New tools and methods for gene expression employing fission yeast

Master's Thesis in Biochemistry and Molecular Biomedicine

by Anna-Maria Hatzl, BSc

Graz, 2014

Deutsche Fassung:

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.

Graz, Mai 2014 .....

Englische Fassung:

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declaredsources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Graz, Mai 2014 .....

Danksagung

An erster Stelle möchte ich mich bei meiner Familie bedanken, die mich immer bedingungslos unterstützt hat. Im Besonderen danke ich meinen Eltern, Petra und Rudolf, für die finanzielle und emotionale Unterstützung in den letzten Jahren. Großer Dank gebührt auch Michael, der mich in der Entstehungszeit dieser Arbeit immer unterstützt und motiviert hat.

Des Weiteren möchte ich mich bei Prof. Anton Glieder für die Betreuung meiner Masterarbeit und für die Möglichkeit bedanken, in diesem Projekt mitgewirkt haben zu dürfen. Ein großer Dank geht an die gesamte Glieder-Gruppe für die vielen interessanten Gespräche und das tolle Arbeitsklima. Ganz besonders danke ich dabei Andrea Mellitzer für ihre tolle Betreuung im Labor und ihre wertvolle Unterstützung im letzten Jahr.

Für die Finanzierung dieser Masterarbeit möchte ich mich zum einen bei der Firma Clariant und zum anderen bei den Unterstützern des Chem21 Projektes, nämlich bei der Innovative Medicines Initiative, der Europäischen Union und bei den EFPIA Firmen bedanken.

Schlussendlich danke ich all meinen Freunden und Studienkollegen für die schöne Zeit während meines Studiums.

Abstract

The fission yeast *Schizosaccharomyces pombe* is an interesting expression host for heterologous proteins derived from higher eukaryotes. Therefore, two strategies for the expression of heterologous proteins in the fission yeast *S. pombe* had been developed. Both strategies are based on the pGAZ2 vector (1), which is suitable for episomal expression. Due to stability problems and the moderate expression, new integrative vectors and approaches were chosen. The first strategy is based on integrative PCR cassettes, which have 80 basepair homologous extensions to the targeted gene locus and lead to a gene deletion. Three different nutritional integration markers and integration sites and two antibiotic resistance genes were used to test which parts are most suitable for the PCR cassettes. As second strategy a new bifunctional vector capable of chromosomal integration but also of episomal expression was created. The bifunctional vector can easily be changed to an integration cassette by restriction digestion. Here, the homologous ends were at least 220 basepairs long and the right integration leads to an insertion of the expression cassettes in the *leu1* or the *leu2* locus and thereby creates an auxotrophy. A reporter gene, enhanced green fluorescent protein (eGFP), was used to investigate the different expression constructs and different transformation protocols were compared.

Finally the new tools were evaluated and used for the expression of new genes for enzymes of so far unknown function which might be involved in lignocellulose degradation by fungi. Due to the unknown enzymatic function of these proteins, the cultivation supernatant of the transformants was used for a dot blot screening to identify secreted proteins. The detection of the proteins was based on an antibody binding to a C-terminal his-tag, which was added to the genes. Eight targets were secreted by *S. pombe* and could be detected in the supernatant. In order to improve heterologous protein expression and to make an expression host comparison interesting target were redesigned and optimized for high expression in *S. pombe* as well as *P. pastoris*.

Zusammenfassung

Die Spalthefe *Schizosaccharomyces pombe* eignet sich gut zur Expression von heterologen aus höher entwickelten Eukaryonten stammenden Proteinen, deren Produktion sich in anderen Wirten als schwierig erweist. Deshalb wurden zwei Expressionsstrategien für *S. Pombe* entwickelt und getestet. Beide basieren auf dem pGAZ2 Vektor (1), welcher ausschließlich zur episomalen Expression von Proteinen verwendet werden kann. Aufgrund von Stabilitätsproblemen und der moderaten Expression dieses Vektors wurden integrierbare Systeme gewählt. Die erste Strategie basiert auf PCR Kassetten, die zu einer Gen Deletion führen. Dabei wurden 80 Basenpaar lange Sequenzen, die homolog zum gewünschten Integrationslocus sind, an beiden Seiten der Expressionskassette hinzugefügt. Es wurden drei verschiedene Integrationsloci getestet. Die zweite Strategie basiert auf bifunktionellen Vektoren, die sowohl zur episomalen Expression als auch als integrierbare Expressionskassetten verwendet werden können. Die homologen Sequenzen sind in diesem Fall mehr als 300 Basenpaare lang. Eine richtige Integration der Kassetten führt zur Insertion im gewünschten Genlocus (*leu1, leu2*) und zerstört dabei dessen Funktion. Die ausgeknockten Gene sind in beiden Strategien essentiell und involviert in die Aminosäuresynthese. Das Reportergen eGFP (enhanced green fluorescent protein) wurde verwendet, um die Expression der zwei integrierbaren Systeme als auch die episomale Expression zu evaluieren. Zudem wurden verschiedene Transformationsmethoden verglichen

Abschließend wurden diese neuen Tools für die Expression von neuen Proteinen eingesetzt. Da diese neuen Proteine möglicherweise eine noch unbekannte enzymatische Funktion im Lignocelluloseabbau haben, hat man die Detektion mittels Dot Blot Screening durchgeführt. Dafür wurde ein C-terminaler His –tag an die Proteine fusioniert. Acht Targets wurden von *Schizosaccharomyces pombe* sekretiert und konnten im Kulturüberstand detektiert werden. Um die heterologe Proteinexpression zu optimieren wurden zwei unterschiedliche Wirte verwendet, *Pichia pastoris* und *Schizosaccharomyces pombe.* Dafür wurden sowohl die Codons an den jeweiligen Wirt angepasst als auch eine Veränderung der Sequenzen vorgenommen.

Contents

1. Introduction

1.1. One yeast is not like another: Fission vs. budding yeast

In contrast to higher eukaryotic cells, like plant or mammalian cell lines, yeasts are easy to handle from a scientific point of view. Therefore, yeasts gained huge popularity in basic research as well as in industrial application. In general, yeasts appear as unicellular organisms – in rare cases they show dimorphic appearance – they belong to the kingdom of fungi and therefore possess a fungal-type cell-wall. From the phylogentic point of view yeasts do not form a homogenous group; they can be assigned to basidiomycetes as well as ascomycetes (2). Yeasts mainly reproduce asexually, either by budding like the baker's yeast *Saccharomyces cerevisiae* or by fission like *Schizosaccharomyces pombe* (3). Budding is defined as the process of cell division, where a smaller sized daughter cell is formed on the surface of the parent cell (4). In Figure 1A budding of a *S. cerevisiae* cell culture is shown. After dividing a so called bud scar remains on the surface of the parent cell.

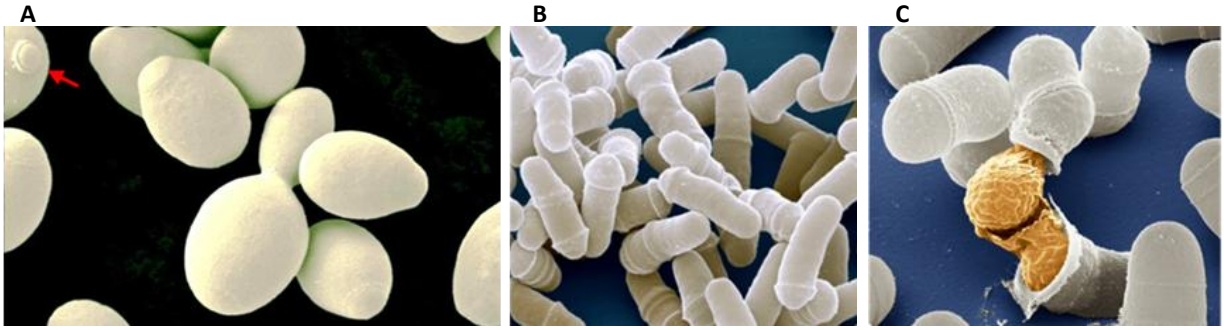


Figure 1: Scanning electron micrograph of budding yeast *S. cerevisiae* (A) and fission yeast S. pombe (B and C). In the middle of picture A there is a *S. cerevisiae* cell which is just about to bud into a small daughter cell and a bigger parental cell. In the upper left corner of the picture a bud scar is indicated on the surface of a cell with a red arrow. Picture B shows a snapshot of a propagating *S. pombe* culture whereas C shows the fission yeast as ascospores. Here, the ascus in the middle of the picture is just about to release its four spores. (5; 6; 7)

Fission, on the other hand, leads to the development of two identical daughter cells (Figure 1B). Therefore, yeasts propagating in this way are called fission yeasts. In principle, fission is a process more similar to the proliferation of higher eukaryotic cells (8). In 1893, the first fission yeast was discovered by Paul Lindner in an east-African beer called "Pombe". He named this fermenting organism after that beverage – *Schizosaccharomyces pombe*. *S. octasporus*, *S. japonicus* and the recently found *S. cryophilus* complete the group of fission yeasts known today (9; 2). Fission yeasts grow preferentially as haploid cells, but under nitrogen depleting conditions two of them with opposite mating types fuse and a zygote

(diploid) is formed. The zygote then undergoes meiosis in order to form four spores, which are able to outlast harsh times (Figure 1C). *S. japonicas* and *S. ocatsporus* develop eight spores by meiosis II, a process very similar to mitosis (9). In contrast to fission yeasts, the budding yeast *S. cerevisiae* propagates rather in diploid as in haploid form. However, under nutrient depletion *S. cerevisiae* also undergoes sporulation to ensure its survival (10).

Due to the long history of *S. cerevisiae* and *S. pombe* as fermenting agents, it is no surprise that their genomes are among those few, which are completely sequenced. In 2002 Wood *et al*. report the annotation of the *S. pombe* genome. They found around 5000 protein-coding genes distributed on three chromosomes, 43 % of these genes have introns. The completion of the *S. cerevisiae* genome sequencing project was announced by Gouffeau *et al*. in 1996. The 6000 open reading frames (ORFs) from budding yeast are distributed on 16 chromosomes. In contrast to *S. pombe* only 4 % of these ORFs have introns. Moreover, the region upstream of a protein-coding gene in *S. pombe* is longer than in *S. cerevisiae*, indicating more sophisticated control mechanisms in the fission yeast. (11; 12)

Overall, fission yeasts seem to have more in common with metazoans than with their yeast relatives. With *S. pombe* as model organism Leland Hartwell, Tim Hunt and Paul Nurse (13) studied the cell cycle and even won the Nobel Prize in physiology and medicine for their work (14). Thereby, one breakthrough was the finding that the *S. pombe* homologous cyclin-dependent kinase 1 (CDK1, *cdc2*) – a protein important in cell cycle progression – was successfully replaced with the human CDK1 in *S. pombe* (13; 15). Chromosome structure, splicing machinery and signal pathway (e.g. KDEL sequence is identical to mammals) are further components that fission yeasts share with metazoans (16; 12). In Figure 2 the evolutionary distance of *S. cerevisiae* and *S. pombe* is shown. These two yeasts diverged 573 million years ago; a broad distance is separating them and *S. pombe* seems to be two steps ahead on the evolutionary ladder (17).

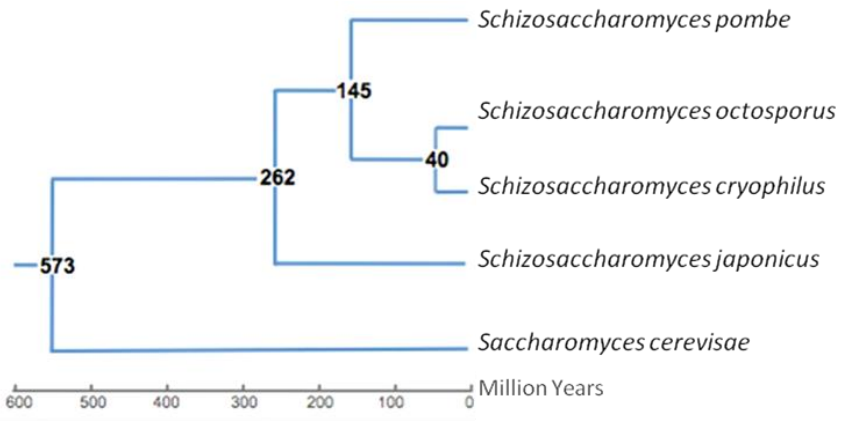


Figure 2: Evolutionary tree and distance separating the four fission yeasts (*Schizosaccharomyces*) and the budding yeast *S. cerevisiae* (Rhind, et al. 2011 modified by Allshire lab)

1.2. Expression hosts for heterologous eukaryotic protein production

Unicellular organisms, prokaryotes as well as eukaryotes, are the preferred hosts for heterologous protein production. Their relatively fast growth, the easy manipulation of their genomes and the possibility to work in big scales favours unicellular organisms over multicellular cell lines as protein producers. Bacteria are by far the fastest producers, but they lack the eukaryotic post-translational processing machinery (18). Moreover, the gram-negative bacterium *Escherichia coli* has problems with secreting proteins; they stuck in the periplasmic space. Gram-positive bacteria, like *Bacillus subtilus*, are indeed able to secrete proteins directly into the medium, but cannot perform typical eukaryotic modifications after translation (19).

Yeasts – as unicellular eukaryotes – on the other hand are equipped with the tools for post-translational modifications like glycosylation, protein folding, phosphorylation and secretion of biofunctional proteins. *Saccharomyces cerevisiae*, *Pichia pastoris* and *Schizosaccharomyces pombe* are only a few examples for widely used yeasts in industry (20). Humans have been using *S. cerevisiae* since ages for making bread fluffier and alcohol-containing beverages. Therefore, it is hardly surprising that *S. cerevisiae* was the first organism used for foreign protein production and secretion (21; 22).

Turning to *P. pastoris*, a yeast which gained enormous popularity as expression host over the last three decades. *P. pastoris* is among those few yeasts, which are capable of growing on methanol as their only carbon source. Behind this feature of methylotrophic yeasts stands a row of tightly regulated genes. For example, the promoter of the alcohol oxidase 1 gene (*AOX1*) is used in many applications to express heterologous proteins in *P. pastoris*. The *AOX1* promoter is highly repressed on non-limiting glucose concentrations and very strong activated when glucose is depleted and methanol is added into the medium (23). Furthermore, *P. pastoris* can be grown to very high cell densities (more than 500 OD_{600} units/ml in fermenters) and secretion of foreign proteins is also very efficient (several g/l are possible) (24).

The fission yeast *S. pombe* is more similar to higher eukaryotes and therefore very interesting for the production of mammalian proteins and proteins derived from higher eukaryotes. Although the protein yield is smaller than in other expression system – e.g. *P. pastoris* – the gained protein quality is very good (18). *S. pombe* is not as often used as

S. cerevisiae or *P. pastoris*, but there are already several applications available for this yeast: For instance, the high-yield expression of human lipocortin was successfully accomplished – 150 mg pure protein could be obtained per 10 g wet cell weight (25) – or the expression of cytochrome P450 enzymes in *S. pombe* as it was done by a number of work groups (26). In addition fission yeast was also used to express heterologous pathways, for e.g. the pathway for the biosynthesis of vanillin (27).

In summary yeasts represent a good choice for heterologous eukaryotic protein expression: They can reach high protein yields, grow to high cell densities and they can handle post-translational modifications.

1.3. Construction of *S. pombe* expression strains

The way from a heterologous gene to an expressed and maybe also secreted biofunctional protein by *S. pombe* is complex. First, an expression system has to be chosen, which can be either a vector (episomal or integrated) or an expression cassette obtained by polymerase chain reaction (PCR). The most important components for the expression system are a selection marker, an origin of replication (for episomal expression) and a promoter. For the integration into a defined locus an integration sequence with at least 60 homologous nucleotides to the genome are needed.

1.3.1. Selection markers

In *S. pombe* the most common selection markers are prototrophic markers. Therefore, the used *S. pombe* strain also needs the corresponding genetic background. One of the most often used prototrophic markers is *ura4*, which encodes an orotidine 5'-phosphate decarboxylase (ODCase). Positive selection is obtained by growing transformed cells on medium without uracil. When 5-fluoroorotic acid (5-FOA) is added to the medium, the ODCase converts the 5-FOA into a toxic compound. This causes a negative selection, because only *ur4*$^{-}$ mutants, without functioning ODCase, can survive. (28)

Another very popular prototrophic marker is the *leu1* marker. The *LEU2* gene from *S. cerevisiae* is able to complement *leu1*$^{-}$ mutants when multi copies are present in the *S. pombe* cell. However, for single copy episomes as well as for integration and disruption constructs *leu1* should be used instead of *LEU2* (29).

Ade6, on the other hand, is not very often used as selection marker due to its big size and the many restriction sites inside its gene sequence. Hence, the selection with an a*de6* knock out strain is more popular. This gene dysfunction results in pink or red colonies, similar to *ade2*$^{-}$ mutants in *S. cerevisiae*. The reason for the colour is the accumulation of a red pigment which cannot be further metabolized. Positive selection of the *ade6*$^{-}$ mutants can be achieved by adding a *sup3-5* gene to the expression construct. The *sup3-5* gene encodes a suppressor tRNA, which reads nonsense codons and suppresses mutations in the genome. Although the suppressor tRNA encoded by *sup3-5* is slightly toxic in high copy numbers, a nice prescreen by the coloration of the colonies can be done: Red colonies are indicating the absence of the *sup3-5* gene whereas pink colonies are formed when *sup3-5* gene is episomally expressed. The pink coloration, which is just a fainter red, is appearing due to

plasmid instability. When *sup3-5* gene is integrated into the genome, the positive colonies appear white. (30)

There are many other nutritional markers, like *his3*, *his7* and *arg3*, which had been used for selection in *S. pombe*. In general, drug markers are not as popular as nutritional ones, but the aminoglycoside antibiotics G418, hygromycin, chloramphenicol, phleomycine and bleomycine had been used for selection in fission yeast as well (31). The combination of minimal medium with Zeocin™ has been tested in *S. pombe* and was not recommended (32; 33), because high ionic strength, acidity or basicity inhibit the antibiotic(34).

1.3.2. Origins of replications

While the research on *S. pombe* was still in its early stages, there were already a huge set of working techniques and material available for *S. cerevisiae*. Hence, it is no surprise that the first plasmids used in *S. pombe* had the *2µ* origin of replication (ori), which derives from *S. cerevisiae*. Also epsiome studies without origin of replication were made, but most of the transformants were highly unstable or the construct integrated randomly into the genome (35). Later the autonomously replicating sequence (*ars1*) in *S. pombe* was identified. Compared to the *2µ* ori the a*rs1* sequence is far better regarding the stability of the episomes. Around 15 to 80 copies can be achieved by transforming *ars1* plasmids into *S. pombe*, but they are mitotically unstable. If *ars1* is combined with the stabilizing sequence (*stb* sequence) mitotically and meiotically stable plasmids are formed with a copy number of around 80 and a lost rate per generation without selection of 13 % (31).

In *S. cerevisiae* their relative simple and small centromers – around 125 base pairs – provide an additional element, which can be used for extrachromosomal replication. By combining *ARS/CEN* in *S. cerevisiae*, very stable single copy epsiomes can be obtained. In fission yeasts, on the other hand, the size and repeated DNA sequences within their large centromers – many kilo base pairs – are not suitable for the construction of such centromeric plasmids (36).

1.3.3. Promoters

The homologous *nmt1* promoter (no message in thiamine) is by far the most often used and as well the strongest known promoter in *S. pombe*. The *nmt1* gene is highly transcribed when the thiamine concentration drops to a level lower than 0.5 µM. Media, containing yeast extract or YNB, are not suitable for protein expression driven by *nmt1* promoter, because the thiamine concentration in these components is sufficient to repress the promoter. A thiamine concentration of 15 µM (vitamine B1, Sigma #4625) in minimal medium shuts off expression within 3 hours. A disadvantage of the *nmt1* promoter is its long induction time. Protein expression starts not before 16 to 20 hours after the removal of thiamine. Moreover, this requires an additional washing step of the cell, which is followed by continued growth in minimal medium. The relative long time period of around 18 hours is apparently necessary to deplete the cell's own thiamine reserves. (37; 38) In order to have a more extended set of inducible pREP expression vectors, the TATA box of the *nmt1* promoter was mutated. Thus, the *nmt41* and *nmt81* promoter – an intermediate and a weak version of *nmt1* promoter – were generated. (39)

Another strong promoter in *S. pombe* is the constitutive promoter of the alcohol dehydrogenase gene $adh1^+$ (40). More recent expression systems are based on the $urg1^+$ promoter (inducible by uracil), the $ctr4^+$ promoter (copper repressible) or the constitutive $tif51^+$ promoter (41; 42; 43)

There are also some mammalian promoters which work in *S. pombe* and have been used for heterologous protein expression. This is not surprising, because the fission yeast's TATA box is 25 to 30 bp upstream of the transcription start – that is also the case in mammalian promoters. The promoter of the simian virus 40 (SV40) small-T antigen transcript was the first mammalian promoter used in *S. pombe* that worked well (44). However, the later used human cytomegalovirus and gomadotropin α promoter show an 11- and 9-fold higher activity than the SV40 promoter (45).

From the cauliflower mosaic virus 35S (CaMV 35S) – a plant virus – derives a promoter that normally functions as a constitutive promoter in fission yeast. But Faryar and Gatz (1991) presented an expression system, where CaMV 35S promoter can be induced by tetracycline. They constitutively coexpressed a tet repressor with the $adh1^+$ promoter. This tet repressor binds near to TATA box of the CaMV 35S promoter and thereby represses the expression of

the reporter gene. Its half-maximal induction is achieved 5 hours after tetracycline addition. (46)

A study from Forsburg in 1993 compares different episomal expression system for *S. pombe*. The conclusion of this study is presented in Figure 3. Table 1 provides an overview of available expression systems for *S. pombe*.

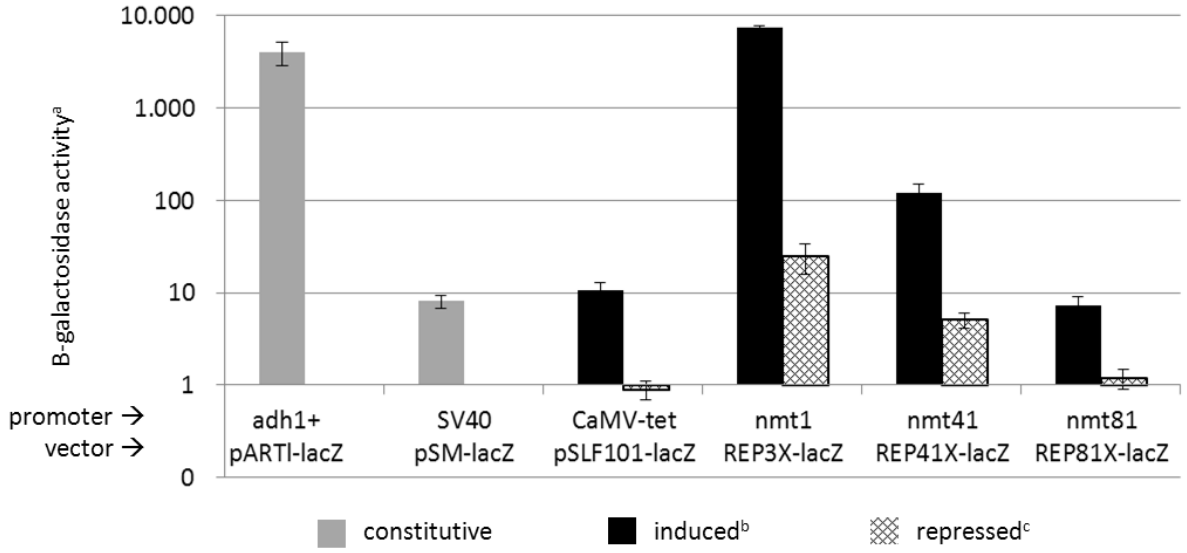


Figure 3: Comparison of different *S. pombe* expression systems. [a]Units = $(100 \cdot OD_{420}) \cdot (min \cdot ml \cdot OD595)^{-1}$, y-axis is in logarithmic scale. [b]Induced conditions: For pSLF101: EMM + 10µg/mL tetracycline; for REP3X, 41X and 81X: EMM. [c]Repressed conditions: For pSLF101: EMM; for REP3X, 41X and 81X: EMM + 5 µg/ml thiamine. Based on Forsburg (47).

Table 1: Promoters and corresponding expression sytems for heterologous protein production in *S. pombe*.

Promoter	Origin	Regulation	Protein	Comments	Vector	Reference
nmt1	homologous	repressible (thiamine)	N-myristoyltransferase	strong promoter	REP1/2/3	(Maundrell 1990)
nmt41				weaker	REP41/42	(Basi, Schmid and Maundrell 1993)
nmt81				weakest	REP81/82	
ctr4	homologous	repressible (copper)	copper transporter	rapid and tight inductions	p*ctr4*$^{+}$*-X* vector series	(Bellemare, et al. 2001)
inv1	homologous	repressible (glucose)	invertase	induction upon carbon source change from glucose to sucrose, rapid and tight	pINV1-*spc1*$^{+}$-HA6his pINV1-*spc1*$^{+}$-12myc pINV1-GST pREP1-*spc1*$^{+}$-GFP (*LEU2*)	(Iacovoni, Russell and Gaits 1999)
pho1	homologous	repressible (adenine)	acid phosphotases	constitutive in adenine auxotrophic strains	-	(Schweingruber, et al. 1992)
fbp1	homologous	repressible (glucose)	fructose-l,6-bisphosphatase	high-level of expression, but maximal activity if cells are joining stationary phase	pCHY21 (*URA3*)	(Hoffman and Winston 1991)
urg1	homologous	inducible (uracil)	GTP cyclohydrolase II family of enzymes involved in riboflavin biosynthesis (gene belongs to "uracil-regulable genes")	quick induction (2 h), but does not work well outside its locus	pFA6a-kanMX6-P*urg* (PCR tagging construct)	(Watt, et al. 2008)
gld1	homologous	inducible (ethanol)	glycerol dehydrogenase	repressed by high glucose concentrations, useful under glucose-limiting conditions and glycerol as carbon source	pEG1	(Matsuzawa and Tohda 2013)
adh1	homologous	constitutive	alcohol dehydrogenase	strong promoter	pART (LEU2)	(Russell and Hall 1982)
cam1	homologous	constitutive	encoding calmodulin	weak (low-level expression)	pDUAL vector series and pDUAL2 vector series (vectors for genome integration or for episomal expression)	(Matsuyama, Shirai and Yashiroda, et al. 2008, Matsuyama and Yoshida 2012)
tif51	homologous	constitutive	eIF5A, eukaryotic translation initiation factor 5A-1	strong (high-level expression)		
ef1α-c	homologous	constitutive	EF-1α, eukaryotic elongation factor	strong (high-level expression)		
CaMV	plant	inducible (tetracycline), constitutive	gene-switch from the cauliflower mosaic virus	moderate constitutive but under coexpression with tet repressor and subsequential induction with tetracycline 400-fold induction	pSLF101 (*LEU2*)/102 (*ura4*$^{+}$) in the presence of pSLF104 pEVP-Tet-Cat	(Forsburg 1993) (Faryar and Gatz 1991)
SV40	human	constitutive	simian virus 40 T-antigen	used for first splicing studies in *S. pombe*, SV40 transcript was successfully spliced	pSM1/2 (*LEU2*)	(Käufer, Simanis and Nurse 1985)
CMV	human	constitutive	human cytomegalovirus	11 times more effective than SV40	pTR	(Toyoma and Okayama 1990)
CGα	human	constitutive	chorionic gonadotropin a-subunit	9 times more effective SV40	pTR	(Toyoma and Okayama 1990)

1.3.4. Vectors

In general, there are two classes of vectors used in *S. pombe*. The first class is a conventional vector for episomal expression and the second class is an integrative one, where the expression cassette stays stable in the genome. The difference between those two classes is that the integrative vectors do not have an element for extrachromosomal maintenance; they lack the *ars1* sequence.

One of the earliest, but also most successful vector series were the pRep vectors with the thiamine repressible *nmt1* promoter, the *ars1* element and various selection markers (38). The change from homologous selection marker *ura4* to the heterologous *URA3* of *S. cerevisiae* increases the copy number significantly. The *URA3* gene does not replace the homologous gene effectively; hence more copies are necessary to complement the uracil deficiency (48).

Another remarkable episomal expression system is the pTL2M described by Giga-Hama and Kumagai (1999). Unlike many other plasmids, this one has no origin of replication or a prototrophic selection marker, although *S. pombe* strains with auxotrophic background are used for the transformation, but it has a neomycin resistance gene. A cotransformation strategy is chosen, were the pTL2M vector is transformed together with another vector, which contains *ars1*, *stb* and a nutritional marker (*LEU2* or *URA3* from *S. cerevisiae*). During the cotransformation the two vectors recombine and they fuse to one episome. In general, it would be possible to combine *ars1,* the *stb* sequence and the nutritional marker in advance, but then there would have been a problem with the available restriction sites on the plasmids. The nutritional marker from *S. cerevisiae* increases the copy number and allows selection without drug resistance. Additionally, the copy number increases with the concentration of G418 in the medium – a concentration of 200 µg/ml G418 leads to a copy number of around 200.

In one of the first integrative studies, Grallert et al. (1993) examined the disruption of *suc1* gene in various conditions. By using two different promoters – the strong constitutive *adh1* and the moderate SV40 promoter – for the expression of *ura3* gene, they could conclude that the disruption efficiency was not affected by the strength of the promoter. Nonetheless, the transformation rate was slightly increased if using the stronger *adh1* promoter.

Furthermore, they compared two different transformation methods, finding that lithium acetate yields a 10-fold higher disruption efficiency and a 5- to 10-fold higher transformation efficiency in contrast to protoplast method. In addition, a third experiment was designed to test illegitimate integration by subsequently adding non-homologous sequences to the flanking regions of the expression cassette. This process decreased disruption, but not transformation efficiency (49).

Matsuyama et al. (2004) designed the pDUAL vectors series. This vector either expresses the protein of choice in multi copies – by transforming the whole circular plasmid – or in single copy – by transforming a linear expression cassette which is obtained by restriction digestion. The digestion removes the *ars1* and the pUC sequence, a prokaryotic origin of replication, and the linear cassette can be used for gene targeting.

Based on different genetic backgrounds of the used *S. pombe* strains, there are two different versions of the vector: pDUAL and pDUAL2. The first one needs an auxotrophic strain. For example, when *leu1* is the integration marker on the plasmid, it will complement a *leu1-32* mutant. In contrast, pDUAL2 works with the exact opposite strategy: The transformation of the expression cassette in a *S. pombe* wild type strain will lead to a leucine auxotrophic strain. (43) Later the series was extended and the vectors are now available with many different promoters, integration as well as selection markers and tags (50; 51).

1.3.5. PCR based gene targeting

In 1998 Bähler et al. used a PCR based method for gene targeting in order to identify the functions of genes in *S. pombe* the first time. This method allows deleting, tagging or overexpressing any desired gene easily. In principle, a homology between 60 and 80 nucleotides is sufficient to drive homologous recombination and indeed, Bähler et al. (1998) report with almost all tested genes integration efficiency of greater than 50 %. Eight years later the work group around Bähler presented PPPP (Pombe PCR Primer Programs) – a useful webtool for the design of primers for the deletion, tagging and expression of genes in *S. pombe*. The handling of this webtool is simple and straight-forward: For instance, to simulate a gene deletion the target gene is selected. The program outputs several possible primers pairs, which bind in different distances to the untranslated region of the gene. Moreover, the G+C content, Tm and sequencing primers are displayed of each primer (52).

1.4. Posttranslational modifications in *S. pombe*

1.4.1. Secretion signals

There are only a few signal peptides in *S. pombe* that had been used for the secretion of proteins. Two of these signal peptides are derived from the mating pheromones P-factor and M-factor, which belong to the mating types h+ and h-, respectively. The leader peptide of the P-factor precursor – encoded by *map2* – is 12 amino acids long and had been used for the secretion of GFP (48) and the human growth hormone (53). The P-factor is secreted over the endoplasmic reticulum (ER) and the golgi apparatus whereas the M-factor – encoded by *mfm1*, *mfm2* and mfm3 – is secreted over an ABC transporter (*mam4*), which is located in the plasma membrane. Compared to the mating pheromones in *S. cerevisiae* the a- resembles the M-factor and the α- the P-factor(54). The Mam1 ABC transporter exports GFP with the prosequence of the M-factor attached to the N-terminus and a CaaX-box attached to the C-terminus successfully. Hence, the transporter is not specific for the amino acid sequence of the M-factor. This ER independent way allows secreting without unwanted modifications which can be useful in certain applications; however the secretion efficiency of GFP via the Mam1 ABC transporter was relatively low. (55)

Another secretion sequence comes from the *pho1* acid phosphatase. It was successfully applied for the secretion of GFP and human papilloma virus type 16 E7 (HPV)(56). The signal peptide from the carboxypeptidase Y (*cpy1*) is the most efficient one known today. It has been tested for the secretion of GFP as well as for a tumor necrosis factor (TNF) and was compared to other signal peptides like the P- and the α-factor. Interestingly, the α-factor did not work at all in *S. pombe,* although it is a potent and efficient signal peptide in other yeasts like *S. cerevisiae*, *P. pastoris* and *H. polymorpha*. (48) In Table 2 the most important native signal peptides of *S. pombe* are listed.

Alternatively, it is still possible to secrete a protein with its own signal peptide. This strategy works not for all proteins, but there is already a row of proteins like human gastric lipase(57), *S. cerevisiae* invertase (58) or mouse α-amylase (59) where it was successful.

Table 2: This table summarizes the most important signal peptides used so far in *S. pombe*. The table was adapted from Petrescu-Danila, et al. 2009 by the author.

Protein name	**Acces number in NCBI database**	**Amino acid sequence of signal peptide**	**References and applications**
P-factor	*map2* NM 001023030	MKITAVIALLFSLAAASPIPVAD	GFP (48), hGH(53)
Cpy1	*cpy1* NM 001023870	MLMKQTFLYFLLTCVVSAQFNGYVPPE	GFP and TNF (48)
Pho1	*pho1* NM 001023870	MFLQNLFLGFLAVVCANAQFAE	GFP and HPV(56)
M-factor	*mfm1* NM 001020207	MDSMANSVSSSSVVNAGNKPAETLNKTVK **N**-**Y**TPKVPYMC-VIA	GFP (55)

1.4.2. Chaperones and other proteins involved in the secretion in *S. pombe*

Secretion upon the ER dependent pathway is necessary for the functionality of many eukaryotic proteins, because many post-translational modifications only occur in the ER and the golgi apparatus. Many chaperones and foldases are involved in this process. The overexpression of a single protein can increase the secretion efficiency already several times as it was shown by the overexpression of Kar2p (homologous to binding protein BiP in mammals), of Ssa1p (essential chaperone), or of protein disulfide isomerase (PDI) in *P. Pastoris*(60). In *S. pombe* similar overexpression studies of chaperones were performed. One of them reports that the secretion efficiency of human transferrin could be increased up to 30 times, because of the co-overexpression of two putative protein disulfide isomerases – SpPdi2p and SpPdi3p – and additives to the growth medium (61; 62). Furthermore, the endoplasmic reticulum oxidase (ERO) – in *S. pombe* encoded by *ero1-1* – is involved in the regeneration of PDI and therefore also an interesting target for co-overexpression (18).

Another important chaperone is calnexin, a membrane-bound protein and highly conserved chaperon in the ER. In general, calnexin and its soluble counterpart calreticulin are involved in the retention of misfolded secretory proteins. Interestingly, *S. pombe* does not express the soluble calreticulin like mammalian cells. Nevertheless, it seems that the endoplasmic protein quality control in *S. pombe* is more similar to mammalian cells than to other yeasts like *S. cerevisiae* (63). In contrast to *S. cerevisiae*'s calnexin (*CNE1*) the *cnx1* from *S. pombe* is essential for the viability of the cell. A mammalian calnexin version did not complement the *cnx1* gene disruption. However, the expression of the ER lumenal part of *cnx1* was able to complement this gene disruption (64; 65). In 2007 Hajjar, Beauregard and Rokeach

investigated the function of calnexin in *S. pombe*. They screened various *cnx1* mutants and found that a shortened version – consisting of the 160 N-terminal amino acids of the mature protein – led to a 10-fold increased secretion efficiency of a cellulase. Besides, like other viable versions of calnexin, also this version forms complexes with the BiP chaperone.

1.4.3. Glycosylation

An advantage of yeast expression hosts is their ability to modify heterologous proteins in the ER and Golgi apparatus as it is necessary for many proteins in order to function properly. On the other hand one of these modifications in yeasts, the glycosylation, tends to be excessive compared to other organisms. This hyperglycosylation issue can set limits for yeasts as protein producers. Slight changes in the environment can influence the glycosylation pattern of a protein and enzymes in their native host and features like solubility, enzymatic activity, receptor binding, pH- and thermostability are influenced by these glycosylation changes (66). For example, the in vivo activity of human erythropoietin – a hormone which regulates red blood cell production – is severely decreased when its sialic acid residues are not present due to exchange of Asn to Gln in the tripeptide Asn-X-Ser/Thr. Moreover, the desialylated form of erythropoietin shows less affinity to its receptor (67). These implicates that a complete different glycosylation pattern of another species can also have quite an impact on the properties of the expressed protein.

Therapeutic enzymes produced in yeasts can cause immunogenic reactions in the human body and are therefore also less effective than the proteins with native glycosylation (68). At the moment many projects deal with the “glyco engineering” of yeasts and other organisms to overcome these glycosylation assets. The human N-linked glycosylation pathway was for instance completely transferred into *P. pastoris*, leading to a humanized N-linked glycosylation pattern of produced proteins in the yeast (69).

The first step in N-linked glycosylation is conserved in yeasts and higher eukaryotes like insects, plants, and humans. This step takes place in the ER by transferring a $Gl_3Man_9NAc_2$ unit to an asparagine or arginine. One mannose and three glucoses are removed from this core unit, leading to a carbohydrate structure, which is still conserved between the species. This newly formed $Man_8GlcNAc_2$ unit initiates the transmission of the glycoprotein to the Golgi apparatus. Further carbohydrates are added, which are varying from species to species. Most yeasts, for example, transfer several mannose residues – more than 50 are possible – to the N-linked oligosaccharides when the glycoprotein enters the Golgi apparatus (70). Here again, fission yeasts are an exception among other yeasts, because instead of just adding mannose residues they form an outer carbohydrate chain consisting of galactose and mannose residues.

The α-1,6-mannosyltransferase, which is encoded by *och1* in *S. pombe*, is responsible for the mannosylation of glycoprotein's in the Golgi apparatus. An *och1* deficient strain was generated by Takegewa et al. (2009). This strain produces glycoproteins with galactosylated core structures, which differs them from known other species. In order to produce a more humanized glycosylation the workgroup of Takegewa suggest the deletion of the α-galactosyltransferase from the *och1* Δ strain. (71)

Similar to the first step of N-linked glycosylation, the first step of O-linked glycosylation is highly conserved among species and occurs in the lumen of the ER. In *S. cerevisiae* a set of enzymes, known as protein O-mannosyltransferases (Pmt1-7), is responsible for it. These enzymes transfer a dolichyl phosphate-activated mannose to the hydroxyl group of serine and threonine residues (72). In the following steps of O-linked glycosylation up to four mannoses are attached to the first one by mannsoyltransferases resulting in a linear mannosyl chain. In *S. cerevisiae* there are two types of enzymes responsible for this elongation process: The α-1,2-mannsoyltransferases are encoded by the *KTR* gene family and the α-1,3- mannsoyltransferases by the *MNN1* gene (73). In contrast to *S. cerevisiae*, *S. pombe* does not only attach mannoses but also galactoses, resulting in branched glycoside units (Gal_{0-1}-Man_{1-3}-Ser/Thr). The major galactosyltransferase involved in the O-linked glycosylation is Gma12 (*gma12*$^{+}$) (72).

Ikeda *et al.* (2008) characterized six genes (*omh1* to *omh6*) with sequence similarity to the *KTR* gene family. By the analysis of *S. pombe* deletion strains of these six genes, they found

that these enzymes indeed play a major role in the O-linked mannosylation. The O-glycosylation pattern in the *omh1* deficient strain contains mainly galactosylmannaose disaccharides. The phenotype of this strain does not differ from the wild type indicating that the disaccharide does not influence the cell wall function. The *omh1* Δ is interesting for the expression of foreign proteins, because the heterogeneity of the O-linked glycosylation pattern is reduced without visible dysfunctions of the cell wall. (26; 74)

1.5. Process optimization of heterologous protein production in *S. pombe*

In 2007 Giga-Hama *et al*. presented in a review their collected work on the *S. pombe* minimum genome factory (MGF), which includes various approaches like the deletion of large regions in the genome or the generation of protease deficient strains. The MGF concept is based on the theory that many genes are obsolete in nutrient-rich growth conditions and that the deletion of these genes saves energy for the production of heterologous proteins in *S. pombe*. The aim of their work is the development of enhanced and efficient strains for the industrial production of heterologous proteins with *S. pombe* as expression host.

One of their approaches deals with the deletion of 100 kb regions in the genome of *S. pombe*. These regions are supposed to contain only non-essential genes. In this case non-essential means that they are not required under specific laboratory growth conditions. In order to find such regions, various expression profiles have to be analysed. This is done with microarrays, where transcriptome during sexual differentiation or environmental stress is investigated. Further characterizations of these strains have not been published yet. (75)

Another approach is the generation of protease deficient *S. pombe* strains. Single disruptants of 52 different proteases were successfully generated and the obtained strains were tested for improved secretion of human growth hormone (hGH) and changed growth behaviour (53). Later, multiple protease deficient strains were generated and it was possible to create strain where the hGH secretion is 30-fold higher compared to the control strain(76).

Codon optimization of a heterologous protein is another way to improve the expression yield. For *S. pombe* Forsburg created a codon usage table based on three highly expressed genes, *fyi1+/spi1+, adh1+,* and *tpi1+*, which could be used to improve expression levels of heterologous genes(77).

1.6. Degradation and conversion of lignocellulosic material by fungi

S. pombe has been extensively used as expression host for recombinant proteins (18). The fission yeast is an interesting host, due to its evolutionary stand: *S. pombe* actually shares more features with higher eukaryotes than with other yeasts and is therefore especially suited for the expression of proteins deriving from higher eukaryotes. One such application is the expression of lignocellulolytic proteins, which are produced in nature by higher developed fungi (ascomycetes, basidiomycetes) and are often difficult targets for recombinant expression. Lignocellulolytic proteins are involved in the degradation of lignocellulose, the material that builts up the plant's cell walls.

But why is the degradation of lignocellulose an interesting matter for research and industry? Since the beginning of the industrialization, our society is exploiting coal, oil stocks and gas reservoirs, resulting in environmental issues like climate change, air and water pollution. Moreover, due to the limitation of these reservoirs, the constant growth of the population and the newly industrializing countries, we head towards energy and resource shortness. Therefore, we need sustainable alternatives for energy and chemical production. One promising approach is the utilization of lignocellulosic material. Lignocellulose is the major part of plant's biomass and therefore one of the most abundant raw materials in the world. It is a non-uniform material, which consists of three different polymeric structures: lignin, hemi-cellulose and cellulose. Although lignocellulose is a very heterogeneous material, some organisms are able to break it down into smaller compounds. This is mainly done by wood decaying fungi. By taking the biocatalytic tools from these organisms, it is possible to use them for industrial purposes like the production of industrial sugars which can be converted for example to bioethanol.(78)

At the moment biofuel is mainly obtained by the fermentation of food sources like corn and starch in the US or sugarcane juice and molasses in Brazil. Food derived biofuel can never be produced in such capacity that it will become possible to replace fossil fuels completely. Moreover, there are moral considerations concerning the usage of edible raw material for the bioethanol production. Using lignocellulosic biomass for the fermentation of bioethanol is therefore a desirable approach. The food stocks remain untouched and additionally a big-scale conversation is possible. The lignocellulosic material is cheap, because it often derives from residues of other industrial processes or from fast growing plants. (79)

Unfortunately, lignocellulose is not easy to process and therefore an economically viable procedure has not established yet. Thereby the main problem is the heterogenous composition of lignocellulose.

1.6.1. Composition of Lignocellulose

Lignocellulose consists of three polymers with a different degree of heterogeneity. Cellulose is the most homogenous polymer, which is composed only of glucose monomers by a 1,4-β glycosidic bonds. Per cellulose molecule several thousand of these units are beaded on one linear chain. Cellulose comprises the greatest part of lignocellulose with a percentage of 30 to 50 % depending on the plant species. The second polymer is hemicellulose, which consists of a mixture of different carbohydrates. In contrast to cellulose, the structure of hemicelluloses is more variable. It is a branched heteropolymer of hexoses like glucose, mannose and galactose as well as of pentoses like xylose and arabinose. The hemicellulose composition is depending on the wood type (soft wood or hard wood). The third polymer is lignin the most recalcitrant part of lignocellulose. It consists of many aromatic monomers resulting in a very heterogeneous macromolecule. Lignin serves as a glue to stick together hemicellulose and cellulose and all three together built up a barrier to protect the plant.(80; 79)

1.6.2. Biodegradation of lignocelluloses by fungi

Fungi have developed very efficient strategies for the degradation of lignocellulose. Due to the fact lignocellulose is insoluble and cannot be imported into the cell these strategies rely on secretion of enzymes and proteins. In principle there are two types: The first strategy degrades the polysaccharides by the secretion of hydrolytic enzymes and the second one concerns the lignin degradation by an oxidative set of enzymes.

Brown-rot fungi, for example, are very rapid cellulose degraders whereas their influence on lignin is limited to the modification of already degraded lignin. The enzymes for cellulose degradation are classified into three categories. The first one embraces the carboxymethylcellulases (or endo-1,4-β-glucanases), which hydrolyze 1,4-β glycosidic bonds somewhere in the middle of the cellulose molecule. The second category covers the cellobiohydrolases (or exoglucanases). These enzymes represent the largest group of proteins in the fungal cellulase system. Cellobiohydrolases are removing glucose monomers

and dimers from the end of the cellulose chain. The last group of enzymes are the β-glucosidases. They are responsible for the hydrolyzation of glucose dimers to monomers.

For the hemicelluloses degradation similar enzymes as for the cellulose degradation are needed. There are endo-1,4-β-xylanases, xylan esterases, ferulic and p-coumaric esterases as well as α-1-arabinofuranosidases. All these enzymes work together in order to efficiently hydrolyze cellulose and hemmicellulose. There are already many cellulases and hemi-cellulases, which are used in industrial applications. Some of them are deriving from *Trichoderma reesei,* a well studied fungus belonging to the ascomycetes. (81; 78)

Many microorganisms are capable of degrading cellulose and hemicelluloses, but only a few are capable of degrading lignin, the most recalcitrant component of plant cell walls. White-rot fungi can efficiently degrade lignin into CO_2 and are even capable of degrading many long-living environmental pollutants like chlorinated aromatic compounds, heterocyclic aromatic hydrocarbons, various dyes and synthetic high polymers. *Phanerochaete chrysosporium* is the best studied white-rot fungus and belongs to the phylum of basidiomycetes. Enzymes of this species are involved in the oxidative degradation of lignin and show little substrate specificity. Among these enzymes there are laccases, which can abstract single electrons from aromatic substrates, peroxidases, oxidases, reductases and dehydrogenases and thus perform degradation of lignin effectively. (82)

1.6.3. Bioconversion of lignocellulose

The conversion of lignocellulose into other products requires several steps:

1. Pretreatment (mechanical, chemical or biological)
2. Hydrolysis (chemical or biological)
3. Production of aimed chemical (e.g. bioethanol by fermentation of sugars)
4. Separation and purification

The pretreatment of lignocellulosic material is necessary to decentralize the lignin and hemicellulose and further obtain cellulose for fermentation to bioethanol. Thermochemical and physical pretreatments have high energy requirements and therefore result in high costs. The environmental best solution is a biological pretreatment. It is based on lignocellulose decaying microorganisms like brown-, white- and soft-rot fungi. These fungi secrete enzymes for the degradation of lignin and hemicellulose. Thus, cellulose is

uncovered and is hydrolyzed in the next step. The hydrolyzation can be done either chemically with acids or again biologically with enzymes. The obtained sugars can be used as nutrient source for *S. cerevisiae* – producers of bioethanol. In the case of bioethanol, the separation and purification occurs by distillation. However, the application is not constrained to fuels; there are many different possibilities for the utilization of lignocellulosic material as shown in Figure 4 (83; 78).

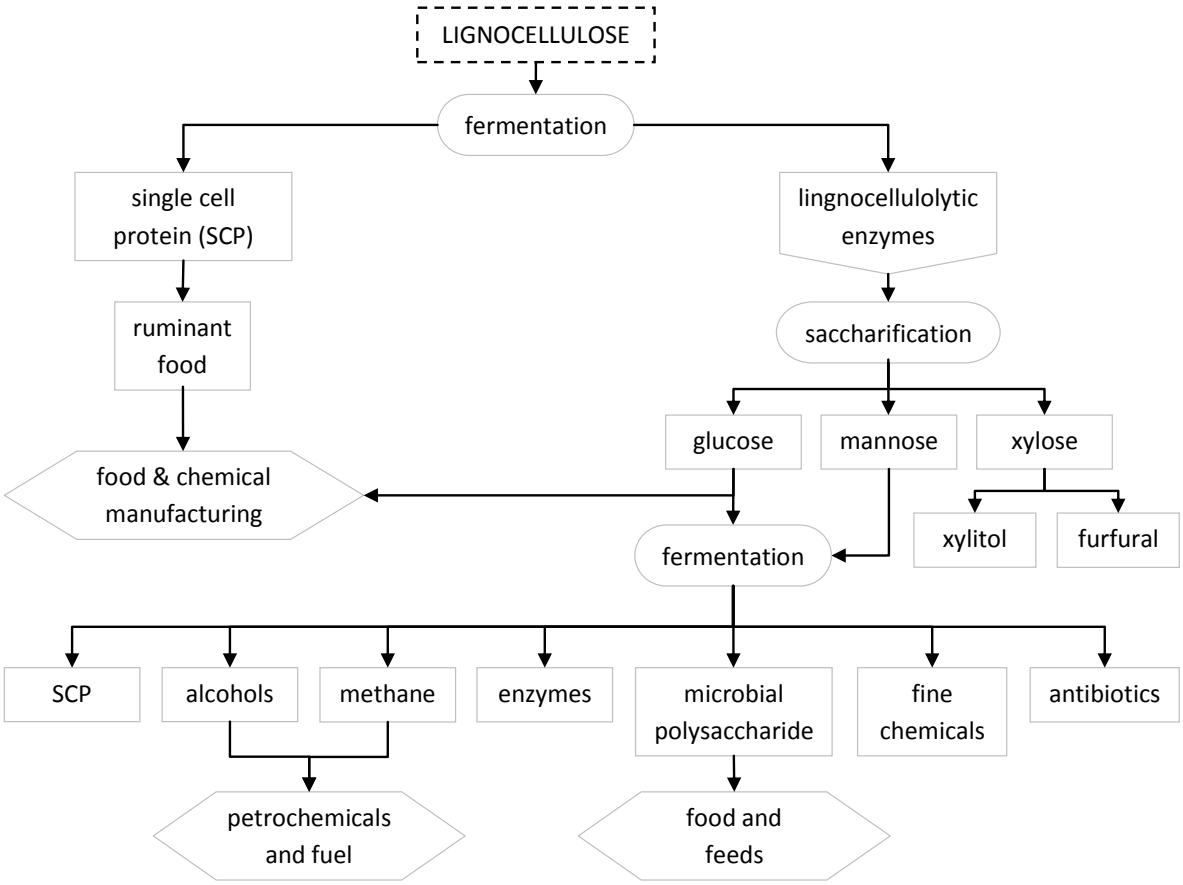


Figure 4: Possibilities for the utilization of lignocellulosse. (78)

1.7. *Fomes fomentarius* – a white-rot fungus

Fomes fomentarius belongs to the basidiomycetes and is located in Africa, North America, Asia and Europe. It is a pathogen for hardwoods mainly and it is most often growing on birch and beech trees. *F. fomentarius* is a so called primary colonizer and is able to keep off other fungi from the wood it is growing on. Humans found many applications for this fungus over the time: For the production of tinders (Zunderschwamm) the fruiting body of the fungus is soaked in salpeter. Even the ice mummy Ötzi carried small pieces of this fungus, probably also for the purpose of igniting fire. Besides it is used for the production of clothes and even in the traditional Chinese medicine(84). Indeed, polysaccharides produced by *F. fomentar*ius showed antitumor effects (85). Moreover, *F. fomentarius* had been subject of some other investigations like the autofluorescence of its fruiting body (86), its feature to absorb dyes (87) and how heavy metals influence its growth (88).

F. fomentarius belongs to the white-rot fungi and is able to degrade the lignin as well as the polysaccharide part of lignocellulose. A laccasse from *F. fomentarius* was already purified and characterized. In general, a laccase is a copper-containing enzyme with broad substrate specificity, which is able to oxidize phenolic substrates by reducing oxygen to water. The properties of this laccase are similar to those of other already characterized laccases(89).

In 2006 Jaszek *et al*. triggered successfully the expression of ligninolytic enzymes of *F. fomentarius* by causing oxidative stress due to the addition of menadion to the fungi culture. A more recent study reports the production of several hydrolytic and oxidative enzymes from *F. fomentarius:* 1,4-β-glucosidase is produced in high quantities. Therefore, they purified and characterized this enzyme (90).

1.8. Codon usage

In general, 18 amino acids out of 20 can be decoded by two or more codons. Although codons may encode the same amino acid, they are not equally used. This phenomenon is called codon usage or codon bias and describes the discrimination or preference of codons by one species. The reason for this bias are the concentrations of the codon related tRNA, which differ depending on the species. This mechanism represents a regulation point for the biosynthesis of proteins: High expression codons, for example, are often used in house-keeping genes (Hatfield and Roth 2007).

For the generation of a codon usage table there are different methods available (Davis and Olsen 2010). Alternatively, there are already a lot of codon usage tables available. The codon usage database provides information for 35799 organisms, which corresponds to 3027973 complete protein coding genes (Codon Usage Database 2007).

The heterogeneity in the codon usage of two different species can result in problems for the heterologous gene expression. Therefore, a rational gene design, where the recombinant gene is codon optimized for the expression host can increase the protein amount. As Mellitzer *et al*. (2012) showed, codon optimized lignocellulolytic enzymes in the methylotrophic yeast *P. pastoris*, achieved a 5- to 8-fold higher protein yield for the optimzed xylanase A of *Thermomyces lanuginosus* compared to similar studies without optimized genes.

1.9. Domain of unknown function

Although DNA sequencing technologies are already very advanced and sequence data of genes accumulate rapidly there are still many sequences found in databases where the function of corresponding proteins is still unknown. A domain of unknown function (DUF) is a protein or a protein domain, which is uncharacterized and has not been registered in the Pfam (protein family) database. There are still many protein functions, which have not been discovered yet. For example, although *S. cerevisiae* belongs to the best studied organisms, around 17 % of the genome (corresponds ~ 1000 proteins) are still uncharacterized. So, there is still a lot of potential to discover new enzymatic activities and protein functions (91). However pure bioinformatics approaches usually are not sufficient to identify natural activities of such proteins but functional analyses are necessary.

1.10. Aims of the thesis

The first part of this thesis aims to generate new integrative expression systems for *S. pombe*, which should lead to mitotically stable expression of foreign proteins. Two strategies were developed: First, PCR cassettes that are integrating into the *S. pombe* genome resulting in the deletion of the targeted gene. The second approach was relying on a bifunctional vector: It can be used for episomal expression and if transformed into a linear expression construct by restriction digestion it is suitable for integration into the *S. pombe* genome. Such integration results in an insertion of the expression construct into the genome and a disruption of the targeted gene.

Moreover, a comparison between episomal expression and expression from the integrated construct was a major goal of these studies. To clarify if the new systems have an advantage over the so far used episomal system, the pGAZ2 vector (1) – expressing eGFP as well – was used as control.

The major aim in the second part of the thesis was to apply these new tools and methods for foreign protein expression in *S. pombe*, especially for potential new lignocellulolytic proteins. The wood decaying fungi *F. fomentarius* was cultivated under inducing conditions – i.e. grown on wood – in order to find new enzymes involved in the degradation of lignocellulose. The mRNA of this fungus was isolated, transformed into a cDNA library and then sequenced already before. A bioinformatical analysis of the obtained sequences resulted in a list of interesting targets, which are tested for secretion in *S. pombe.* Moreover an expression host comparison between *S. pombe* and *P. pastoris* should be done by redesigning the target genes for the respective hosts.

2. Material and Methods

2.1. Instruments, devices and software

2.1.1. Centrifuges and associated materials

- 5810 R centrifuge, Eppendorf
- 415 R centrifuge, Eppendorf
- JA-10 rotor, Beckman Coulter

2.1.2. Shakers and incubators

- Titramax 1000 platform shaker, Heidolph
- Thermomixer comfort, Eppendorf
- RS 306 rotary shaker, Infors
- Drying oven, Binder
- HT Multitron incubator shaker, Infors
- GFL-3013 shaker, Gesellschaft für Labortechnik

2.1.3. Polymerase chain reaction cyclers

- GeneAmp® 2720 thermal cycler, Applied Biosystems
- Mastercycler® personal, Eppendorf

2.1.4. Photometers and plate readers

- SynergyMx plate reader, Bioetek
- Spectrophotometer DU 800, Beckman Coulter
- Plus 384 Spectramax, Molecular Devices
- Semi-micro cuvette, polystyrene 10x4x45 mm, Sarstedt
- Gemini XS Spectramax, Molecular Devices
- NanoDrop 2000c spectrophotometer, Thermo Scientific
- G:BOX imaging system, Syngene

2.1.5. Electroporation devices

- Gene Pulser, Bio-Rad Laboratories
- Electroporation cuvettes, 2 mm, Bridge Bioscience

2.1.6. Pipettes and pipette tips

- Pipetman, adjustable single channel pipettes 0.2 - 1000 µl, Gilson
- Research® pipette, adjustable 0.1-2.5 µl, Eppendorf
- Pipette tips, 10, 200 and 1000 µl, with and without filter, GBO Greiner Bio-One
- Proline® multichannel pipettor, 8 channels 5-50 µl, Biohit
- Proline® multichannel electronic pipettor, 8 channels 50-1200 µl, Biohit
- Tips 300 µl, Single Tray, Biohit
- Tips 1200 µl, Bulk, Biohit

2.1.7. Reaction tubes

- Microcentrifuge tubes, 1.5 ml with lid, GBO Greiner Bio-One
- Polypropylene tubes, sterile 15 ml, GBO Greiner Bio-One
- Polypropylene tubes, sterile, 50 ml, GBO Greiner Bio-One
- Polypropylene tubes, sterile, 12 ml volume, GBO Greiner Bio-One
- Polypropylene tubes, sterile, 15 ml volume, GBO Greiner Bio-One

2.1.8. Microplates

- 96 well plates, black and white polystyrene with clear and solid bottom, GBO Greiner Bio-One
- 96 well polypropylene microplates for PCR, GBO Greiner Bio-One
- Silicone sealing mat for 96 well PCR plates
- 96 deep well plate and cover, polypropylene, Scienceware

2.1.9. Web tools and software

- *Pombe* PCR primer programs (PPPP) (92)
- Sequence alignment with ClustalW2 (93)
- Expasy tool to translate DNA sequences (94)
- Basic local alignment search tool (BLAST) (95)
- RNA structure prediction: (96)
- GC-distribution(97)
- Primer3Plus (98)
- Conversion of DNA strand to antiparallel (99)
- SignalP 4.0 Server, prediction of signal peptides (100)
- Gene designer 2.0, DNA2.0
- Lasergene® 7, DNASTAR

2.1.10. Other devices

- DASGIP bioreactor, DASGIP
- DFC 350 FX microscope, Leica
- Vortex-Genie 2, Scientific Industrie
- inoLab® pH720 pH meter, WTW
- Polyplast Temp Din pH electrode, Hamilton
- MR 3000 and MR 2002 magnetic stirrers, Heidolph
- BL 120S scale, Sartorius
- PG12001-S delta range scale, Mettler-Toledo
- arium® basic ultrapure water system, Sartorius
- Äkta purifier, GE Healthcare
- Äkta prime, GE Healthcare

2.2. Material for electrophoresis and Dot Blot

2.2.1. Material for agarose gel electrophoresis

- Biozym LE agarose, Biozym Scientific
- GeneRuler™ 1kb DNA-ladder, Thermo Scientific
- 6x DNA loading dye, Thermo Scientific
- Sub-cell GT, Bio-Rad Laboratories
- Power Pac™ Basic, Bio-Rad Laboratories
- Chroma 43 mittelwellig 302 nm, Laborgeräte Vetter
- GelDoc-It™ Imager, UVP
- BioImaging Systems Gel HR Camera 6100, UVP

2.2.2. Material for sodium dodecyl-Polyacrylamid gel electrophoresis (SDS-PAGE)

- NuPage® 4-12% Bis-Tris gel, 1.0 mm x 15 well, Invitrogen
- PageRuler™ prestained protein ladder, Thermo Scientific
- NuPage® MOPS SDS running buffer (20x), Invitrogen
- NuPage® sample reducing agent (10x), Invitrogen
- NuPage® LDS sample buffer (4x), Invitrogen
- XCell SureLock™, Invitrogen
- Power Ease 500, Invitrogen

2.2.3. Material for Dot blot and Western blot

- XCell II Blot Module™, Invitrogen
- Amersham™ HybondTM-ECL membrane, GE Healthcare
- Bio-Dot® and Bio-Dot SF microfiltration apparatus, Bio-Rad
- 25x Tris-glycine transfer buffer: 12 mM Tris base, 96 mM glycine, the pH of the solution is 8.3 (do not adjust with acid or base), store up to 6 months at room temperature
- Transfer buffer: 40 ml/l 25x Tris-glycine transfer buffer, 200 ml/l methanol, 760 ml/l deionized water
- Tris-buffered saline (TBS) buffer: 10 mM Tris-HCl (pH7.5), 150 mM NaCl

- TBS-Tween/triton buffer: 20 mM Tris-HCl (pH7.5), 500 mM NaCl, 0.05 % (v/v) Tween20, 0.2 % (v/v) Triton X-100
- Blocking buffer: 3 % (w/v) BSA (Sigma, Cat. No. A7906) in TBS buffer
- TBS-BSA1%: 1 % (w/v) BSA (Sigma, Cat. No. A7906) in TBS buffer
- Tetra His antibody, mouse monoclonal IgG1, Qiagen (34670):The lyophilized antibody was, as recommended, resuspended in 500 µl water to an end concentration of 1 mg/ml. 25 µl aliquots of this stock solution were stored at -20°C
- Goat anti-mouse IgG1 (γ1), horseradish peroxidase conjugate, Invitrogen (A10551): The lyophilized antibody was resuspended in 500 µl sterile PBS buffer to an end concentration of 1 mg/ml. This stock solution was stored at 4 °C
- 1^{st} antibody solution (1:1000): 25 µl of Tetra His antibody (mouse monoclonal IgG1) stock solution in 25 ml TBS-BSA1% solution
- 2^{nd} antibody solution (1:000): 5 µl of Goat anti-mouse IgG1 (γ1) stock solution in 25 ml TBS-BSA1% solution
- SuperSignal West Pico chemiluminescent substrate, Thermo Scientific (34079)

2.3. Media and solutions

2.3.1. Antibiotics

- Ampicillin stock, 100 mg/ml: 5 g ampicillin / 50 ml dH2O, filter-sterilize (Ø 0,2 µm)
- Kanamycine stock, 100 mg/ml: 5 g kanamycine / 50 ml dH2O, filter-sterilize (Ø 0,2 µm)
- Zeocin® stock, 100 mg/ml: 2 g Zeocin® / 20 ml dH2O, filter-sterilize (Ø 0,2 µm)
- Geneticin stock, 100 mg/ml: 2 g geneticin / 20 ml dH2O, filter-sterilize (Ø 0,2 µm)
- Antibiotic concentrations for E. coli: 25 mg/l Zeocin®, 100 mg/l kanamycine (1 ml stock solution)
- Antibiotic concentrations for yeasts: 100 mg/l Zeocin®, 300 mg/l geneticin

2.3.2. Media and solutions for *E.coli*

- Low salt Luria Bertani (LB): 10 g/l Bacto™ tryptone, 5 g/l Bacto™ yeast extract, 5 g/l NaCl, autoclave sterilization
- LB agar: 35 g/l LB agar (Lennox, Roth)
- Super optimal broth with gluose (SOC): 5 g/l Bacto™ yeast extract, 20 g/l Bacto™ tryptone, 0.58 g/l NaCl, 2 g/l $MgCl_2$, 2.46 g/l $MgSO_4$, 0.18 g/l KCl, 3.81 g/l α-D(+)-glucose monohydrate, autoclave sterilization

2.3.3. Media and solutions for *S. pombe*

- 10x dextrose (D): 220 g/l α-D-glucose·H_2O (20 %)
- Yeast extract supplemented medium (YES) – rich medium: 5 g/l yeast extract (0.5%), 1 g/l supplements (Formedium), 15 g/l Difco Bacto Agar for solid medium, add 150 ml/l 10x D (3 %) after autoclaving
- Edinburgh minimal medium (EMM2): 3 g/l KH-phthalate (14.7 mM), 2.2 g/l Na_2HPO_4 (15.5 mM), 5 g/l NH_4Cl (93.5 mM), 15 g/l Difco Bacto Agar for solid medium, for 1 l medium dissolve in 880 ml dH_2O and autoclave; add 100 ml/l 10x D (2 %), 20 ml/l salt stock, 1 ml/l vitamin stock and 0.1 ml/l mineral stock and if needed 20 ml amino acid stock (50x) after autoclaving

- 50× salt stock solution : 52.5 g/l $MgCl_2 \cdot 6H_2O$, 0.735 g/l CaCl2·$2H_2O$, 50 g/l KCl, 2 g/l Na_2SO_4
- 1000× vitamin stock solution: 1 g/l panthothenic acid, 10 g/l nicotinic acid, 10 g/l inositol, 10 mg biotin
- 10000× mineral stock solution: 5 g/l boric acid, 10 g/l citric acid, 4 g/l MnSO4, 4 g/l $ZnSO_4 \cdot 7H_2O$, 20 g/l $FeSO_4 \cdot 7H_2O$, 0.4 g/l H_2MoO_4, 1 g/l KI, 0.4 g/l $CuSO_4 \cdot 5H_2O$
- 50x leucine: 11,25 mg/ml
- 50x histidine: 11,25 mg/ml
- Dithiothreitol (DTT) buffer: 20 mM Hepes, 0.6 M sorbitol, filter-sterilize (Ø 0,2 µm)
- 10x lithium acetate (LiAc): 1 M LiAc, pH 7.5 (adjust with diluted acetic acid), filter-sterilize (Ø 0,2 µm)
- 10x Tris-HCl/EDTA (TE): 0.1 M Tris–HCl, 0.01 M EDTA, pH 7.5 (adjust with HCl), filter-sterilize (Ø 0,2 µm)
- LiAc-TE buffer: Mix from 10x stocks and fill up with sterile water
- PEG4000-LiAc-TE buffer (20 ml): dissolve 8 g of PEG 4000 in 2 ml of 10xLiAc, 2 ml of 10xTE and 9.75 ml water, filter-sterilize (Ø 0,2 µm), can be stored up to 1 month
- Carrier DNA: 10 mg/ml lyophilized herring sperm DNA (Boehringer) in TE buffer, sheared with a syringe and cannula (pull up and down several times, ~ 15 minutes), filter-sterilize (Ø 0,2 µm), aliquote and store at -20°C, before usage cook for 20 minutes

2.3.4. Media and solutions for *P. pastoris*

- 10x D: 220 g/l α-D-glucose·H2O (20 %)
- 500x biotin: 200 mg/l D-Biotin, filter-sterilize (Ø 0,2 µm)
- 10x yeast nitrogen base (YNB): 134 g/l Difco™ yeast nitrogen base w/o amino acids, autoclave sterilization
- 1 M potassium phosphate buffer (PPB): 30.0 g/l K_2HPO_4, 118.0 g/l KH_2PO_4, pH 6.0, autoclave sterilization
- Yeast extract/peptone/dextrose (YPD) medium: 10 g/l Bacto™ yeast extract, 20 g/l Bacto™, 15g/l Difco Bacto Agar for solid medium, for 1 l medium dissolve in 900 ml dH2O and autoclave; add 100 ml/l 10x D (2 %)

- Basal medium with 1 % dextrose (BMD1): 200 ml/l 1 M PPB, 100 ml/l 10x YNB, 50 ml/l 10x D, 2 ml/l 500x biotin
- Basal medium with 1 % methanol (BMM2): 200 ml/l 1 M PPB, 100 ml/l 10x YNB, 10 ml/l methanol, 2 ml/l 500x biotin
- Basal medium with 5 % methanol (BMM10): 200 ml/l 1 M PPB, 100 ml/l 10x YNB, 50 ml/l methanol, 2 ml/l 500x biotin

2.3.5. Media and solutions for bioreactor cultivation in DASGIP system

- YES batch medium (YB): 30 g/l glucose, 5 g/l yeast extract (0.5 %), 1 g/l supplements (Formedium), 1 ml/l 1000x vitamins stock solution, 100 µl/l 10000x minerals stock solution, 0.010 mg/l biotin
- YES feed medium (YB): 300 g/l glucose, 5 g/l yeast extract (0.5 %), 1 g/l supplements (Formedium), 1 ml/l 1000x vitamins stock solution, 100 µl/l 10000x minerals stock solution, 0.010 mg/l biotin
- SMF23 (101): 30 g/l glucose, 1.9 g/l $(NH_4)_2SO_4$, 0.39 g/l KH_2PO_4, 0.13 g/l $MgSO_4 \cdot 7H_2O$, 0.04 g/l Na_2HPO_4, 0.2 g/l $CaCl_2 \cdot H_2O$, 1 ml/l 1000x vitamins stock solution, 100 µl/l 10000x minerals stock solution, 100 µl/l 500x biotin
- FSS05 (101): 80 g/l yeast extract, 300 g/l glucose, 19.4 g/l $(NH_4)_2SO_4$, 3.9 g/l KH_2PO_4, 1.3 g/l $MgSO_4 \cdot 7H_2O$, 0.04 g/l Na_2HPO_4, 0.2 g/l $CaCl_2 \cdot H_2O$, 100 µl/l 10000x minerals stock solution, 1 ml/l 1000x vitamins stock solution, 100 µl/l 500x biotin
- Antifoam, Bussetti & Co (LM1207072): autoclave sterilization
- 12.5% (v/v) NH3

2.3.6. Buffers for protein purification on an ÄKTA purifier

- Wash buffer: 20 mM sodium phosphate, 500 mM NaCl, pH 7.4
- Elution buffer:20 mM sodium phosphate, 500 mM NaCl, 500 mM imidazole, pH 7.4
- Desalting buffer: 20 mM sodium phosphate, 150 mM NaCl, pH 7.4

2.3.7. Other buffers, solutions and chemicals

- dNTP mix: 10 mM each, dATP, dTTP, dCTP, dGTP, Thermo Scientific

- 50x Tris-HCl/acetate/EDTA (TAE) buffer: 242 g/l Tris base, 57.1 ml/l glacial acetate, 100 ml/l 0.5 M EDTA
- 1 M dithiothreitol (DTT): 0,16 g/ml DTT, filter-sterilize (Ø 0,2 µm)
- Bicine/ethylene glycol/dimethyl sulfoxide (DMSO)/sorbitol (BEDS): 10 mM bicine-NaOH (pH 8.3), 3% (v/v) ethylene glycol, 5% DMSO (v/v) and 1 M sorbitol filter-sterilize (Ø 0,2 µm), freeze aliquots at - 20 °C
- SDS gel staining solution: 2 g Coomassie Blue R, 500 ml ethanol, 100 ml acetic acid (100 %), 400 ml dH2O
- SDS gel destaining solution: 10 % absolute ethanol, 10 % acetic acid
- Yeast lysis buffer: 2 % Triton X-100, 1 % SDS, 100 mM NaCl, 10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0)
- BioRad, BCA Protein Assay Kit, Thermo Scientific
- 5 X ISO buffer: 25 % (w/v) PEG-8000, 500 mM Tris/Cl (pH 7.5), 50 mM $MgCl_2$, 50 mM DTT, 1 mM dATP, 1 mM dCTP,1 mM dGTP 1 mM dTTP, 5 mM NAD; filter-sterilize (Ø 0,45 µm)
- Lysis buffer: 2 % TritonX-100, 1 % SDS, 100 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0)

2.4. Enzymes

The enzymes were applied according to their manuals.

- BamHI, 5'G^GATTC3', Thermo Scientific (# ER0051)
- BglII, 5'A^GATCT3', Thermo Scientific (# ER0082)
- EcoRI, 5'G^AATTC3', Thermo Scientific (# ER0271)
- HindIII, 5’A^AGCTT3‘,Thermo Scientific (# ER0501)
- FastDigest BamHI, 5'G^GATTC3', Thermo Scientific (# FD0054)
- FastDigest BglII, 5'A^GATCT3', Thermo Scientific (# FD0083)
- FastDigest EcoRI, 5'G^AATTC3', Thermo Scientific (# FD0274)
- FastDigest HindIII, 5’A^AGCTT3‘, Thermo Scientific (# FD0504)
- FastDigest NotI, 5'GC^GGCCGC3', Thermo Scientific (# FD0594)
- FastDigest SmiI (SwaI), 5'ATT^AAAT3', Thermo Scientific (# FD1244)
- FastDigest SpeI (BcuI), 5'A^CTAGT3', Thermo Scientific (# FD1254)

- FastDigest XhoI, 5'C^TCGAG3', Thermo Scientific (# FD0694)
- NotI, 5'GC^GGCCGC3', Thermo Scientific (# ER0595)
- SmiI (SwaI), 5'ATT^AAAT3', Thermo Scientific (# ER1241)
- SpeI (BcuI), 5'A^CTAGT3', Thermo Scientific (# ER1251)
- XhoI, 5'C^TCGAG3', Thermo Scientific (# ER0695)
- T4 DNA Ligase, Thermo Scientific (# EL0014)
- EndoH, New England Biolabs (P0702S)
- FastAP thermosensitive alkaline phosphatase, Thermo Scientific (# EF0654)
- GoTaq® DNA polymerase, Promega (M3005)
- Phusion high-fidelity DNA polymerase , Thermo Scientific (# F-530L)
- T5 exonuclease, Biozyme (162340)

2.5. Substrates for enzymatic assay

- 4-Methylumbelliferyl-alpha-D-galactopyranoside, Sigma (M7633)
- 4-Methylumbelliferyl-beta-D-mannopyranoside, Sigma (M 0905)
- 4-Methylumbelliferyl-alpha-L-arabinofuranoside, Sigma (M9519)
- 4-Methylumbelliferyl-beta-D-cellobioside, Sigma (M 6018)
- 4-Methylumbelliferyl-beta-D-galactopyranoside, Sigma (M 3633)
- 4-Methylumbelliferyl-beta-D-lactopyranoside, Sigma (M2405)
- 4-Methylumbelliferyl-beta-D-glucuronide hydrate, Sigma (M 9130)
- 4-Methylumbelliferyl-alpha-L-rhamnopyranoside, Sigma (M 8412)

2.6. Strains

- *S. pombe,* wild type strain, culture collection number (CC #) 3370, , mating type h-s, retrotransposon Tf2, production of gycerol dehydrogenases, isolated from grape most (102)
- *P. pastoris*, CBS-7435 mutS (Δ*AOX1*), CC # 3445

2.7. Primers

Table 3: Primers used for the construction of the PCR expression cassettes. The universal linker (**CTAGGTACTTCGAAACGAGGACTTCA**) allows connecting two fragments in an OE-PCR or a Gibson assembly.

leu2_fwd	CTAACAAATTTGAGGGGAGAGGTTAATAATATCTTTAATTTCTTAATATTTTGATTTAGC
leu2-adh1_fwd	CTTAATATTTTGATTTAGCCACATAGTAGAGATACTGCTGCATGCCCTACAACAACTAAG
leu2-tefTT_rev	GTAAATAAACAGTATATGACAACATATTACAAAGCATGCTGGATGGCGGCGTTAGTATCG
leu2_rev	ATATGAAGCTATTAGTATGATATTTCCGACAGGCTCTTATTGTAAATAAACAGTATATGA
leu1_fwd	CGATATCCCAATCTGTAGTATAAAAAGTGCAACAGTGTTGAGTTTCCCGAAACCAGGCAT
leu1-adh1_fwd	GTTTCCCGAAACCAGGCATTTTTGCGGTAGGGTCTGAAAGCATGCCCTACAACAACTAAG
leu1-tefTT_rev	GTAAATCAATTCCATGCTTTTGCAACTTAATACGCATGCTGGATGGCGGCGTTAGTATCG
leu1_rev	AGTACACAGCGACAACTCGGTCATAAAGTTGAACGGATGTCGTAAATCAATTCCATGCTT
his3_fwd	GAGCTTTATTTCTCTTTCCTCTTCAGGTTTCTGAATTTATTGGATATAATGACACGGTGA
his3-adh1_fwd	GGATATAATGACACGGTGATACACGTTGTAATGAATTAAGCATGCCCTACAACAACTAAG
his3-tefTT_rev	CATCAGAAAAGAAATGTGACGAGTCCGGAAGGGGCATGCTGGATGGCGGCGTTAGTATCG
his3_rev	GGAATGCAAGCGCAATTAGCGAACACAAAATATATTGGGTACATCAGAAAAGAAATGTGA
EcoRV-eGFP_fwd	CTTTGATATCATGGCTAGCAAAGGAGAAGAAC
SpeI-eGFP_rev	CGAGACTAGTTTACTTGTACAATTCATCCATGCC
overlap-rev	**CGTGAAGTCCTCGTTTCGAAGTACCTAG**AGCTTGTGATATTGACGAAAC
overlap-fwd	**CTAGGTACTTCGAAACGAGGACTTCACG**GATCTAACGACATGGAGGCCC

Table 4: Primers for the amplification of the integration vectors.

Adh1_leu2-3‘_fwd	TCTTGTACCAATTCTCGTATTGAGGATTTGCGTCTTGCTGCATGCCCTACAACAACTAAG
leu2-5‘_tefTT_fwd	TCGCTATACTGCTGTCGATTCGATACTAACGCCGCCATCCAAGTAAAAACATGCGCATTC
leu2-5‘_pUC_rev	GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGATCCAGATCTGTTCCCAATCATCGCC
pUC_leu2_fwd	CCTCTTGCTCCCAAGGGCGATGATTGGGAACAGATCTGGATCCCGTAGAAAAGATCAAAG
pUC_ars1_rev	TTAGCTATAATTATAGACGATGCCCTTGTATCCCATTCTGTCTGTGAGCAAAAGGCCAGC
ars1_pUC_fwd	CCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACAGACAGAATGGGATACAAG
Ars1_leu2-3‘_rev	TTCAACTTCAATATCGTACTTAGCATTCTCATCGGATCCAGATCTCCTTAACCACTCGGA
leu2-3'_ars1_fwd	GTCAGATAAGTCACTATGTCCGAGTGGTTAAGGAGATCTGGATCCGATGAGAATGCTAAG
leu2-3_adh1_rev	GCTTCCGCATGATAGCCATTTTCTTAGTTGTTGTAGGGCATGCAGCAAGACGCAAATCCT
Adh1_leu1-3‘_fwd	GGTTCATTGTTTGGTTGAGCCCATTCACGGTAGCGCTCCCGCATGCCCTACAACAACTAA
tefTT_leu1-5‘_rev	AAAGTTGCAAGGTCGAAGGTTGGCCCAAACACCCATACTCTGGATGGCGGCGTTAGTATC
Leu1-5‘_tefTT_fwd	CGCTATACTGCTGTCGATTCGATACTAACGCCGCCATCCAGAGTATGGGTGTTTGGGCCA
Leu1-5‘_pUC_rev	AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGATCCAGATCTTGCGCCAAAGACGA
pUC_leu1-5‘_fwd	AATGTTTTGGCAACTTCTCGTCTTTGGCGCAAGATCTGGATCCCGTAGAAAAGATCAAAG
Ars1_leu1-3‘_rev	TGACCAAAAGCATGGCTGCGGAGTCAATGAGCTGGATCCAGATCTCCTTAACCACTCGGA
Leu1-3‘_ars1_fwd	GTCAGATAAGTCACTATGTCCGAGTGGTTAAGGAGATCTGGATCCAGCTCATTGACTCCG
Leu1-3‘_adh1_rev	CTTCCGCATGATAGCCATTTTCTTAGTTGTTGTAGGGCATGCGGGAGCGCTACCGTGAAT

Table 5: Primers for cDNA target amplification. The forward primers are introducing a *Hind*III restriction site (AAGCTT) and the Kozak sequence (**GAAACG**) to the amplified cDNA target. The reverse primer attaches a *Spe*I restriction site (ACTAGT) and a 6x his tag (*gtgatggtgatggtgatg*).

IDN5_F	TTTAAGCTT**GAAACG**ATGTGCTACAGGACGACTTG
IDN5_R	TTTACTAGTTTA*gtgatggtgatggtgatg*CCCGAGATATTTGTACAC
IDN6_F	TTTAAGCTT**GAAACG**ATGCTGAGGTTTACAATCTCC
IDN6_R	TTTACTAGTTTA*gtgatggtgatggtgatg*TTTGATCTGGGATTCAAG
IDN7_F	TTTAAGCTT**GAAACG**ATGAAGTCCTTTACAACCCTC
IDN7_R	TTTACTAGTTTA*gtgatggtgatggtgatg*CGCCGAGACGTACGTCTG
IDN8_F	TTTAAGCTT**GAAACG**ATGCGCTTTTCTCTCTTTCTC
IDN8_R	TTTACTAGTTTA*gtgatggtgatggtgatg*CGTCGCCTTCCCGCCCGTCG
IDN9_F	TTTAAGCTT**GAAACG**ATGTTCGCCCAATCTCTTGTG
IDN9_R	TTTACTAGTTTA*gtgatggtgatggtgatg*CTGGCACTGGTAGTAGTAC
IDN10_F	TTTAAGCTT**GAAACG**ATGCAGCTCTTTACCGCAGTC
IDN10_R	TTTACTAGTTTA*gtgatggtgatggtgatg*AGTGAGCTTGGGCGTGGTC
IDN13_F	TTTAAGCTT**GAAACG**ATGTTCGCTCGCTCTGTCGTC
IDN13_R	TTTACTAGTTTA*gtgatggtgatggtgatg*GAGCTCATCGCACCAGG
IDN16_F	TTTAAGCTT**GAAACG**ATGAAGTTCACGTACGCTCTTG
IDN16_R	TTTACTAGTTTA*gtgatggtgatggtgatg*CAGCATGATCGGGATGCAAC
IDN24_F	TTTAAGCTT**GAAACG**ATGACGCGCATCCTGCTTCC
IDN24_R	TTTACTAGTTTA*gtgatggtgatggtgatg*TTGAGGTCCAACGACG
IDN40_F	TTTAAGCTT**GAAACG**ATGTTCGCTATCATCTCTG
IDN40_R	TTTACTAGTTTA*gtgatggtgatggtgatg*GTTCCCGCCAGTCCAAAC
GH61_1_F	TTTAAGCTT**GAAACG**ATGATGAACCTCTCTGTTATTGC
GH61_1_R	TTTACTAGTTTA*gtgatggtgatggtgatg*GCACGTGAAAAGAGATGGTC
GH61_2_F	TTTAAGCTT**GAAACG**ATGAAGACCTTCGCTTCTCT
GH61_2_R	TTTACTAGTTTA*gtgatggtgatggtgatg*CAGGCACTGGCTGTAGTATTC
GH61_3_F	TTTAAGCTT**GAAACG**ATGTTCTCGTCTACTTTCTC
GH61_3_R	TTTACTAGTTTA*gtgatggtgatggtgatg*GTGGTGCCGCTTGGCGTGGTG
GH61_4_F	TTTAAGCTT**GAAACG**ATGAAGCTCGCGCTCCTCTTC
GH61_4_R	TTTACTAGTTTA*gtgatggtgatggtgatg*TCCGGACCAGACAGCGGGCC
GH61_5_F	TTTAAGCTT**GAAACG**ATGAAGTACTTCTCTGTTCTTGC
GH61_5_R	TTTACTAGTTTA*gtgatggtgatggtgatg*ATGGCTCCCGTGCAGGATGC
GH61_6_F	TTTAAGCTT**GAAACG**ATGCTCGCCCCCGTCCTTGTCC
GH61_6_R	TTTACTAGTTTA*gtgatggtgatggtgatg*TCACCCGCTCCAGACAGCAGGG

Table 6: Other used primers.

Zeo_fw (P08688)	GCTGGTGCTGTTGAGTTCTG
Kan_Mitte_fw (P09156)	GCAATCACGAATGAATAACGG
B1GAP_PUC_fw	CACACAGCCCAGCTTGGAGC
KpnI_ADH_F(P09170)	ATATAGGTACCGCATGCCCTACAACAACTAAG
seq.EGFP.fwd (P12012)	CGGACACAAACTTGAGTACAAC
seq1_GAZ2	CATGCTCCTACAACATTACCA
seq2_GAZ2	GGCTCATGGCGTTAAAGAGG
seq3_GAZ2	CCGTACATTTAGCCCATACAT
seq4_GAZ2	CGCCTCGACATCATCTGCC
pGAZ2 fw-SEQ_A	GGGAGAGGTGCGATAATC
pGAZ2_rv	GCATCGACTTTTTCAATAACCA
PUC_fw	CCTGGCCTTTTGCTGGCCTTTTGCTCAC
GA_LEU2_fw	GCCACGTTGTCGACCTTCAAGAGG
GA_LEU2_rv	GAGCGTTGAATAGACGCAGCAC
GA_LEU1_fw	CGTTGAGAAGAAGCGACCTGAG
GA_LEU1_rv	GATAGCTTCAGCCTCCTTAGGAGC
GT_LEU1_fw	AAGCGCTGATAATTGAGGCTAC
GT_LEU1_rv	TCGGGTTGTCATCGTTACAATA
GT_LEU2_fw	CCATAACCAAACACCAAGGATT
GT_LEU2_rv	AGCTTCCACAAAGTCGAAAGAC
GT_HIS3_fw	AATATCTATTTCTCGGACGGCA
GT_HIS3_rv	CGAACTATACGGAAACACACCA

2.8. Kits

- GeneJET™ Plasmid Miniprep Kit, Thermo Scientific. Changes to manual: Cells were taken off an agar plate, grown over night and directly resuspended in the resuspension solution.
- Wizard® SV Gel and PCR Clean-Up System, Promega
- CloneJET™ PCR Cloning Kit, Thermo Scientific: For the transformation into *E.coli* 5 µl ligation mix were directly used.

2.9. Methods

2.9.1. Agarose gel electrophoresis

The gels contained 1.5 % agarose (Biozym) and were dissolved in 1x TAE by heating up in a microwave. Ethidium bromide was added to the cooled down agarose solution. As standard the GeneRuler™ 1 kb ladder (Figure 5) was used. The settings for the electrophoresis were 120 V for 50 minutes in case of a control gel and 90 V for 1.5 hours to 2 hours in case of a preparative gel. The bands were visualised with the GelDoc-It™ imaging system. The Chroma43 UV screen was used to cut out bands from the gel.

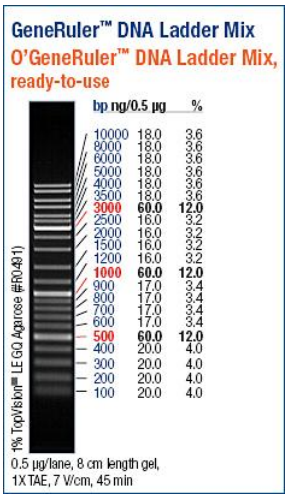


Figure 5: GeneRuler™ DNA ladder mix

2.9.2. PCR

The standard PCR mixture and temperature program is shown in Table 7 and Table 8.

Table 7: PCR-mixture

Component	Volume [µl]
5x Phusion HF buffer	10
dNTPs (2mM)	5
Forward primer (5 µM)	2
Reverse primer (5 µM)	2
Template (1ng/µl)	2
Phusion™ polymerase (2 U/µl)	0.3
ddH2O	Up to 50

Table 8: PCR temperature program

Temperature [°C]	Time [sec]	
98	30	
98	10	35 cyles
adjusted to primers	20	
72	30/1 kb	
72	7	
4	∞	

2.9.3. Construction of PCR cassettes

For the construction of PCR cassettes suitable for integration into the *S. pombe* genome an overlap extension PCR (OE-PCR) was done. The primer design was performed with the help of the *Pombe* Primer PCR Program (PPPP) webtool (92). First the fragments used for the OE-PCR were amplified to introduce 40 nucleotides homology to the *S. pombe* genome as well as a linker sequence. Therefore, the expression fragment – consisting of *adh1* promoter, eGFP reporter gene and *ura4* terminator – and the selection marker – either against Zeocin™ or kanamycine – were amplified. In Table 9 the primers and templates needed for the fragment amplification are listed. The introduced linker combines the expression construct and the selection marker and its sequence is shown in Table 3.

Table 9: Templates and primers for the amplification of the expression and the selection marker fragment.

Amplified fragment	**Template**	**Homology**	**Forward primer**	**Reverse primer**
expression construct	pGAZ-eGFP	*leu2*	leu2-adh1_fwd	overlap_rev
expression construct		*leu1*	leu1-adh1_fwd	
expression construct		*his3*	his3-adh1_fwd	
selection marker against Zeocin™		*leu2*	overlap_fwd	leu2-tefTT_rev
selection marker against Zeocin™		*leu1*		leu1-tefTT_rev
selection marker against Zeocin™		*his3*		his3-tefTT_rev
selection marker against kanamycine	pGAK2	*leu2*		leu2-tefTT_rev
selection marker against kanamycine		*leu1*		leu1-tefTT_rev
selection marker against kanamycine		*his3*		his3-tefTT_rev

Two fragments, which combine to a ready cassette, were mixed in a 1:1 molar ratio and an OE-PCR was performed: In the first step of the OE-PCR no primers were added to the PCR mixture (Table 10). The same temperature program was used as described in Table 8, but the PCR cycles were reduced to 15 and the extension time was orientated at the lengths of both fragments.

Table 10: PCR mixture used in the first step of the OE-PCR

Component	**Volume [µl]**
5x Phusion HF buffer	10
dNTPs (2mM)	5
Expression cassette	Equimolare amounts of each (1 to 10 ng of the largest)
Selection marker	
Phusion™ polymerase (2 U/µl)	0.3
ddH_2O	Up to 50

In the second step of the OE-PCR primers were added to amplify the expression cassettes. The primers, listed in Table 11, introduced another 40 nucleotides of homology, resulting in 80 nucleotides of homology to the *S. pombe* genome in total. They were added with fresh polymerase and dNTPS in mixture listed in Table 12 and the temperature program was the same as for a usual PCR (Table 8).

Table 11: Primers used for OE-PCR.

Amplified fragment	**Forward primer**	**Reverse primer**
Cass_leu1-zeo	leu1_fwd	leu1_rev
Cass_leu2-zeo	leu2_fwd	leu2_rev
Cass_his3-zeo	his3_fwd	his3_rev
Cass_leu1-kan	leu1_fwd	leu1_rev
Cass_leu2-kan	leu2_fwd	leu2_fwd
Cass_his3-kan	his3_fwd	his3_rev

Table 12: PCR-mixtuer added in the second step of the OE-PCR.

Component	Volume [µl]
5x Phusion HF buffer	4
dNTPs (2mM)	2
Forward primer (5 µM)	2.8
Reverse primer (5 µM)	2.8
Phusion™ polymerase (2 U/µl)	0.3
ddH2O	Up to 20

A blunt end ligation of the OE-PCR products into the pJET vector was done. The vectors – containing, PCR cassettes as inserts, were transformed into *E. coli* Top 10 F' cells. The transformants were checked for right inserts and the sequence was verified by sequencing.

2.9.4. Construction of the integration vectors

For the construction of the integration vectors, the fragments for Gibson assembly (103) were amplified with the templates and primers listed in Table 14. The sequences of the primers are shown in Table 4. The DNA fragments had a 28 to 40 bp overlapping homology region, which enables efficient recombination.

Table 13: Templates and primers for the amplification of the fragments used for the assembly of the integration vectors. The features are as followed: The prokaryotic ori pUC (673 bp), the eukaryotic ori ars1 (868 bp), the two integration markers *leu1/2*-3'(400 bp/365 bp) and *leu1/2*-5'(300 bp/356 bp), the *adh1* promoter (P(*adh1*), 720 bp), the gene of Interest (GoI), the *ura4* terminator (T(*ura4*), 425 bp), and the Zeocin™ (zeo, 375 bp) or kanamycinie (kan, 810 bp) resistance gene under the control of the *tef* promoter (P(*tef*) 352 bp) and the *tef* terminator (T(*tef)*, 229 bp).

Number	Amplified fragment	Template	Homology to	Forward primer	Reverse primer
A1	P(*adh1*)-T*(ura4)*	pGAK2	*leu1*-3', *P(tef)*	Adh1_leu1-3'_fwd	overlap_rev[1]
A2	P(*adh1*)-T*(ura4)*	pGAK2	*leu2*-3', *P(tef)*	Adh1_leu2-3'_fwd	
A1-eGFP	P(*adh1*)-eGFP-T*(ura4)*	pGAZ2-eGFP	*leu1*-3', *P(tef)*	Adh1_leu1-3'_fwd	
A2-eGFP	P(*adh1*)-eGFP*T(ura4)*	pGAZ2-eGFP	*leu2*-3', *P(tef)*	Adh1_leu2-3'_fwd	
B1-zeo	*P(tef)*-zeo-T(*tef*)	pGAZ2	linker, *leu1*-5'	overlap_fwd[1]	teftt_leu1-5'_rev
B2-zeo	*P(tef)*-zeo-T(*tef*)	pGAZ2	linker, *Leu2*-5'		teftt_leu2-5'_rev
B1-kan	*P(tef)*-kan-T(*tef*)	pGAK2	linker, *leu1*-5'		teftt_leu1-5'_rev
B2-kan	*P(tef)*-kan-T(*tef*)	pGAK2	linker, *leu2*-5'		teftt_leu2-5'_rev
C1	*leu1*-5'	*S. pombe* gDNA	*teftt*, pUC	leu1-5'_teftt_fwd	Leu1-5'_pUC_rev
C2	*leu2*-5'	*S. pombe* gDNA	*teftt*, pUC	leu2-5'_teftt_fwd	leu2-5'_pUC_rev
D1	pUC	pGAK2	*leu1*-5', *ars1*	pUC_leu1-5'_fwd	pUC_ars1_rev
D2	pUC	pGAK2	*leu2*-5', *ars1*	pUC_leu_fwd	
E1	*ars1*	pGAK2	pUC, *leu1*-3'	ars1_pUC_fwd	Ars1_leu1-3'_rev
E2	*ars1*	pGAK2	pUC, *leu2*-3'		Ars1_leu2-3'_rev
F1	*leu1*-3'	*S. pombe* gDNA	*ars1*, P(*adh1*)	leu2-3'_ars1_fwd	leu2-3_adh1_rev
F2	*leu2*-3'	*S. pombe* gDNA	*ars1*, P(*adh1*)	leu1-3'_ars1_fwd	Leu1-3'_adh1_rev

[1]Sequences for overlap_fwd and overlap_rev can are the same ones as for the construction of the PCR cassettes and can be found in Table 3.

The master mix for Gibson assembly was prepared as described in Table 14. The mix was stored in 15 µl aliquots at -20 °C.

Table 14: Gibson assembly master mix. The 5x ISO buffer was prepared as described in 2.3.7.

Component	**Volume [µl]**
5x ISO buffer	320
T5 exonuclease (10 U/µl)	0.64
Phusion® High-Fidelity DNA Polymerase, (2 U/µl)	20
Taq DNA ligase, 40 U/µl	160
sterile ddH_2O	fill up to 1200

A 15 µl aliquot of the assembly master mix was thawed on ice. The fragments resulting in integration vectors (listed in Table 15) were mixed in a 1:1 molar ratio (50 to 100 ng of largest fragment were used) up to a maximal volume of 5 µl and were added to the master mix. The assembly was performed in a thermomixer at 50 °C for 60 minutes. The assembly mix (5 µl) was then directly used for transformation into *E. coli* Top10 F' cells. Plasmids were checked for right sizes by restriction digestion and their sequences were verified by sequencing.

Table 15: Fragments used for Gibson assembly and resulting vectors.

PCR fragment number[1]						**Resulting vector**
A1	B1-zeo	C1	D1	E1	F1	pGAZ2_*leu1*
A1-eGFP	B1-zeo	C1	D1	E1	F1	pGAZ2_*leu1*-eGFP
A1	B1-kan	C1	D1	E1	F1	pGAK2_*leu1*
A1-eGFP	B1-kan	C1	D1	E1	F1	pGAK2_*leu1*-eGFP
A2	B2-zeo	C2	D2	E2	F2	pGAZ2_*leu2*
A2-eGFP	B2-zeo	C2	D2	E2	F2	pGAZ2_*leu2*-eGFP
A2	B2-kan	C2	D2	E2	F2	pGAK2_*leu2*
A2-eGFP	B2-kan	C2	D2	E2	F2	pGAK2_*leu2*-eGFP

[1]For designation see Table 13.

2.9.5. Cloning of the cDNA targets of *F. fomentarius*

Primers were designed in the way that they introduce the Kozak sequence, a terminal his-tag as well as *Hind*III and *Spe*I resriction sites (see). As a template for the amplification the cDNA library of *F. fomentarius* was used. After the purification of the PCR products, the DNA fragments as well as the pGAZ2_*leu1* vector were cut with the two restriction enzymes HindIII and *Spe*I. The cut and purified DNA fragments were ligated with the T4 ligase. 5 µl of

this ligation mixture were directly used for the transformation into *E. coli* Top10 F'. Positive clones were either checked by colony PCR or directly by restriction digestion, depending on how the transformation control (ligated empty vector) had worked. In order to check the inserts the positive clones were sent to sequencing.

2.9.6. Rapid gDNA isolation from single colony

This protocol is an adapted version of the gDNA isolation protocol by Harju *et al*.(104). A single yeast colony was resuspended in 100 µl lysis buffer in a 1.5 ml reaction tube. The cell suspension was frozen in liquid nitrogen for 2 minutes and then incubated in a thermomixer at 99°C for 1 minute. This freezing and cooking step was repeated twice. The tube was vortexed for 30 seconds and then 100 µl chloroform were added. Another vortexing step for 2 minutes followed. The cell suspension was centrifuged for 3 minutes at maximum speed. The upper aqueous phase was transferred into 400 µl ice cold 100 % ethanol and precipitated for 15 minutes at -20°C. The tubes were centrifuged for 3 minutes at maximum speed and the supernatant was discarded. The resulting pellet was washed with ice cold 70 % ethanol. Then the pellet was dried at 37 °C and resuspended in 30 µl water.

2.9.7. Colony PCR for *E. coli*

For colony PCR of *E. coli* single colonies were picked with 10 µl pipette tip, streaked on a master plate and then resuspended in 20 µl PCR mixture (Table 16). The temperature profile for the PCR was applied as shown in Table 17.

Table 16: PCR-mixture for colony PCR using Go Taq polymerase.

Component	Volume [µl]
5x Go Taq buffer green or colourless	4
dNTPs (2mM)	2
Forward primer (5pM)	0.8
Reverse primer (5pM)	0.8
Go Taq polaymerase	0.1
ddH_2O	Up to 20

Table 17: PCR temperature program used for Go Taq polymerase.

Temperature [°C]	Time [sec]	
95	30	
95	30	25 cyles
58 (adjusted to primers)	30	
72	60/1 kb	
72	5	
4	∞	

2.9.8. Colony PCR for yeast

For colony PCR 50 to 100 ng gDNA of a single colony was taken. Apart from that the PCR was performed as described in 2.9.2.

2.9.9. Electrocompetent *E. coli* cells

A preculture was inoculated in LB medium with fresh streaked *E. coli* cells (50 ml LB in 250 ml glass flask, 28°C-30°C, 120 rpm). After at least 12 hours the main culture was inoculated with 15 ml of the preculture and grown to an OD of ~0.8 (500 ml in 2 l flasks, 37°C, 200 rpm). Cells were collected in sterile centrifugation tubes (500 ml) and centrifuged (2000x g, 15 minutes, 4°C). The supernatant was discarded and washed 2x (resuspend gently) in 500 ml of sterile, ice-cold 10 % (w/v) glycerol (2000x g, 15 minutes, 4°C). The pellet was resuspend in 1 ml of ice-cold 10 % glycerol (1/50 of original culture volume). 80 µl aliquots were frozen at -80°C.

2.9.10. Electroporation of *E. coli*

80 µl electrocompetent cells were thawed on ice and mixed with 25-100 ng of the plasmid DNA. The mixture was transferred to a pre-chilled electroporation cuvette and incubated on ice for 5 minutes. Immediately after pulsing in the electroporator (2.50 kV, 1 pulse) 1 ml of SOC media was added and the cells were regenerated at 37°C for 1 hour at 700 rpm. After regeneration, cells were plated on selective medium.

2.9.11. Electrocompetent cell preparation of *S. pombe*

A preculture was inoculated in YES medium with preferably fresh streaked cells (50 ml in 250 ml baffled glass flask, 28°C-30°C, 120 rpm). After at least 12 hours the main culture was inoculated to an OD of 0.2 and grown for ~5 hours to an OD of 0.8 (200 ml in 2 l flasks, 28-30 °C, 120 rpm). Cells were collected in sterile centrifugation tubes (500 ml) and centrifuged (1600x g, 5 minutes, 4°C). The supernatant was discarded and the cells were resuspended in ice cold DTT buffer (1/10 of the original culture volume). The cell suspension was transferred into Greiner tubes (50 ml) and incubated for 15 minutes at 30°C and 150 rpm. The resulting pellet was washed 3 times with ice cold 1 M sorbitol (1600x g, 5 minutes). The pellet was finally resuspended in ice cold 2 M sorbitol (1/50 of original culture volume) and 300 µl aliquots were frozen at -80°C.

2.9.12. Electroporation protocol for *S. pombe*

The frozen electrocompetent cells were quick thawed in a thermomixer (2 minutes, 30°C). The cells were pelleted and washed twice in ice cold 1 M sorbitol by centrifugation (1600x g, 2 minutes, 4°C). The final pellet was resuspended in 80 µl ice cold 1 M sorbitol. The cell suspension was mixed with plasmid DNA (1-2 µg) and transferred to a precooled electroporation cuvette (0.2 cm electrode gap). A high electric pulse (2 kV) was applied and 0.5 ml ice cold 1 M sorbitol was immediately added. The transformed cells were transferred to a 1.5 ml reaction tube and incubated at 4°C for 1.5 h on ice. 0.5 ml YES and was added and incubated for 1.5 h in a thermomixer (30°C, 700 rpm). Cells were plated on YES or EMM2 with adequate antibiotic.

2.9.13. Lithium acetate (LiAc) method adapted by Bähler *et al.*(105)

An overnight preculture was grown in YES full medium with fresh streaked *S. pombe* wild type cells (50 ml in 250 ml baffled glass flask 28°C-30°C, 120 rpm). If the cells were older than one week, the preculture was inoculated about 24 hours in advance. The next day a main culture was started with an optical density (OD) of 0.2 which was then grown for about 5 hours to an OD of 0.8 (50 ml in 250 ml baffled glass flask, 28°C-30°C, 120 rpm). A 50 ml culture was sufficient for approximately 8 transformations. For harvesting the cells were transferred to a sterile Greiner tube (50 ml) and centrifuged (1600x g, 4°C). The supernatant

was discarded and the pelleted cells carefully washed with cold sterile water. The cell suspension was centrifuged and the pellet was washed carefully in ~1 ml LiAc-TE buffer. The cells were resuspended in 800 µl LiAc-TE buffer. 100 µl cells were mixed with 2 µl carrier DNA into a 1.5 ml reaction tube and up to 10 µl of DNA was added. The cells were incubated for 10 minutes at room temperature. 260 µl of the PEG-LiAc-TE solution was added and everything was mixed. This cell suspension was incubated in a thermomixer for 30-60 minutes at 29°C-30°C. 43 µl pre-warmed DMSO were added and everything was mixed gently. Afterwards the cells were heat shocked in a thermomixer at 42°C for 5 minutes. The cells were pelleted and washed once with 1 ml water. Then the cells were resuspended in 1 ml YES media and regenerated for 1 to 4 hours at 30°C. The cells were plated on selective medium.

The adapted version differs from the original protocol in the regeneration step: Transformed cells were grown on YES plates and regenerated there for ~18 hours before replated onto YES plates containing the antibiotic.

2.9.14. Electrocompetent cell preparation of *P. pastoris*

The preparation of electrocompetent *P. pastoris* cells was performed according to Lin-Cereghino *et al.*(24)

2.9.15. Electroporation protocol for *P. pastoris*

The cells were always freshly prepared. 100 µl of the electrocompetent cells were mixed with linearized plasmid DNA or PCR product (1-2 µg) and transferred to a precooled electroporation cuvette (0.2 cm electrode gap). A high electric pulse (1.5 kV) was applied and 0.5 ml ice cold 1 M sorbitol was immediately added. The transformed cells were transferred to a 12 ml regeneration tube and incubated at ice for 15 minutes. 0.5 ml YPD was added and incubated for 2 hours at 28 °C. Cells were plated on selective YPD medium.

2.9.16. 96 deep well cultivation of *S. pombe*

Colonies were picked from the selection plates and cultivated in 300 µl YES containing 100 mg/L Zeocin™ for 60 hours in 96 deep well plates. EMM2 is also suitable for deep well cultivation, but for leucine and histidine auxotrophic strains supplementation has to be regarded and the minimal medium is also not suitable in combination with Zeocin™.

2.9.17. eGFP fluorescence measurement

The cells were transferred in a microplate suitable for fluorescence measurements and diluted in order to reach an optical density between 0.2 and 0.8. In most cases 20 µl yeast culture were diluted into 180 µl deionised water. Since the construct harbours an eGFP reporter gene, the fluorescence of the transformants was measured with an excitation wavelength of 485 nm and detected at an emission wavelength of 520 nm.

2.9.18. Shaking flask cultivation of *S. pombe*

Transformants, which gave a positive signal in the screening, were streaked on YES plates containing 100 mg/l Zeocin™. Single colonies were inoculated in 5 to 10 ml YES medium and grown over night in 50 ml Greiner tube. This overnight culture was used to inoculate 200 ml YES in a 2 l baffled glass flask and then the cells were grown for 60 hours at 30°C at 110 rpm.

For harvesting the cells were centrifuged for 5 minutes at 4000x g and the supernatant was decanted and again centrifuged in order to ease the filtration. Since the proteins are supposed to be involved in lignocellulose degradation, interactions with a cellulose membrane were suspected. Therefore, filters with a 0.2 µm PES membrane were used instead of those with cellulose membranes.

2.9.19. TCA-acetone precipitation

90 µl protein solution was mixed with 10 µl 100 % TCA (w/v) and was then precipitated for 1 hour at 4°C. Afterwards 1 ml ice cold acetone was added and centrifuged for five minutes at maximum speed. The pellet was washed with ice old acetone and then dried at 37°C for at least half an hour.

2.9.20. Protein purification on an ÄKTA purifier system

For protein purification a cooled ÄKTA purifier system with sample collector was used. For the nickel affinity purification three buffers were used: Wash buffer, elution buffer and desalting buffer. If the supernatant was directly loaded on the column (100 ml) an external pump was used with flow rate of 2 ml/min. If the supernatant was concentrated, the buffer was exchanged with binding buffer and then the protein solution (~2 ml) was loaded by loop injection onto the column. The flow rate during the purification process was 5 ml/min and the maximal pressure was set to 0.3 MPa. The method was performed as described in Table 18. For the purification of the proteins two different nickel affinity columns were used: First the 5 ml HisTrap FF column (GE Healthcare) and secondly the 5 ml HisTrap™ excel (GE Healthcare).

Table 18: Ni-chelate purification protocol for ÄKTA purifier.

Step	Column volumes	Buffer	Comment
Equilibration	7	Wash buffer	-
Sample load	-	-	-
Wash	7	2 % elution buffer	15 ml fractions
Gradient	7	From 10 to 100 % elution buffer	1.5 ml fractions

2.9.21. SDS-PAGE

For sample preparation, 4x loading dye was added to maximal 18 µl protein sample and heated up to 70°C for 10 minutes to denature the proteins. Precast gels and MOPS buffer were put in the electrophoresis chamber. The samples were loaded in the slots next to 5 µl of PageRuler™ prestained protein ladder (Figure 6). The gel was run for 1 hour and 10 minutes at 200 V and was stained for 30 minutes with Coomassie blue SDS gel staining solution. For destaining the gel was either put in SDS gel destaining solution for 12 hours or heated for 2 minutes in the microwave and incubated on a shaker.

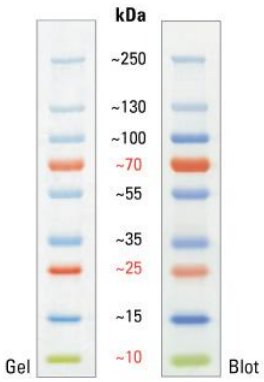


Figure 6: Band profile of the PageRuler Plus prestained protein ladder

2.9.22. cDNA target screening

The colonies for the screening were picked from the transformation plate and cultivated in a 96 deep well plate. The cells were grown in 300 µl YES medium and harvested after 48 hours. The first antibody binds to the his-tag and the second antibody is conjugated with a horse radish peroxidase, which converts luminol to a chemoluminescent product. As positive control an *E. coli* cell lysate from a strain that expresses a his-tagged protein as well, was used. To clarify if a background signal is caused by the medium or the cells two negative controls had been applied: First a *S. pombe* strain having the empty integrated vector control (vec ctrl) and second the medium alone.

2.9.23. Dot Blot

The membrane for the dot blot was soaked for 5 minutes in TBS buffer before placed into the dot blot microfiltration apparatus. The membrane was dried by applying vacuum to the dot blot apparatus. 15 µl supernatant of a 96 deep well cultivation were transferred onto the membrane and the whole apparatus was incubated at 60°C for 1 hour. Then the supernatant was removed by applying vacuum and the membrane was extracted out of the apparatus. Washing steps followed: First the membrane was washed twice in ~20 ml buffer and then twice in TBS buffer each time for 7 minutes. Afterwards the membrane was incubated for 1 hour in blocking buffer. The membrane was washed as before and then incubated for 1 hour in the first antibody solution. Again the membrane was washed and the second antibody solution was added. Finally the membrane was washed a last time and afterwards soaked for detection in 4 ml SuperSignal West Pico substrate. The first antibody binds to the C-terminal his-tag and the second antibody is conjugated with a horse radish peroxidase,

which converts luminol to a chemoluminiscent product, which was detected in the g-box. The preparation of the solutions is shown in 2.2.3.

2.9.24. Western Blot

A SDS-Gel was assembled as followed in a sandwich set-up: Three pieces of felt, filter paper, SDS-gel, membrane, filterpaper and again three pieces of felt were put in the Xcell II module. The settings for the blot were 25 V, 160 mA for 1 hour. The transfer buffer and other materials necessary for the Western Blot are listed in 2.2.3. The detection followed the same protocol as for the dot blot.

2.9.25. eGFP fluorescence measurement

Relative fluorescence units were determined by transferring 20 µl yeast culture into 180 µl deionised water in black microplates with clear bottom. The fluorescence of the transformants harbouring eGFP as reporter gene was measured with an excitation wavelength of 485 nm and was detected at an emission wavelength of 520 nm.

2.9.26. Enzymatic assay

A 50 µl substrate mix was added to 55 µl protein solution (if needed diluted with 25 mM Na-acetate buffer, pH 4.8) in a black microplate and incubated at 37 °C. As controls, buffer and supernatant of a cultivation of a *S. pombe* strain carrying the empty pGAZ2_*leu1* vector, was used. The fluorescence is measured at different time points and the experiment was stopped after 16 hours. The fluorescence was determined in the Spectramax GeminiXS plate reader. It was measured with an excitation wavelength of 395 nm and was detected at an emission wavelength of 507 nm.

2.9.27. Complementary assay

The substrate solution for the complementary assay consists of 24.5 µl enzyme mix (exact composition unknown, provided by Clariant), 2.5 % straw (provided by Clariant) and 875.5 µl 100 mM sodium acetate buffer (pH 5). This mixture represents one aliquot, which was

stored at -20°C. One aliquot was thawed and mixed with 50 to 200 µl purified protein und filled up to a final volume of 1800 µl with 100 mM sodium acetate buffer (pH 5). The mix was incubated at 50°C and 1000 rpm for 48 hours and samples were regularly drawn as followed: The tubes were shortly centrifuged in order to pellet the straw and 250 µl were transferred into another tube. This sample was inactivated at 99°C for 7 minutes. The samples were analyzed for released reducing sugars by DNS assay.

2.9.28. DNS assay

40 µl of 50 mM sodium acetate buffer (pH 5) and 20 µl sample were put in a 96 well PCR plate. Afterwards 120 µl DNS reagent were added and everything was mixed via pipetting. The plate was carefully sealed with a silicon mat and incubated in a PCR cycler at 95°C for 5 minutes. 36 ml of the incubated sample was transferred into a microplate and mixed with 160 µl H_2O. The absorption was measured at 540 nm. For a calibration curve glucose dissolved in 50 mM sodium acetate buffer (pH 5) in a concentration range of 1 to 10 mg/ml was measured.

2.9.29. Bioreactor cultivation

The bioreactor cultivation was performed in the DASGIP system with a maximal volume up to 1.4 l. A pH as well as a pO_2 electrode were connected to the system. The fermentation was base regulated (12.5 % (v/v) NH_3-solution) in order to maintain a pH of 6 in the fermentation vessels. The partial pressure of oxygen (pO_2) – set to a minimum concentration of dissolved oxygen of 30 % – was controlled with the stirrer velocity. The fermentation had two phases: In the batch phase biomass was built up and in the fed batch phase glucose was fed into the reactors in order to reach high cell densities for the protein production. The media used for the two phases is described in 1.2.6.

The fermentation was performed with the *S. pombe* wild type strain expressing IDN10 E6 with its native DNA sequence and leader peptide. A preculture was used to inoculate batch medium. Each medium pair was used twice and treated equally until the fed batch was initiated. Then the feeding was fixed to a pulsed mode supply, where the fed batch medium

was either brought in with 4 ml/h or 12 ml/h for half an hour. After a pause of one and a half hour the cycle started again.

2.9.30. Codon optimization

The codon usage tables used for the redesign of the gene sequences are based on highly transcribed genes in both yeasts and are shown in Abad *et al.* (106) for *P. pastoris* and in Forsburg *et al*. (77) for *S. pombe*. The program Gene Designer from DNA 2.0 was used for the codon optimization. Parameters like G-C content (30-70 %) and few AT repeats were considered. Moreover, rarely used codons and energy rich RNA secondary structures were as far as possible avoided. Native leader sequences were predicted by using the webtool SignalP (100) and were replaced with ones which already proofed to work in both expression hosts: For *P. pastoris* the α-factor leader peptide was used whereas for *S. pombe* the Carboxypeptidase Y leader peptide sequence was chosen. For both hosts the same Kozak sequence (cgaaacg) was attached. The sequences were ordered at GeneArt, the sequences are shown in the appendix in the section 7.4 and 7.5.

In the case of *P. pastoris* the native occurring *Xho*I restriction site in the alpha factor and the *Not*I restriction site on the multiple cloning site were used to clone the synthetic genes into the pPpB1_GAP_α vector. Thereby, the α-factor was attached to the synthetic genes. The optimized genes for *S. pombe* on the other hand were ordered with the *Cpy* leader peptide. HindIII and SpeI into the pGAZ2_leu1 vector.

3. Results and discussion: Expression systems in *S. pombe*

Episomal protein expression in *S. pombe* is known to be mitotically unstable (31). Therefore, we aimed to find methods for the integration of expression constructs into the *S. pombe* genome. Two integration strategies were chosen: On the one hand, linear PCR cassettes for direct transformation into *S. pombe* and vectors suitable for chromosomal as well as episomal integration on the other hand (105; 107). Both strategies are based on homologous recombination. Integration loci are associated with amino acid synthesis and correct integration leads to auxotrophic strains.

3.1. Integrative PCR cassettes for gene deletion in *S. pombe*

3.1.1. Cloning of PCR-derived integration cassettes

Primers were designed with the help of the PPPP web tool (52). This program facilitates the design of primers for the deletion or tagging of any gene in the *S. Pombe* genome. In Figure 7 the amplification steps of the integrative PCR cassettes are schematically depicted. As templates for the PCR cassettes the pGAZ2 and the pGAK2 vectors were used(1). The complete cassette is assembled by two fragments. The first fragment consists of the gene of interest (GoI) – in this case eGFP under the control of the *adh1* promoter and the *ura4* terminator – and the second fragment consists of the resistance gene, which is either aginst kanamycin (kan) or Zeocin^{TM} (zeo) under the control of the *tef* promoter and terminator. As first step 40 nucleotides (nt) of homology to the untranslated region (UTR) of the integration locus (*leu1, leu2 or his3)* were introduced to the integration marker on the flanking sides of the cassettes. The 28 base pairs long linker region in the middle of the construct proved to be a good universal linker for overlap extension PCRs (OE-PCRs) beforehand. The OE-PCR fulfils two functions: On the one hand, it combines the fragments to the finished cassette and on the other hand, it introduces another 40 nt of homology resulting in a total amount of 80 nt of homology on each side of the construct. According to Bähler *et al.* (1993) 60 to 80 nt should be sufficient to drive homologous recombination in *S. pombe*. The PCR cassette deletes the targeted gene in the genome of *S. pombe* and leads to an auxotrophy.

Cass_*leu1*-zeo (3114 bp)	Cass_*leu1*-kan (3549 bp)
Cass_*leu2*-zeo (3114 bp)	Cass_*leu2*-kan (3549 bp)
Cass_*his3*-zeo (3114 bp)	Cass_*his3*-kan (3549 bp)

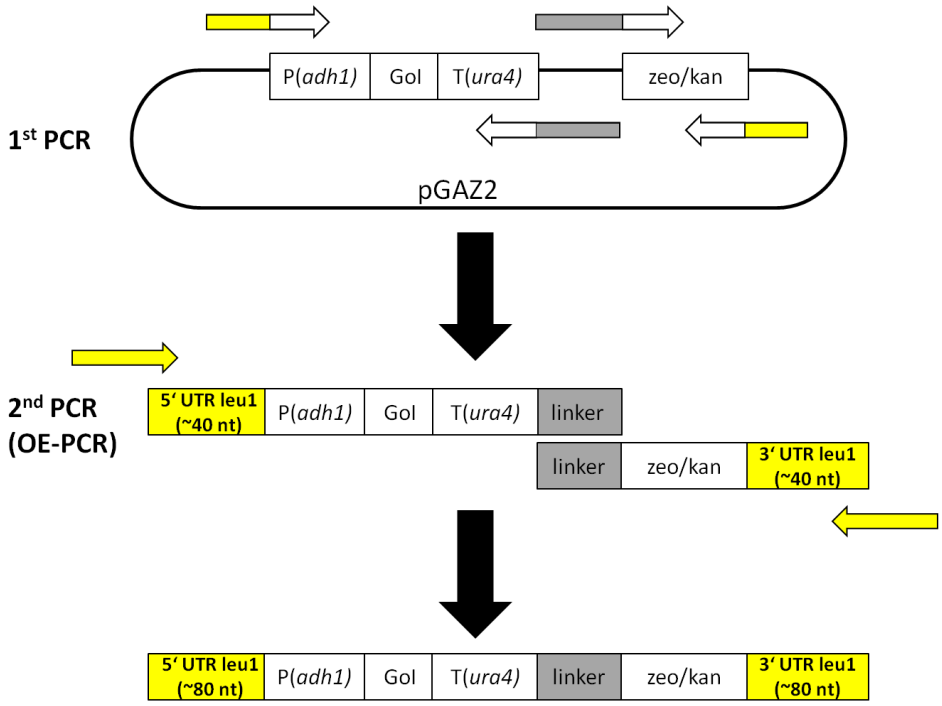


Figure 7: Construction of the linear PCR cassette for gene targeting of *leu1* in *S. pombe*. Arrows represent primers and the colours on the primers indicate homologies. The overlapping region is a universal linker with 28 base pairs. The features are the adh1 promoter (P(*adh1*), 720 bp), the gene of Interest (GoI), the *ura4* terminator (T(*ura4*), 425 bp), the linker (28 bp) and the Zeocin™ (zeo, 375 bp) or kanamycin (810 bp) resistance gene under the control of the *tef* promoter (352 bp) and the *tef* terminator (229 bp).

PCR products of the correct size (Figure 8A) were used for the OE-PCR. Afterwards, fully assembled PCR cassettes (Figure 8B), indicated by the correct size, were cloned into the pJET™ vector and transformed into *E. coli*. Positive transformants were chosen by restriction digestion analysis and the correctness was verified by sequencing.

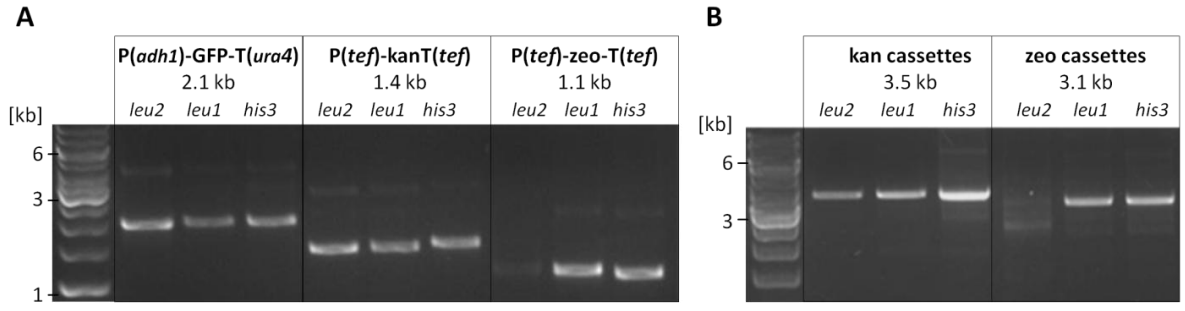


Figure 8: Agarose gel (1 %) picture from: **A** the amplified fragments for integrative PCR cassettes and **B** complete PCR cassettes obtained by OE-PCR.

3.1.2. Transformation and integration of PCR cassettes into *S. pombe*

The PCR cassettes have an 80 nt long homology to the UTR of the targeted gene (*leu1, leu2* or *his3*) and if correctly integrated the cassettes substitutes this gene. This recombination event has been described as a so called ends-out recombination and requires two cross overs, which are marked as dashed crosses in Figure 9 (108; 109). In this case "ends-out" refers to the regions of homologies, which are pointing out during the integration event. In contrast the later described integration vectors show an ends-in configuration, where the regions of homology are pointing inwards when they are integrated (see Figure 15).

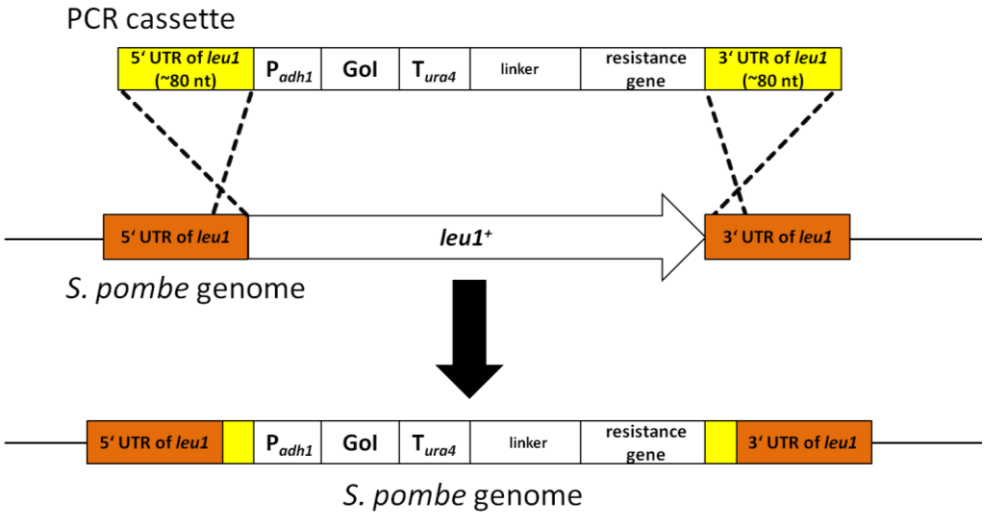


Figure 9: Strategy for chromosomal integration of the PCR-derived expression cassettes. The purified PCR product can be directly used for the transformation. The 80 nt long flanking regions of the PCR cassette are homologous to the UTR of the *leu1* gene. The integration of the PCR cassette leads to the deletion of *leu1* and leucine auxotrophy.

The three expression cassettes with Zeocin™ resistance gene were amplified from the sequenced pJET vectors. The purified PCR cassettes could be directly used for the transformation in a *S. pombe* wild type strain. The lithium acetate method described by Bähler *et al.* (1998) was used. After 3 days colonies were visible on the selection plates: In case of the *leu2* integration locus no transformants could be obtained. However, the *leu1* performs better than the *his3* locus, because 130 colony forming units per µg DNA (cfu/µg DNA) were obtained with cass_*leu1*_zeo whereas only 31 cfu/µg DNA with cass_*his3*_zeo. After 60 h growing, the transformants were stamped on different agar plates to verify correct integration. Positive integrands are growing on full (Figure 10A) but not on minimal medium plates (Figure 10B). Integration into the correct locus could be determined by

supplementing minimal medium (EMM2) agar plates with histidine or leucine. All *his3* deficient clones could grow on minimal medium supplemented with histidine (Figure 10C) and all *leu*$^{-}$ deficient clones on minimal medium supplemented with leucine (Figure 10D). This corresponds to an integration efficiency of 100 %.

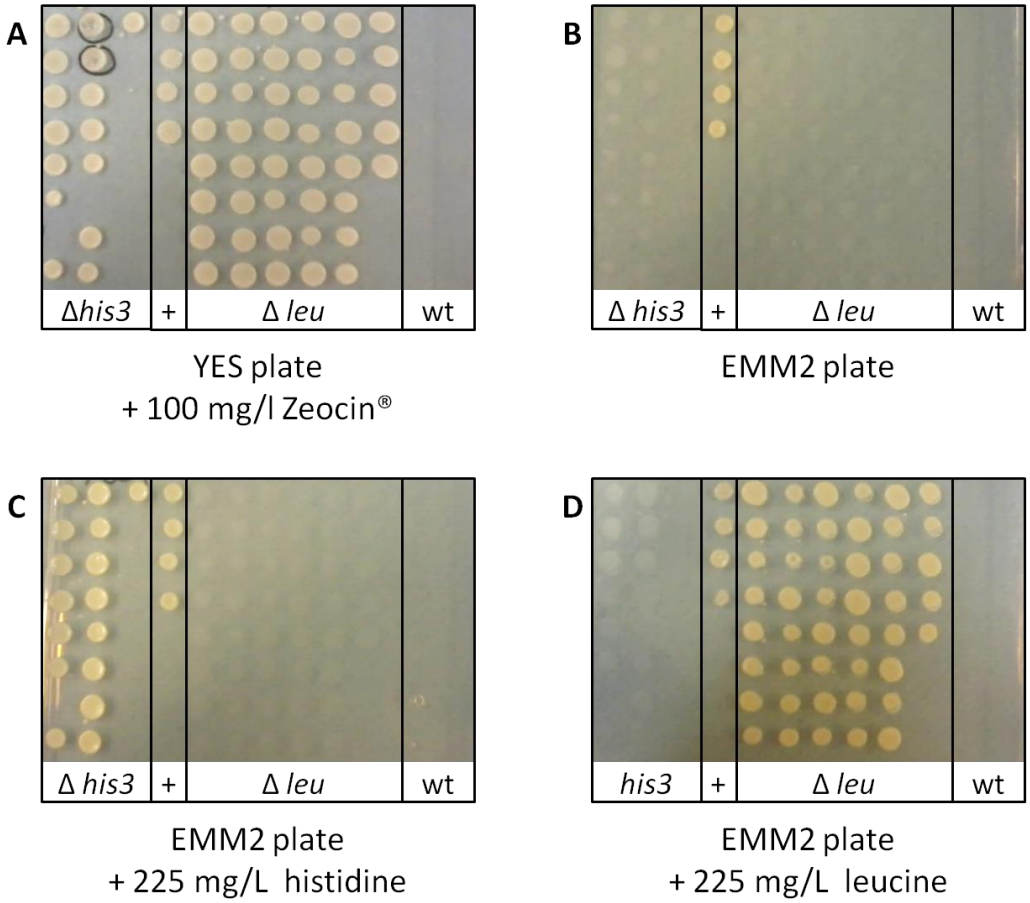


Figure 10: Integration of cass_*his3*_zeo and cass_*leu1*_zeo into *S. pombe* wild type strain. First the colonies were grown in a 96 deep well plate in 400 µl rich medium (YES) containing 100 mg/L Zeocin™ for 60 h. Then they were stamped on different agar plates to verify the integration into the correct locus. As a control a *S. pombe* strain containing the pGAZ2 vector (+) was cultivated on the same plate. The transformation of the cass_*his3*_zeo leads to a *his3* knock out (Δ*his3*) and a histidine auxotrophic strain. The transformation of the cass_*leu1*_zeo leads to a *leu1* knock out (Δ*leu1*) and a leucine auxotrophic strain. **A** shows the transformants on a YES agar plate containing 100 mg/L Zeocin™. All transformants with a functional expressed Zeocin™ resistance gene are growing on this plate. In **B** the same plate on an EMM2 agar plate (minimal medium) is shown. Strains with an auxotrophy are not growing and as expected only the control strain is growing. In **C** the EMM2 agar plate is supplemented with 225 mg/L histidine. All Δ*his3* transformants, which grew on plate A, are here again growing. The same can be seen in **D**, where the EMM2 plate with 225 mg/L leucine is shown. All Δ*leu1* are again growing on the leucine supplemented minimal medium. This corresponds to an integration efficency of 100 %.

3.1.3. Fluorescence measurements of *S. pombe* strains harbouring the integrated PCR cassette

The cells from the previous cultivation were used for further experiments. Since the constructs contained eGFP as reporter gene the fluorescence was measured. The RFU/OD values of the transformants were multiple times higher than the empty pGAZ2 vector (vec ctrl). The integration locus does not seem to influence the expression rate of eGFP, because the fluorescence intensity was in both cases in the same range.

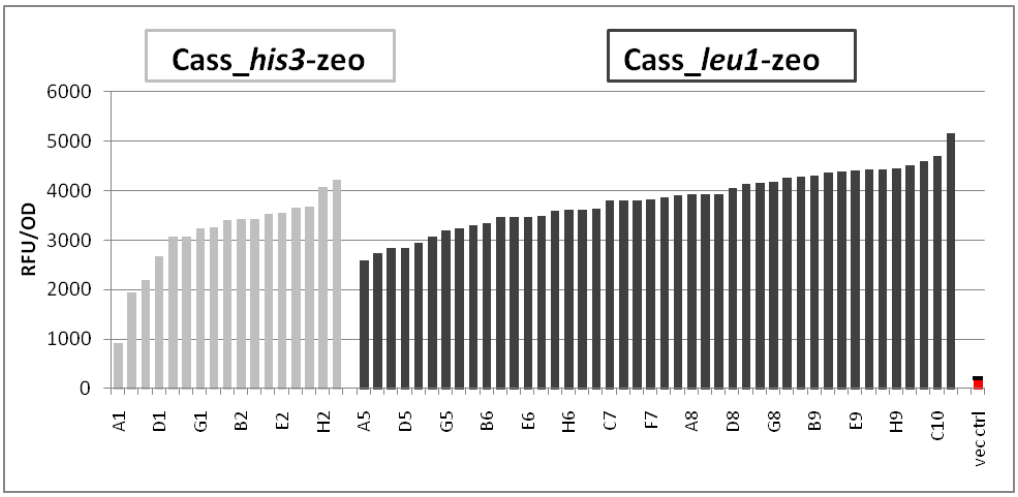


Figure 11: Fluorescence measurement (eGFP) of *S. pombe* strains with PCR cassettes cass_*his3*-zeo and cass_*leu1*-zeo. As a control a *S. pombe* strain containing the pGAZ2 vector (vec ctrl) was cultivated on the same plate.

3.2. Integration vectors

3.2.1. Construction of the pGAZ2 and pGAK2 integration vectors for *S. pombe*

The pGAZ2 and the pGAK2 vectors served as starting point for the new integration vectors. In these vectors the prokaryotic ori pUC and the eukaryotic ori *ars1* are separated by both, the GoI and the resistance gene. However, for the later integration vectors a different assembly was chosen, where the two oris are side by side, in order to have the possibility to cut them out in a single restriction digestion. To accomplish this structure the pGAZ2 as well as the pGAK2 vector had to be reassembled. In addition it was necessary to insert homologous sequences for the integration into the *S. pombe* genome. This reconstruction was a 2-step process: First the single fragments were amplified and then assembled with Gibson cloning to obtain the integration vectors. The construction is shown in Figure 12. The resulting integration vectors still have the *ars1* – the autonomous replication sequence of *S. Pombe* – and are therefore still able to function as episomes. Integration of the vector requires a digestion step (*Bgl*II or *BamH*I). The *leu1* and the *leu2* loci were chosen as integration markers resulting in four new vectors:

- pGAZ2_*leu1* (Zeocin™ resistance)
- pGAZ2_*leu2* (Zeocin™ resistance)
- pGAK2_*leu1* (kanamycin resistance)
- pGAK2_*leu2* (kanamycin resistance)´

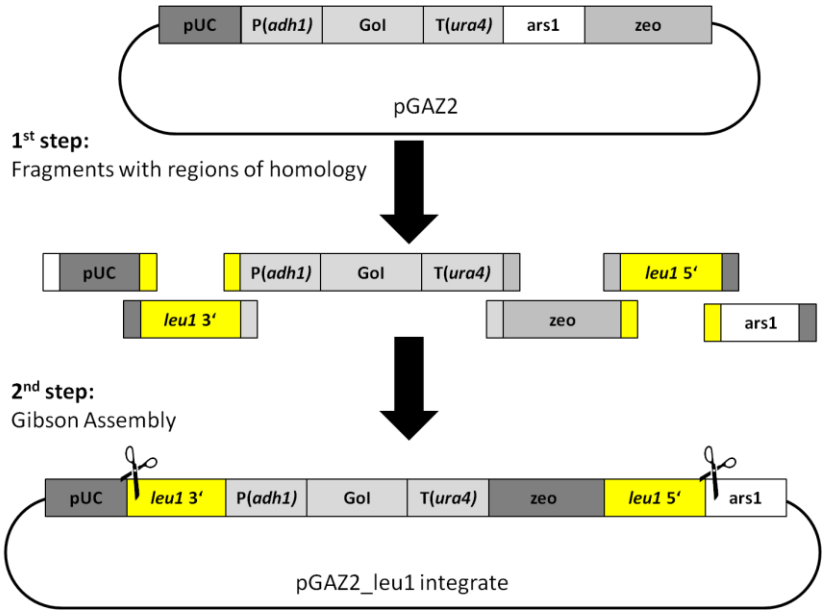
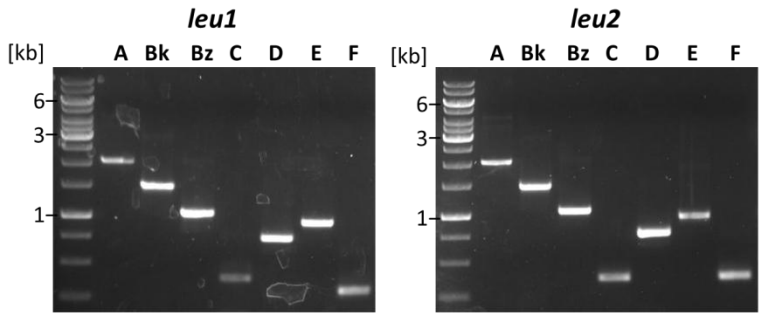


Figure 12: Construction sheme of the pGAZ2_*leu1* integration vector for *S. pombe*. In the first step the single fragments of the pGAZ2 vector and the integration marker (*leu1*) are amplified. Overlapping boxes indicate homologies to the adjacent fragments. In the second step the fragments are connected with Gibson assembly. The integration construct can be cut by restriction digestion, which is indicated by the scissors. The features are as followed: The prokaryotic ori pUC (673 bp), the eukaryotic ori *ars1* (868 bp), the two integration markers *leu1*-3'(400 bp) and *leu1*-5'(300 bp), the *adh1* promoter (P(*adh1*), 720 bp), the gene of Interest (GoI), the *ura4* terminator (T(*ura4*), 425 bp), and the Zeocin™ (zeo, 375 bp) or kanamycin (810 bp) resistance gene under the control of the *tef* promoter (352 bp) and the *tef* terminator (229 bp).

For the construction of the new integration vectors the pGAZ2 vector was used as template to amplify the prokaryotic ori *pUC*, the *ars1* and the Zeocin™ resistance gene. For the amplification of the kanamycin resistance gene the pGAK2 vector was used. Both resistance genes are under the control of the *tef* promoter and terminator. The GoI – in this case the reporter gene eGFP – is under the control of the *adh1* promoter and the *ura4* terminator. The integration markers are at least 356 bp long and the gDNA of the *S. pombe* wild type strain was used as template for their amplification. The fragments were purified and monitored on an agarose gel (Figure 13).



Fragment		Size [bp]
A	P(*adh1*)-eGFP-T(*ura4*)	2021
Bz/Bk	P(*tef*)-zeo/kan-T(*tef*)	1056/1491
D	pUC	753
E	ars1	948
C	*leu1/leu2*-5‘	300/356
F	*leu1/leu2*-3‘	400/365

Figure 13: Purified Fragments ready for Gibson assembly on 1 % agarose gel. The picture on the left shows the single fragments necessary for the assembly of the pGAZ2_*leu1* and the pGAK2_*leu1* vectors. The agarose gel picture in the middle shows the single fragments necessary for pGAZ2_*leu2* and pGAK2_*leu2* assembly. In the table on the right side the fragments and their sizes are listed.

The amplification of the fragments introduces at least 32 bp of homology to the sequences of the adjacent fragments. These homologies are necessary for Gibson assembly, a fast one pot technique. Multiple fragments can be assembled by the cooperation of three different enzymes: the T5 exonuclease, the Phusion polymerase and the T4 ligase (103). The assembled vectors were isolated and tested with restriction digestion (Figure 14). Only one plasmid out of 16 did not show right sizes. This proofs that Gibson assembly is a reliable method for the assembly of many fragments. Plasmids showing right fragment sizes on the control gel were sent to sequencing to LGC genomics. The results proved the correctness of the vector sequences.

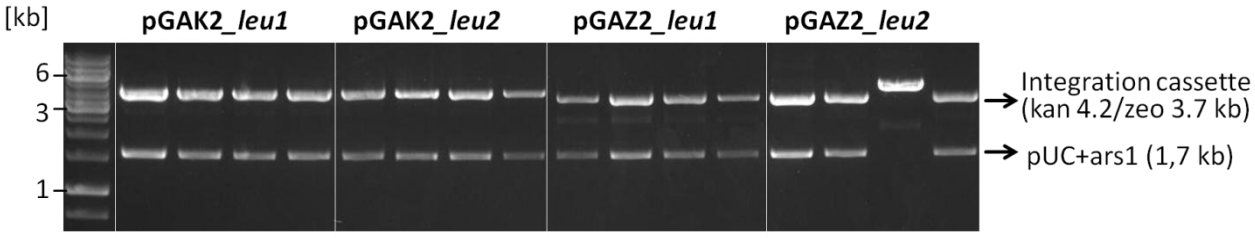


Figure 14: Restriction digestion of integrative vectors with *Bgl*II. The integration cassette and the oris should be checked. 15 out of 16 picked plasmids were correct assembled. Only the plasmid next to the last does not show right fragment sizes.

3.2.2. Transformation of the *S. pombe* wt strain with integration vectors

In the case of the integration vectors a so called ends-in integration occurs. In contrast to the PCR cassettes that utilizes the ends-out integration strategy – where the expression construct substitutes the whole targeted gene (see Figure 9). – here a single cross over leads to an insertion of the expression construct at the targeted locus (109). As discussed before "end-ins" refers to the regions of homologies, which are pointing inwards during the integration event.

The homologous patches on the integration vector are incomplete sequences of the *leu1* gene, resulting in two dysfunctional genes, named *leu1*Δ5' and *leu1*Δ3' (Figure 15). For this integration strategy no specific genetic background is needed and a laboratory wild type strain was used.

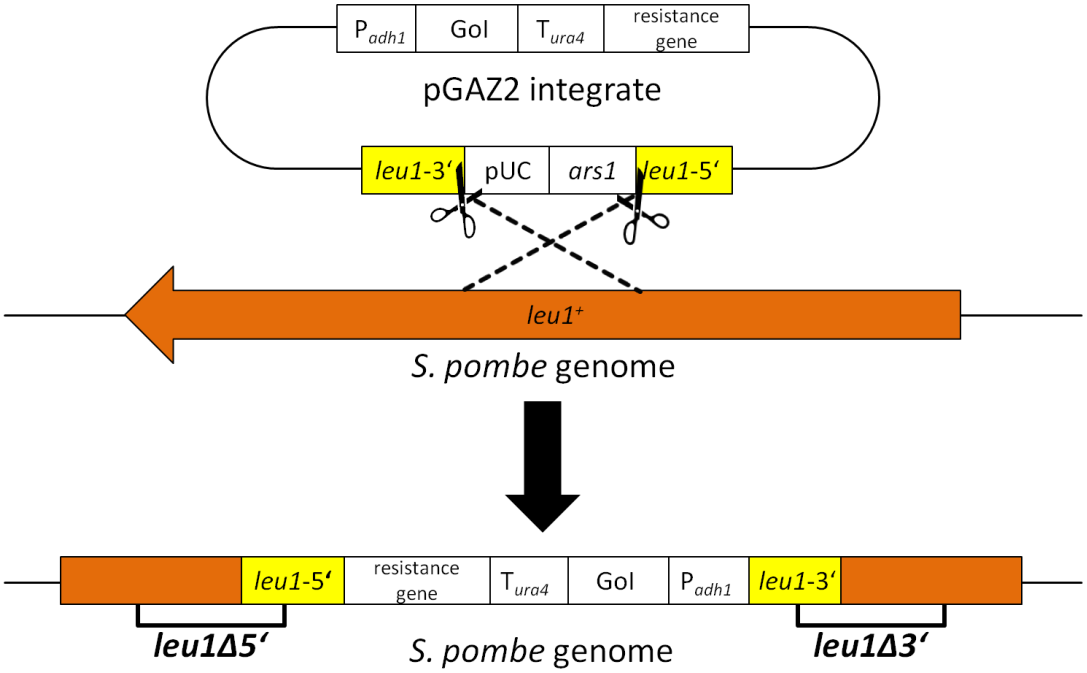


Figure 15: Strategy for chromosomal integration of the pGAZ2_*leu1* integration vector. The eukaryotic and prokaryotic oris, *ars1* and pUC respectively, can be cut out by restriction digestion, which is indicated by the two scissors. The targeted gene is *leu1* and the resistance gene is either against Zeocin™ or kanamycine. For this integration strategy a laboratory wild type strain was used. The integration of the expression construct results in *leu1* disruption and leucine auxotrophy. The *leu1* gene is transcribed from the lagging strand.

The *S. pombe* wild type strain was transformed with the pGAZ2_*leu1*-eGFP as cut linear construct. An adapted version of the lithium acetate transformation protocol, which is less

time-consuming, was used. The transformants were cultivated in 96 deep well plates for 60 hours. As controls two additional *S. pombe* strains were cultivated on the same plate: On the one hand the empty pGAZ2_*leu1* (vec ctrl) and on the other hand an eGFP positive control (+) – a strain that contains the integrated pGAZ2_*leu1*-eGFP but derives from a previous transformation that already proofed to show high fluorescence. In order to check the integration efficiency the cells were stamped on different agar plates. Among these plates were also EMM2 plates with and without leucine supplementation (Figure 16). The integration of the expression cassette leads in the ideal case to a *leu1* dysfunction and hence to a leucine auxotrophy. Therefore, the transformants should not grow on the EMM2 plates without added leucine. 78 out of 88 picked transformants were auxotrophic. This corresponds to an integration efficiency of 89 %. The remaining ten transformants – showing normal growth on the minimal medium plate – appeared greenish and the fluorescence measurement confirmed this visual impression. This transformants come either from heterologous recombination, which is a very common event in *S. pombe* (110) or from not entirely cut plasmids. An isolation of the gDNA and a PCR for checking the integration is necessary to confirm the integration efficiency of 89 %.

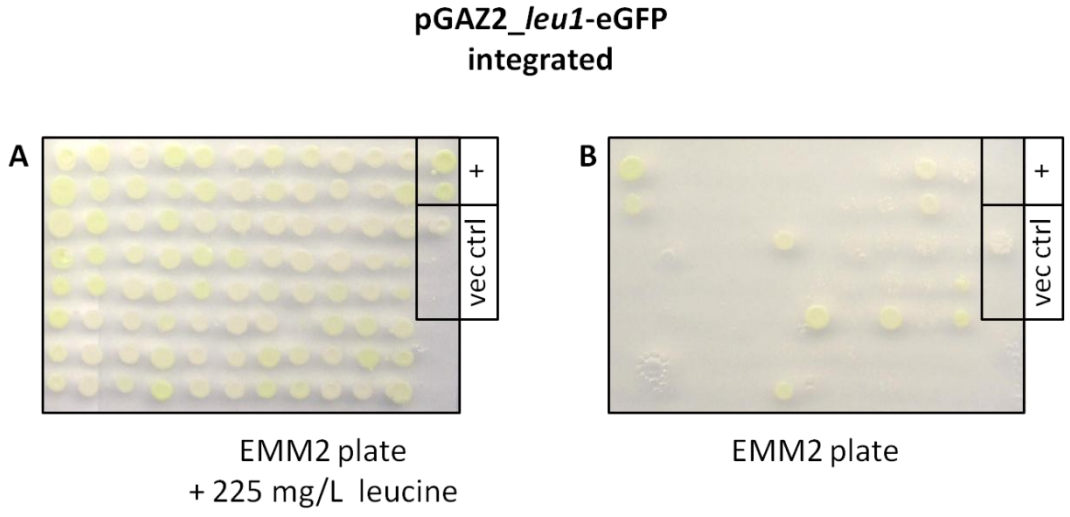


Figure 16: The pGAZ2_*leu1*-eGFP was transformed as linear construct into *S. pombe* wild type strain. As controls a *S. pombe* strain containing the pGAZ2_*leu1* vector (vec ctrl) and an eGFP control strain (+), which is the same construct from a previous transformation, were cultivated on the same plate. The transformation of the pGAZ2_*leu1*-eGFP leads to a *leu1* disruption and a leucine auxotrophic strain. **A** shows the transformants on a EMM2 agar plate supplemented with 225 mg/L leucine. All transformants are supposed to grow on this plate. In **B** the same plate on an EMM2 agar plate is shown. Strains with a leucine auxotrophy should not grow on this plate. Only 10 out 88 picked clones do not exhibit a leucine auxotrophy. This corresponds to an integration efficiency of 89 %.

The pGAZ2_*leu1*-eGFP was also transformed in its episomal form and was processed the same way as linearized plasmid above. The cells were stamped on different plates to see if integration of the circular plasmid into the *leu1* locus took place. In contrast to Figure 16, where the majority of the transformants was not growing on the EMM2 plate, Figure 17 shows that all transformants were growing with leucine supplementation and also without it. This indicates that the origin of replication, the *ars1*, prevented the integration of the plasmid into the genome.

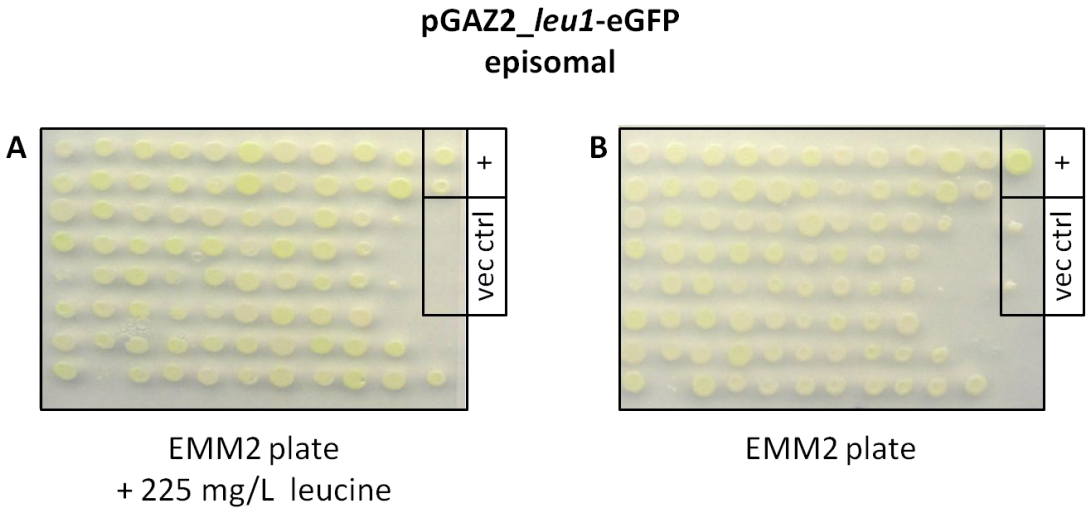


Figure 17: *S. pombe* wild type was transformed with pGAZ2_*leu1*-eGFP (episomal). As controls a *S. pombe* strain containing the pGAZ2_*leu1* vector (vec ctrl) and an eGFP control strain (+), which is the same construct from a previous transformation, were cultivated on the same 96 deep well plate. **A** shows the transformants on a EMM2 agar plate supplemented with 225 mg/L leucine and **B** the same plate only on an EMM2 agar plate (minimal medium). No leucine auxotrophy of the transformants can be seen.

This plate experiment was also performed with the pGAZ2_*leu1* empty vector and the pGAZ2_*leu2*-eGFP vector – both integrated and episomal – as well as with the pGAZ2-eGFP vector. In all experiments similar results were obtained: Linear constructs led mainly to auxotrophic strains, whereas episomal ones did not. The pictures of all plates can be found in the appendix in section 7.1 and the integration efficiency is listed in Table 19.

Table 19: Summary of transformation and integration efficiencies. Transformants were determined as integrated, when they exhibited a leucine auxotrophy.

Construct name	Transformation	Integration efficiency [%]
pGAZ2_leu1 (integrated)	$2.0 \cdot 10^3$	65 (30/44)
pGAZ2_leu1 (episomal)	$1.2 \cdot 10^3$	-
pGAZ2_leu1-eGFP (integrated)	$4.0 \cdot 10^3$	89 (78/88)
pGAZ2_leu1-eGFP (episomal)	$1.9 \cdot 10^3$	-
pGAZ2_leu2-eGFP (integrated)	$1.0 \cdot 10^3$	85 (75/88)
pGAZ2_leu2-eGFP (episomal)	$6.4 \cdot 10^3$	-
pGAZ2-eGFP (episomal)	$4.3 \cdot 10^4$	-

3.2.3. Screening of eGFP expression

The transformants of the 96 deep well cultivations were not only used to stamp on different plates but also to measure the fluorescence of the expressed eGFP. The results for the integrated and the episomal expressed pGAZ2_*leu1*-eGFP are shown in Figure 18. The green bar indicates the positive control (eGFP control). As before the integrated construct of a previous transformation was used. The vector control is represented by the small red bar and contained the integrated pGAZ2_*leu1*, which showed fluorescence in the range of the background only.

The cells were cultivated in rich medium containing Zeocin™ to maintain the episomal expressed vector stably in the cell. The plasmid is supposed to be present as multicopies whereas the integrated construct is expected to be single copy (107). The fluorescence should correlate with the copy number. Hence, it was expected to get higher values in case of the episomal expressed eGFP. Nevertheless, the integrated version results in the same range of relative fluorescence units per OD (RFU/OD).

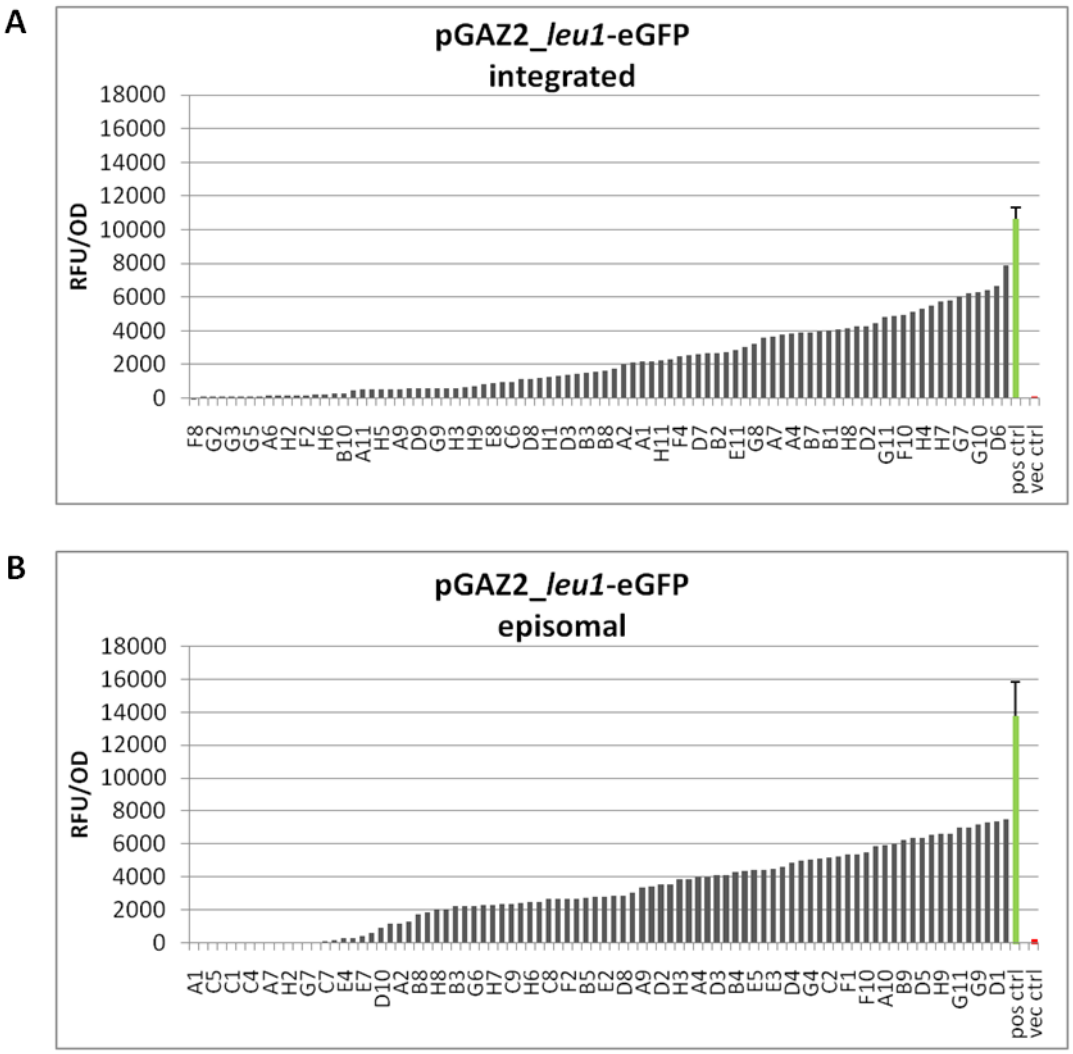


Figure 18: Fluorescence measurement of the *S. pombe* transformants. **A** shows the integrated pGAZ2_*leu1*-eGFP and **B** the episomally expressed pGAZ2_*leu1*-eGFP transformants. As controls a *S. pombe* strain containing the pGAZ2_*leu1* vector (vec ctrl, red) and an eGFP control strain (pos ctrl, green), which is the same construct from a previous transformation, were cultivated on the same plate. The controls were picked four times, whereas the picked transformants were cultivated in one well.

Noticeable is also the appearance of the two landscapes – integrated and episomal – which was very similar. The landscapes are both showing a broad variety of transformants: The values are constantly increasing from actually no fluorescence to high RFU/OD. However, for the integration a more even distributed landscape was expected, because the literature states that single copy integration occurs(107). However, the data indicate that a multicopy integration of the linear construct took place. The eGFP screening was also done for the other constructs and the landscapes can be found in the appendix.

In order to compare all constructs, rescreens with chosen transformants in different media were performed. As media EMM2 – supplemented with leucine – and YES, both with and without Zeocin™ were used. The overall pattern is in all cases the same, but the standard deviations were the lowest with leucine supplemented EMM2 as shown in Figure 19. On the other hand, the RFU/OD values but also the standard deviations are generally higher if grown in YES. The addition of Zeocin™ distresses the cells additionally and hence also influences the reliability of the cultivation in the 96 deep well plate.

The constructs chosen for the rescreens were the pGAZ2-eGFP and both integration vectors (*leu1* and *leu2*) which as well expressed eGFP. The light gray bars in Figure 19 are indicating those strains where the eGFP is episomally expressed, whereas the dark grey bars and the only green one are showing strains where the construct is expressed from the genome of *S pombe*. Considering only the episomal expressed constructs there was no significant difference between the expression levels. That indicates that the rearrangement of the pGAZ2 vector and the insertion of the integration markers had no influence on the expression of eGFP. However, looking on the whole rescreen result there was a difference between episomal and integrated expressed values. The integrated levels are at least two times higher than the episomal ones. As already mentioned before, this was not expected. The result from the rescreen 1 is another indication that several copies of the expression cassette are integrated into the genome. One explanation for the lower values of the episomal expressed eGFP, might be the missing selection pressure. However, the same trend can also be observed in the medium with Zeocin™.

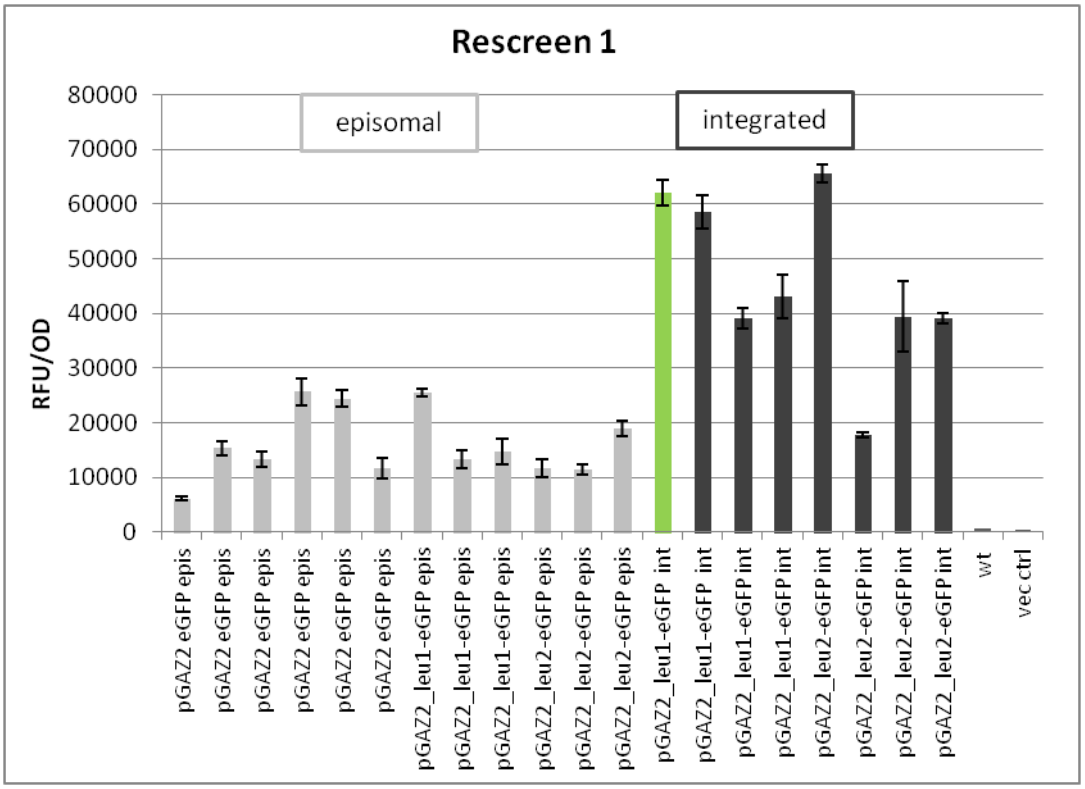


Figure 19: Rescreen 1 in EMM2 medium supplemented with leucine. The fluorescence of the expressed eGFP was measured after 42 hours. The constructs chosen for the rescreen were the pGAZ2-eGFP and both integration vectors pGAZ2_*leu1*-eGFP and pGAZ2_*leu2*-eGFP. Episomally expressed constructs are designated with "epis" and constructs from genome are designated with "int". As controls served the wild type strain (wt) and the pGAZ2 empty vector control (vec ctrl). The green bar is the transformant, which was used as positive control used in later experiments.

In the rescreen 1 only the best performing transformants from the screening were chosen. However, as figures above indicate for the integrated constructs single as well as multi copy integration most probably occurs. Therefore, a second rescreen was done with transformants more representative over the whole range of transformants.

For the rescreen 2 shown in Figure 20, one weak, one mediocre and the two strongest (W, M and S) eGFP expressing transformants per construct were chosen. In order to see if those transformants behave reproducible, one transformant was inoculated four times in EMM2 and the plate was measured after 36 hours. The transformants are derived from the screening shown in 0. As controls the empty integrated pGAZ2-*leu1* (vector ctrl) and the positive control (green bar), from the rescreen 1, were added.

For the pGAZ2-eGFP construct – the first four bars from the left in Figure 20 – the weak-mediocre-strong-strong pattern maintained in the rescreen. However, looking on the next

four bars, the pGAZ2_*leu1*-eGFP vector (episomal), the pattern is uniformly and the values seem to even out on the same level. This is probably due to plasmid loss, because there is no selection pressure in the medium. For the pGAZ2_*leu2*-eGFP vector (episomal) the two strongest strains show also the highest values. The integration of the expression constructs on the other hand should lead to stable expression eGFP and indeed, they behave as expected: The weakest give the weakest signal and the strongest the strongest. However, remarkable is the great difference between the weak transformant and the higher ones. The question occurs, if the weak ones are single copy and the other multi copy integrants. Moreover, it has to be clarified how these possible multi copy integrants are obtained either by homologous recombination or by heterologous recombination somewhere in the genome. The next section deals with the colony PCRs, of some chosen transformants in order to clarify these questions.

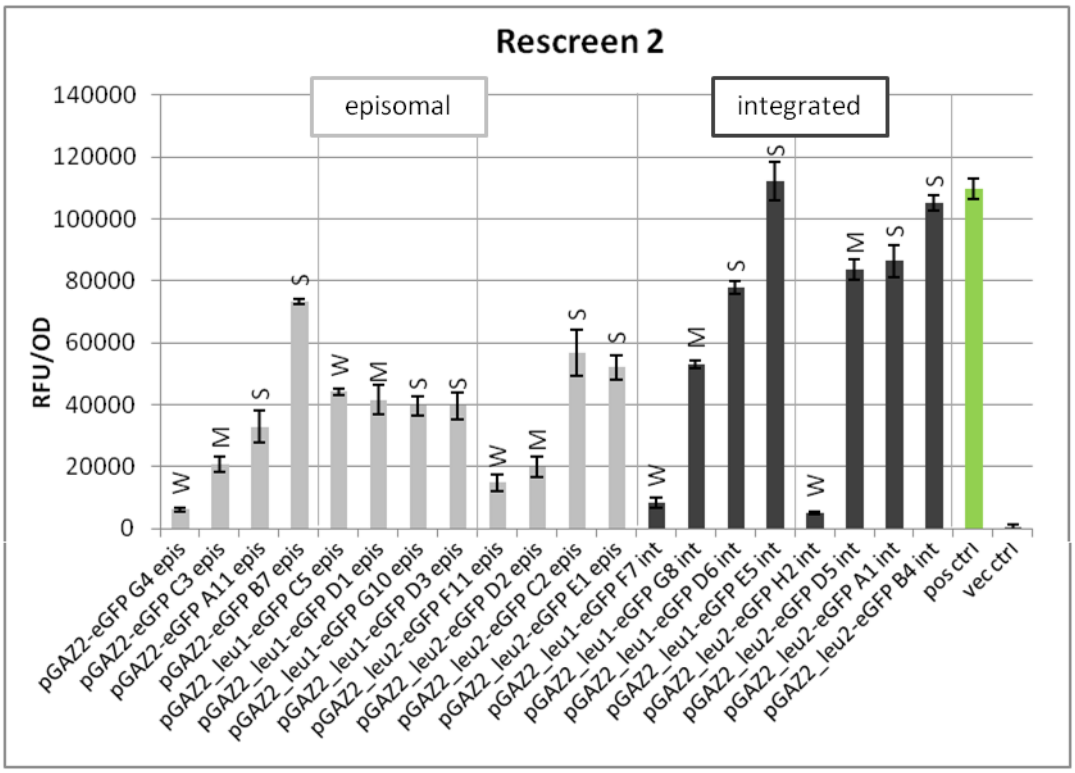


Figure 20**:** Rescreen 2 in EMM2 medium supplemented with leucine. The fluorescence of the expressed eGFP was measured after 36 hours. The constructs chosen for the rescreen were the pGAZ2-eGFP and the both integration vectors pGAZ2_*leu1*-eGFP and pGAZ2_*leu2*-eGFP. Episomally expressed constructs are designated with "epis" and constructs from genome are designated with "int". Controls were a strain having the empty the integrated pGAZ2-*leu1* (vector ctrl) and a positive control harbouring the integrated pGAZ2-*leu1*-eGFP (green bar). "W" stands for weak, "M" for mediocre and "S" for strong fluorescence in the screening.

3.2.4. Colony PCR

To clarify if multicopy integration occurs, the four pGAZ2_*leu1*-eGFP transformants from the rescreen 2 in Figure 20 were chosen for PCRs: One transformant exhibiting weak (W), one with mediocre (M) and the two with strong (S) RFU/OD values. The four transformants all show a leucine auxotrophy if grown on an EMM2 plate. In order to get reliable results the gDNA was isolated from a single colony with an adapted version of the rapid “Bust n' Grab” protocol (111). The obtained gDNA was used as template for the PCRs.

In Figure 21 the genomic situation of a single copy integrant and the agarose gel pictures of the colony PCRs are shown. As the orange triangles indicate two different primer pairs were chosen. To examine if the linear expression cassette has integrated at the right locus in the genome at all, primer pair 1 was chosen. One primer binds within the expression cassette and the other primer binds outside of the construct in the UTR of the *leu1* gene. The then performed PCR – shown in Figure 21B – gave results as expected: No fragments for the episomal and the wt strain, but a 2800 bp large fragment for each integrated transformants were detected. Hence, the expression construct integrated for all four transformants at the right locus.

To check the copy number of the expression cassettes primer pair 2 was chosen. The primers bind outside of the targeted locus and the size of the obtained fragment should indicate the copy number. In Figure 21B the outcome of this colony PCR is shown: As expected the two controls resulted in 2000 bp. A fragment with the size of 5300 bp would indicate a single copy integration. For every additional integration of an expression cassette 3500 bp is added to this fragment (2 copies: 8800 bp, 3 copies: 12300 bp,…). Only for the weak transformant a fragment with a size of 5300 bp was obtained. This explains the weak fluorescence signal that is indicating a single copy integration. The other three transformants show no band that could confirm a multi copy integration. However, due to the fact that the extension time of the PCR was only 5 minutes – which resembles 10 kb – most likely it was just too short to obtain evidence. Hence, it is still possible that more than 2 copies have been integrated. In order to determine the copy number a second PCR with a longer extension is necessary. If there is still no result the copy number can be determined for example with a quantitative PCR and heterologous recombination by genome walking.

Interestingly there is an unexpected 2000 bp band for the weak and the mediocre transformant with primer pair 2. As mentioned above a 2000 bp band corresponds to the size, you would obtain with the wt strain. It has to be clarified by sequencing this fragment if it derived from unspecific binding in the genome.

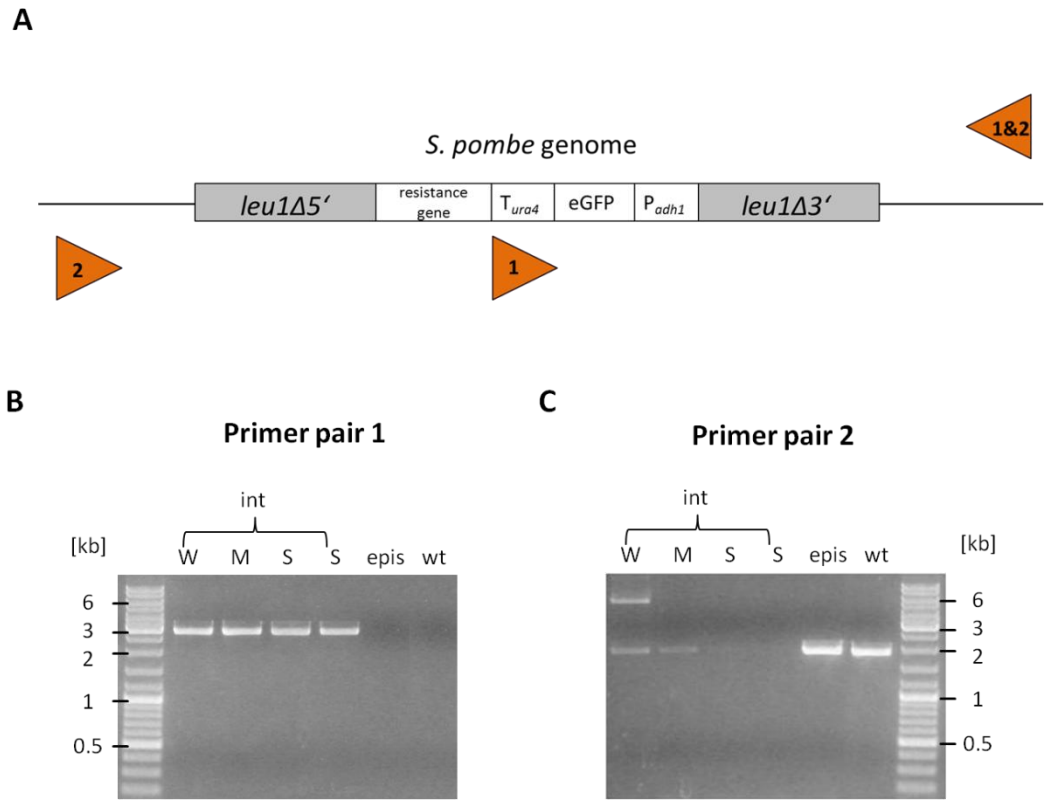


Figure 21: In **A** the *S. pombe* genomic situation at the *leu1* locus is shown after the pGAZ2_*leu1*-eGFP construct has integrated as a single copy. The orange triangles indicate the primers and are labelled according to the primer pairs. Primer pair 1 is binding to the 3' UTR and to the construct, whereas primer pair 2 is binding to the flanking UTRs of the *leu1* gene. Beneath, the PCR products obtained by primer pair 1 (**B**) and primer pair 2 (**C**) are shown. The abbreviations W are standing for weak, M for mediocre and S for strong fluorescence signal. All transformants have the integrated pGAZ2_*leu1*-eGFP (int). As controls a transformant containing the pGAZ2_*leu1*-eGFP as episome (epis) and the wild type strain (wt) were chosen. Expected sizes for primer pair 1 are 2800 bp for int and no band for wt and epis. Expected sizes for the primer pair 2 are 5300 bp for single copy integration and 2000 bp for wt and epis.

3.2.5. Extended rescreen

By extending the cultivation of rescreen 1 the plasmid stability of transformants containing the episomal construct was further tested. Thereby, 100 µl fresh medium was added after each measurement to prevent the cells from dying. Due to the plasmid loss without selection pressure decreasing values only for the episomal expressed eGFP are expected. The integrated expression on the other hand should stay the same. However, this assumption was falsified. As shown in Figure 22, the values – not only for the "episomal" ones but for all transformants – were significantly decreasing. Due to the fact that the 42 hours' values of the positive control are very similar to the ones obtained in later performed experiments, it is unlikely that a further loss of the integrated expression cassette occurs.

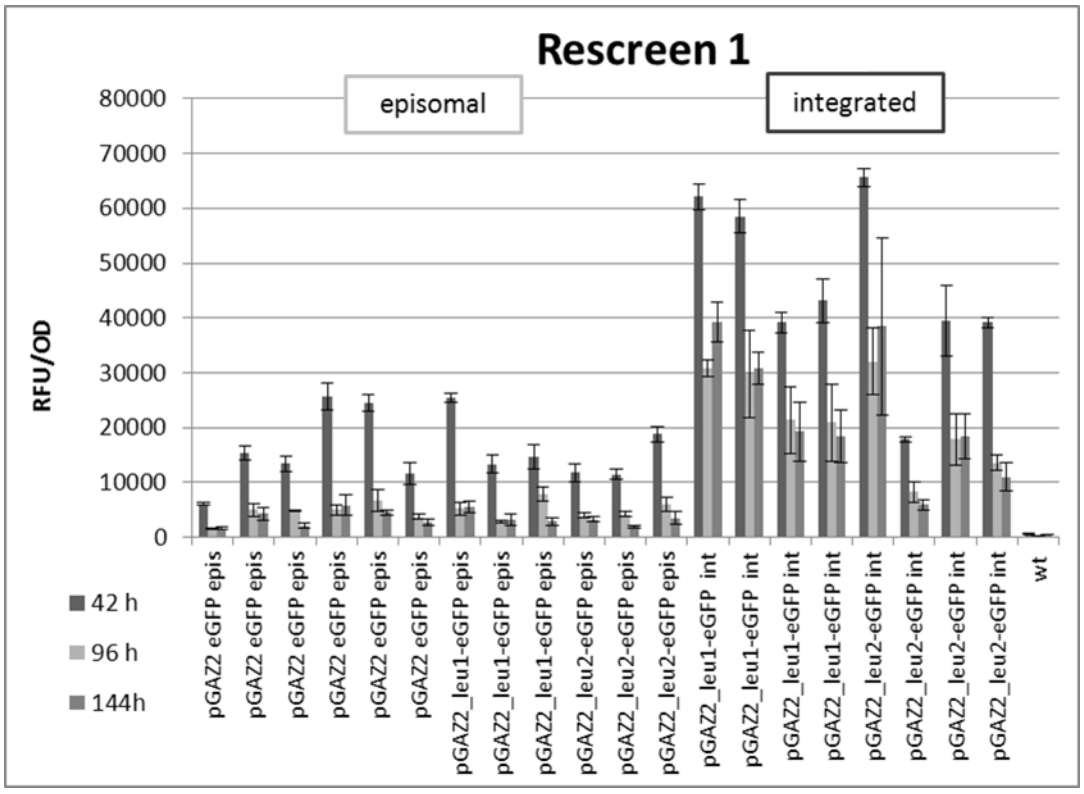


Figure 22: Time resolved diagram of rescreen plate 1. On the left side the transformants harboring an episome (epis) and on the right side the transformants with the integrated expression construct (int).

Although the *adh1* promoter is known to be constitutive (112), there might be regulation mechanisms that are responsible for the observed decrease of fluorescence achieved for all constructs after 42 hours. *S. pombe* has been designated as a Crabtree positive yeast, and produces ethanol under high glucose concentration. At the time of the first measurement –

after 42 hours of growth – the glucose in the medium is depleted. Later added medium did also contain glucose in its normal concentration (20 g/l). The volume of the cell culture, which had been already in the wells, was not considered. Therefore, the sugar concentration was diluted 3- to 4-fold – depending on evaporation and removed volume for the fluorescence measurement – to approximately 5 g/l. However, this is still a very high glucose concentration. Therefore, it is unlikely that the glucose concentration alone is responsible for the regulation of the *adh1*$^{+}$ promoter.

Looking on the situation in *S. cerevisae,* there are four alcohol dehydrogenases known (ADH1 to ADH4). ADH1 is resident in the cytosol and is the major enzyme in ethanol production under high glucose concentrations (113). Normally, the promoter is down-regulated in the late exponential phase, when ethanol has accumulated and the glucose had been consumed. A study revealed that the deletion of an 800 base pair stretch in the *ADH1* promoter results in constitutively expression throughout all growth phases (114). The full *ADH1* promoter is therefore said to be repressed by ethanol in the medium.

Maybe the *adh1*$^{+}$ promoter of *S. pombe* is similarly regulated as the *ADH1* promoter of *S. cerevisiae*. In this case *S. pombe* has to have at least one other alcohol dehydrogenase, which is responsible for the catabolization of ethanol in the later diacuxic growth phase. And indeed, a second alcohol dehydrogenase, encoded by the *adh4*$^{+}$ gene, had been reported by Sakurai *et al*.(115): Moreover, it seems that those two alcohol dehydrogenases are the only ones expressed by *S. pombe*: Knockouts of the two genes revealed that *adh1*$^{+}$ contributes the greatest part of alcohol dehydrogenase activity in *S. pombe* cell lysates, whereas the double knockouts revealed that both enzymes together constitute almost all activity.

The proposed regulation mechanism of the *adh1*$^{+}$ promoter is repression by ethanol. This would also explain the results shown in Figure 22. In the first 42 hours cells were expressing eGFP until ethanol had been accumulated in the cell culture. Then the amount of the very stable eGFP remained more or less the same until the end of the cultivation. The RFU/OD values decreased because the freshly added medium triggered further cell growth whereas the eGFP production stopped.

3.3. Methodical improvements

From experience in our group it is known that the pGAZ2 vector was not stable without selection pressure in the medium. Moreover, the overall expression of the reporter gene eGFP was very low. As it turned out, these two problems were mainly caused by using the electroporation protocol for transforming the DNA. This protocol is faster than the lithium acetate protocol and also the transformation efficiencies are higher, yielding 10^5 to 10^6 colonies per µg DNA in contrast to 10^3 colonies per µg DNA for the lithium acetate method (116). However, by using the lithium acetate method it seems that the copy number of the plasmid in the cell is increasing whereas the transformation efficiency is decreasing. The pGAZ2-eGFP plasmid was transformed by using different transformation protocols: Firstly, the electroporation protocol suggested in literature for the transformation of DNA in *S. pombe* (*S. p.* EP) (117). For this method electrocompetent cells are needed, which can be aliquoted and stored at -80°C for several months. The second method is the standard in-house electroporation protocol for the transformation of DNA into *P. pastoris* (*P.p*. EP). The third protocol is the lithium acetate protocol adapted by Bähler *et al*. (105) with the only difference that the same electrocompetent cells as for the *S.p.* EP were used. Finally, the fourth method was the usual lithium acetate protocol, where the cells are usually freshly prepared on the same day as the transformation is performed. All protocols are generally described in the section 2.9. In all cases, 1 µg plasmid DNA was used. The transformed cells were plated on YES medium plates containing 100 mg/l Zeocin™ and 20 transformants of each protocol were picked and cultivated together in EMM2 on the same 96 deep well plate. After 60 hours the cells were stamped on agar plates with and without Zeocin™. In Figure 23A the one without Zeocin™ is shown, where no selection pressure is put on the cells Therefore, it was expected that all transformants are growing. However, the transformants obtained by lithium acetate method using the electro competent cells (LiAc 1) did not grow at all, neither with nor without Zeocin™. This is no surprise, because as the OD measurements revealed the cells were not growing in the deep well plate in the first place. The reason for this is unknown, especially because the transformations and the later cultivations of cells using the same protocol were successful and gave similar results to the LiAC2 method. However, more interesting is that most of the transformants obtained by the two electroporation protocols are losing the plasmid during cultivation without selection pressure as the empty spots in Figure 23B are showing. The lithium acetate method (LiAc 2),

on the other hand, is maintaining the pGAZ2-eGFP vector throughout a cultivation period of 60 hours without selection pressure. This is also reflected in the measured fluorescence values, which are in average 5.4 times higher compared to *S. p.* EP and 8.4 times higher to *P. p.* EP at the end point measurement.

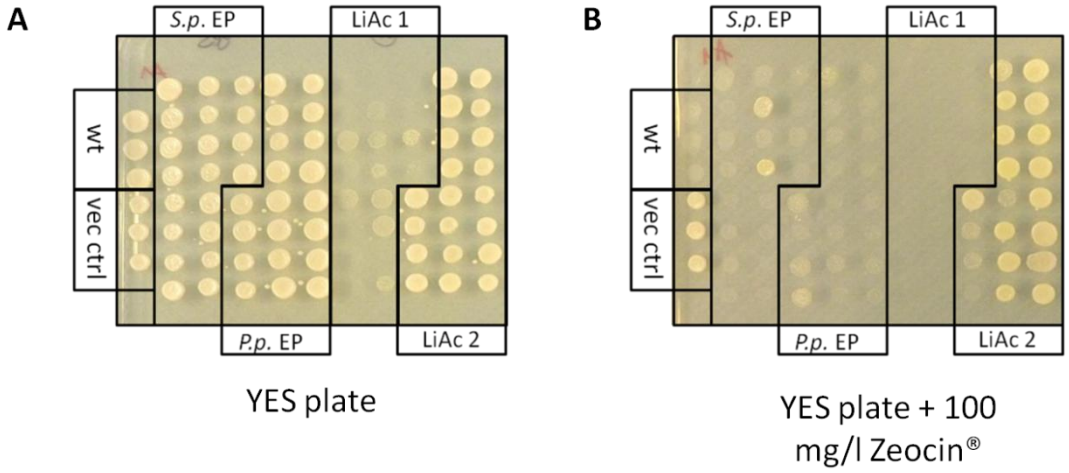


Figure 23: pGAZ2-eGFP transformed into *S. pombe* using different transformation protocols. Colonies were picked and cultivated in a 96 deep well plate without selection pressure and stamped after 60 hours on A a YES plate as well as on B a YES plate containing 100 mg/l Zeocin™. The transformants obtained by the *S. pombe* electroporation protocol (*S.p.* EP) and the *P. pastoris* electroporation (*P.p.* EP) are growing on the YES plate but hardly any on the plate with Zeocin™. The lithium acetate method with before frozen electrocompetent cells (LiAC 1) did not grow at all. Most transformants of the lithium acetate method with fresh cells (LiAc 2) on the other hand seem to maintain the Zeocin™ resistance.

Although the electroporation protocol used to transform DNA in *P. pastoris* is normally successful, the linearized pGAZ2 integration vectors were tested with both electroporation protocols in *S. pombe* and did not work. However, the transformants obtained by the lithium acetate method were stably integrated into the genome as expected.

4. Results and discussion: Expression of lignocellulolytic proteins in *P. pastoris* and *S. pombe*

4.1. Choosing cDNA targets

The aim of the project was the discovery of new enzymes for the degradation of lignocellulose. Therefore, the wood decaying fungus *Fomes fomentarius* was cultivated under inducing conditions in previous work in the group – i.e. the fungus was growing on wood. In order to obtain the proteins and enzymes expressed under these conditions, the RNA of the fungus had been isolated by colleagues and sent to the company LGC genomics, which sequenced the transcriptome and generated a cDNA library.

To discover new enzymes and proteins out of this cDNA library two approaches had been used: A screening of the library in the fission yeast *Schizosaccharomyces pombe* on the one hand and a bioinformatical analysis of the obtained cDNA sequences on the other hand. The project structure is depicted in Figure 24.

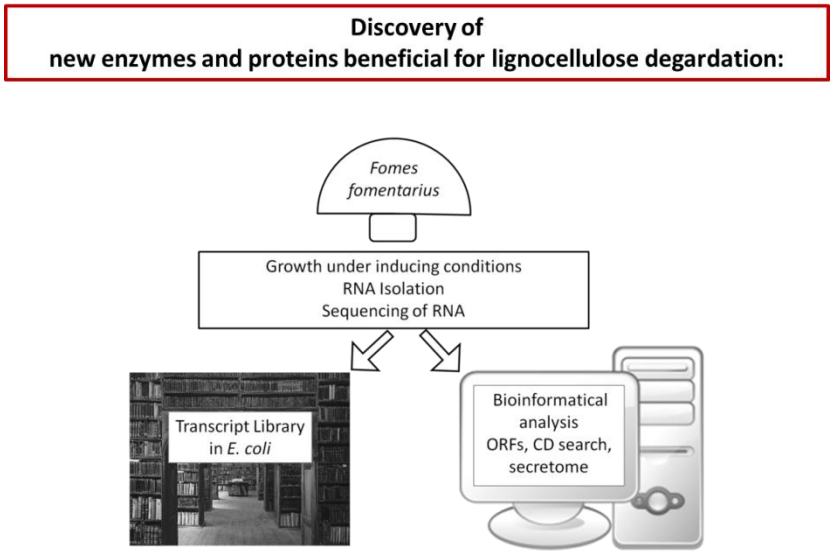


Figure 24: Strategies for the discovery of new enzymes and proteins for lignocellulose degradation.

The first approach – the screening of the cDNA library – was utilized by Mitrović (118). For the second approach a bioinformatical analysis was done in-house and an assembly was

provided (119). The assembly 2.60 was used for further working steps. The sequences were analysed and sorted according to criteria like complete open reading frames (ORFs), predicted signal peptides, conserved domains (CDs) and domains of unknown function (DUFs). To find new enzymes our approach rather focuses on DUFs than on domains with known enzymatic functions. Table 20 lists cDNA targets which were tested for secretion in *S. pombe* and *P. pastoris.* The screening in *S. pombe* was part of this thesis; hence the following chapter will focus on the expression in *S. pombe* in detail. To clarify if these proteins have an influence on the degradation of lignocellulose in general, a complementation assay was performed.

Table 20: *F. fomentarius* cDNA targets that were chosen to be expressed. Sequences can be found in the appendix.

Name	**Hosts**	**Contig**	**Conserved domain**
IDN5	*P. pastoris*	Isotig04428	Carbohydrate-binding domain (2x), at 5’ and 3’ end
IDN6	*S. pombe, P. pastoris*	Contig01698	WSC domain, this domain may be involved in carbohydrate binding,at 5’ end
IDN7	*P. pastoris*	Isotig04926	Glycosyl hydrolase catalytic core, at 5’ end
IDN8	*S. pombe*	Isotig05085	Glycine-rich protein domain (DUF2403) and putative TOS1-like glycosyl hydrolase (DUF2401), both at 3’ end
IDN9	*P. pastoris*	Isotig05070	Fungal-type cellulose-binding domain, at 3’ end
IDN10	*S. pombe*	Contig02057	Fungal-type cellulose-binding domain, at 5’ end
IDN13	*S. pombe, P pastoris*	Isotig02488	Lysin domain, at 5’ end
IDN16	*S. pombe*	Isotig03896	Fungal hydrophobin, at 3’ end
IDN24	*P. pastoris*	Isotig03302	Cupin, at the middle of ORF
IDN40	*S. Pombe, P. pastoris*	Contig01925	Glycoside Hydrolase family 61 (GH61), main part of ORF
GH61_1	*S. Pombe*	Isotig04623	GH 61, main part of ORF
GH61_2	*S. Pombe*	Isotig05034	GH 61, at the 5’ end
GH61_3	*S. Pombe*	Isotig03735	GH 61, 5’ end
GH61_4	*S. Pombe*	Isotig06156	GH 61, main part of ORF
GH61_5	*S. Pombe*	Contig01420	GH 61, at the 5’ end
GH61_6	*S. Pombe*	Contig01118	GH 61, main part of ORF

4.2. Cloning of the cDNA targets into the pGAZ-leu1 integration vector and transformation into *S. pombe*

The chosen targets were amplified form the *F. fomentarius* cDNA library and the PCR fragments are shown in Figure 25. As seen in Table 21, amplified genes were correlating to the expected sizes. After restriction digestion the fragments were ligated with the pGAZ2_*leu1* integration vector where the adh1 promoter drives transcription of the target gene. For the transformation into *S. pombe* the vectors were linearized and the adapted lithium acetate method of Bähler *et al*. (105) was used to transform them into *S. pombe* wild type strain.

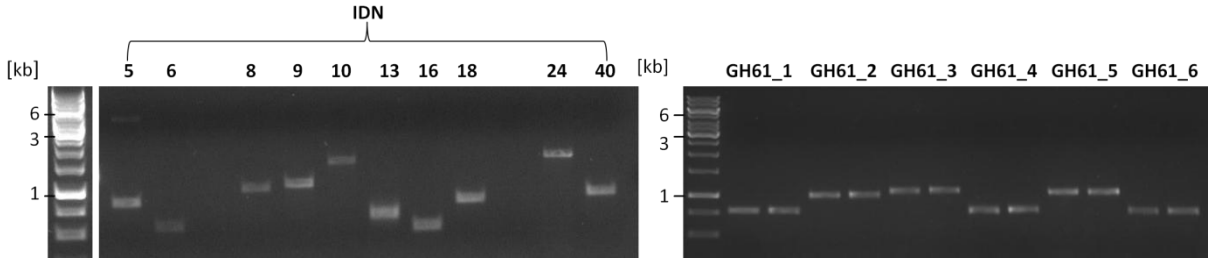


Figure 25: PCR products of cDNA targets from *F. fomentarius* amplified from the cDNA library.

Table 21: Sizes of cDNA targets in bp.

cDNA targets	Size [bp]
IDN5	732
IDN6	465
IDN8	858
IDN9	915
IDN10	981
IDN13	555
IDN16	429
IDN18	711
IDN24	1458
IDN40 (GH61_1)	752
GH61_1	732
GH61_2	948
GH61_3	1014
GH61_4	717
GH61_5	996
GH61_6	697

4.3. Dot blot screening

The screening for enzymes with unknown function is difficult, due to the fact that there is no activity or other property known of these proteins to screen for. Nevertheless, his-tags that had been attached to all proteins to the 3'end, made it possible to screen for proteins, which had been secreted with their native leader peptides.

For each target around 88 transformants were picked and screened for secreted his-tagged proteins in the supernatant. This was done by cultivation in a 96 deep well plate for 48 hours. As positive control an *E. coli* cell lysate from a strain that expresses a his-tagged protein was used. To clarify if a background signal is caused by *S. pombe*, without expressing a heterologous protein or by the medium, two negative controls had been applied: First a strain having the empty integrated vector (vec ctrl) and secondly the medium alone.

Table 22 gives an overview of all dot blot analyses. As shown the best targets in *S. pombe* were IDN9, IDN10, IDN13, IDN24 and GH61_2. In these cases, the controls all worked well and the signals of the proteins were relatively intense. As example the dot blot analyses of IDN10 and IDN13 are depicted in Figure 26.

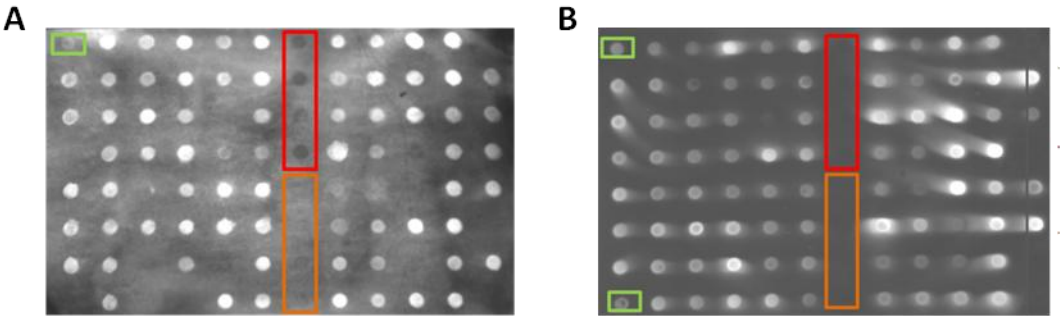


Figure 26: Dot blot analyses of IDN10 (**A**) and IDN13 (**B**) based on his tag detection. *S. pombe* were cultivated for 48 hours and subsequently 15 µl supernatant from each well was transferred on the membrane. The green box shows the positive control – *E. coli* cell lysate – the red box the *S. pombe* empty vector control and the orange box only the medium.

Table 22: Dot Blot results for all tested targets. The minus (-) stands for no signal with the anti His antibody or not distinct enough, whereas one or more plus (+) are indicating a positive signal and the intensity of this signal.

Name	**Contig**	***S. pombe***	***P. pastoris***
IDN5	Isotig04428	not tested	+
IDN6	Contig01698	+	+++
IDN7	Isotig04926	not tested	++
IDN8	Isotig05085	-	not tested
IDN9	Isotig05070	+++	+
IDN10	Contig02057	+++	not tested
IDN13	Isotig02488	+++	+++
IDN16	Isotig03896	-	-
IDN24	Isotig03302	+++	-
IDN40	Contig01925	-	-
GH61_1	Isotig04623	+	-
GH61_2	Isotig05034	+++	++
GH61_3	Isotig03735	-	not tested
GH61_4	Isotig06156	++	-
GH61_5	Contig01420	-	not tested
GH61_6	Contig01118	-	not tested

4.4. Cultivation, purification and characterization ofIDN6, IDN10 and IDN13

The work load to process all positive targets would go beyond the scope of this thesis. Therefore, the shake flask cultivations were exemplarily done for IDN6, IDN10 and IDN13.

For the purification of the proteins two different nickel affinity columns were used: First the 5 ml HisTrap FF column and second a 5 ml HisTrap™ excel. The latter column is packed with a special resin, which allows loading the filtered media directly onto the column, without buffer exchange in advance. Normally, media would cause a stripping of the Ni^{2+} and thus lead to a decreased binding of the protein or no binding at all. However, the HisTrap™ excel did not prove to work as promised in its description. The obtained protein amount was always very low and the peak heights of all tested proteins had the same height as the elution peak of the empty vector control (~20 mAU).

Therefore, the purification was done with the HisTrap FF column. As for the HisTrap™ excel, the filtered culture medium was directly loaded onto the column. In contrast to the purification with the HisTrap™ excel, peaks were 20 times higher than the negative control (shown in Figure 27). However, compared to other his-trap purifications, the obtained 400 mAU are still not very high. Therefore the approximately 200 ml supernatant were concentrated with Viva spin columns (PES, 5 kDa cut off) beforehand and the buffer was exchanged with desalting buffer in order to avoid Ni^{2+} stripping. The volume was reduced to ~2 ml and was then loaded onto the column. However, this could not improve the yield. The results obtained were similar to those shown in Figure 27 and considering the time loss due to the additional concentration step at the beginning and the fact that actually more protein should have been loaded on the column did not prove this method to be viable.

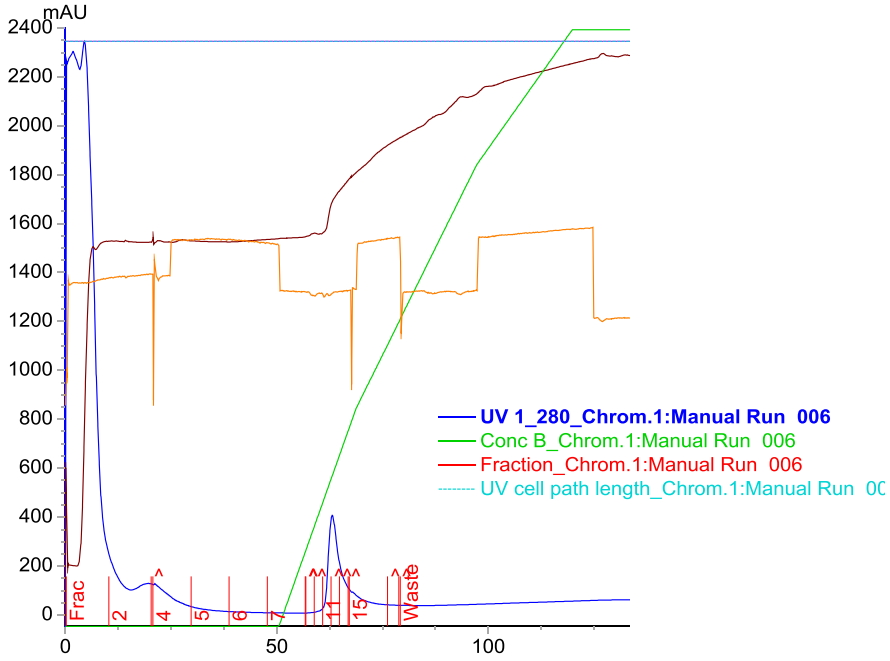


Figure 27: Äkta purification of IDN10 expressed in *S. pombe* by a 5 ml HisTrap column. A peak height of around 400 mAU can be observed at an imidazole concentration of around 25 % and is supposed to be the heterologously expressed protein. (A sample of an empty vector control which was similarly treated got a peak height of about 20 mAU).

The eluted peak fractions were pooled and concentrated with Viva spin columns and also the buffer was exchanged in order to get rid of the imidazole in the fractions. The concentrations of the proteins were determined with Bradford assay and BCA assay in microplates. The determined protein concentrations of the samples were very low and in most cases even too low to be in the linear range of the calibration (0.1 to 1.5 mg/ml). The highest amount was obtained for IDN13 D11 with a concentration of 0.71 mg/ml (~1.5 ml)

determined with Bradford assay. Besides the microtiter procedure there are other protocols to determine the protein concentration for BCA as well as for Bradford Assay. Therewith, it is possible to measure lower concentrations (down to 5 µg/ml), but more sample is required. Considering that these samples were further used for SDS-PAGE, Western Blot and activity assays, this option was not feasible.

4.4.1. SDS-PAGE and Western Blot

The purified samples were further used for SDS-PAGE and Western blot analyses. In Figure 28 the SDS-PAGE of IDN10 E6 and IDN13 D11 is shown. Because *S. pombe* tends to hyperglycosylate its secreted proteins, the bands in a SDS-PAGE are often appearing rather blurry than as sharp bands. Therefore, the samples were additionally deglycoslated with EndoH to get rid of N-glycosylation. This enzyme cleaves high mannose structures and hence reduces the volume of the protein. The EndoH treatment thus results in sharper bands on the gel. This is also true for IDN13 D11 and IDN 10 E6. In Figure 28 lane 1 shows the deglycosylated and concentrated supernatant of IDN13 D11. There are several sharp bands, whereas the same sample without the EndoH treatment in lane 2 appears smeary. The size of the EndoH itself is about 30 kDa and visible in lane 6. A similar pattern is shown in lane 4 obtained by IDN13 and it is very likely that this is EndoH alone. There are two possibilities: First that the protein concentration is too low to be seen on the gel or second that the EndoH and IDN13 per chance have the same size. The calculated size of IDN13 is 19 kDa and is too low on its own. However, this size does not consider the glycosylation of the protein. Therefore, a higher size can be expected and 30 kDa for IDN13 on the SDS-PAGE could be possible.

The situation for the second protein, IDN10, is clearer. In lane 11 the deglycosylated, purified and concentrated sample is shown. There is a faint band around 70 kDa, whereas the native sample shows a smear on the top of lane 12. The calculate size for IDN10 is 49 kDa. Also in this case – not unexpected – the calculated and real sizes are different.

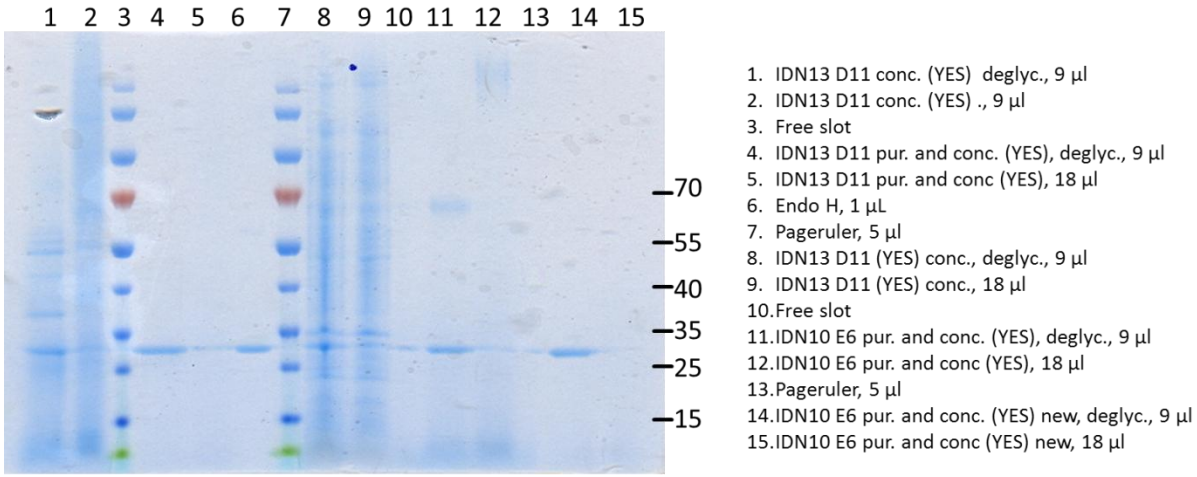
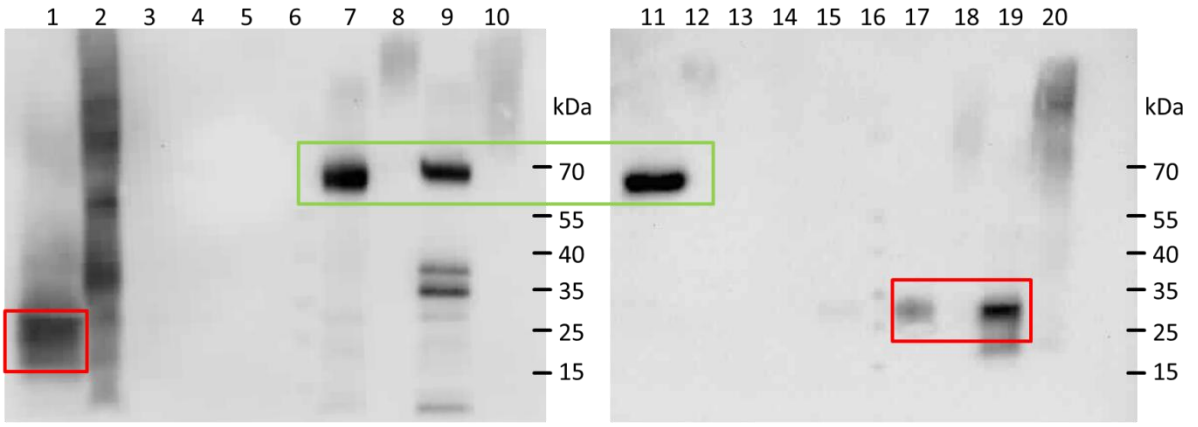


Figure 28: SDS-PAGE analysis of IDN10 and IDN13. Abbreviations: conc…supernatant was concentrated with a Viva pin column (5 kDa cut off); deglyc…sample was treated with EndoH; cpur…purified with 5 ml his trap column; YE…grown in YES medium; EMM…grown in EMM2

As next step Western blot analyses of the protein samples were performed, resulting in the two blots shown in Figure 29. The red boxes are indicating the bands for IDN13 at a size of 30 kDa. In lane 3 and 13 only EndoH was loaded but no signal was detected. Therefore, it is likely that EndoH and deglycosylated IDN13 have the same protein size of 30 kDa. The green boxed bands show the deglycosylated IDN10 at a height of around 70 kDa. Thereby, the band in lane 9 seems to be slightly higher. This is due to the fact that the samples shown in lane 7 and lane 11 are deriving from the same cultivation, whereas the sample shown in lane 9 is from a different one, where EMM2 instead of YES medium was used. It seems that the minimal medium had an effect on the glycosylation pattern and is the reason for the higher protein size. Moreover, the EMM2 seems to lead to more degradation of the protein, as the smaller bands in lane 9 indicate.



1. IDN13 D11 conc. (YES), deglyc., 9 µl
2. IDN13 D11 conc. (YES., 18 µl
3. Endo H, 1 µL
4. Purification fraction next to protein peak IDN10 E6 (YE), deglyc., 9 µl
5. Purification fraction next to protein peak IDN10 E6 (YE), 18 µl
6. Pageruler, 5 µl
7. IDN10 E6 pur. and conc. (YES), deglyc., 9 µl
8. IDN10 E6 pur. and conc (YES), 18 µl
9. IDN10 E6 conc., deglyc. (EMM2), 9 µl
10. IDN10 E6 (EMM), 18 µl
11. IDN10 E6 conc. (YES) , deglyc., 9 µl
12. IDN10 E6 conc. (YES) , 18 µl
13. Endo H, 1 µL
14. Purification fraction next to protein peak IDN13 D11(YE), 18 µl
15. Purification fraction next to protein peak IDN13 D11 (YE), deglyc., 9 µl
16. Pageruler, 5 µl
17. IDN13 D11 pur. and conc. (YES), deglyc., 9 µl
18. IDN13 D11 pur. and conc (YES), 18 µl
19. IDN13 D11 conc., deglyc. (EMM2), 9 µl
20. IDN13 D11 (EMM2), 18 µl

Figure 29: Western Blot analysis of IDN13 D11 and IDN10 E6. The colours of the picture were inverted. The green highlighted bands are deglycoslylated IDN10 and the red highlighted bands are deglycosylated IDN13. Abbreviations: conc…supernatant was concentrated with a Viva Spin column (5 kDa cut off); deglyc…sample was deglycosylated with EndoH; pur…purified with 5 ml HisTrap column; YE…grown in YES medium; EMM…grown in EMM2.

4.4.2. Enzymatic assays

The proteins obtained by the shake flask cultivations were also tested for different enzymatic activities. This was done with different glycosidic substrates, which were all linked with 4-Methylumbelliferyl (4-MU). When the bond between the 4-MU and the glycosidic substrate is cleaved, the 4-MU becomes a fluorescent molecule and can be measured. First the proteins were tested in a substrate mix to clarify if any activity can be detected. Because the protein concentration was not exactly known, different dilutions of the purified and concentrated protein solutions were tested. The fluorescence was measured at several time points and the experiment stopped after 16 hours.

In Figure 30 the results of the enzymatic assay for IDN6 D11, IDN10 E6 and IDN13 D11 – all diluted 1:5 – is shown. The negative control (neg ctrl, red bars) is desalting buffer with a pH of 7.6. This buffer had been also used for the buffer exchange of the protein solutions after their purification. The maximal value derives from the plate reader which reaches its detection limit when the fluorescence exceeds 99999 units. This overflow was reached after

16 hours for IDN10 as well as for IDN13. All three enzymes showed at least a 2 fold higher fluorescence than the negative control. However, IDN13 is the only protein, which shows constantly increasing values. IDN6 and IDN10 on the contrary are – except for the last measurements – always in the same range as the negative control. To clarify if the proteins indeed have an enzymatic activity and especially which one, they were further tested with single substrates.

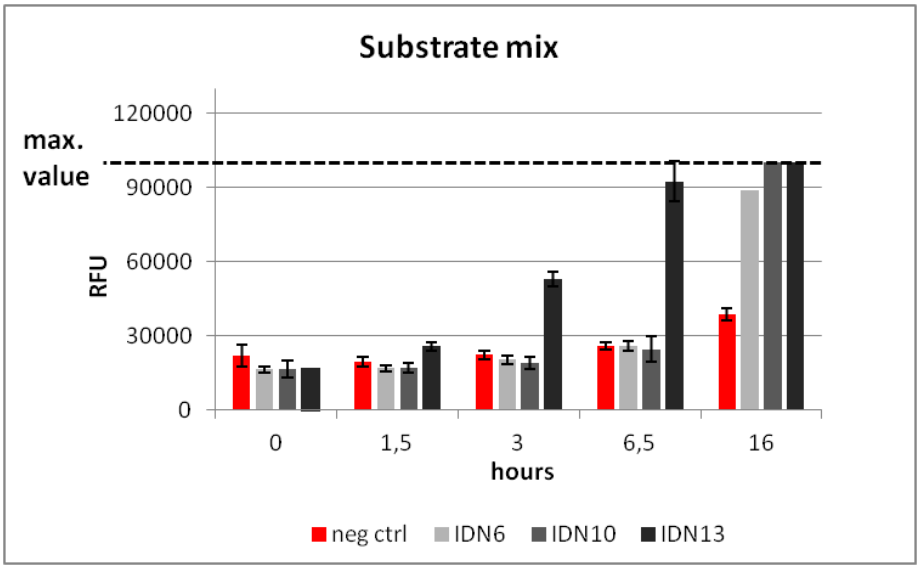


Figure 30: Enzymatic activity DN6 D11, IDN10 E6 and IDN13 D11 (1:5).

IDN6 D11 and IDN13 D11 tested on single substrates showed no enzymatic activity at all. However, the tested protein samples were already a few weeks old, because at that time no fresh samples were available. To examine the aging effect on the enzymatic activity, fresh IDN10 E6 was tested in a single substrate solution. In Figure 31, the fresh IDN10 was active on the 4-MU-β-D-glucopyranoside (Glc), whereas the two week old sample showed no comparable activity.

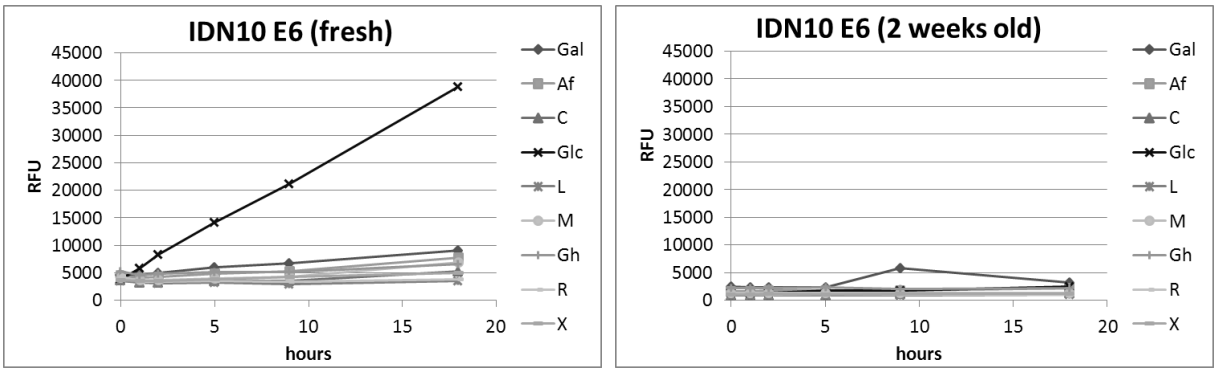


Figure 31: Emzymatic asseay with protein samples IDN10 E6, one fresh and one two weeks old. Different substrate were used: Gal…4-MU-α-D-galactopyranoside, Af…4-MU-α-L-arabinofuranoside, C…4-MU-β-D-cellobioside, Glc…4-MU-β-D-glucopyranoside, L….4-MU-β-D-lactopyranoside, M….4-MU-β-D-mannopyranoside, Gh….4-MU-α-L-glucuronide hydrate, R….4-MU-α-rhamnopyranoisde, X….4-MU-β-D-xylobioside.

Although the preliminary results with IDN13 looked promising, the enzymatic activity for IDN13 as well as for IDN6 remains unknown. IDN10 on the other hand seems to have a β-glucosidase activity. Besides the clear motif of a cellulose binding domain at the N-terminus of the protein no clear conserved domains have been recognized. Most homologous proteins in databases were annotated as putative proteins with potential 4 O-methyl glucuronyl esterase activity or as members of the carbohydrate esterase family 15 but not as glycosyl hydrolases. One related protein is even involved in the hemicellulose disconnection of lignin (120). However, for further characterization the protein yield has to be increased. Codon optimized genes and the removal of the native leader peptide of *F. fomentarius* would probably lead to an improvement of the secretion.

4.4.3. Complementation assay

There is the possibility that the proteins themselves do not have an enzymatic activity, but an enhancing effect on the degradation of lignocellulose – e.g. due to protein-protein interactions (scaffold proteins), helper proteins or surface active proteins. Therefore the proteins IDN9, IDN10 and IDN13 were further tested with a complementary assay. The assay does not aim to prove a specific enzymatic activity but the complementing effect on the lignocellulose degradation process. By testing a lignocellulolytic enzyme mix together with the produced proteins, their effect on the degradation of straw can be examined. The enzyme mix was kindly provided by Clariant (the exact composition of the protein mix is due to company concealment reason unknown). The assay was performed at 50°C for at least 24

hours and several samples were drawn during the incubation time. Afterwards the released reducing sugars were measured with a DNS assay. By comparing the results with the negative control, effects due to the added DUF proteins could be evaluated.

However, no significant effect of the lignocellulose degradation could be detected for any of the tested proteins IDN9, IDN10 and IDN13. The development of reducing sugars was in all samples – with and without added proteins – at the same rate.

4.5. Small scale bioreactor cultivations (DasGip) of a *S. pombe* strain

To compare the two expression hosts *P. pastoris* and *S. pombe* small scale bioreactor cultivations were performed. In the case of *P. pastoris* there is already an established protocol for bioreactor cultivations in-house. The cultivation conditions of *S. pombe* on the other hand still needed improvement. Therefore, a pretrial for the *S. pombe* bioreactor cultivation was necessary and conducted as described below.

Two different media pairs – batch media and fed batch media – were used for this fermentation. The first pair is based on YES medium, which is the most common rich medium used for *S. pombe*. Additionally minerals and vitamins were added. The glucose concentration is ten times higher in the fed batch medium than in batch medium. The second media pair, SMF23 and FSS05, are more complex synthetic media (SM) (121). This medium pair was chosen, because a similar expression system was used: A constitutive promoter (hCMV promoter) drove the expression of a heterologous protein, which was secreted into the supernatant. Moreover, high cell densities (OD_{600} >250) were reported.

Another possibility is to use EMM2. Minimal medium may reduce the amount of other proteins in the supernatant. However, shake flask cultivations indicate that the protein yield is significantly decreased in contrast to the yield reached with full medium. Therefore, YES full media and the media described by Ikeda et al. (121) were chosen for the bioreactor pretrial.

The fermentation was done with the same strain – *S. pombe* expressing native IDN10 – in four vessels. The first two vessels, YES1 and YES2, were treated equally until the fed batch phase started. The same was true for the other two vessels, SM1 and SM2. In fed batch phase the feeding was always three times higher in the second vessel. The reason for applying different feeding rates is the crabtree effect. This phenomenon describes the behaviour of many yeasts under high glucose concentrations, where energy is produced rather by ethanol production than by the citric acid cycle even under aerobic condition. This leads to faster energy consumption but also decreases biomass growth. In order to reach the optimal glucose concentration during fed batch growth of crabtree positive yeasts, the optimal feed rate where the respiratory quotient ($RQ = CO_{2\ eliminated} / O_{2\ consumed}$) is equal to 1 has to be found. However, the used bioreactors did not have suitable instruments to measure the exhaust fumes of the cell culture during the fermentation. Therefore, the

feeding was fixed to a pulsed mode supply, where the fed batch medium was brought in with a certain feed rate for half an hour. After a pause of one and a half hours the cycle started again. In Table 23 the four fermentation vessels and their key data are shown.

Table 23: Characteristics of the regarding bioreactor vessel. Abbreviations: DO…dissolved oxygen

Vessel 1 – YES1	**Vessel 2 – YES2**
pH 6, 30 % DO, 30°C **Batch medium:** YES medium + minerals and vitamins **Fed batch medium:** YES medium + minerals and vitamins + 30 % glucose (10 times higher than in the batch medium) **Feed start:** 36 h **Feed type:** pulsed mode supply: 4 ml/h for half an hour, one and a half hour pause **Fermentation end:** 60 h	pH 6, 30 % DO, 30°C **Batch medium:** YES medium + minerals and vitamins **Fed batch medium:** YES medium + minerals and vitamins + 30 % glucose (10 times higher than in the batch medium) **Feed start:** 36 h **Feed type:** pulsed mode supply: 12 ml/h for half an hour, one and a half hour pause **Fermentation end:** 60 h
Vessel 3 – SM1	**Vessel 3 – SM2**
pH 6, 30 % DO, 30°C **Batch medium:** SMF23 **Feed medium:** FSS05 (30 % glucose) **Feed start:** 44 h **Feed type:** pulsed mode supply: 4 ml/h for half an hour, one and a half hour pause **Fermentation end:** 60 h	pH 6, 30 % DO, 30°C **Batch medium:** SMF23 **Feed medium:** FSS05 (30 % glucose) **Feed start:** 44 h **Feed type:** pulsed mode supply: 12 ml/h for half an hour, one and a half hour pause **Fermentation end:** 60 h

During the cultivation process samples were taken in regular intervals and analysed. As seen in Figure 32 it was not possible to achieve as high cell densities (OD_{600} >250) as stated in literature(121). Moreover, the biomass of the *S. pombe* strain cultivated in the synthetic minimal medium was far lower than the strain cultivated in the rich medium.

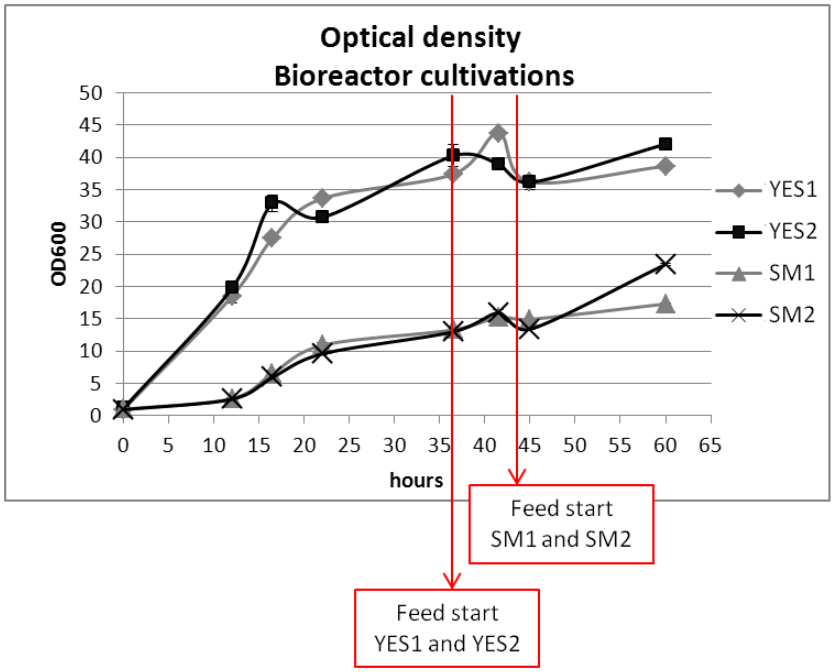


Figure 32: Measured optical density during the bioreactor cultivation procedure of IDN10. Rich medium (YES), Synthetic Medium (SM). Higher feeding rates for YES2 and SM2.

The taken end point samples of the supernatants were either directly or after TCA precipitation loaded on a SDS-PAGE as shown in Figure 33A. For the Western blot analysis as shown in Figure 33B instead of TCA precipitation samples were concentrated by Viva Spin columns (5 kDa cut off). Although a lot of different bands were detected on the SDS-PAGE for SM2 no target protein was detected on the Western blot. Hence, it is likely that the cell culture in SM2 was contaminated. As Figure 33B indicates vessel YES1 contains more degraded product and less protein compared to YES2. In general, the expression worked better in rich media than in synthetic media.

For expression experiments it is therefore recommended to use rich medium (YES) combined with the higher feeding rate. The influence of a further increased feeding rate on the expression remains unknown, but it probably improves the protein yield. Furthermore, compared to shaking flask cultivations the found data show that the protein expression performance in *S. pombe* is better in conditions with pH 5.5 (or pH5) than pH 6.

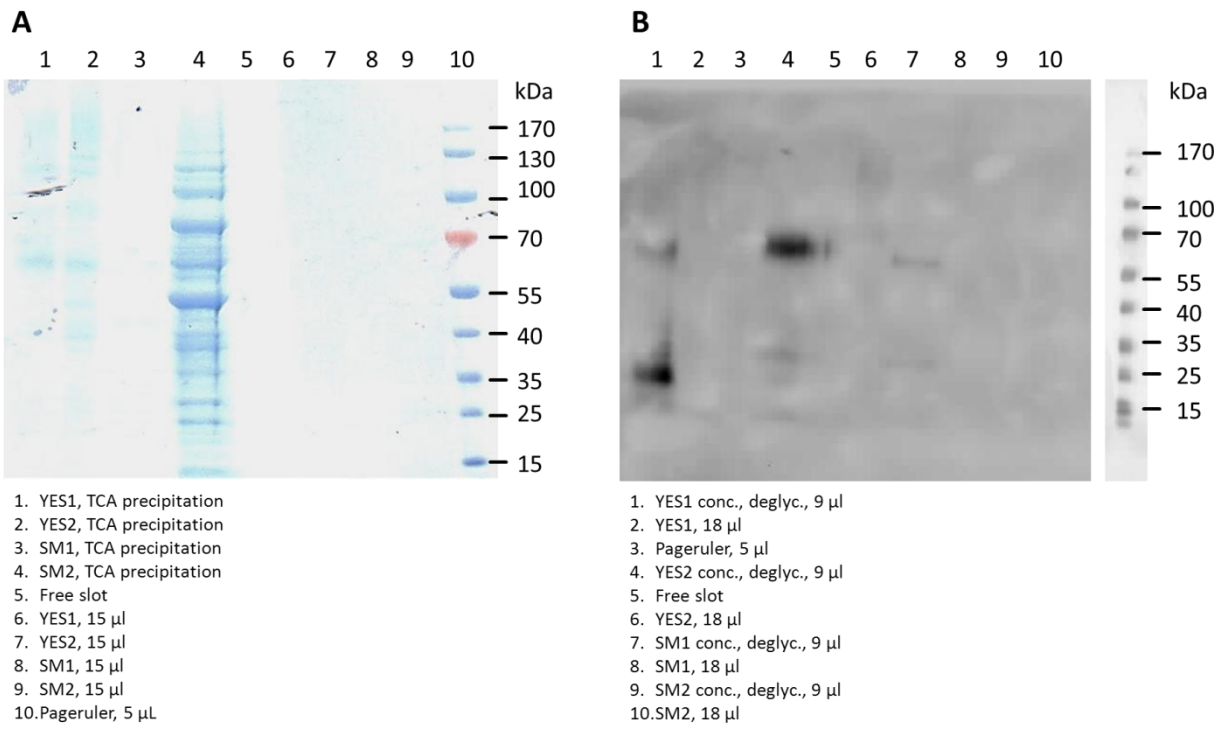


Figure 33: **A** shows SDS-PAGE of samples taken at the end of the bioreactor cultivation (after 60 hours). For the TCA precipitation 500 µl supernatant was used. **B** shows the Western Blot analyses of the end point samples. Abbreviations: conc…supernatant was concentrated with a Viva Spin column (5 kDa cut off); deglyc…sample was treated with EndoH; pur…purified with 5 ml HisTrap column; YES1 and YES2…refers to vessels with rich media; SM1 and SM2…vessels with synthetic media pair (SMF23 and FSS05).

4.6. Codon optimization of various cDNA targets

To compare the two expression hosts *P. pastoris* and *S. pombe*, interesting targets from the described cDNA target screening were used. Additionally, the laminarinase gene (clone 30H12) from Mitrovic´s work (118) was included to the expression host comparison. This gene was found during the screening of the *F. fomentarius* transcript library in *S. pombe* and revealed to be active on the MUGal (4-MU-α-D-galactopyranoside) substrate. All used genes are listed in Table 24 and the related sequences can be found in the appendix.

To guarantee similar conditions in both organisms the genes were codon optimized for *S. pombe* and *P. pastoris*. The codon usage tables used for the redesign of the gene sequences are based on highly transcribed genes in both yeasts and are shown in Abad *et al.* (106) for *P. pastoris* and in Forsburg *et al*. (77) for *S. pombe*. The native leader sequences were predicted by using the webtool SignalP. They were replaced with approved ones: For *P. pastoris* the complete α-factor signal sequence and for *S. pombe* the Carboxypeptidase Y leader peptide sequence (18; 48). For both hosts the same Kozak sequence was used.

Table 24: The cDNA targets chosen for codon optimization. The cDNA targets were either optimized for *S. pombe* and *P. pastoris*. The contig and the isotig numbers are referring to the bioinformatical analysis of the transcriptome of *F. fomentarius.*

cDNA targets	**optimized for**	**new designation**	**contig**	**conserved domain**
IDN6	*S. pombe,* *P. pastoris*	IDN6_Sp, IDN6_Pp	contig01698	Carbohydrate-binding domain
IDN9	*S. pombe,* *P. pastoris*	IDN9_Sp, IDN9_Pp	isotig05070	Fungal-type cellulose-binding domain
IDN10	*P. pastoris*	IDN10_Pp	contig02057	Fungal-type cellulose-binding domain
Laminarinase (clone 30H12)	*S. pombe,* *P. pastoris*	La_Sp, La_Pp	contig01418	GH16

4.7. Cloning of the synthetic genes into expression vectors and transformation into the yeasts

The synthetic genes were cloned into the respective *E. coli* shuttle vector systems, either suitable for expression in *S. pombe* or *P. pastoris*. In Figure 35 the three synthetic genes optimized for *S. pombe* cloned into the pGAZ2-*leu1* vector can be seen. The vector was cut with *Hind*III and *Spe*I in order to check the right size of the insert. Afterwards the correct sequences were verified by sequencing and transformed as linear expression cassettes into *S. pombe*.

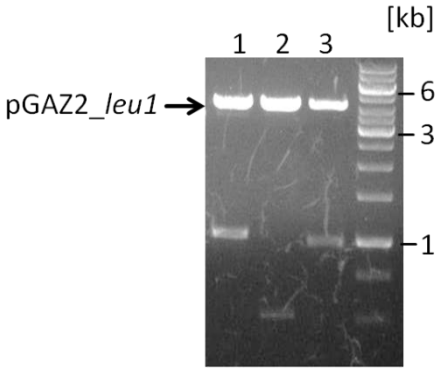


Figure 34: Restriction digestion (*Hind*III and *Spe*I) of pGAZ_*leu1* containing the synthetic genes optimized for *S. pombe*. Lane 1 shows IDN9_Sp (1045 bp), lane 2 IDN6_Sp (517 bp) and lane 3 La_Sp (1006 bp)

For *P. pastoris* the synthetic genes were cloned at first into the pPpB1_GAP-α_S vector. However, the transformation in *P. pastoris* did not work for any of the synthetic genes. The reason for the failure of the transformation might have been the metabolic burden. This phenomenon describes an overload of the metabolic machinery due to the expression of a heterologous protein (122). The expression system used in the first place, has the strong constitutive *GAP* promoter and the pPpB1_S vector, leads to transformants with varying gene dosage – i.e. to a range of single to multi copy transformants. Moreover, the genes are optimized for high-level expression in *P. pastoris*, which plays probably the most contributing factor in the metabolic burden.

For a host comparison between *S. pombe* and *P. pastoris*, the expression systems should be as similar as possible. Therefore, the genes were recloned into the pPpT4-GAP_S vector, which leads in contrast to the pPpB1_GAP-α_S to single copy integrants. And indeed, the vector change seemed to ease the metabolic burden, because transformants for IDN6 could

be obtained. However, the transformation of the other three synthetic genes remained unsuccessful. Hence, the genes were recloned a second time into the pPpB1_AOX1_S vector – shown in Figure 35 – which contains the strong methanol inducible *AOX*1 promoter. This should also release the metabolic burden, because the expression of the genes does not start until the cells get induced. The transformations of these constructs into *P. pastoris* on the other hand had been successful and the obtained transformants which were handed over to colleagues for further analysis.

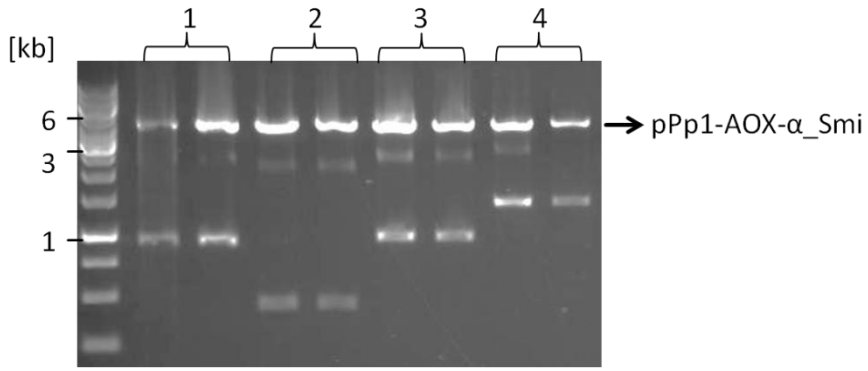


Figure 35: Restriction digestion (*EcoR*I and *Not*I) of pPp1-AOX-α_S containing the synthetic genes optimized for *P. pastoris*. Lane 1 shows La_Pp (938 bp), lane 2 IDN6_Pp (449 bp), lane 3 IDN9_Pp (977 bp) and lane 4 IDN10_Pp (1403 bp).

4.8. Dot blot analysis of synthetic gene transformants

The screening procedure was the same as described above. In Figure 36 the Dot blots are shown. In A to C the screening of the *S. pombe* strains is depicted, whereas D shows the only *P. pastoris* screening, where the *GAP* promoter worked. The strong intensity of this blot indicates that the expression level of IDN6-Pp is very high.

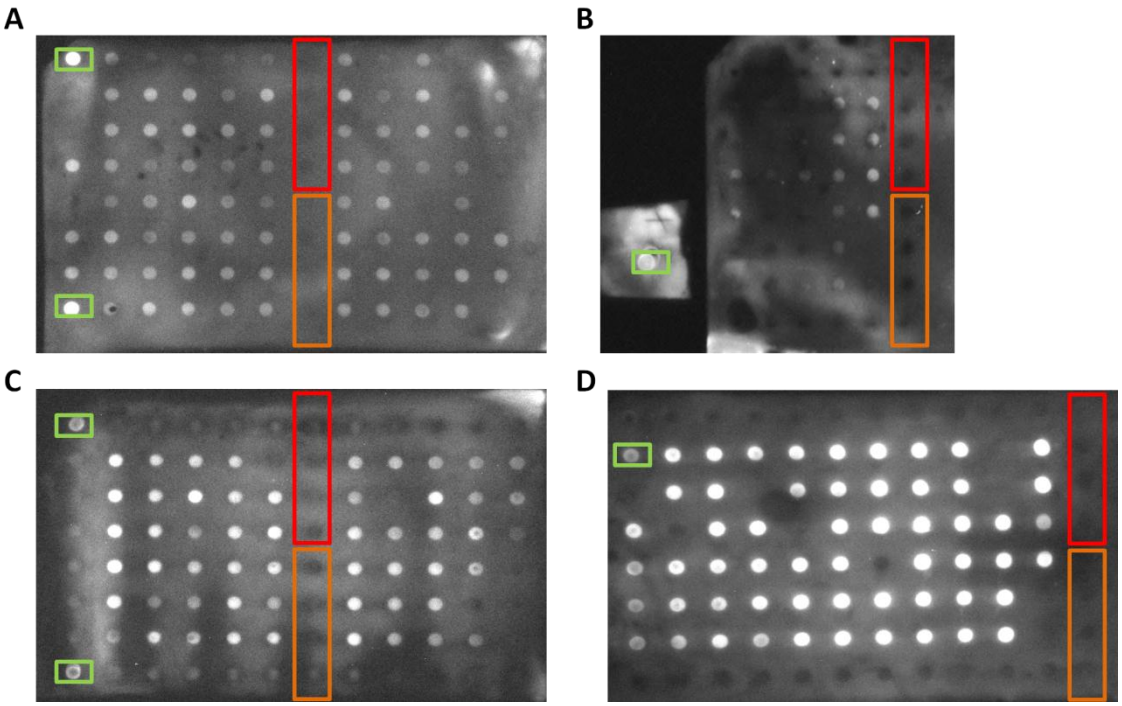


Figure 36: Dot blot analyses of *S. pombe* and *P. pastoris* strains with codon optimized genes. A shows the pGAZ2_*leu1*-IDN9_Sp, B the pGAZ2_*leu1*-IDN6_Sp, C the pGAZ2_*leu1*-30H12_Sp and D thepPpT4_GAP_S-IDN6_Pp Dot blot. . Green box shows the positive control – *E. Coli* cell lysate – the red box the *S. pombe* empty vector control and the orange box only the medium.

The screening of the synthetic genes optimized for *P. pastoris* in the pPp_AOX1_S vector and also following steps go beyond the scope of this thesis.

5. Conclusion and Outlook

The construction of PCR cassettes is a fast and straight forward approach: It is easily possible to tag, overexpress or delete any desired gene in the *S. pombe* genome. The fluorescence measurement – with eGFP as reporter gene – indicated that a single copy integration event occurs. Thereby, the targeted region in the *S. pombe* genome should be replaced by the expression construct. A colony or a quantitative PCR can verify this assumption. The integration efficiency obtained for the tested construct was 100 %. However, the reproducibility of this result has not been checked. A good application for these PCR cassettes can be the enhancement of heterologous protein expressing strains by targeted overexpression of helper proteins in their own loci. Another possibility to enhance heterologous protein expression is by targeted integration of foreign helper proteins. This integration deletes any desired gene, for example, a gene involved in amino acid synthesis. In this case a negative selection becomes possible, without using an antibiotic selection marker. Furthermore, protein producing strains are also often enhanced by gene knockouts, e.g. to influence the glycosylation. Overall, this method might be suitable for many applications.

The designed integration vectors can not only be used for episomal expression of proteins but also for integration into the *S. pombe* genome. This is possible because a*rs1* – the eukaryotic origin of replication –together with the pUC – the prokaryotic origin of replication – can be easily cut out by restriction digestion, leaving a linear expression construct that is suitable for integration. In contrast to the PCR cassettes, the integration vectors led to a spectrum of different transformants which seem to exhibit varying gene dosage. This is reflected by the fluorescence landscapes: The integration vectors are leading to steeply rising landscapes with high RFU/OD values. Interestingly the same trend can be observed for the integrated construct as well as for the episomal (Figure 18). This result was expected for the episomal expressed eGFP because up to 80 copies are possible with plasmids having the *ars1* sequence (31). The integration of the construct with a similar recombination system on the other hand has been reported to result in single copy transformants (107). However, if the integration event is occurring as in Figure 15 proposed, every insertion of the linearized expression construct may result in a new insertion site. This fact is shown in Figure 37: Thereby, every successful insertion increases the possibility for further integrations. Colony PCR of possible multi copy integrants could not clearly confirm this possible integration

mechanism. In order to determine the copy number, it is necessary to repeat the colony PCR with a longer extension time. Moreover, a quantitative PCR could clarify the findings. However, there is no protocol for *S. pombe* available at the moment and hence has to be established first.

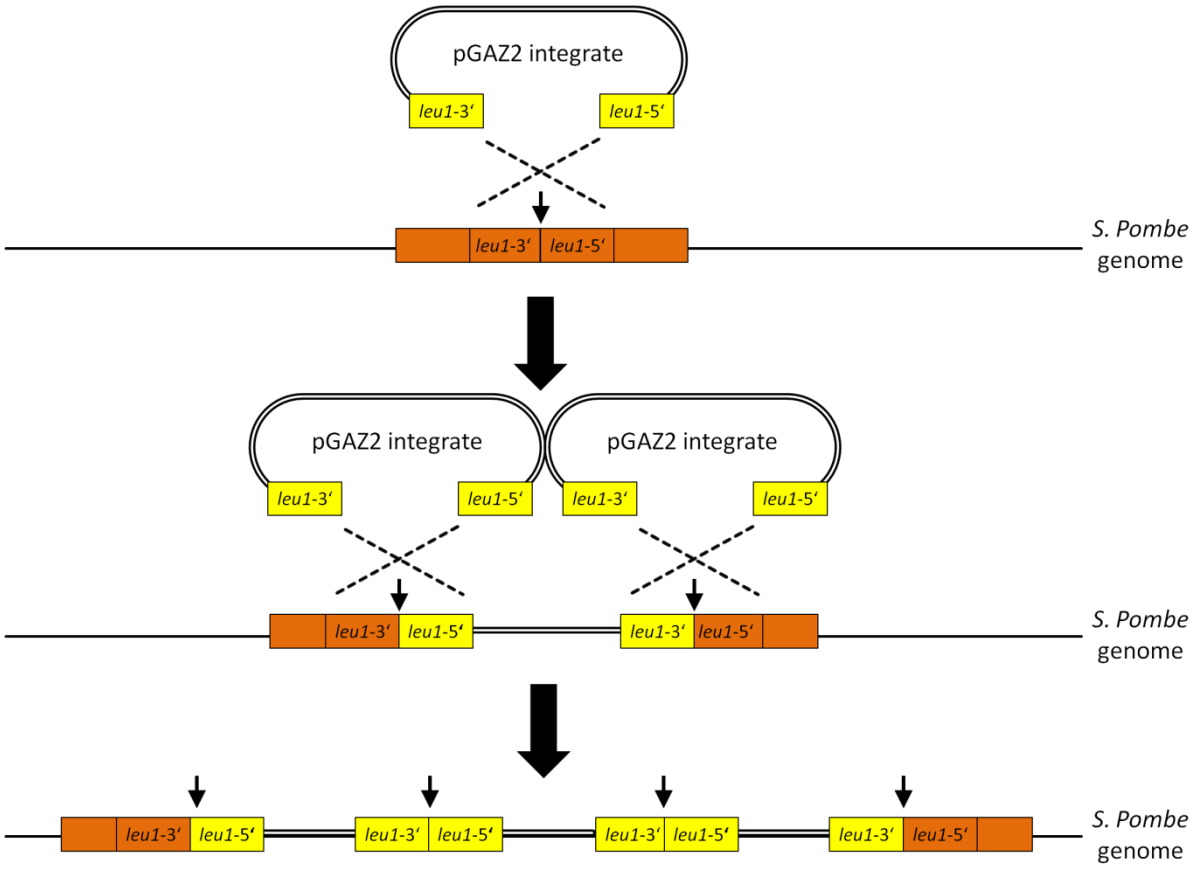


Figure 37: Proposed mechanism for multi copy integration of the pGAZ2 integrate. Every successful insertion establishes another insertion site. The integrated pGAZ2 vector is shown as double line in the *S. pombe* genome and the insertion sites are marked by the small arrows.

The *adh1*$^{+}$ promoter of *S. pombe* is said to be a constitutive promoter (112), but the experiment in 3.2.5 proposes that there is a possible regulation mechanism. Similar to the *ADH1* promoter in *S. cerevisiae* the *adh1*$^{+}$ promoter is probably repressed by ethanol in the medium. To verify this assumption, control experiments with different C-sources are necessary: Glucose, glucose-ethanol mixtures and glycerol-acetate mixtures are good choices for the growth comparison. The cultivation with a glucose medium leads to decreasing values over time as the experiment in 3.2.5 shows. Although fresh medium was added and the growth of the cells again triggered, the produced ethanol seems to repress the *adh1*$^{+}$ promoter even after the addition of new glucose. If the assumption of the ethanol

repression is true, general lower RFU/OD values are expected for the growth of these strains on a glucose-ethanol mixture. Thereby, an adequate glucose-ethanol ratio has to be found. The glycerol-acetate mixture is a non fermentable C-source and therefore allows observing the expression level in the respiratory growth mode. The reason for choosing a mixture of glycerol and acetate is that *S. pombe* hardly grows on glycerol as its only carbon source. The two key enzymes involved in glycerol metabolization are expressed at very low levels and cannot efficiently convert glycerol to ensure the cell's survival. These two enzymes are the glycerol dehydrogenase (*gld1*$^{+}$) and the fructose-1,6-bisphosphatase (*fbp1*$^{+}$). (123). Moreover, *S. pombe* cannot grow on ethanol as its only carbon source, because it has no functional glyoxylate cycle. Therefore, ethanol C-source should only be tested with others like glucose in low concentrations (124).

The stability of the episomal and integrated plasmids in *S. pombe* has to be still tested. Therefore, shaking flask cultivation with strains containing the circular plasmids without Zeocin™ in medium can be done. To check the stability, samples from these cultivations are drawn at several time points and a fixed amount of cells are plated on agar plates with and without selection pressure. The ratio of count colonies that either kept or lost the plasmid then will determine the plasmid loss. As additional control, the same cultivations can be done with Zeocin™ to see if the plasmid is maintained. Moreover, by testing strains with the integrated expression construct it is possible to compare their performance to strains habouring the episomes.

Moreover, a comparison of an electroporation and a lithium acetate transformation protocol was done. Interestingly, by using lithium acetate method all transformants show an increased fluorescence of the eGFP reporter. The integration vector combined with the lithium acetate transformation method leads to an 8.4 times increased fluorescence values compared to the pGAZ2-eGFP vector used with electroporation protocol. However, the finding that the lithium acetate method seems to be superior compared to the two tested electroporation methods should be verified by repeating the transformations and the cultivation of the transformants in 96 deep well plates.

In the last part of this thesis the above described tools and methods for the expression of *S. pombe* were applied in order to find and characterize new proteins and enzymes for lignocellulose degradation. Therefore, the transcriptome of the brown-rot fungus *F.*

fomentarius – grown under inducing conditions – was isolated, sequenced and converted into a cDNA library before. Afterwards a bioinformtic analysis was done. Interesting targets were identified by the project team and had to have parameters like an existing secretion signal, a complete ORF as well as DUFs. These targets were successfully amplified and cloned into a shuttle vector, suitable for integration into the *S. pombe* genome. In my thesis 14 targets were screened for secretion in the fission yeast and 8 could be detected in the supernatant of the cell culture by a dot blot screening.

The targets IDN6, IDN10 and IDN13 were cultivated in shake flasks, purified with nickel affinity chromatography and further investigated via SDS-PAGE and Western Blot. The expression level of the proteins seemed to be very low, because the protein bands on the SDS-PAGE were in all cases very faint or not detectable with Coomassie® Blue staining. The proteins were also tested for different glycolytic activities and for a complementing effect in the degradation of lignocellulose. However, due to the low protein yield, the characterization of the proteins was without any significant result, except for IDN10, which showed a β-glucosidase activity.

In order to reach higher expression levels and to compare the two expression hosts, targets were codon optimized for *S. pombe* and *P. pastoris*. While the transformation of the synthetic genes worked well in *S. pombe* the transformations in *P. pastoris* was not that easy: Apparently, the expression level of the optimized genes was too high and therefore toxic for the cells. It was not possible to express the proteins as planned with the strong constitutive *GAP* promoter; instead, the methanol inducible *AOX1* promoter was used.

Next steps (not part of this thesis) have been DasGip bioreactor cultivations for the expression host comparison. Moreover, the proteins obtained by this cultivation had been characterized by the enzymatic and the complementary assay. Overall our results indicate, that *F. fomentarius* targets are challenging targets to be expressed in yeast systems

6. Literature

1. **Andryushkova, Alexandra.** *Establishing fission yeast as a host for eukaryotic genomic libraries.* Graz : s.n., 2009. Dissertation.

2. **Sipiczki, Matthias.** Where does fission yeast sit on the tree of life? *Genome Biology.* 2000, Vol. 1, 2, pp. 1011.1–1011.4.

3. **Kurtzman, Cletus.** Taxonomy and phylogenetic diversity among the yeasts. [ed.] Per Sunnerhagen and Jure Piškur. *Using Fungi as Models.* Berlin : Springer, 2006, Vol. Topics in Current Genetics 15, pp. 29-46.

4. **Yeong, Foong May.** Severing all ties between mother and daughter: cell separation in budding yeast. *Molecular Microbiology.* 2005, Vol. 55, 5, pp. 1325-1331.

5. **Langegger, Maria and Berger, Jürgen.** S. pombe ascus releasing spores. [Online] [Cited: 2013 05-September.] http://www.fml.tuebingen.mpg.de/typo3temp/pics/b1102a0224.jpg .

6. **Berger, Jürgen.** Scanning electron micrograph of fission yeast (Schizosaccharomyces pombe). [Online] [Cited: 2013 05-September.] http://tuebingen.mpg.de/typo3temp/pics/52b13aa786.jpg.

7. **SciMAT.** Saccharomyces cerevisiae. (Scanning electron micrograph (SEM) of saccharomyces cerevisiae yeast (fungal) cells. [Online] [Cited: 2013 09-September.] http://www.visualphotos.com/image/1x3740051/saccharomyces_cerevisiae_scanning_electron.

8. **Balasubramanian, Mohan, Bi, Erfei and Glotzer, Michael.** Comparative Analysis of Cytokinesis in Budding Yeast, Fission Yeast and Animal Cells. *Current Biology.* 2004, Vol. 14, pp. R806–R818.

9. **Helston, Rachel, et al.** Schizosaccharomyces cryophilus sp. nov., a new species of fission yeast. *FEMS Yeast Research.* 2010, Vol. 10, 6, pp. 779-7786.

10. **Herskowitz, Ira.** Life Cycle of the Budding Yeast Saccharomyces cerevisiae. *Microbiological Reviews.* 1988, Vol. 52, 4, pp. 536-553.

11. **Goffeau, André et al.** Life with 6000 Genes. *Science.* 1996, Vol. 274, 5287, pp. 546+563-567.

12. **Wood, Val.** The genome sequence of Schizosaccharomyces pombe. *Nature.* 2002, Vol. 415, pp. 871-880.

13. **Simanis, Viesturs and Nurse, Paul.** The cell cycle control gene cdc2+ of fission yeast encodes a protein kinase potentially regulated by phosphorylation. *Cell.* 1986, Vol. 45, 2, pp. 261-268.

14. **Nobel Media AB.** [Online], 2014. [Cited: 2014 26-April.] http://www.nobelprize.org/nobel_prizes/medicine/laureates/2001/.

15. **Lee, Melanie and Paul, Nurse.** Complementation used to clone a human homologue of the fission yeast cell cycle control gene cdc2. *Nature.* 1987, Vol. 327, pp. 31-35.

16. **Parlati, Frank et al.** The calnexin homologue cnx1+ in Schizosaccharomyces pombe, is an essential gene which can be complemented by its soluble ER domain. *The EMBO Journal.* 1995, Vol. 14, 13, pp. 3064-3072.

17. **Rhind, Nicholas, et al.** Comparative Functional Genomics of the Fission yeasts. *Science.* 2011, Vol. 332, 6032, pp. 930–936.

18. **Petrescu-Danila, Elena, et al.** Fission yeast Schizosaccharomyces pombe as a producer and secretor of heterologous proteins. *Romanian Biotechnological Letters.* 2009, Vol. 14, 2, pp. 4201-4210.

19. **Westers, Lidia, Westers, Helga and Quax, Wim J.** Bacillus subtilis as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochimica et biophysics Acta.* 2004, Vol. 11, 1694, pp. 299-310.

20. **Çelik, Eda and Çalık, Pınar.** Production of recombinant proteins by yeast cells. *Biotechnology Advances.* 2012, Vol. 30, 5, pp. 1108-1118.

21. **Hitzeman, Ronals, et al.** Expression of a human gene for interferon in yeast. *Nature.* 1981, Vol. 293, 5835, pp. 717-722.

22. **Bitter, Grant, et al.** Secretion of foreign proteins from Saccharomyces cerevisiae directed by alpha-factor gene fusions. *Proc. Natl. Acad. Sci. USA.* 1984, Vol. 8, 17, pp. 5330–5334.

23. **Hartner, Franz and Glieder, Anton.** Regulation of methanol utilisation pathway genes in yeasts. *Microbial Cell Factories.* 2006, Vol. 5, 39.

24. **Cereghino, Geoff, et al.** Production of recombinant proteins in fermenter cultures of the yeast Pichia pastoris. *Current Opinion in Biotechnology.* 2002, Vol. 13, 4, pp. 329–332.

25. **Giga-Hama, Yuko, et al.** High-level expression of Human Lipocortin I in the Fission Yeast Schizosaccharomyces pombe Using a Novel Expression Vector. *Nature Biotechnology.* 1994, Vol. 12, 4, pp. 400-404.

26. **Takegawa, Kaoru, et al.** Production of heterologous proteins using the fission-yeast (Schizosaccharomyces pombe) expression system. *Biotechnology and Applied Biochemistry.* 2009, Vol. 53, 4, pp. 227-235.

27. **Hansen, Esben, et al.** De Novo Biosynthesis of Vanillin in Fission Yeast (Schizosaccharomyces pombe) and Baker's Yeast (Saccharomyces cerevisiae). *Applied and Environmental Microbiology.* 2009, Vol. 75, 9, pp. 2765-2774.

28. **Grimmm, Christian, et al.** Genetic enigineering of Schizosaccharomyces pombe: A system for gene disruption an replacemnet using the ura4 gene as a selectable marker. *Molecular and General Genetics.* 1988, Vol. 215, 1, pp. 81-86.

29. **Siam, Rania, Dolan, William and Forsburg, Susan.** Choosing and using Schizosaccharomyces pombe plasmids. 2004, Vol. 33, 3, pp. 189-198.

30. **Hottinger, Herbert, et al.** Nonsense suppression in Schizosaccharomyces pombe: the S. pombe Sup3-e tRNASerUGA gene is active in S. cerevisiae. *Molecular and General Genetics.* 1991, Vol. 188, 2, pp. 219-224.

31. **Giga-Hama, Yuko and Kumagai, Hiromichi.** Expression system for foreign genes using the fission yeast Schizosaccharomyces pombe. *Biotechnology and Applied Biochemistry.* 1999, Vol. 30, 3, pp. 235-244.

32. **Benko, Zsigmond and Zhao, Richard.** Zeocin for selection of bleMX6 resistance in fission yeast. *BioTechniques.* 2011, Vol. 51, 1, pp. 57-60.

33. **Hentges, Pierre, et al.** Three novel antibiotic marker cassettes for gene disruption and marker switching in Schizosaccharomyces pombe. *Yeast.* 2005, Vol. 22, 13, pp. 1013-1019.

34. **Invitrogen.** ZeocinTM Selection Reagent. [Online] 2012. [Cited: 2014 9-February.] http://tools.lifetechnologies.com/content/sfs/manuals/zeocin_man.pdf.

35. **Wright, Anthony, Maundrell, Kinsey and Shall, Sidney.** Transformation of Schizosaccharomyces pombe by non-homologous, unstable integration of plasmids in the genome. *Current Genetics.* 1986, Vol. 10, 7, pp. 503-508.

36. **Clarke, Lousie.** Centromers of budding and fission yeasts. *Trends in Genetics.* 1990, Vol. 6, 5, pp. 150-154.

37. **Maundrell, Kinsey.** nmt1 of fission yeast. A highly transcribed gene completely repressed by thiamine. *JBC - Journal of Biological Chemistry.* 1990, Vol. 265, 19, pp. 10857-10864.

38. **Maundrell, Kinsey.** Thiamine-repressible expression vectors PREP and pRIP for fission yeast. *Gene.* 1993, Vol. 123, 1, pp. 127-130.

39. **Basi, Gabriele, Schmid, Elisabeth and Maundrell, Kinsey.** TATA box mutations in the Schizosaccharomyces pombe nmtl promoter affect transcription efficiency but not the transcription start point or thiamine repressibility. *Gene.* 123, 1993, Vol. 1, pp. 131-136.

40. **Russell, Paul and Hall, Benjamin.** The Primary Structure of the Alcohol Dehydrogenase Gene from the Fission Yeast Schizosaccharomyces pombe. *The Journal of Biological Chemistry.* 1982, Vol. 258, 1, pp. 143-149.

41. **Watt, Stephen, et al.** urg1: A Uracil-Regulatable Promoter System for Fission Yeast with Short Induction and Repression Times. *PLoS One.* 2008, Vol. 3, 1, p. e1428. doi: 10.1371/journal.pone.0001428.

42. **Bellemare, Daniel, et al.** A novel copper-regulated promoter system for expression of heterologous proteins in Schizosaccharomyces pombe. *Gene.* 2001, Vol. 273, 2, pp. 191-198.

43. **Matsuyama, Akihisa, et al.** A novel series of vectors for chromosomal integration in fission yeast. *Biochemical and Biophysical Research Communications.* 2008, Vol. 374, 2, pp. 315-319.

44. **Käufer, Norbert, Simanis, Viesturs and Nurse, Paul.** Fission yeast Schizosaccharomyces pombe correctly excises a mammalian RNA transcript intervening sequence. *Nature.* 318, 1985, Vol. 6041, pp. 78-80.

45. **Toyoma, Reiko and Okayama, Hiroto.** Human chorionic gonadotropin a and human cytomegalovirus promoters are extremely active in the fission yeast Schizosaccharomyces pombe. *Federation of European Biochemical Societies - FEBS.* 1990, Vol. 268, 1, pp. 217-221.

46. **Faryar, Khaschayar and Gatz, Christiane.** Construction of a tetracycline-inducible promoter in Schizosaccharomyces pombe. *Current Genetics.* 1991, Vol. 21, 4-5, pp. 345-349.

47. **Forsburg, Susan.** Comparison of Schizosaccharomyces pombe expression systems. *Nucleic Acids research.* 1993, Vol. 21, 12, pp. 2955-2956.

48. **Kjærulff, Søren and Jensen, Martin.** Comparison of different signal peptides for secretion of heterologous proteins in fission yeast. *Biochemical and Biophysical Research Communications.* 2005, Vol. 336, 3, pp. 974-982.

49. **Grallert, Beáta, Nurse, Paul and Patterson, Thomas.** A study of integrative transformation in Schizosaccharomyces pombe. *Molecular and General Genetics.* 1993, 238, pp. 26-32.

50. **Matsuyama, Akihisa, Shirai, Atsuko and Yoshida, Minoru.** A novel series of vectors for chromosomal integration in fission yeast. *Biochemical and Biophysical Research Communications.* 2008, Vol. 374, 2, pp. 315–319.

51. **Matsuyama, Akihisa and Yoshida, Minoru.** Heterologous gene expression by chromosomal integration in fission yeast. *Methods in Molecular Biology.* 2012, Vol. 824, pp. 433-450.

52. **Penkett, Christopher, Birtle, Zoe and Bähler, Jürg.** Simplified primer design for PCR-based gene targeting and microarray primer database: two web tools for fission yeast. *Yeast.* 2006, Vol. 23, 13, pp. 921-928.

53. **Idiris, Alimjan, et al.** Construction of a protease-deficient strain set for the fission yeast Schizosaccharomyces pombe, useful for effective production of protease-sensitive heterologous proteins. *Yeast.* 2006, Vol. 23, 2, pp. 83–99.

54. **Imai, Yoshiyuki, et al.** Genes encoding farnesyl cysteine carboxyl methyltransferase in Schizosaccharomyces pombe and Xenopus laevis. *Molecular and Cellular Biology.* 1997, Vol. 14, 3, pp. 1543-1551.

55. **Kjærulff, Søren, Müller, Sven and Jensen, Roland.** Alternative protein secretion: The Mam1 ABC transporter supports secretion of M-factor linked GFP in fission yeast. *Biochemical and Biophysical Research Communications.* 2005, Vol. 338, 4, pp. 1853-1859.

56. **Braspenning, Joris, et al.** Secretion of heterologous proteins from Schizosaccharomyces pombe using the homologous leader sequence of pho1+ acid phosphatase. *Biochemical and Biophysical Research Communications.* 1998, Vol. 245, 1, pp. 166-171.

57. **Smerdon, Gary, Aves, Stephen and Walton, Fintan.** Production of human gastric lipase in the fission yeast Schizosaccharomyces pombe. *Gene.* 1995, Vol. 165, 2, pp. 313-318.

58. **Zárate, Victoria and Belda, Fernando.** Characterization of the heterologous invertase produced by Schizosaccharomyces pombe from the SUC2 gene of Saccharomyces cerevisiae. *Journal of Bacteriology.* 1996, Vol. 80, 1, pp. 45-52.

59. **Tokunaga, Masao, et al.** Secretion of mouse alpha-amylase from fission yeast Schizosaccharomyces pombe: Presence of chymostatin-sensitive protease activity in the culture medium. *Yeast.* 1993, Vol. 9, 4, pp. 379-387.

60. **Zhang, Wei, et al.** Enhanced secretion of heterologous proteins in Pichia pastoris following overexpression of Saccharomyces cerevisiae chaperone proteins. *Biotechnology Progress.* 2006, 22, pp. 1090-1095.

61. **Mukaiyama, Hiroyuki, et al.** Dextran sodium sulfate enhances secretion of recombinant human transferrin in Schizosaccharomyces pombe. *Applied Microbiology and Biotechnology.* 2009, Vol. 85, 1, p. 155.164.

62. **Mukaiyama, Hiroyuki, Tohda, Hideki and Takegawa, Kaouro.** Overexpression of protein disulfide isomerases enhances secretion of recombinant human transferrin in Schizosaccharomyces pombe. *Applied Microbiology and Biotechnology.* 2010, Vol. 86, 4, pp. 1135-1143.

63. **Parodi, Armando.** Reglucosylation of glycoproteins and quality control of glycoprotein folding in the endoplasmic reticulum of yeast cells. *Biochimica et Biophysica Acta.* 1999, Vol. 1426, 2, pp. 287-295.

64. **Parlati, Francesco, et al.** The calnexin homologue cnx1+ in Schizosaccharomyces pombe, is an essential gene which can be complemented by its soluble ER domain. *The EMBO Journal.* 1995, Vol. 14, 13, pp. 3064-3072.

65. **Maréchal, Alexandre, et al.** Cell viability and secretion of active proteins in Schizosaccharomyces pombe do not require the chaperone function of calnexin. *Biochemical Journal.* 2004, Vol. 380, Pt 2, pp. 441-448.

66. **Demain, Arnold, Vaishnav Preeti.** Production of recombinant proteins by microbes and higher microbes. 2009, Vol. 27, 3, pp. 297-306.

67. **Yamaguch, Kyoji, et al.** Effects of Site-directed Removal of N-Glycosylation Sites in Human Erythropoietin on Its Production and Biological Properties. 1991, Vol. 266, 30, pp. 20434-20439.

68. **Gerngross, Tillmann.** Advances in the production of human therapeutic proteins in yeats and filamentous fungi. *Nature Biotechnology.* 2004, Vol. 22, 12, pp. 1409-1414.

69. **Hamilton, Stephen and Gerngross, Tillman.** Glycosylation engineering in yeast: the advent of fully humanized yeast. *Current Opinion in Biotechnology.* 2007, Vol. 18, 5, pp. 387-92.

70. **Betenbaugh, Michael, et al.** Biosynthesis of human-type N-glycans in heterologous syystems. *Current Opinion in Structural biology.* 2004, Vol. 14, 5, pp. 601-66.

71. **Ohashi, Takao, et al.** The och1 Mutant of Schizosaccharomyces pombe Produces Galactosylated Core Structures of N-Linked Oligosaccharides. *Bioscience, Biotechnology, and Biochemistry.* 2009, Vol. 73, 2, pp. 407-414.

72. **Strahl-Bolsinger, Sabine, Gentzsch, Martina and Tanner, Widmar.** Protein O-mannosylation. *Biochimica et Biophysica Acta.* 1999, Vol. 1426, 2, pp. 297-307.

73. **Lussier, Marc, Sdicu, Anne-Marie and Bussey, Howard.** The KTR and MNN1 mannosyltransferase families of Saccharomyces cerevisiae. *Biochimica et Biophysica Acta.* 1999, Vol. 1426, 2, pp. 323-334.

74. **Ikeda, Yuka, et al.** Identification and characterization of a gene required for alpha1,2-mannose extension in the O-linked glycan synthesis pathway in Schizosaccharomyces pombe. *FEMS.* 2008, Vol. 9, 1, pp. 115-125.

75. **Giga-Hama, Yuko, et al.** Schizosaccharomyces pombe minimum genome factory. *Biotechnology and Applied Biochemistry.* 2007, Vol. 46, Pt 3, pp. 147-155.

76. **Idris, Alimjan, et al.** Enhanced protein secretion from multiprotease-deficient fission yeast by modification of its vacuolar protein sorting pathway. *Applied Microbiology and Biotechnology.* 2010, Vol. 85, 3, pp. 667-677.

77. **Forsburg, Susan.** Codon Usage Table for Schizosaccharomyces pombe. *Yeast.* 1994, Vol. 10, 8, pp. 1045–1047.

78. **Sánchez, Carmen.** Lignocellulosic residues: Biodegradation and bioconversion by fungi. *Biotechnology Advance.* 2009, Vol. 27, 2, pp. 185-194.

79. **Limayem, Alya and Ricke, Steven.** Lignocellulosic biomass for bioethanol production: Current perspectives, potential issues and future prospects. *Progress in Energy and Combustion Science.* 2012, Vol. 38, 4, pp. 449-467.

80. **Petterson, Roger and Gray, Richard.** The Chemical Composition of Wood. [book auth.] Roger Rowell. *The Chemistry of Solid Wood.* Washington, DC : American Chemical Society, 1984, pp. 57-126.

81. **Pérez, J., et al.** Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview. *International Microbiology.* 2002, Vol. 5, 2, pp. 53-63.

82. **Hammel, K E.** Fungal degradation of lignin. [book auth.] G Cadisch and K E Giller. *Driven by nature : plant litter quality and decomposition.* s.l. : CAB-International, 1997, pp. p. 33–46.

83. **Parveen, Kumar, et al.** Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production. *Industrial & Engineering Chemistry Research.* 2009, Vol. 48, 8, pp. 3713-3729.

84. **Schmidt, Olaf.** *Wood and Tree Fungi:Biology, damage, protection, and use.* Berlin, Heidelberg : Springer, 2006.

85. **Chen, Wei, et al.** Optimization for the production of exopolysaccharide from Fomes fomentarius in submerged culture and its antitumor effect in vitro. *Bioresource Technology.* 2008, Vol. 99, 8, pp. 3187–3194.

86. **Zizka, Z and Gabriel, J.** Primary fluorescence (autofluorescence) of fruiting bodies of the wood-rotting fungus Fomes fomentarius. *Folia Microbiologica.* 2006, Vol. 51, 2, pp. 109-113.

87. **Maurya, Nityanand Singh, et al.** Biosorption of dyes using dead macro fungi: effect of dye structure, ionic strength and pH. *Bioresource Technology.* 2006, Vol. 97, 3, pp. 512-521.

88. **Gabriel, Jirí, et al.** Heavy metal content in wood-decaying fungi collected in Prague and in the National Park Sumavain the Czech Republic. *Bulletin of Environmental Contamination and Toxicology.* 1997, Vol. 59, 4, pp. 595-602.

89. **Neifar, Mohamed, et al.** Purification, characterization and decolourization ability of Fomes fomentarius laccase produced in solid medium. *Journal of Molecular Catalysis B: Enzymatic.* 2010, Vol. 64, 1-2, pp. 68-74.

90. **Větrovský, Tomás, et al.** Extracellular enzymes of the white-rot fungus Fomes fomentarius and purification of 1,4-β-glucosidase. *Applied Biochemistry and Biotechnology.* 2012, Vol. 169, 1, pp. 100-9. doi: 10.1007/s12010-012-9952-9. Epub 2012 Nov 14.

91. **Bateman, Alex, Coggill, Penny and Finn, Robert.** DUFs: familie in search of function. *Acta Crytallographia Section F.* 2010, Vol. 66, Pt 10, pp. 1148-1152.

92. **Bähler, Jürg.** PPPP: Pombe PCR Primer Programs. [Online] 2014. [Cited: March 15, 2014.] http://www.bahlerlab.info/resources/.

93. **EMBL-EBI.** ClustalW2 web-tool for sequence alignment. [Online] 2014. [Cited: March 15, 2014.] http://www.ebi.ac.uk/Tools/msa/clustalw2/.

94. **Bioinformatics, Swiss Institute of.** Expasy Translate tool. [Online] 2014. [Cited: March 15, 2014.] http://web.expasy.org/translate/.

95. **Medicine, National Library of.** BLAST Basic Local Alignment Search Tool. [Online] 2014. [Cited: March 15, 2014.] http://blast.ncbi.nlm.nih.gov/Blast.cgi.

96. **Vienna, Institute for Theoretical Chemistry University of.** RNAfold Web Server. [Online] 2014. [Cited: March 15, 2014.] http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi.

97. **BitGene.** Gene analysis GC-distribution. [Online] 2010. [Cited: March 15, 2014.] http://www.bitgene.com/cgi/gene_analysis.cgi .

98. **Andreas Untergasser, Harm Nijveen.** Primer. [Online] 2007. [Cited: March 15, 2014.] http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi.

99. **Pawlowski, Nikolaus.** Sequence editor - convert DNA and RNA sequences. [Online] [Cited: March 15, 2014.] www.fr33.net/seqedit.php.

100. **Analysis), Henrik Nielsen (Center of Biological Sequence.** SignalP 4.1 Server. [Online] 2013. [Cited: March 15, 2014.] http://www.cbs.dtu.dk/services/SignalP/.

101. **Ikeda, Soichiro, et al.** Production of recombinant human lysosomal acid lipase in Schizosaccharomyces pombe: development of a fed-batch fermentation and purification process. *Journal of Bioscience and Bioengineering.* 2004, Vol. 98, 5, pp. 366-373.

102. **Overmann, Jörg.** Leibniz Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen. [Online] 2014. [Cited: 2014 25-April.] http://www.dsmz.de/catalogues/details/culture/DSM-70576.html?tx_dsmzresources_pi5[returnPid]=304INfo.

103. **Gibson, Daniel, et al.** Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods.* 2009, Vol. 6, 5, pp. 343-345.

104. **Harju, Susanna, Fedosyuk, Halyna and Peterson, Kenneth.** Rapid isolation of yeast genomic DNA: Bust n' Grab. *BMC Biotechnology.* 2004, Vol. 4, 8, pp. 10.1186/1472-6750-4-8.

105. **Bähler, Jürg, et al.** Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. *Yeast.* 1998, Vol. 14, 10, pp. 943-951.

106. **Abad, Sandra, et al.** High-level expression of Rhodotorula gracilis d-amino acid oxidase in Pichia pastoris. 2010, Vol. 33, 3, pp. 557-563.

107. **Matsuyama, Akihisa, et al.** pDUAL, a multipurpose, multicopy vector capable of chromosomal integration in fission yeast. *Yeast.* 2004, Vol. 21, 15, pp. 1289–1305.

108. **Astromskas, Eimantas and Cohn, Marita.** Ends-in vs. ends-out targeted insertion mutagenesis in Saccharomyces castellii. *Current Genetics.* 2009, Vol. 55, 3, pp. 339-347.

109. **Hastings, Philip J., et al.** Ends-In vs. Ends-Out Recombination in Yeast. *Genetics Society of America.* 1993, Vol. 135, 4, pp. 973-980.

110. **Chua, Gordon, et al.** Insertional mutagenesis based on illegitimate recombination in Schizosaccharomyces pombe. *Nucleic Acids Research.* 2000, Vol. 28, 11, p. E53.

111. **Harju, Susanna, Fedosyuk, Halyna and Peterson, Kenneth.** Rapid isolation of yeast genomic DNA: Bust n' Grab. *BMC Biotechnology.* 2004, Vol. 4, 8, pp. doi:10.1186/1472-6750-4-8.

112. **Forsburg, Susan.** Comparison of Schizosaccharomyces pombe expression systems. *Nucleic Acids research.* 1993, Vol. 21, 12, pp. 2955-2956.

113. **Piškur, Jure, et al.** How did Saccharomyces evolve to become a good brewer? *Trends in Genetics.* 2006, Vol. 22, 4, pp. 183-186.

114. **Ruohonen, Laura, Aalto, Markku K and Keränen, Sirkka.** Modifications to the ADH1 promoter of Saccharomyces cerevisiae for efficient production of heterologous proteins. *Journal of Biotechnology.* 1995, Vol. 39, 3, pp. 193-203.

115. **Sakurai, Masao, et al.** A distinct type of alcohol dehydrogenase, adh4+, complements ethanol fermentation in an adh1-deficientstrain of Schizosaccharomyces pombe. *FEMS Yeast Research.* 2004, Vol. 4, pp. 649–654.

116. **Forsburg, Susan and Rhind, Nicholas.** Basic methods for fission yeast. *Yeast.* 2006, Vol. 23, 3, pp. 173-183.

117. **Suga, Minoru and Hatakeyama, Toyomasa.** High efficiency transformation of Schizosaccharomyces pombe pretreated with thiol compounds by electroporation. *Yeast.* 2001, Vol. 18, 11, pp. 1015-1021.

118. **Mitrović, Aleksandra.** *Discovery and Optimization of Fungal Lignocellulolytic Enzymes.* TU Graz. Graz : s.n., 2013. Dissertation.

119. **Thallinger, Gerhard.** Cluster Control. [Online] 2014. [Cited: 2014 10-March.] https://acib.genome.tugraz.at/clustercontrol/.

120. **Foreman, Pamela, et al.** Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus Trichoderma reesei. *The Journal of Biological Chemistry.* 2003, Vol. 278, 4, pp. 31988-31997.

121. **Ikeda, Soichiro, et al.** Production of recombinant human lysosomal acid lipase in Schizosaccharomyces pombe: development of a fed-batch fermentation and purification process. *Journal of Bioscience and Bioengineering.* 2004, Vol. 98, 5, pp. 366-373.

122. **Bentley, Williarm, et al.** Plasmid-encoded protein: the principal factor in the "metabolic burden" associated with recombinant bacteria. *Biotechnology and Bioengineering.* 1990, Vol. 35, 7, pp. 668-6681.

123. **Matsuzawa, Tomohiko, et al.** Promotion of glycerol utilization using ethanol and 1-propanol in Schizosaccharomyces pombe. *Applied Microbiology and Biotechnology.* 2012, Vol. 95, 2, pp. 441–449.

124. **de Jong-Gubbels, Patricia, van Dijken, Johannes and Pronk, Jack.** ethanol, Metabolic fluxes in chernostat cultures of Schizosaccharomyces pombe grown on mixtures of glucose and. *Microbiobgy.* 1996, Vol. 144, pp. 1399-1407.

125. **Lin-Cereghino, Joan, et al.** Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast Pichia pastoris. *BioTechniques.* 2005, Vol. 38, 1, pp. 44–48.

126. **Hatfield, G. Wesley and Roth, David A.** Optimizing scaleup yield for protein production: Computationally Optimized DNA Assembly (CODA) and Translation Engineering. *Biotechnology Annual Review.* 2007, Vol. 13, pp. 27-42.

127. **Davis, James and Olsen, Gary.** Modal Codon Usage: Assessing the Typical Codon Usage of a Genome. *Molecular Biology and Evolution.* 2010, Vol. 27, 4, pp. 800-810.

128. **Nakamura, Yasukazu.** Codon Usage Database. [Online] 2007. [Cited: 2013 05-September.] http://www.kazusa.or.jp/codon/.

129. **Kaeberlein, Matt.** Lessons on longevity from budding yeas. 2010, Vol. 464, pp. 513-519.

130. **Matsuyama, Akihisa and Yoshida, Minoru.** Heterologous Gene Expression by Chromosomal Integration in Fission Yeast. *Methods in Molecular Biology.* 824, 2012, pp. 433-450.

131. **Schweingruber, Anne-Marie, et al.** Regulation of pho1-encoded acid phosphatase of Schizosaccharomyces pombe by adenine and phosphate. *Current Genetics.* 1992, pp. 289-292.

132. **Iacovoni, Jason, Russell, Paul and Gaits, Frédérique.** A new inducible protein expression system in fission yeast based on the glucose-repressed inv1 promoter. *Gene.* 1999, Vol. 232, 1, pp. 53-58.

133. **Matsuzawa, Tomohiko and Tohda, Hideki.** Ethanol-inducible gene expression using gld1+ promoter in the fission yeast Schizosaccharomyces pombe. *applied geneticsand Molecular biotechnology.* 2013, p. [Epub ahead of print].

134. **Matyuyama, Akihisa, et al.** pDUAL, a multipurpose, multicopy vector capable of chromosomal integration in fission yeast. *Yeast.* 2004, 21, pp. 1289–1305.

135. **Sakurai, Masao, et al.** A distinct type of alcohol dehydrogenase, adh4+, complements ethanol fermentation in an adh1-deficient strain of Schizosaccharomyces pombe. *FEMS Yeast Research.* 2004, Vol. 4, 6, pp. 649-654.

136. **Strahl-Bolsinger, Sabine and Scheinost, Alexandra.** Transmembrane topology of pmt1p, a member of an evolutionarily conserved family of protein O-mannosyltransferases. *The Journal of Biological Chemistry.* 1999, 274, pp. 9068-90675.

137. **Hajjar, Fadi, Beauregard, Pascale and Rokeach, Luis.** The 160 N-terminal residues of calnexin define a novel region supporting viability in Schizosaccharomyces pombe. *Yeast.* 2007, Vol. 24, 2, pp. 89-103.

138. **Cregg, James, et al.** Recombinant protein expression in Pichia pastoris. *Molecular Biotechnology.* 2000, Vol. 16, 1, pp. 23-52.

139. **Jaszek, Magdalena, et al.** Ligninolytic enzymes can participate in a multiple response system to oxidative stress in white-rot basidiomycetes: Fomes fomentarius and Tyromyces pubescens. *International Biodeterioration & Biodegradation.* 2006, Vol. 58, 3-4, pp. 168-175.

140. **Mellitzer, Andrea, et al.** Expression of lignocellulolytic enzymes in Pichia pastoris. *Microbial Cell Factories.* 2012, Vol. 11, 61, pp. doi:10.1186/1475-2859-11-61.

141. **Hoffman, Charles and Winston, Fred.** Glucose repression of a of the Schizosaccharomyces pombe fbpl gene occurs by a cAMP signaling pathway. *Genes & Development.* 1991, Vol. 5, 4, pp. 561-571.

7. Appendix

7.1. Plate pictures of *S. pombe* strains harboring integration vectors

The transformants were picked and cultivated in 300 µL YES containing 100 mg/L Zeocin™ in a 96 deep well plate. The cells were stamped on different plates to check integration and the function of the Zeocin™ resistance gene.

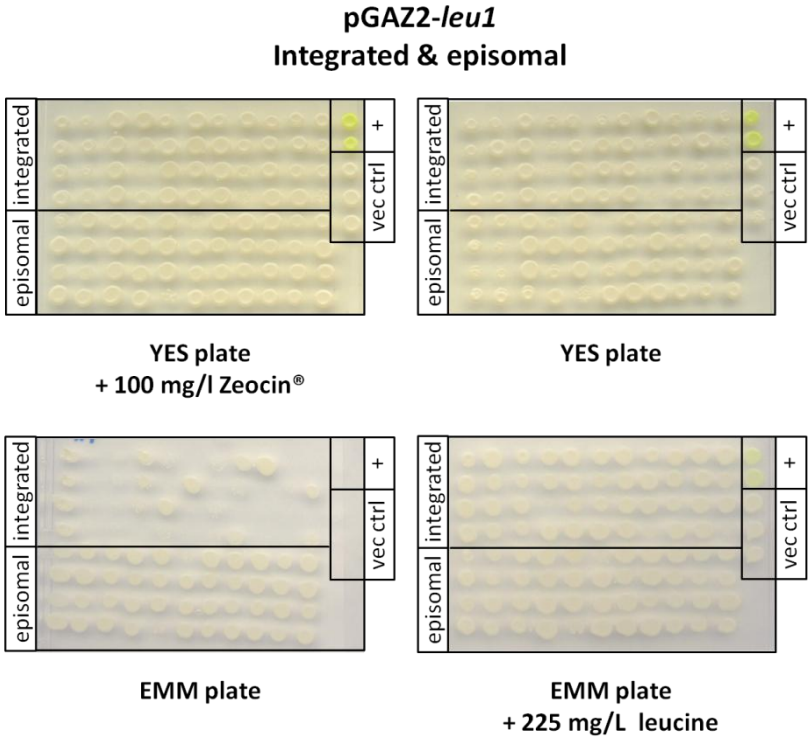


Figure 38: pGAZ2_*leu1* (empty vector) was transformed as linear construct and in its circular form in a *S. pombe* wild type strain.

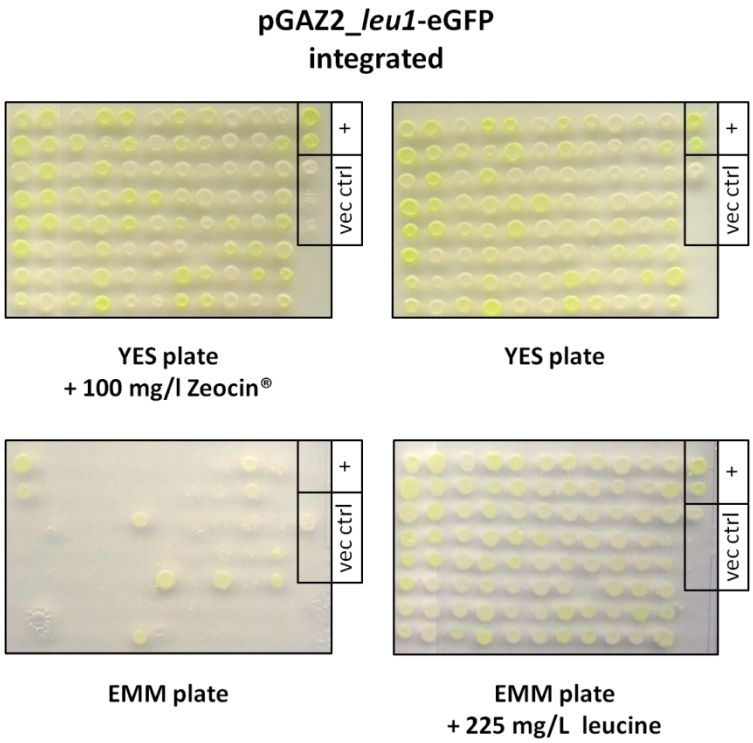


Figure 39: pGAZ2_*leu1*-eGFP was transformed as linear expression construct into *S. pombe* wild type strain.

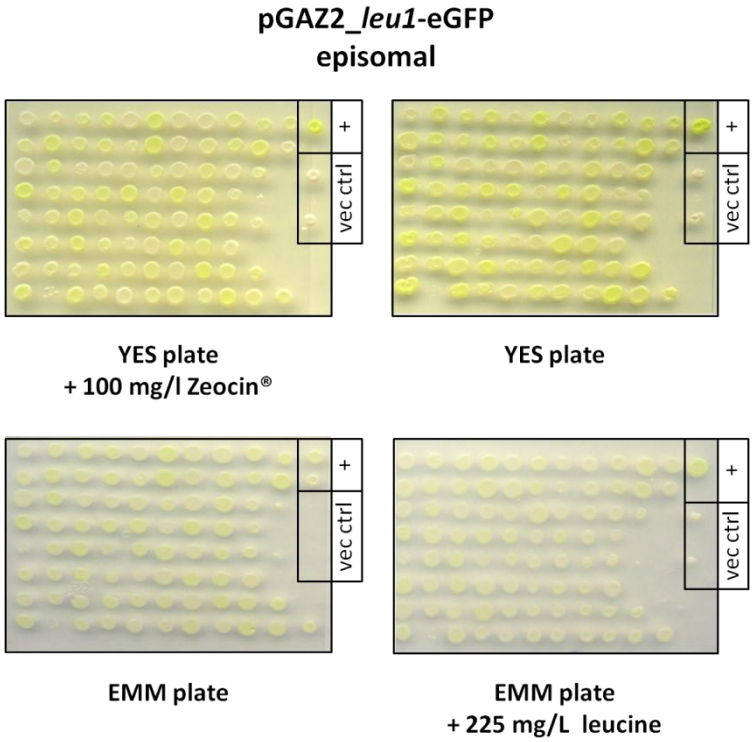


Figure 40: pGAZ2_*leu1*-eGFP was transformed as episome into *S. pombe* wild type strain.

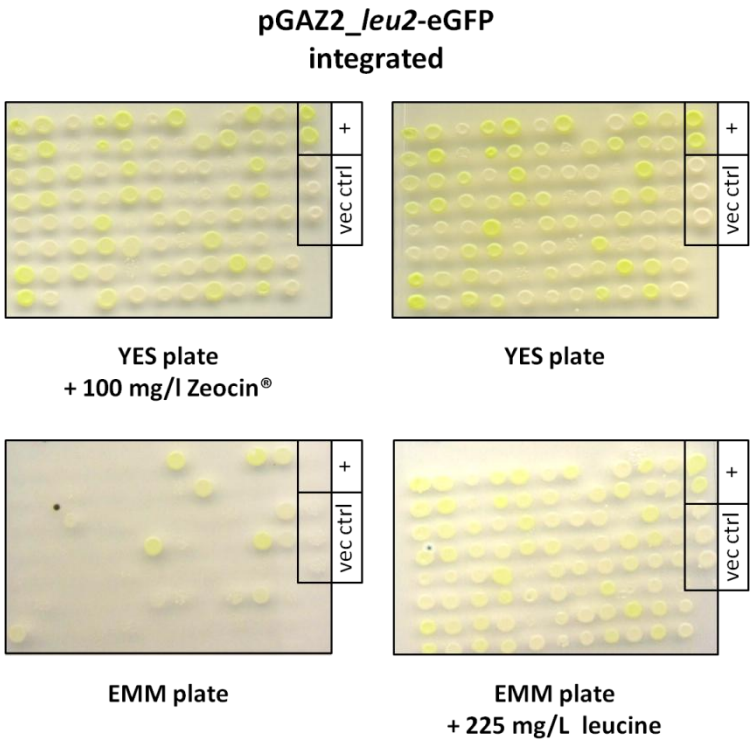


Figure 41: pGAZ2_*leu2*-eGFP was transformed as linear expression construct in a *S. pombe* wild type strain.

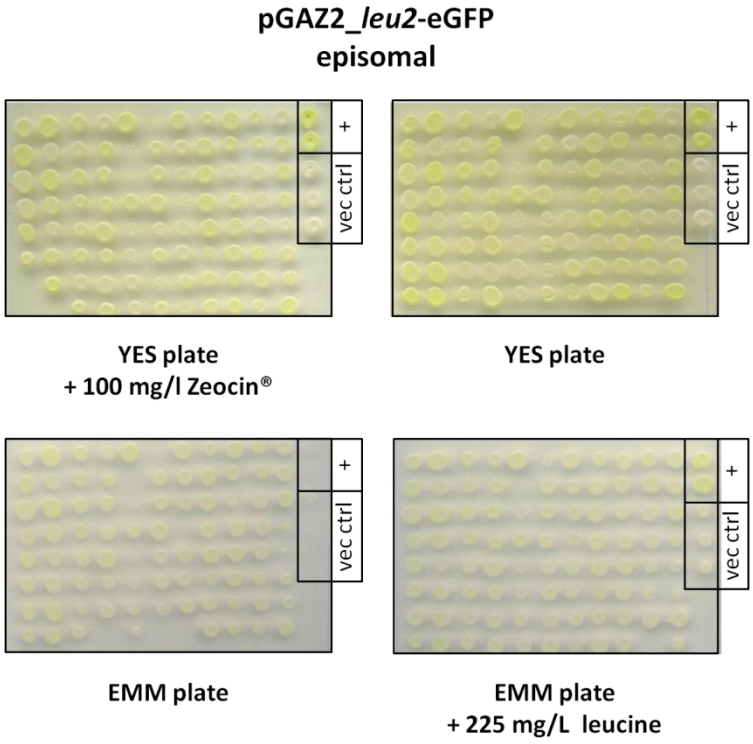


Figure 42: pGAZ2_*leu2*-eGFP was transformed as episome in a *S. pombe* wild type strain.

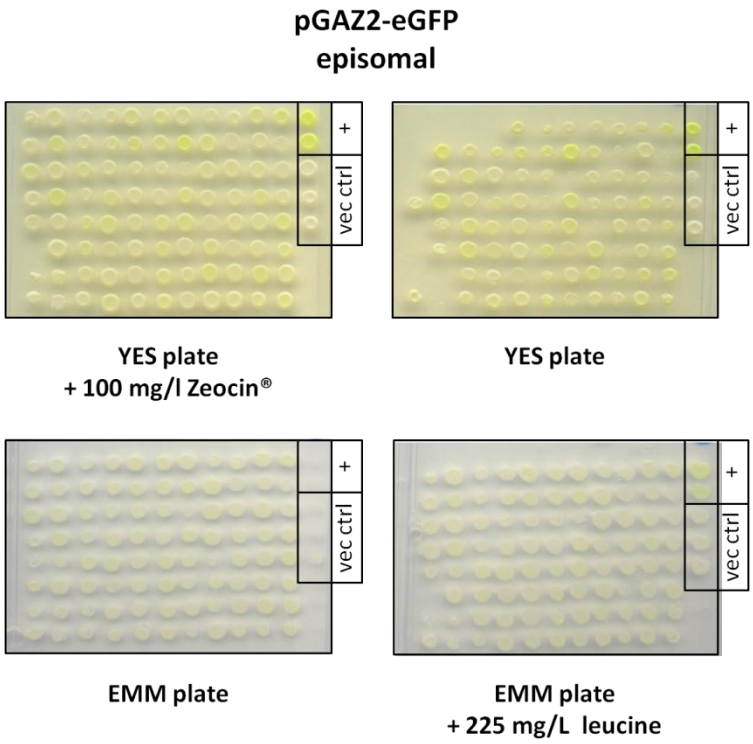


Figure 43: pGAZ2-eGFP was transformed as linear expression construct in a *S. pombe* wild type strain.

7.2. Fluorescence measurements of *S. pombe* strains harboring pGAZ2-eGFP and pGAZ2_*leu2*-eGFP

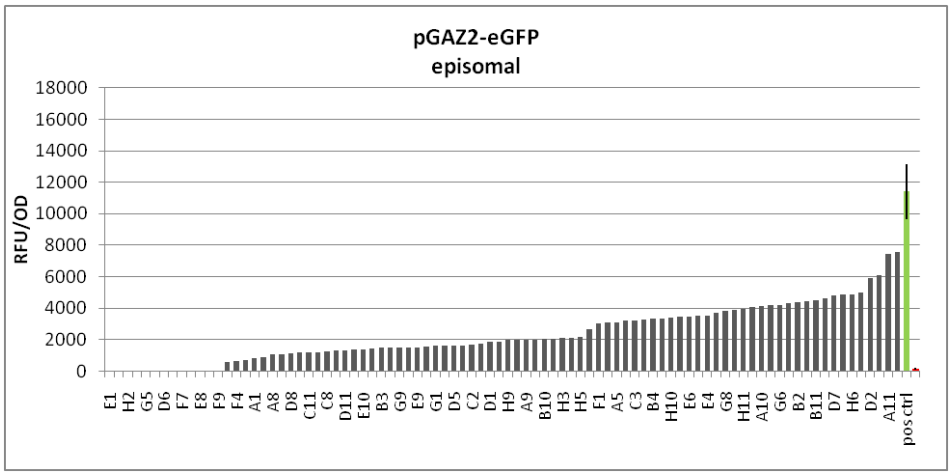


Figure 44: Fluorescence measurement of the *S. pombe* transformants showing pGAZ2-eGFP**.** As controls a *S. pombe* strain containing the pGAZ2 vector (vec ctrl, red) and an eGFP control strain (pos ctrl, green. The controls were picked four times, whereas the picked transformants were cultivated in one well.

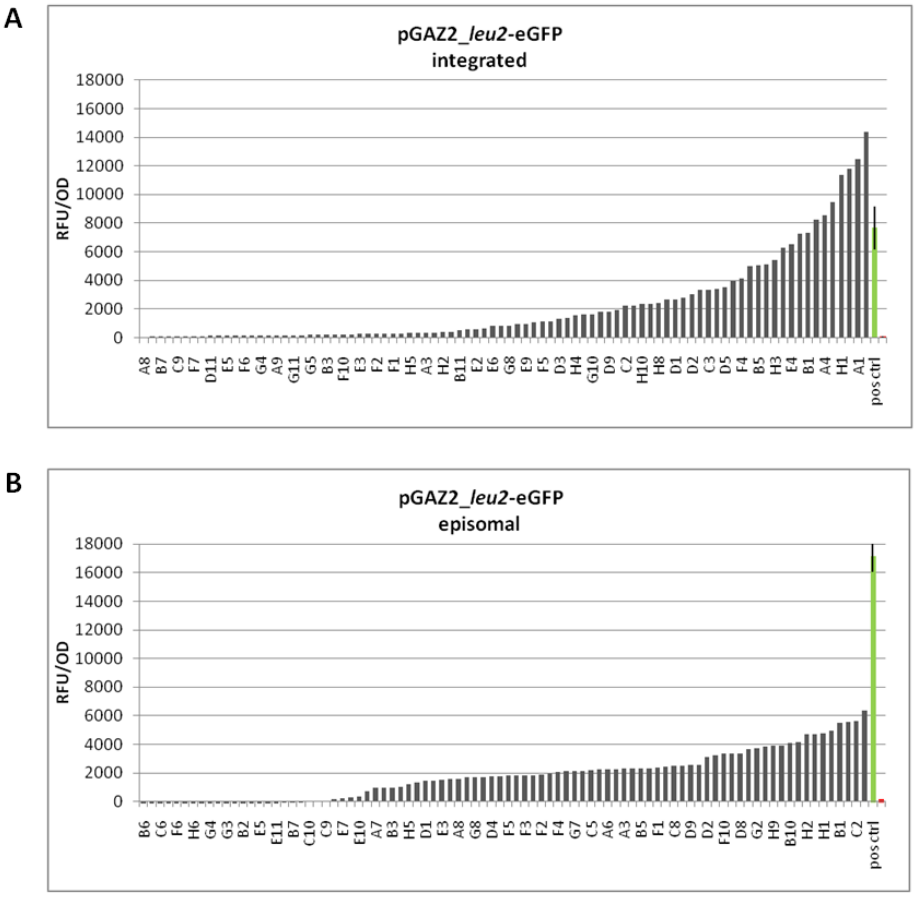


Figure 45: Fluorescence measurement of the *S. pombe* transformants showing **A** the integrated pGAZ2_*leu2*-eGFP and **B** the episomally expressed pGAZ2_*leu2*-eGFP transformants. As controls a *S. pombe* strain containing the pGAZ2_*leu2* vector (vec ctrl, red) and an eGFP control strain (pos ctrl, green. The controls were picked four times, whereas the picked transformants were cultivated in one well.

7.3. DNA sequences of cDNA targets

IDN5 (isotig04428)

ATGTGCTACAGGACGACTTGTCGACTTGTATCCAGCACGATGAGCATCGTTCCTCTAGCAGTCACCGCGATCATCGGACTG
GCCCATTCAGCTTGCGCATTTGACCCTCCCTCACTCCCACTCAACTGGACCACGCAGTTCGCCTGCGCCGTCGACATCCCGT
CCCGCATCATCACCAGTGACGTGACCACGCAGTACACCGACAACACCCCAGCGACATGCATTCAACGCTGCGATGCTGCGA
ATTTCACGTACGCCGGCGTTGAGTACACCAATGAGTGCCACTGCGGCACGGGTCTCAAGGGCACTCCGACTGCCGCCCCG
ACCTCCGAGTGCGACATGCCGTGCTCTGGCGATCCAGACTTGTCATGCGGAGGGGCCTTCCGCATCCAGATCTACAAATCC
CCCGCCTTGCCTCCAGGCGCATGGTCCCTGCAAGGCTGTTTCGTCGATACCGTGGACTCGCCAACGTTTGGCAGCCCGGTC
GTGCACCATACCTTCGCCTCGAACCTGAACCTCGTTGACCAGTGTATCAACTTCTGCCAACACATTGGCCTCCCGTTCTCGG
GTGTAGAGAGCGCGAATGACTGCCAATGCAGCCTCGGGCTCACTCCTACCGTCGTGCGTGCGCCCCAATCGGAGTGCAAC
TCAACGTGTCCCCTTCCACCGGGGGCGGGGAACGAGTTCTGTGGAGGTGTGCAGAGGGTGCAGGTGTACAAATATCTCG
GGTAA

IDN6 (contig01698)

ATGCTGAGGTTTACAATCTCCTTCTTACTCCTCACCATTGCGTACCTCCTATCATCCACGATCGCAATCCCTTTCAACCCTGGT
CTACTCCCCTACGGGTGGTTCACGGAATACGAATGCGCCACCGACACGCCCTCTCGTGTCTTGGTCAATATCCAATCTAACG
ACATGGGCAGCAACAACACCGCCTTCCTCTGCACCACGAAGTGCGCGTTCAATAACTTCATCTACGCCGGCGTCGAGGCCA
CCAACTGGTGCTATTGCGGCACTGGCCTCAACGGGTCGTCCCTGCCACAGACCGCCCTCATCACGGACTGCAGTGAACCGT
GCACCGGCGATCCCAACCGCTCGTGCGGAAACCATTGGAGGATCCAGATTTATAGGAATTACATAACCTCAACGGTCCCAA
ACCCGCTACCGACGAAAGCGGCAAGCATGGAGCTTCTTGAATCCCAGATCAAATAG

IDN7 (isotig04926)

ATGAAGTCCTTTACAACCCTCATCTCCATCGCACTCCTTGCTACCTCTGCACTGGGCTACGGGGCCGGAATGCAGAAGGTC
CATCGCCGTCAGGCTCCCTCCGCCGTGCCAGCACCAGCACCGGCTTCTGTTGAGGTGTCTATATCACCTCGTGCGGCAGGG
AATAAGCGTGGCCTGTCTTTTAACGATGCAGGCCTGACGAGGAGTTTTAACTCTGGCCAAGTCTCGTGGGCGTACAACTG
GGATTCTGCCTACCAGGGTACGCTCCCTGCCGGAGTGACCTATTTCCCCATGCTCTGGAGTGCTGCTCCGGACCACAGCAA
CCAGTGGGCAGCCAACGCAAACGCTGCGATTGCCAAGGGAGCCACGCATCTCCTCTTTATAAACGAGCCCGACCTCGCCC
AGCAGTCCAACACGAGCCCCGCAGACGCTGCGAGGCAATGGATGCAGTTCATGCAGCCATTCGCCGGCCGCGCCACGCTC
GTCTCCCCTGCGATCACCAACGGCGCTCCGCCGGCGATGGGCACGGGTTGGATGGATTCGTTCCTGAACGAGTGCAACAA
GCTCGGGTGCAGGGTCGACGCCATTGCCGTGCACATCTACGACTCGGCGACGAACATCGCGTACTACAAGCAGTACGTCA
CCGACCTCGGCAAGTACAACCGTCCGGTGCTCCTCACCGAGTTCGGCGCGGCGGGCGCGCTCGACCAGCAGCAGCAGTTT
TTGAGGGAGTTCATCCCCTTCCTCAACGGCGTGTCCCACGTTAGCCACTTTGCGTGGTTTATGACTGCCGTCGGCAACCTTG
TCAACCAGGACGCGAGCCTCACGGCTCTCGGCCAGACGTACGTCTCGGCGTAG

IDN8 (isotig05085)

ATGCGCTTTTCTCTCTTTCTCAAAGCTCTCTCTCTGGTTGGGGTTCTCGCTGCGGCGCCGCATGTCCACGCGCAGCCGAGGC
CCGCGAACCTCGCAGCGACGTCGACGTCCGGGTCACTCAGCGCGCTCAAGTACACCAAGGTCGGTGGTGCGGGCTCGTAC
GACCAGGTTGTGGACATCAAGCCGGGCACGTTCCCCAGCTGTGACGTGAGCCCGTCTTGCGTCACCCAGCCCAAGGCGAT
CAGCGGGAACTTGGCACCGTTCGATGACGAGGTGACGGTGGCGTTCCGTGGGCCGATGAATATCTACAATATCGCGGTGT
ACCAGCCGACGAGCACGAATTCGAGCTCGGGTAGTTGGAAGCGCGTCAGTTCGTGGGCGGCGGGCGAGCAGCCCGAGA
ACGTCGTATTTATGAATAACATGGGCGGGAGTAAGAGCGGGGAGTGGTCGATCTGCGCGGGATTGAGTCAGTCGTATGC
CAATGGCGACTGGACGGACGCGGCGAGCAAGCCGAACGCAGACATCGCGAAGGGCTATCTTGACCAGGACCACGAGATC
AACATCATGACGTCCGAGACCTGCGCCTCCCGCTCGTGTGCCGGCTTCACGCGCGGCACGGCGAACCACGGCTGGGCAGG
CAGCAAGCTCTTCGTGTTCACCTTTGACATGCCCGCCTCCTCCGACGCATCCAAGACGCCCGCGATCTGGGCTCTGAACGC
ACAGATCGTGCGCGCGGCGCAGTATGGGTGTAACTGCCGCGGCGTCGGGCCCGGGGGGTGCGGCGAGCTCGATATTGTC
GAGACCGTCTCCTCCGTGGACCCCGCGAACAAGGGCATCTCGGAGATCTACTCGCCCAAAGGCGCGACGGGCTCGGGTG
CGAGCTTCTTTGCACGCCCGACGGGCGGGAAGGCGACG

IDN9 (isotig05070)

ATGTTCGCCCAATCTCTTGTGTCTTGCGCTCTCTTCGCGGCCGCGGCTCGCGCCGGCAGCATTGTTTGGGATGGGTCATTCA
ACCCGTTCAACACTGCGGCTGACTTCGACAAATGGTCGTGGGCGAATCAGACCGGAAGCTACCAATGGTACATCCATGGT
AGCCAGCCCACCTCCCACTACCTGACGCTCGACTCGGCCTACAAGAACCCTGCGCTGACAAATGAGGAGCATGGTCTAAAA
CTGACAATCGACTCTACCGCGACTTGGAATTCCAACATGGAGAGAGCCGAGCTCATCCCTCAGACCAGCCAGAACCTCGGC
ACGGGCAACCTCTTTTACCACTTCTCCATCTCGCATCCCGACACAAACACACCAGACTCCACTCTAGAGCACCAGATCCTGT
TCTTCGAGTCTCACTTCACAGAGCTCAAGTATGGCGTCGGCCCTAACACCACTAACCTCGAGTGGCACGTCGGCGGCCAGC
CCCAGTGGGGCGTCCCCTTCGCCGCGGGAACGTGGTTCAACTTCGCGTACGACATCGACTTTTCGGCCAGCACCGTCGGAC
TGTGGGCCTCCACCGGCTCTGCGCCGCTCACCAAGGTCGCACAGAACATCAAGGCATCGACCTCGACTAACTCTGCTGACT
TCCACGTCGGTGTGCTCCGCATTGTCAACACTCCCGCGGCCGAGGACTGGTTCATCTCTGGTGTGTACATCGAGCAGGGCC
CCATCACGACGACCATCGGCAGCGGCTCTTCTGCACCCAGCCCCCCCGTCACGTCGGCTCCGGTGACGACTGTGCCGCCGA
CCTCTGCTCCGTCCACCTCCGCGCCGCCTTCCACGACCTCGTCTGCAAGTGGTCCTCTCGCATCTCAGTGGGCCCAGTGCGG
AGGTACTGGGTACACTGGCCCCACCGCCTGCGCGAGCCCCTACACATGCACGGCTGTCTCGCCCCCGTACTACTACCAGTG
CCAGTAG

IDN10 (contig02057)

ATGCAGCTCTTTACCGCAGTCAGTGTTGCCTTCTCTCTTCTGCACATTGTCCGTGCGCAGGCTCCCGTTTGGGGCCAGTGCG
GTGGTCAGGGCTGGACTGGAGCAAGTACCTGTGTTTCTGGTACGGTATGCACGGTCCTGAACGCATACTACTCGCAATGT
CTTCCTGGTACCGCTTCGTCTGCCCCTGCGCCCCCTTCGTCCACCACCGCCACTGTCCCCACTACTACGAGCAACCCTGCTCC
TCCCGTGACGACGAGCAACCCCGTTCCTGGCTCTTGTTCCACCCCCAGCACAGTCTCCGGCTTTAGCAACACCAAACTCCCG
AACCCTTTCAAATTCAACGACGGTAGCGCGGTCACGACCCAGGATGACTGGGAGTGCAGGCGTGCACAGATCCTCAGCTT
GGTCCAGGGCTACGAAGGTGGCGCCCTCCCGGGCAAACCCTCTTCGCTTACTGCACAGCTCACCAAGTCCGGCACCTCCGC
GACGCTCGCTATCACCGCGAGCGATGGCGGCAAGTCGATCACCTTTTCGCCGAAGATCACCTATCCGTCTGGCAACCCGCC
CGCTGGTGGCTGGCCCGTCATCATCGTCTATGAGGGACTCAGCTTGCCCGTCCCTTCTGGTGTTGCCCAACTCGTGTACCA
GAACTCGGAGATGGGTGCCCAGGACTCCACGGCTAGCCGCGGCCAGGGTCTCTTCTTCAACCTGTATGGCAGCAACGCAA
CCGCCTCTGCGATGACCGCCTGGGCATGGGGTGTCAGCCGCATCATTGATGTACTCGAGTCGACGCCCAACGCGCAGCTC
AACCCACAGAAGGTCGCCGTCACTGGCTGCTCGCGCAACGGCAAGGGCGCGCTCGTGGCTGGCGCACTGGACACCCGTA
TCGCGCTCACGCTCCCGCAGGAGTCCGGCTCGGGAGGTGACGCATGCTGGCGTCTGTCCAAGTTCGAGCAGGATTCAGGT
AGTGTGGTGCAGACGGCAACAGAGATCGTCGGCGAGAACGTCTGGTTCTCGACGAACTTCAACAACTACGTCAACCAGCT
GAGCGTGCTCCCGTACGACCACCACCTGCTCGCGGCGCTCGTCGCGCCTCGCGCGATGATCTCGTTCGAGAACACCGACTT
CGTGTGGCTGAGCCCGCTCTCGTCCTTCGGATGCATGACCGCGGCGCACACCGTCTGGGACGCGCTTGGCGTAGCGGACC
ACCACGGGTTCGAGCAGGTCGGCGGACACGCGCACTGCGCGTGGCCGTCGTCGCTCAACAACCAGCTGAGCGCGTTCTTC
AATCGGTTCCTGCTCGACCAGAACGTCAGCACGACTTTCTTTGAGTCGAACAAGATCTTCAACGGCGTCACCTGGAACCAG
GCCAACTGGATCGACTGGACCACGCCCAAGCTCACTTGA

IDN13 (isotig02488)

ATGTTCGCTCGCTCTGTCGTCGCTGCCTTCGTTGCCGTTCCCTTCCTTGCTCGCTCTGCCTTCGCGGCGGACTGCACTCGCAC
GTACACCGTCAAGGAGGGTGACTGGTGTGACACGATCTCCGCTGCTCACAATGTCTCCTCCTACCAGCTCGCCGCTGTGAA
CCCGACAATCGATGACCTCTGCCACAACCTCCAGATCGACCAGGTTCTCTGCCTGGGTACCGAGGGAGAGGATTGCACCA
ACACCCACATCGTTGCCGAGAATGATTCCTGCGAGTTCATCTCCAGCACTTATGGTCTTAACTCGACTGTCCTCTTTCACAAC
AACCCTCAGATCGATGCCGAGTGCAGCAACATCTACACCGGCGAGGTTCTCTGCGTTGCTACCGCGGCTCTCGCTCCCGCT
GTCCCGGCGGGCTTCGTTGCCCCTTCCGCGAGCGCCGCTGCTCCCGCTTCCACCAAGGCGGCGGCTCCTACTCAGCCCGCT
GCTCCTACTCAGCCCGCTGATGACGATGATGATGACGAGGACCTCCCCTGGTGCGATGAGCTCTAA

IDN16 (isotig03896)

ATGAAGTTCACGTACGCTCTTGCTGCCGTTGTCGCTGCCGCTGCCTCTGTGCAGGCGACTGAGACCAATGGCCAACGCTTG
GCCCGTGGTCTCAACCCTCTTCCTCCTCGCAACCTCCATCACCACGGGTCCCGCACTTCCAACGCGAAGCCGGCTACCCCAT
CTGGCTCCTCTGGTGACTACAACTGCAACACTGGCCCCGTACAGTGCTGCGAGTCGGTGACGAAGGCACACGATAATGTC
ATCACTGCCATTCTCGGTCTTCTCAACATCCCCATTGTCCCCGATCTGCTCGTTGGCCTCAAGTGTTCGCCTCTGAGTGTTGT
CGGTCTGGGAAGTGGAAACGCTTGTCGCCAGCGCCCAGTCTGCTGTGAAAACAACAGCCATGGTGGTCTCATTTCTATTG
GTTGCATCCCGATCATGCTGTGA

IDN24(isotig03302)

ATGACGCGCATCCTGCTTCCCCTCGTCGCCGCTTTTCTCTTCTCCGATGTGTCAACTGCCGCACCGGCATCGGCTGCGTCGA
GTGCTGCGGCCGCTGTTCCCCTGAGTGCAGCATCGAGCGTTGCAGCCTCCGCTATGTCGGGTGCCGTAGTGCCAGTTGCAT
CCGGCGCCGTATCCTCCTTTGCGTCGAGTCCTTTGGCCACTCGCACAAGTGCCGTTGCCTCTAGTGCCGGTGCACCACCACC
ACCGGCTGCGACTGTACCACTGGCCAGCGACGACCCGAACCCTATCCTCTTCCTGCCTGACACCGATATCATCCCTACCGCT
CAACGTGGACAGCTCGGAGCGAGCGTCATCGGCCCACAGAACGTCCCTATCGACCGCCAGAATGCAGATCTACTTGCACC
TCCGACTACTGACCACGGCACCGTGATGAATGCCAAATGGCCGTTCGCTCTGAGCCACAACCGTCTGCAGACGGGAGGTT
GGGCCCGAACTCAGAACCGGGACCAGATGCCCGTTGCGCAGTCAATGGCTGGTGTGAACATGCGGCTCCTACCTGGTGCC
ATCCGCGAGCTCCATTGGCACACTACTTCTGAGTGGGCGTACATTCTCAAAGGGACCGTTCAAGTCACCGCAATAAACACG
GACGGTCAGAACTTCATCGGCACTGTGAATCCAGGCGACCTATGGTACTTCCCTGCGGGCATGCCGCACAGCTTGCAAGCT
ACCAACGACACGGAGGACGGGGCTGAATTTGTTATTGCGTTTGATGACGGCGCGTTCGACGAAGATTCGACGTTCTTGCT
CACTGATTGGTTGGCGCACGTTCCCAAGGAAGTAATTGCCAAAAACTTCCAGACCAGTGTCTCGGCCTTCGATCGCGTCCC
GAACCAACAACTATACATATTCCCCGGCGTTCCACCTGCGGACAACGCACCCGCTCCGCCAAGCCCACAAGGCACAATTCC
AGACCCGTTCACCTTCGCTCTCTCGCAGATCAAGCCTACTCAGCTTTCCGGCGGCACAATCAAGATTGTCGATGACAAGAT
GTTCAAGGCCGCCACTGGAATTTCTGCGATAGAGGTCACAGTCGAGCCTGGCGCGATGAGGGAACTTCATTGGCATCCCA
ATGCAGATGAATGGAGCTTCTTCATTAGCGGCCAAGGTCGCATGACGATCTTCGCAACGAATGGGAACGCCCGGACCTTC
GACTACGAGGCCGGAGATGTTGGCTTCGTTCCCCAGTCGTTCGGCCACCACGTCGAGAACACTGGGAATGATACCTTGCA
CTTCTTGGAAATTTTCAACTCGGGTACCGTTCAAGATGTTAGCTTAAGCCAGTGGCTCGCTTTGACGCCTCCCGAGCTCGTG
CAAGCGCACCTAGCACTTCCGGACGATGTGCTCAGCAAACTCAGCAATTTCAGGAAAAAGGCATTCGTCGTTGGACCTCAA
TAG

IDN40 (contig01925)

ATGTTCGCTATCATCTCTGTATTCCTCGTTGCTGCATTCACCGCAAAGTCCGTTAGTGCCCACGGCTACGTACAGGACGTTG
TCGTCGAAAACACTCATTACCCCGGTTACCTGCCTTATTCTGATCCTTACTACAACCCCGTCCCCGAGAGAATCATTCGTCCC
ATTCCTGGGAACGGCCCTGTGCAAGACCTGACTTTAATCGACCTGCAGTGCAATGGTTACTCTGAGGGTGGTGTCGTGGG
CTCTCAACCCGCTCCCATCTTCGCGACTGTGCCCGCTGGATCGCAGGTGGGGCTGAATTGGACCACTTGGCCCGACTCACA
TGTCGGACCCATCATCACCTACCTCGCTCGTGCCCCGAGCGATATCACCAATTGGCAGCCCGGCAATGCTGCGATTTGGTT
CAAGATCGCCGAGTCCGGCAAAACCGCCGACGGCAAGTGGGCGGCGACCGACGGGCTCTATGCTACTAACAGCATCTAC
ACGTTCACCGTCCCCCCAAAGCTCAAACCGGGCCAGTACATCGTTCGCCATGAGATCATCGCGCTGCACGCCGCGTTCCAG
TACCCCGGAGCGCAGTTCTACCCCTCGTGCATCCAAATCGAGGTGACGGGCTCTGGCACTGCGCTCCCTACCTCCGGCTTG
GTTTCCTTCCCTGGCGCGTATACCCCCAGCACCCCCGGAATCGTCTACGACGCTTACCAGAACACGAGCGCATACCCCATCC
CTGGCCCCGCCGTTTGGACTGGCGGGAACTGA

GH61_1 (isotig04623)

ATGATGAACCTCTCTGTTATTGCGAcCGCCGTGATCCTCGCTGGCTCTGCGTCAGCGCACACCATCTTCCAAAGGCTCtACG
TCAATGGCGCGGACCAAGGCCAACTGACCGGTATCCGTGCCCCCGATTCTAACAACCCCATAACGGACGTCAACGTGCAG
GACATCATATGCAATGGAATCAGTAACCCCTTCCACCAACCCGTGTCCAAGACCGTGATTGACGTGCCTGCTGGCGGACAA
GTCACTGCTGAGTGGCACCACATGCTATACTCGGTCCCGAACGACCAGGACGATCCCATTGCTGCGAGCCACCACGGCCC
GATCATCTCTTACCTTGCTAAGGTCCCtGACGCGACCCAGAGCGACGTCTCTGGCCTGAAGTGGTTCAAGATCTACGAGGA
CGGGTACTCGAACGGAGtCTGGGGTGTCGAtAAgATGGTTGCCaACAAGGGCAAGGTGACATTCAACATCCCGAGCTGCA
TCGAGTCTGGCCAGTACCTTCTCCGCCACGAGATCATCGCCCTCCACTCCGCCTACAACTACCCGGGTgCGCAGTTCTACAT
GGAGTGCGCGCAGATTAACGTCGTTGGAGGCTCTGGCTCGAAGACGCCTTCGAATTACGCCTCCTTCCCGGGCGCGTACA
AGGGCACAGACCCCGGCGTCAAGATCAACATCTACGACCCTCTATCAAACTACACGATTCCAGGACCATCTCTTTTCACGTG
CTGA

GH61_2 (isotig05034)

ATGAAGACCTTCGCTTCTCTCCTCGTTGCGACTGCACTTGCGACTTCTACCTTCGCGCACGCCATCTTCCAGGAGCTATACGT
GAATGGTGTTAGCCAGGGCCACCTGACGGGTATCCGTGTTCCCGACTATGACGGGCCGATAACGGACGTAACGTCCAACG
ACGTCATCTGCAATGGTGGTATCAATCCCTATCACCAGCCCATTTCGACCACTGTCATCCCTGTCCCCGGGGGTGCGCAGGT
CACCGCTGAGTGGCATCATACTCTTCAGGGTGCCGACCCGTCCGATAGTGCGGACCCTATTGATGCAAGCCACAAGGGAC
CGGTCATGGCCTACCTCGCGAAGGTCAGCAACGCGACCGACTTGAACGTGACTGGTTTGAAGTGGTTCAAGATCTGGGAA
GATGGTCTCGACAGCAACAAGGTGTGGGGTGTCGACCGTCTGATCGCCAACAAGGGCAAGGTCTCCTTCACTATCCCCAG
TTGCATTGAGGCTGGCCGATACCTCCTCCGTGTTGaACTTATcGCTCTCCACGGTGCCCAGAGCTACCCTGGTGCTCAGTTA
TACATGGAATGCGCTCAGCTCCAGATCACCGGTGGAGGCAGCACTAATCCTGCTACGGTTAACTTCCCCGGCGCGTACTCC
GGCTCGGACCCGGGCATCAAAATCAACATTTACCAGAACCTGCCCTCGTACACCATCCCTGGCCCCTCTGTaTTTTCCTGCG
ACGGTGGCTCGAGCCCTGAGCCGCCTGTCACGAGCAACCCCTCCGGTCCTACCTCTGTCCCGCCCACAACCACCTCAGTGC
CACCTGTGAGCACTGCTCCCGCTGGTACCGTTCCGCAATTTGGTCAGTGCGGTGGCATCGGCTACAGCGGCTCCACCGGTT
GCGTGGCGCCATACGCTTGCGTCAAGTCGAACGAATACTACAGCCAGTGCCTGTAG

GH61_3 (isotig03735)

ATGTTCTCGTCTACTTTCTCCCTCGTTGCCCTCGCGGCTGCCCTCGTTGCGCCCACCGCCGTCAACGCGCATGGCTACGTCC
ACTCACTCGATGTTGGCGGCAAGAACTACACAGGATGGCTTCCGTTTACCGATCCGTACGAATCGCCTGTTCCTAAGCGAA
TTGTCAGGAAGGTCCCTGATGACGGTCCCGTCATGTCTGCGACTTCTGCCGATATTGCTTGCAATAAGGGTGGCGAGTCTC
CTGCGGGCGCCATTGCGACCGCTCCTGCCGGAAGCAAGATCACTTTCGACTGGGTGAACTGGCCTAGCGACCATATGGGT
CCGGTTGACACCTACATGACTTCTTGCAACGGTGACTGCTCCAATTTCGATGCTTCGAACGCCAGGTGGTTCAAGATCGAC
GCTGCCGGTTATTCCAACGGCAAGTGGGCCGCGACTCAGCTCATCGAGAACGGCGCCCAGTGGACCAGCACCATTCCTGC
CGAGCTCAAGGCCGGTGAATACCTAGTCCGTCACTCAATCGTTGCTCTTCACTCTACTGGCGAGCCTCAGTACTACCCCAGC
TGCGCTCAGGTCAAGGTGACGGGCGGTGGTGGCAAGGTTCCGTCTGGCGATGAGCTTGTCTCCATCCCGGGCCTATACAA
CAGCGTCAAGTGGCCCGACATCTGGGACAACAACTTCAAGAAGTTCACTATCCCTGGCCCCGCAGTCGCATTCTCGGGCtC
CAGCAGCGGTTCCGGTTCCAACTCTGGCTCCGAGGAACCCGAATCCTCTTCCTCCGCAGGACATGCGTCCTCTACCCACGT
GTCCTCTACCCACGCGTCGGTTACCCCTTCAGCTTCATCGACGAAGGTCGCCCAGCCTACGTCGTCTCCGACCGAGTCTGAA
GGTTCCCAGGCGACGTCGATCGCCACCTCCTCGGCCCCCGCGTCGACGGGCAGGTGCTCGACCAAACGCCGTCGCGGCAT
GGTCAAGCGCCGTGTCTCCCACCACGCCAAGCGGCACCACTGA

GH61_4 (isotig06156)

ATGAAGCTCGCGCTCCTCTTCTCCCTGGCTACGGCCGCCATTGGCGCCAATGCGCACTTCACGATGCAGTACATCTGGATC
AACGGCGTTGACCAGGGTCAAAACAACTTCATCCGAATCCCGCCAAACAACAACCCAGTCACCGACGTCACCTCGACGGA
CATCACCTGCAACGTAAACGGCCTCTCGGGCGCGGGCGTGCAGACCGCCTCCATACCTGCGGGTGCGAACATCACCTTTG
AGTGGCgCCAACACGCTCaGCGTACTGGCGAGGACGCGATcTCGGGCGGTCACAAGGGACCAGTCCAGGTCTACATCGCC
AAGGCGCCGTCCACCGCCGCCTCCTTCGACGGCCAGGGCGCTGTCTGGACCAAAATCTATTCTTCTGGTCTCCTCAACGCG
CAGTCACAGGAGTGGGCTACCGATGTCGTCAACACTGACAGCGGCAAGCACAGCGTTGTGATTCCCTCATCCCTCGCTGCC
GGCGAGTACCTCCTCCGTGCTGAGGTCATCGCCCTCCACGTCGCGCAGAGTTACCCTGGTGCACAATTCTACATTGGCTGT
GCTCAGGTTAAGATTGCCAGTGGCGGCAATGCTAACCCCCcGAAGGTTGCTCTTCCGGGTGCGTACCGGAGCTCGGACCC
CGGCATCACTGTCAACATCTACAACAACCTCCAGAGCTACACTGCTCCTGGGCCCGCTGTCTGGTCCGGATGA

GH61_5 (contig01420)

ATGAAGTACTTCTCTGTTCTTGCCGCTCTTTGCGCCCTGCCCTATGTTGCTGCCCACGGCTTCGTGAGCTCAGTATCCATTGA
CGGAACGGAGTACGCCGGTAACGAGCCCAATCAGTACAAGGGTCCGAGCCCGATCCGCCTCATTTCTGACATCAGCCCTG
TTAAGGGCGCCTCCAACGCCGATCTCTTCTGCGGCTTGAAGGCCAAGCCCGCCGAGCTCGTTGTTCCCGCGAATGCTGGCA
GCGAAGTTACGTTCCAATGGTCTGGTGGAGGTGGCCAGAAGTGGCCGCACAACACCGGTCCTCTCATGACCTACATGGCC
TCATGCGGGTCGACGACCTGTGACAAGTTCGATGAGAGGGACGCACAATGGTTCAAGATTGATGAGGCGGGCAAGAAGC
CGAATGATGACTCGACCTGGATTCAGGCGGATATCATGAAAGGCGACTCGTACTCCCTCAAGCTTCCCGAAGACCTTGCCC
CCGGAGACTACCTCATTCGTCACGAGATCATCGCTCTCCACCTTGCCGTGAGCAAAGGTGGCGCCGAGTTCTACCCCTCCT
GCACCCAAGTGCGTGTCAGCGGCGATGGCAGCGGCAAGCCTCAGGACACGGTCACGTTCCCCGGTGCTTACAGCGACAG
CGACCCTGGCATCTTCGATCCTGAGGTCTTTGACTCGGGTGCCAAATACACCTTCCCTGGCCCTGCCATCTCCAACCTCGCT
GCCTCCGACGGCGCGATCGGCGCACCTGCTAGCAGTGCTACGTTCCCGTCGGCAACCGCTACCTCGGCCTCTGGCAAGGG
CAATGGATCGAGCACACGTTCCTCACAAGTCGCTACTCCATCTGCATCTGCTACCGCGTCGTCCAGTTCCAAGTCGACATGT
CACCTCACGAAGAGCGCCAGCTCGAGCAAGACTGTCTACCCTCGGCATTATAGCCGCATCATGGCCCGCATCCTGCACGG
GAGCCAT

GH61_6 (contig01118)

ATGCTCGCCCCCGTCCTTGTCCTCGCAGCCCTCGTGGCCCAGGCCGTCGCGCATTACACCTTCCCTGCCCTCATCGTTGGCG
GTAcCCGCACTGGGGACTGGGTGAAcGTCCGCATGACCGACAATCACTACTCCAACGGACCCGTCACCGACGTGACGAGT
GAAGCATTTCGCTGCTACACCGCTGAGACGCACGCGACCGCACAGACGATAAACGTCGCTGCGGGCTCGGAGATCAGCAT
TGCCGCCAACGGCAACATCTCTCACCCCGGTGTCGCCAACGTTTACCTCGCCAAGGCACCCACCAGCACCGACGTCACGTac
TGGGATGGCTCGGGCAACGTGTGGTTCAAGGTCTACCAGATCTCCGCGGTGACGGACGGCGGGCAGAGTATCACCTTCCC
CTCTCAGGGTCTCCCCGCAATCACCTTCACACTTCCGAAAAACCTCCCCAGCGGCCAGTACCTCGTCCGTGTTGAACACATC
GCCCTGCATTCCGCCAGTCAGCCTGGTGGTGCACAGTTCTACATCTCTTGCGGTCAGATCAACGTTACTGGGGGCgGAAGC
GGCACGCCTGGCCCCTTGGTCGCCATcCCTGGCGTATATTCTGGCAACGAGCCCGGcATCGTCATTAACATCTACTACCCCG
TcCCCGcCACCTACACCCAGCCCGGcCCTGCTGTCTGGAGCGGGTGA

7.4. Sequences of cDNA targets optimized for *P. pastoris*

Following features are highlighted in the shown sequences:

- *Xho*I restriction site: CTCGAG
- his-tag: *catcaccatcaccatcac*
- *Not*I restriction site: GCGGCCGC

IDN6 optimized for *P. pastoris* (449 bp)

CTCGAGAAGAGAGAGGCCGAAGCTATTCCTTTCAATCCAGGACTGCTTCCATACGGTTGGTTTACTGAGTACGAATGTGCT
ACTGATACTCCATCTCGTGTCTTGGTCAACATCCAGTCTAACGACATGGGTTCCAACAACACCGCTTTCTTGTGCACTACCA
AGTGTGCCTTCAACAATTTCATCTACGCCGGTGTTGAAGCTACTAACTGGTGCTACTGTGGAACCGGTTTGAATGGTTCTTC
CCTTCCACAAACTGCTTTGATTACCGATTGTTCTGAACCTTGTACTGGTGATCCAAACAGATCCTGTGGTAATCATTGGCGT
ATCCAAATCTACCGTAACTACATCACTTCTACCGTTCCAAATCCATTGCCTACTAAGGCTGCCTCCATGGAATTGCTTGAATC
CCAAATCAAG*catcaccatcaccatcac*TAAGCGGCCGC

IDN9 optimized for *P. pastoris* (977 bp)

CTCGAGAAGAGAGAGGCCGAAGCTGGATCCATCGTCTGGGATGGATCCTTCAATCCTTTCAACACTGCTGCCGATTTTGAT
AAGTGGTCTTGGGCCAATCAAACTGGTTCTTACCAGTGGTACATTCATGGATCCCAGCCAACCTCTCATTACTTGACTCTTG
ATTCCGCTTACAAGAACCCTGCCCTTACTAACGAAGAGCACGGTTTGAAGTTGACCATTGATTCCACTGCTACTTGGAACTC
CAACATGGAAAGAGCTGAGTTGATCCCTCAAACCTCCCAGAACCTGGGTACTGGTAACTTGTTCTACCATTTCTCAATCTCC
CACCCAGATACCAATACTCCAGACTCTACCTTAGAGCACCAAATCTTGTTCTTTGAATCCCATTTCACTGAGTTGAAGTACG
GAGTTGGTCCAAACACCACTAATCTTGAATGGCATGTTGGTGGACAACCTCAATGGGGTGTGCCATTCGCAGCCGGAACTT
GGTTCAACTTCGCTTACGACATTGACTTTTCTGCTTCTACTGTTGGTTTGTGGGCTTCCACTGGATCTGCTCCTCTTACCAAG
GTTGCCCAGAACATCAAGGCCTCCACTTCAACCAATTCCGCTGACTTTCACGTTGGAGTTTTGAGAATCGTTAACACTCCAG
CTGCCGAAGATTGGTTCATCTCTGGTGTCTACATTGAACAAGGACCAATCACTACCACTATTGGTTCTGGATCTTCCGCTCC
TTCCCCTCCAGTTACTTCAGCTCCTGTTACTACCGTCCCACCTACTTCCGCTCCATCTACCTCCGCCCCTCCATCAACTACCTCT
TCCGCTTCAGGACCTCTTGCCTCCCAATGGGCTCAGTGTGGTGGAACTGGTTACACCGGACCAACTGCCTGTGCTTCTCCTT
ACACTTGTACCGCTGTTTCCCCACCTTACTACTATCAATGTCAG*catcaccatcaccatcac*TAAGCGGCCGC

IDN10 optimized for *P. pastoris* (1403 bp)

CTCGAGAAGAGAGAGGCCGAAGCTCAGGCCCCTGTTTGGGGACAATGTGGTGGACAGGGTTGGACTGGAGCTTCAACCT
GTGTCTCCGGTACTGTCTGCACCGTTCTTAACGCCTACTATTCACAATGCTTGCCTGGTACTGCTTCCTCTGCTCCAGCCCCT
CCATCTTCCACTACCGCAACTGTTCCTACTACCACTTCTAATCCAGCTCCTCCAGTCACCACTTCCAACCCTGTTCCAGGATCT
TGTTCTACCCCTTCCACTGTCTCTGGATTTTCCAACACCAAGTTGCCAAATCCATTCAAGTTCAACGATGGATCTGCTGTCAC
TACCCAAGACGATTGGGAGTGTCGTAGAGCCCAGATTCTTTCTTTGGTCCAAGGATACGAAGGAGGTGCTCTTCCAGGTA
AGCCATCTTCCTTAACTGCTCAATTGACCAAATCTGGAACTTCCGCCACCCTTGCTATCACTGCCTCTGATGGTGGAAAGTC
CATTACCTTTTCCCCAAAGATTACTTATCCTTCTGGTAATCCTCCAGCCGGTGGATGGCCTGTCATCATTGTCTACGAAGGTC
TTTCATTGCCAGTCCCATCCGGAGTTGCTCAACTTGTCTACCAAAACTCTGAAATGGGTGCCCAGGATTCAACTGCTTCTCG
TGGTCAAGGATTGTTCTTCAACTTGTATGGTTCCAATGCTACTGCCTCTGCTATGACTGCCTGGGCTTGGGGTGTCTCCAGA
ATCATTGACGTTCTTGAATCTACCCCTAACGCCCAATTGAATCCACAGAAAGTTGCTGTTACCGGTTGTTCCAGAAACGGAA
AGGGTGCCTTGGTCGCTGGTGCCCTTGATACTCGTATTGCTTTGACCCTTCCACAGGAATCTGGATCCGGTGGAGACGCCT
GTTGGAGACTTTCTAAGTTCGAGCAAGACTCTGGTTCCGTTGTCCAGACTGCTACCGAAATCGTTGGAGAGAACGTCTGGT
TTTCCACCAACTTCAATAACTACGTCAACCAGTTGTCCGTCTTACCTTACGATCACCATCTTTTGGCTGCCCTTGTTGCTCCAC
GTGCCATGATTTCCTTTGAAAACACCGATTTCGTTTGGTTGTCCCCACTTTCTTCCTTTGGATGTATGACTGCCGCTCACACT
GTCTGGGATGCCTTGGGTGTTGCTGACCACCATGGTTTTGAGCAAGTCGGAGGTCATGCTCACTGCGCCTGGCCATCCTCC
TTGAACAATCAGCTTTCTGCCTTCTTTAACCGTTTCTTGCTTGATCAGAATGTTTCCACTACCTTCTTTGAGTCCAACAAGATT
TTCAATGGAGTTACCTGGAATCAAGCCAACTGGATTGACTGGACTACCCCAAAGTTGACT*catcaccatcaccatcac*TAAGCGG
CCGC

Laminarinase (clone 30H12) optimized for *P. pastoris* (938 bp)

CTCGAGAAGAGAGAGGCCGAAGCTGTGACTTATCAGCAATCCGACTCACACCAGGGAAACGGTTTCTTGAAGTCCTTCAA
GCATATGGCTATCTCTGATCCAACTCATGGTAGAGTTAACTACGTTGATCAAGACACTGCTTTGGCCCAGAACCTTACCTTC
TACTCCGGCGATCACTTTGTCATTAGAGCTGACCACAAAACTAAGTTGTCCTCTTCCGGTCCAGGTCGTAACTCTGTTAGAT
TGCAGTCCAACAAGAAGTATACTACCCACGTTACCGTTTGGAACATCAGACATATGCCAGTCGGTTGTGGTACTTGGCCTG
CTGTTTGGGAGGTTGGAGACAACTGGCCAAATGAAGGTGAGATTGATATCTTGGAAGGAGTTAACGACCAATCCCCAAAC
CAAGCTACTCTTCATACCAATTCTGGTTGTACCATGTCTGCCTCAAGATCCCAAACTGGATCCTCTACCGGTAACAATTGCG
ACATTGCTGCCACTAACAATGCTGGTTGTGGAGTCGAAGCCCCTTCTTCCAACTCTTATGGACCAGCTTTCAATTCTGCTGG
TGGAGGTTGGTACGCCATGGAGCGTACCAACAACTTCATCAAGGTTTGGTTTTGGTCTCGTAATGCTGGTAACGTCCCATC
CGATGTTAAGAATGGTGCCACTTCTGTCAACACCGACAAATGGGGAACCCCATTTGCTTACTTTCCATCTCAGTCCTGCTCT
ATCTCCGATCATTTCGGTGAGCACAACTTGATCATCAACCTTACCTTGTGTGGAGACTGGGCTGGTGCTGTTTTCGGTTCTG
ATGGATGCCCAGGTGACTGTGTCTCCTACGTCAACAACAATCCTTCTGCTTTCACTAACGCCTACTTTGACATCCAGTGGTT
GAAGATTTACCAG*catcaccatcaccatcac*TAAGCGGCCGC

7.5. Sequences of cDNA targets optimized for *S. pombe*

Following features are highlighted in the shown sequences:

- *Hind*III restriction site: AAGCTT
- Kozak sequence: **CGAAACG**
- Cpy leader sequence: **ATGCTTATGAAGCAGACTTTCTTGTATTTTCTTTTGACCTGTGTCGTTTCTGCTCAATTCAACGGATACGTTCCACCTGAA**)
- his-tag: *catcaccatcaccatcac*
- *Spe*I restriction site: ACTAGT

IDN6 optimized for *S. pombe* (517 bp)

AAGCTT**CGAAACG**<u>**ATGCTTATGAAGCAGACTTTCTTGTATTTTCTTTTGACCTGTGTCGTTTCTGCTCAATTCAACGGATACGTTCCACCTGAA**</u>ATTCCTTTCAATCCCGGCCTCTTGCCTTACGGTTGGTTTACTGAGTATGAATGTGCTACTGACACCCCTTCCCGTGTCCTTGTTAACATTCAGTCTAATGATATGGGTTCCAACAACACCGCTTTCCTTTGCACTACCAAGTGTGCCTTCAACAACTTCATCTACGCTGGTGTTGAGGCTACTAACTGGTGCTATTGTGGCACCGGTTTGAATGGTTCTTCCTTGCCTCAGACTGCCCTTATCACCGACTGTTCTGAACCTTGCACTGGCGATCCCAATCGTTCCTGTGGTAACCACTGGCGCATTCAAATCTACCGTAACTACATCACTTCTACTGTTCCTAATCCCCTTCCTACTAAGGCTGCCTCCATGGAATTGCTTGAATCTCAGATCAAG*catcaccatcaccatcac*TAA<u>ACTAGT</u>

IDN9 optimized for *S. pombe* (1045 bp)

AAGCTT**CGAAACG**<u>**ATGCTTATGAAGCAGACTTTCTTGTATTTTCTTTTGACCTGTGTCGTTTCTGCTCAATTCAACGGATACGTTCCACCTGAA**</u>GGTTCTATCGTTTGGGATGGCTCCTTCAATCCTTTCAACACCGCTGCCGACTTTGATAAGTGGTCCTGGGCTAATCAAACCGGTTCTTACCAGTGGTACATTCATGGTTCCCAGCCTACCTCTCACTATCTTACTCTTGATTCCGCTTACAAGAACCCTGCCCTTACTAATGAGGAACACGGTCTTAAGTTGACTATTGATTCTACCGCTACTTGGAACTCCAACATGGAACGTGCTGAGTTGATTCCTCAAACCTCCCAGAACTTGGGTACTGGCAACCTTTTCTATCATTTCTCTATCTCCCACCCCGATACTAACACCCCCGATTCTACCTTAGAGCATCAAATCTTGTTCTTTGAATCCCATTTCACTGAGCTTAAGTATGGTGTCGGCCCTAACACCACTAATCTTGAATGGCATGTTGGTGGACAACCTCAATGGGGTGTTCCTTTTGCTGCCGGCACTTGGTTCAACTTTGCTTACGACATTGATTTCTCTGCCTCCACTGTTGGTCTTTGGGCTTCTACTGGTTCTGCCCCCTTGACCAAAGTTGCCCAGAACATCAAGGCCTCCACTTCTACCAATTCCGCTGACTTTCACGTTGGTGTCTTGCGCATTGTTAACACCCCAGCTGCCGAAGATTGGTTCATCTCCGGCGTCTACATTGAACAAGGTCCCATCACTACCACTATTGGCTCTGGTTCTTCCGCTCCTTCTCCACCTGTTACTTCCGCCCCAGTCACTACCGTTCCCCCAACTTCTGCTCCATCCACTTCTGCCCCCCCATCCACCACTTCTTCCGCTTCCGGTCCCCTTGCCTCTCAATGGGCTCAGTGTGGTGGCACTGGTTACACCGGTCCCACTGCCTGCGCTTCTCCCTATACCTGTACTGCCGTTTCCCCACCCTACTACTATCAATGCCAG*catcaccatcaccatcac*TAA<u>ACTAGT</u>

Laminarinase optimized for S. pombe (1006 bp)

AAGCTT**CGAAACG**<u>**ATGCTTATGAAGCAGACTTTCTTGTATTTTCTTTTGACCTGTGTCGTTTCTGCTCAATTCAACGGATACGTTCCACCTGAA**</u>GTTACTTATCAGCAATCCGATTCTCATCAAGGTAACGGCTTTCTTAAGTCCTTCAAACACATGGCCATTTCTGACCCTACTCACGGTCGTGTTAACTATGTCGATCAAGACACTGCCTTGGCTCAGAATCTTACCTTCTACTCCGGTGATCACTTTGTTATCCGTGCCGACCATAAGACTAAACTTTCTTCTTCAGGTCCTGGCCGTAACTCCGTTCGTCTTCAATCTAACAAGAAGTACACCACTCACGTCACTGTTTGGAACATTCGTCATATGCCTGTTGGTTGTGGCACATGGCCTGCTGTTTGGGAAGTCGGTGACAACTGGCCCAATGAGGGCGAAATTGATATCCTTGAGGGTGTTAACGACCAATCCCCAAATCAAGCTACTCTTCATACTAACTCTGGTTGTACCATGTCAGCCTCTCGTTCCCAAACTGGTTCTTCCACTGGTAACAATTGCGATATTGCTGCCACTAACAATGCTGGCTGTGGTGTCGAAGCCCCCTCTTCAAACTCTTACGGCCCTGCTTTCAATTCTGCTGGTGGCGGTTGGTACGCCATGGAACGTACTAACAACTTCATCAAGGTCTGGTTTTGGTCTCGCAACGCTGGTAATGTTCCTTCCGATGTCAAGAACGGCGCCACTTCCGTCAATACTGACAAATGGGGTACCCCTTTTGCTTACTTCCCCTCTCAATCCTGCTCTATCTCCGATCACTTTGGTGAACACAATCTTATCATTAACTTGACTCTTTGTGGTGATTGGGCTGGTGCTGTCTTTGGCTCTGATGGTTGTCCAGGTGATTGCGTTTCTTATGTTAACAACAA*catcaccatcaccatcac*TAA<u>ACTAGT</u>

7.6. Plasmids

Table 25: Used and produced plasmids in this work. The first two plasmids pGAZ2 and pGAK2 were part of Andryushkova's work (1). The others are part of this work.

Plasmid	Insert	Description of the plasmid	Description of the insert
pGAZ2 (CC # 6027)	-	Shuttle vector, suitable for episomal expression in *S. pombe*	-
pGAK2 (CC# 6028)	-		-
pGAZ2	eGFP		Reporter gene
pGAK2	eGFP		
pJET	Cass_*leu1*-zeo-	From CloneJET™ PCR Cloning Kit, Thermo Scientific	PCR cassettes cloned into pJET vector, linear insert is suitable for targeted integration into *S. pombe* genome
pJET	Cass_*leu2*-zeo-		
pJET	Cass_his3-zeo-		
pJET	Cass_*leu1*-kan-		
pJET	Cass_*leu2*-kan-		
pJET	Cass_his3-kan-		
pGAZ2 *leu1* (CC# ST876)	-	Shuttle vector, suitable for integration in *S. pombe* as well as episomal expression	-
pGAZ2 *leu2* (CC# ST877)	-		-
pGAK2 *leu1* (CC# ST878)	-		-
pGAK2 *leu2* (CC# ST879)	-		-
pGAZ2 *leu1*	eGFP		Reporter gene
pGAZ2 *leu2*	eGFP		
pGAK2 *leu1*	eGFP		
pGAK2 *leu2*	eGFP		
pGAZ2_*leu1*	IDN5		cDNA from *F.* fomentarius
pGAZ2_*leu1*	IDN6		
pGAZ2_*leu1*	IDN8		
pGAZ2_*leu1*	IDN9		
pGAZ2_*leu1*	IDN10		
pGAZ2_*leu1*	IDN13		
pGAZ2_*leu1*	IDN16		
pGAZ2_*leu1*	IDN14		
pGAZ2_*leu1*	IDN40 (GH6		
pGAZ2_*leu1*	GH61_1		
pGAZ2_*leu1*	GH61_2		
pGAZ2_*leu1*	GH61_3		
pGAZ2_*leu1*	GH61_4		
pGAZ2_*leu1*	IDN6_Sp	Shuttle vector, suitable for integration in *P. pastoris*	
pGAZ2_*leu1*	IDN9_Sp		
pGAZ2_*leu1*	La_Sp		
pPpB1_GAP_α_S	IDN6_Pp		
pPpB1_GAP_α_S	IDN9_Pp		
pPpB1_GAP_α_S	IDN10_Pp		
pPpB1_GAP_α_S	La_Pp		
pPpT4_GAP_S	IDN6_Pp		
pPpT4_GAP_S	IDN9_Pp		
pPpT4_GAP_S	IDN10_Pp		
pPpT4_GAP_S	La_Pp		
pPpB1_AOX_α_S	IDN6_optPp		
pPpB1_AOX_α_S	IDN9_optPp		
pPpB1_AOX_α_S	IDN10_optPp		
pPpB1_AOX_α_S	La_Pp		

7.7. Vector charts

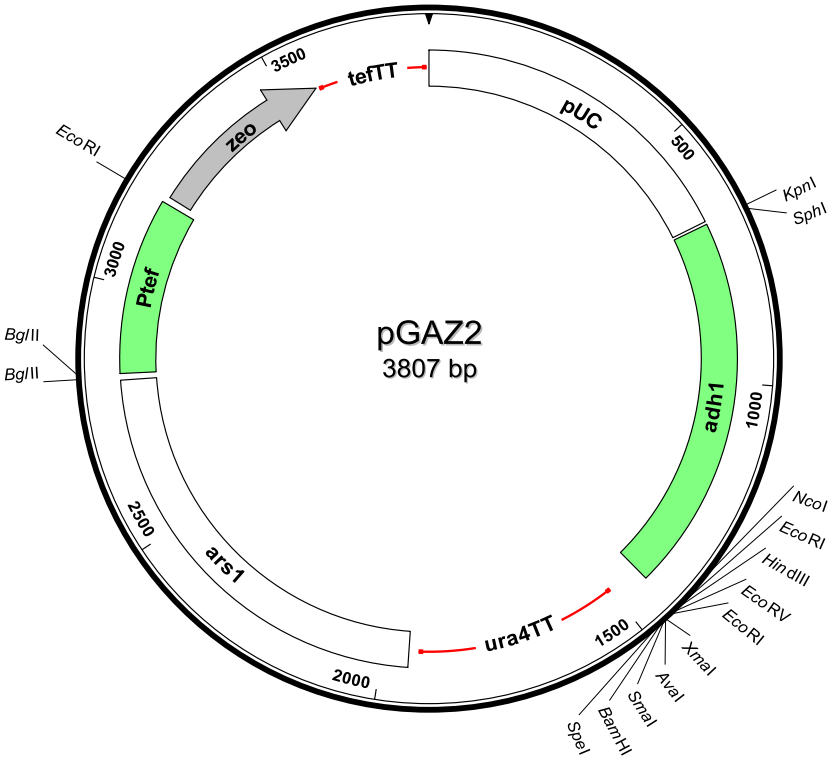


Figure 46:pGAZ2/pGAK2 vector. The features are as followed: The prokaryotic ori pUC (673 bp), the eukaryotic ori *ars1* (868 bp), the *adh1* promoter (P(*adh1*), 720 bp), the gene of Interest (GoI), the *ura4* terminator (T(*ura4*), 425 bp), and the Zeocin™ (zeo, 375 bp) or kanamycinie (810 bp) resistance gene under the control of the *tef* promoter (352 bp) and the *tef* terminator (229 bp).

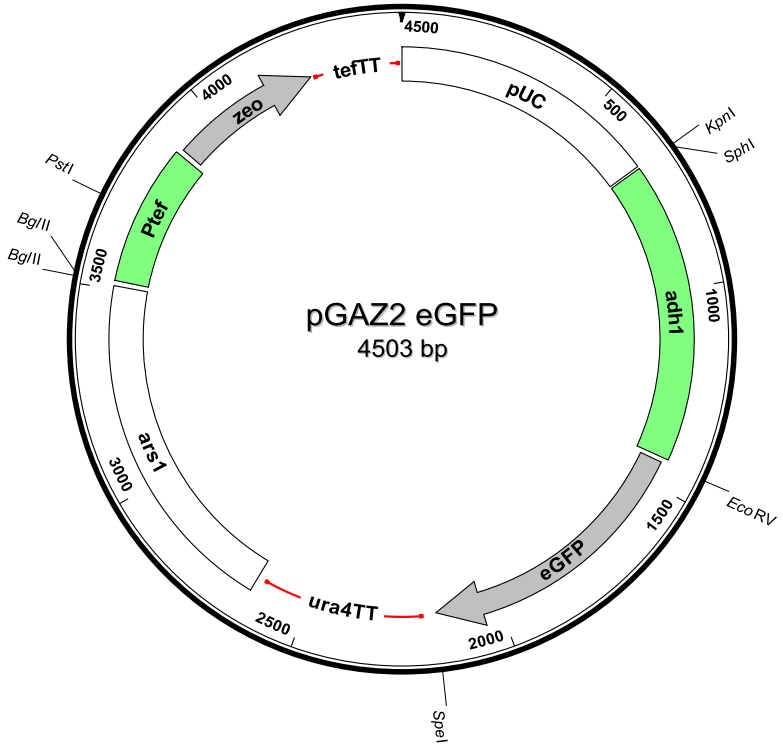


Figure 47: pGAZ2/pGAK2-eGFP vector. The features are as followed: The prokaryotic ori pUC (673 bp), the eukaryotic ori *ars1* (868 bp), the *adh1* promoter (P(*adh1*), 720 bp), eGFP (720 bp), the *ura4* terminator (T(*ura4*), 425 bp), and the Zeocin™ (zeo, 375 bp) or kanamycinie (810 bp) resistance gene under the control of the *tef* promoter (352 bp) and the *tef* terminator (229 bp).

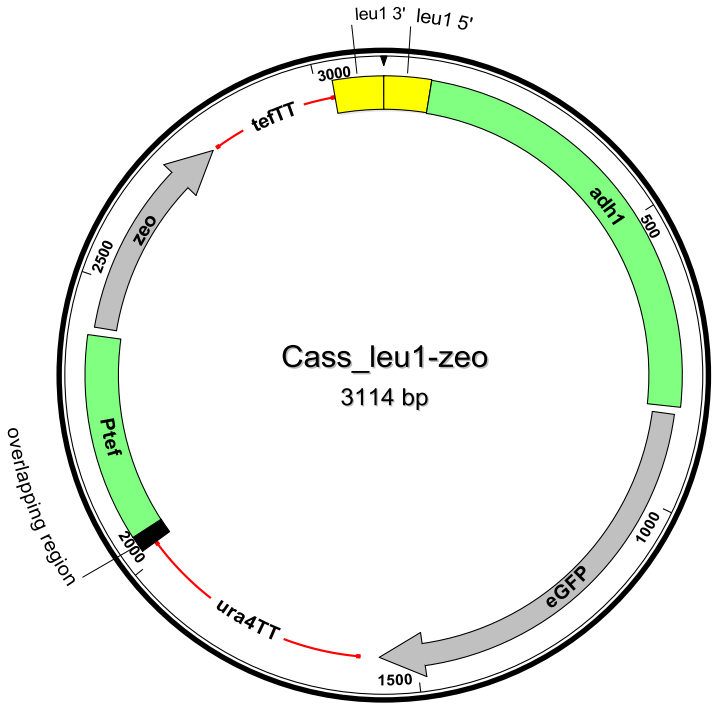


Figure 48: PCR cassette. The features are as followed: yellow sequences are integration markers (80 bp, *leu1, leu1* or *his3*), the *adh1* promoter (P(*adh1*), 720 bp), eGFP (720 bp), the *ura4* terminator (T(*ura4*), 425 bp), and the Zeocin™ (zeo, 375 bp) or kanamycinie (810 bp) resistance gene under the control of the *tef* promoter (352 bp) and the *tef* terminator (229 bp).

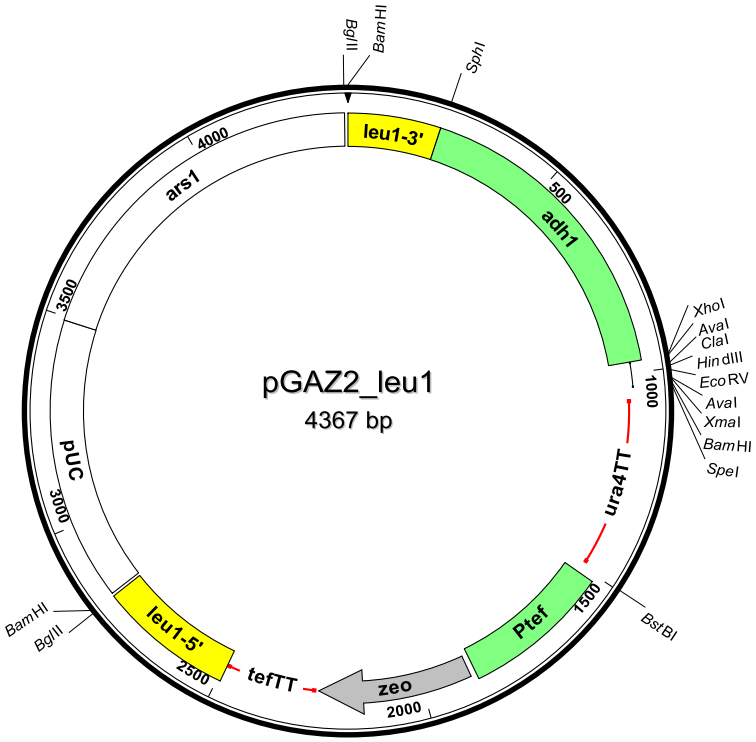


Figure 49: pGAZ2/pGAK2_*leu1*. The features are as followed: The prokaryotic ori pUC (673 bp), the eukaryotic ori *ars1* (868 bp), the two integration markers *leu1*-3'(400 bp) and *leu1*-5'(300 bp), the *adh1* promoter (P(*adh1*), 720 bp), the gene of Interest (GoI), the *ura4* terminator (T(*ura4*), 425 bp), and the Zeocin™ (zeo, 375 bp) or kanamycinie (810 bp) resistance gene under the control of the *tef* promoter (352 bp) and the *tef* terminator (229 bp).

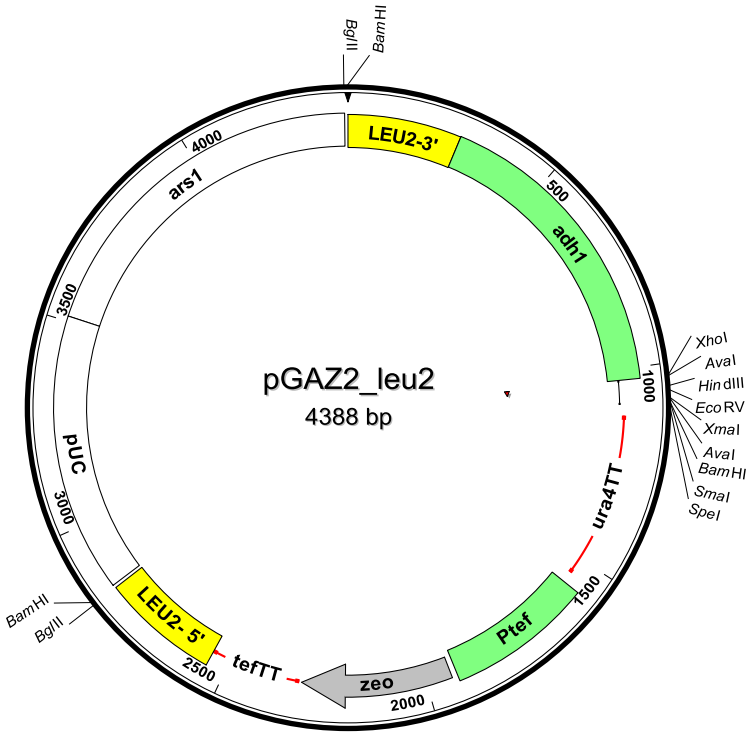


Figure 50: pGAZ2/pGAK2_*leu2* The features are as followed: The prokaryotic ori pUC (673 bp), the eukaryotic ori *ars1* (868 bp), the two integration markers *leu2*-3'(365) and *leu2*-5'(356 bp), the *adh1* promoter (P(*adh1*), 720 bp), the gene of Interest (GoI), the *ura4* terminator (T(*ura4*), 425 bp), and the Zeocin™ (zeo, 375 bp) or kanamycinie (810 bp) resistance gene under the control of the *tef* promoter (352 bp) and the *tef* terminator (229 bp).