation for the industrial enzyme production are hardly known publically the following chapter should give an impression about typical examples of industrial enzymes produced by eukaryotic microorganisms.

When it comes to classification there is a difference between homologous and heterologous proteins. Homologous proteins are proteins, which are naturally produced by the organism. Self-cloning allows to increase the expression yields without introducing genetic changes which cannot take place by nature. Heterologous proteins are not native to the organism and are encoded by genes introduced to organism by molecular biological methods.

1.7.1 Homologously produced proteins

Using “waste” material from food production (e.g. soy straw) containing cellulose, hemicellulose and lignin, the most abundant polymers on our planet, for production of enzymes is a goal to reach. The problem is that such resources have to be pre-treated for the most organisms, which is either be done chemically or physically, to fulfil the function as carbon source. [61]

One organism, which can use such natural polymers after pre- treatment for growing without any genetic engineering, is *Trichoderma reesei*. *T. reesei* is known for the effective production of biomass degrading enzymes such as cellbiohydrolases (CBH), endo-β-1,4 gluconases (EG) and β-glucosidases. This cocktail of enzymes, referred to as cellulase has an enormous value for modern industry. They are used textile industry, paper industry, for laundry applications and, last but not least, in bio refineries. [62] Further *T. reesei* has excellent protein secretion properties; modern strains can secrete up to 100 g/l of total protein.

For this reason, *T. reesei* is assumed to be one of the most efficient organisms for protein production.

Myceliophthora thermophila C1 is another lignin cellulose degrading strain which was established for industrial enzyme production by Dyadic (now DuPont/ Genecore). Like *T. reesei* the secretion potential of *M. thermophile* is enormous and reaches up to 100g/l with a purity up to 80%. [63] *M. thermophila* shows in contrary to *T. reesei* low protease activity during continuous cultivation processes, which makes it a good production strain for recombinant protein production. [64]

Table 4 Selected examples of organisms with corresponding enzymes, yields and promoters for industrial homologous protein production adapted from Teuschler, 2017 [65]

Expression Host	Enzyme	Yield	Activity	Reference
Aspergillus oryzae	Glycomylase	20 g/L		Finkelstein, 1987
Aspergillus niger	Amylase	20 g/L		Wang *et al*., 2016
Aspergillus niger	Acid Protease		3600 U/L	O´Donnel *et al*., 2001
Aspergillus niger	Pectinase	2.15 g/L		Suhaimi *et al*., 2016
Aspergillus niger	Catalase	16.5 g/L		O'Donnell *et al*., 2001
Aspergillus flavus	Glucosidase	21 mg/L	0.82 IU/mL	Dutt *et al*., 2012
Penicillum notatum	Glucose oxidase		112 U/mL	Sabir *et al*., 2007
Trichoderma reesei	Cellulase	100 g/L		Cherry *et al*., 2003

1.7.2 Heterologous produced proteins

Compared to native proteins, the secretion of recombinant proteins is relatively low [66]. The main cause for this low secretion yield could be the high proteases activity and the incorrect folding of recombinant proteins. Over the years many different strategies, such as optimizing the codon usage [67], multiple random integration and the use of engineered promoters [68] have been developed. Next to those strategies on DNA level, approaches to increase the amount of secreted recombinant protein also included optimization of the cultivation conditions. It was observed that, the pH of the cultivation media has a major influence on the proteolytic degradation of the secreted proteins. Acid aspartic protease is the main protease active at pH 5 or lower. [69] At pH 6 it could be shown that trypsin-like alkaline serine protease was the main secreted protease. [70] Based on this, one of the most promising

approaches for recombinant protein production would be pH control (pH 6) and a knockout on the alkaline protease gene. This strategy was shown to bring some success and recombinant proteins showed improved stability when the alkaline protease was knocked out [71].

Table 5 List of organisms with corresponding enzymes and yields and promoter for industrial heterologous protein production adapted from Teuschler, 2017 [65] **and expanded by data from Nevalainen et al., 2014** [43]

Expression Host	Enzyme	Origin of gene	Promoter	Yields	Reference
Aspergillus oryzae	Aspartyl protease	*R. miehei*	amyA	3.3 g/L	Christensen *et al.* (1988)
Aspergillus oryzae	Laccase	*Scytalidium thermophilum*	amyA	1.8 g/L	Berka *et al.* (1995)
Aspergillus oryzae	Xylanase	*Thermomyces lanuginosus*	-	-	X2753 - Novozymes
Aspergillus niger	Lysozyme	*Chicken*	glaA	0.18 g/L	Wongwicharn *et al.* (1999)
Aspergillus niger	Phytase	*Aspergillus ficuum*	glaA	2.8 g/L	van Gorcom *et al.* (1991)
Aspergillus niger	Human interleukin-6	*Human*	glaA	0.3 g/L	Punt *et al.* (2002)
Aspergillus niger	Lipase	*Candida antarctica*	-	-	L3170 - Novozymes
Aspergillus awamori	Chymosin	*Calf*	glaA	1.3 g/L	Dunn-Coleman *et al.* (1991)
Acreomonium chrysogenum	Alkaline protease	*Fusarium sp.*	Alp	4 g/L	Morita *et al.* (1994)
Myceliophthora thermophila	Hydrolytic enzymes	*Trichoderma reesei*	cbhII	-	Gusakov *et al.* (2005)
Neurospora crasse	HT186-D11 (scFv)	Human antibody fragment	Pccg1nr	3 mg/L	Havlik *et al.* (2017)
Pichia pastoris	Glucose oxidase	*Penicilum notatum*	AOX1	2.5 g/L	Gao *et al.* (2012)
Pichia pastoris	Cellobiohydrolase II	*Trichoderma reesei*	AOX1	6.55 g/L	Mellitz *et al.* (2012)
Trichoderma reesei	Lignin Oxidase	*Phelbia radiata*	cbhI	0.02 g/L	Saloheimo *et al.* (1991)
Trichoderma reesei	Acid Phosphatase	*Aspergillus niger*	cbhI	0.5 g/L	Miettinen-Oinonen *et al.* (1997)
Trichoderma reesei	Chymosin	*Calf*	cbhI	0.04 g/L	Uusitalo *et al* , 1991

Trichoderma reesei	Xylanase	pdc	1.52 g/l	Nevalainen et al., 2014

While focussing on preferred hosts, vector design and strain construction the important aspects of bioreactor cultivation and downstream processing have been neglected in this chapter. However, these aspects are for sure as important as the production host itself as they cause a major cost fraction on the total manufacturing process. Finally product recovery, polishing and formulation are important factors in commercial enzyme manufacturing.[72]

2. Objectives

This master thesis built up on the previous work described in the master thesis of Nico Teuschler [65]. The basic vector design of Teuschlers work was maintained as it worked for the expression of recombinant proteins with monodirectional promoters. As no vector for recombinant protein production with bidirectional promoters was available, it was necessary to design and construct a new vector.

The DNA sequences of the vector of Teuschler should be used as basis of the new vector as only a few changes within the area of the expression cassette would be necessary. Discussing the results and checking the final sequencing results of his work showed that many point mutations and incorrect assembly occurred. Furthermore, although resulting in functional expression of reporter protein the synthetic DNA for the assembly of his plasmid showed also some incorrect/ unknown sequences. For this reason, my supervisor Dr. Christoph Reisinger and I decided to start from scratch as it would be less time consuming to assemble a new plasmid than debugging the pre-existing vector. The existing differences in the existing DNA compared to the designed vector might have some significant influence on transformation efficiency as well as expression strength.

A successful use of bidirectional promoters in *T. reesei* has not been shown before, what made it necessary to find a suitable efficient bidirectional promoter sequence. In the work of Teuschler [65] bidirectional promoters transcribing in *Trichoderma reesei* for histones are described. Nevertheless, a systematic approach to identify potentially efficient new bidirectional promoters in the *T. reesei* genome had never be performed before and new sequences originating from such search can enhance diversity and choice for future projects.

A third and key objective was to establish an efficient protocol for the transformation of *T. reesei* as the protocol of the previous master thesis resulted in low transformation efficiency of 3.5 mycelia forming units (mfu) per µg of DNA. [65]

3. Material and Methods

3.1 Equipment and devices

All devices, which were used during the master thesis are listed in Table 6. Plastic consumables like pipet tips, reaction tubes and Petri dishes (not listed) were obtained from Greiner Bio-one and Biozymes.

Table 6 List of all devices used during the master thesis

	Device	Company
Centrifuge	Eppendorf Centrifuge	Eppendorf AG
	Eppendorf Centrifuge	Eppendorf AG
PCR-Cycler	Applied Biosystems 2720 Thermal Cycler	Applied Biosystems, Inc
Plate Reader	FLUOstar	BMG LABTECH
Pipettes	VWR Signature™ Ergonomic High Performance Single-Channel Pipettors 100-1000µl	VWR
	VWR Signature™ Ergonomic High Performance Single-Channel Pipettros 20-200µl	VWR
	VWR Signature™ Ergonomic High Performance Single-Channel Pipettors 2-20µl	VWR
	Denville Xl 3000i single channel pipette 0.1-2 µL	Denville Scientific Inc
Gel-electrophoreses	SubCell GT	BioRad
	PowerPac Basic	BioRad
	GelDoc-It Imaging System	UVP
	Microwave	BOMANN
SDS-PAGE	Mini GEL Tank	Invitrogen
	Power-Ease 300W	life Technologies
	Shaker	GFL
Others	CertoClav EL	CertoClav
	Hei- Standard (Mixer)	Heidolph
	MR 2002 (Mixer)	Heidolph
	SenTix 21 PLUS® pH electrode	WTW, Weilheim, Germany
	NanoDrop 2000c Spectrophotometer	peqlab Biotechnologie GmbH
	Leica DM LB microscope	Leica Mikrosysteme GmbH
	Leica DM LB2 microscope	Leica Mikrosysteme GmbH
	Leica DFC 350 FX camera	Leica Mikrosysteme GmbH
	UV-Lamp ebq 100 isolated	Leica Mikrosysteme GmbH

Neubauer cell count chamber	Paul Marienfeld GmbH & Co. KG,
Miracloth roll filter material 22 - 25 µm	Millipore Sigma-Aldrich GmbH
Glass wool	Roth
Eppendorf BioPhotometer plus	Eppendorf AG, Hamburg, Germany
Incubation Water Bath Model 1003	Lactan Vetriebsgesellschaft m.b.H. & Co KG

3.2 Media, Buffer and Chemicals

All media used in the master thesis are listed in Table 7. The amount of used ingredients is for 1 litre of medium. All media can also be solidified by using 15g of Agar-Agar, KobeI (Roth).

Table 7 List of all media used during the master thesis. The amount of ingrediants is for 1 litre of medium. All media can be solidified by adding 15g of Agar-Agar, KobeI.

Name	Component	Amount/l	Company
LB Medium (Lennox)	Yeast Extract	5g	Roth
	Tryptone	10g	BD
	NaCl	5g	Roth
	Ampicillin	100mg	
PDA	Potato-infuse	6.5g	Roth
	Glucose	20g	Roth
Mandel's Medium Glucose	Tween 80	0,2ml	Merck
	Peptone	0.2g	BD
	Sol1 MM 100x	10ml	
	Sol2 MM 100x	10ml	
	Sol3 MM 100x	10ml	
	Sol4 MM 100x	10ml	
	Sol5 MM 100x	10ml	
	Glucose	1%	Roth
Mandel's Medium Avicel	Tween 80	0,2ml	Merck
	Peptone	0.2g	BD
	Sol1 MM 100x	10ml	
	Sol2 MM 100x	10ml	
	Sol3 MM 100x	10ml	
	Sol4 MM 100x	10ml	
	Sol5 MM 100x	10ml	
	Avicel	1%	Roth
Mandel's Medium Lactose	Tween 80	0.2ml	Merck
	Peptone	0.2g	BD
	Sol1 MM 100x	10ml	
	Sol2 MM 100x	10ml	
	Sol3 MM 100x	10ml	
	Sol4 MM 100x	10ml	

	Sol5 MM 100x	10ml	
	Lactose	1%	Roth
Bottom-Medium	Sucrose	85.6g	Roth
	Glucose	10g	Roth
	Sol1 MM 100x	10ml	
	Sol2 MM 100x	10ml	
	Sol3 MM 100x	10ml	
	Sol4 MM 100x	10ml	
	Sol5 MM 100x	10ml	
	HygromycinB	100mg	Formedia Ltd
Overlay-Medium	Sucrose	342,3g	Roth
	Glucose	10g	Roth
	Sol1 MM 100x	10ml	
	Sol2 MM 100x	10ml	
	Sol3 MM 100x	10ml	
	Sol4 MM 100x	10ml	
	Sol5 MM 100x	10ml	
	Agarose	10g	VWR
Sporulation medium	PDA	39g	Roth
	Glycerine	100g	Roth
	Peptone	1g	BD
Spore wash solution	NaCl	8g	Roth
	Tween 80	0.5ml	Merck

Stocks Sol1 MM 100x- Sol5 MM 100x had to be added after autoclaving.

The buffers listed in Table 8 were used during the thesis. The amount of ingredients is for 1 litre of buffer.

Table 8 List of buffers used during the master thesis The amount of ingrediants is for 1 litre of buffer

Name	Component	Amount/l		Company
Lysis Buffer	SDS	3%		Roth
	EDTA	0.5mM		Roth
	NaCl	1M		Roth
	Tris HCl	100mM		Roth
Lysing Solution	Lysing Enzyme	15g	use solution A for solving	Sigma-Aldrich
Protoplast Washing Solution	Bis-Tris	10mM	pH 5,8	Roth
	$MgSO_4$ $7H_2O$	1,2M		Roth
Transformation Solution A	Tris	100mM	pH 7	Roth
	Sorbitol	0,6M		Roth

Transformation Solution B	Tris	10mM	pH 7,5	Roth
	$CaCl_2$ $2H_2O$	50mM		Roth
	Sorbitol	1M		Roth
Solution for Protoplast focusing	Tris	10mM	pH 7,5	Roth
	$CaCl_2$ $2H_2O$	50mM		Roth
	Sucrose	1M		Roth
PEG-Solution	PEG6000	250g	pH 7,5	Roth
	Tris	10mM		Roth
	$CaCl_2$ $2H_2O$	50mM		Roth

Autoclaving could be used to sterilize all buffers mentioned above.

The stocks listed in Table 9 were used to prepare the different buffers and media for this thesis.

Table 9 List of stocks which were made to prepare buffers and media. The amount of ingrediants is for 1 litre of stock.

Name	Component	Amount		Company
1M Tris	Tris	121g		Roth
	HCl		pH 8	Roth
0.5M EDTA	Na-EDTA	186.2g		Roth
Sol1 MM 100x	$(NH_4)_2SO_4$	140g		Roth
Sol2 MM 100x	KH_2PO_4	200g		Roth
Sol3 MM 100x	$MgSO_4$ $7H_2O$	30g		Roth
Sol4 MM 100x	$CaCl_2$ $2H_2O$	40g		Roth
	$FeSO_4$ $7H_2O$	0.5g		Roth
	$MnSO_4$ $7H_2O$	0.16g		Roth
	$ZnSO_4$ $7H_2O$	0.14g		Roth
	$CoCl_2$ $7H_2O$	0.2g		Roth
Sol5 MM 100x	Urea	30g		Roth

With exception of Sol4 MM 100x which is containing the different salts for the Mandel's media all the stocks can be autoclaved. Sol4 MM 100x should be sterilized by sterile filtration using a membrane with a pore size of 0.2µm.

3.3 Plasmid storing

The plasmids produced during this master thesis were stored in two different ways.

1000ng of all plasmids were dried on filter paper and stored including the correct plasmid maps in the general plasmid collection of the group.

E. coli glycerol stocks bearing the final vectors for integration and expression in *T. reesei* were prepared and added to the IMBT strain collection (Complete list including the IMBT numbers is attached to the appendix (Table 28)). For the glycerol stocks, a fresh ONC was prepared and mixed it with glycerol to a final concentration of 25% before storing at -80°C.

3.4 Strains and Codon usage

3.4.1 *Trichoderma reesei* strains

During this master thesis, five *T. reesei* strains: QM6a, QM9414, RutC30, NRRL15499, NRRL15500 and NRRL15502 (Table 10) were used. Dr. Christoph Reisinger ordered the strains NRRL15499, NRRL15500 and NRRL15502 in May 2017 at NRRL, USA. The NRRL strains invented and patented by Seigo and Yasushi in 1989 [73] have in common, that they have been modified to be inducible by L-sorbose to produce cellulases. The strains origin is in *T. reesei* QM9414. Compared to QM9414, the NRRL strains show slower growth and have also lost their colour when cultivated on PDA media. NRRL15499 and NRRL15500 have, a worse spore formation ability compared to QM9414. The strain NRRL15502 has stronger spore formation ability than QM9414 [73]. Unfortunately, no document which lists the mutations could be found.

Table 10 List of different *T. reesei* strains used in the master thesis.

Organism	Strain	IMBT number	Entry date
Trichoderma reesei	QM6a	2133	17.03.1987
Trichoderma reesei	QM9414	2041	06.04.1977
Trichoderma reesei	RutC30	2111	02.04.1985
Trichoderma reesei	NRRL15499	8054	23.05.2017
Trichoderma reesei	NRRL15500a	8055	23.05.2017
Trichoderma reesei	NRRL15500b	8056	23.05.2017
Trichoderma reesei	NRRL15500c	8057	23.05.2017
Trichoderma reesei	NRRL15502	8058	23.05.2017

3.4.2 *Escherichia coli* strains

E. coli XL1- Blue from the strain collection of the Graz University of Technology (IMBT number: 3095; Entry date: 04.02.2010) was used for molecular biological applications.

3.4.3 Codon usage

Table 11 the codon usage of *T. reesei* with the different codons, the corresponding amino acids and the frequency with which the codon is used in nature (in percentage) is listed. In the last column, the codons for manual optimization are listed. Optimization was done by hand and all codons which were rarely used in the wild type strains were exchanged in order to enhance translation efficiency. The codon usage was further adapted to avoided the restriction sites for the enzymes: *Hpa*I, *Spe*I, *Not*I, *Pst*I, *Fse*I and *Sph*I. All sequences, which were codon optimized, are listed in the appendix.

Table 11 Codon usage table of *T. reesei* which was used for adapting the coding sequences of the genes expressed in *T. reesei*. Based on publication of S.J. Te'o et al. [67]

	Codon	Codon usage in *T. reesei* (%)	Used
Gly	GGG	8	
	GGA	14	
	GGT	18	
	GGC	60	GGC
Glu	GAG	82	GAG
	GAA	18	
Asp	GAT	31	
	GAC	69	GAC
Val	GTG	23	
	GTA	4	
	GTT	17	
	GTC	56	GTC
Ala	GCG	17	
	GCA	12	
	GCT	22	
	GCC	49	GCC
Arg	AGG	13	
	AGA	7	
	CGG	12	
	CGA	16	
	CGT	17	
	CGC	35	CGC
Ser	AGT	5	
	AGC	25	AGC
	TCG	22	
	TCA	8	
	TCT	16	

	TCC	25	
Lys	AAG	92	AAG
	AAA	8	
Asn	AAT	19	
	AAC	81	AAC
Met	ATG	1	
Ile	ATA	4	
	ATT	32	
	ATC	64	ATC
Thr	ACG	29	
	ACA	10	
	ACT	20	
	ACC	41	ACC
Trp	TGG	100	TGG
Stop	TGA	19	
	TAG	44	TAG
	TAA	37	
Cys	TGT	24	
	TGC	76	TGC
Tyr	TAT	25	
	TAC	75	TAC
Leu	TTG	10	
	TTA	1	
	CTG	36	CTG
	CTA	2	
	CTT	14	
	CTC	38	CTC
Phe	TTT	38	
	TTC	62	TTC
Gln	CAG	81	CAG
	CAA	19	
His	CAT	17	CAT
	CAC	83	CAC
Pro	CCG	22	
	CCA	12	
	CCT	25	
	CCC	40	CCC

3.5 Vectors for transformation

One objective of the master thesis was to design and assemble a new vector for homologous integration at the *CBH1* locus by double crossover. The plasmid should also be compatible with bidirectional promoters.

3.5.1 Design

The plasmid design is based on Teuschlers work [65]. The insilco work including primer design was done with SnapGene. In difference to the monodirectional promoters in the thesis of Teuschler this thesis should proof that bidirectional promoters (BDPs) work for recombinant protein production.

The designed vector is a shuttle plasmid. For replication and selection of the plasmid during cloning procedures in *E. coli*. a pUC-origin of replication and an ampicillin resistance marker had to be included.

For heterologous protein production in the final host *T. reesei* the expression cassette should be integrated by double crossover, facilitated by sequences homologous to the *T. reesei* genome. CBH1 (Cel7a) is the best secreted protein in *T. reesei* and was therefore a preferred site for directed integration of the expression cassette. Directed integration by double crossover is facilitated using homologous regions corresponding to parts of the promoter and terminator of *cel7a*. As upstream homologous site, 1500 base pairs starting at position -150 based on *cbh1* were included.

1500 base pairs starting from +2024 based on the *cbh1* were assumed as terminator region of the *cbh1*. Next to the function as homologous integration site, the terminator had a double function as terminator sequence for the selection marker.

Selection in *T. reesei* is done by Hygromycin B (*hph*). To ensure constitutive expression of the *hygromycin* transcription was controlled by the pyruvate kinase promoter of *T. reesei* which was shown to be successfully by previous experiments of Teuschler [65] and published by Mach et al.[74]

Due to the fact, that the genomic DNA of eukaryotic organisms is located in the nucleus, the plasmid had to be linearized to ensure the transport into the nucleus. The restriction enzyme *Hpa*I, which was introduced to the flanking regions linearized the plasmid in a way that exposes the integration sites at both ends. On both sides, a blunt overhang of three base pairs remained on the overhangs.

Codon optimized *eGFP* and *sTomato,* were used to test our new expression vectors. Two different marker proteins were needed to evaluate promoters, which were able to transcribe two coding sequences. The expressed proteins should show green (eGFP) and red (sTomato) fluorescence if everything works as assumed.

The vector design included a possible exchange of the fluorescence proteins by using

restriction enzymes. eGFP can be exchanged using *Not*I and *Sph*I. *Spe*I and *Fse*I allow the exchange of sTomato. For possible industrial applications lox sites for marker recycling were introduced upstream and downstream of the selection cassette. The lox sites 66 and 71were chosen as recognition sequence for the Cre recombinase as described in the publication of Zhang and Lutz. [75]

Another objective of the thesis was to see if bidirectional promoters work at all and if they can be used for heterologous protein production. For easy exchange of the promoter sequences by homologous recombination the plasmid could be linearized with *Sbf*I. A second method for promoter exchange was to use *Not*I and *Fse*I.

3.5.2 Promoter sequence

The first task was to find additional bidirectional promoter sequences to the ones already shown by Teuschler. Therefore, a new search tool was developed by Dr. Reisinger and me which used Excel spreadsheets and customized VBA (Visual Basic for Application) codes to extract bidirectional promoter sequences from the annotated genome sequence of *Trichoderma reesei* QM6a published by Martinez et al [76] in connection with RNA sequencing data published by Kubicek et al. [77]. The tool, is able to identify DNA regions, by applying two conditions: First all open reading frames (ORF) are identified which are next to each other, pointing in opposite (Figure 10) directions and which are, second, not separated by an open reading frame as shown in Figure 11. These sequences are hypothetical bidirectional promoter sequences. (Figure 12).

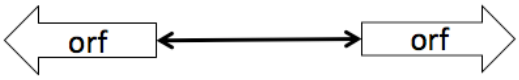


Figure 10 Condition 1 of the tool looking for bidirectional promoters in the genome of *T. reesei*. Hypothetical bidirectional promoter sequence implied by the black arrow

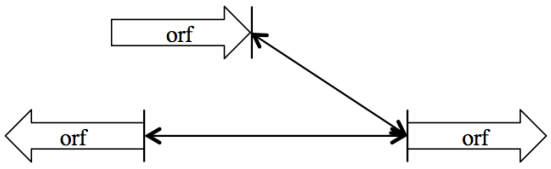


Figure 11 Condition 2 which implies that the hypothetical promoter sequence is no promoter sequence because another other open reading frame is located within the hypothetical promoter sequence.

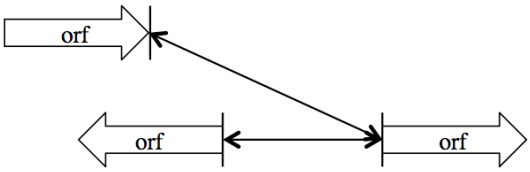


Figure 12 Condition 2 which implies that the hypothetical promoter sequence is a promoter sequence because no another other open reading frame is located within the hypothetical promoter sequence.

A second promoter variant, was an artificial bidirectional promoter. Two small, strong monodirectional promoters were used to form a bidirectional promoter element, by combining one promoter tail-to-tail with the reverse compliment of the other promoter.

With both, "natural" and "artificial" bidirectional promoters, three different constellations of Kozak-sequences were tested to fine-tune expression levels. Variant one had the Kozak-Sequence from the promoter itself and an additional CBH1 Kozak-Sequence (later on indicated by "n"). The second variant applied the native Kozak-sequences from the bidirectional promoter elements (later on indicated by "ok"), while the third variants carried the CBH1-Kozak sequence positioned between promoter and open reading frame (later on indicated by "wl"). In Figure 13 schemes of the different Kozak-Sequence versions are shown.

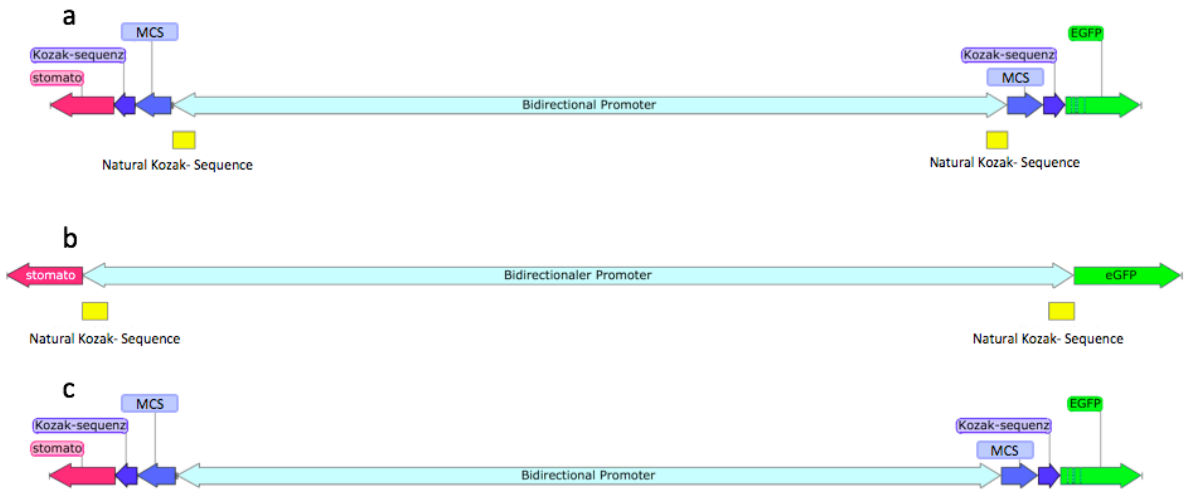


Figure 13 Scheme of the different Kozak-Sequence variations tested for the fine tuning of the expression. (A) shows the variant with two Kozak-Sequences (natural and CBH1) indicated during the theses as "n". (B) shows the variant with the Kozak-Sequence originating from the promoter sequence this variant is indicated as "wl" later on. (C) shows the third variant with the CBH1 Kozak- sequence only which is located within the vector, indicated as "ok"

3.6 Primer

All primers were ordered at IDT, diluted to 100 pmol/µl and stored at -20°C.

The primers in Table 12 were used for amplification of the DNA fragments for the construction of the *T. reesei* expression plasmid. The primers were designed to have an overhang of 30 base pairs to the flanking DNA fragments in order to be used for overlap extension PCR or homologous recombination cloning.

Table 12 Complete list of primers used for amplifying the fragments from the genome or gBlock for further cloning procedures. Primers were designed to have an overhang of 30 base pairs to the flanking fragments for PCR or homologous recombination cloning.

Name	Sequence 5' --> 3'	Internal number
pBSYA2G_HpaI_FWD	ccgaagtgttgttaacttaccaatgcttaatcagtgaggcacc	P17364
pBSYA2G_HpaI_REV	gagccacgtgcgttaactttccataggctccgccccctg	P17365
Cbh1P_pBSYA2G_FWD	cggagcctatggaaagttaacgcacgtggctcaccgaaaagcaag	P17362
Cbh1P_REV	gaggcatggccgacaatactccttgaatgtcaacatttc	P17263
Cbh1T_FWD	ataacttcgtataatgtatgctatacgaacggtaagctccgtggcgaaagcctgacg	P17264
Cbh1T_pBSYA2G_REV	taagcattggtaagttaacaacacttcggtggaggtgtcgag	P17363
Hpa1_PyrT_FWD	gcctccgcgttaagtccgggatgccgatttgtaggtac	P17266
Hpa1_PyrT_REV	tcggcatcccggacttaacgcggaggctctcatcgccg	P17267
HygR-FWD_BstX1	agaaccctcttaaccatgaaaaagcctgaactcaccgccacgtctgtggagaagtttc	P17594
HygR_REV	taccgttcgtatagcatacattatacgaagttatctattcctttgccctcggacgag	P17269
NotI_PkiP_FWD	agttggagcaaagcgaccgccatgggagcagcgaacc	P17270
NotI_PkiP_REV	ctcccatggcggtcgctttgctccaactcaggcgatg	P17271
PkiP_FWD_kurz	tcttaccgttcgtataatgtatgctatacgaagttatcggagggatgccgtgctttgtcg	P17453
PKIP_REV_BSTX1	gaaacttctccacagacgtggcggtgagttcaggcttttcatggttaagagggttcttc	P17567
PkiT_FWD	ctcgggtatgcatgcacgaaatgggaaaggatacacaaaatggaaatctgagc	P17274
PkiT_REV	ataacttcgtatagcatacattatacgaacggtaagagatttgagattaagaacttg	P17275
PyrT_FWD	gacattcaaggagtattgtcggccatgcctcttctctg	P17276
PyrT_REV	tgtcgaggactagtggcagttgtcgacgatatcagcttc	P17277

Table 13 lists all primers which were used for the amplification of the bidirectional promoters. The primers are designed for homologous recombination as well as restriction cloning which should allow an easy exchange in the final construct.

Table 13 List of primers used to amplify the promoter sequences form the genome. Primers were designed for homologous recombination cloning to the final expression plasmid.

Name	Sequence 5' --> 3'	Internal number
PC_CBH1_oK_REV	ggccggcccacgacacctaccgcggttgactattgggtttctgtgcctc	P17335
PC_CBH1_REV	ggccggcccacgacacctagatgcgcagtccgcggttgactattg	P17336
PC_CBH1_FWD	cacctctgacttcacttagccagggatgcttgagtgtatc	P17337
PC_PDC_FWD	gcggccgcaggacgacctgattgtgctgtagctgcgctgc	P17338
PC_PDC_oK_FWD	gcggccgcaggacgaccttagctgcgctgctttgatcgttttgaggtg	P17339
PC_PDC_REV	agcatccctggctaagtgaagtcagaggtggtcgttaattg	P17340
PC_PDC_WL_FWD 2	gcccttgctcaccatgattgtgctgtagctgcgctgctttg	P17341
PC_PDC_WL_REV 2	gcccttcgagaccatgatgcgcagtccgcggttgactattgg	P17342
GHiston_FWD	gcggccgcaggacgacctttgaaaataacgttgatttc	P17343
GHiston_oK_FWD	gcggccgcaggacgacccgttgatttcaaatagatttcagagactg	P17344
GHiston_oK_REV	ggccggcccacgacaccggggtttttgcttaacgtggaaaaagtaac	P17345
GHiston_REV	ggccggcccacgacacctttgaagttgggggtttttgcttaac	P17346
GHiston_WL_FWD 2	gcccttgctcaccattttgaaaataacgttgatttcaaatag	P17347
GHiston_WL_REV 2	gcccttcgagaccattttgaagttgggggtttttgcttaacg	P17348
Ribosom_BiDi_76939	gcggccgcaggacgacctcgcgaccgattaattgccgtcg	P17349
Ribosom_BiDi_oK_FWD	gcggccgcaggacgaccttaattgccgtcgatattcgttcgaggattg	P17350
Ribosom_BiDi_oK_REV	ggccggcccacgacacctgtcagaggttatggctcggacgggtaag	P17351
Ribosom_BiDi_WL_FWD 2	gcccttgctcaccatcgcgaccgattaattgccgtcgatattc	P17352
Ribosom_BiDi_WL_REV 2	gcccttcgagaccattctgtagaattgtcagaggttatggctc	P17353
RibosomBiDi_REV	ggccggcccacgacacctctgtagaattgtcagaggttatg	P17354

Table 14 lists the primers which were used for the sequencing of the final vector constructs. Primers were designed to have an GC content of 50% and a melting temperature of about 55°C. Sequencing was done by GATC, LGC and Microsynth.

Table 14 List of primers used for sequencing the final plasmid.

Name	Sequence 5' --> 3'	number
Seq_1	gtgatgacggtgaaaacctc	P17278
Seq_2	cttcgacaagcaaagcgttc	P17279
Seq_3	ggtaagaatgtctgactcgg	P17280
Seq_4	cgacacaaatagccatcagg	P17281
Seq_5	gaactcatggccgttcatgc	P17282
Seq_6	gaccactaccagcagaacac	P17283
Seq_7	gtgaaagttagcgtcaaggg	P17284
Seq_8	aaaatccaaccactgacggc	P17285
Seq_9	agacctgcctgaaaccgaac	P17286

Seq_10	tactcgccgatagtggaaac	P17287
Seq_11	acaagggacgcaaagttgtc	P17288
Seq_12	tggtggagatctctctatcg	P17289
Seq_13	tcagttcggtgtaggtcgttc	P17290
Seq_14	tcagcaataaaccagccagc	P17291

3.7 Enzymes

Table 15 lists all enzymes used in the master thesis.

Table 15 List of enzymes used in the master thesis.

Enzyme	Type	Company
*Hpa*I	RE	Thermo Fisher Scientific Inc.
*Spe*I	RE	Thermo Fisher Scientific Inc.
*Not*I	RE	Thermo Fisher Scientific Inc.
*Pst*I	RE	Thermo Fisher Scientific Inc.
*Fse*I	RE	Thermo Fisher Scientific Inc.
*Sph*I	RE	Thermo Fisher Scientific Inc.
Q5 DNA Polymerase	Polymerase	New EngIand Biolabs
Taq DNA Polymerase	Polymerase	New EngIand Biolabs
T4 DNA Ligase	Ligase	Thermo Fisher Scientific Inc.
Lysing Enzyme from Trichoderma harizanum		Sigma-Aldrich

3.8 Kits and Protocols

3.8.1 Kits

All kits used during the thesis are listed in Table 16.

Table 16 List of all kits used during the master thesis

Name	Company
Wizard Plus SV Minipreps DNA Purification System	Promega GmbH, Mannheim, Germany
Wizard SV Gel and PCR Clean-Up System	Promega GmbH, Mannheim, Germany
CloneJET PCR Cloning Kit	Thermo Fisher Scientific Inc.
Mix & Go! *E. coli* Transformation Buffer Set	Zymo Research

3.8.2 Plasmid isolation from *E. coli*

Wizard Plus SV Miniprep and DNA Purification System from Promagea GmbH was used for isolating the plasmid DNA from the cells. Purification was performed in accordance with the manufactures manual. Deviating from the manual 100 µl of nuclease free water was used for the elution of the DNA. For better results the water was preheated at 55°C and an incubation

step of 5 minutes on 55°C was performed before the last centrifugation step.

3.8.3 DNA isolation from gels

Wizard SV Gel and PCR Clean-Up System from Promega GmbH was used for DNA isolation from agarose gels. Isolation was performed in accordance with the manufactures manual. Deviating from the manual, a centrifugation step after washing, 10 min at top speed, was added and the sample was incubated for 10 min on 37 °C to be sure all ethanol was removed before eluting with 50 µl of ddH_20.

3.8.4 Transformation of *E. coli*

We used chemical competent *E. coli* XL1blue cells for all cloning steps. The competent cells were made with the Mix & Go! *E. coli* Transformation Buffer Set. All steps were performed in accordance with the manufactures manual, using low temperature cultivation (22°C) of the main culture overnight. The competent cells were stored at -80 °C in aliquots of 50 µl until use.

50 µl of competent *E. coli* XL1 blue cells and 500 ng of pure plasmid DNA were used for the transformation procedure. The mixture was at first incubated on ice for 10 min, then transferred to 42°C for 1 min for transformation. Afterwards the cells were regenerated using 500 µl of SOC-media and incubated for 30 min before plating on agar plates with the correct antibiotic.

3.8.5 Isolation of genomic DNA from *T. reseei*

For genomic DNA isolation, the cell material was frozen in liquid nitrogen, grinded and finally used for a chloroform/phenol extraction.

An overnight culture of *T. reesei* Q6MA, in 50 ml PDA-media was incubated for 24 h at 28°C in shake flasks. The next day the culture was transferred to a 50ml tube and centrifuged for 10 min at maximum speed. After discarding the supernatant, cells were dried on filter paper to remove as much cultivation media as possible.

The mortar was pre-cooled with liquid nitrogen before the mycelium was transferred form the filterer paper to the mortar and processed to a fine powder. In the next step, 10 ml lysis buffer per gram of cell-powder was added.

An equal amount of chloroform:phenol (1:1), was added before a 10 min centrifugation step. The resulting aqueous phase was transferred to a new tube and an equal amount of isoamylalcohol was added to remove the remaining proteins. 10 min of centrifugation at

maximal speed was performed before transferring the aqueous phase to a new tube. Ice-cold isopropanol to precipitate the DNA and centrifuged for 10 min was used to get a DNA pellet. After discarding the supernatant, the DNA was washed with 75% ethanol and centrifuged at the same conditions as before. After discarding the supernatant, the DNA was dried at 37 °C until no pellet was seen. 500 µl of ddH_2O was used to dissolve the DNA. The final DNA concentration was determined with the NanoDrop.

3.8.6 Restriction analysis

For a preliminary and quick analysis of the plasmids, analytical restriction digest and subsequent gel electrophoreses was used. Components for the reaction mixtures and the corresponding amounts are shown in Table 17. About 200 ng of DNA were used for the restriction analysis. The reaction was incubated for 1 hour at 37 °C and stopped by adding 4 µl of 6x Loading Dye before being loaded on an 1% agarose gel. As reference marker, the GeneRuler 1KB from Thermo Fischer scientific was used.

Table 17 General Mix for restriction analysis

Components	Volume
Buffer 10X	2µl
RE 1	0.2µl
Re 2	0.2µl
DNA	x µl
ddH20	17.6-x
Total	20µl

3.8.7. Ligation

For ligation, two different protocols were used. For ligation into the pJET1.2 backbone the protocol shown in Table 18, which is in accordance with the producer manual was applied.

Table 18 Ligation mix for pJET1.2 cloning

Components	Volume
Buffer 2x	10µl
T4 Ligase	1µl
pJET1.2 plasmid	1µl
Insert (gel purified PCR product)	1µl
ddH2O	7µl
Total	20µl

For all other ligations, the ratio of insert and backbone was 3:1 and the protocol in Table 19

was applied

Table 19 Ligation mix for general ligation

Components	Volume
Buffer 2x	10µl
T4 Ligase	1µl
Plasmid Backbone	x1 µl
Insert	x2 µl
ddH20	9µl -x1-x2
Total	10µl

Ligation reactions were incubated for 60min at room temperature or overnight at 18 °C. For amplification of the assembled constructs, *E. coli* XL1 cells were transformed with 5 µl of the reaction mixture.

3.8.8 Polymerase chain reaction

All PCR reactions were performed using the Q5 DNA Polymerase from New England Biolabs (NEB) with the compatible buffers. For standard reactions, the mix listed in Table 20 was used. In general, between 15 and 25 ng of DNA was used as template.

Table 20 PCR reaction mix

Components	Volume
Template	x µl
5x Q5 Reaction Buffer	10 µl
5x Q5 high GC-enhancer	10 µl
dNTPs 2mM	5 µl
Primer FWD 100pmol/µl	0.5µl
Primer REV 100pmol/µl	0.5µl
ddH20	23.8-x
Q5 DNA Polymerase	0.2µl
Total	50µl

For overlap extension PCR (oePCR) the mix listed in Table 21 was used. The ratio of the fragments was 1:1 and the maximum amount of DNA used in the reactions was 40ng.

Table 21 Reaction mix for oePCR

	Components	Volume
	Template 1	x1
	Template 2	x2
	Primer FWD 100pmol/µl	0.5µl
	Primer REV 100pmol/µl	0.5µl
	5x Q5 Reaction Buffer	10µl
	5x Q5 high GC-enhancer	10µl
	dNTPs 2mM	5µl
	ddH20	23.8-x1-x2
	Q5 DNA Polymerase	0.2µl
Total		50µl

Three different PCR profiles were used during this work. The standard profile used during this thesis was a 2-step PCR-protocol (Profile 1). Fragments which could not be amplified with profile 1 were amplified either with profile 2 (3-step protocol with low annealing temperature) or profile 3 (standard 3-step protocol as described in the manual NEB)

Table 22 Primary 2 step PCR reaction profile 1

Profile 1			
	Temperature	Time	Cycles
Initial Denaturation	98°C	50sec	1
Denaturation	98°C	20sec	35
Annealing& Extension	72°C	30sec/kb+20sec	
Final Extension	72°C	7min	1
Holding	4°C	∞	∞

Table 23 Alternative PCR reaction profile 2

Profile 2			
	Temperature	Time	Cycles
Initial Denaturation	98°C	50sec	1
Denaturation	98°C	20sec	35
Annealing	55°C	20sec	
Extension	72°C	30sec/kb	
Final Extension	72°C	7min	1
Holding	4°C	∞	∞

Table 24 Alternative PCR reaction profile 3

Profile 3	Temperature	Time	Cycles
Initial Denaturation	98°C	50sec	1
Denaturation	98°C	20sec	
Annealing	60°C	20sec	35
Extension	72°C	30sec/kb	
Final Extension	72°C	7min	1
Holding	4°C	∞	∞

3.9 Homologous recombination cloning

The protocol for homologous recombination cloning was based on the method of Gibson et al. The reaction mix was prepared like described in the publication. [78] As the described one pot reaction did not work the protocol was adapted to a 2-step protocol as explained at the SGI-DNA lecture in our department.

For the 2-step protocol 4 tubes with 15µl reaction mix was mixed with 5 µl of DNA. Tube 1 contained the fragments A and B, tube 2 the fragments B and C, tube 3 the fragments C and D and tube 4 contained the fragments D and A. The reaction was incubated at 50 °C for 1 hour. In a second step, 10 µl of tube 1 and tube 3 and 10 µl of tube 2 and tube 4 were combined in a new reaction tube before incubating 1 hour at 50°C. The reaction mix was spun down and used for the transformation of *E. coli* XL1 which was done in quadruplets.

3.10 Gel-Electrophoreses

1% agarose gels containing ethidium bromide were used for DNA separation and analysis. Samples were prepared by adding 6x Loading dye in appropriate amount. Electrophoreses was done at 145V and 400mA until adequate band separation was achieved (30-60 minutes).

3.11 SDS-PAGE

The samples for the SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis) were prepared like described in the manual from NuPAGE® for reductive conditions.

For running the gels, 1x MOPS buffer and 200V for about 40min were used. 10µl of sample were loaded on the gel. The reference for estimating the protein size was PageRuler prestained Protein Ladder from Thermo Fischer scientific

The proteins were stained with InstantBlue™ overnight under constant shaking.

3.12 Construction of a new *T. reesei* expression vector for protein coexpression

Using bidirectional promoters for recombinant protein production in *T. reesei* was absolutely novel and it was also the first time this was done in our department. For this reason, a new vector was needed.

The parts for the plasmid were either amplified from the genome of the *T. reesei* wild type strain QM6a or from gBlocks which were already had stock from the master thesis of Nico Teuschler. [65] One part of the plasmid, the part containing the cloning site, was ordered anew as gBlock at IDT. The gBlock also included the codon optimized *eGFP* and *sTomato*.

Altogether, eight parts had to be amplified for the final vector. For cloning procedures in *E. coli* an ampicillin resistance and pUC origin was amplified from the bisy plasmid pBSYA2G by using the primer pair: pBSYA2G_HpaI_FWD and pBSYA2G_HpaI_REV.

As already discussed above, the *CBH1* promoter and terminator sequence were used for homologous integration. The promoter region was amplified from the genome of *T. reesei* QM6a by using the primers: Cbh1P_pBSYA2G and Cbh1P_REV. The PCR amplifying the terminator region from the genome was done with the primer pair: Cbh1T_pBYSA2G_REV and Cbh1T_FWD. Both parts are located on chromosome VII of *T. reesei*.

The Hygromycin B resistance gene for selection was amplified from the gBlock with the internal number P17171 by using the primer pair: HygR-FWD_BstX1 and HygR_REV. The constitutive pyruvate kinase promoter on chromosome II was amplified from genomic DNA of *T. reesei* QM6a with the primer pair: PkiP_FWD_kurz and PKIP_REV_BstX1.

Bidirectional promoters, need two terminators flanking each side of the cloning site. The terminator region of the *PYR* on chromosome V was amplified from genomic DNA of *T. reesei* QM6a with the primer pair: PyrT_FWD and PyrT_REV. Within the amplified fragment a *Hpa*I restriction site was eliminated by using the primer pair: HpaI_PyrT_FWD and HpaI_PyrT_REV. This restriction site had to be eliminated because the *Hpa*I was used for the linearization of the plasmid for transformation procedure.

The second terminator, the *PKI* terminator, was amplified form genomic DNA of *T. reesei* QM6a by the primer pair: PkiT_FWD and PkiT_REV.

Both terminator sequences (*PYR* and *PKI)* were used and tested in the work of Teuschler and as the design of his work should be maintained they were used again. [65]

The amplified DNA fragments were blunt-end cloned into the pJET1.2 plasmid and the sequence was verified using the primers pJET1.2 forward (P17471) and pJET1.2 reverse (P17472).

Clones showing the correct sequence were used as templates for the subsequent amplifications, with the corresponding primer pair, to get the correct DNA parts for the plasmid.

In the next step, the DNA fragments, which had overhangs to each other, were used for overlap extension PCR (oePCR), using the outer forward and revers primers of the fragments which should be linked. Resulting from this, we had 4 constructs, which were cloned into the pJET2.1 and send for sequence verification.

To assemble the final plasmid, the correct sequences were amplified from the pJET2.1 plasmids with the fragments of correct sequence. For final homologous recombination cloning we used the protocol as described above.

For transformation of the newly assembled expression constructs, chemically competent *E. coli* XL1 cells were transformed with DNA in quadruplicates as explained above.

3.13 Promoter DNA amplification and cloning

Naturally occurring bidirectional promoters from *T. reesei* were tested as well as artificial bidirectional promoters, that we designed and constructed ourselves from known monodirctional promoters of *T. reesei*.

3.13.1 Homolog bidirectional promoters

The first promoter is a bidirectional promoter, which in nature is directing the transcription of a ribosomal protein belonging to the S9, S4 family and an unknown protein. The template sequence for amplification of the promoter from the genome of *T. reesei* QM6ais located on chromosome III. The primer pair Ribosom_BiDi-oK_FWD and Ribosom_BiDi_oK_REV was used for the amplifiaction. The final promoter has a size of 372 base pairs.

3.13.2 Heterologous bidirectional promoters

The second promoter was an artificial construct, where two mono directional promoters were combined head to head in order to form a bidirectional promoter. One side was a constitutive promoter, the *PDC* (pyruvate decarboxylase) promoter located on chromosome II with 740 base pairs. As second part of the promoter, the *CBH1* (cellbiohydrolase1) promoter located on chromosome VII of *T. reesei* QM6a with 506bp was used. The *PDC* part was amplified with

the primer pair: PC_PDC_oK_FWD and PC_PDC_REV. For the *CBH1* part the primer pair: PC_CBH1_FWD and PC_CBH1_oK_REV was used. The primers PC_PDC_REV and PC_CBH1_FWD introduced overhangs of 30 base pairs to the adjacent fragment, which were needed for overlap extension PCR to form the final promoter with a size of 1246 base pairs and a head to head orientation of the two linked DNA fragments.

As third promoter, a histone promoter located on chromosome II of *T. reesei* should be used. The promoter had a size of 2114 base pairs and presumably regulates the transcription of the Histones H2A and H2B. This promoter showed the best transcription levels of the histone promoters in both directions according to our promoter finding tool. The primer pair: GHiston_oK_FWD and GHiston_oK_REV was used for the amplification of the promoter from genomic DNA of *T. reesei* QM6a.

3.14 Linearization and ethanol precipitation

The DNA used for *Trichoderma* transformation had to be linearized and concentrated. For restriction digest, 20 µg of DNA and 2.5 µl of *Hpa*I (5000 units/ml) restriction enzyme was used. The mix was incubated for about 5 hours at 37°Cto ensure total linearization. To be sure linearization has worked correctly an aliquot of 5µl was used for analysing agarose gel electrophoresis.

Clones showing the correct pattern on the gel were used for ethanol precipitation. DNA was precipitated by addition of 1:10 of the DNA volumes of 3M sodium acetate (pH 5.5) and 2.5 volume of ice-cold ethanol (99%) and incubation over night at -20°C.

The next day in the morning the suspension was centrifuged for 20min at 4°C and 16500xg to form a pellet. The supernatant was discarded and the pellet washed twice. The first washing step was done with 1ml cold ethanol (75%) followed by 5 minutes of centrifugation. For the second washing step, 250µl of ethanol (75%) were added followed by another centrifugation step for 5min. To be sure no ethanol is left the pellet was incubated on 37°C for 10min before dissolving the precipitated DNA in 20µl of ddH_2O.

3.15 Using the Neubauer Counting chamber

A Neubauer- Counting Chamber was used to count the protoplasts for the *T. reesei* transformation procedure.

Before every use the chamber and the coverage plate were cleaned with 75% ethanol. Ethanol (75%) was put on the bars flanking the counting part in the middle of the device before

installing the cover plate on the chamber.

10µl of the sample was pipetted into the small gap between chamber and coverage plate. The sample was soaked into the gap by the capillary forces. 9 big squares were counted in order to get a statically realistic protoplast number.

3.16 Spore stocks

When working with *T. reesei* Spore stocks have to be prepared. Spore stocks can be stored at -80°C and are necessary for all future cultivation procedures.

To prepare spores *T. reesei* was cultivated on agar plates under spore forming conditions (Results). After three days of incubation the spores were harvested by adding 5ml of spore wash solution on the plate and for scrubbing with Drigalski spatula and low pressure over the mycelium. By that the spore wash solution turned green by suspended spores (at least at the wild type strains). In the next step, a stock with an OD_{600} of 10 and a glycerol concentration of 25% was prepared before making aliquots of 1ml and freezing at -80°C.

3.17 Trichoderma transformation

The protocol below explains how to produce competent protoplasts and how to transform *Trichoderma reesei* by a so-called PEG-mediated protoplast transformation. The number of competent protoplast is high enough to transform two different plasmids.

The method is based on the work of Penttilä et al. ,1987. In difference to the original protocol the fungi were not grown on cellophane disks on agar plate as Teuschler described it to be easier and faster in liquid culture [65]. Further 10g/l instead of 5g/l of lysing enzyme was used and incubation was for at least three hours. [79]

Before the transformation the buffers and media listed in Table 7 and Table 8 had to be prepared: Lysing Solution (day of transformation), Protoplast Washing Solution, Transformation Solution A, Transformation Solution B, Solution for Protoplast focusing, PEG-Solution, Pre-culture medium, Overlay-Medium (day of transformation) and plates with Bottom-Medium. PEG-Solution could be stored at -20°C and should be put on room temperature in the morning of use.

Following material is needed: two 300ml flasks (sterile), frit (porosity 1), ice bath, Myracloth, glass wool, a funnel, 15ml Falcons (sterile), 50ml Falcons (sterile), Neubauer counting chamber, microscope slides and cover slip, water bath, 250ml flask, 1l flask and spatula (sterile). In the funnel, we put a sheet of Myracloth and the glass wool is put into the frit. Both

parts are autoclaved.

Preparing the protoplasts was the first part of the transformation procedure. A pre-culture of the strain which should be transformed was prepared the day before the actual transformation procedure. 40ml of PDA media was inoculated with 1ml of spore stock (OD_{600} 10) solution and incubated at 30°C and 150rpm.

On the day of transformation 22ml overlay-media per plasmid was autoclaved and stored at 60°C and 500ml of Protoplast Washing Solution, Transformation Solution A, Transformation Solution B and seven 50ml falcons were pre-cooled.

A funnel with Myracloth was used to separate the mycelium from the culture media. The filter cake was washed with 20ml of Protoplast Washing Solution. The mycelium was suspended in 40ml of Lysing solution and incubated at 30°C for at least 3 hours at 250rpm the progress of protoplast formation was observed with the microscope. Protoplast formation was finished when almost no mycelium was seen.

The next steps were performed without exceptions on ice. A frit (porosity 1) with glass wool was used to separate the spores from extant mycelium. To be sure no protoplasts were left in the frit it was washed with max. 20ml of cold Protoplast Washing Solution. The solution which contained the protoplasts was divided into 4 precooled 50ml tubes to a maximum of 15ml.

10ml of cold Transformation Solution A was put in each of the 50ml tube containing the protoplast solution. This step had to be performed really carefully as a sharp boundary should arise between the solutions. (The best way to do it is to hold the tube at a slight angle.) A centrifugation step at 3220xg at 4°C for 30min followed.

The protoplast should be in a murky layer between the solutions. A pipette was used to transfer the protoplast from the layer into a new tube if no layer was seen the top 7ml were used. The protoplasts from two tubes were pooled in one new, pre-cooled 50ml tube (max. 15ml/tube). Transformation Solution B was used to fill the tubes to 45ml before a centrifugation step at 450xg and 4°C was done for 10min.

The supernatant was discarded and the pellet in the remaining liquid suspended. The protoplasts were pooled in one new tube and filled up to 10ml with ice-cold Transformation Solution B before another centrifugation step at 450xg and 4°C for 20min was done.

The supernatant was discarded and the protoplasts were suspended in the remaining liquid,

additionally we added 150µl of Transformation Solution B.

The Neubauer counting chamber was used to count the protoplasts. In the final protoplast solution should have 2x10^7 protoplasts per ml. If concentration of the protoplast is to high Transformation Solution B can be used for dilution.

$$\frac{Protoplasts}{Chambers} * 400 * 10^4 * dilution = protoplasts/ml$$

For transformation, 80µl of protoplast solution, 800ng of linearized DNA and 20µl of PEG-Solution were mixed in a 15ml tube before the mixture was incubated on ice for 20min. Within this incubation time the overlay media was divided into aliquots of 22ml (50ml tubes) and store at 50°C in the water bath.

In the next step 800µl of PEG-Solution was added to the transformation mix and carefully mixed before incubation for 5min at room temperature. After 5 minutes, 1600µl of Solution for Protoplast focusing was added and mixed carefully.

The transformation mix and a 22ml Overlay-media aliquot was combined, mixed and 12.5ml were evenly spread on a Bottom-media plate. After the top- agar was solid the plates were incubated for at least 5 days at 30°C.

After 5 days of incubation, a tweezer was used (sterilization with buns burner after every use) to isolate the transformants from the regeneration media. The transformants were transferred to new plates containing media for spore formation (Results) important was that mycelium was in contact with the new plate before another incubation step for 3 days at 30°C was done.

Spore stocks were prepared from the single colonies from the first isolation round and dilutions prepared to get single colonies for another round of isolation.

Single colonies were picked and transferred to new plates with sporulation media. The spores were harvested and used for the final assays.

The transformation procedure of Trichoderma was very time consuming as two rounds of colony isolations had to be done to be sure we had homokaryotic protoplasts.

The “real” transformation takes about 2 days including overnight culture, protoplast formation and introducing DNA to the protoplasts. It took about 3 days until the first transformants could be seen and 5 days until the colonies could be transferred to new plates. To harvest the spores, it took another 3-4 days of incubation for the cell growth. As this process had to be done twice it took about 4 weeks to be sure the transformants were stable. A basic scheme of

the transformation procedure is shown in Figure 14. A detailed SOP including a complete list of chemicals and troubleshooting is attached to the appendix.

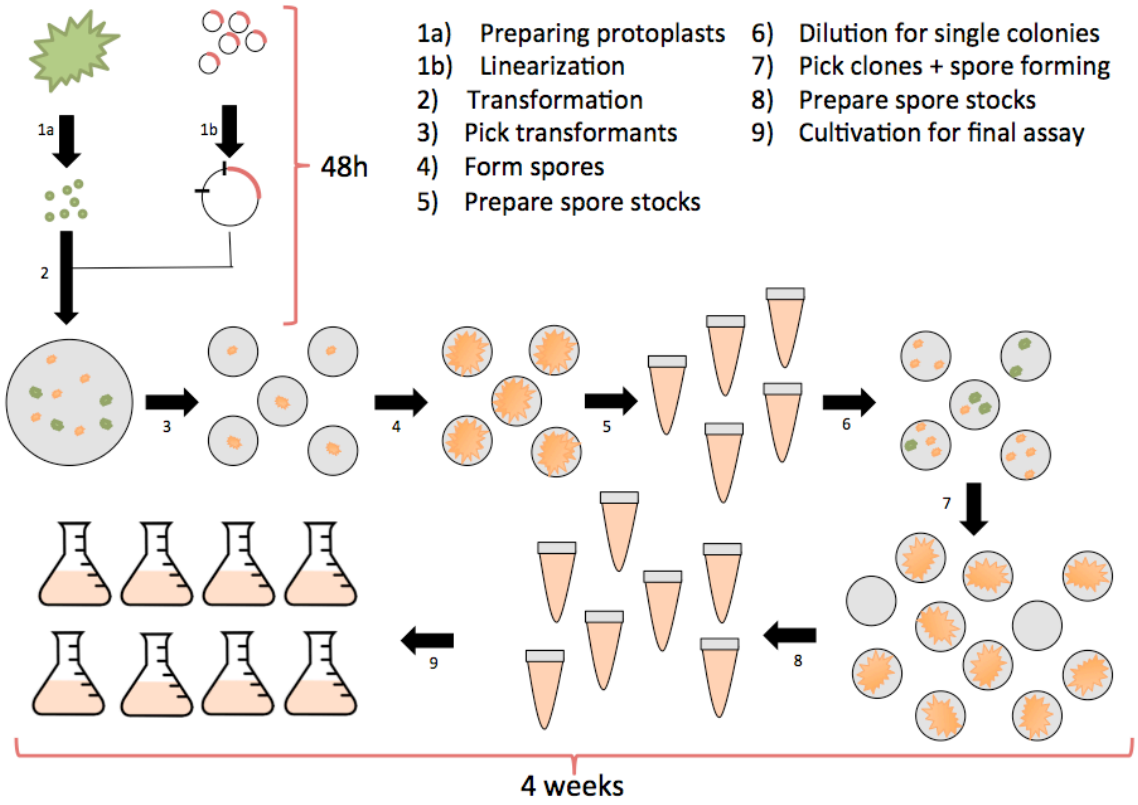


Figure 14 Quick succession for PEG mediated protoplast transformation of *T. reesei.* After preparation of protoplasts and the linearization with included ethanol precipitation of the integration vector. *T. reesei* can be transformed using the protocol above. After 5 days of growing the colonies can be isolated and spore formation can be induced by cultivation on sporulation media plates. After a growing period, the spores can be harvested and diluted for another single colony isolation process. Single colonies are picked and transferred to media enhancing spore formation before spore harvesting. The collected spores are cultivated in 300ml shake flasks with Mandel's media for further testing.

3.18 Isolation of genomic DNA from *T. reesei*

For genomic DNA isolation cell material was frozen in liquid nitrogen, grinded and finally used for a chloroform/phenol extraction.

An overnight culture of *T. reesei* Q6MA, in 50 ml PDA-media was incubated for 24 h at 28°C in shake flasks. The next day the culture was transformed to a 50ml tube and centrifuged for 10 min at maximum speed. After discarding the supernatant, cells were dried on filter paper to remove as much cultivation media as possible.

The mortar was pre-cooled with liquid nitrogen before the mycelium was transferred form the

filterer paper to the mortar and processed to a fine powder. In the next step, 10 ml lysis buffer (3% SDS) per gram of cell-powder was added.

An equal amount of chloroform:phenol (1:1), was added before a 10 min centrifugation step. The resulting liquid phase was transferred to a new tube and an equal amount of isoamylalcohol to remove the remaining proteins was added. 10 min of centrifugation at maximal speed was performed before transferring the liquid phase to a new tube. Ice-cold isopropanol was used to precipitate the DNA and centrifuged for 10 min to get a DNA pellet. After discarding the supernatant, the DNA was washed with 75% ethanol and centrifuged at the same conditions as before. After discarding the supernatant, the DNA was dried at 37 °C until no pellet was seen. 500 µl of ddH_2O was used to dissolve the DNA.

3.19 Phenotypical analysis of transformants

Different methods to detect fluorescence were used to proof the expression of *eGFP* and *sTomato* in *T. reesei*.

Fluorescence microscopy is a fast, easy and stable method the detect fluorescence which allowed also to differentiate between the Trichoderma clones for expression and contaminations. (Figure 15).

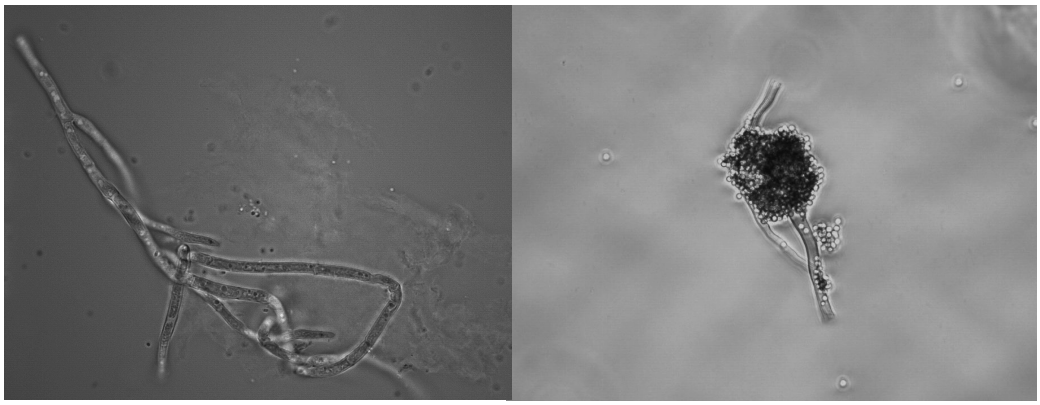


Figure 15 Comparison of morphology between *T. reesei* (left) and most likely a Penicilium strain (right). The supposed Penicillium contamination shows typical bulk of spores which cannot be found in *T. reesei*

To prepare the samples for microscopy an inoculation loop was used to scratch over the surface of the mycelium and to suspend the cells in 100µl of ddH_2O. 10µl of the mixture was transferred to a microscope slide and covered with a cover plate. To see differences in fluorescence different filters for the microscope were used. eGFP was stimulated at a wavelength of 325-435nm and sTomato with filtered light of 515-560nm. Fluorescence of

both proteins could be seen at 340-380nm.

Trichoderma cultures were grown by inoculation of 50ml media with 100µl of spore stocks. Because inhomogeneous growth was observed for the different strains the whole culture was harvested and homogenized with liquid N_2 and a mortar. For each strain, a defined amount (about 100mg) of the cell-powder was divided into two reaction tubes. The sample in one tube was immediately suspended in 800µl in Tris-HCl (pH7.4). 250µl were used for fluorescence measurement (triplicates) and 10µl for protein quantification by Bradford-Assay (triplicates). The sample in the other reaction tube was lyophilized to determine the cell dry weight and used for protein analysis by SDS-PAGE.

4. Results and Discussion

4.1 Cultivation of NRRL strains

Culturing the originally delivered NRRL15500 stock resulted in distinctly different colonial phenotypes when cultured on PDA media, indicating the presence of more than one strain. Three different phenotypes were found which were divided into the three new sub strains NRRL15500a, NRRL15500b, NRRL15500c. Phenotypes of the strains NRRL15500a, NRRL15500b and NRRL15500c, were as followed: yellow coloured with a white edge, green coloured with a yellow spot in the middle and green coloured with a white edge. Figure 16

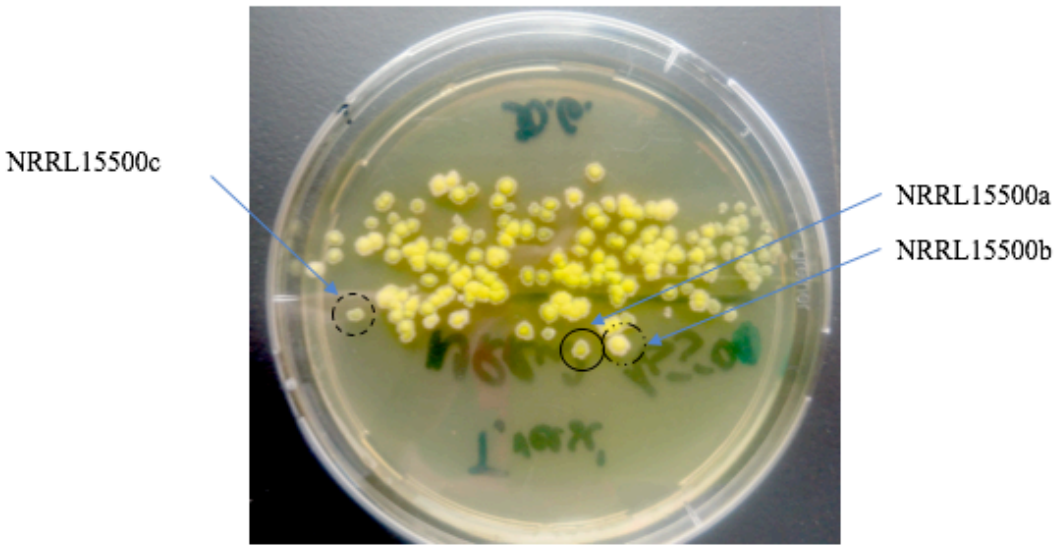


Figure 16. Cultivation of *Trichoderma reesei* NRRL15500 strains ordered from the NRRL (USA) on PDA plates showed three distinctly different phenotypes. The corresponding colonies were isolated and divided into three new sub-strains NRRL15500a, NRLL1550b and NRRL15500c. Strain NRRL15500a had a yellow colour with a white edge while NRRL15500b was green with a yellow spot in the middle and NRRL15500c was green with a white edge

Further experiments shown in this chapter will proof that the three substrains NRRL15500a, NRLL1550b and NRRL15500c behaved dinstictly differner and therfore subclassification of the strains was not obly justified but also neccasery.

4.2. Sporulation behaviour of *Trichoderma reesei* strains cultivated on Agar plates

It is known, that the conditions needed for spore formation can vary tremendously between different *T. reesei* strains.[80] Additional, literature states, that conidiation is triggered by light. Special blue light receptors are responsible for an increase of transcription factors for enhanced spore formation. [81] Therefore, all of the different *T. reesei* strains used in this thesis were in a first step cultivated on different agar plates under varying light and

temperature conditions to find the optimal conditions for sporulation. Specialised media for enhanced spore formation (Table 7, Sporulation media) as well as PDA was used. 100 µl of a spore stock were used for plating. The conditions were set to 28°C with light and 30°C without light. Both sets were incubated for 72 hours.

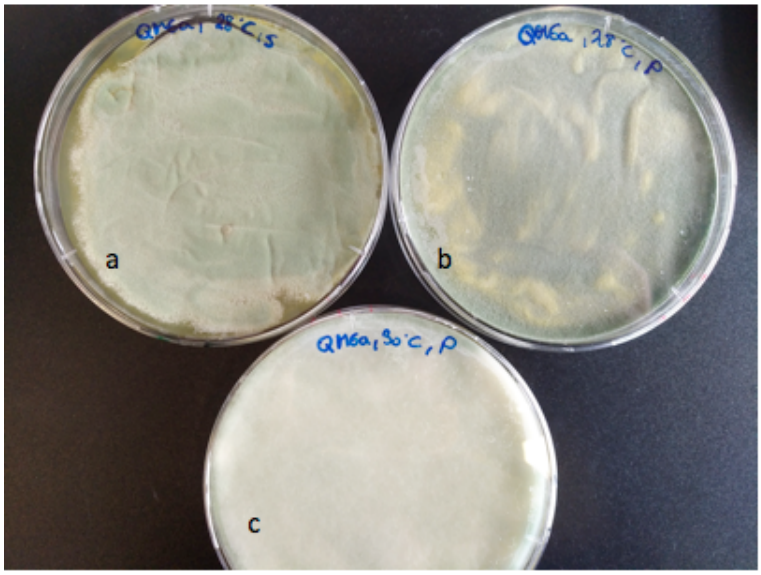


Figure 17 *T. reesei* QM6a on PDA (b,c) and sporulation media plates (a). Plate a and b were incubated at 28°C and day light conditions and plate c at 30°C in the dark. Plate a and b show spore formation which could indicate that QM6a needs day light conditions for spore formation and does not relay on specific media.

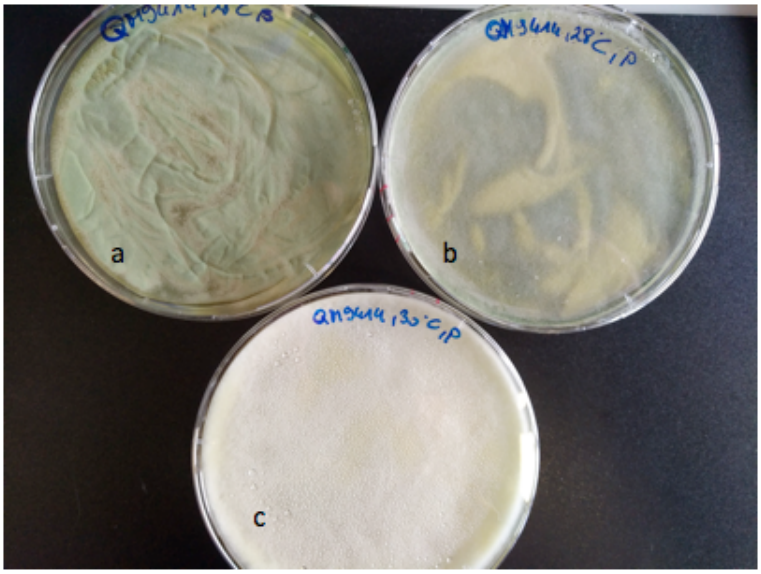


Figure 18 *T. reesei* QM9414 on PDA (b,c) and sporulation media plates (a). Plate a and b were incubated at 28°C and day light conditions and plate c at 30°C in the dark. Plate a and b show spore formation which could indicate that QM9414 needs day light conditions for spore formation and does not relay on specific media.

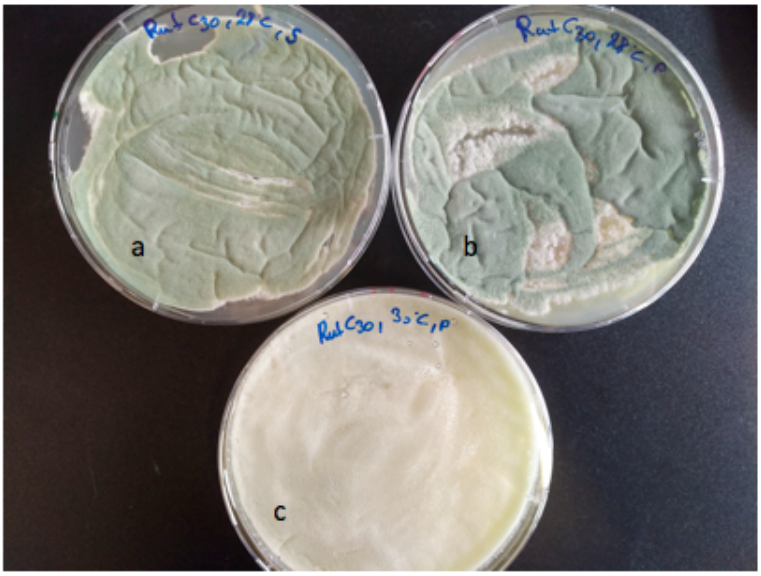


Figure 19 *T. reesei* RutC30 at PDA (b,c) and sporulation mediaplates (a). Plate a and b were incubated on 28°C and day light conditions and plate c at 30°C in the dark. Plate a and b show spore formation, which indicates RutC30 needs most probably day light conditions for spore formation and does not relay on specific media.

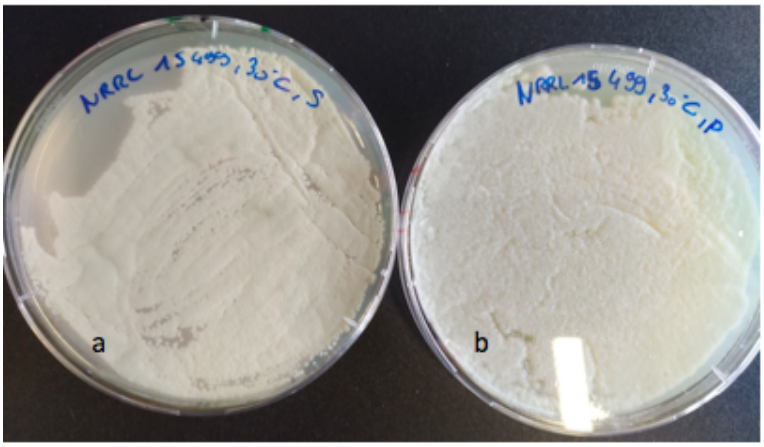


Figure 20 *T. reesei* NRRL15499 at PDA (b) and sporulation media plate (a). Plate a and b were incubated on 30°C in the dark. Plate a shows spore formation while plate b shows white mycelium, which indicates NRRL15499 needs sporulation media. Further it could be seen that spores show no green color.

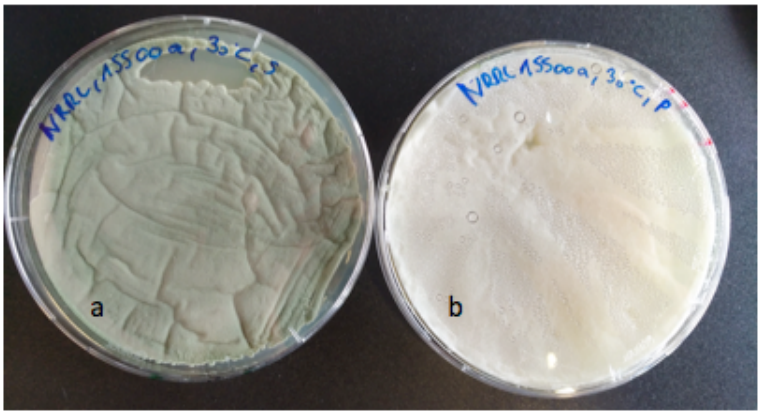


Figure 21 *T. reesei* NRRL15500a on PDA (b) and sporulation media (a). Plate a and b were incubated on 30°C in the dark. Plate a shows spore formation while plate b shows white mycelium, which indicates NRRL15500a needs sporulation media.

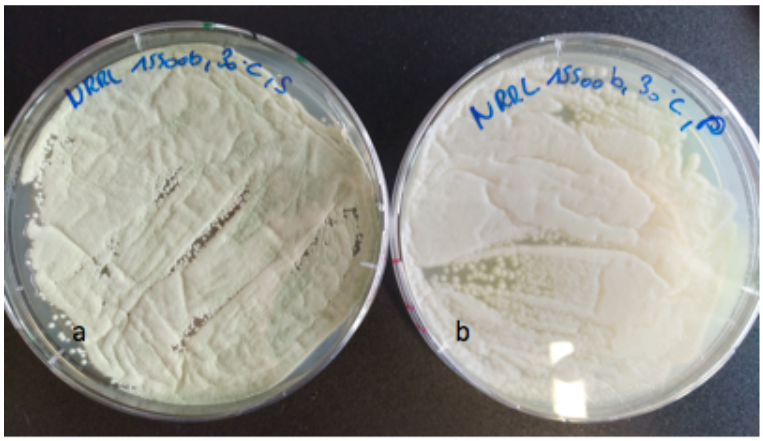


Figure 22 *T. reesei* NRRL15500b on PDA (b) and sporulation media (a). Plate a and b were incubated on 30°C in the dark. Plate a shows spore formation while plate b shows white mycelium which indicates NRRL15500b needs sporulation media.

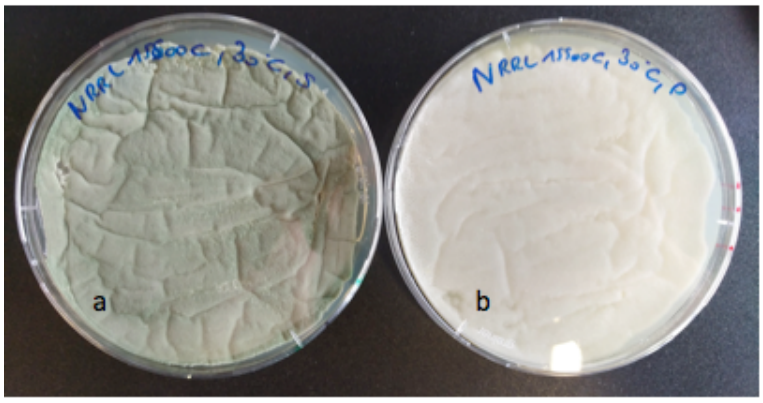


Figure 23 *T. reesei* NRRL15500c on PDA (b) and sporulation media (a). Plate a and b were incubated on 30°C in the dark. Plate a shows spore formation while plate b shows white mycelium which indicates NRRL15500c needs sporulation media.

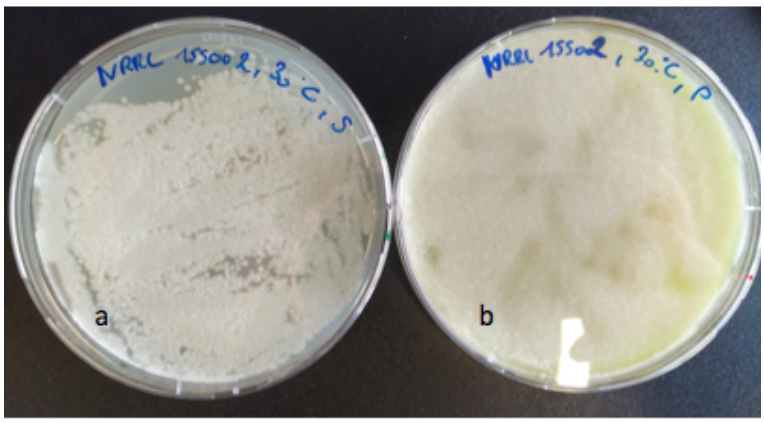


Figure 24 *T. reesei* NRRL155002 on PDA (b) and sporeulation media (a). Plate a and b were incubated on 30°C in the dark. Plate a shows spore formation while plate b shows white mycelium which indicates NRRL155002 needs sporulation media.

The optimal sporulation conditions for the different *T. reesei* strains are summarized in Table

25. As already explained, the focus of this experiment was to find conditions which could be applied for the different strains to achieve good spore formation. As it was described in literature that light settings are the most crucial factor for conidiation I focused on this. Further Anuradha Singh et al. showed that the ideal temperature for cultivating *Trichoderma reesei* is between 25°C- 30°C. [82]

Table 25 Influence of cultivation condition on the formation of conidio-spores

	PDA dark	PDA light	SM dark	SM light
QM6a	-	+++	n.a.	++
QM9414	-	+++	n.a	++
RutC30	-	+++	n.a	++
NRRL15499	++	n.a	+++	n.a.
NRRL15500a	-	n.a.	+++	n.a.
NRRL15500b	+	n.a.	++	n.a.
NRRL15500c	-	n.a.	+++	n.a.
NRRL15502	-	n.a.	+	n.a.

Therefore, although the temperature used for the cultivation was different, variation of 2°C should make no difference [82], for the comparison of the light and dark cultivation experiments. The differences in conidiation depending on growth media and light settings leads in my point of view to a clear result which goes along with the research of Schmoll et al. [81]. The strains QM6a, QM9414 and RutC30 need a light-source during incubation to form spores (Figure 17, Figure 18 and Figure 19).

The *T. reesei* strains NRRL15499, NRRL15500a, NRRL15500b, NRRL15500c and NRRL15002 show a different behaviour compared to the strains QM6a, QM9414 and RutC30. Here the influence of the cultivation media is the determining factor while light seems to be no longer required for the conidiation which could implement that the pathway for light depended conidiation is no longer intact. Spore formation worked the best when using sporulation media. (Figure 20, Figure 21, Figure 22 and Figure 23)

4.2 Cellulase induction in shake flasks – Influence of the carbon source

The main benefit of *T. reesei* is the high titer of secreted protein which can be reached in submersed cultures.

As already discussed, the main fractions of proteins secreted by *Trichoderma reesei* are cellulases and hemicellulases which are used for degrading cellulosic compounds.

In order to see, if there were differences in the amount of secreted protein depending on the primary carbon source, different sugars as sole carbon sources for growth were used. Avicel, lactose and glucose as primary carbon sources was tested. 300ml Erlenmeyer flasks containing 50ml of Mandel's media containing 1% of primary carbon source were prepared and inoculated with the different *T. reesei* strains. Each strain was cultivated in triplicates. The first sample was taken after 16 hours of cultivation and the last sample after 120 hours with intermediate time point samples taken in between. For the quantification of the secreted protein the Bradford assay was used.

The following figures 25-32, show the amount of protein secreted by the different *T. reesei* strains, when cultivated with Avicel as sole carbon source. With this carbon source, the largest amount of secreted protein for all strains was expected, since this substrate has the same structure as cellulose

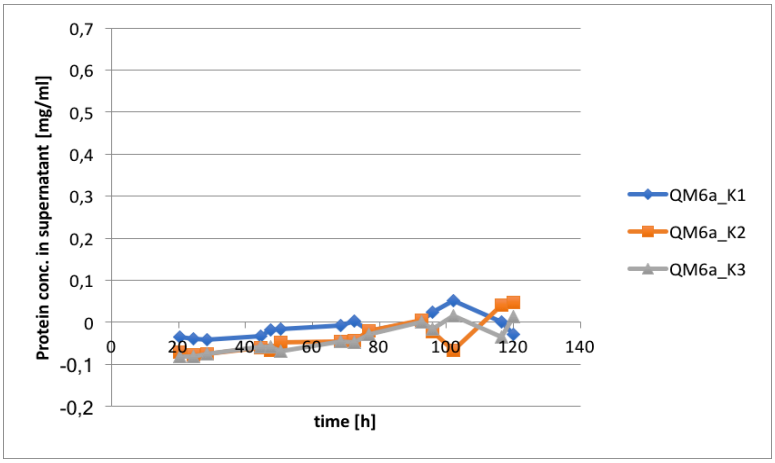


Figure 25 Concentration of secreted protein from *T. reesei* QM6a after cultivation with Avicel as primary carbon-source. Three flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. Almost no secretion was detected

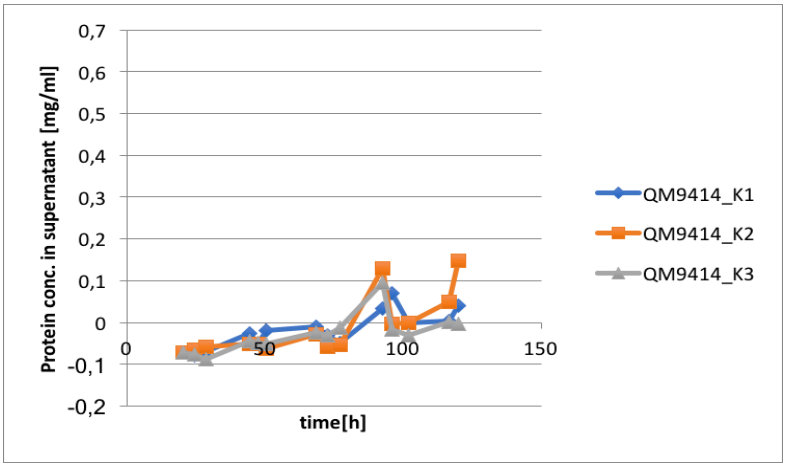


Figure 26 Concentration of secreted protein from *T. reesei* QM9414 after cultivation with Avicel as primary carbon-source. Three flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. Low amounts of secreted protein was detected.

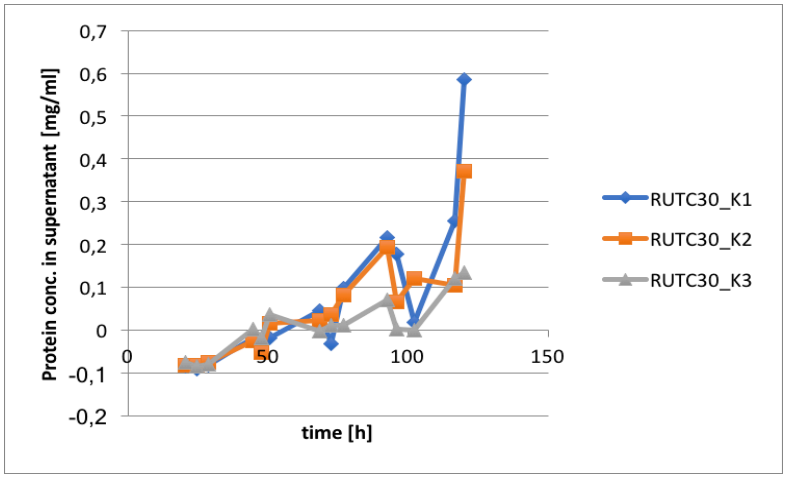


Figure 27 Concentration of secreted protein from *T. reesei* RUTC30 after cultivation with Avicel as primary carbon-source. Three flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. 0.6 mg/ml of secreted protein was detected.

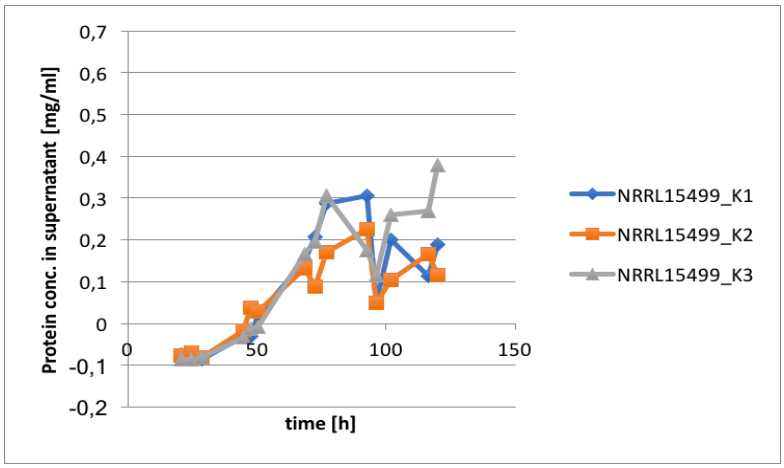


Figure 28 Concentration of secreted protein from *T. reesei* NRRL15499 after cultivation with Avicel as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours to 120hours. About 0.4 mg/ml of secreted protein was detected.

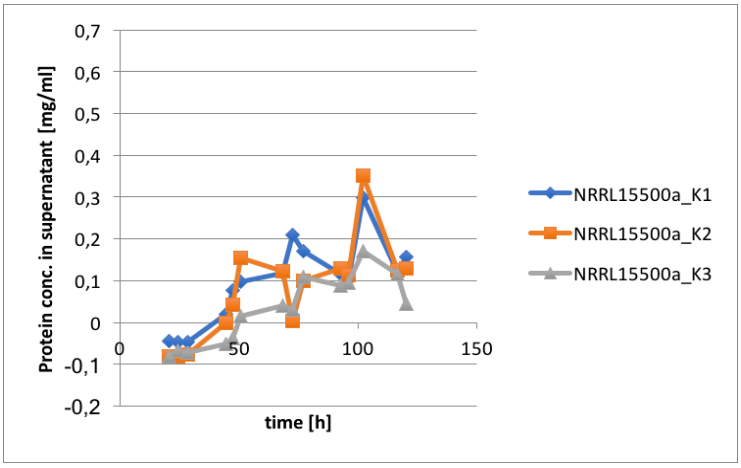


Figure 29 Concentration of secreted protein from *T. reesei* NRRL15500a after cultivation with Avicel as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. About 0.4 mg/ml of secreted protein was detected.

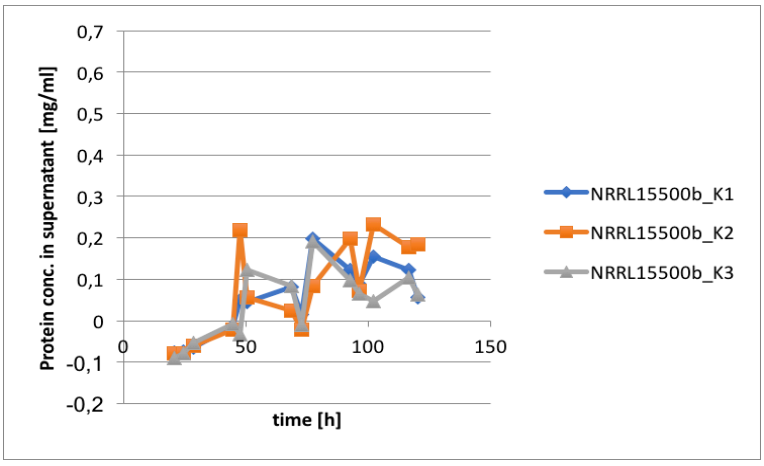


Figure 30 Concentration of secreted protein from *T. reesei* NRRL15500b after cultivation with Avicel as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. About 0.2 mg/ml of secreted protein was detected.

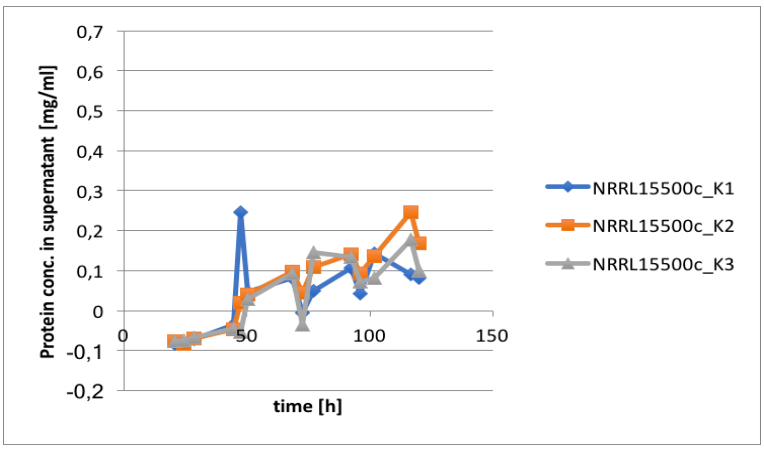


Figure 31 Concentration of secreted protein from *T. reesei* NRRL15500c after cultivation with Avicel as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. About 0.2 mg/ml of secreted protein was detected.

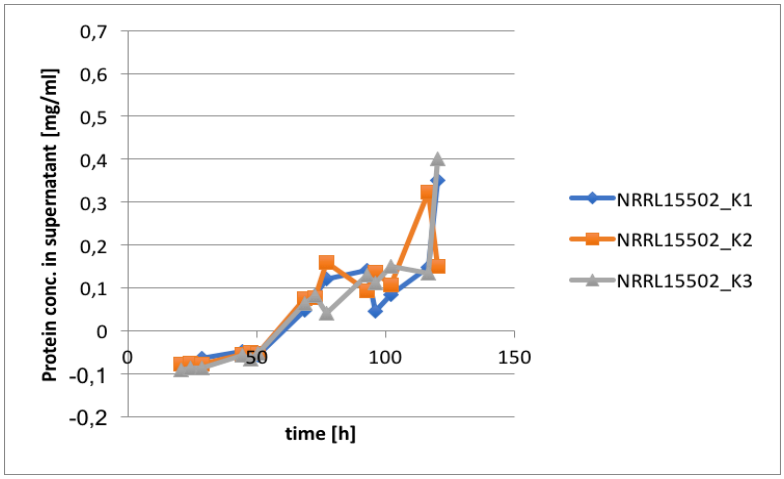


Figure 32 Concentration of secreted protein from *T. reesei* NRRL15502 after cultivation with Avicel as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. About 0.4 mg/ml of secreted protein was detected.

The following figures 33-40, show the amount of protein secreted by the different *T. reesei* strains, when cultivated with lactose as sole carbon.

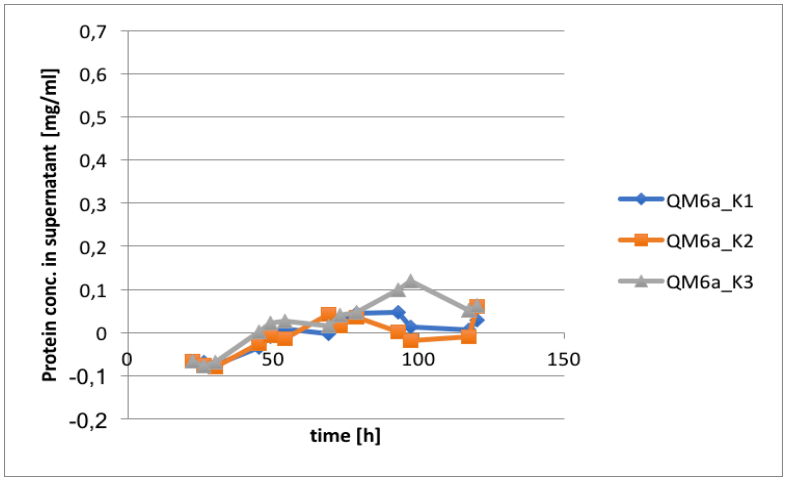


Figure 33 Concentration of secreted protein from *T. reesei* QM6a after cultivation with lactose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. About 0.1 mg/ml of secreted protein was detected.

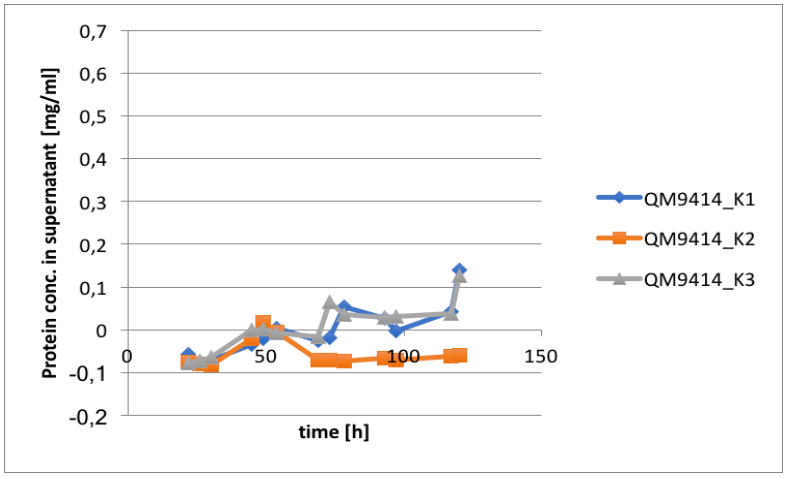


Figure 34 Concentration of secreted protein from *T. reesei* QM9414 after cultivation with lactose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. About 0.15 mg/ml of secreted protein was detected.

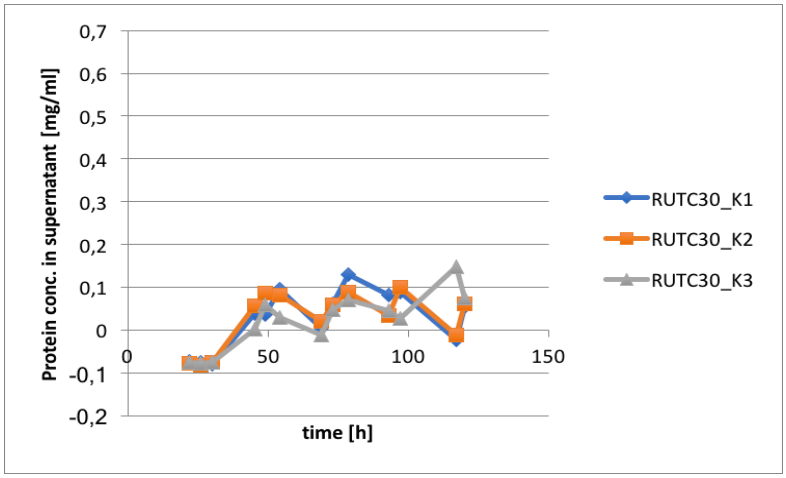


Figure 35 Concentration of secreted protein from *T. reesei* RutC30 after cultivation with lactose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. About 0.15 mg/ml of secreted protein was detected.

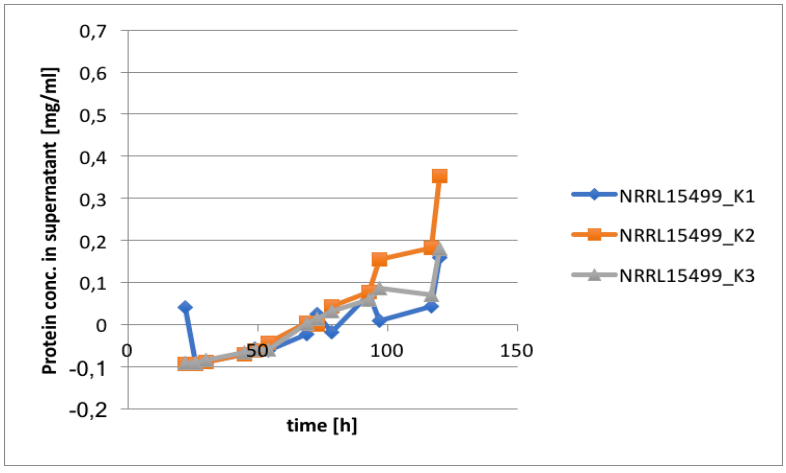


Figure 36 Concentration of secreted protein from *T. reesei* NRRL15499 after cultivation with lactose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. About 0.4 mg/ml of secreted protein was detected.

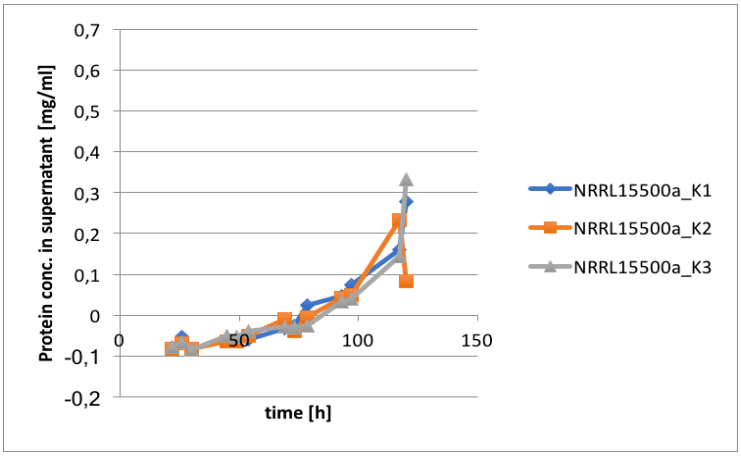


Figure 37 Concentration of secreted protein from *T. reesei* 15500a after cultivation with lactose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. About 0.4 mg/ml of secreted protein was detected.

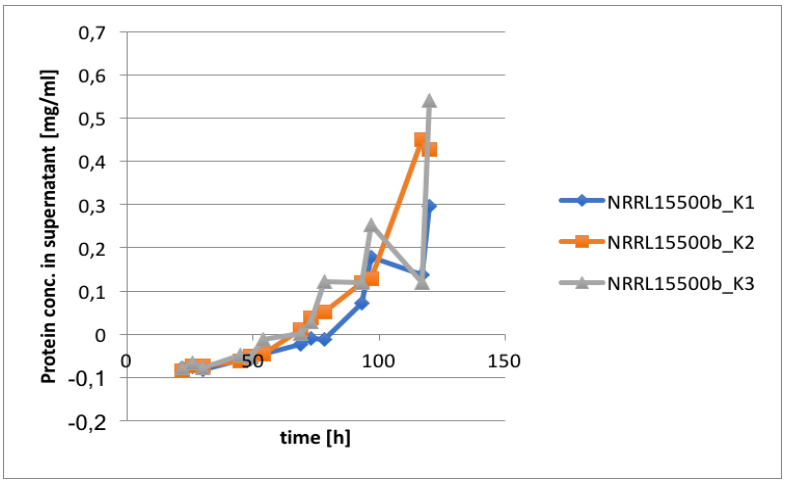


Figure 38 Concentration of secreted protein from *T. reesei* NRRL15500b after cultivation with lactose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. About 0.55 mg/ml of secreted protein was detected.

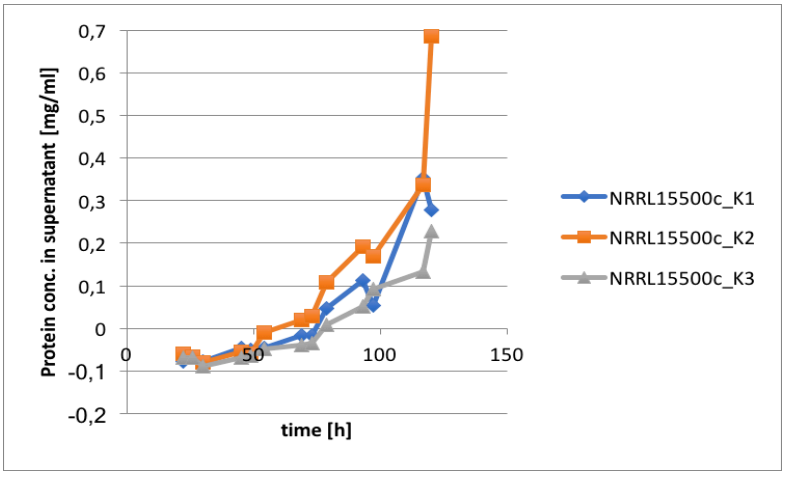


Figure 39 Concentration of secreted protein from *T. reesei* NRRL15500c after cultivation with lactose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. About 0.7 mg/ml of secreted protein was detected.

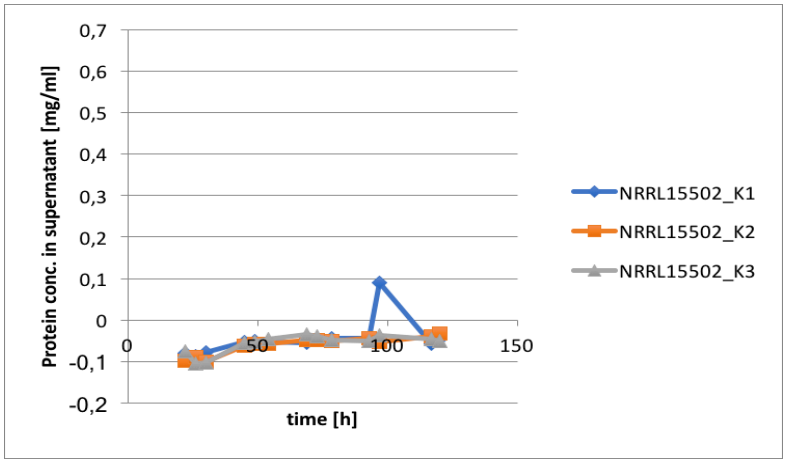


Figure 40 Concentration of secreted protein from *T. reesei* NRRL15502 after cultivation with lactose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. Almost no secreted protein was detected.

The following figures 41-48, show the amount of protein secreted by the different *T. reesei* strains, when cultivated with glucose as sole carbon source. Seeing as the main fraction of the naturally secreted proteins belong to enzymes involved in degrading of complex polysaccharides, we expected to observe bad secretion behaviour when cultivating with glucose as primary carbon source, since the enzyme cocktail needed to make these sugar molecules available is not required.

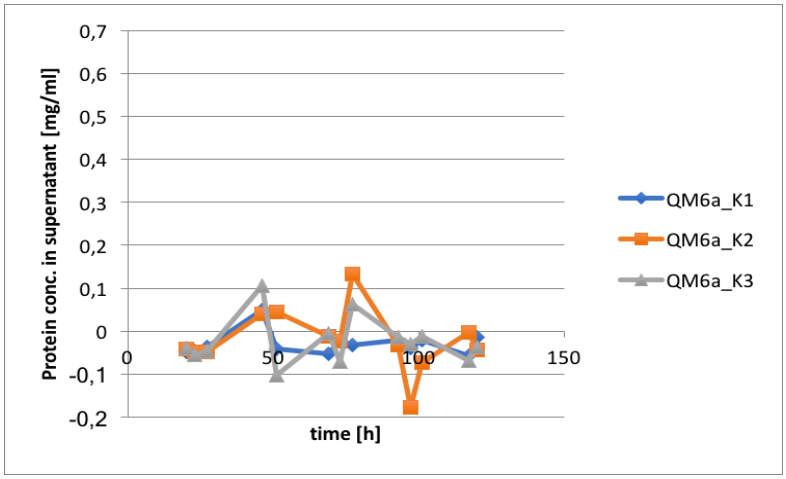


Figure 41 Concentration of secreted protein from *T. reesei* QM6a after cultivation with glucose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. Almost no secreted protein was detected.

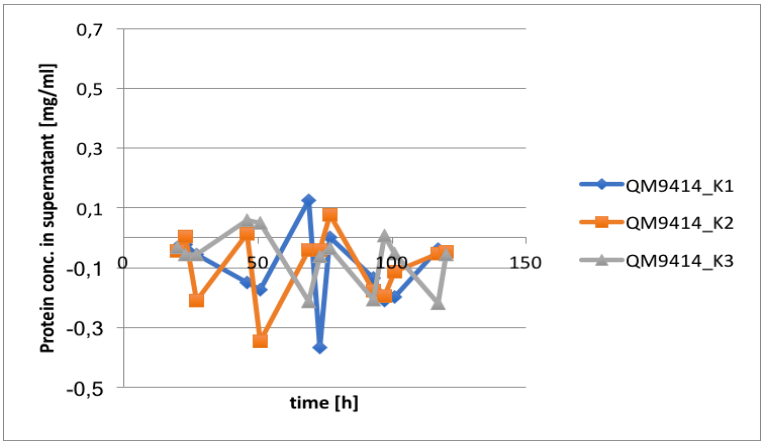


Figure 42 Concentration of secreted protein from *T. reesei* QM9414 after cultivation with glucose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. Almost no secreted protein was detected.

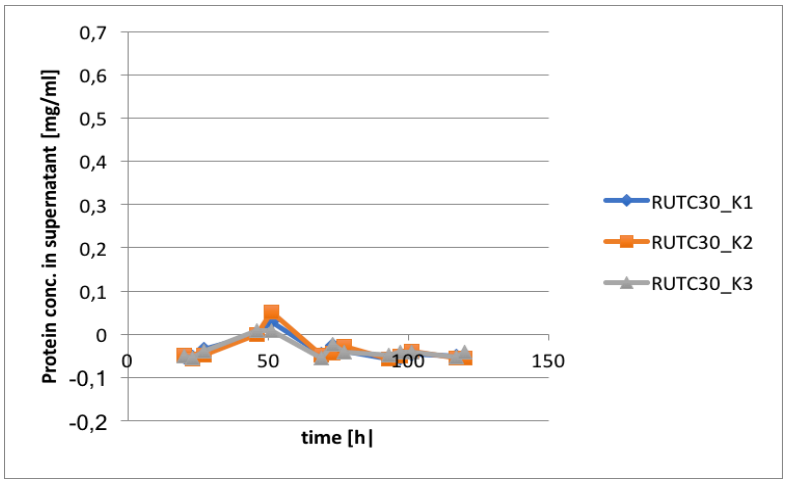


Figure 43 Concentration of secreted protein from *T. reesei* RutC30 after cultivation with glucose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. Almost no secreted protein was detected.

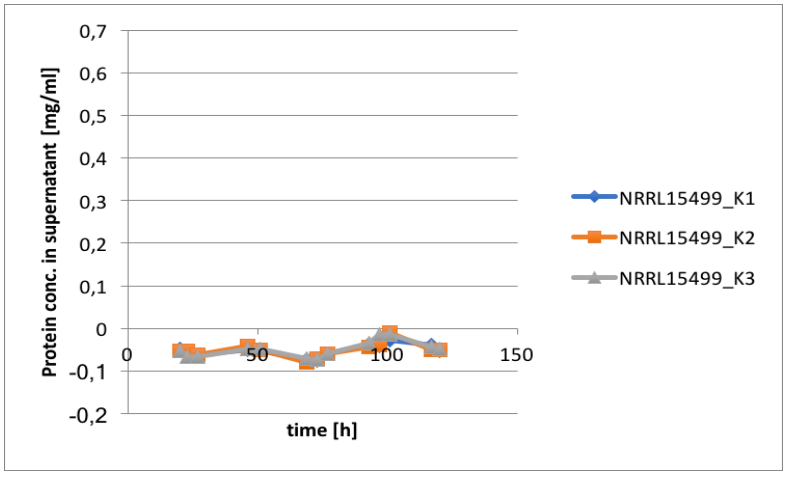


Figure 44 Concentration of secreted protein from *T. reesei* NRRL15499 after cultivation with glucose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. Almost no secreted protein was detected.

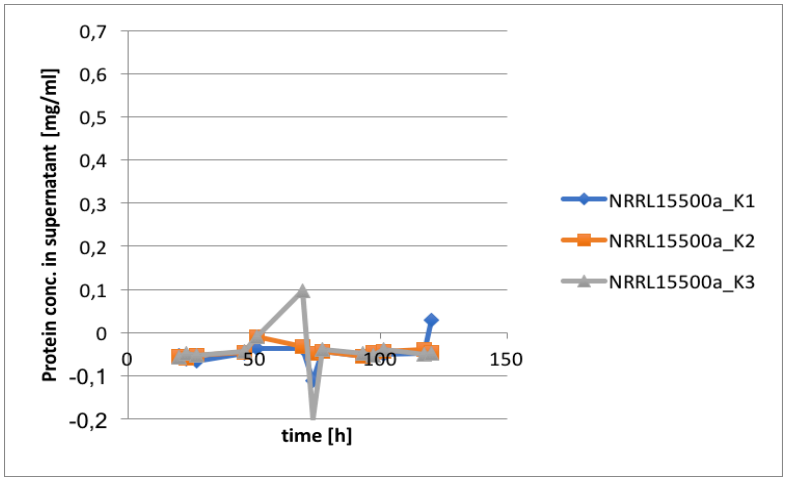


Figure 45 Concentration of secreted protein from *T. reesei* NRRL15500a after cultivation with glucose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. Almost no secreted protein was detected.

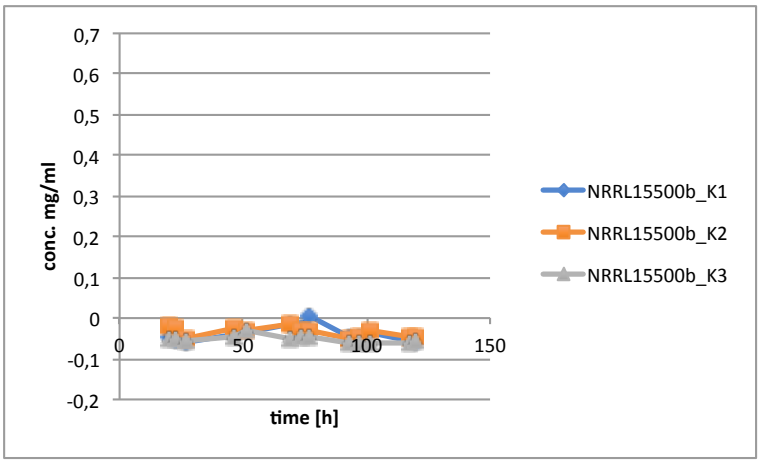


Figure 46 Concentration of secreted protein from *T. reesei* NRRL15500b after cultivation with glucose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning at 24 hours and ending at 120hours. No secreted protein was detected.

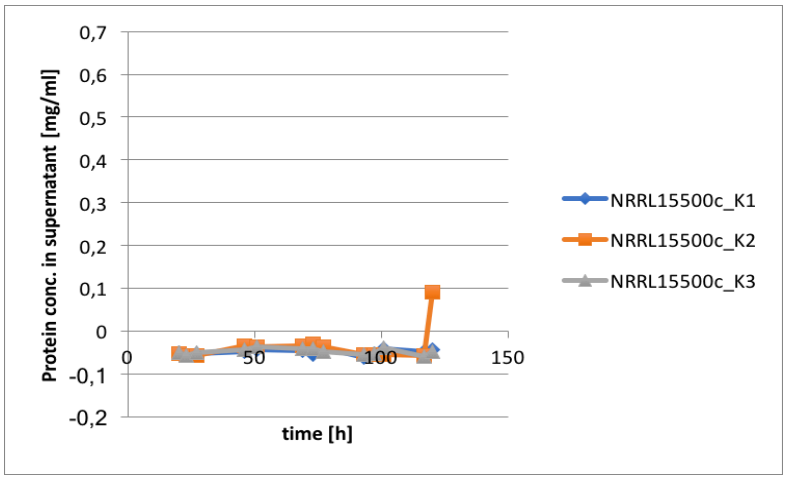


Figure 47 Concentration of secreted protein from *T. reesei* 15500c after cultivation with glucose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning and ending at 24 hours to 120hours. Almost no secreted protein was detected.

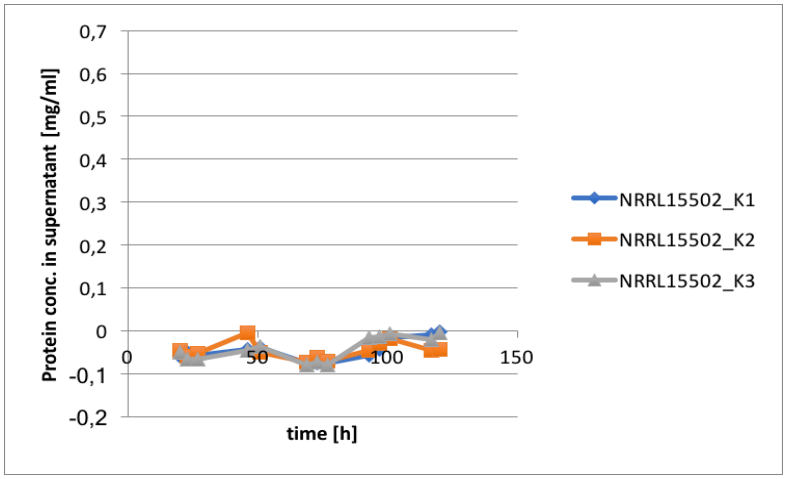


Figure 48 Concentration of secreted protein from *T. reesei* NRRL15502 after cultivation with glucose as primary carbon-source. 3 flasks were cultivated in parallel containing 50ml of media (1% c-source) and 100µl of spore stock. All points were measured in triplicates on a regular base beginning and ending at 24 hours to 120hours. Almost no secreted protein was detected.

Evaluation of the results confirmed the expectations about the secretion behaviour of the individual strains on different carbon sources. It was seen, that the samples from the cultivation with Mandel's media containing Avicel showed the highest amount of secreted protein. Other inducers for cellulose production like lactose also triggered enhanced secretion in all strains with exception of the *T. reesei* strains RutC30 (Figure 35) and NRRL15502 (Figure 40) which showed much less secreted protein.

Glucose causes, like already explained in the introduction, the repression of the cellulase transcription by the Cre1 repressor molecule. This explains why only small amounts of secreted protein could be detected, although the highest biomass formation was observed. (Figure 41, Figure 42, Figure 43, Figure 44, Figure 45, Figure 46, Figure 47 and Figure 48)

Compared the other strains *T. reesei* strains QM6a and QM9414 showed a clear difference in secretion ability upon cultivation with Avicel and lactose (Avicel: Figure 25, Figure 26; Lactose: Figure 33, Figure 34) as sole carbon source (Avicel: Figure 27, Figure 32; Lactose: Figure 35, Figure 40).

For the cultivation with Avicel as sole carbon source the starting point for protein secretion was distinctly different for the *T. reesei* strains QM6a, QM9414 and RutC30 (Figure 25, Figure 27) compared to the NRRL strains (Figure 28, Figure 32) It took the NRRL strains about 50 hours and the other strains about 80 hours of cultivation until the first significant amounts of protein could be measured by the Bradford assay.

Upon cultivation on Mandel's media containing lactose the NRRL strains needed 80 hours until detectable secretion started (Figure 36, Figure 40) compared to 50 hours for the

cultivation with Mandel's media containing Avicel

The amount of secreted protein for cultivations with Avicel as sole carbon source might be even higher than indicated by the protein quantification, because a significant fraction of the enzymes hydrolysing the large sugar molecules might be attached to the surface of the Avicel.

The results generated in the experiment using the different carbon sources go along with other publication which showed that the secretion of the enzyme mix for degrading complex polysaccharides is triggered by the presence these complex polysaccharides and repressed by the presence of glucose. [83], [84]

In a second experiment SDS-PAGE was used to see the composition of the secreted proteins of the individual culture supernatants. Only the four major cellulases, which make up 90% of the total secretom of Trichoderma can be detected, while 50-60% of the total secreted protein belong to the CBH1. [10] Therefore, it was expected to see one sharp band at 65kDA the size of the CBH1 [85]

For each strain, the samples of the time points 50 (t1) and 120 (t2) hours were loaded onto the gel. Based on the protein quantification by the Bradford assay these time points represented the beginning of the secretion and the end of the cultivation, respectively. Figure 49 shows the SDS gel with the supernatant of the cultivation with Avicel as sole carbon source. A band at 65kDa indicating the secretion of the CBH1 was detected at every time point for the cultivation with Avicel as sole carbon source.

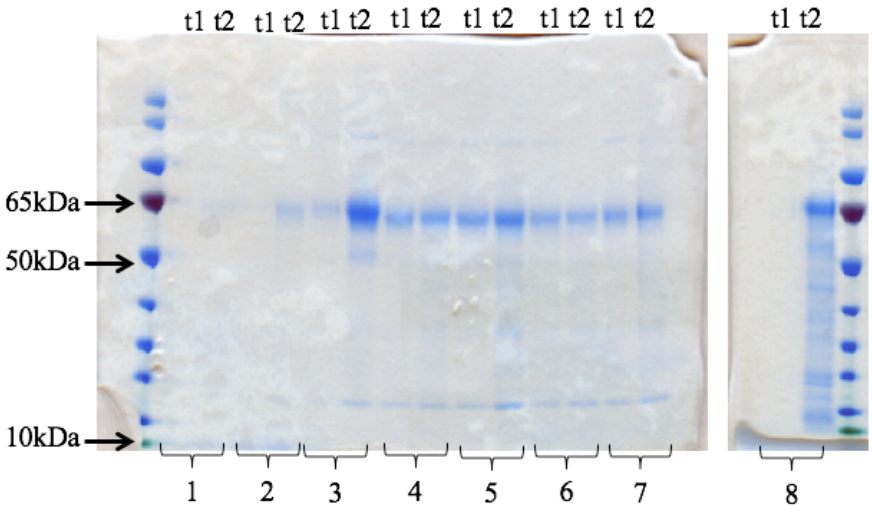


Figure 49 SDS-PAGE analyzing the culture supernatant from cultivation of different *T. reesei* strains with Avicel as primary carbon source. The strains RutC30 (Lane 3) NRRL15499 (Lane 4), NRRL15500a (Lane 5), NRRL 15500b (Lane 6), NRRL 15500c (Lane 7) and NRRL 15502 (Lane 8) showed already secreted protein after 50 hours (t1) and an increase in protein up to the final samples taken at 120 hours (t2). The most prominent band at the size of 65kDA belongs most likely to the CBH1 the strongest secreted protein of Trichoderma. The other bands, could not be assigned clearly, maybe they were just a product from protease activity.

The SDS gel with the supernatant of the cultivation with lactose as sole carbon source (Figure 50) was fully compatible with the results from the protein quantification. The characteristic band at 65kDA indicating the cellobiohydrolase 1 was also detected for lactose as sole carbon source. The difference to the cultivation with Avicel as sole carbon source was the time it took until the first amount of secreted protein could be detected. This time-shift for the secretion starting point from 50 to 80 hours was also indicated by the Bradford assay.

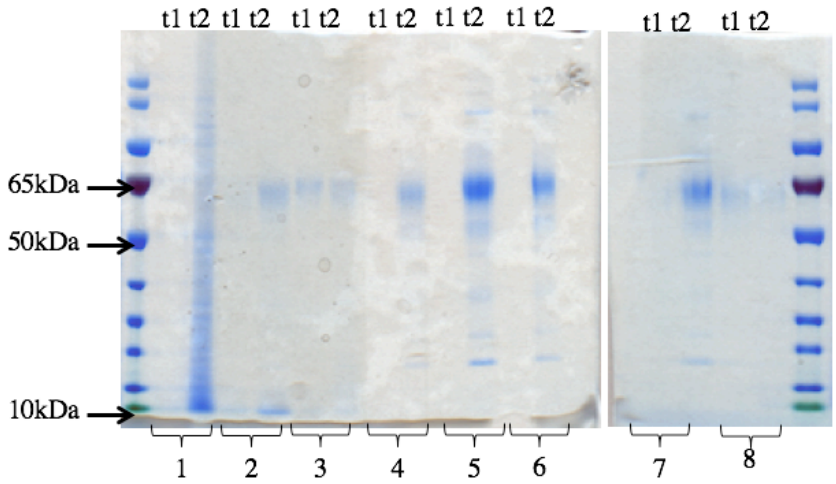


Figure 50 SDS-PAGE analyzing the culture supernatant from cultivation of different *T. reesei* strains with lactose as primary carbon source. The strains RutC30 (Lane 3) NRRL15499 (Lane 4), NRRL15500a (Lane 5), NRRL 15500b (Lane 6), NRRL 15500c (Lane 7) and NRRL 15502 (Lane 8) showed already secreted protein after 50 hours (t1) and increase in protein up to the final samples taken at 120 hours (t2). The most prominent band at the size of 65kDA belongs most likely to the CBH1 the strongest secreted protein of Trichoderma. The other bands, could not be assigned clearly, maybe they were just a product from protease activity.

Figure 51 shows the SDS gel with the supernatant of the cultivation with glucose as sole carbon source. As the results of the protein quantification already indicated, no bands could be found on the SDS-gel verifying, that no proteins were secreted, when glucose was used for cell growth. Possible reasons have already been discussed above.

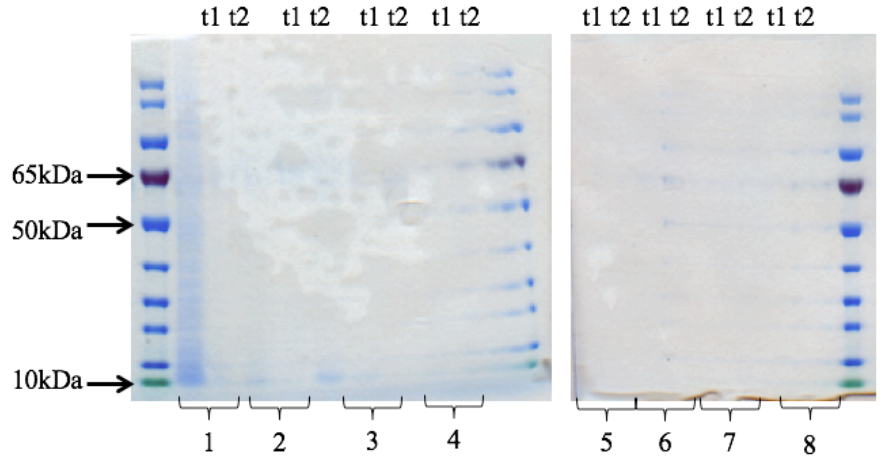


Figure 51 SDS-PAGE analyzing the culture supernatant from cultivation of different *T. reesei* strains with glucose as primary carbon source. No secreted protein was found in the different samples. (Lane 1-8) at any time point.

The results of the SDS-PAGE analysing the different culture conditions and strains confirmed the results of protein quantification by the Bradford assay.

Generally, the bands detected for the *Trichoderma reesei* strains QM6a (Lane1) and QM9414 (Lane2) were much weaker than the bands for the strains RutC30 and the NRRL strains. (Figure 49, Figure 50 and Figure 51) Further, it was proven, that secretion of the proteins started at about 50 hours, as we saw already strong bands in all lanes.

The results showed, that *T. reesei* secretes a mixture of enzymes degrading larger molecules into glucose and cellobiose. Avicel and lactose both have ß-1,4 glycosidic bonds, which had to be cleaved for using the carbon source for growth. [86]

The cultivation processes with Avicel as sole carbon source had been used to decide which strains were used for the following experiments for recombinant protein production. *Trichoderma reesei* RutC30 and NRRL1500b showed the largest amount of secreted protein with 0,6mg/ml and 0.3mg/ml, respectively.

The fact that no or low secretion of proteins occurred for the cultivation processes with glucose as sole carbon source lead to the conclusion, that new promoters for recombinant protein production should be not repressed by glucose. In comparison to the promoters, which are normally used for industrial enzyme production e.g. P_{CBH1} (repressed by glucose) a glucose-active promoter used in glucose containing media and combined with a strong secretion signal (CBH1, listed in the appendix) would lead to high secretion of the favoured protein along with almost no other proteins secreted.

4.2 Bidirectional promoters

Generally, heterologous protein production in *T. reesei* is not easily achieved and the principle of working with bidirectional promoters was already explained in the introduction part. Most frequently, when working with *T. reesei*, P_{CBH1} is used as promoter, as it is the strongest known promoter in this organism. Using other promoters, for example P_{CBHII}, has been described and delivers almost the same results [62], but working with a bidirectional expression system has not been published so far.

The tool to find bidirectional promoter sequences described above was applied using the genome data of *Trichoderma reesei* QM6a published by Martinez, D. et al. [76] and transcriptome data from Kubicek et al. [77] The results depicted in Figure 52 showed many possible bidirectional promoter sequences of *T. reesei*, which still needed to be validated by experiments to be sure, that these sequences really initiate transcription.

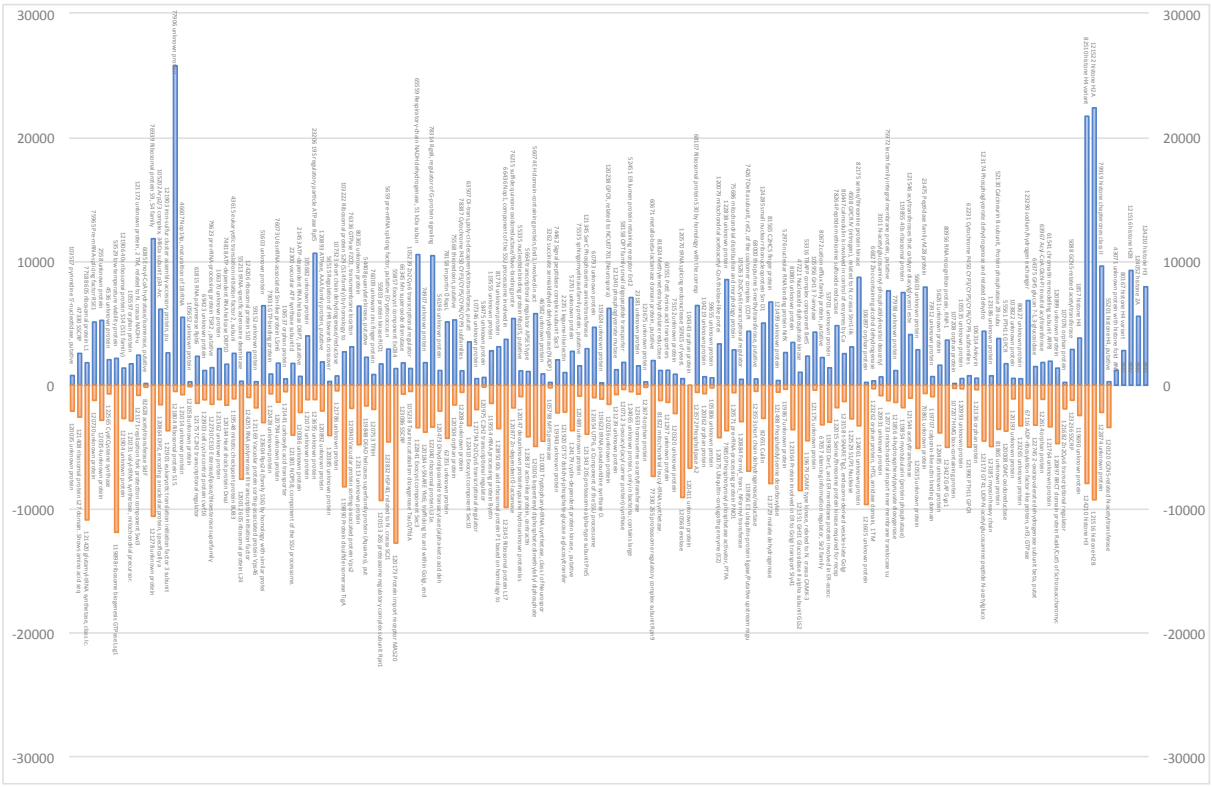


Figure 52 Overview of bidirectional promoter sequences from *T. reesei*. The blue and orange bar show the transcription strength to both directions of the promoter. Sequences are ordered with increasing size form left to right.

The different coloured bars in Figure 52 represent the transcription levels in both directions of the hypothetical promoter sequence. The height of the bars in both directions helped to decide which two natural promoters should be used in the final plasmid.

As a new vector system was designed from which nobody knew if it works or not the first

bidirectional promoters were picked in because of their transcription strength according to the developed tool.

Knowing or at least having an idea about the transcription strength of the promoters is in my point of view an essential part for testing a new expression system of has not been validated before. On the other hand, mRNA stability largely influences the outcome of current transcriptomics experiments and using an isolated promoter in a new context might significantly change the view on the strength of industrial promoters.

The first tested bidirectional promoter sequence regulates transcription of a ribosomal protein belonging to the S9, S4 family and an unknown protein on chromosome III. The promoter had a length of 427 base pairs and transcriptomics data indicated strong transcription on both sides.

The second promoter tested was a bidirectional histone promoter regulating transcription of the histones H2A and H2B in the organism. This promoter was already described in the work of Teuschler [65] which was another reason this promoter was selected for testing. This sequence is located on chromosome II and has 2100 base pairs with the strongest transcription levels of all histone promoters according to our tool.

4.3 Vector backbone for bidirectional promoter testing

The plasmid for recombinant protein production by integration into the cbh1 locus for *Trichoderma reesei* (Figure 53) was assembled by using the method described in the material and methods part.

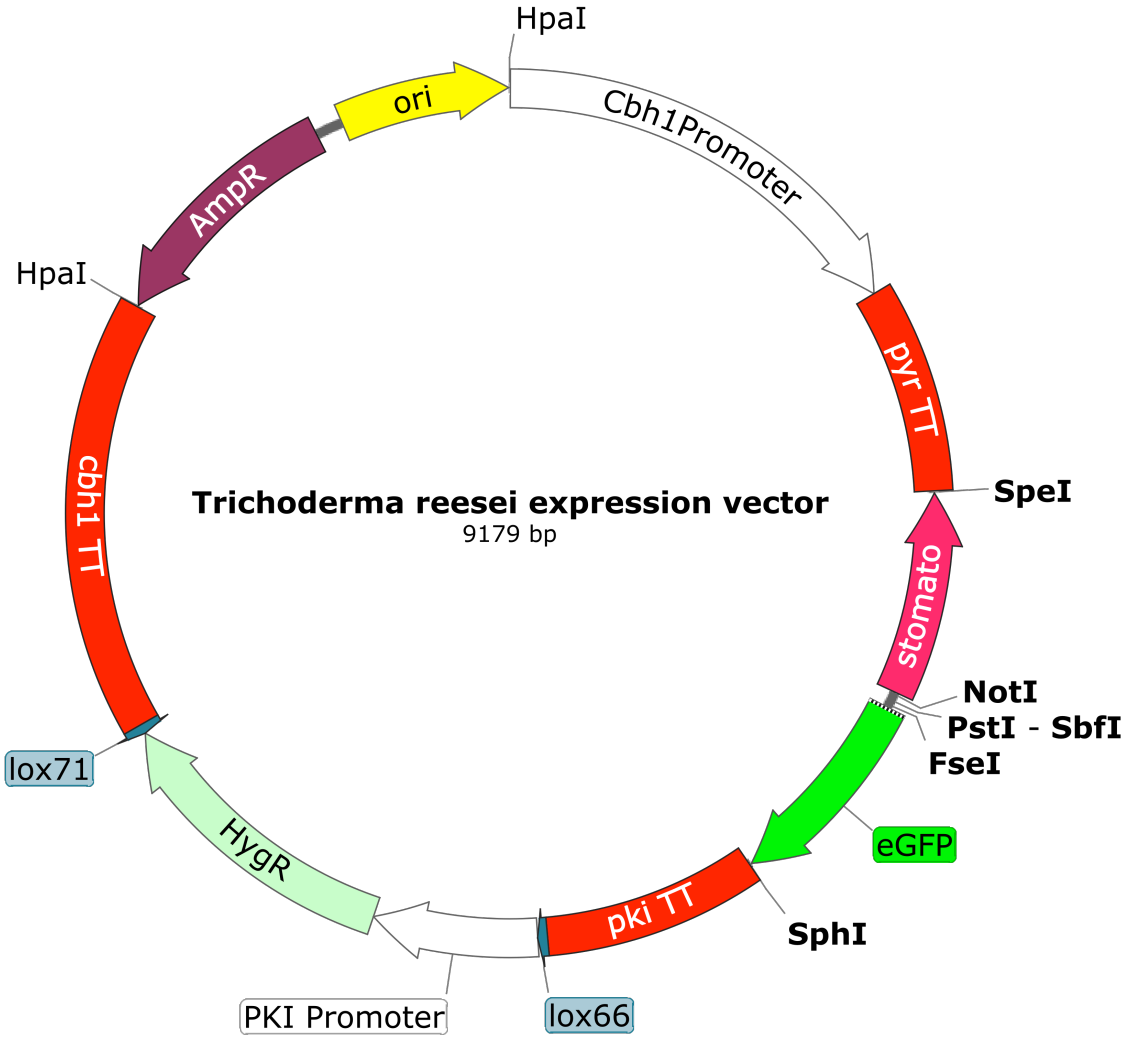


Figure 53 The new basic vector without promoter sequences for recombinant protein production in *T. reesei.* The parts for autonomous replication and selection in *E. coli* are highlighted in violet (selection marker) and yellow (ORI). Terminator regions are highlighted in red and promoter regions in white. Hygromycin B resistance gene (green) served as selection marker. For integration to the genome of *T. reesei* the construct was flanked by the CBH1 Promoter and Terminator sequence flanking the construct. For Marker recycling a Lox-sites were implemented on the vector.

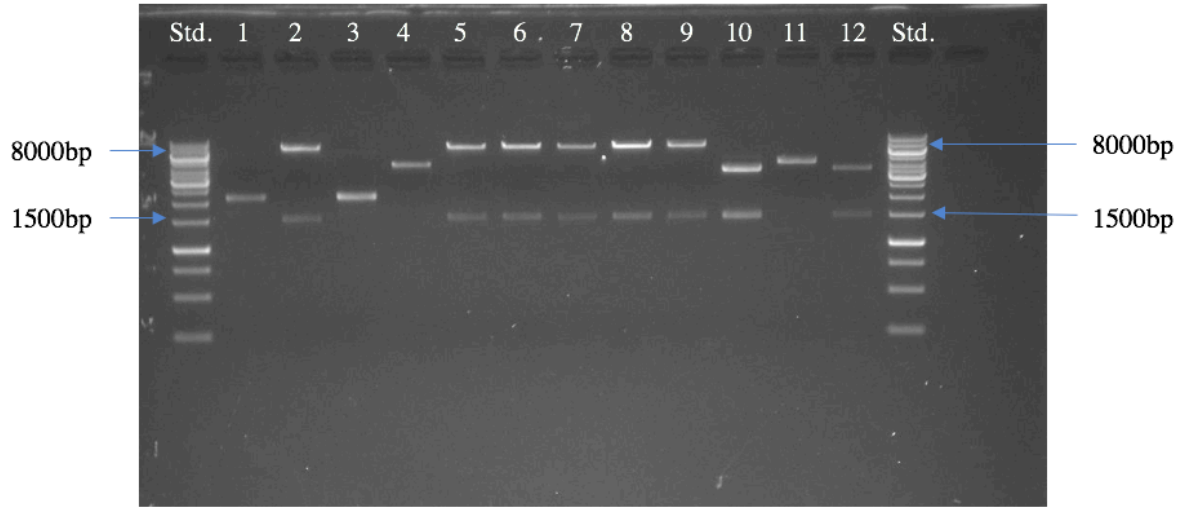


Figure 54 Restriction analysis of the backbone for later protein production in *T. reesei*. GeneRuler 1KB from Thermo Fischer scientific was used as reference in the lanes marked with Std. Lanes 1-12 show the linearized plasmids. Plasmids 5-9 show the correct sized fragments with the sizes 1.5kb and 8.0kb

The DNA fragments in lane 5, 6, 7, 8, 9 showed the correct size. The plasmids in lane 6 and 9 were sent for sequencing to confirm the DNA sequence. The results of the sequencing showed that the plasmid had two mutations at the positions G3339A and G3365T. Both mutations are within the coding sequence of the *eGFP*.

The mutation G3339A on the plasmids is synonymic to the mutation G351A on the coding sequence of the *eGFP*. The mutation causes a change from aspartic acid to asparagine within the amino acid sequence. The mutation G3365T on the plasmid is synonymic to the mutation on position G377T within the CDS of *eGFP* but as this mutation is silent no change in the amino acid sequence is introduced.

In addition, clone 8 was sent for sequencing but also this clone showed the mutations mentioned before. The research of Li et al. [87] describes the important amino acids for the functionality of the eGFP. The mutation G351A in the DNA sequence or D116N on amino acid level should not to make any difference in the functionality of the protein which lead to the decision to use clone 9 for all following experiments.

4.4 Cloning of BIDI promoters

The three variants of the $P_{Ribosome}$ and $P_{CBH1/PDC}$ were amplified for following promoter characterization experiments and after confirming the correct sequence cloned into the basic backbone.

Amplification worked for all selected promoters except the histone promoter. Since,

amplification of the $P_{Histone}$ did not work with three different primer pairs another primer pair binding to the genes *h2a* and *h2b* flanking the desired promoter was tried. A fragment of approximately 2500 base pairs was expected and as Figure 55, (lane 3 and 4) proofed this amplification worked.

This lead to the assumption that the DNA sequence for the histone promoter was not correct, which would explain the problems that occurred during amplification. To be sure the fragment was correctly annotated in the genome the sequence was verified by sequencing.

Although the sequencing results showed no deviation to the map which was used for the primer design the amplification of the histone promoter did not work (Figure 56), neither from the genome (Lane 1,3 and 5) nor from plasmid with the sequence from *h2a* and *h2b* cloned into a pJET1.2 blunt backbone (Lane 2, 4 and 6). In addition, also different temperature profiles were tested without success. Profile 1: Lane 1 and 2; Profile 2: Lane 3 and 4; Profile 3: Lane 5 and 6. The different profiles are explained in the material and methods part.

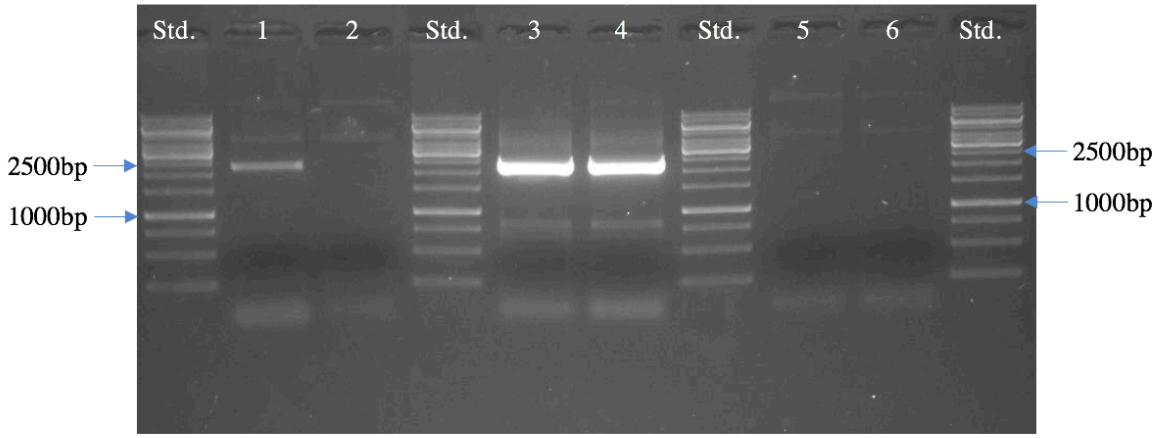


Figure 55 PCR to amplify the fragment from *h2a* to *h2b*. GeneRuler 1KB from Thermo Fischer scientific was used as reference in the lanes marked with Std. For reactions in lane 1 and 2 we used PCR profile 1, for lane 3 and 4 PCR profile 2 and for lane 5 and 6 PCR profile 3. The expected fragment should have about 2.0kb which can be seen in lane 3 and 4.

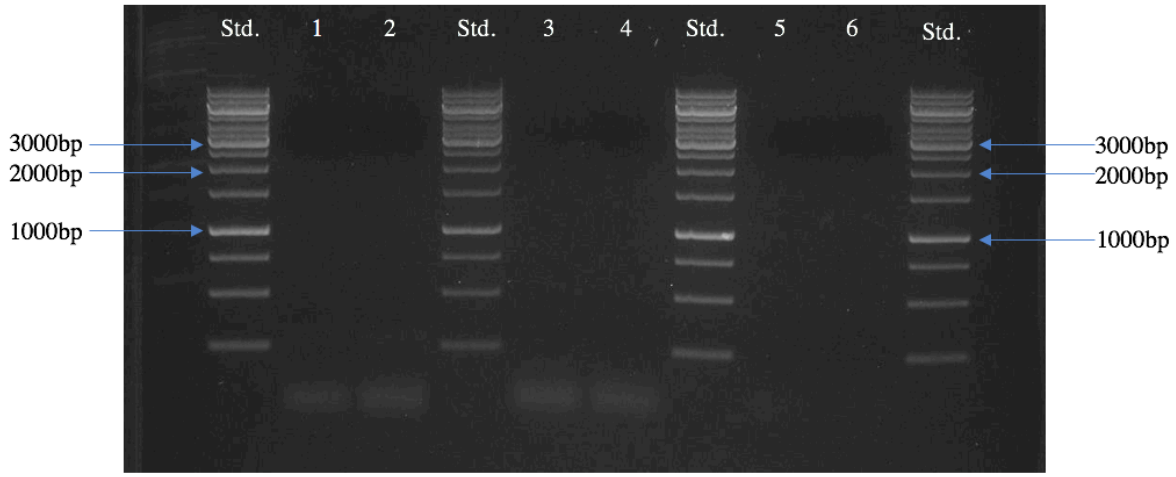


Figure 56 PCR for amplification of the bidirectional histone promoter from the genome of *T. reesei* and a pJET plasmid. PCR neither worked with the genomic DNA as template (Lane1, 3, and 5) nor with the h2a-h2b fragment in a pJET1.2 plasmid (Lane 2,5 and 7). For reactions in lane 1 and 2 we used PCR profile 1, for lane 3 and 4 PCR profile 2 and for lane 5 and 6 PCR profile 3. GeneRuler 1KB from Thermo Fischer scientific was used as reference in the lanes marked with Std

In a last attempt, I also tried to order $P_{Histone}$ as synthetic fragment from IDT but because of the AT-rich sequence and the poly A repetitions its synthesis was not possible.

In the work of Teuschler [65] one further native histone promoter from *T. reesei* was described, controlling the transcription of the histones H3 and H4. However, compared to the other promoters chosen, this bidirectional promoter shows low transcription rates to both sides according to our tool. The other bidirectional promoters mentioned in the work of Teuschler [65] were from *Saccharomyces cerevisiae* and *Myceliophthora thermophila.* Since using non-native promoters with a newly developed expression vector in combination with such a complex and demanding expression host as *T. reesei* results in too many variables that could negatively influence the expression. I in accordance with my supervisor decided to focus my work on the other already aforementioned promoters.

Both promoters where amplification worked $P_{Ribosome}$ and $P_{CBH1/PDC}$ were combined with all three different Kozak-Sequence variants, which were introduced by primers during the amplification. All combinations were successfully cloned into the basic backbone by homologous recombination as well as restriction cloning, to compare the efficiency of both strategies.

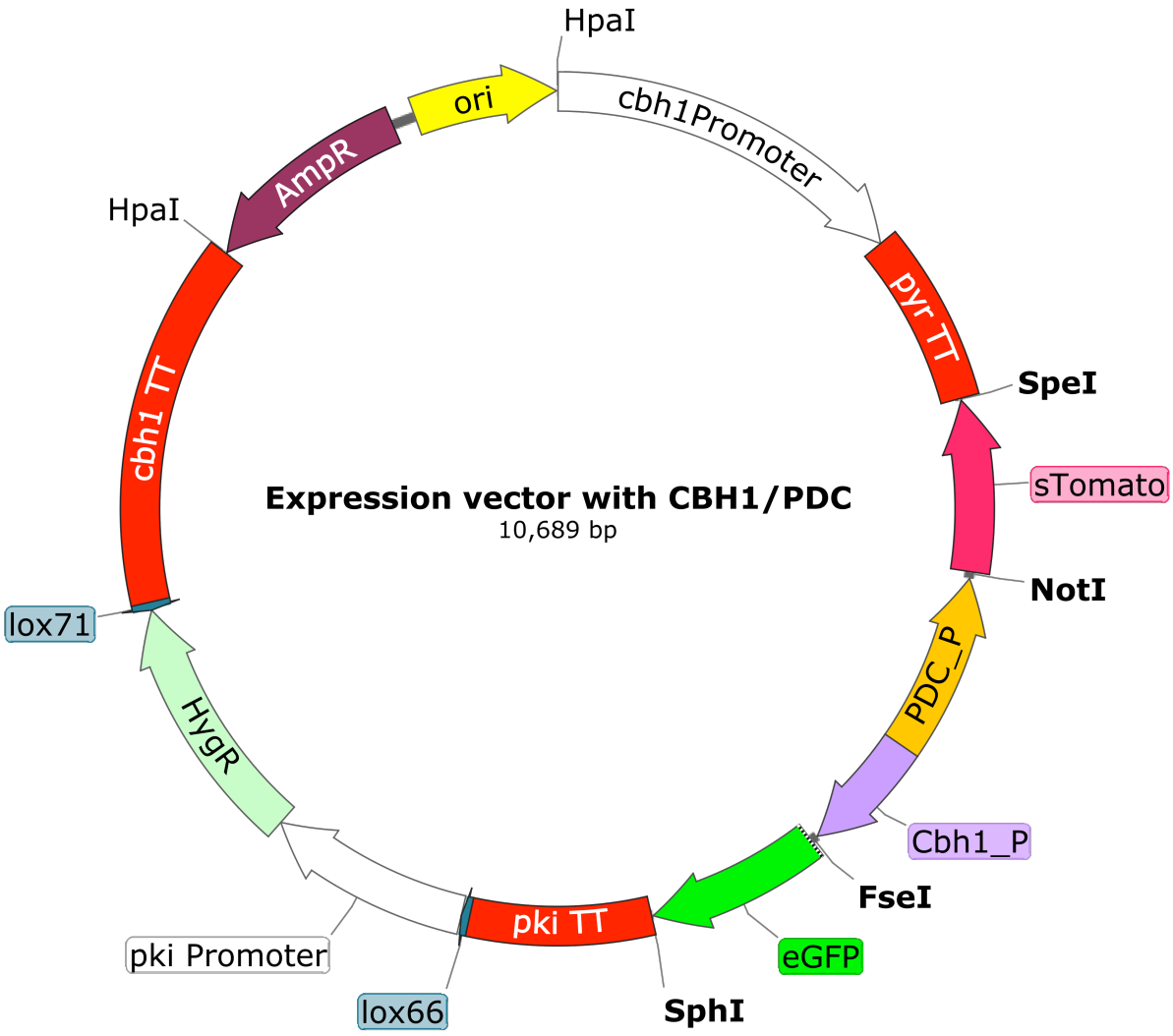


Figure 57 Expression vector for testing the artificial bidirectional promoter $P_{CBH1/PDC}$ (purple/orange) for co expression of eGFP and sTomato in *T. reesei*. The vector is based on the basic vector for recombinant protein production and was assembled in three different Kozak-Sequence versions.

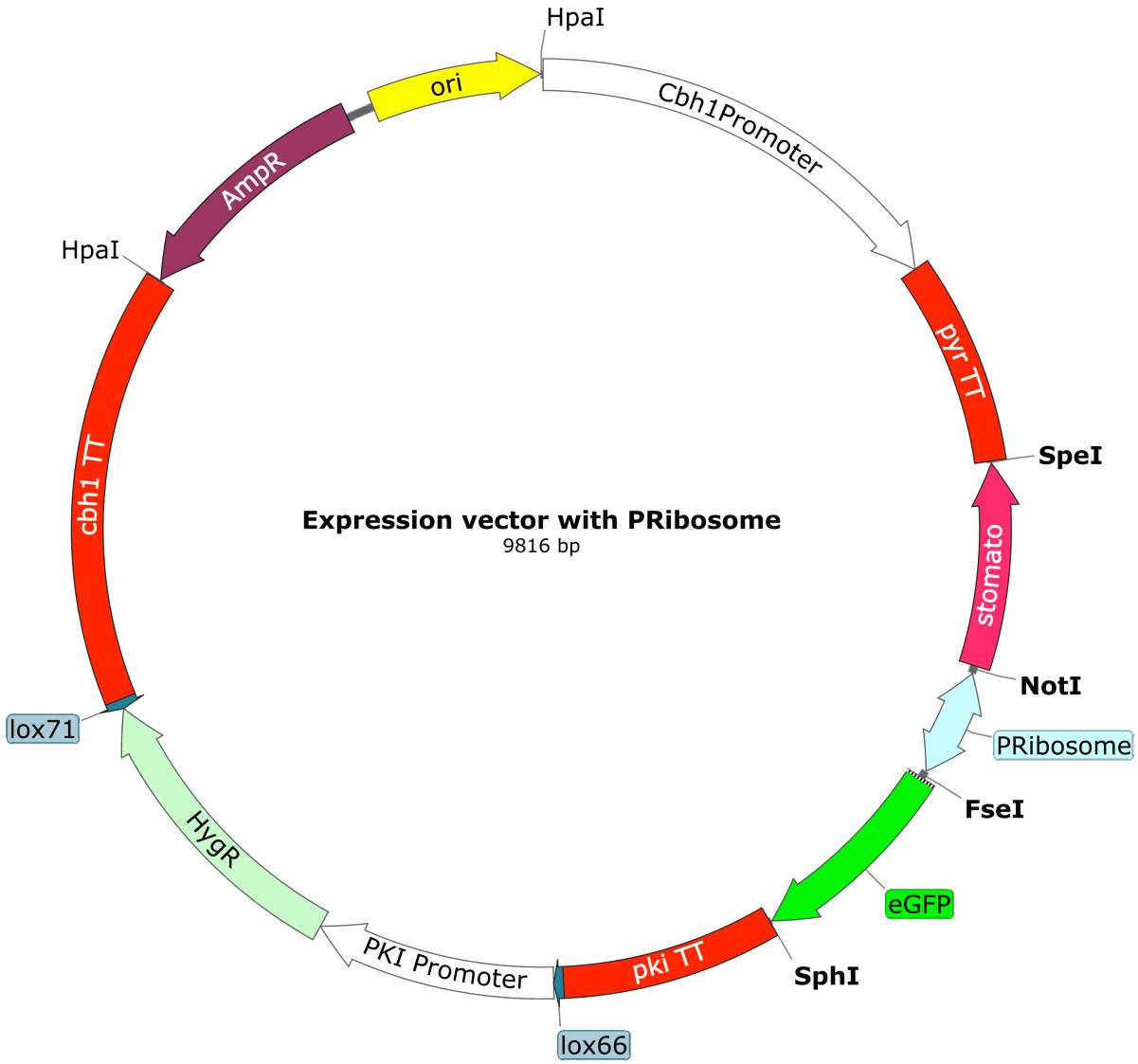


Figure 58 Expression vector for testing the native bidirectional promoter $P_{Ribosome}$ (blue) for co expression of eGFP and sTomato in *T. reesei*. The vector is based on the basic vector for recombinant protein production and was assembled in three different Kozak-Sequence versions.

The result of the cloning procedure were six plasmids. Three variants of the $P_{Ribosome}$ (Rn, Rok, Rwl) and three variants of the $P_{CBH1/PDC.}$ (Pn, Pok, Pwl)

The "n" variants had the promoter sequence which occurred naturally on the bidirectional promoter and in addition the *CBH1* Kozak-Sequence, the "ok" variants had the *CBH1* Kozak-Sequence and the "wl" variants had only Kozak-Sequence which occurs naturally on the bidirectional promoter. The different variants are explained more precisely in the material and methods part.

The clones were checked by restriction analysis using the enzymes *Spe*I and *Sph*I. The

plasmids with $P_{Ribosome}$ showed two bands at the size of about 8000 base pairs and another band at 1900 base pairs. Restriction analysis for the plasmids with the $P_{CBH1/PDC}$ showed one band at the size of 8000 base pairs and one at the size of 2700 base as expected.

If restriction analysis was successful the plasmid DNA was sent for DNA sequencing.

During cloning procedure, I realized that the cloning efficiency is not as high as I was used to for homologous recombination cloning or restriction cloning in other projects. The most probable reason for this is that the size of the construct is already demanding for the cells to handle and causes a lot of metabolic pressure. This is also indicated by the slow growth of the *E. coli* cells after transformation.

4.5 *Trichoderma reesei* transformations

One of the main objectives of this master thesis was to find a good protocol for the transformation of *Trichoderma reesei.*

Transformation was only done with the *Trichoderma reesei* strains RutC30 and NRRL15500b since these strains were the most efficient for secretion of cellulose degrading enzymes when cultivated with Avicel and lactose as carbon source. In both cases protoplast formation worked fairly well resulting in about 1x10^9 protoplasts per ml. Protoplasts were diluted in order to achieve a concentration of 2x10^7 as required by the protocol.

For both strains a transformation efficiency of about 45 mfu/µg DNA was achieved in 3 independent rounds of transformation. This transformation rate was more than 10 times better as in the previous work of Teuschler. [65] About 50% of the transformants showed fluorescence activity for the other 50% most likely only the selection cassette was integrated. (Table 26) The positive clones also showed a slight orange colour most probably due to the overexpressed sTomato.

A statement of the ratio of eGFP and sTomato could not be made as the expression of the fluorescence proteins was different within the same mycelium which is also shown by the pictures form fluorescence microscopy. Reasons for this different expression levels within the cells of the mycelium is also described later on.

Table 26 Number of colonies after transformation and the number of colonies which showed fluorescence by florescence microscopy

Strain	Transformation	Colonies	Colonies showed fluorescence
	1	39	15
T. reesei RutC30	2	46	23
	3	53	28
	1	42	21
T. reesei NRRL15500b	2	41	19
	3	57	24

Most probable reason for this improvement is most likely the mutation free vector especially the mutation free selection cassette. As detected during the last sequencing of the selection cassette of the vector of Teuschler a fragment of unknown sequence between the promoter and the coding sequence of the *hygromycin B* was detected.

The other possibility was that the protocol used in the work of Teuschler is simply not that good although he also worked with a PEG-mediated protoplast transformation protocol.

A simple experiment to detect the problem would be to use both plasmids and both protocols and perform them parallel. If the *Trichoderma reesei* transformation with Teuschler's and the plasmid assembled in this thesis works equally good/bad for the protocols the transformation protocol is the reason for the differences in the transformation efficiency. On the other hand, if the same protocol leads to a different transformation efficiency the plasmid is most probably the deciding reason.

4.5 New *Trichoderma reesei* expression system

As already described in detail, the newly developed expression vectors were designed for homologous integration into the *CBH1* locus.

To test the functionality of the newly constructed expression vectors for recombinant protein expression, the *Trichoderma reesei* strains RutC30 and NRRL1550b were used as hosts in combinations with the two fluorescence proteins eGFP and sTomato (both codon optimized) cloned into the basic expression vector.

The fastest method to proof if the system works was to use fluorescence microscopy. The

Trichoderma strains were transformed with the six different promoter plasmids as explained above.

For fluorescence microscopy transformants were cultivated on PDA agar plates for five days prior to sampling the mycelium for further analysis

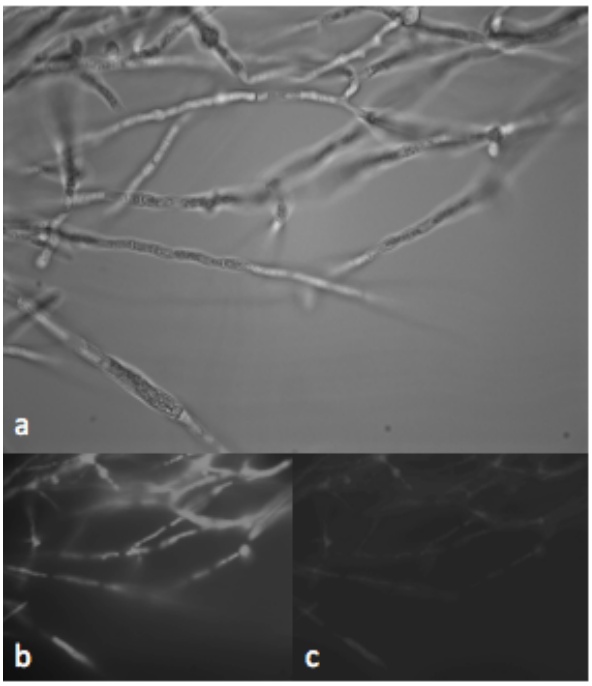


Figure 59 Negative control of non-transformed *T. reesei* RutC30 using different filters: a… Bright field, b…355-425nm, c…515-560nm. The mycelium was cultivated for 5 days on PDA plates at 28°C.

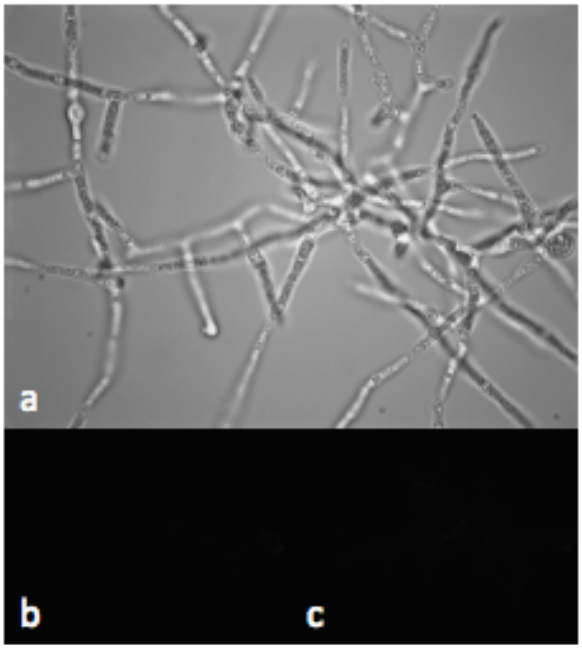


Figure 60 Negative control of non-transformed *T. reesei* NRRL1550b using different filters: a… Bright field, b…355-425nm, c…515-560nm. The mycelium was cultivated for 5 days on PDA plates at 28°C.

The non-transformed RutC30 strain (Figure 59) showed some autofluorescence which had to be considered for later microscopy pictures.

Figure 61 showed *T. reesei* RutC30 with transformed construct for the expression of eGFP and sTomato under control of $P_{CBH1/PDC}$ while Figure 62 shows the identically transformed construct with $P_{Ribosome}$. The promoter had the native promoter Kozak-Sequence and additionally the newly introduced *CBH1* Kozak-Sequence.

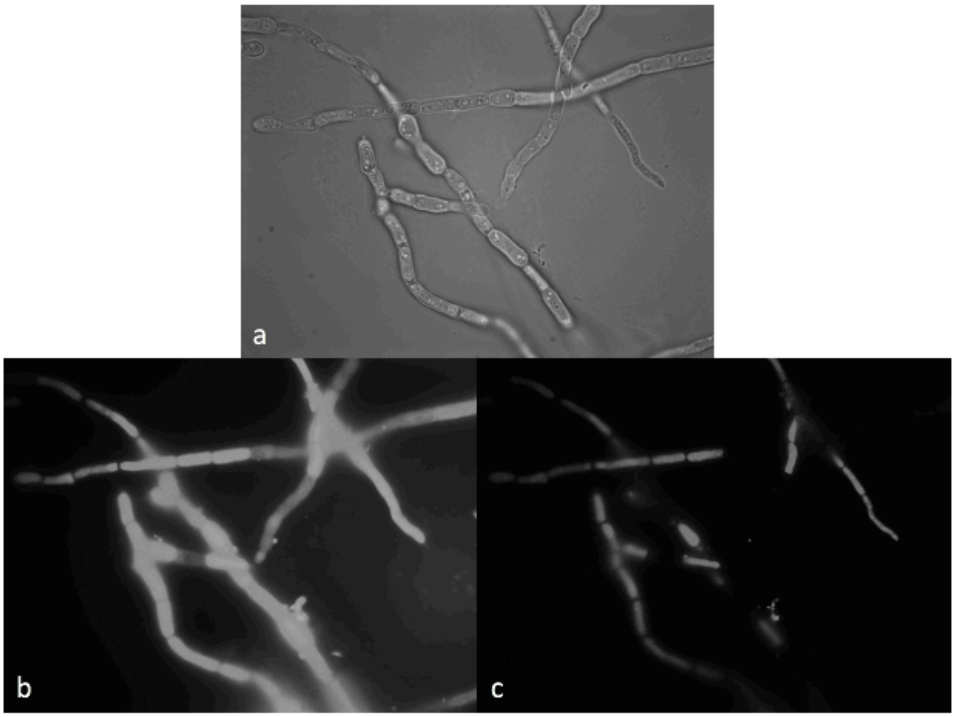


Figure 61 Pictures of *T. reesei* RutC30 with integrated construct for expression of eGFP and sTomato controlled by $P_{CBH1/PDC}$. The promoter had the native promoter Kozak-Sequence and additionally the newly introduced *CBH1* Kozak-Sequence. The mycelium was cultivated for 5 days on PDA plates on 28°C. (a) bright field; (b) using filter for 355-425nm, for stimulating green fluorescence; (c) using filter for 515-560nm, stimulating red fluorescence. Strong fluorescence from eGFP and sTomato was detected indicating that the vector in combination with the strain as well as the promoter was functionally.

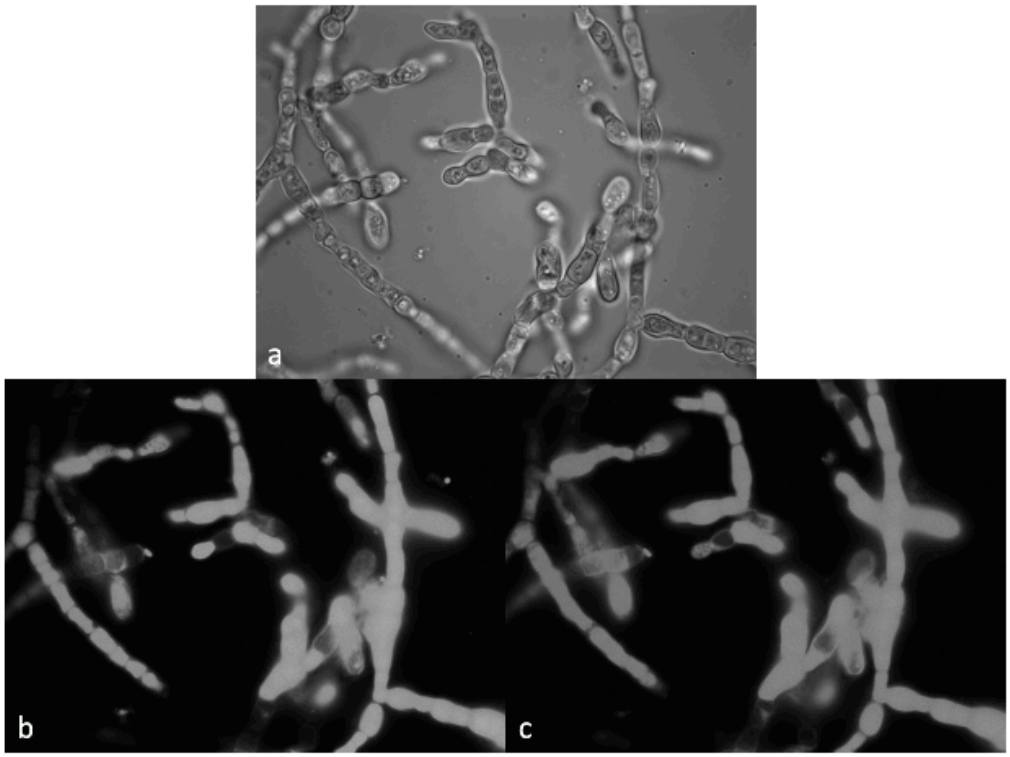


Figure 62 Pictures of *T. reesei* RutC30 with integrated construct for expression of eGFP and sTomato controlled by $P_{Ribosome}$. The promoter had the native promoter Kozak-Sequence and additionally the newly introduced *CBH1* Kozak-Sequence. The mycelium was cultivated for 5 days on PDA plates on 28°C (a)bright field; (b) using filter for 355-425nm, for stimulating green fluorescence; (c) using filter for 515-560nm, stimulating red fluorescence. Strong fluorescence from eGFP and sTomato was detected indicating the vector in combination with the strain as well as the promoter was functionally.

Figure 63 showed transformed *T. reesei* NRRL15500b with the construct for the expression of *eGFP* and *sTomato* under control of $P_{CBH1/PDC}$. In Figure 64 the pictures of NRRL1550b which was transformed with the construct for expression of eGFP and sTomato under control of $P_{Ribosome}$ was shown. The promoter had the native promoter Kozak-Sequence and additionally the newly introduced *CBH1* Kozak-Sequence.

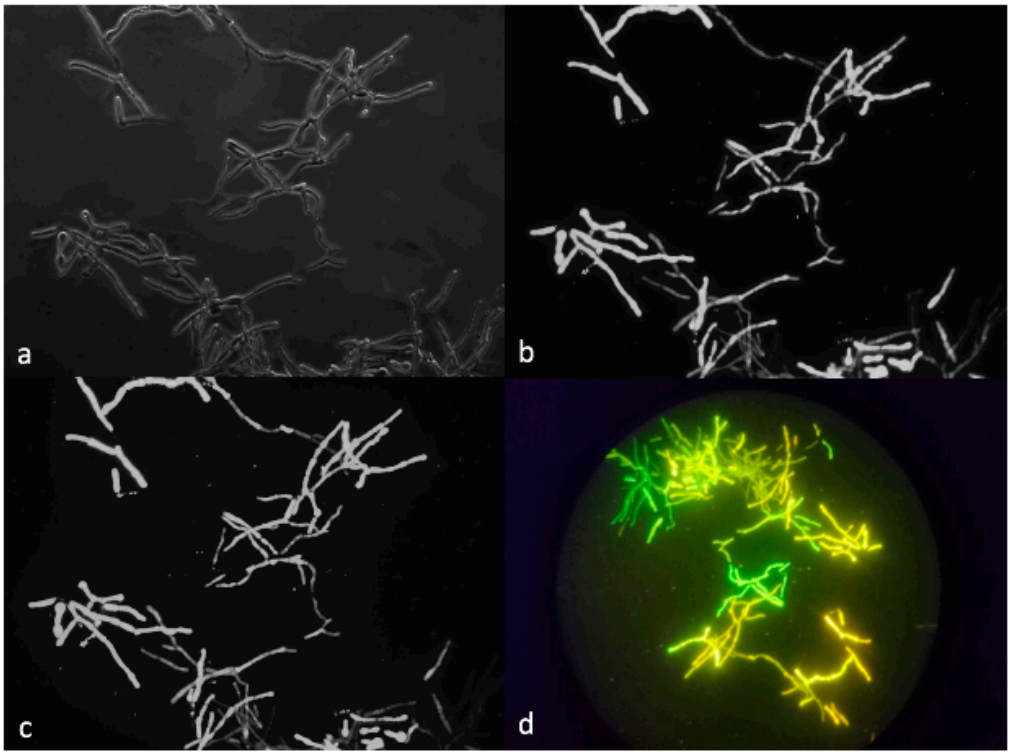


Figure 63 Pictures of *T. reesei* NRRL15500b with integrated construct for expression of eGFP and sTomato controlled by $P_{CBH1/PDC}$. The promoter had the native promoter Kozak-Sequence and additionally the newly introduced *CBH1* Kozak-Sequence. The mycelium was grown for five days on PDA plates at 28°C. (a) bright field; (b) using filter for 355-425nm, stimulating green fluorescence; (c) using filter for 515-560nm, stimulating red fluorescence. (d) shows the mycelium stimulated by UV light showing red and green fluorescence. Strong fluorescence from eGFP and sTomato was detected for the correct wavelengths indicating the vector worked well in combination with this strain as well as with the new promoter.

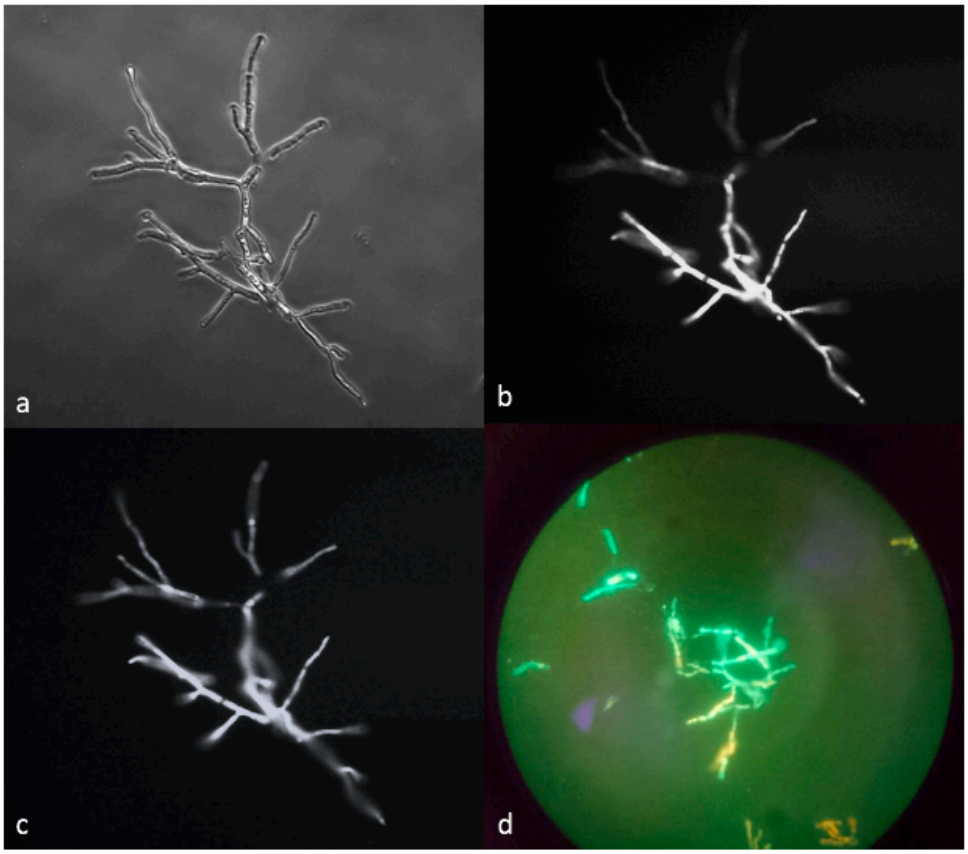


Figure 64 Pictures of *T. reesei* NRRL15500b with integrated construct for expression of eGFP and sTomato controlled by $P_{Ribosome}$. The promoter had the native promoter Kozak-Sequence and additionally the newly introduced *CBH1* Kozak-Sequence. The mycelium was grown for five days on PDA plates at 28°C (a) bright field; (b) using filter for 355-425nm, stimulating green fluorescence; (c) using filter for 515-560nm, stimulating red fluorescence. (d) shows the mycelium stimulated by UV light showing red and green fluorescence. Strong fluorescence from eGFP and sTomato was detected for the correct wavelengths indicating the vector worked well in combination with this strain as well as with the new promoter.

Figure 61, Figure 62, Figure 63 and Figure 64 gave a strong indication, that the system which was designed and assembled during this master thesis was functional and can be used to drive the co-expression of two proteins simultaneously employing new bidirectional promoters in *Trichoderma reesei*.

The bidirectional promoter sequences resulting from the tool for finding natural bidirectional promoters in the genome of *T. reesei* worked as well. It was also possible to link two small monodirectional promoters form *T. reesei* to get a novel bidirectional promoter for protein expression.

When comparing Figure 61b and Figure 61c it could be seen, that not all subparts of the mycelium showed the same level of fluorescence. The most probable reason for this is that the protoplasts used for transformation had several nuclei, and these heterokaryotic transformants exhibit in different expression levels. For this reason the two rounds of clone separation as

described in the transformation protocol were of essential importance to get homokaryotic strains and therefore equal expression levels. [88]

However, Figure 63d and Figure 64d showed, that expression of eGFP (green) and sTomato (red/orange) differs for different parts of the mycelium. Simultaneous expression of both fluorescence proteins (yellow) indicates, that expression is not uniform although the clones were separated and isolated twice during the transformation process.

The functionality of the artificial promoter, which was built by tail-to-tail fusion of the *CBH1* promoter with the *PDC* promoter, was confirmed by the observed fluorescence of eGFP and sTomato. The original plan was to have one side of the bidirectional construct giving constitutive expression (P_{PDC}) and the other side inducible one (P_{CBH1}). Cultivation of all transformants was done on PDA (full media with 20g/l glucose) plates and although no fluorescence was expected, because the eGFP was under control of the P_{CBH1}, which is subjected to glucose repression, a distinct green fluorescence could be detected (Figure 63d).

Possible explanations therefore are, that either the promoter sequence which was chosen was too short and that the sequences making up the regulatory elements which lead to glucose repression were cut off, or that the strains used for transformation were insensible to glucose repression. The result from SDS-PAGE support the first theory as no cellulases were secreted during cultivation experiments with glucose as sole carbon source.

The general consensus sequence for the binding of the repressor molecule is 5'-SYGGRG-3'. The letters S,Y and R are so called wobble base pairs which can be exchanged by C or T for Y, G or C for S and R can be exchanged by G or A [46]. Combination of all possibilities leads to eight different cre1 repressor consensus sequences, which are listed in Table 27

Table 27 Possible variations of Cre1 repressor consensus sequence. [46]

	Sequence 5' --> 3'
Var1	gcggag
Var2	ccggag
Var3	gcgggg
Var4	ccgggg
Var5	gtggag
Var6	ctggag
Var7	gtgggg
Var8	ctgggg

These possible repressor consensus sequences were aligned to the *CBH1* locus of *Trichoderma reesei* QM6a. The alignment showed that none of the repressor consensus sequences was within the sequence, which was chosen as promoter sequence. The alignment also revealed, that all consensus sequences were within 2500 base pairs up- and downstream of the promoter. The consensus sequences Var2, Var6, Var7, Var8 are located in 5'→ 3' upstream and Var1, Var3, Var4, Var5 are located in 3'→5' downstream of the promoter (Figure 65). Two of these short consensus sequences were even within the coding region of the *CBH1*.

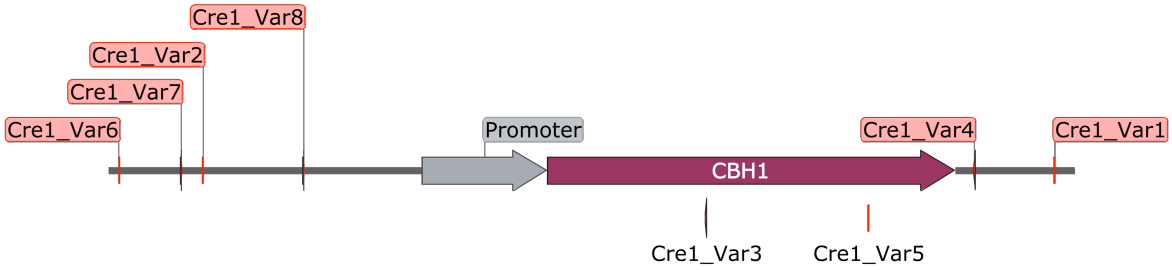


Figure 65 Variations 1-8 of cre1 repressor consensus sequence aligned to the CBH1 locus. The consensus sequences Var2, Var6, Var7, Var8 are located in 5'→ 3' upstream of the promoter sequence and the repressor consensus sequences Var1, Var3, Var4, Var5 are located in 3'→5' down steam of the promoter.

Although this could be a coincidence as statistically a fragment of the same six base pairs occurs every 4096 base pairs, one might speculate that the localisation of the various consensus sequences implies that the cre1 repressor molecule binds to two or more consensus sequences up and down stream of the promoter at the same time. Consequently, the resulting loop structure would lead to an inaccessibility of parts of the promoter sequence for the RNA polymerase (Figure 66, grey) and therefore lowering transcription.

It was shown that the most enzymes for biomass degradation are controlled on transcription level. The Cre1 repressor is the main repressor molecule using a zinc-finger motife for binding to the consensus sequences. [83] Literature as well as the results from this work showed that in the presence of glucose no cellulase expression/secretion occur.[89] Most probably two CRE1 biding sites upstream of the promoter were needed for transcription regulation as shown for the *CBH1* and *XYN1* when cultivated with glucose as sole carbon source. Mutations within these sites lead to constitutive expression of the genes although cultivation was done with glucose as primary carbon source which should have led to strong repression. Nevertheless the exact regulation mechanism had not been detected yet and recent

research showed that loop formation is a widely spread control mechanism for transcription initiation in eukaryotic organisms, which can also be found for the control of BDPs. [84], [90]

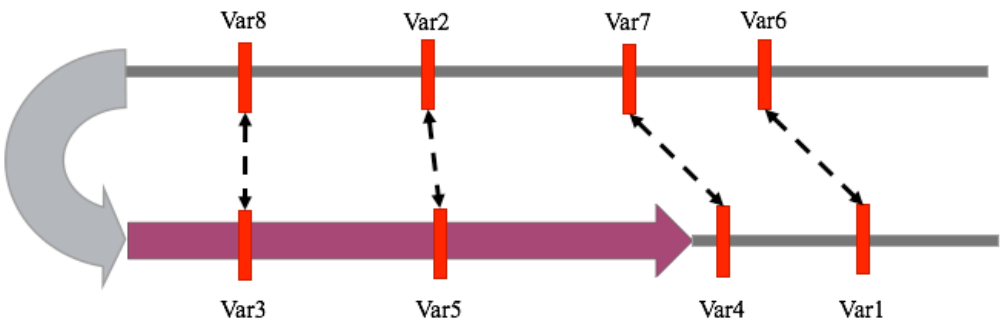


Figure 66 Possible transcription control mechanism of *cbh1*. One possible theory would be that the Cre1 repressor molecule binds to one or more consensus sequences and introduces a loop that prevent the RNA-Polymerase from binding to the core promoter region (grey). This would fit too current publications describing loop formation as a widely spread control mechanism for eukaryotic organisms. [90]

To verify this hypothesis the bases which form the consensus sequence proposed to regulate e.g. the transcription of the *CBH1* gene of Trichoderma, could be exchanged systematically. If our theory is true CBH1 should be expressed in increasing amounts as more bases are changed.

Expression levels are probably dependent on the locus of integration within the genome of *Trichoderma reesei*. Protoplasts prepared in the same batch, which were transformed with identical constructs showed different strengths of fluorescence. The constructs which was transformed hat the $P_{CBH1/PDC}$ to control the transcription of the marker protein. *sTomato* was controlled by the constitutive P_{PDC} while eGFP transcription was regulating by the P_{CBH1}.

For clones the expression levels were so high, that the colour of the mycelium switched from white to red. However, the clone which showed the strongest expression of *sTomato* also showed impaired spore formation abilities.

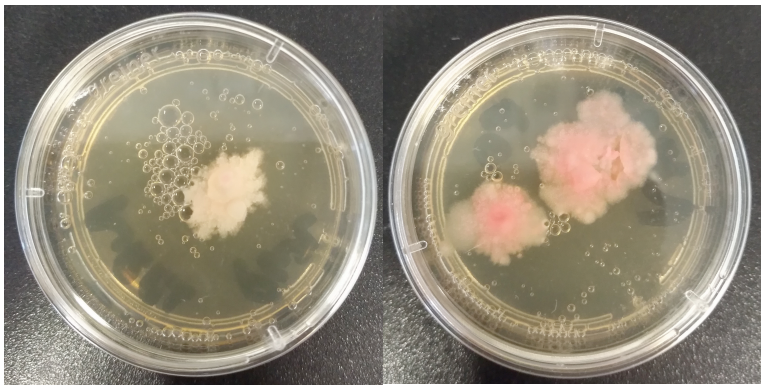


Figure 67 Pictures of *Trichoderma reesei* NRRL1550b with the construct for expression of eGFP and sTomato controlled by $P_{CBH1/PDC}$ integrated into the genome. Transcription of the sTomato was controlled by the constitutive PDC promoter while the eGFP was controlled by the CBH1 promoter. The left pictures show a white mycelium while the left mycelium has changed its color to red. Both strains showed fluorescence, which indicates that the integration locus is an important factor for the expression level.

For the final experiment the *T. reesei* strain NRRL15500b was transformed with the expression cassettes controlling transcription of the fluorescence genes by $P_{PDC/CBH1}$ and $P_{Ribosome}$ in all Kozak-Sequence variations. This experiment should also show if it is possible to isolate intracellularly produced proteins in an easily applicable downstream process. Further this experiment should reveal the best Kozak-Sequence for fine-tuning of transcription.

In order to analyse if the carbon source influences the expression of the fluorescence marker proteins, 50ml of Mandel's medium was supplemented either with 1% Avicel, lactose or glucose as sole carbon source. Each strain was cultivated in all different media and inoculation was done with the same number of spores. Prior to media inoculation fluorescence of the strains was tested via fluorescence microscopy to exclude inactivity of the fluorescence proteins in the spores.

At first problems occurred during inoculation, as it was not possible to use 100 µL of spore stock with an OD_{600} of 10. As already discussed, we observed that some Trichoderma transformants showed impaired spore formation abilities or that the spores lost the colour, which made it nearly impossible to inoculate each flask with the same number of spores.

The Cultivation processes showed that strains with an integrated expression cassette seemed to have altered growth behaviour compared to the wild type. (Figure 68) This could either be because not enough spores were used for inoculation or because the integration of our constructs lead to a change in morphology. Similar things had been observed for *Pichia pastoris* [8]. As the expression of the marker proteins is achieved by integration into the

genome of *T. reesei* another gene or ORF is disrupted. In case the construct integrates by double cross over into the *CBH1*locus the major cellulase got knocked out which would most probably lead to a slower growth on media containing complex polysaccharides such as Avicel.

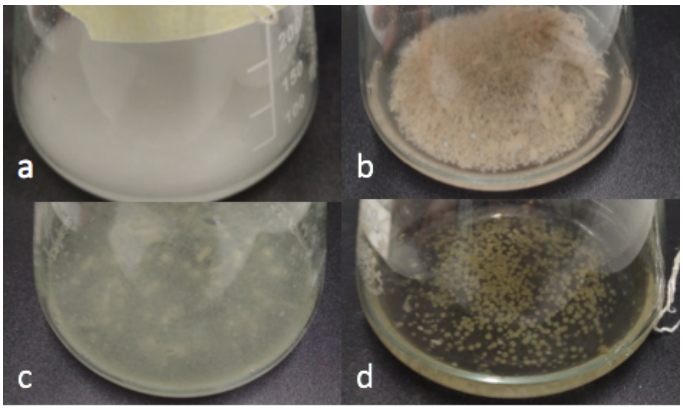


Figure 68 Changed morphology of strains expressing eGFP and sTomato. The WT NRRL1550b strain on Avicel (a) and on lactose (c) shows almost heterogenic growth. After integration of the expression cassette for the expression of eGFP and sTomato morphology of the strains changed to small spherically cell pellets on Avicel (b) and lactose (d). Alternatively, formation of these aggregates could also be a result of using not enough spores for inoculation of the media.

The different constellations of the promoter variants were marked with the suffix "n", "ok" and "wl" indicating an additional *CBH1*-Kozak-Sequence, the native Kozak-Sequence and the combination of both Kozak-Sequences. The prefix "P" and "R" is indicated the different promoters $P_{CBH1/PDC}$ and $P_{Ribosome}$.

Because inhomogeneous growth was observed for the different strains the whole culture was harvested and homogenized with liquid N_2 and a mortar. For each strain, a defined amount (about 100mg) of the cell-powder was divided into two reaction tubes. The sample in one tube was immediately suspended in 800µl in Tris-HCl (pH7.4). The other sample was used to determine the cell dry weight (CDW) and for following SDS-PAGE analysis.

To calculate the CDW about 100mg of cell lysate was added to a reaction tube and the weight determined. The samples were lyophilized overnight and weight was determined again. By calculating the difference of the weight before and after lyophilisation, the percentage of liquid within the sample could be determined.

The total protein concentration of the cell lysate was determined by the Bradford assay in triplicates. To compare the amounts of isolated protein for the different strains and different cultivation conditions the results were normalized to 100 mg of lyophilized cell lysate (Figure 69).

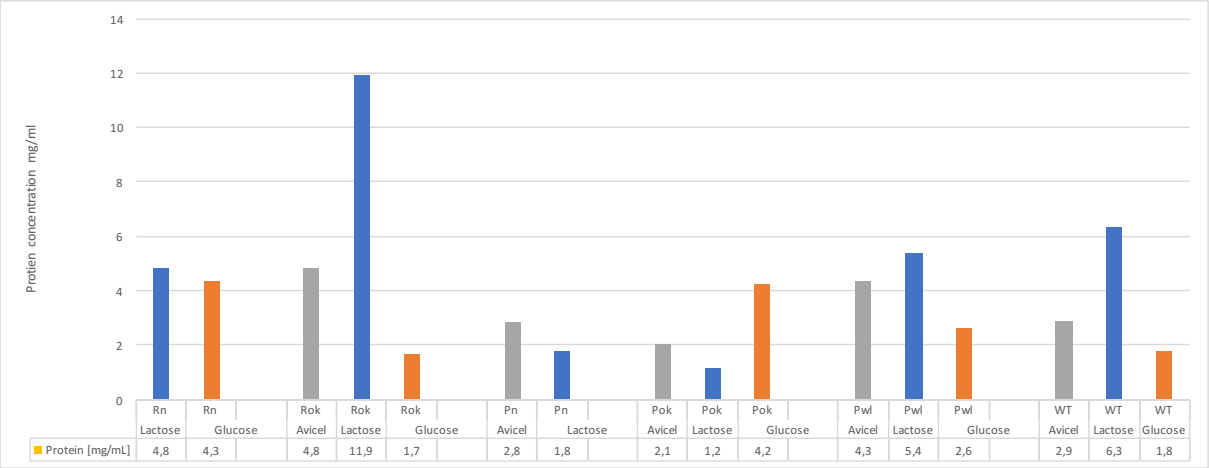


Figure 69 Comparison of protein concentrations measured with Bradford assay from cell lysate of different *T. reesei* expression strains normalized to 100 mg of CDW. Differences could be detected for the different constructs with the different promoters, Pn, Pok, Pwl, Rn, Rok and Rwl. The suffix "n", "ok" and "wl" indicating an additional CBH1-Kozak-Sequence, the native Kozak-Sequence and the combination of both Kozak-Sequences. The prefix "P" and "R" is indicating the different promoters $P_{CBH1/PDC}$ and $P_{Ribosome}$.

The protein concertation determined by the Bradford assay after cell disruption was total different for each strain and each carbon source used for cultivation.

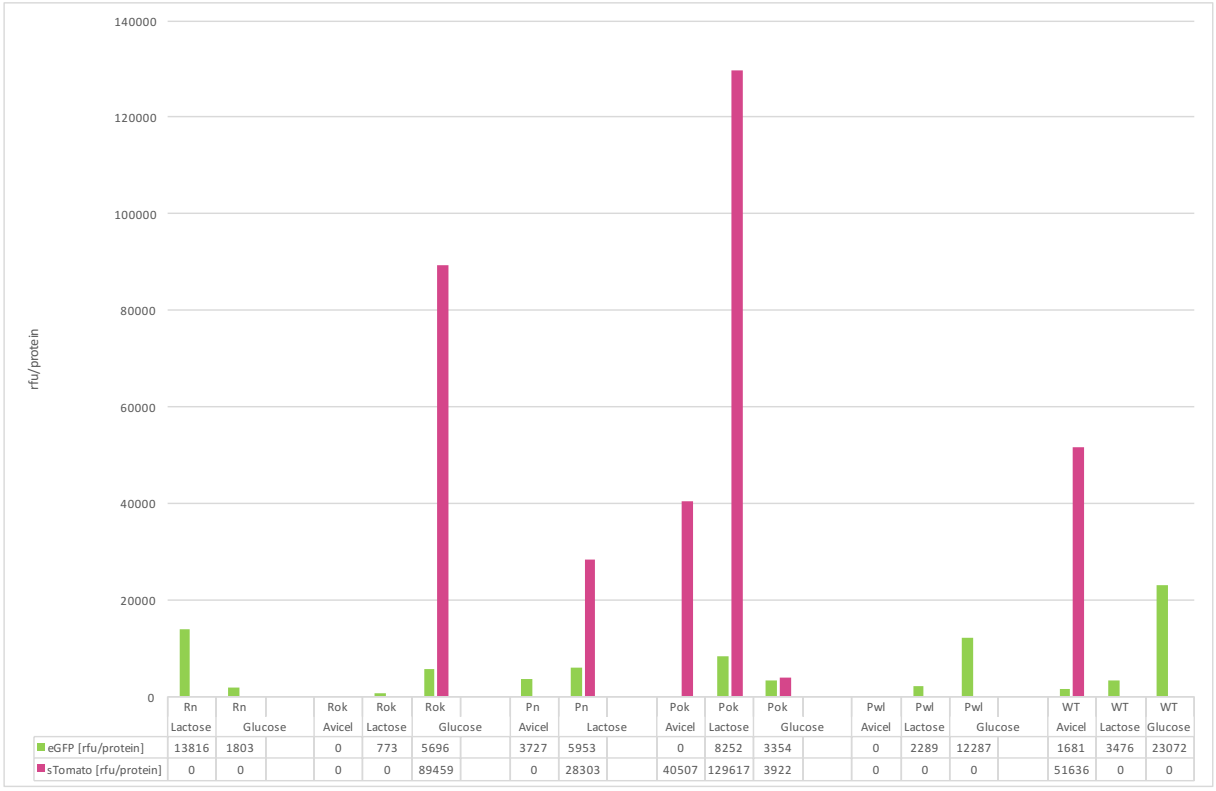
The strain with the P_{Rok} vector showed the highest protein concentration for lactose and a about one third of the protein concentration for the cultivation on Avicel. A possible reason could be that *CBH1,* the main cellulase of the organism was knockout out, but never the less higher protein concentration for cultivation on glucose where expected as the highest biomass formation was observed. Similar results were observed for the strain were transcription of the fluorescence proteins was controlled by P_{Pwl} although the difference in protein concentration for cultivation with lactose and Avicel as carbon source was not that high cultivation process with glucose as carbon source was much lower.

The other clones with the different promoter did not show the same correlation of protein concentration and culture media. At least the protein concertation of the negative control (Trichoderma NRRL15500b without integrated expression cassette) cultivated on different carbon sources were expected to be comparable but this data was also not really convincing.

Additionally, to protein quantification by Bradford assay we used the suspended cell lysate for fluorescence measurement. 250 µl lysate were transferred to a microtiter plate, and measured in triplicates. Fluorescence was determined at 325-435nm (eGFP) and at 515-560 nm (sTomato). The fluorescence was normalized to the amount of protein (rfu/ [mg] protein) in order to be able to compare the results from the different promoters and cultivation conditions.

As the plasmid was designed to integrate to the *CBH1* locus and it is well described that the

CBH1 is controlled by the external concentration of complex polysaccharides and other plant polymers [83]. Different carbon sources were used for cultivation to be sure no limitations for example from epigenetic effects occur.



	Rn Lactose	Rn Glucose	Rok Avicel	Rok Lactose	Rok Glucose	Pn Avicel	Pn Lactose	Pok Avicel	Pok Lactose	Pok Glucose	Pwl Avicel	Pwl Lactose	Pwl Glucose	WT Avicel	WT Lactose	WT Glucose
eGFP [rfu/protein]	13816	1803	0	773	5696	3727	5953	0	8252	3354	0	2289	12287	1681	3476	23072
sTomato [rfu/protein]	0	0	0	0	89459	0	28303	40507	129617	3922	0	0	0	51636	0	0

Figure 70 Fluorescence measurement referred to the protein concentration. The suffix "n", "ok" and "wl" indicating an additional CBH1-Kozak-Sequence, the native Kozak-Sequence and the combination of both Kozak-Sequences. The prefix "P" and "R" is indicating the different promoters $P_{CBH1/PDC}$ and $P_{Ribosome}$. The transformed *T. reesei* clones with the integrated plasmid with the $P_{CBH1/PDC}$ controlled the exprrssion of eGFP with the P_{CBH1} while the constitiutive P_{PDC} controlled the expression of the sTomato. Measurement for eGFP did not work well as high levels of autofluorescence in negative control were detected. RFU values from sTomato after cultivation on lactose and glucose were more reliable at least for Rok and Pok although also high RFU values for sTomato could be detected in the negative control. Expectations would have been low fluorescence for the wild type and clear signals for eGFP and sTomato.

For the different strains, eGFP and sTomato a of comparable intensity of fluorescence (rfu/[mg]protein) was expected as the clones which were used for inoculation showed about the same fluorescence in the fluorescence microscopic analysis.

However, the data received from fluorescence spectrophotometry were not reliable as the wildtype control showed an unreasonable high autofluorescence. Some clones showed higher fluorescence values (rfu/[mg]protein) for GFP but as the transcription is controlled by the constitutive PDC promoter about the same values for all cultivation conditions were expected. However, the exact regulation of the PDC promoter by different carbon sources is not known.

As the data delivered from the plate reader measurement were not reliable the lyophilized cell

lysate was used to perform SDS-PAGE analyses. Successful protein isolation from the cells should also include the coexpressed sTomato and eGFP proteins showing their characteristic bands at the size of about 25kDa.

Lanes 5 and 13 in Figure 71 and lane 16 in Figure 72 showed a band at the correct size of 25kDa but as cell lysate was used as sample this band could also be something else as the band is quite weak. Additionally, the wildtype negative control (lane 17-19, Figure 72) gave no evidence if the fluorescence proteins were expressed or not.

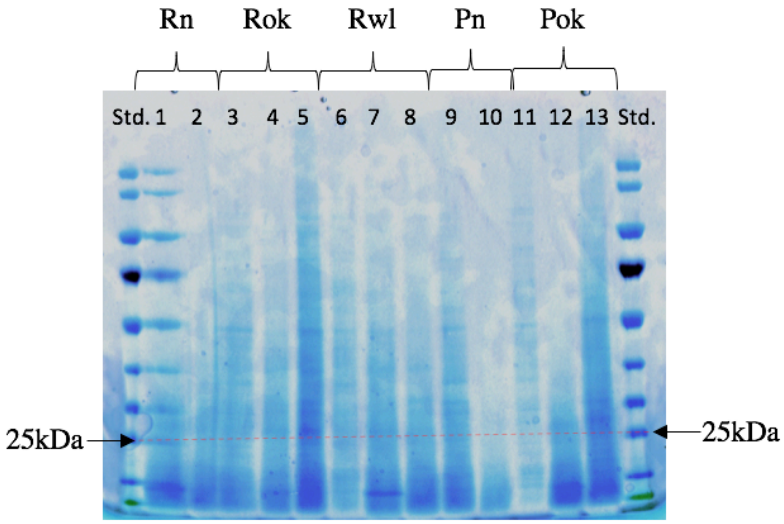


Figure 71 SDS-PAGE from the transformed *T. reesei* strain NRRL 15500b with the different promoters variations. Rn 1-2; Rok 3-5, Rwl 6-8, Pn 9-10, Pok 11-13 Cultivations with Avicel were the lanes 3, 6, 9,11 with lactose 1,4,7, 10, 12 and with glucose 2, 4, 8,13. Lanes 5 and 13 may had a band at the correct size of 25kDa.

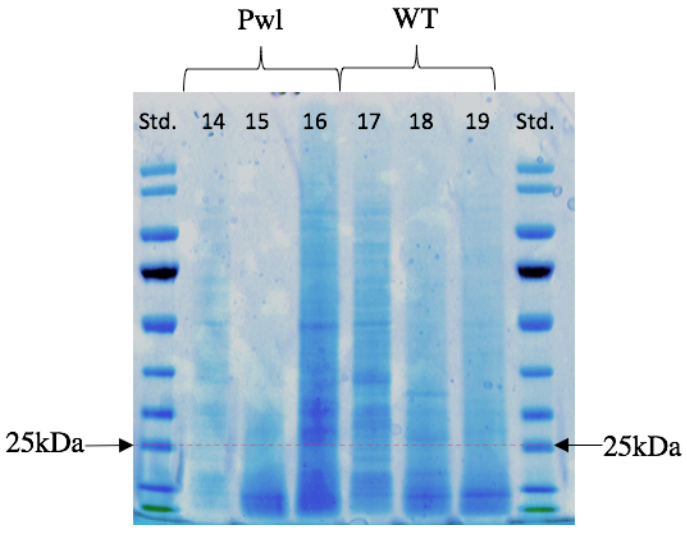


Figure 72 SDS-PAGE from the transformed *T. reesei* strain NRRL 15500b with the promoter Pwl 14-16 and the wild type (17-19). Cultivation with avicel was in lane 14 and, 17 with lactose in lane 15 and 18 and and with glucose in lane 16-19. Lane 16 may showed a band 25kDa.

Although the SDS-PAGE analysis showed some slight bands with the correct size e.g. 5, 13 and 16 these results were not totally conclusive. It is not certain if the bands were due to overexpressed eGFP/sTomato or if another protein in the cell lysate caused this signal. These bands could also be a result of degradation of various proteins by enhanced protease activity, which is a well described phenomenon for recombinant protein production in Trichoderma. [10]

The inconsistency of these final results implied that protein isolation from the mycelium did not work very well. Although the *Trichoderma reesei* NRRL15500b clones showed fluorescence under the microscope the fluorescence proteins in the cell lysate could not be confirmed by the additional analysis.

Problems occurred most probably during cell disruption, as it was maybe not the correct method to use liquid nitrogen and a mortar to extract the proteins from the mycelium. A further reason why it was maybe not possible to detect the proteins in the different experiments was the enhanced protease activity of genetically modified Trichoderma strains.

6. Conclusion and Outlook

Within the six months this thesis lasted it was possible to fulfil the objectives we defined at the start of the project.

Due to identified mistakes in the pre-existing vector and the need for a bidirectional entry vector a new vector was designed on base of the previous work and the DNA parts were amplified form genomic DNA of *T. reesei* QM6a. The new assembly of an expression vector for recombinant protein production in *T. reesei,* took about four months.

Supported by Christoph Reisinger a new tool was created showing us all hypothetical bidirectional promoters of *T. Reesei* QM6a. One such bidirectional promoter responsible for transcription of a ribosomal and an unknown protein ($P_{Ribosome}$) was tested. Additionally, a new artificial bidirectional promoter was generated ($P_{CBH1/PDC}$) by linking two monodirectional promoters, P_{CBH1} and P_{PDC}. Amplification of a histone promoter which was already shown in the work of Teuschler, did not work. Although a larger fragment was amplified and the sequence was confirmed, amplification of the exact promoter region was not possible. In the end, two new bidirectional promoters ($P_{Ribosom}$ & $P_{CBH1/PDC}$) were tested with three different variants of the Kozak-sequence.

Additionally, to the cloning work and creation of recombinant *T. reesei* strains, 8 different available commercial applied *T. reesei* strains were tested for the optimal conditions for spore formation and their secretion behaviour with three different sole carbon sources was analysed. The result was that different strains needed different growth media and light conditions for good spore formation. The strains *T. reesei* QM6a, QM9414 and RutC30 showed the best sporulation when incubated under daylight conditions independent of the growth media while the most crucial factor for the strains NRRL15499, NRRL15500a, NRRL 15500b, NRRL15500c and NRRL15502 was the cultivation on sporulation media.

Further, a SOP for the transformation of *T. reesei* leading to a transformation rate of 40-50 mfu/µg DNA was tested and described as new reliable protocol for future *Trichoderma reesei* transformations. Like in the work of Teuschler a PEG-mediated protoplast transformation protocol was used in this thesis. For this reason, the most probable reason for the higher transformation efficiency is the mutation free selection cassette which was part of the newly constructed vector.

Pictures of fluorescence microscopy demonstrated the functionality of the new plasmids and

showed that both promoter types worked. Unfortunately, it was not possible to say which Kozak-Sequence variant works the best, for intracellular fluorescence protein expression which should have been determined by fluorescence spectroscopy. Due to inhomogeneous growth and non-optimized protein isolation for the intracellularly produced reporter proteins, eGFP and sTomato, from the mycelium no useful data could be generated. The conclusion would be, that eGFP and sTomato can be used to test the activity of different promoters by fluorescence microscopy, but they are not well suited to characterize the strength of promoters. Secreted proteins, like GOx or HRP, are most likely more suitable at least to identify promoters leading to secreted proteins. Codon optimized sequences of these proteins including the CBH1 secretion signal were ordered as synthetic DNA and are already in the primer library of the Glieder-Group. (Internal numbers: P17814 and P17815). Codon usage was done like described in the work of Te'o et al. [67].

The next experiments should be to clone HRP and GOx into the expression construct. Best would be to use both promoters ($P_{CBH1/PDC}$ and $P_{Ribosome}$) and the RutC30 strain, as this one so far showed, the most favourable secretion behaviour.

In addition, a comparison of all used strains would be very interesting in order to see if only secretion of the well-known cellulose degrading enzymes such as CBH1 are strongly secreted or if some of these strains shows a generally more efficient secretion of heterologous proteins too. Cultivation should be done on Mandel's media with Avicel and glucose as sole carbon source in parallel to see if secretion works under cellulase-expression repressing conditions.

Further experiments should include fine-tuning the transformation protocol for the different strains, as protoplast formation is really important for transformation efficiency. First attempts could be done by increasing the amount of lysing enzyme from 15 to 30 g/l, as time of incubation and shaking speed had no influence on the protoplast formation in the performed experiments of this thesis.

I would definitely wait with further promoter testing experiments, seeing as, at the moment, the expertise in the handling of *T. reesei* as recombinant protein producer is not sufficient. Experiments to determine the efficiency of site specific integration are missing which could lead to a smaller vector design which could in turn enhance efficiency during cloning procedure.

7. References

[1] A. D. Bandaranayake and S. C. Almo, “Recent adcances in mammalian protein production,” *Febs Lett.*, vol. 588, no. 2, pp. 253–260, 2015.

[2] M. Paloheimo, T. Haarmann, and S. Mäkinen, *Gene Expression Systems in Fungi: Advancements and Applications*. 2016.

[3] A. R. Gomes, S. M. Byregowda, V. B. MUNIVEERAPPA, and V. BALAMURUGAN, “An Overview of Heterologous Expression Host Systems for the Production of Recombinant Proteins,” vol. 4, no. 4, pp. 346–356, 2016.

[4] A. Lakowitz, T. Godard, R. Biedendieck, and R. Krull, “Mini review: Recombinant production of tailored bio-pharmaceuticals in different Bacillus strains and future perspectives,” *Eur. J. Pharm. Biopharm.*, 2016.

[5] J. A. Glaser, “White Biotechnology,” *Clean Technol. Environ. Policy*, vol. 7, no. 4, pp. 233–235, 2005.

[6] J. H. Song, R. J. Murphy, R. Narayan, and G. B. H. Davies, “Biodegradable and compostable alternatives to conventional plastics,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 364, no. 1526, pp. 2127–2139, 2009.

[7] B. Max, J. M. Salgado, N. Rodríguez, S. Cortés, A. Converti, and J. M. Domínguez, “Biotechnological production of citric acid.,” *Braz. J. Microbiol.*, vol. 41, no. 4, pp. 862–75, 2010.

[8] R. Weis, P. Poechlauer, R. Bona, W. Skranc, R. Luiten, M. Wubbolts, H. Schwab, and A. Glieder, “Biocatalytic conversion of unnatural substrates by recombinant almond R-HNL isoenzyme 5,” *J. Mol. Catal. B Enzym.*, vol. 29, no. 1–6, pp. 211–218, 2004.

[9] J. Lippincott-Schwartz, T. H. Roberts, and K. Hirschberg, “SECRETORY PROTEIN TRAFFICKING AND ORGANELLEDYNAMICS IN LIVING CELLS,” *Rev. Lit. Arts Am.*, 2003.

[10] W. Wang, F. Meng, P. Liu, S. Yang, and D. Wei, “Construction of a promoter collection for genes co-expression in filamentous fungus Trichoderma reesei,” *J. Ind. Microbiol. Biotechnol.*, vol. 41, no. 11, pp. 1709–1718, 2014.

[11] J. Berg, J. Tymoczko, and L. Stryer, *Biochemistry. 5th edition.* New York: W. H. Freeman, 2002.

[12] D. J. Voet, J. G. Voet, and C. W. Pratt, *Lehrbuch der Biochemie*, 2. Wiley-VCH, 2002.

[13] M. Shintani, Z. K. Sanchez, and K. Kimbara, "Genomics of microbial plasmids: Classification and identification based on replication and transfer systems and host taxonomy," *Front. Microbiol.*, vol. 6, no. MAR, pp. 1–16, 2015.

[14] L. A. Palomares, S. Estrada-Mondaca, and Ramírez Octavio T., "Production of Recombinant Proteins," vol. 267, 2010.

[15] F. Baneyx, "Recombinant protein expression in Escherichia coli."

[16] E. Acosta-Cruz, F. Wisniewski-Dyé, Z. Rouy, V. Barbe, M. Valdés, and P. Mavingui, "Insights into the 1.59-Mbp largest plasmid of azospirillum brasilense CBG497," *Arch. Microbiol.*, vol. 194, no. 9, pp. 725–736, 2012.

[17] H. Lodisch, A. Berk, and S. Ziprusky, *Molecular Cell Biology. 4th edition*. New York: W. H. Freeman, 2000.

[18] K. Van Craenenbroeck, P. Vanhoenacker, and G. Haegeman, "Episomal vectors for gene expression in mammalian cells.," *Eur. J. Biochem.*, vol. 267, no. 18, pp. 5665–5678, 2000.

[19] Y. Y. M. Chen, H. R. Shieh, C. T. Lin, and S. Y. Liang, "Properties and construction of plasmid pFW213, a shuttle vector with the oral Streptococcus origin of replication," *Appl. Environ. Microbiol.*, vol. 77, no. 12, pp. 3967–3974, 2011.

[20] G. L. Rosano and E. A. Ceccarelli, "Recombinant protein expression in Escherichia coli: Advances and challenges," *Front. Microbiol.*, vol. 5, no. APR, pp. 1–17, 2014.

[21] A. M. Deshpande and C. S. Newlon, "The ARS consensus sequence is required for chromosomal origin function in Saccharomyces cerevisiae.," *Mol. Cell. Biol.*, vol. 12, no. 10, pp. 4305–4313, 1992.

[22] K. M. H. Nevalainen, V. S. J. Te'o, and P. L. Bergquist, "Heterologous protein expression in filamentous fungi," *Trends Biotechnol.*, vol. 23, no. 9, pp. 468–474, 2005.

[23] N. B. Jensen, T. Strucko, K. R. Kildegaard, F. David, J. Maury, U. H. Mortensen, J. Forster, J. Nielsen, and I. Borodina, "EasyClone: Method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae," *FEMS Yeast Res.*, vol. 14, no. 2, pp. 238–248, 2014.

[24] B. Twyman, R.M. and Whitelaw, "Genetic engineering: Animal cell technology, in

The Encyclopedia of Cell Technology," 2000.

[25] P. Olson, Y. Zhang, D. Olsen, a Owens, P. Cohen, K. Nguyen, J. J. Ye, S. Bass, and D. Mascarenhas, "High-level expression of eukaryotic polypeptides from bacterial chromosomes.," *Protein Expr. Purif.*, vol. 14, pp. 160–166, 1998.

[26] S. Bin Zaman, M. A. Hussain, R. Nye, V. Mehta, K. T. Mamun, and N. Hossain, "A Review on Antibiotic Resistance: Alarm Bells are Ringing," *Cureus*, vol. 9, no. 6, 2017.

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

[28] "plasmids-101-everything-you-need-to-know-about-antibiotic-resistance-genes @ blog.addgene.org." .

[29] M. Mülleder, K. Campbell, O. Matsarskaia, F. Eckerstorfer, and M. Ralser, "Saccharomyces cerevisiae single-copy plasmids for auxotrophy compensation, multiple marker selection, and for designing metabolically cooperating communities," *F1000Research*, vol. 5, no. 0, p. 2351, 2016.

[30] J. H. Nett, N. Hodel, S. Rausch, and S. Wildt, "Cloning and disruption of the Pichia pastoris ARG1 , use as auxotrophic markers," pp. 295–304, 2005.

[31] L. Näätsaari, B. Mistlberger, C. Ruth, T. Hajek, F. S. Hartner, and A. Glieder, "Deletion of the pichia pastoris ku70 homologue facilitates platform strain generation for gene expression and synthetic biology," *PLoS One*, vol. 7, no. 6, 2012.

[32] J. Nielsen, "Production of biopharmaceutical proteins by yeast," *Bioengineered*, vol. 4, no. 4, pp. 207–211, 2013.

[33] J. Reyrat, V. Pelicic, and B. Gicquel, "MINIREVIEW Counterselectable Markers : Untapped Tools for Bacterial Genetics and Pathogenesis," vol. 66, no. 9, pp. 4011–4017, 1998.

[34] D. Solis-Escalante, N. G. A. Kuijpers, N. Bongaerts, I. Bolat, L. Bosman, J. T. Pronk, J. M. Daran, and P. Daran-Lapujade, "amdSYM, A new dominant recyclable marker cassette for Saccharomyces cerevisiae," *FEMS Yeast Res.*, vol. 13, no. 1, pp. 126–139, 2013.

[35] M. Ahmad, M. Hirz, H. Pichler, and H. Schwab, "Protein expression in Pichia pastoris:

Recent achievements and perspectives for heterologous protein production," *Appl. Microbiol. Biotechnol.*, vol. 98, no. 12, pp. 5301–5317, 2014.

[36] A. Kanhere and M. Bansal, "Structural properties of promoters: Similarities and differences between prokaryotes and eukaryotes," *Nucleic Acids Res.*, vol. 33, no. 10, pp. 3165–3175, 2005.

[37] D. Omotajo, T. Tate, H. Cho, and M. Choudhary, "Distribution and diversity of ribosome binding sites in prokaryotic genomes," *BMC Genomics*, vol. 16, no. 1, p. 604, 2015.

[38] a G. Pedersen, P. Baldi, Y. Chauvin, and S. Brunak, "The biology of eukaryotic promoter prediction--a review.," *Comput. Chem.*, vol. 23, no. 3–4, pp. 191–207, 1999.

[39] S. Nakagawa, Y. Niimura, T. Gojobori, H. Tanaka, and K. ichiro Miura, "Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes," *Nucleic Acids Res.*, vol. 36, no. 3, pp. 861–871, 2008.

[40] N. Sonenberg and A. G. Hinnebusch, "Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets," vol. 136, no. 4, pp. 731–745, 2009.

[41] M. Kozak, "Influences of mRNA secondary structure on initiation by eukaryotic ribosomes.," *Proc. Natl. Acad. Sci.*, vol. 83, no. 9, pp. 2850–2854, 1986.

[42] T. Weenink, R. M. Mckiernan, and T. Ellis, "Rational Design of RNA Structures that Predictably Tune Eukaryotic Gene Expression," 2017.

[43] H. Nevalainen and R. Peterson, "Making recombinant proteins in filamentous fungi-Are we expecting too much?," *Front. Microbiol.*, vol. 5, no. FEB, pp. 1–10, 2014.

[44] J. Li, J. Wang, S. Wang, M. Xing, S. Yu, and G. Liu, "Achieving efficient protein expression in Trichoderma reesei by using strong constitutive promoters," *Microb. Cell Fact.*, vol. 11, no. 1, p. 84, 2012.

[45] W. C. Raschke, B. R. Neiditch, M. Hendricks, and J. M. Cregg, "Inducible expression of a heterologous protein in Hansenula polymorpha using the alcohol oxidase 1 promoter of Pichia pastoris," vol. 10, no. 8, pp. 179–189, 1996.

[46] T. Portnoy, A. Margeot, R. Linke, L. Atanasova, E. Fekete, E. Sándor, L. Hartl, L. Karaffa, I. S. Druzhinina, B. Seiboth, S. Le Crom, and C. P. Kubicek, "The CRE1 carbon catabolite repressor of the fungus Trichoderma reesei: a master regulator of carbon assimilation," *BMC Genomics*, vol. 12, no. 1, p. 269, 2011.

[47] T. Vogel, T. Kickenweiz, J. Pitzer, L. Sturmberger, E.-M. Köhler, A. Weninger, A. Baumschlager, J. Fischer, P. Hyden, M. Geier, and A. Glieder, "A library of bidirectional promoters facilitates fine-tuning of gene coexpression."

[48] R. W. West, S. M. Chen, H. Putz, G. Butler, and M. Banerjee, "GAL1-GAL10 divergent promoter region of Saccharomyces cerevisiae contains negative control elements in addition to functionally separate and possibly overlapping upstream activating sequences.," *Genes Dev.*, vol. 1, no. 10, pp. 1118–1131, 1987.

[49] T. S. Lopes, J. Klootwijk, A. E. Veentra, P. C. van der Aar, H. van Heerikhuizen, H. A. Raue, and R. J. Planta, "High-copy-number integration into the ribosomal DNA of Saccharomyces cerevisiae: a new vector for high-level expression," *Annu. Rev. Genet.*, vol. 99, no. 1215, pp. 261–311, 1999.

[50] Z. Zhou, Y. Dang, M. Zhou, L. Li, C. Yu, J. Fu, S. Chen, and Y. Liu, "Codon usage is an important determinant of gene expression levels largely through its effects on transcription," *Proc. Natl. Acad. Sci.*, vol. 113, no. 41, pp. E6117–E6125, 2016.

[51] Y. J. Choi, D. Bourque, L. Morel, D. Groleau, and C. B. Míguez, "Multicopy integration and expression of heterologous genes in methylobacterium extorquens ATCC 55366," *Appl. Environ. Microbiol.*, vol. 72, no. 1, pp. 753–759, 2006.

[52] M. Schmoll and A. Schuster, "Biology and biotechnology of Trichoderma," *Appl. Microbiol. Biotechnol.*, vol. 87, no. 3, pp. 787–799, 2010.

[53] R. Peterson and H. Nevalainen, "Trichoderma reesei RUT-C30 - Thirty years of strain improvement," *Microbiology*, vol. 158, no. 1, pp. 58–68, 2012.

[54] A. Ghosh, B. k. Ghosh, H. Trimino-Vazquez, D. E. Eveleigh, and B. S. Montencourt, "Cellulase secretion from a hyper-cellulolytic mutant of Trichoderma reesei Rut-C30," 1984.

[55] C. P. Kubicek, "Systems biological approaches towards understanding cellulase production by Trichoderma reesei," *J. Biotechnol.*, vol. 163, no. 2, pp. 133–142, 2013.

[56] C. P. Kurtzman, "Biotechnological strains of Komagataella (Pichia) pastoris are Komagataella phaffii as determined from multigene sequence analysis," *J. Ind. Microbiol. Biotechnol.*, vol. 36, no. 11, pp. 1435–1438, 2009.

[57] L. Sturmberger, T. Chappell, M. Geier, F. Krainer, K. J. Day, U. Vide, S. Trstenjak, A. Schiefer, T. Richardson, L. Soriaga, B. Darnhofer, R. Birner-Gruenberger, B. S. Glick,

I. Tolstorukov, J. Cregg, K. Madden, and A. Glieder, "Refined Pichia pastoris reference genome sequence," *J. Biotechnol.*, vol. 235, pp. 121–131, 2016.

[58] R. T. Rowlands, "Industrial strain improvement: Mutagenesis and random screening procedures," *Enzyme Microb. Technol.*, vol. 6, no. 1, pp. 3–10, 1984.

[59] G. Yuen, F. J. Khan, S. Gao, J. M. Stommel, E. Batchelor, X. Wu, and J. Luo, "CRISPR/Cas9-mediated gene knockout is insensitive to target copy number but is dependent on guide RNA potency and Cas9/sgRNA threshold expression level," *Nucleic Acids Res.*, vol. 45, no. 20, pp. 12039–12053, 2017.

[60] S. Bin Kim and S. Timmusk, "A Simplified Method for Gene Knockout and Direct Screening of Recombinant Clones for Application in Paenibacillus polymyxa," *PLoS One*, vol. 8, no. 6, 2013.

[61] W. Kricka, J. Fitzpatrick, and U. Bond, "Metabolic engineering of yeasts by heterologous enzyme production for degradation of cellulose and hemicellulose from biomass: A perspective," *Front. Microbiol.*, vol. 5, no. APR, pp. 1–11, 2014.

[62] F. Meng, D. Wei, and W. Wang, "Heterologous protein expression in Trichoderma reesei using the cbhII promoter," *Plasmid*, vol. 70, no. 2, pp. 272–276, 2013.

[63] "There's a saying that 'Expression systems aren't everything, but everything is nothing without a good expression system' | Mark A. Emalfarb | Dyadic International, Inc., USA | Euro Biosimilars 2018 | Conferenceseries Ltd." [Online]. Available: https://biosimilars-biologics.pharmaceuticalconferences.com/europe/abstract/2015/there-s-a-saying-that-expression-systems-aren-t-everything-but-everything-is-nothing-without-a-good-expression-system. [Accessed: 21-Nov-2017].

[64] A. Karnaouri, E. Topakas, I. Antonopoulou, and P. Christakopoulos, "Genomic insights into the fungal lignocellulolytic system of Myceliophthora thermophila," *Front. Microbiol.*, vol. 5, no. JUN, 2014.

[65] N. Teuschler, "Vectorsystem studies for recombinant protein production in Trichoderma reesei and Myceliophthora thermophila," 2017.

[66] J. R. Cherry and A. L. Fidantsef, "Directed evolution of industrial enzymes: An update," *Curr. Opin. Biotechnol.*, vol. 14, no. 4, pp. 438–443, 2003.

[67] V. S. J. Te'o, A. E. Cziferszky, P. L. Bergquist, and K. M. H. Nevalainen, "Codon

optimization of xylanase gene xynB from the thermophilic bacterium Dictyoglomus thermophilum for expression in the filamentous fungus Trichoderma reesei," *FEMS Microbiol. Lett.*, vol. 190, no. 1, pp. 13–19, 2000.

[68] D. Lv, W. Wang, and D. Wei, "Construction of two vectors for gene expression in Trichoderma reesei," *Plasmid*, vol. 67, no. 1, pp. 67–71, 2012.

[69] D. Haab, K. Hagspiel, K. Szakmary, and C. P. Kubicek, "Formation of the extracellular proteases from Trichoderma reesei QM 9414 involved in cellulase degradation," *J. Biotechnol.*, vol. 16, no. 3–4, pp. 187–198, 1990.

[70] E. V. Eneyskaya, A. A. Kulminskaya, K. A. Shabalin, K. N. Neustroev, A. N. Savel'ev, and N. V. Savel'eva, "Acid protease from Trichoderma reesei: Limited proteolysis of fungal carbohydrases," *Appl. Microbiol. Biotechnol.*, vol. 52, no. 2, pp. 226–231, 1999.

[71] G. Zhang, Y. Zhu, D. Wei, and W. Wang, "Enhanced production of heterologous proteins by the filamentous fungus Trichoderma reesei via disruption of the alkaline serine protease SPW combined with a pH control strategy," *Plasmid*, vol. 71, no. 1, pp. 16–22, 2014.

[72] A. L. Zydney, "Continuous downstream processing for high value biological products: A Review," *Biotechnol. Bioeng.*, vol. 113, no. 3, pp. 465–475, 2016.

[73] "No Title," vol. 99, no. 1.

[74] R. Mach, M. Schindler, and C. P. Kubicek, "Transformation of Trichoderma reesei based on hygromycin B resistance using homologous expression signals," 1994.

[75] Z. Zhang and B. Lutz, "Cre recombinase-mediated inversion using lox66 and lox71: method to introduce conditional point mutations into the CREB-binding protein," *Nucleic Acids Res.*, vol. 30, no. 17, pp. e90–e90, 2002.

[76] D. Martinez, R. M. Berka, B. Henrissat, M. Saloheimo, M. Arvas, S. E. Baker, J. Chapman, O. Chertkov, P. M. Coutinho, D. Cullen, E. G. J. Danchin, I. V Grigoriev, P. Harris, M. Jackson, C. P. Kubicek, C. S. Han, I. Ho, L. F. Larrondo, A. L. de Leon, J. K. Magnuson, S. Merino, M. Misra, B. Nelson, N. Putnam, B. Robbertse, A. A. Salamov, M. Schmoll, A. Terry, N. Thayer, A. Westerholm-Parvinen, C. L. Schoch, J. Yao, R. Barabote, M. A. Nelson, C. Detter, D. Bruce, C. R. Kuske, G. Xie, P. Richardson, D. S. Rokhsar, S. M. Lucas, E. M. Rubin, N. Dunn-Coleman, M. Ward,

and T. S. Brettin, “Genome sequencing and analysis of the biomass-degrading fungus Trichoderma reesei (syn. Hypocrea jecorina),” *Nat. Biotechnol.*, vol. 26, no. 5, pp. 553–560, 2008.

[77] R. Bischof, L. Fourtis, A. Limbeck, C. Gamauf, B. Seiboth, and C. P. Kubicek, “Comparative analysis of the Trichoderma reesei transcriptome during growth on the cellulase inducing substrates wheat straw and lactose.,” *Biotechnol. Biofuels*, vol. 6, no. 1, p. 127, 2013.

[78] D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, and H. O. Smith, “Enzymatic assembly of DNA molecules up to several hundred kilobases,” *Nat. Methods*, vol. 6, no. 5, pp. 343–345, 2009.

[79] E. Salminenb, “versatile transformation system for the cellulolytic filamentous fungus Z’riclroderma,” vol. 61, pp. 155–164, 1987.

[80] S. D. Said, “Spore Production by Biocontrol Agent Trichoderma Harzianum in Submerged Fermentation : Effect of Agitation and Aeration,” *J. Rekayasa Kim. dan Lingkung.*, vol. 6, no. 2, pp. 71–76, 2007.

[81] M. Schmoll, E. U. Esquivel-Naranjo, and A. Herrera-Estrella, “Trichoderma in the light of day - Physiology and development,” *Fungal Genet. Biol.*, vol. 47, no. 11, pp. 909–916, 2010.

[82] M. S. M. S. S. P. A. S. and V. K. Anuradha Singh, “Optimal Physical Parameters for Growth of Trichoderma Species at\r\nVarying pH, Temperature and Agitation,” *Virol. Mycol.*, vol. 3, no. 1, pp. 1–7, 2014.

[83] P. K. Foreman, D. Brown, L. Dankmeyer, R. Dean, S. Diener, N. S. Dunn-Coleman, F. Goedegebuur, T. D. Houfek, G. J. England, A. S. Kelley, H. J. Meerman, T. Mitchell, C. Mitchinson, H. A. Olivares, P. J. M. Teunissen, J. Yao, and M. Ward, “Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus Trichoderma reesei,” *J. Biol. Chem.*, vol. 278, no. 34, pp. 31988–31997, 2003.

[84] A. Amore, S. Giacobbe, and V. Faraco, “Regulation of Cellulase and Hemicellulase Gene Expression in Fungi,” *Curr. Genomics*, vol. 14, no. 4, pp. 230–249, 2013.

[85] P. Tomme, H. Van Tilbeurgh, G. Pettersson, J. Van Damme, J. Vandekerckhove, J. Knowles, T. Teeri, and M. Claeyssens, “Studies of the Cellulotic system of Trichoderma reesei QM9414. Analysis of domain function in the two

cellobiohydrolases by limited proteolysis," *Febs*, vol. 581, pp. 575–581, 1988.

[86] S. K. Brady, S. Sreelatha, Y. Feng, S. P. S. Chundawat, and M. J. Lang, "Cellobiohydrolase 1 from Trichoderma reesei degrades cellulose in single cellobiose steps," *Nat. Commun.*, vol. 6, no. May, p. 10149, 2015.

[87] X. Li, G. Zhang, N. Ngo, X. Zhao, S. R. Kain, J. B. Chem, and C. Huang, "Deletions of the Aequorea victoria Green Fluorescent Protein Define the Minimal Domain Required for Fluorescence Deletions of the Aequorea victoria Green Fluorescent Protein Define the Minimal Domain Required for Fluoresc," *J. Biol. Chem.*, vol. 272, no. 45, pp. 28545–28549, 1997.

[88] C. P. Kubicek and G. E. Harman, *Trichoderma& Gliocladium*. 2002.

[89] M. Ilmen, A. Saloheimo, M. L. Onnela, and M. E. Penttila, "Regulation of cellulase gene expression in the filamentous fungus Trichoderma reesei," *Appl. Environ. Microbiol.*, vol. 63, no. 4, pp. 1298–1306, 1997.

[90] P. Grzechnik, S. M. Tan-Wong, and N. J. Proudfoot, "Terminate and make a loop: Regulation of transcriptional directionality," *Trends Biochem. Sci.*, vol. 39, no. 7, pp. 319–327, 2014.

Appendix

SOP: Transformation of *T. reesei* by PEG-mediated protoplast transformation

1. Objective

Isolation of protoplasts from *Trichoderma reesei* mycelium as base for PEG-mediated protoplast transformation of linear DNA-constructs.

2. Field of application

One batch of protoplasts isolated by this protocol can be used for two transformations

3. Experimental

3.1 Chemicals

Table 1 List of all chemicals which are needed for the transformation procedure.

	Company	Item number	Lot number
$(NH_4)_2SO_4$	Roth	3746.1	470160069
KH_2PO_4	Roth	6875.1	496245825
$MgSO_4$ $7H_2O$	Roth	P027.2	267260137
$CaCl_2$ $2H_2O$	Roth	T885.1	424220363
$FeSO_4$ $7H_2O$	n.a	n.a.	n.a.
$MnSO_4$ $7H_2O$	Roth	X890.1	24670680
$ZnSO_4$ $7H_2O$	n.a	n.a.	n.a.
$CoCl_2$ $7H_2O$	n.a	n.a.	n.a.
Urea	Roth	7638.1	147255758
Lysing Enzyme	Sigma-Aldrich	L1412-25G	SLBJ0553V
Bis-Tris	Roth	6999.2	386249696
Tris	Roth	AE15.3	226230298
Sorbitol	Roth	6213.2	486251177
Sucrose	Roth	9286.2	247254570
PEG6000	Sigma-Aldrich	81253-250G	BCBQ5666V
HygromycinB	Formedia Ltd	HYG5000	17/MFM/1931

PDA	Roth	CP74.2	336248400
Glucose	Roth	6780.2	037253715
Agarose	VWR	35-1020	D00057
Glycerine	Roth	7530.4	047252665
Peptone	BD	211820	6215719
NaCl	Roth	3957.1	094209577
Tween 80	Merck	822187	0278913

3.1 Stock solutions, buffer solutions and growth medium

Table 2 List of stocks for buffers and media needed for the transformation procedure.

	Component	Amount/l
Sol1 MM 100x	$(NH_4)_2SO_4$	140g/l
Sol2 MM 100x	KH_2PO_4	200g/l
Sol3 MM 100x	$MgSO_4$ $7H_2O$	30g/l
Sol4 MM 100x	$CaCl_2$ $2H_2O$	40g/l
	$FeSO_4$ $7H_2O$	0.5g/l
	$MnSO_4$ $7H_2O$	0.16g/l
	$ZnSO_4$ $7H_2O$	0.14g/l
	$CoCl_2$ $7H_2O$	0.2g/l
Sol5 MM 100x	Urea	30g/l

With exception of Sol4 MM 100x which is containing the different salts for the Mandel's media all the stocks can be autoclaved. Sol4 MM 100x should be sterilized by sterile filtration using a membrane with a pore size of 0.2µm.

Table 3 List of buffers used for the transformation procedure.

	Component	Amount/l	
Lysing Solution	Lysing Enzyme	15g/l	Solved in solution A
Protoplast Washing Solution	Bis-Tris	10mM	pH 5,8
	$MgSO_4$ $7H_2O_4$	1,2M	
Transformation Solution A	Tris	100mM	pH 7
	Sorbitol	0,6M	

Transformation Solution B	Tris	10mM	pH 7,5
	$CaCl_2$ $2H_2O$	50mM	
	Sorbitol	1M	
Solution for Protoplast focusing	Tris	10mM	pH 7,5
	$CaCl_2$ $2H_2O$	50mM	
	Sucrose	1M	
PEG-Solution	PEG6000	250g/l	pH 7,5
	Tris	10mM	
	$CaCl_2$ $2H_2O$	50mM	

Autoclaving can be used to sterilize all buffers mentioned above

Table 4 List of media for the *T. reesei* transformation

	Component	Amount
Potato Extract Glucose Broth	Mix	26.5g/l
Bottom-Medium	Sucrose	85.6g/l
	Glucose	10g/l
	Sol1 MM 100x	10ml/l
	Sol2 MM 100x	10ml/l
	Sol3 MM 100x	10ml/l
	Sol4 MM 100x	10ml/l
	Sol5 MM 100x	10ml/l
	Hygromycin	100mg/l
Overlay-Medium	Sucrose	342,3g/l
	Glucose	10g/l
	Sol1 MM 100x	10ml/l
	Sol2 MM 100x	10ml/l
	Sol3 MM 100x	10ml/l
	Sol4 MM 100x	10ml/l
	Sol5 MM 100x	10ml/l
	Agarose	10g/l
Sporulation medium	PDA	39g/l
	Glycerine	100g/l
	Peptone	1g/l
Spore wash solution	NaCl	8g/l
	Tween 80	0.5m/l

Stocks Sol1 MM 100x- Sol5 MM 100x have to be added after autoclaving.

3.2 Materials

Table 5 List of martials needed for transformation

Amount	Device
1x	300ml flasks (sterile)
1x	Frit (porosity 1)
1x	Funnel
1x	Myracloth
1x	Glass wool
2x	15ml tube with screw cap (sterile)
8x	50ml tube with screw cap (sterile)
1x	Neubauer counting chamber
Depending on microscopy	Microscope slides and cover slip
1x	Spatula (sterile)
1x	Water bath
1x	Ice bath

4. Experimental (including Troubleshooting)

4.1 DNA linearization and ethanol participating:

For the transformation 20ng of DNA was linearized with the correct restriction enzyme according to the manual of the enzyme. Linearization efficiency is checked by agarose gel.

Ethanol precipitation:

- 1:10 of the volume of 3M natriumacetate (pH 5.5) and 2.5 volumes of ice cold EtOH (99%) is added
- Incubate at least 10 min at -20°C
- Centrifuge at top speed for 20min at 4°C.
- Wash with 1ml ice cold ethanol (75%)
- Centrifuge at top speed for 5min at 4°C
- Wash with 250µl ice cold ethanol (75%)
- Centrifuge at top speed for 5min at 4°C

- Incubation for 30min at 37°C
- Solve DNA in 20µl ddH_2O at 55°C

4.2 Protoplast formation Day 1:

Prepare medium under sterile conditions. A funnel with Myracloth is prepared and autoclaved.

- 40ml of PDA media is inoculated with 1ml Spore stock of OD_{600} 10
- Incubate for 24 hours at 30°C and 150rpm
- Mycelium growth can be controlled with a microscope.
- When enough biomass is formed a funnel with Myracloth is used to separate the mycelium from culture broth.
- The biomass is transferred to the protolyis solution and incubated for at least 5 hours at 30°C and 150rpm.

Protoplast formation is checked with the microscope.

<u>If no protoplasts can be found under the microscope reasons could be</u>

- wrong osmotic pressure due to wrong buffers → protoplasts burst
- lysing enzyme is inactive

4.3 Transformation procedure Day 2

Top media is prepared and stored at 60°C. Glass wool is put in the frit and autoclaved.

Transformation steps are performed without exception on ice.

- Protoplast solution is filtrated by the frit with glass wool and washed with max 20ml of cold solution A. The spores are collected in a 300ml flask.
- Protoplast solution is divided into 4 50ml tubes (precooled) with max. 15ml/tube
- 10ml of cold solution B is added to each of the tube. This step has to be done really carefully. A sharp boundary layer should rise between solution A and B.
- Centrifuge at 3220xg at 4°C for 30min.
- A murky layer containing the protoplast should be visible, if not take the top 7ml, and pool two phases with protoplasts into one precooled 50ml tube. (max. 15ml/tube)
- Fill the tube with cold solution C to 45ml.
- Centrifuge at 450xg at 4°C for 10min.

- The supernatant is discarded and the protoplasts are suspended in the remaining liquid. The solution in the falcons is pooled and filled up to 10ml with cold Solution C.
- Centrifuge at 450xg at 4°C for 20min.
- The supernatant is discarded and the protoplasts are suspended in the remaining liquid and 250µl of solution C is added.
- The protoplasts are counted with the Neubauer counting chamber and if necessary diluted to 2x10^7 protoplasts/ml with solution c
- 80µl of protoplasts, 800ng of linearized DNA and 20µl solution E are carefully mixed in a 15ml tube and incubated on ice for 20min.
- 22ml aliquots of Top-agar are prepared in 50ml tubes and stored at 50°C (water bath).
- 800µl of solution E is added to the protoplasts, carefully mixed and incubated for 5min at RT.
- 1600µl of solution D is added and carefully mixed.
- The whole transformation mix (3ml) is put into one aliquot of overlay media and carefully mixed
- 12.5ml of the protoplast agar mix is spreaded evenly on a Bottom media plate. The solidified agar plates are incubated for 5 days at 30°C
- Check for transformants.

Troubleshooting transformation procedure:

- If no murky layer containing the protoplasts can be seen most likely the buffers had the wrong concentrations
- If no protoplasts are visible under the microscope they may have bursted due to wrong buffers or got lost during suspending procedure, especially when no murky layer was formed

4.4 Separating the transformants over two rounds

Prepare agar plates 9.2mm and 35mm with media for enhanced spore formation and selection substrate.

- A tweezer is used to pick all single transformants and to transfer the mycelium to 9.2mm agar plates. Important is that the mycelium is directly put on the new plates.

Top down.

- The colonies are incubated to form spores. Conidiation is depending on the organism. However, growth on sporulation media in presence of light should always work
- Wash spores from mycelium by adding 5ml of spore wash solution and scrubbing with a Drigalski spatula over the mycelium.
- The solution containing the spores is diluted and plated on large agar plates to get single colonies.
- Single colonies are picked with a tweezer and transferred to small agar plates, mycelium down.
- Collect spores and use for further experiments.

Troubleshooting spore separating

- It may occur that some colonies do not grow anymore after several rounds selection.

  → normal, pick enough colonies at the beginning

List of figures:

List of tables:

List of strains

Table 28 lists all *E. coli* strains bearing the plasmids for recombinant protein expression in *T. reesei* which were stored in the strain collection of the IMBT. Corresponding vector maps are in Figure 73- 75 were saved on the external hard drive of the Glieder- Group.

Table 28 List of all of all *E. coli* strains bearing the plasmids for the recombinant protein expression in *T. reesei*. List includes the IMBT numbers and selection markers.

#	IMBT Number	Vector name	Host organism	Selection marker *E. coli*	Selection marker *T. reesei*	Date
1	8059	Basic vector without promoter sequences with eGFP and sTomato	*E. coli* XL1 Blue	Amp	Hph	24.11.2017
2	8060	Basic vector with ribosomal promoter "Rn" (CBH1 and natural Kozak-Sequence)	*E. coli* XL1 Blue	Amp	Hph	24.11.2017
3	8061	Basic vector with ribosomal promoter "Rok" (CBH1 Kozak-Sequence)	*E. coli* XL1 Blue	Amp	Hph	24.11.2017
4	8062	Basic vector with ribosomal promoter "Rwl" (natural Kozak-Sequence)	*E. coli* XL1 Blue	Amp	Hph	24.11.2017
5	8063	Basic vector with CBH1/PDC promoter "Pn" (CBH1 and natural Kozak-Sequence)	*E. coli* XL1 Blue	Amp	Hph	24.11.2017
6	8064	Basic vector with CBH1/PDC promoter "Pok" (CBH1 Kozak-Sequence)	*E. coli* XL1 Blue	Amp	Hph	24.11.2017
7	8065	Basic vector with CBH1/PDC promoter "Pwl" (natural Kozak-Sequence)	*E. coli* XL1 Blue	Amp	Hph	24.11.2017

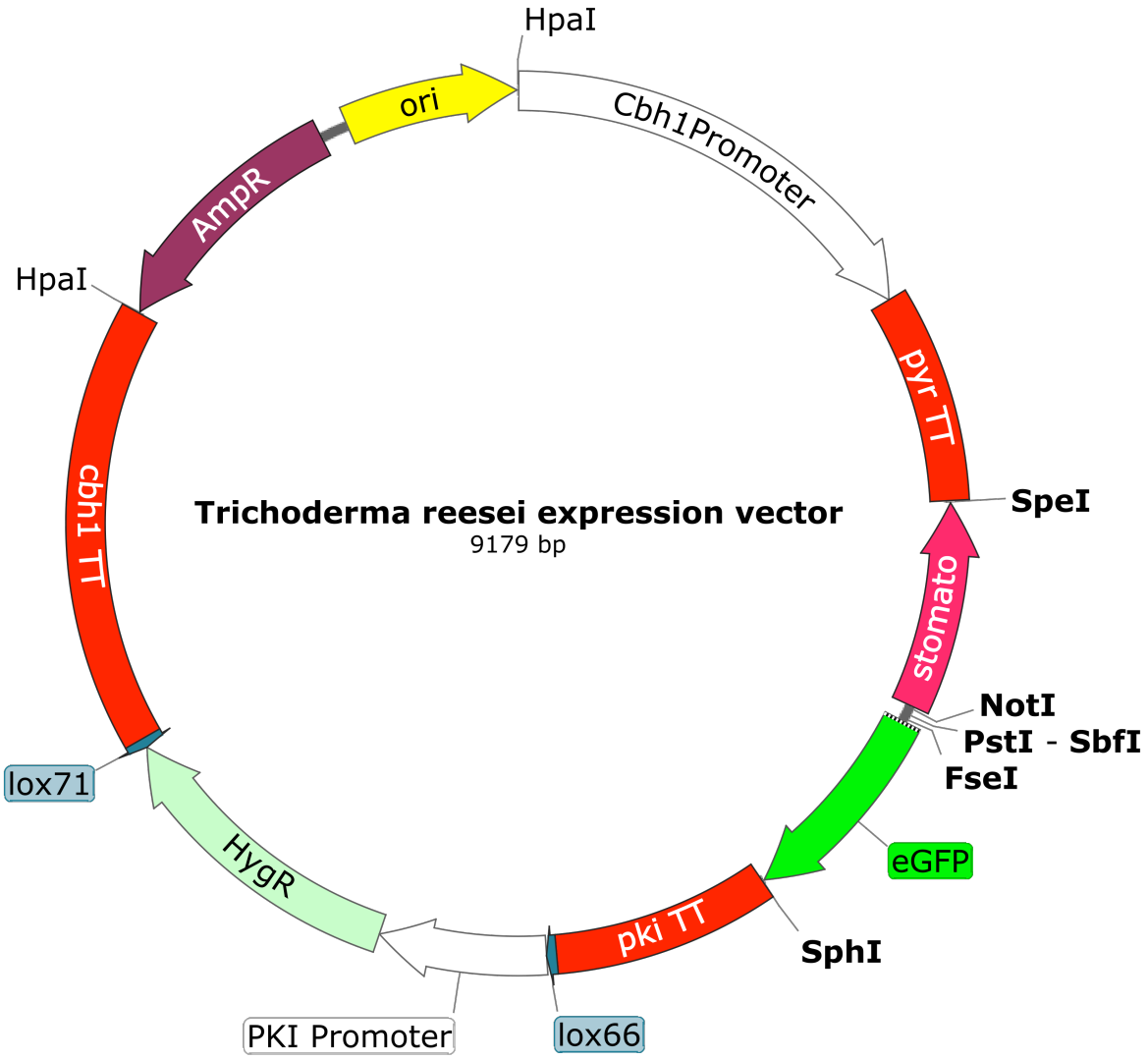


Figure 73 Plasmid map of the basic vector without promoter including the codon optimized genes of *eGFP* and *sTomato*. Map belongs to the plasmid #1 in Table 28. An electronic version of the map can be found on the hard drive of the Glieder-Group.

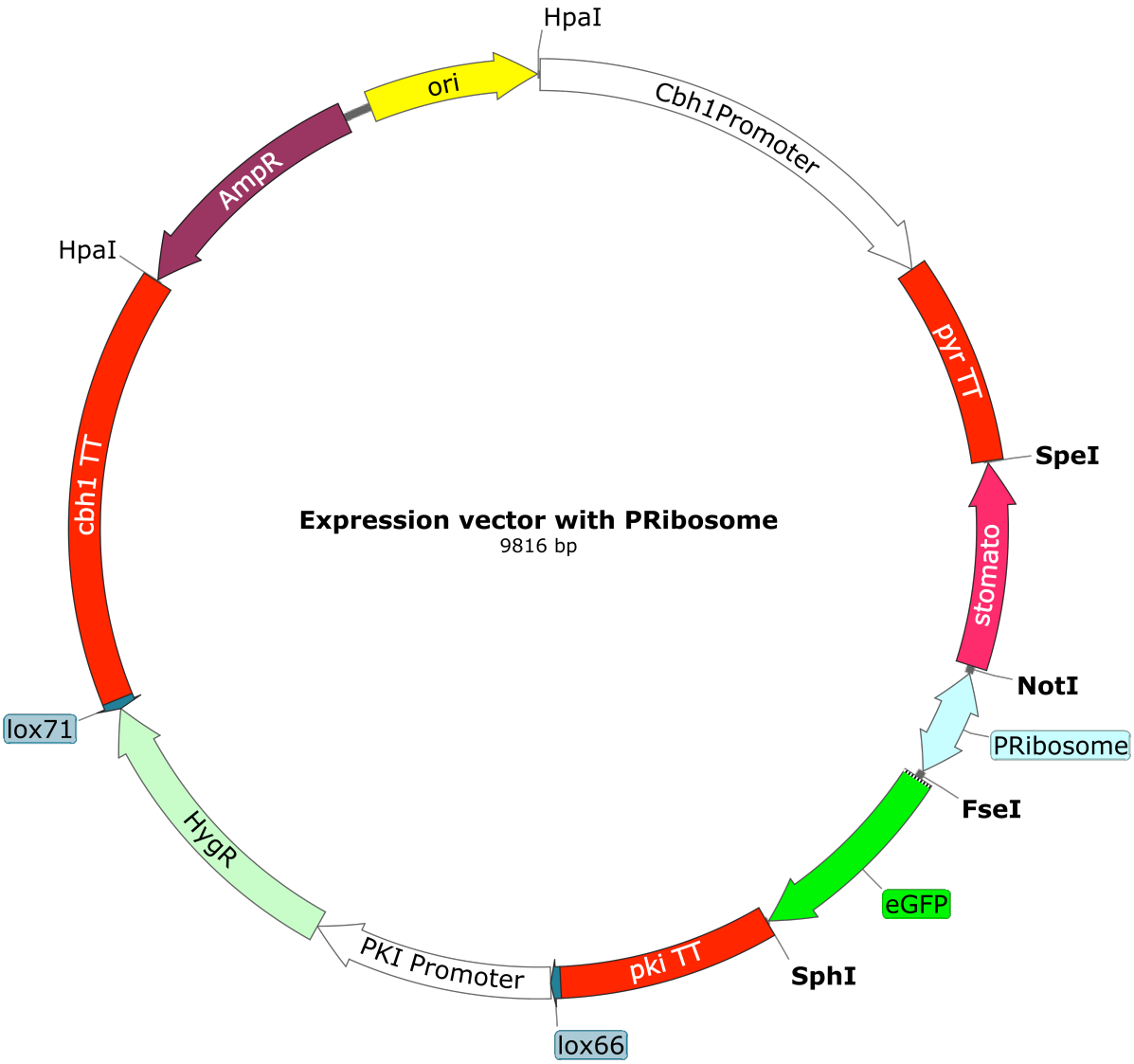


Figure 74 Plasmid map of the basic vector with ribosomal promoter including the codon optimized genes of *eGFP* and *sTomato*. Map belongs to the plasmid #2, 3 and 4 (depending on the Kozak-Sequences) in Table 28. The electronic versions of the maps can be found on the hard drive of the Glieder-Group.

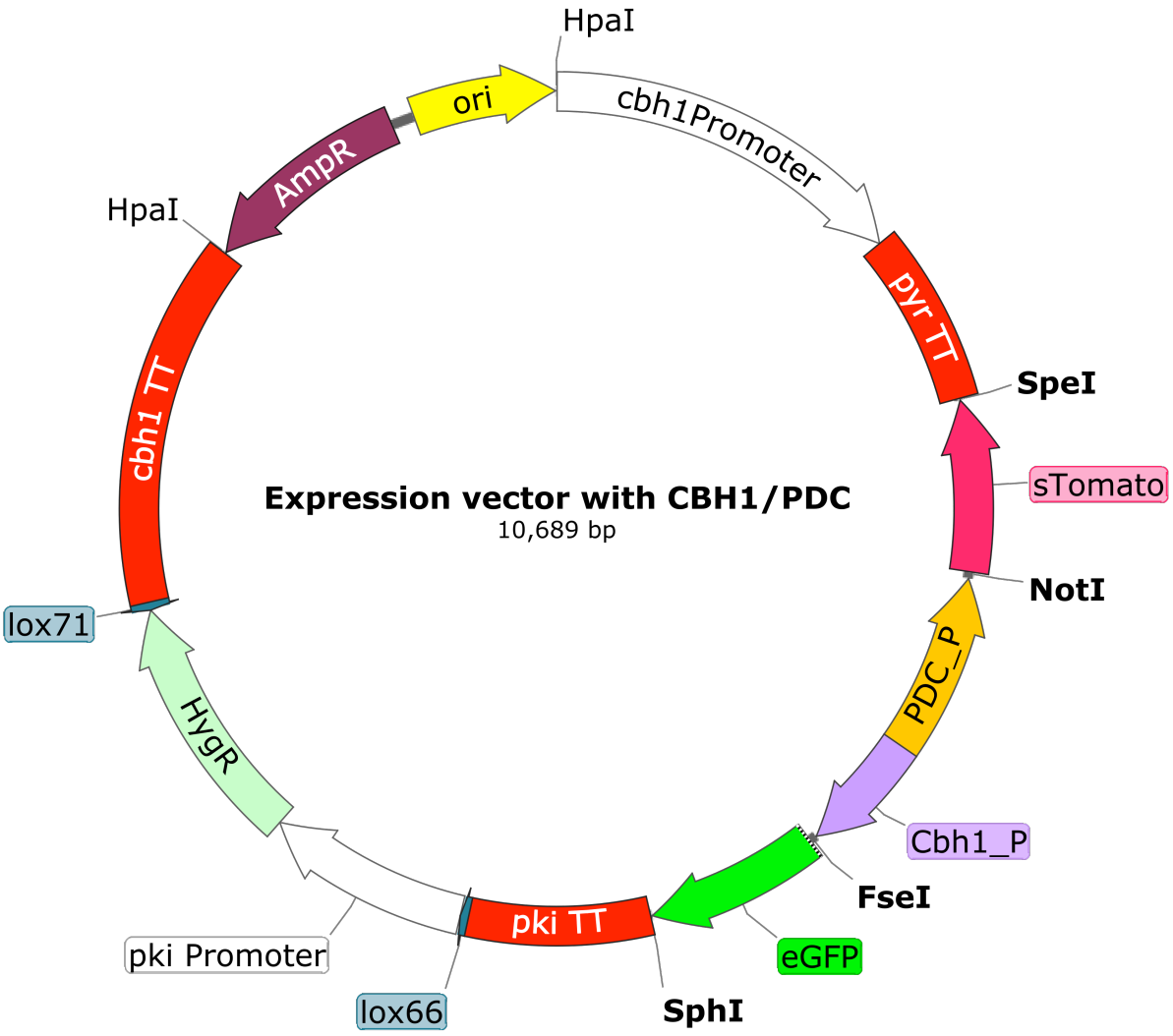


Figure 75 Plasmid map of the basic vector with artificial CBH1/PDC promoter including the codon optimized genes of *eGFP* and *sTomato*. Map belongs to the plasmid #5, 6 and 7 (depending on the Kozak-Sequences) in Table 28. The electronic versions of the maps can be found on the hard drive of the Glieder-Group.

Sequences:

All sequences are listed in 5'→ 3'

Basic expression vector:

aacgcacgtggctcaccgaaaagcaagatgtttgcgatctaacatccaggaacctggatacatccatcatcacgcacgaccactttgatctgctggtaaactcgtattcgccctaaaccg
aagtgcgtggtaaatctacacgtgggccccttttcggtatactgcgtgtgtcttctctaggtgccattcttttcccttcctctagtgttgaattgtttgtgttggagtccgagctgtaactacctctg
aatctctggagaatggtggactaacgactaccgtgcacctgcatcatgtatataatagtgatcctgagaaggggggtttggagcaatgtgggactttgatggtcatcaaacaaagaacga
agacgcctcttttgcaaagttttgtttcggctacggtgaagaactggatacttgttgtgtcttctgtgtatttttgtggcaacaagaggccagagacaatctattcaaacaccaagcttgctcttt
tgagctacaagaacctgtggggtatatatctagagttgtgaagtcggtaatcccgctgtatagtaatacgagtcgcatctaaatactccgaagctgctgcgaacccggagaatcgagatg
tgctggaaagcttctagcgagcggctaaattagcatgaaaggctatgagaaattctggagacggcttgttgaatcatggcgttccattcttcgacaagcaaagcgttccgtcgcagtagc
aggcactcattcccgaaaaaactcggagattcctaagtagcgatggaaccggaataatataataggcaatacattgagttgcctcgacggttgcaatgcaggggtactgagcttggaca
taactgttccgtaccccacctcttctcaaccttggcgtttccctgattcagcgtacccgtacaagtcgtaatcactattaacccagactgaccggacgtgttttgcccttcatttggagaaata
atgtcattgcgatgtgtaatttgcctgcttgaccgactggggctgttcgaagcccgaatgtaggattgttatccgaactctgctcgtagaggcatgttgtgaatctgtgtcgggcaggacac
gcctcgaaggttcacggcaagggaaaccaccgatagcagtgtctagtagcaacctgtaaagccgcaatgcagcatcactggaaaatacaaaccaatggctaaaagtacataagttaat
gcctaaagaagtcatataccagcggctaataattgtacaatcaagtggctaaacgtaccgtaatttgccaacggcttgtggggttgcagaagcaacggcaaagccccacttccccacgt
ttgtttcttcactcagtccaatctcagctggtgatcccccaattgggtcgcttgtttgttccggtgaagtgaaagaagacagaggtaagaatgtctgactcggagcgttttgcatacaaccaa
gggcagtgatggaagacagtgaaatgttgacattcaaggagtattgtcggccatgcctcttctctgtcttcttcttctctctctctctctctctctctctgcgtctcgttcagctcgtaatcgtatatggggg
gaggggaggtcaggtcgctgttccgggggggacggttgattgttgccgtccgtttcgtgtctcggtggggagcgggccggcaaacgggatgtgagagtgagatgtaagaaagagggt
cggaaggaggaaaaggaatgcggaacctcaatgcgaccagcatccgcgttggagcctttgcgcccgattagcattgatatgtcgaattggtggagacgcccggcgatgagagcctc
cgcgttaagtccgggatgccgatttgtaggtaccttaggttaggctgcatagtaacgactccacgatgcatatgatataatggtatatattaatactctgttcgcttggacacgaaacgagta
gcacaaacgatcaatgatcgtgcaatttcctctcctggagacatatcgccttggttgccaccagcgatcaatgggtttctcacagcacaagtactgcttatctgaatattgtagaggggag
aacatgccgccgtgctctgaagatgttacgctgcatacatttagaggtacaggcgccatcacatgtcaatgtcacaaaggatcgacacaaatagccatcaggtaatactcaattgaacaa
ttccgtatatatggaagctgatatcgtcgacaactgccactagttacttgtagagctcgtccatgccgtacaagaagaggtggtgtctgccctcggagcgctcgtactgctcgacgatggt
gtagtcctcgttgtgagaggtgatgtccagcttagtgtcgacgtagtagtagccaggcagctggacgggcttcttggccatgtagatggtcttgaactcgaccaggtagtggccgccgt
ccttgagcttgagggcctggtggatctcgcccttgaggacgccgtcgcgggggtagagtctctcagtgctggcctcccagcccatggtcttcttctgcatgacggggccgtcgggggg
gaagttggtgccgcgcatcttgaccttgtagatgagggtgccgtcctggaggctagagtcctgggtgacagtcacgaggccgccgtcctcgaagttcatgacgcgctcccacttgaag
ccctcggggaaggagagcttcttgtagtcaggaatgtcggcaggatgcttgacgtaggccttagagccgtacatgaactggggagagaggatatcccaagcgaaggggagagggc
cgcccttggtgaccttgagcttggcggtctgggtgccctcgtatgggcggccctcgccctcgccctcgatctcgaactcatggccgttcatgctgccctccatgcggaccttgaagcgc
ataaactccttaatgacttcctcgcccttgctcaccatgatgcgcagtgcggccgcaggacgacctgcaggtgtcgtgggccggccactgcgcatcatggtctcgaagggcgaggag
ctcttcaccggcgtcgtccccatcctggtcgagctggacggcgacgtcaacggccacaagttcagcgtctccggcgagggcgagggcgacgccacctacggcaagctgaccctga
agttcatctgcaccaccggcaagctgcccgtccctggcccaccctcgtcaccaccctgacctacggcgtccagtgcttcagccgctaccccgaccacatgaagcagcacgacttctt
caagagcgccatgcccgagggctacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgccgaggtcaagttcgagggcgacaccctggtcaaccg
catcgagctgaagggcatcgacttcaaggaggacggcaacatcctgggccacaagctggagtacaactacaacagccacaacgtctacatcatggccgacaagcagaagaacgg
catcaaggtcaacttcaagatccgccacaacatcgaggacggcagcgtccagctcgccgaccactaccagcagaacaccccatcggcgacggccccgtcctgctgcccgacaac
cactacctgagcacccagagcgccctgagcaaagaccccaacgagaagcgcgaccacatggtcctgctggagttcgtcaccgccgccgggcatcaccctgggcatggacgagctgt
acaagtaggcatgcacgaaatgggaaaggatacacaaaatggaaatctgagctttttttttttttctttttctttgttatgacggaactcgggggcggcaaagggaacattcgaagaccaaa
acgaacaaagagtgcttgaccacggccatgacggaacaaggcaaaaggccctaggtacggcctgttattaggttatgaggcggttgcatacacgccggtatcattaatgaccgaaga
ataaaactatcacgccttacgatgcatacctatttttatttcattttcaattgcagcagtgtttcagtttatttgtgatttgatttggtttcaacgaatatttgacgatggtgatgactccagacatgtgt
gtccatttccgattcgataaacctctgtatgaagatctgtatatacctgagctccattgctataacttggaatggattttggatcgcttcgttgtgctgctgagtttgatgcgagggtgcggctg
ctgagtttacaggtaggcatggatatatgtgaaagttagcgtcaagggcaagccaaaaggaggcagcctcataacttcatcaggaagaagaagtacaataacaacgagaacatcaac
aagaaataagatagtcattaaatcccgtctatacagatctcaaacgccatcaaacgcatatttgaacaaattcagcagaaaagagaaagaaagagtgaaagagagcaggaaaactcttc
actactacaagttcttaatctcaaatctcttaccgttcgtataatgtatgctatacgaagttatcggagggatgccgtgctttgtcgtggctgctgtggccaatccgggcccttggttggctca
cagagcgttgctgtgagaccatgagctattattgctaggtacagtatagagagaggagagagagagagagagagagagagagggaaaaaaggtgaggttgaagtgagaaaaaaaa
aaaaaaaaaaaaaatccaaccactgacggctgccggctctgccacccccctccctccaccccagaccacctgcacactcagcgcgcagcatcacctaatcttggctcgccttcccgc
agctcaggttgttttttttttctctctccctcgtcgaagccgcccttgttcccttatttatttccctctccatccttgtctgccttggtccatctgcccctttgtctgcatctcttttgcacgcatcgcct
tatcgtcgtctcttttttcactcacgggagcttgacgaagacctgactcgtgagcctcacctgctgatttctctccccccctcccgaccggcttgacttttgtttctcctccagtaccttatcgc
gaagccggaagaaccctcttaaccatgaaaaagcctgaactcaccgccacgtctgtggagaagtttctgatcgaaaagttcgacagcgtctccgacctgatgcagctctcggagggc
gaagaatctcgtgctttcagcttcgatgtaggagggcgtggatatgtcctgcgggtaaatagctgcgccgatggtttctacaaagatcgttatgtttatcggcactttgcatcggccgcgct
cccgattccggaagtgcttgacattggggaattcagcgagagcctgacctattgcatctcccgccgtgcacagggtgtcacgttgcaagacctgcctgaaaccgaactgcccgctgttc
tacagccggtcgcggaggccatggatgcgatcgctgcggccgatcttagccagacgagcgggttcggcccattcggaccgcaaggaatcggtcaatacactacatggcgtgatttca
tatgcgcgattgctgatccccatgtgtatcactggcaaactgtgatggacgacaccgtcagtgcgtccgtcgcgcaggctctcgatgagctgatgctttgggccgaggactgccccgaa
gtccggcacctcgtgcacgcggatttcggctccaacaatgtcctgacggacaatggccgcataacagcggtcattgactggagcgaggcgatgttcggggattcccaatacgaggtc
gccaacatcttcttctggaggccgtggttggcttgtatggagcagcagacgcgctacttcgagcggaggcatccggagcttgcaggatcgccgcggctccgggcgtatatgctccgc
attggtcttgaccaactctatcagagcttggttgacggcaatttcgatgatgcagcttgggcgcagggtcgatgcgacgcaatcgtccgatccggagccgggactgtcgggcgtacac
aaatcgcccgcagaagcgcggccgtctggaccgatggctgtgtagaagtactcgccgatagtggaaaccgacgccccagcactcgtccgagggcaaaggaatagataaacttcgtat
aatgtatgctatacgaacggtagctccgtggcgaaagcctgacgcaccggtagattcttggtgagcccgtatcatgacggcggcgggagctacatggcccgggtgatttatttttttttgt
atctacttctgacccttttcaaatatacggtcaactcatctttcactggagatgcggcctgcttggtattgcgatgttgtcagcttggcaaattgtggctttcgaaaacacaaaacgattcctta
gtagccatgcattttaagataacggaatagaagaaagaggaaattaaaaaaaaaaaaaaaacaaacatcccgttcataacccgtagaatcgccgctcttcgtgtatcccagtaccacgg
caaaggtatttcatgatcgttcaatgttgatattgttcccgccagtatggctccacccccatctccgcgaatctcctcttctcgaacgcggtagtggcgcgccaattggtaatgacccatag
ggagacaaacagcataatagcaacagtggaaattagtggcgcaataattgagaacacagtgagaccatagctggcggcctggaaagcactgttggagaccaacttgtccgttgcgag
gccaacttgcattgctgtcaagacgatgacaacgtagccgaggaccgtcacaagggacgcaaagttgtcgcggatgaggtctccgtagatggcatagccggcaatccgagagtagc
ctctcaacaggtggccttttcgaaaccggtaaaccttgttcagacgtcctagccgcagctcaccgtaccagtatcgaggattgacggcagaatagcagtggctctccaggatttgactgg
acaaaatcttccagtattcccaggtcacagtgtctggcagaagtcccttctcgcgtgcgagtcgaaagtcgctatagtgcgcaatgagagcacagtaggagaataggaacccgcgagc

acattgttcaatctccacatgaattggatgactgctgggcagaatgtgctgcctccaaaatcctgcgtccaacagatactctggcaggggcttcagatgaatgcctctgggcccccagat
aagatgcagctctggattctcggttacgatgatatcgcgagagagcacgagttggtgatggaggggacgaggaggcataggtcggccgcaggcccataaccagtcttgcacagcatt
gatcttcctcacgaggagctcctgatgcagaaactcctccatgttgctgattgggttgagaatttcatcgctcctggatcgtatggttgctggcaagaccctgcttaaccgtgccgtgtcat
ggtcatctctggtggcttcgtcgctggcctgtctttgcaattcgacagcaaatggtggagatctctctatcgtgacagtcatggtagcgatagctaggtgtcgttgcacgcacataggccg
aaatgcgaagtggaaagaatttcccggcgcggaatgaagtctcgtcattttgtactcgtactcgacacctccaccgaagtgttgttaacttaccaatgcttaatcagtgaggcacctatctc
agcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccctagtgctgcaatgataccgcgagacccacgctca
ccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaag
tagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatg
atcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatg
ccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacata
gcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagc
atcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatgtttagtcctccttacaccttg
tcgtattatacgagccggaagtataaagtaacaacactctagacgcgtcccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaa
accaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagttag
gccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagtt
accggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgc
cacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgt
cgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaagtt

Bidirectional promoters

$P_{CBH1/PDC}$

gcggccgcaggacgacctgattgtgctgtagctgcgctgctttgatcgttttgaggtgcttcttggattgtggaacgaagaaggacgagagcttgatgcggttatttatactggtgacgac
actctgaggcgcgaggagatgtggatcaacgatgatgcaaatgatgttgctcgtataaggtggttgatgaagaagctccttgtgatggagacttgcaggattccagacgtgtttttgagaa
ggtatttacttttgcaggaggaggggcggaggattggagacgtgagccaacacactgacagggcaatgaatgaggatatatagatgtttttttatctacctagacttcgaagtgttgaatcat
ccatcggtctcggtttcaaccatcaacagctcggacaaactcaaatgggcagcggcaccggagacggcaggcaggagccccgccagttgcagctgctcgtgccatcggcaagcat
gagctcactttttggacgcgccgctgtcatgatagtgatgtaatggggccatttctgcgcgggtgggcagccgctaatgcagctgcttctacagcgctcgacgatagtggacggctgctc
ctccagtgatactagatagcctcaggtccactatagctactagagagttggaacctatccatcgtaacatggatcccgggtggcttcatgcacggtacatacagcagagccggcaatcc
gaaggagcggccgattcgctcgtcggacaattaacgaccacctctgacttcacttagccagggatgcttgagtgtatcgtgtaaggaggtttgtctgccgatacgacgaatactgtatag
tcacttctgatgaagtggtccatattgaaatgtaagtcggcactgaacaggcaaaagattgagttgaaactgcctaagatctcgggccctcgggccttcggcctttgggtgtacatgtttgt
gctccgggcaaatgcaaagtgtggtaggatcgaacacactgctgcctttaccaagcagctgagggtatgtgataggcaaatgttcaggggccactgcatggtttcgaatagaaagaga
agcttagccaagaacaatagccgataaagatagcctcattaaacggaatgagctagtaggcaaagtcagcgaatgtgtatatataaaggttcgaggtccgtgcctccctcatgctctccc
catctactcatcaactcagatcctccaggagacttgtacaccatcttttgaggcacagaaacccaatagtcaaccgcggactgcgcatctaggtgtcgtgggccggcc

$P_{Ribosome}$

gcggccgcaggacgacctcgcgaccgattaattgccgtcgatattcgttcgaggattgacgaggatggaggtgaagctttgctaatttggcgtgggtgaaattacgcacggcgtgtgg
aagcgaatgtgcttagataatccacggatttttttttcccgccgcgacgatagggcaggattgtgggttttgggggggaacggcacatgttgcccggcttagggctcgtggctagaaggtc
acgtgcatggtttccagacgccaaatttcctatccacacaaccttgacgattcctcgagatcacccgcgcaaaagttcaccagcagccaccactttaggaagcaagtcaagtgagtatct
catatactcaattcttacccgtccgagccataacctctgacaattctacagaggtgtcgtgggccggcc

$P_{Histone}$

gcggccgcaggacgacctttgaaaataacgttgatttcaaatagatttcagagactgaaaaggttgagtgttgatgatgcaagatgtcagaagggcggactggcagcgggtgatgatg
ggaaaagtgaatgggtgcgggaacgacgttcttggggacgagctgggcttttgtagctggcacgcgttggcggccggtggattactgatttggcctggagcaggaaggggtggcgc
gagcgaggccaggcgcgtgcacgccctcgcgggtacgggccttgcgctgcgaaaatgtgcggtaaggctaggtaccgctcgctggaccactcagaattgagcgccacaggcaag
agggggggggggggagggcaaaaacatgagagggaaaaaatgtggtggaattcactgctgccgctagcgctggtatttgcaatgtacccagaaaatgctactgcttgtgttgcagct
ctgctcacgctaggcatttctcggcaaggcgcaacactgccttttttttttttttctcacgcatttcttctggtgactgggggcggtggtccaggtgaattgacggcctgcaccgaatgatgt
gatgcgattgccttgcccgctccatagcttctagtcgcgtcgagctcttttgttgttttttctcccgccttttttttgtacctctgctcggtattactagccttttcctctcgcaggcagagagcc
acgtgattggctgacgccaagtctctctttgttgcatcattttctccccacgagcccgccttgcagctgagggccaatgtacctttttctgccgtgatctcagcggcacatgtacctcccgcc
ttgcgtctgcgcgctactgatgttttgacggagcgcgcctggatactagcctcatcttgatgcttcggcctgcttaccctgtaatgtgaatagaactgctttgtgcgcaattgtagcttcagtc
tactacgcatcatatgaaattctgtttattcgcttctctggtctatttggattatatgtatgctcgtgatgtgatagtgagacggctttggtgaggtaaatactcttaggccaaatgctcaccagtg
gtgttgttggataacctaaaggcaacgaagggattccatagtaactgaacccatctccttacgtataacccaccatttggagcagacattgcatggacaaagagaaaccgggcacagac
cagttgcctgatcaagatcaaagtcacgatgagagtccgcaaaatgatatccatctgtataacggggcatatctgcgcgtatactgcaagcacatgaagcggccgggttcacgcaagat
ctcatggcatcaagattgacccagtctggcccagattggacccccatttagatctcgttattggttccgtcgtcttcgagactgcaatcgtaattttcactgttgtctaccatctaaacgtggagt
tgccgctgtcttctgtttacagtacaatacgctaccataccagcacgcgtgtcaccaaggaagagctttgctgccgcgtaggaggagtcaacgagtgccaaagggcccagccaggag
tcctggtgccatcatcagatggcaaatcactcccagccatcgggggactcgtatcgtctaacaaaggaatttgcgcaacactcgtgcatgcgccgttgaagctcatcccagccgtgtca
atacatggtgcagccctgagacggggcggagatcacgacgagagcctgctgagggccgcgatatcgagcatgtaccgtcaaggtacccggagcgagcattagccgcatcattctc
gggacccttaccttctggaggtactccctgcgtgtacctttttttgcgcacgccccccgcaccctcaaattgtcgcaattgttttgctggaacacgtccatttccccaacaggcctagcgc
cgcaacagctgcgagagggagggtgattgctgatctcgcgcggcgagcaccaaaaagtacttaactttggagccccccgccaaatccaaacctctgcatttacactgtttctctcaact
ctccacgctgcacgattaccattctgacatttgccgcatctcgcgcatctcgcatcaacactttacgttacttttccacgttaagcaaaaacccccaacttcaaaggtgtcgtgggccggc
c

Fluorescence proteins

Codon optimized *eGFP*:

atggtctcgaagggcgaggagctcttcaccggcgtcgtccccatcctggtcgagctggacggcgacgtcaacggccacaagttcagcgtctccggcgagggcgagggcgacgcc
acctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtccctggcccaccctcgtcaccaccctgacctacggcgtccagtgcttcagccgctaccccgac
cacatgaagcagcacgacttcttcaagagcgccatgcccgagggctacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgccgaggtcaagttcga
gggcgacaccctggtcaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctgggccacaagctggagtacaactacaacagccacaacgtctacatcatg
gccgacaagcagaagaacggcatcaaggtsaacttcaagatccgccacaacatcgaggacggcagcgtccagctcgccgaccactaccagcagaacaccccccatcggcgacgg
ccccgtcctgctgcccgacaaccactacctgagcacccagagcgccctgagcaaagaccccaacgagaagcgcgaccacatggtcctgctggagttcgtcaccgccgccggcatc
accctgggcatggacgagctgtacaagta

Codon optimized *sTomato*:

atggtgagcaagggcgaggaagtcattaaggagtttatgcgcttcaaggtccgcatggagggcagcatgaacggccatgagttcgagatcgagggcgagggcgagggccgccca
tacgagggcacccagaccgccaagctcaaggtcaccaagggcggccctctcccttcgcttgggatatcctctctcccagttcatgtacggctctaaggcctacgtcaagcatcctgc
cgacattcctgactacaagaagctctccttccccgagggcttcaagtgggagcgcgtcatgaacttcgaggacggcggcctcgtgactgtcacccaggactctagcctccaggacgg
caccctcatctacaaggtcaagatgcgcggcaccaacttcccccccgacggccccgtcatgcagaagaagaccatgggctgggaggccagcactgagagactctacccccgcgac
ggcgtcctcaagggcgagatccaccaggccctcaagctcaaggacggcggccactacctggtcgagttcaagaccatctacatggccaagaagcccgtccagctgcctggctacta
ctacgtcgacactaagctggacatcacctctcacaacgaggactacaccatcgtcgagcagtacgagcgctccgagggcagacaccacctcttcttgtacggcatggacgagctcta
caagta

CBH1 secretion signal

atgtatcggaagttggccgtcatctcggccttcttggccacagctcgtgct