Characterization of novel promoters in *Pichia pastoris*

Diplomarbeit

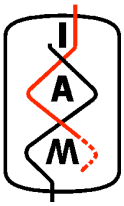
eingereicht von

Astrid Mecklenbräuker

Betreuer:

Ao. Univ. Prof. DI Dr. Diethard Mattanovich

Dr. Brigitte Gasser



Institut für Angewandte Mikrobiologie
Department für Biotechnologie
Universität für Bodenkultur Wien

Wien, im Januar 2011

TABLE OF CONTENTS

ABSTRACT

The yeast *Pichia pastoris* is commonly used as host system for heterologous protein production. However, compared to other yeast species like *Saccharomyces cerevisiae* the number of useable promoter sequences is rather small and also limited to promoters having a very strong activity. For some applications, such as metabolic engineering, the co-expression of secretion helpers, or the separate expression of multimeric proteins, e.g. the light (LC) and heavy chain (HC) of antibodies, it might be of interest to use different promoters with different transcription levels. It is therefore useful to have a selection of different promoter sequences suitable for recombinant expression of a heterologous or homologous gene, varying from strong promoter activity to weak or reduced promoter activity.

In the present study, clones expressing either an intracellular reporter (enhanced green fluorescent protein, eGFP) or a secreted heterologous protein (human serum albumin, HSA) were cultivated. Promoter activities were indirectly determined by measurement of the amount of gene product expressed from the promoter. Transcription levels of selected constructs were analyzed using quantitative real-time PCR (qRT-PCR), but also gene copy numbers were determined, as neglecting this factor can lead to false interpretations of experimental results. Often higher gene copy numbers lead to higher expression levels independently of promoter activity.

Especially P_{PET9} encoding an ADP/ATP carrier of the mitochondrial membrane resulted in high relative expression levels of intracellular eGFP, whereas growth rate dependent P_{TEF1} gave stronger results with HSA as reporter. Fed batch cultivation of selected constructs expressing HSA strengthened the correlation of promoter activity to specific growth rates also for the P_{THI1}. Generally, a higher transcript level resulted in a higher amount of expressed reporter protein, with P_{GAP} as the promoter with the highest transcriptional strength with both reporters. It is important to consider the reporter gene, as different genes have different expression levels under the control of the same promoter.
In conclusion, this allows the regulation of the expression level of a protein of interest by selection of a suitable promoter sequence according to the experimental situation.

ZUSAMMENFASSUNG

Pichia pastoris wird häufig als Wirtsorganismus für die Herstellung heterologer Proteine verwendet. Jedoch, verglichen mit anderen Hefen, wie *Saccharomyces cerevisiae,* ist die Anzahl verwendbarer Promotorsequenzen relativ klein und auch limitiert auf Promotoren mit sehr starker Aktivität. Für manche Anwendungen, wie Metabolic Engineering, der Co-Expression von Sekretions-Helfern, oder der getrennten Expression von multimeren Proteinen, wie z.B. leichte (LC) und schwere Kette (HC) von Antikörpern, kann es sinnvoll sein verschiedene Promotoren mit unterschiedlichen Transkriptionsstärken zu verwenden. Es ist daher nützlich, eine Auswahl verschiedener Promotorsequenzen, geeignet für die rekombinante Exprimierung von heterologen oder homologen Genen zu haben, variierend von starker Promotoraktivität zu schwacher oder reduzierter Promotoraktivität.

In der vorliegenden Studie wurden Klone kultiviert, die entweder ein intrazelluläres Reporterprotein (verbessertes grünfluoreszierendes Protein, eGFP) oder ein sekretiertes Fremdprotein (menschliches Serumalbumin, HSA) exprimieren. Promotoraktivitäten wurden indirekt durch Messung der Genproduktmenge bestimmt. Nicht nur die Transkriptionsstärke der ausgewählten Konstrukte, sondern auch die Genkopienzahl wurde mittels quantitativer real-time PCR (qRT-PCR) bestimmt, da eine Vernachlässigung dieses Faktors zu Fehlinterpretationen der Ergebnisse führen kann. Oft führen höhere Genkopienzahlen, unabhängig von der Promotoraktivität, zu höheren Expressionsraten.

Besonders P_{PET9}, einen ADP/ATP Carrier der mitochondrialen Membran kodierend, führte zu hohen relativen Expressionsraten des intrazellulären eGFP, stattdessen gab der wachstumsabhängige P_{TEF1} höhere Ergebnisse mit HSA als Reporter. Fed batch Kultivierung von ausgewählten Konstrukten bestärkte die Korrelation von Promotoraktivität zu spezifischer Wachstumsrate auch für P_{THI1}. Allgemein resultierte eine höhere Transkriptmenge in größeren Mengen an exprimiertem Reporterprotein, mit P_{GAP} als den Promotor mit der höchsten Transkriptmenge für beide Reporter. Es ist außerdem wichtig das Reporterprotein zu berücksichtigen, da verschiedene Gene unterschiedliche Expressionsraten unter der Kontrolle des gleichen Promotors haben.
Abschließend erlaubt dies, die Expression eines gewünschten Proteins, durch die Auswahl geeigneter Promotorsequenzen abhängig von der Situation, zu regeln.

AIM OF THE STUDY

As the number of useable promoter sequences is rather small for the yeast *Pichia pastoris*, the aim of this study was the characterization of novel potential regulatory sequences. *P. pastoris* clones expressing either an intracellular reporter (enhanced green fluorescent protein, eGFP) or a secreted heterologous protein (human serum albumin, HSA) were cultivated under the control of different promoter sequences. Promoter activities should de determined indirectly by measurement of the gene product as well as directly by analyzing transcription levels applying quantitative real-time PCR (qRT-PCR).

1. INTRODUCTION

For *Pichia pastoris*, one of the most commonly used yeast expression system for recombinant protein production, only a very limited number of regulatory elements is available at present. To close this gap potential regulatory sequences were identified and tested (Stadlmayr et al. 2010a). A very broad range of relative promoter activity (measured indirectly by reporter protein production) compared to P_{GAP} could be seen. Generally promoters with relative expression levels spanning the range from approx. 0 up to the strong *GAP* promoter were found. These novel promoter sequences for *P. pastoris* derive from the heterologous microarray hybridization of *P. pastoris* cDNA to *Saccharomyces cerevisiae* cDNA microarray (Sauer et al. 2004).

1.1. The *Pichia pastoris* expression system

The methylotrophic yeast *P. pastoris* is being developed as a widely used host organism for recombinant protein production (Macauley-Patrick et al. 2005).
Not only rapid growth, reaching high cell densities in cultures and a high secretory capacity, but also proper folding including the formation of disulfide bonds and a variety of post-translational modifications are among the advantages of the host. Furthermore, genetic manipulation of *P. pastoris* can be easily achieved with commercially available kits (Cereghino et al. 2000), making it easy to operate with this expression system.
Genetic engineering led to strains capable of producing glycoproteins with humanized N-glycans (reviewed by Hamilton et al. 2007).
Meanwhile also the sequence of *P. pastoris* is published and publically available (De Schutter et al. 2009; Mattanovich et al. 2009).

1.1.1. Optimization of production strains

In the last years more and more effort has been invested into the optimization of production strains. The main goal of the different strategies is improving the productivity.
Genetic and physiological factors determine the productivity of a recombinant system. Improving the expression level of a heterologous protein, means to systematically observe the synthesis pathway of the protein, from transcription over translation (Sreekrishna et al. 1997) to proper folding and excretion. Among many other factors potential bottlenecks are the efficient transcription by using strong promoters and the gene copy number of the inserted gene of interest. Not only on DNA level rate limiting factors occur, but

also on RNA level, further on in the processing and folding of the protein and finally in the secretion out of the cell.
Co-transformation of helper factors for secretion (Gasser et al. 2007), chaperones and foldases (Gasser et al. 2008), and other cellular host proteins (Stadlmayr et al. 2010b) have led to improved secretion.
Numerous studies on copy number effects (Sreekrishna et al. 1988; Vassileva et al. 2001; Chen et al. 2006) led to the development of multicopy strategies. Targeted integration into the rDNA locus and post-transformational vector amplification by repeated selection on increased antibiotic concentrations as presented by Marx (2009) led to a higher number of multicopy integrants, further leading to a higher productivity of the desired protein.

1.2. Promoter elements in *P. pastoris*

A promoter is a region of DNA that regulates the transcription of a particular gene. Gene expression is the process by which information from a gene is used in the synthesis of a functional gene product, often a protein. Gene expression levels determine the number of copies of mRNA of a particular gene in a cell or tissue.

Even though *P. pastoris* is one of the most commonly used yeast expression systems for recombinant protein production, only a very limited number of regulatory elements is available at present.
Most of the promoters used for expression in *P. pastoris* have been derived from genes that code for enzymes involved in the methanol metabolism, which are usually present at high concentrations in the cell (Tschopp et al. 1987; Shen et al. 1998).
In comparison to model yeast species such as *S. cerevisiae* the number of useable promoter sequences is rather small in *P. pastoris* and also limited to promoters with a strong activity. However, for some applications, such as coexpression of chaperones or other secretion helper factors, it might be of more purpose to use a promoter with a lower transcriptional activity.

1.2.1. *AOX1* – promoter

In *P. pastoris*, heterologous genes are usually expressed under control of the promoter of the alcohol oxidase I gene (*AOX1*), the P_{AOX1}, which encodes the first enzyme in the methanol utilization pathway and is tightly regulated (Tschopp et al. 1987; Cregg et al. 1989).
The P_{AOX1} expression is fully repressed by carbon sources like glucose but is maximally induced during growth of cells on methanol. A typical advantage of this regulation is that

strains expressing heterologous genes, whose products are toxic for the cell, can be maintained by growing this organism under repressing conditions prior to induction of expression.

In some cases, however, the use of methanol to induce the production of foreign proteins is inconvenient or inappropriate, as induction can be too strong, stressing the production strain and therefore leads to no desirable results. For example, although necessary for the full induction of P_{AOX1}, methanol can be a potential fire-hazard. Also the use of methanol to induce gene expression may not be appropriate for the production of food products since methanol is toxic. Additionally high concentrations of methanol cause stress to the cells, cell viability is decreased, cell lysis and proteolytic degradation occur (Hohenblum et al. 2004).

Therefore, promoters that are not induced by methanol are attractive for expression of certain genes.

Efforts were also made to positively regulate the expression of *AOX1* by inserting positive cis-elements in the *AOX1* promoter (Xuan et al. 2009).

1.2.2. *GAP* – promoter

The constitutive glyceraldehyde-3-phosphate dehydrogenase (*GAP*) promoter also has a high level of expression, as already described in 1997 (Waterham et al.).

Cells under the control of the P_{GAP} show strong and constitutive expression on glucose or glycerol and accumulate higher levels of foreign proteins during shorter production times (Zhang et al. 2009).

1.2.3. Alternative promoters

To circumvent such problems as handling with a hazardous substrate or having too high levels of expression, which can result in overwhelming the post-translational machinery in the cell as well as the secretory capacity of the host cell by a too strong induction on methanol (Gasser et al. 2006), other promising promoters are being tested.

An alternative promoter with a high level of expression is the P_{FLD1} (formaldehyde dehydrogenase promoter, Shen et al. 1998).

Other constitutive promoters include P_{TEF1} (translation elongation factor 1-α, Ahn et al. 2007) showing growth-related expression behaviour and the P_{PGK1}, promoter for the glycolytic enzyme 3-phosphoglycerate kinase (de Almeida et al. 2005).

Also moderately expressing promoters have been developed such as P_{PEX8}, the promoter of the peroxisomal matrix protein Pex8 (Liu et al. 1995) and P_{YPT1}, a GTPase involved in secretion (Sears et al. 1998).

Furthermore the use of the ethanol inducible P_{ICL1} (isocitrate lyase, Menendez et al. 2003) has been described, as well as a phosphate responsive promoter of the putative sodium (Na^+)-coupled phosphate symporter Pho89, which is highly active in phosphate limited conditions (Ahn et al. 2009).

Also recently, a focus on the development of new promoters, by creating a promoter library by deletion and duplication of putative transcription factor-binding sites within the P_{AOX1} sequence was set by Hartner and his coworkers (2008).

But only a limited number of these enumerated promoters have been applied for routine production processes so far.

1.2.4. Novel approaches

To identify novel promoter sequences for use in *P. pastoris* a novel approach was conducted by Stadlmayr and colleagues (2010a). Not only data derived from the heterologous microarray hybridization of *P. pastoris* cDNA to *S. cerevisiae* cDNA (Sauer et al. 2004) were evaluated in a specific manner, but additionally a selection of promoters based on literature data from other yeast species were selected and considered for further analysis.

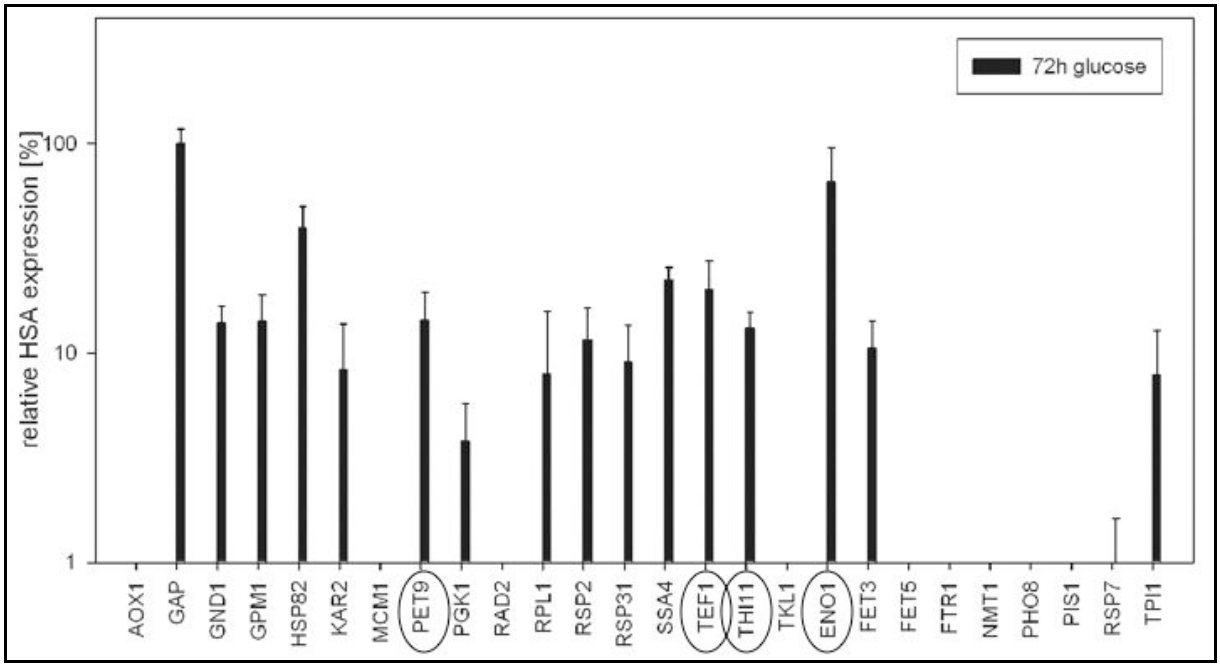


Figure 1: Promoter dependent relative expression level of human serum albumin (HSA) (figure from Stadlmayr et al. 2010a) Each bar represents the mean value (+/- standard error of the mean) of 10 individual single clones per promoter construct grown in shake flask culture. The expression of each promoter was normalized using the expression of HSA under the control of P_{GAP} as reference. HSA accumulation in the supernatant was quantified by ELISA after 72h cultivation on complex glucose medium (YPD).

Based on the promising results of the comparative analysis (Fig. 1) three novel promoters were chosen for further studies. In the present work the promoter activities of *PET9* – a major

ADP/ATP carrier in the mitochondrial inner membrane, *TEF1* – the translation elongation factor EF-1α, and the promoter of *ENO1* – the enolase I in *P. pastoris*, should be characterized. Additionally, the promoter of *THI11* – the thiamine biosynthesis gene, which exhibited expression in thiamine deficient media, was considered for ongoing experiments.

1.3. Gene copy number effects

One of the early bottlenecks affecting heterologous protein productivity is the copy number of the expression cassette introduced into the host organism. Not surprisingly, the isolation of multicopy integrants has resulted in higher yields (Romanos et al. 1998).

To achieve an increase in product titers the systematic increase in copy number is performed in many studies. An increased production of heterologous proteins (glucoamylase from yeast and mammalian interleukin-1β) was observed by Morlino and coworkers (1999) by inducible amplification of the gene copy number (GCN) in the yeast *Kluveromyces lactis*. In 2001 a direct correlation between GCN and the produced amounts of HBsAg (Hepatitis B surface antigen) in *P. pastoris* was described (Vassileva et al.). As could be seen in the work of Marx (2009) the protein production is tightly correlated with gene copy number if producing an intracellular protein such as hSOD (human superoxide dismutase), while using the secretory pathway with HSA (human serum albumin) a high clonal variability can be seen, making a direct correlation of secretion to gene copy number valid only for low gene copy numbers. This leads back to bottlenecks in association with secretion out of the cell (Rossini et al. 1993). Also Hohenblum and colleagues described that the increase of GCN correlates only to a certain limit with higher yields of the human protein trypsinogen (2004). Other limitations of the entire expression pathway counteract the positive effect of higher GCN.

Furthermore GCN can only have a positive effect on the product yield if the gene can be transcribed at efficient levels (Reisinger et al. 2008).

Neglecting gene copy numbers can easily lead to false interpretations of experimental results from promoter studies and co-expression of helper factors as demonstrated in an example by Abad very recently (2010).

1.4. Quantitative real-time PCR (qRT-PCR)

Polymerase Chain Reaction (PCR) has secured its place in molecular biology history. Realtime PCR or qRT-PCR is an emerging technology following the same principle as PCR, but combining it with high sensitivity and specifity, good reproducibility and a wide dynamic quantification range (Heid et al. 1996).

Real-time PCR gives the possibility to detect and measure the PCR products as they are accumulating because they are bound by a fluorescence dye.
An increase in double stranded DNA product during PCR therefore leads to an increase in fluorescence intensity and is measured in the exponential phase of the amplification, when it is most efficient and least affected by reaction-limiting conditions (Ginzinger 2002). The increase in fluorescence is directly proportional to the increase in the amplified product during the PCR. A fluorescence signal threshold (cycle threshold, C_t) is determined at that point, at which a comparison of all samples becomes possible. The fewer cycles are necessary to exceed the defined C_t the greater is the number of target sequence in the sample (Higuchi et al. 1993). Theoretically, the amount of DNA doubles every cycle during the exponential phase, but this can be affected by the efficiency of the used primers.

1.4.1. Detection chemistries

The main basic detection systems commonly used in real-time PCR are well described in the work of Giulietti (2001). All of these methods utilize fluorescent dyes.

- Fluorescent dyes

Intercalating fluorescent dyes, like SYBR green, are the simplest and cheapest way to monitor a real-time PCR. Intercalating dyes fluoresce only when bound to double-stranded DNA. During the PCR the target sequence is amplified and the fluorescence is increasing with every cycle. As the increase of the intensity of fluorescence is proportional to the amount of built PCR-product the amplification can be observed in real time. A major disadvantage of this system is the lack of specifity as it binds to any double-stranded DNA present in the mix, such as primer-dimers or inappropriate PCR products.

- Fluorescent probes

The probe is designed to be sequence specific and will only bind to the specific PCR product, so this method is more accurate. The probe specificity allows for quantification even in the presence of non-specific PCR product. Also multiplexing, assaying several genes in the same reaction, is made possible by using probes with different coloured labels.

  - TaqMan Probes

TaqMan probes are the most commonly used.
Typically, this method uses an RNA-based probe with a fluorescent reporter at one end and a quencher of fluorescence at the other end (Fig.2a). Breakdown of the probe by 5’ to 3’ exonuclease activity of taq polymerase removes the quencher and allow the PCR product to be detected (Fig. 2b).
TaqMan probes have been widely used for real-time RT-PCR for research purposes or diagnosis (Martell et al. 1999; Overbergh et al. 1999; Killgore et al. 2000)

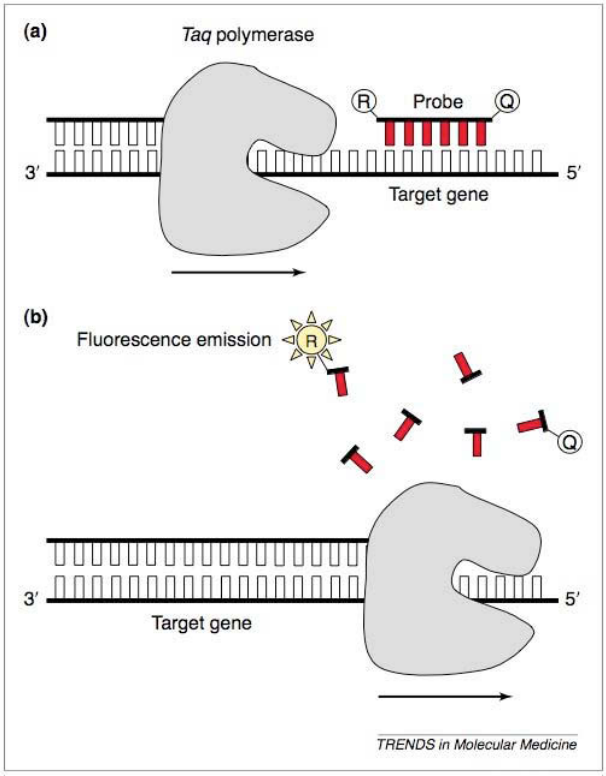


Figure 2: Principle of TaqMan probes (taken from Mocellin (2003))

- Molecular Beacons

Molecular Beacons are small pieces of DNA complimentary to the gene of interest labeled with a fluorescent reporter and a quencher molecule on opposite ends (Tyagi et al. 1996).

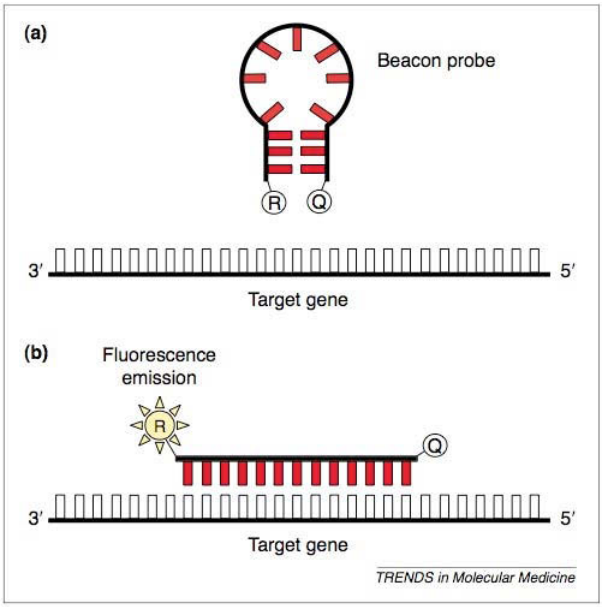


Figure 3: Molecular Beacons as depicted by Mocellin (2003) are hairpin-shaped molecules. The fluorescence of the internally quenched fluorophore is restored when they bind to a target nucleic acid.

These probes are designed to fold onto themselves to bring the reporter and the quencher in to closer proximity and minimize fluorescent emission (Fig. 3a). When the probe binds to the

gene of interest the probe takes up a linear conformation and reporter and quencher are separated resulting in the desired increase in fluorescence (Fig. 3b). Molecular beacons are especially suitable for identifying point mutations (Mhlanga et al. 2001).

Also Scorpions (Whitcombe et al. 1999) and hybridization probes (Bellin et al. 2001) are frequently employed in real-time analysis.

1.4.2. Normalization

The data have to be normalized by a reference gene, to overcome variations caused by differences in different amounts of initial sample, sample preparation, RNA/DNA-extraction, and reverse transcription (Pfaffl et al. 2004).
A gene sequence has to be chosen that is contained in all samples and as far as possible constant in expression and unaffected by experimental conditions (Heid et al. 1996; Neuvians et al. 2005). A number of reference (housekeeping) genes have been described in the literature and are used at different frequencies. According to Suzuki and colleagues the most frequently used reference gene is GAPDH (33%), followed by β-actin (32%) and 18S rRNA (14%) (2000). β-actin is a cell-structure element found in all eukaryotes and encodes a ubiquitous cytoskeleton protein. GAPDH is an abundant glycolytic enzyme, present in most cell types, but there is plenty of evidence that this is not a suitable endogenous control for quantification assays (Bustin 2000; Suzuki et al. 2000)
The amount of the target gene (the reporter protein) is normalized by dividing by the amount of reference gene.

1.4.3. Quantification

Two different methods are commonly used to quantify the results obtained by real-time PCR, the standard curve method and the comparative threshold method.

- Standard Curve Method

A sample of known concentration is used to construct a standard curve (Bustin 2000). The concentration of these samples can be measured and converted to the number of copies using the molecular weight of DNA or RNA.

- Comparative Threshold Method

Relative expression levels compared with a calibrator, which can be a control sample, are calculated with arithmetic formulas as described by Livak (2001).

$$\Delta\Delta C_T = \left(C_{T,\ target\ gene} - C_{T,\ housekeeping\ gene}\right)_{calibrator} - \left(C_{T,\ target\ gene} - C_{T,\ housekeeping\ gene}\right)_{sample}$$

The amount of target is then calculated as:

$$comparative\ expression\ level = 2^{-\Delta\Delta C_T}$$

The efficiency of PCR amplification for the target gene must be approximately equal to that of the housekeeping gene. This has to be tested for every individual target gene, and if the efficiencies of the two are not the same, which according to the report of Giulietti (2001) is the case quite often, another method for quantification has to be chosen. Hence, the advantages of this system are that no standards have to be constructed, and that more space on the instrument is available for unknown samples, saving time and money.

1.5. Reporter proteins

Reporter genes are attached to a regulatory sequence of a gene of interest to indicate successful uptake by expression of the reporter protein. It is therefore important to use a reporter gene that is not natively expressed in the cell or organism under study and can be easily identified and measured. Furthermore the use of different reporter proteins, like intra- or extracellular proteins, is very important, to scrutinize varying behaviour.

1.5.1. Green fluorescent protein – eGFP

Green fluorescent protein (GFP) is a spontaneously fluorescent protein isolated from the pacific jellyfish, *Aequorea victoria.* It transduces the blue chemiluminescence into green fluorescent light. Since the molecular cloning of GFP cDNA and demonstration of GFP as a functional transgene, GFP has become a powerful tool with exciting applications in developmental, cell and molecular biology (Tsien 1998).

In 2008 three scientists were awarded the Nobel Prize in Chemistry for isolation, characterizing, expressing the protein in *E. coli* and developing a palette of fluorescent proteins that could be used in many applications (Zimmer 2009).

GFP fluorescence is not species specific and can be expressed in bacteria, yeast, plant and mammalian cells. GFP can fuse with proteins of interest without interfering significantly with their assembly and function. Based on the structure of the GFP molecule, many GFP variants have been created with much improved fluorescence emission, or shifted excitation or emission spectra that are well suited for fluorescence microscopy and flow cytometry.

Although GFP expression can be easily detected under a fluorescence microscope, GFP fluorescence intensity varies from cell to cell because of the heterogeneity nature of GFP expression. In order to quantitate the GFP expression in cells, flow cytometric analysis is usually employed.

1.5.2. Human serum albumin – HSA

Human serum albumin (HSA) is the major component of human plasma and consists of a single non-glycosylated polypeptide chain of 585 amino acids, with a molecular weight of 66.5 kDa (Minghetti et al. 1986). HSA plays an important role in clinical therapy of severe hypoalbuminemia or traumatic shock (Peters 1985).

The non-glycosylated feature of HSA has made possible the wide range screening for host organisms with the correct structure and potential productivity. Secretion systems of many organisms, such as *Bacillus subtilis* (Saunders et al. 1987) and *S. cerevisiae* and other yeasts, such as *Kluyveromyces lactis* and *Hansenula polymorphis* (Sleep et al. 1991), have been widely studied as well as production in *P. pastoris* (Kobayashi et al. 2000). *P. pastoris* serves as an attractive host organism for the production of heterologous proteins, especially due to the fermentation properties for industrial process development.

Quantification of secreted HSA is possible with a commercially available ELISA-Kit (Bethyl laboratories, USA).

2. MATERIAL AND METHODS

2.1. *P. pastoris* strains

X-33 (wild type)

SMD1168 (a *his4*, *pep4* mutant)

SMD1168H (a *pep4* mutant)

Strains like SMD1168H show reduced capability to degrade foreign proteins due to a deletion in their *PEP4* gene (vacuolar protease A, Brierley 1998).

All these *P. pastoris* strains are from Invitrogen (Carlsbad, CA, USA).

2.2. pPUZZLE-vector system

A novel versatile vector backbone called pPUZZLE was constructed to allow for easy promoter exchange and equal genetic background in the present study. The vector backbone pPuzzle_zeoR_AOXTT, containing an *E. coli* origin of replication, an antibiotic (Zeocin) resistance marker cassette for both *E. coli* and *P. pastoris,* and a multiple cloning site (MCS) flanked by a transcription terminator of the *S. cerevisiae CYC1* gene region, is depicted in figure 4. Additionally the vector backbone contained a 800bp fragment of the 3´ region of *P. pastoris AOX1* gene for genome integration via homologous recombination (Stadlmayr et al. 2010a).

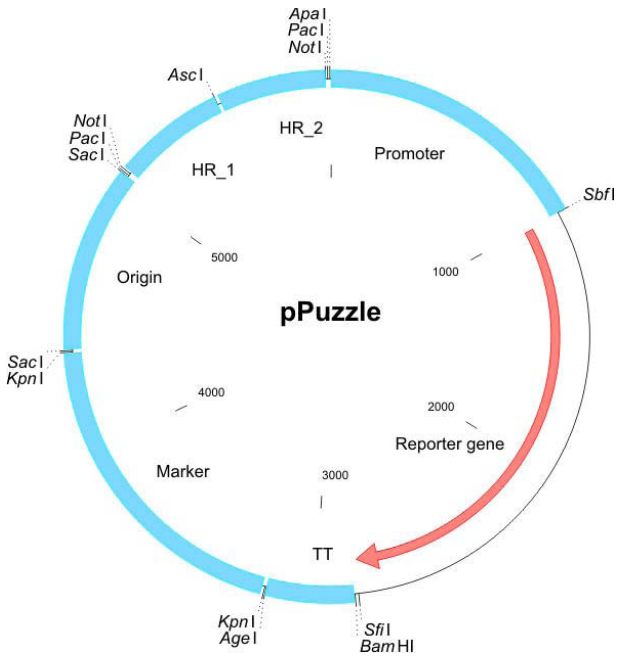


Figure 4: Schematic map of a pPUZZLE expression vector used in the present work

2.3. Media and stock solutions for yeast cultivation

All media and solutions have been heat-sterilized by autoclaving at 121°C and 2 bar for 20 min, exceptions are mentioned.

2.3.1. YPD-medium

10 g yeast extract
20 g soy peptone
20 g agar-agar if solidifying is required for plate cultivation
ddH_20 up to 900 mL
added after sterilization: 100 mL separately sterilized 10X glucose

- **Preparation:**

The required amounts of yeast extract and soy peptone were dissolved in water, if needed agar-agar was added, and sterilized by autoclaving. Before using the medium separately sterilized 10X glucose as well as a suitable selection marker if needed were added aseptically.
The liquid medium was stored at 16°C, whereas the agar-plates were kept at 4°C.

2.3.2. BM-medium

10 g yeast extract
20 g soy peptone
ddH_20 up to 700 mL
added after sterilization: 100 mL 1M phosphate buffer pH6
100 mL 10X YNB
2 mL 500X biotin
100 mL 10X glucose

- **Preparation:**

The required amounts of yeast extract and soy peptone were dissolved in water and sterilized by autoclaving.
Prior to use, after the medium has cooled down again, 100 mL of 10X glucose, 100 mL 1 M phosphate buffer (pH6), 100 mL YNB and 2 mL of 500X biotin were supplemented aseptically. The completed medium was stored at 16°C.

2.3.3. SSC-medium (synthetic shake flask medium)

3.15 g $(NH_4)_2HPO_4$
0.4920 g $MgSO_4*7H_20$
0.8040 g KCl
0.0268 g $CaCl_2*2H_20$
22.00 g citric acid monohydrate
22.00 g glucose monohydrate
1470 µL trace salt
370 µL 500X biotin
ddH_20 up to 1000 g

- **Preparation:**

The required amounts were dissolved in water. The pH was adjusted to pH 5.5 by adding KOH dropwise, then the medium was sterilized by filtration with Millex-GP filter (0.22 µm) and stored at 16°C protected from light.

2.3.4. RDB-agar

182.2 g sorbitol
20 g agar-agar
ddH_20 up to 700 mL
added after sterilization: 100 mL 10X glucose
100 mL 10X YNB
2 mL 500X biotin
10 mL 100X AA (amino acid stock solution)

- **Preparation:**

The required amount of sorbitol was dissolved in water, agar-agar was added and the solution was autoclaved. After cooling to about 60°C the needed amounts of also prewarmed stock-solutions were added. Prewarming of the stock-solutions was necessary to prevent the medium from becoming to thick for mixing. After carefully mixing the solution the plates were immediately poured and after solidifying stored at 4°C.

2.3.5. Stock solutions

- 10X glucose (20 %): 220 g D(+)-glucose-monohydrat / 1000 mL
- 10X glycerol (10 %): 100 mL (126 g) glycerol water free / 1000 mL

- 10X YNB (13.4 %): 134 g yeast nitrogen base / 1000 mL
  the solution was heated until YNB was completely dissolved, then sterilization by Millex-GP filter (0.22 µm)
- 500X biotin (0.02%): 20 mg D-biotin / 100 mL
  The solution was sterilized by filtration with Millex-GP filter (0.22 µm). Storage at 4°C.
- 100X AA : 0.5 % of each amino acid
  500 mg each of L-glutamic acid, L-methionine, L-lysine, L-leucine and L-isoleucine / 100 mL
  The solution was sterilized by filtration with Millex-GP filter (0.22 µm) and stored at 4°C.
- trace salt solution (PTM1): per 1000 mL:
  5.0 mL H_2SO_4 (95-98 %)
  65.0 g $FeSO_4*7H_2O$
  20.00 g $ZnCl_2$
  6.00 g $CuSO_4*5H_2O$
  3.36 g $MnSO_4*H_2O$
  0.82 g $CoCl_2*6H_2O$
  0.20 g $Na_2MoO_4*2H_2O$
  0.08 g NaI
  0.02 g H_3BO_3
  dissolved in HQ-water and sterilized with a 0.22 µm membrane filter. The solution was stored at 4°C.

2.3.6. Selection marker

Zeocin is a member of the bleomycin/phleomycin family of antibiotics. Zeocin causes cell death by intercalating into DNA and cleaving it. This antibiotic is effective on most aerobic cells and is therefore useful for selection in bacteria, eukaryotic microorganisms, plant and animal cells. The gene product of the *She ble* gene (included in the pPuzzle-vector) provides resistance against Zeocin by binding stochiometrically to the drug and therefore inhibiting its function.

The *She ble* gene derives from the *Streptoalloteichus hindustanus* bleomycin gene.

Zeocin™ stock solution: 100 mg / mL → used end concentration 25 µg / mL.

2.4. Preparation, purification and analysis of nucleic acids

2.4.1. Polymerase chain reaction – PCR

The polymerase chain reaction (PCR) is a technique widely used in many applications, ranging from molecular biology to forensic science, and various other fields.

The PCR method unzips a DNA double helix into two complementary strings, which are replicated by the DNA polymerase enzyme. During this process the number of DNA molecules is doubled, in repeating steps – a chain reaction – the original DNA template is exponentially amplified.

During one incubation cycle, the two DNA strands are separated (denaturation step) before the two primers can hybridize to the single stranded DNA (annealing step) where they serve as starting point for the DNA polymerase which synthesizes the complementary DNA strand in 5' direction (elongation step). This PCR cycle is repeated usually 20 to 30 times resulting in a high amount of a specific region of a DNA target.

A primer is a strand of nucleic acid that serves as a starting point for DNA replication because DNA polymerases can only add new nucleotides to an existing strand of DNA. Primers are synthetically produced oligonucleotides, which are complementary to the 3' or the 5' end of the target DNA sequence. PCR has also been extensively modified to perform a wide array of genetic manipulations, and for many other uses. Primers can for example contain restriction sites for cloning strategies.

All used primers were designed with appropriate software, to avoid complications such as primer dimers and secondary structures.

To overcome a frequently encountered problem in PCR amplification of target sequences, the appearance of smaller bands in the product spectrum, "Touchdown PCR" was applied in all PCR experiments (Don et al. 1991). During touchdown PCR, the annealing temperature is decreased in each step from a higher starting point down some degrees below the melting point of the primers.

Hotstart and touchdown PCR reduce accumulation of unspecific and therefore undesired PCR-product and therefore often improve amplification. Many polymerases require a hotstart at 95°C for activation. Usually 10min of elongation are added at the end of a PCR program to fill up unfinished polymerization products.

For control of the amplification reaction, 2-5 µL of the PCR-mixtures were analyzed and quantified by agarose gel electrophoresis after the PCR. If amplification was successful the rest of the mixture containing the template was purified with the mi-PCR Purification or the mi-Gel Extraction Kit as described in 2.4.4.

2.4.2. Restriction digest of DNA

Restriction enzymes are DNA-cutting enzymes found in bacteria. Bacteria use these enzymes to defend themselves against invading bacteriophages and viruses. Because restriction enzymes cut within the nucleic acid molecule they are also called restriction endonucleases. They recognize and cut DNA only at a particular sequence of nucleotides – the recognition sequence.

Restriction endonucleases are categorized into three general groups based on their composition and enzyme cofactor requirements, the nature of their target sequence, and the position of their DNA cleavage site relative to the recognition sequence

Restriction enzymes can be used for many applications, such as linearising plasmid DNA. During the cloning process they assist with the directed insertion of genes into vectors. To clone a gene fragment into a vector, both plasmid DNA and gene insert are typically cut with the same restriction enzymes, and then put together with the assistance of a DNA ligase.

Also restriction enzymes are used to digest genomic DNA for gene analysis by Southern blot. With the use of this technique it can, for example, be identified how many copies of a gene are present in one genome, as described in chapter 2.14.

2.4.3. Agarose gel electrophoresis

Agarose gels allow separation and identification of nucleic acids based on charge migration. Electrophoresis is a simple and highly effective method for separating, identifying, and purifying DNA fragments up to 25 kb (Voytas 2001). DNA fragments can be separated according to their size. This is achieved by moving negatively charged nucleic acid molecules through an agarose matrix with an electric field (electrophoresis) towards the positively charged anode. The relationship between the fragment size and rate of migration is non-linear, since larger fragments have greater frictional drag and are less efficient at migrating through the polymer. Smaller molecules travel farther because they are not retained by the net structure of the agarose gels, but conformation of the DNA molecule is also an important factor. To avoid this problem mostly linear molecules are separated, such as DNA fragments from a restriction digest, PCR products or also RNA fragments. The nucleic acid is visualized in the gel by addition of ethidium bromide. This binds strongly to DNA by intercalating between the bases and is fluorescent. The length of the sample fragments can be estimated by comparison to a DNA ladder containing DNA fragments of known size (e.g. MassRuler^{TM} DNA Ladder Mix, Fermentas).

Agarose gels are made between 0.7 % and 2 % dependent on the desired purpose. A 0.7 % gel shows good separation of large DNA fragments and a 2 % gel shows good resolution for small fragments.

- **Material and procedure**

- Ethidium bromide stock: 10 mg / mL
- 50X TAE buffer: 242 g Tris base
  57.1 mL glacial acetic acid
  0.5 M EDTA pH 8
  AD up to 1000 mL
- 1% agarose gel: 1 g / 100 mL agarose
  2 mL / 100 mL 50X TAE buffer
  4 µL / 100 mL ethidium bromide
- DNA loading buffer: 6X MassRuler DNA loading dye (Fermentas)
  containing bromphenol blue in glycerol
- molecular weight standard: MassRuler DNA ladder mix (Fermentas)

The basic protocol can be divided into three stages:

- a gel is prepared with an agarose concentration appropriate for the size of DNA fragments to be separated
- the DNA samples are loaded into the sample wells and the gel is run at a voltage and for a time period that will achieve optimal separation
- the gel is visualized directly upon illumination with UV light as the ethidium bromide has been incorporated into the gel.

For one gel 1.5 g of agarose was dissolved in 150 mL 1X TAE (3 mL of 50X TAE in 150 mL AD) and melted in the microwave. After cooling the complete clear solution to approximately 50°C, 7 µL of the ethidium bromide stock were added. The solution was then slightly shaked in order to achieve a homogenous distribution of the ethidium bromide but at the same time air bubbles had to be avoided. After that the gel was cast into a rinsed chamber with an appropriate comb in place. The completely solidified gel can be stored in 1X TAE at 4°C for several weeks.

For an electrophoresis run, the samples were mixed with 6X loading buffer (final concentration of loading dye is 1X) and loaded into the slots of the gel. The run is carried out in an electrophoresis chamber containing 1X TAE with ethidium bromide for 45 – 120 min at a constant voltage of 80 – 130 V. (Running time is dependent on length of fragments)

After the run had finished the gel was exposed to UV light on a transilluminator to make the DNA bands visible.

Exposure times were kept as short as possible to avoid denaturation of DNA if it was separated for preparative use and was intended to be further manipulated. The DNA can then be extracted from the gel by the mi-Gel Extraction Kit (metabion).

2.4.4. Purification of nucleic acid fragments

If PCR amplification or restriction digest were successful the remaining mixture containing the desired fragment had to be cleaned.

A very simple and easy to use method is the DNA purification by silica-membrane technology. The overall principle is the specific and selective adsorption of DNA molecules to the silica-gel–membrane in the spin column. After the binding of the DNA in optimized high-salt buffers to the surface, all contaminants like primers, nucleotides, enzymes, mineral oil, salts, agarose, ethidium bromide, and other impurities from DNA samples, can be removed through specialized wash buffers. The purified DNA can be eluted by low-salt buffer or distilled water.

2.4.4.1. PCR purification

The mi-PCR Purification Kit (metabion # mi-PP200) is designed for rapid purification of PCR products.

500 µL DNA Binding Buffer was added to the PCR reaction or restriction digest and the mixture was applied to a spin column containing silica-based membranes where the double stranded DNA is selectively absorbed.

After two washing steps with 500 µL of alcohol containing Column Wash Buffer, the DNA was eluted with usually 50 µL Tris (10 mM, pH 8) into a fresh microcentrifuge tube.

The centrifugation conditions between the different steps were 13.000 rpm for 1 min at room-temperature.

2.4.4.2. Gel extraction

The mi-Gel Extraction Kit (metabion # mi-GE100) was used to recover pure DNA from agarose gels. This kit also uses silica-based membranes in spin-columns.

For gel band purification the desired DNA band was cut out of the gel with a clean razor blade under the UV-transilluminator and transferred to a microcentrifuge tube. 300 µL of Gel-Extraction Buffer were added and the tube was then incubated for approximately 10min at 60°C under vigorously shaking. As soon as the agarose was completely dissolved the mixture containing the desired DNA band was applied to a spin-column. After following the same washing steps as described above, the DNA was again eluted with 50 µL of 10 mM Tris (pH 8).

2.4.5. Quantification of nucleic acids

The most common method for quantifying DNA or also RNA samples is by conventional absorbance measurements. Nucleic acids have an absorption maximum at 260 nm. Both RNA and DNA absorb UV light very efficiently making it possible to detect and quantify either.
As most samples contain contaminates such as proteins and single stranded DNA and RNA that absorb maximally at 280 nm it is essential to determine the A260/A280 ratio. The higher the ratio, the more pure the sample. (Range: 1.8 – 2.0)

Using a 1 cm light path, the concentration of the DNA or RNA in the sample is calculated, based on the extinction coefficient of 20, as follows:

DNA conc. (µg / mL) = (OD_{260}) x dilution factor x (50 µg DNA / mL) / (1 OD_{260}unit)
RNA conc. (µg / mL) = (OD_{260}) x dilution factor x (40 µg RNA / mL) / (1 OD_{260}unit)

All measurements were taken at a Nanodrop ND-1000 spectrophotometer, which allows UV-Vis measurements to be made from 1 µL of sample with no cuvettes or dilutions. A sample droplet was simply loaded onto the optical pedestal. The sample is then drawn into a column and measured, resulting UV-Vis spectra are displayed.

2.5. Transformation of *P. pastoris*

The yeast cells have to be specially treated for being “competent” for successful transformations.

2.5.1. Linearization of vector DNA

Homologous recombination between the vector and the corresponding site at the *Pichia* genome takes place resulting in stable transformants. For the single cross over event the vector has to be linearized within a given locus, here upstream of the *AOX1* locus. By choosing an appropriate restriction enzyme care should be taken not to apply an enzyme that also cuts within the gene of interest.
The expression vectors of each used construct in this work were linearized with AscI in the 3´-AOX1 integration region and transformed into *P. pastoris* by electroporation.

Table 1: 40 µL restriction digest mix for vector linearization

µL	component
34	vector DNA (100 ng to 5 µg DNA in 10 mM Tris/Cl pH 8)
4	Buffer (NEBuffer4)
2	AscI (NEB)

The restriction digest mix was incubated at 37°C for 2 h. After the incubation time an aliquot (2 – 5 µL) was analyzed on an agarose gel and the rest of the mixture was purified with the mi-PCR Purification kit (metabion) as described before.

2.5.2. Preparation of competent *P. pastoris*

- **Material and Procedure**

- YPD-Medium (250 mL)
- 1 mM sterile HEPES buffer (approximately 1500 mL)
- 1 M sterile sorbitol (50 mL)
- 1 M DTT (1.25 mL)

- BECKMAN centrifuge with rotor JLA 10.000 cooled down to 4°C

An overnight culture was inoculated with a single colony from a fresh *P. pastoris* agar-plate in 10 mL of YPD-medium and incubated at 30°C under vigorous shaking. From this pre-culture 0.1 to 0.5 mL were used to inoculate the main-culture in 200 mL of fresh YPD-medium and also grown overnight at 30°C up to an OD_{600} of 3. The volume of the pre-culture for inoculation is calculated according to the exponential growth formula using $\mu = 0.347\ h^{-1}$.
Once the culture had reached the required cell density it was transferred into sterile centrifuge tubes and centrifuged at 1,500 g and 4°C for 5 min.
The pellet was carefully resuspended in 50 mL of prewarmed YPD with the addition of 1 mL 1 M HEPES and 1.25 mL 1 M DTT. After an incubation time of about 15 min at 28°C under shaking conditions, 200 mL of sterile and ice-cold water were added and the cells were harvested by centrifugation as described above. The resulting yeast pellet was washed in 250 mL of ice-cold, sterile 1 mM HEPES buffer with the same centrifugation settings. After removal of the supernatant the yeast pellet was carefully resuspended in 20 mL of sterile and ice-cold 1 M sorbitol and again centrifuged.
At last the competent yeast cells are taken up in 1 mL of 1 M sorbitol (sterile, ice-cold) and distributed in cooled microcentrifuge tubes in aliquots of 80 µL.
Competent *P. pastoris* cells can be used immediately, or stored at - 80°C.

2.5.3. Electroporation of *P. pastoris*

- **Material and procedure**

- 1 M sterile, ice-cold sorbitol
- YPD-Zeo-agar plates

- BioRad Gene Pulser™

An 80 µL aliquot of the competent *P. pastoris* cells was mixed very gently with 25 µL of the linearized and purified vector DNA. As a negative control instead of DNA Tris-buffer was used. The yeast-DNA mixture should not be vortexed to avoid shearing of the DNA.
The mixture was transferred into a chilled electroporation cuvette (2 mm) and after incubation on ice for about 5 min electroporation was performed using the following parameters:

Table 2: Electroporation parameters for transformation of *P. pastoris*

voltage	2000 V
capacitance	25 µF
resistance	200 Ω

Immediately after transformation approximately 400 µL 1 M sorbitol were added to the cells in the cuvette before they were carefully transferred into a sterile microcentrifuge tube with a total volume of 1 mL 1 M sorbitol. The yeast cells are allowed to regenerate for about 1.5 h at 28°C before they were plated in aliquots on selective agar plates. The agar plates were incubated at 28°C until colonies appeared. Some of the colonies were then chosen for further analysis.
The yeast cells only treated with Tris-buffer should not be able to grow on the Zeocin selection, otherwise the experiment had to be redone.

2.6. Cryocultivation

Cryoculture cultivation is required for long term storage and preservation of strains. For short periods up to several weeks, cells can be stored by repeated passages on appropriate agar plates at 4°C.

- **Procedure**

5 mL of YPD medium were inoculated with a single colony from an agar plate and grown overnight at 30°C under vigorous shaking.

Aliquots of the overnight cultures were transferred into cryo tubes (Nunc) and carefully mixed with 100 % glycerol. The final concentration of glycerol was 20 %. Immediately after mixing the cultures were frozen at -70°C.
To bring cryo-preserved strains in culture again, cultures were taken from the frozen surface with a sterile inoculation loop and used either to inoculate liquid media or agar plates.

2.7. Screening of *P. pastoris* transformants

Screening methods in general are used to examine large numbers of positive transformants in order to find the desired properties, such as production of recombinant proteins. *P. pastoris* screening methods are based on a two-stage cultivation. The first step is a pre-culture in a rather small volume of medium to gain cells in the same growth phase, the second step is the main-culture with expression of the recombinant protein.
Biomass accumulation is followed by measuring the optical density OD_{600} while the production of the recombinant protein is analyzed after harvesting either the cells or the supernatant dependent on the location of the protein.
Shake flask experiments were carried out in 100 mL shake flasks at 28°C with 170 rpm (rotations per minute).

2.7.1. Expression screening

- **Material and procedure**

- YPD-medium (5 mL per clone for pre-culture)
- BM-medium or SSC-medium (10 mL per clone)
- 10X Glucose

Several colonies on the transformation plates were picked with a sterile toothpick and cultured onto a masterplate (selective Zeo-YPD-agar) as well as in 5 mL of YPD-medium and grown overnight at 28°C and vigorous shaking. After measuring the OD_{600} of this pre-culture, applying aseptic handling, a main-culture with 10 mL of BM-medium was inoculated corresponding an optical density of 0.1 in shake-flasks.
After 48 h of cultivation including feeding with 1 % glucose every 12 hours the OD_{600} was measured again to determine cell growth. The culture was then harvested by centrifugation at room temperature. The supernatant as well as the cell pellets were stored at -20°C for further analysis.

The cultures for the RNA isolation were grown in a 3-step mode. The first two steps were the same as described above. The main culture was grown overnight at 28°C and shaking. To be able to harvest cells in the exponential growth phase a second culture was prepared, which was grown for only 6 h. The required volume for an OD of 1 was transferred into 10 mL BM with glucose. After 6 h, the samples had to be immediately quenched for subsequent RNA isolation, therefore 6 mL of culture were cooled on ice. The cells were mixed with an ice-cold mix of 100 % ethanol and 5% phenol at a ratio of 2:1, aliquoted à 2 mL and centrifuged at 4°C and 13,000 rpm. The special treated cell pellet was then directly stored at -70°C.

- **Cell growth measurement by optical density (OD_{600})**

The optical density or turbidometry is used as measurement for the concentration of particles including microorganisms in a solution. Since there is a linear correlation between optical density and dry weight, biomass concentrations of a specific organism can be estimated according to the law of Lambert-Beer by measuring the absorbance at a wavelength of 600 nm using a 1 cm cuvette. Samples showing an absorbance higher than 0.8 have to be diluted.

2.8. Bioreactor cultivation

Fed batch cultivation was carried out in a 3.5 L working volume bioreactor (Minifors, Infors, Switzerland) as described by Stadlmayr and coworkers (2010a).

The inoculum was cultivated in shaking flasks (100 mL YPG) and incubated at 28°C and 180 rpm for 24 hours. The cultures where then washed by centrifugation (3,000 x g, 10 min, 25°C) and resuspended in 100 mL sterile batch medium. This was used to inoculate the starting volume of 1.75 L in the bioreactor to an optical density (600 nm) of 1.0.
Briefly, cells were grown for 24 h in batch on glycerol medium, followed by a fed batch phase of glucose medium, employing an optimized feeding strategy designed to maximize the volumetric productivity as described by Maurer (2006).
Cultivation temperature was controlled at 25°C, pH was controlled at 5.8 by addition of 25% ammonium hydroxide and the dissolved oxygen concentration was maintained above 20% saturation by controlling the stirrer speed between 600 and 1200 rpm, whereas the airflow was kept constant at 240 L h^{-1}.
Samples were taken frequently and analyzed for cell dry weight, HSA product concentration and HSA transcript levels.

2.9. Analysis of reporter proteins

2.9.1. eGFP

2.9.1.1. Flow cytometric analysis

Flow cytometric analysis was performed on a FACSCalibur system (Becton Dickinson, Franklin Lakes, NJ, USA), equipped with a 488 nm excitation 15 mW air-cooled argon-ion laser and a 630 nm diode laser. The fluorescence emission signal (FL1) was acquired in logarithmic mode through a 530/30 band-pass filter. Forward scatter (FSC) and side scatter (SSC) were acquired in logarithmic mode.
For each sample 1.0 x10^4 cells were analyzed. Cell debris was excluded by adjusting the threshold setting. Auto-fluorescence of *P. pastoris* cells was measured for *P. pastoris* wild type cells ($FL1_{blank}$).

Calculation of eGFP expression levels (Hohenblum et al. 2003):

$$rfu = \frac{1}{n} * \sum_{n}^{1} \frac{FL1_{sample} - FL1_{blank}}{\sqrt{FSC^3}} \qquad \text{(Eq. 1)}$$

rfu: relative fluorescent units
FL1: geometric mean of 10^4 events in fluorescence channel 1 BP filter 530/30
FSC: geometric mean of 10^4 events in forward scatter

2.9.1.2. eGFP-ELISA

The enzyme-linked immunosorbent assay (ELISA) is a method to determine the concentration of specific molecules present in a sample. It is based in the specifity of antibodies and the possibility to attach proteins to polystyrol-surfaces of microtiter plates.

For the determination of the content of recombinant protein (eGFP), cell aliquots equivalent to approximately 1.0 mg yeast dry mass of the samples were centrifuged (15,000 x g, 1 min); the supernatants were removed. The cells were resuspended in 500 µL of buffer (18 mM Tris-HCl pH = 8.2, 5 mM EDTA, 0.5 %(v/v) Triton X-100, 7 mM β-mercaptoethanol,) and an equal volume of glass beads (0.50 mm in diameter; Sartorius) was added. For cell disruption, samples were treated in a Thermo Savant Fastprep FP120 three times (speed 6.5 m s^{-1} for 30s). After centrifugation, the supernatant containing eGFP was removed. The GFP content was measured using the enzyme-linked immunosorbent assay (ELISA) method described by Reischer (2004).

- **Material and procedure**

- Coating buffer: $NaHCO_3$-buffer, pH 9.6 – 9.8
- Washing buffer: PBST (1X phosphate based saline (PBS) pH 7.4 + 0.1 % Tween-20)
  per litre: 1.15 g $Na_2HPO_4*2H_20$
  0.2 g KH_2PO_4
  0.2 g KCl
  8.0 g NaCl
  1 mL Tween-20
- Blocking solution / Dilution buffer: PBST + 1 % BSA (bovine serum albumin, Sigma)
- Staining buffer (per litre): 7.3 g citric acid monohydrate
  11.86 g Na_2HPO_4
- Staining solution: 10 mL staining buffer with 100 µL OPD-stock solution (10 mg/mL)
  + 3.3 µL H_2O_2
- Stopping solution: 1.25 M H_2SO_4
- Antibodies: monoclonal mouse anti-GFP (Sigma)
  goat anti-GFP biotin-conjugated antibody (DPC Biermann)
  Avidin-HRP conjugate (Amersham Biosciences)

Briefly, ELISA microwell plates (Nunc, MaxiSorp) were coated with monoclonal mouse anti-GFP (Sigma) 1:1,000 in coating buffer (100 µL/well) at 4°C overnight and then washed three times with washing buffer (200 µL/well) and dried. Blocking solution (100 µL/well) was applied to the plate for 2 h at room temperature. In the mean time samples were diluted on a separate dilution plate (Nunc) in dilution buffer in 1:2 steps (the GFP standard (BD Bioscience Clontech) starts with 20 ng/mL). The coated microwell plates were again washed to get rid of the blocking solution, and 50 µL of the standards and diluted samples were transferred. After 1 h incubation at room temperature the plate was washed and the 1:2,000 diluted biotin-conjugated antibody (in dilution buffer, 50 µL/well) was applied for 1 h to the plate. After washing the Avidin-HRP conjugate (1:1,600 in dilution buffer) was added (100 µL/well) and again incubated for 1 h at room temperature and washed afterwards. For detection a staining solution (100 µl/well) was added to each well and incubated until a yellow colour was visible for the standard. To stop the reaction 100 µL of stopping solution was applied (colour changes from yellow to orange) and read at 492 nm against a reference wavelength of 620 nm in the microtiter photometer (TECAN Sunrise). Data evaluation was performed by using a calibration curve (GFP standard) and 4-parameter analysis.

2.9.2. HSA

2.9.2.1. SDS-PAGE (polyacrylamide gel electrophoresis)

The purpose of SDS-PAGE is to separate proteins according to their size, and no other physical feature. SDS (sodium dodecyl sulfate) is a detergent that can dissolve hydrophobic molecules but also has a negative charge attached to it. The samples have to be heated under denaturing conditions. SDS forces polypeptide chains into extended conformations with similar charge to mass ratios. SDS treatment therefore eliminates the effects of differences in shape so that chain length, which reflects mass, is the sole determinant of the migration rate of proteins in SDS – polyacrylamide electrophoresis. The molecular weight of a protein can be determined by comparison to a protein ladder or molecular weight ladder, which is run on the same gel. Also reducing agents such as dithioerythritol (DTT) can be added to reduce disulfide bonds and thereby further cause dissociation of the protein. After this treatment the polypeptides appear as random coil.
Under these conditions proteins move through the gel according to their molar masses, the resulting bands can be made visible either by Coomassie Blue staining or the more sensitive silver staining method.

- **Material and procedure**

- Invitrogen NuPage System:
  - 12 % Bis-Tris-Gel (1.0 mm, 15 well)
  - MOPS SDS Running buffer 20X (Running buffer)
  - LDS Sample buffer 4X (Sample buffer)
- Invitrogen BenchMarkProtein Ladder

The supernatant of the harvested cells was diluted if necessary and according to table 3 prepared.

Table 3: Sample preparation

non-reducing	reducing	component
conditions		
15 µL	13 µL	sample
-	2 µL	1 M DTT
5 µL	5 µL	LDS sample buffer
-	denature mixture for 10min at 99°C	

15 µL of this mixture were then loaded onto the SDS polyacrylamide gel.
Both chambers of the Invitrogen XCell SureLock Mini-Cell were filled with 800 mL of running

buffer before the pre-cast gels were inserted according to the supplier's instruction.
The gel is then run under constant voltage at 150 V until the tracking dye bromphenolblue has reached the bottom of the gel cassette usually in about 60 min.
After finishing the run, the gel has to be fixed and stained immediately, unless transfer to a Western blot membrane is desired.

- **Coomassie Blue staining – Material and procedure**

Coomassie Blue staining was done using the PageBlue Protein Staining Solution from Fermentas. This is a ready to use, Coomassie G-250 stain for visualizing protein bands on polyacrylamide gels.
The gel was rinsed with distilled water 3 times for 10min each and then incubated in the staining solution under slight shaking until the desired intensity of the respective bands was reached.

- **Silverstaining – Material and procedure**

- Fixing solution: 50 % (v/v) ethanol
  10 % (v/v) acetic acid
- Incubation solution: 30 % (v/v) ethanol
  0.83 M sodium acetate
  13 mM sodium thiosulfate
  0.25 % glutaraldehyde (add before use)
- Staining solution: 6 mM $AgNO_3$
  0.02 % formaldehyde
- Development solution: 0.25 M Na_2CO_3; 0.01 % formaldehyde
- Stop solution: 50 mM EDTA

Immediately after electrophoresis the gel was incubated in fixing solution for 1 hour minimum. All incubations were carried out at room temperature under slight shaking. After 30 min of shaking in incubation solution, the gel was rinsed with distilled water 3 times for 10 min each. Silver ions were allowed to bind to cysteine residues of the separated proteins on the gel during 20 min of incubation in staining solution, then the gel was washed again with distilled water. The silver ions were then reduced to molecular metallic silver in the alkaline, formaldehyde-containing containment of the development solution. Staining was allowed to proceed until the desired intensity of the respective bands was reached, followed by addition of stop solution.
Gels can be stored in distilled water for some days.

2.9.2.2. HSA-ELISA

HSA from culture supernatant was quantified by the Human Albumin ELISA Quantitation Kit (catalog No. E80-129, Bethyl Laboratories Inc. Montgomery, TX, USA) following the suppliers instruction manual.

- **Material and procedure**

- Coating buffer: 0.05 M carbonate-bicarbonate pH 9.6
- Basic Solution (BS) (pH 8): 50 mM Tris-HCl
  140 mM NaCl
  can be stored at 16°C
- Wash Solution: BS + 0.05 % Tween-20
- Blocking Solution: BS + 1 % BSA
- Sample Diluent (pH8): BS + 1 % BSA + 0.05 % Tween-20
- Staining Buffer and Staining Solution: see eGFP-ELISA (2.9.1.2)
- Stopping Solution: 2 M H_2SO_4
- Antibodies: Coating Antibody: Goat anti-human Albumin-affinity
  HRP Detection Antibody: Goat anti-human Albumin-HRP conjugate
- Human Reference Serum (HSA Standard)

Briefly, MaxiSorp Microwell Plates (Nunc) were precoated overnight with 100 µL/well of the 1:1,000 diluted capture antibody at room temperature. The precoated plates were washed three times with Wash Solution to remove unbound antibody. Next step was the application of 200 µL/well Blocking Solution for at least 30min and again washing three times. The HSA standard was used with a starting concentration of 1 µg mL^{-1}. Samples were diluted accordingly in Sample Diluent. 100 µL/well Sample Diluent was applied to each well of row A to G of the plate. In row H the HSA standard and the diluted samples were applicated (200 µL/well). A dilution series was reached by taking 100 µL from row H and transferring it into row G (1:2 dilution), continuing this until row A. The last 100 µL were discarded, so that 100 µL remained in each well. After incubation on a shaker for 1 h at room temperature the plate was washed 5 times. The HRP detection antibody was applied in a 1:75,000 dilution (100 µL/well) and the plates were again incubated for 1 h. After 5 times washing 100 µL of Staining Solution was applied to each well. The plates were developed until the standard showed a good colour gradient and the reaction was then stopped with 100 µL Stopping Solution per well. Using a microtiter photometer (TECAN Sunrise) the plates were read at 492 nm against a reference wavelength of 620 nm. Data evaluation was again performed by using a calibration curve and 4-parameter analysis.

2.10. Preparation, purification and quantification of RNA

One of the major difficulties in working with RNA is its susceptibility to hydrolytic cleavage by the enzyme ribonuclease (RNase). RNase is literally everywhere – it is secreted from the skin and tends to coats all surfaces. RNase is also very stable and cannot be inactivated by autoclaving or ethanol cleaning. Very effective is the chemically inactivation of RNase by the methylating agent diethylpyrocarbonate (DEPC).
DEPC was dispensed in a fume hood and added to water in an autoclaveable bottle (Schott) containing a stir bar. After one hour of continuous stirring the solution was autoclaved to denature the DEPC.
To avoid contaminations by RNase it was important to handle solutions, materials and samples with extra care. (gloves, separate stocks of solutions, RNase free tips...)

- **Material and procedure**

- TRI Reagent (Sigma)
- chloroform
- isopropanol
- 75 % ethanol
- RNase free water

- acid treated glass beads (0.50 mm diameter, Sigma)
- RNase free tips, tubes...
- separate waste for TRI and phenol

The acid-phenol extraction is a simple procedure to extract RNA from cells, and to separate RNA from DNA. After quenching of the cells, harvested after 6 h of incubation (the cells should then be in the exponential growth phase, needed for effective RNA isolation, see chapter 2.7.1) the frozen cell pellet was resuspended in 1 mL TRI Reagent. Cells were disrupted after addition of 500 µL glassbeads with a FastPrep cell homogenizer (Thermo Savant) using 6 m s^{-1} for 2 x 30 s and chilling the cells on ice in between. RNA was extracted using 200 µL chloroform, vigorously shaking and approximately 15 min of incubation at room temperature. After centrifugation at full speed for 15 min and 4°C, the sample was separated in 3 distinct phases. A colourless upper aqueous phase – containing desired RNA, an interphase – containing DNA, and a red organic phase – containing protein. The aqueous phase was transferred to a fresh, RNase free microcentrifuge tube and 500 µL isopropanol were added to precipitate the RNA. After 10 min of incubation at room temperature the sample was centrifuged again at full speed and 4°C. After this step the RNA precipitate should be visible as a pellet on the side or bottom of the tube. The supernatant was removed

and the pellet washed in 1 mL of 75 % ethanol. The final step was the solubilization of the air-dried RNA pellet in an appropriate volume (50 µL) of RNase free water and incubation at 55°C for about 10 min.

To eliminate remaining DNA, samples were treated with DNAfree (Ambion) following the manufacturer's instructions. 5 µL of 10X DNase- buffer and 1 µL rDNase 1 were added to the RNA-solution. After an incubation at 37°C for 20-30 min 5 µL DNase-Inactivation-Reagent was added for 2 min followed by centrifugation at 10,000 g for 1.5 min. The supernatant containing the RNA was transferred in a fresh RNase-free tube.

Extracted RNA was quantified by measuring A230/260/280 (Nanodrop, USA) as described before (Quantification of Nucleic Acids 2.4.5).

2.11. cDNA synthesis (reverse transcription)

2.5 µg of total RNA was reverse transcribed in a 20 µL reaction volume using the Superscript III (Invitrogen, USA) reagent set.

Table 4: Mix and procedure for Reverse Transcription

µL	component
11	2.5 µg total RNA in RNase free water
1	oligo dT (500 ng)
1	dNTPs (10 mM)
65°C, 5 min cool on ice	
4	5X 1st strand buffer
1	0.1 M DTT
1	RNase out (40 U)
1	Superscript III (200 U)
55°C, 60 min	

After inactivation of the enzyme by heating to 70°C for 15 min, template RNA was removed by treating the reaction with 2 µL of 2.5 M NaOH and incubation at 37°C for approximately 10min, to stop this digestion 20 µL of 1 M HEPES were added.

The obtained cDNA was then purified with the mi-PCR-Purification Kit and dissolved in 50 µL 10 mM Tris/Cl with pH 8 and stored at -20°C.

2.12. Preparation, purification and quantification of genomic DNA

Genomic DNA constitutes the total genetic information of an organism.

For preparation of genomic DNA from *P. pastoris* the DNeasy® Blood & Tissue Kit (QIAGEN) was used, following the suppliers supplementary protocol for the purification of total DNA from yeast. A special sorbitol buffer (1 M sorbitol, 100 mM sodium EDTA, 14 mM β-mercaptoethanol) was used together with lyticase for lysis of the yeast cells from an overnight culture prepared with 20 mL YPD. After completing the protocol all samples were visually checked on an agarose gel and the concentration of isolated DNA was measured with a spectrophotometer at 260 nm (Nanodrop, USA) and additionally also by using the Hoechst-dye 33258. Samples displaying 260 nm/280 nm ratios of 1.8 to 2.0 were used for further analysis.

2.12.1. Quantification of genomic DNA with Hoechst 33258

Spectrophotometry and fluorometry are commonly used to measure DNA concentration. Spectrophotometry can be used to measure microgram quantities of pure DNA samples (i.e. DNA that is not contaminated by proteins, phenol, agarose, or RNA). Fluorometry is more sensitive, allowing measurement of nanogram quantities of DNA, and furthermore the use of Hoechst 33258 dye allows specific analysis of DNA, since this dye selectively binds to DNA.

Hoechst 33258, a bis-benzimidazole derivate fluoresces 20 times stronger when bound to DNA compared with ethidium bromide, the most commonly used fluorescent dye for staining DNA in agarose gels and quantification of nucleic acids. The used working pH for Hoechst 33258 is pH 7.4, where the fluorescence of the compound alone is minimized and the dye-DNA complex is maximized (Cesarone et al. 1979). Also it is hardly affected by common laboratory reagents (Paul et al. 1982).

Hoechst 33258 binds to the AT rich regions of double-stranded DNA and exhibits enhanced fluorescence under high ionic strength conditions. The reaction is excited at 365 nm and emits light at 458 nm where the measurement is taking place. The DNA content of the samples is determined according to a calf-thymus DNA standard curve. Calf-thymus DNA often serves as a reference because it is double stranded, highly polymerized and is approximately 58 % AT.

- **Material and procedure**

- HITACHI F-2000 Spectrofluorometer
- Suitable fluorescence quartz cuvette

- 10 x TNE: 1 M Tris pH 7.4 adjusted with HCl
  1 M NaCl
  10 mM EDTA
- Dilution Buffer: 1:10 dilution of 10 x TNE
- Fluorescence reagent: 100 mL dilution buffer and 10 µL of Hoechst 33258 (10 mg / mL)

- Calf-thymus DNA standard:

4 DNA standard solutions (1.25, 2.5, 5 ,10 µg / mL) are made starting from a calf-thymus DNA stock solution (200 µg / mL) (Sigma, D3664)

- Samples:

The isolated DNA samples are diluted 1:10 in order to fit into the calibration range.

- Measurement:

The measurement is taking place at a HITACHI F-2000 Spectroflurometer, which is calibrated with calf-thymus DNA standard dilutions ranging from 0 to 10 µg DNA / mL.
1.9 mL of Fluorescence reagent are provided in a fluorescence quartz cuvette and the cuvette is put into the analyzer with the blue mark in front.
100 µL of DNA-Standard (0, 1.25, 2.5, 5, 10 µg / mL DNA) or DNA-sample respectively are added to the cuvette, mixed by pipetting up and down and the fluorescence measurement is carried out by pressing the start button. Accurate pipetting and thorough mixing, but at the same time avoiding air bubbles, is critical for reproducible results.
The DNA contents are calculated based on a computer generated calibration curve and the sample dilutions.

2.13. Quantitative real-time PCR (qRT-PCR)

- **Material**

- Rotor-Gene 6000^{TM} (Corbett Life Science, Australia)
- 2X SensiMix Plus SYBR (QT-605-02/05) (GenXpress)
- HQ-water
- 0.1 mL PCR tubes with strips (GX-18-3001-002) (GenXpress)

2.13.1. Primer design

To avoid complications such as primer dimers and secondary structures, primers were designed with appropriate software.

Following aspects were considered (as far as possible):

- melting temperature (T_m) between 58°C to 60°C
- primer length 19-24 bps
- amplicon size 100-200 bps
- amplicon content of the bases guanine and cytosine about 45-55 %
- the last 5 bases at the 3 prime end should have no more that two G's or C's

(Quantitative RT-PCR Protocol SYBR Green I, Schnable lab, Iowa State University, USA)

Additionally, amplicons should be in the 3 prime region of the gene.

Primer names, primer sequences and characteristics are summarized in table 5.
Table 6 shows relevant amplicon informations for qRT-PCR.

Table 5: Used primers for qRT-PCR

primer	sequence	T_m(°C)	primer length (bp)
Actin fw	CCTGAGGCTTTGTTCCACCCATCT	61.3	24
Actin rv	GGAACATAGTAGTACCACCGGACATAACGA	61.4	30
HSA fw	AAACCTAGGAAAAGTGGGCAGCAAATGT	62.9	28
HSA rv	ACTCTGTCACTTACTGGCGTTTTCTCATG	61.0	29
eGFP fw	TCGCCGACCACTACCAGCAGAA	61.4	22
eGFP rv	ACCATGTGATCGCGCTTCTCGTT	61.6	23
eGFP NEW fw	CCCGCGCCGAGGTGAAGT	60.5	18
eGFP NEW rv	TTCTTCTGCTTGTCGGCCATGATATAG	60.6	27

Table 6: Product characteristics for qRT-PCR

Primerpair	PCR amplicon size (bp)	Product GC-content (%)
Actin	148	45.9
HSA	135	43.7
eGFP	124	64.5

Actin sequence (1085 bp):

ATGTGTAAGGCCGGATACGCCGGAGACGACGCCCCACACACAGTGTTCCCATCGGTCG
TAGGTAGACCAAGACACCAAGGTGTCATGGTCGGTATGGGTCAAAAGGACTCCTTCGT
CGGTGACGAGGCTCAATCCAAGAGAGGTATCTTGACCTTGAGATACCCAATCGAGCAC
GGTATCGTCACTAACTGGGACGATATGGAAAAGATCTGGCACCACACCTTCTACAACGA
GTTGCGTCTGGCCCCAGAAGAGCACCCAGTTCTTTTGACTGAGGCTCCAATGAACCCAA
AGTCCAACAGAGAGAAGATGACCCAAATCATGTTCGAGACTTTCAACGTTCCAGCCTTC
TACGTTTCTATTCAGGCCGTTTTGTCCCTGTACGCTTCCGGTAGAACCACTGGTATCGTT
TTGGACTCTGGTGACGGTGTTACCCACGTTGTCCCAATTTATGCCGGTTTCTCCTTACC
ACACGCTATTTTGCGTATCGACTTGGCCGGTAGAGATTTGACCGACTACTTGATGAAGA
TCTTGTCTGAGCGTGGTTACACTTTTTCTACCTCTGCTGAGAGAGAAATCGTCCGTGAC
ATCAAGGAGAAGCTTTGTTACGTTGCTCTTGACTTTGACCAGGAATTGCAAACTTCTTCT
CAATCTTCATCCATTGAGAAGTCTTACGAGTTGCCAGATGGCCAAGTTATCACTATCGGT
AACGAGAGATTCAGAGCT**<u>CCTGAGGCTTTGTTCCACCCATCT</u>GTACTTGGCCTTGAGG**
CTTCTGGTATCGACCAAACCACTTACAACTCCATCATGAAGTGTGATGTTGATGTTCGT
AAGGAACTCTACAGTAACATCG<u>TTATGTCCGGTGGTACTACTATGTTCC</u>CAGGTATTG
CTGAGCGTATGCAAAAGGAGCTTACTGCCTTGGCTCCATCTTCGATGAAGGTCAAGATT
TCTGCTCCACCAGAAAGAAAGTACTCCGTATGGATCGGTGGTTCTATCCTCGCTTCTTT
GGGTACTTTCCAACAAATGTGGATCTCAAAGCAAGAGTACGACGAATCTGGACCATCCA
TTGTGCACCTCAAGTGTTTCTAAGT

eGFP sequence (720 bp)

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTG
GACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCC
ACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCT
GGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGA
CCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAG
CGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCG
AGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACG
GCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATG
GCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGG
ACGGCAGCGTGCAGC**<u>TCGCCGACCACTACCAGCAGAA</u>CACCCCCATCGGCGACGGC**
CCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGAC
CCC<u>AACGAGAAGCGCGATCACATGGT</u>CCTGCTGGAGTTCGTGACCGCCGCCGGGATC
ACTCTCGGCATGGACGAGCTGTACAAGTAA

HSA sequence (1830 bp)

ATGAAGTGGGTAACCTTTATTTCCCTTCTTTTTCTCTTTAGCTCGGCTTATTCCAGGGGT
GTGTTTCGTCGAGATGCACACAAGAGTGAGGTTGCTCATCGGTTTAAAGATTTGGGAGA
AGAAAATTTCAAAGCCTTGGTGTTGATTGCCTTTGCTCAGTATCTTCAGCAGTGTCCATT
TGAAGATCATGTAAAATTAGTGAATGAAGTAACTGAATTTGCAAAAACATGTGTTGCTGA
TGAGTCAGCTGAAAATTGTGACAAATCACTTCATACCCTTTTTGGAGACAAATTATGCAC
AGTTGCAACTCTTCGTGAAACCTATGGTGAAATGGCTGACTGCTGTGCAAAACAAGAAC
CTGAGAGAAATGAATGCTTCTTGCAACACAAAGATGACAACCCAAACCTCCCCCGATTG
GTGAGACCAGAGGTTGATGTGATGTGCACTGCTTTTCATGACAATGAAGAGACATTTTT
GAAAAAATACTTATATGAAATTGCCAGAAGACATCCTTACTTTTATGCCCCGGAACTCCT
TTTCTTTGCTAAAAGGTATAAAGCTGCTTTTACAGAATGTTGCCAAGCTGCTGATAAAGC
TGCCTGCCTGTTGCCAAAGCTCGATGAACTTCGGGATGAAGGGAAGGCCTCGAGTGCC
AAACAGAGACTCAAGTGTGCCAGTCTCCAAAAATTTGGAGAAAGAGCTTTCAAAGCATG
GGCAGTAGCTCGCCTGAGCCAGAGATTTCCCAAAGCTGAGTTTGCAGAAGTTTCCAAGT
TAGTGACAGATCTTACCAAAGTCCACACGGAATGCTGCCATGGAGATCTGCTTGAATGT
GCTGATGACAGGGCGGACCTTGCCAAGTATATCTGTGAAAATCAAGATTCGATCTCCAG
TAAACTGAAGGAATGCTGTGAAAAACCTCTGTTGGAAAAATCCCACTGCATTGCCGAAG
TGGAAAATGATGAGATGCCTGCTGACTTGCCTTCATTAGCTGCTGATTTTGTTGAAAGTA
AGGATGTTTGCAAAAACTATGCTGAGGCAAAGGATGTCTTCCTGGGCATGTTTTTGTAT
GAATATGCAAGAAGGCATCCTGATTACTCTGTCGTGCTGCTGCTGAGACTTGCCAAGAC
ATATGAAACCACTCTAGAGAAGTGCTGTGCCGCTGCAGATCCTCATGAATGCTATGCCA
AAGTGTTCGATGAATTTAAACCTCTTGTGGAAGAGCCTCAGAATTTAATCAAACAAAATT
GTGAGCTTTTTGAGCAGCTTGGAGAGTACAAATTCCAGAATGCGCTATTAGTTCGTTACA
CCAAGAAAGTACCCCAAGTGTCAACTCCAACTCTTGTAGAGGTCTCAAG**AAACCTAGGA**
AAAGTGGGCAGCAAATGTTGTAAACATCCTGAAGCAAAAAGAATGCCCTGTGCAGAA
GACTATCTATCCGTGGTCCTGAACCAGTTATGTGTGTTGCATGAGAAAACGCCAGTAA
GTGACAGAGTCACCAAATGCTGCACAGAATCCTTGGTGAACAGGCGACCATGCTTTTCA
GCTCTGGAAGTCGATGAAACATACGTTCCCAAAGAGTTTAATGCTGAAACATTCACCTTC
CATGCAGATATATGCACACTTTCTGAGAAGGAGAGACAAATCAAGAAACAAACTGCACT
TGTTGAGCTTGTGAAACACAAGCCCAAGGCAACAAAAGAGCAACTGAAAGCTGTTATGG
ATGATTTCGCAGCTTTTGTAGAGAAGTGCTGCAAGGCTGACGATAAGGAGACCTGCTTT
GCCGAGGAGGGTAAAAAACTTGTTGCTGCAAGTCAAGCTGCCTTAGGCTTATAA

2.13.2. Preparation of the standards

For quantification using a standard curve, a 10fold dilution series of PCR product ranging from 10^8 to 10^4 copies per reaction was used.

Knowing the size of the PCR-fragment that contains the gene of interest one can calculate the number of copies per µL using Avogadro's number. (N_A= 6.022 x 10^{23} 1 / mol)

molecular weight MW (Da = g / mol) = fragment length (bp) x 650 Da / bp
N_A (1 / mol) x MW (g / mol) = copies / g
copies / ng x conc (ng / µL) = **copies / µL**

The following formula displays the evaluation of the number of copies/µL:

$$Number\ of\ copies/\mu L = \frac{amplicon\ conc.(ng/\mu L) \times 10^{-9} \times N_A}{MW} \qquad \text{(Eq. 2)}$$

Starting from the number of copies in the PCR product a stock solution containing 10^9 copies / µL was established for each gene.
For the actin-standard the sequence was amplified from *P. pastoris* genomic DNA, for the heterologous proteins HSA and eGFP the sequences were amplified from a corresponding vector construct, using the primers described in table 5.

Table 7: 50µL PCR-mix for actin standard

µL	component
5	10X Buffer (Biotools)
4	Mg^{2+}
1.5	dNTPs (10 mM)
1.5	Primer fw (10 pM)
1.5	Primer rv (10 pM)
1	taq-polymerase (Biotools)
30.5	HQ-water
5	template (genomic DNA)

Table 8: 50µL PCR-mix for HSA and eGFP standard

µL	component
5	10X Buffer (Biotools)
4	Mg^{2+}
1.5	dNTPs (10 mM)
1.5	Primer fw (10 pM)
1.5	Primer rv (10 pM)
1	taq-polymerase (Biotools)
34.5	HQ-water
1	template (vector)

For amplification a Touchdown PCR was chosen The PCR was launched with a hotstart and one negative control, without template DNA, was included in each run.

Table 9: Touchdown PCR conditions for preparation of the standards

Step	Temperature	Time	Number of cycles
1	94°C	5 min	
2	94°C	30 sec	
3	60°C	30 sec	
4	72°C	20 sec	
5	go to 2 Tochdown -1°C/cycle		10x
6	94°C	30 sec	
7	50°C	30 sec	
8	72°C	30 sec	
9	go to 6		15x
10	72°C	10 min	
11	10°C	1 min	
12	END		

PCR products were analyzed by gel electrophoresis by a 2% agarose gel and then purified by gel-purification with the mi-Gel Extraction Kit (metabion).
The purified DNA was dissolved in 10 mM Tris/Cl with pH 8 and quantified using the NanoDrop ND-1000 Spectrophotometer.

2.13.3. qRT-PCR procedure

Quantitative real-time PCR (qRT-PCR) was accomplished using the Rotorgene 6000 system. PCR was performed in a 10 µL reaction mix using dsDNA intercalating dye SYBR Green for the detection of amplified DNA.

Table 10: 1x Mix for qRT-PCR

µL	component
0.25	Primer fw (10 pM)
0.25	Primer rv (10 pM)
5	Reaction buffer (2X SensiMix Plus SYBR)
3.5	HQ-water

A special designed metallic rack for the PCR-strips was filled with the tubes and cooled on ice prior to pipetting 9 µL of the mastermix into the small tubes.
1 µL of respectively the diluted sample (for transcript analysis the amount of 2.5 µg reverse transcribed RNA was used, for gene copy number determination 4 ng of genomic DNA served as template) or of the calibration standard was added. Each sample was detected in quadruple for reproducibility. To generate a standard curve each standard concentration was measured in double determination. As negative control served HQ-water instead of template, also in duplicates.
The tubes were then closed and placed into the rotor of the Rotor Gene™ PCR-machine. The continuous spinning with 400 rpm ensures that the samples stay at the bottom of the tubes, it also reduces air-bubbles, condensation and variation.

qRT-PCR was performed under following conditions:

Table 11: qRT-PCR conditions

Step	Temperature	Time	Number of cycles
Hotstart	95°C	10 min	1
Denaturation	95°C	15 sec	45
Primer-annealing	60°C	20 sec	
Elongation	72°C	15 sec	

The acquisition of the signal was detected at 60°C during each cycle.

2.13.4. Melt curve analysis

SYBR Green is a fluorogenic dye that exhibits little fluorescence when in solution, but emits a strong fluorescent signal upon binding to double-stranded DNA (dsDNA).
The advantages of SYBR Green are that it is inexpensive, easy to use, and sensitive. The disadvantage is that SYBR Green will bind to any dsDNA in the reaction, including primer-dimers, and other non-specific reaction products in a sequence-independent way (Lekanne Deprez et al. 2002). This can result in an overestimation of the target concentration. Therefore detection by SYBR Green requires extensive optimization and validation of the results through follow up assays such as a Melt Curve.
By increasing the temperature above the melt temperature (T_m) and simultaneous measuring of the fluorescence a graph can be obtained (Fig. 5). The nucleotide composition is an important factor for the T_m hence the specific signal from the target product can be identified (peak of the dF/dT curve).

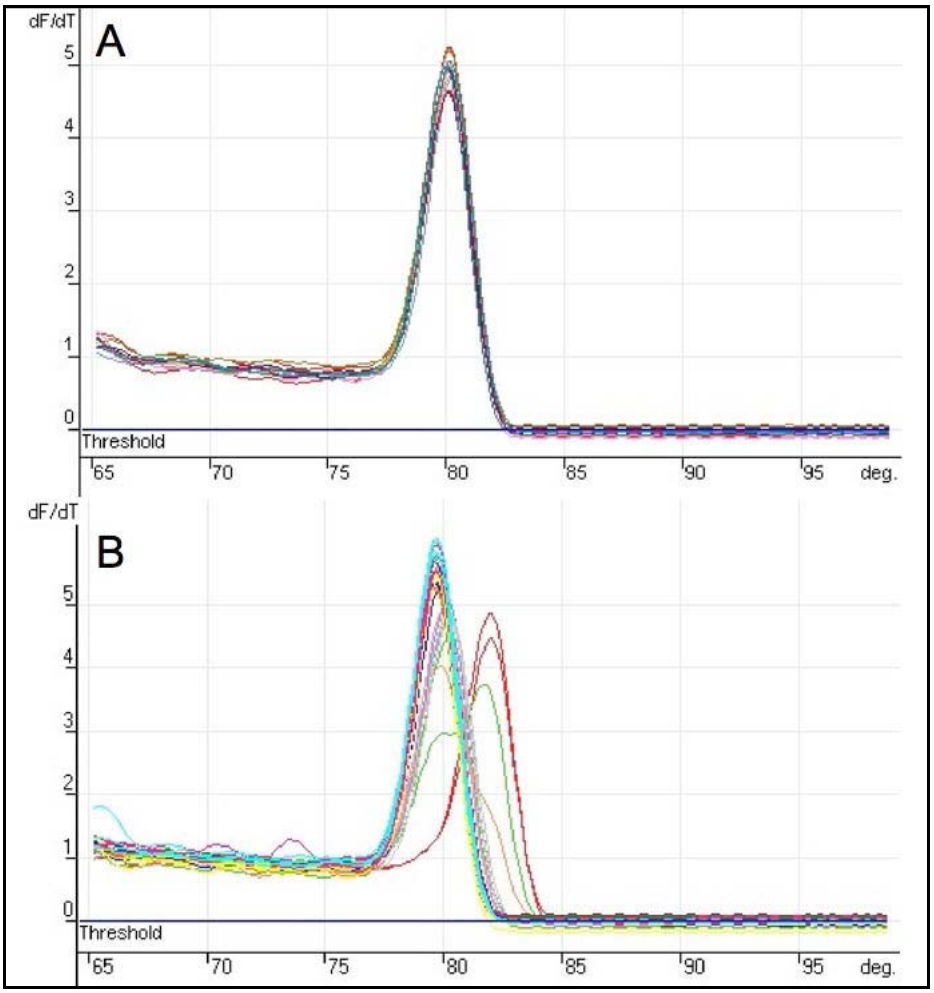


Figure 5: Melt curve of HSA standards and samples after qRT-PCR with HSA primers, plotted as the first derivative of the absorbance with respect to temperature. **A:** normal profile of melt curve analysis **B:** primer-dimers appear as small peaks before the T_m of the PCR product, additionally multiple peaks are observed, they may be the result of amplicon sequence heterogeneity or non-specific amplicons.

The generation of the dissociation curve started subsequently at 65°C for 1 second with a permanent measurement of fluorescence until 99°C, rising by 0.5°C each step and hold for 2 seconds. The characteristic melting peak can be distinguished from unspecific products melting at lower temperature (Bustin 2000). A correct curve consists only of one clearly separated peak (Fig. 5A), additional peaks are an indication for primer-dimers or unspecific PCR-products (Fig. 5B).

2.13.5. Analysis of the realtime PCR experiment

From the fluorescence values obtained through the realtime PCR analysis, the threshold cycles (Ct-values, see Fig. 6) were assessed using the Rotorgene 6000 system software (Version 1.7) and a standard curve was prepared.

The threshold was defined at a point in which the signal generated from a sample was significantly greater than the background fluorescence. The threshold cycle (Ct) is the number of PCR cycles required to exceed this threshold in fluorescence intensity (Ginzinger 2002).

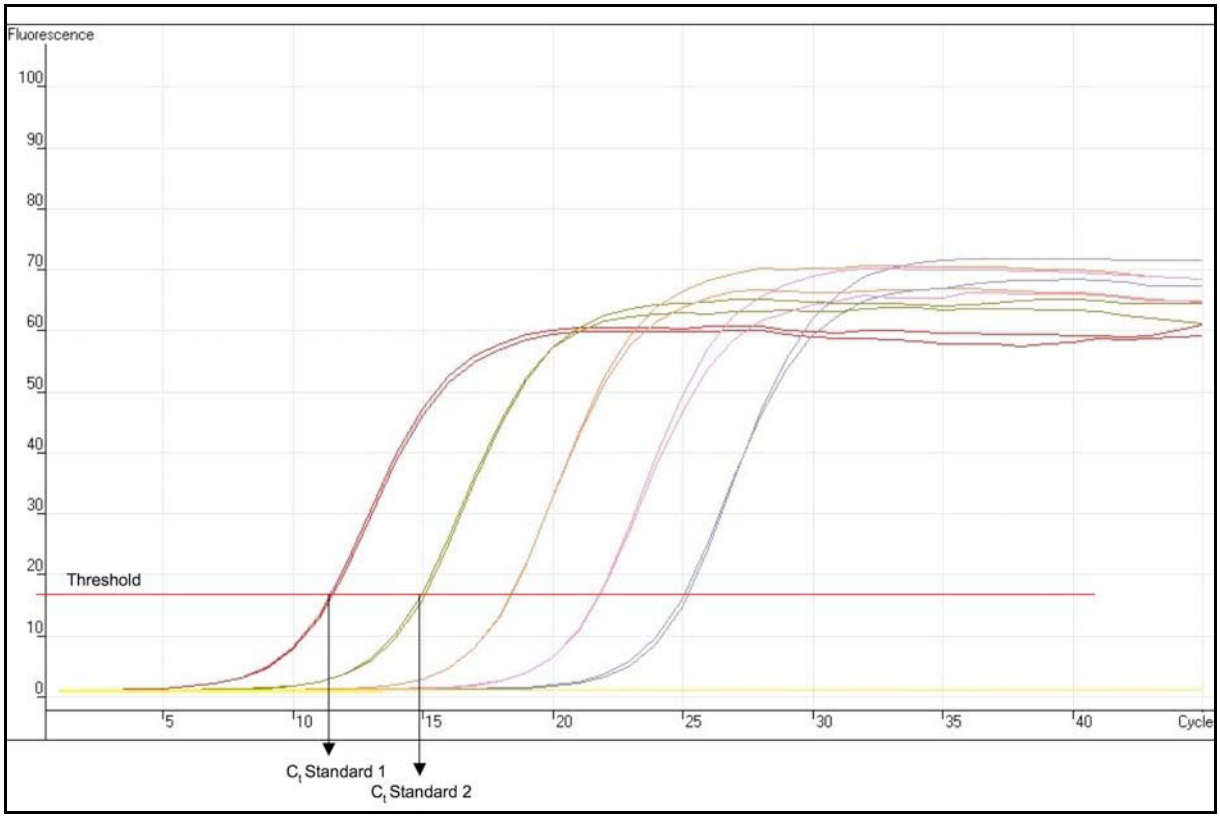


Figure 6: Amplification plot with threshold: fluorescent signals from 45 cycles of qRT-PCR of 10-fold dilution series of the actin PCR product using a linear scale for the y axis. The cycle threshold, depicted as red horizontal line, is the numerical value assigned for each run to calculate the C_t value for each amplification.

The standard curves and the separate curves of the quadruple samples were checked and outliers were discarded and not used in calculation.

Out of these data, exported to MS Excel for further analysis, the quantity of the samples could be determined.

Normalization of the data was achieved using the housekeeping gene actin as a reference. The results are therefore given as a ratio of the calculated target copies to the ones of the reference gene and are displayed in diagrams.

2.14. Southern blot analysis

Southern blotting is one of the cornerstones of DNA analysis since its first description by E.M. Southern in 1975. The advent of the Southern transfer and the associated hybridization techniques made it possible for the first time to obtain information about the physical organization of single and multicopy sequences in complex genome. Southern blotting comprises the separation of DNA fragments by agarose-gel electrophoresis and the transfer of the bands onto a nylon membrane followed by a hybridization-process with a specific probe and finally a detection step. Southern blotting is still a widely used technique that allows analysis of specific DNA sequences.

For non-radioactive labelling and detection the digoxygenin-antidigoxygenin system was used. The digoxygenin-antidigoxygenin system uses digoxigenin (DIG), a cardenolide steroid isolated from *Digitalis* plant.

The hybridized probe was then detected with a chemiluminescent substrate, an alkaline phosphatase conjugate of an anti-DIG-antibody, which was directed against the DIG-labelled DNA-probe, with the enzyme activating the chemiluminescent substrate CPD-Star (NEB) to produce a light signal. This chemiluminescent signal was then detected on a LumiImager (Boehringer Mannheim).

2.14.1. DIG-labelling of probes

DIG probe synthesis is carried out with Taq-DNA polymerase during polymerase chain reaction. It is able to incorporate Digoxigenin-11-dUTP instead of normal dNTPS – 1 every 10-20 nucleotides. The used PCR DIG Probe Synthesis Mix (Roche) features a 1:2 DIG-11-dUTP:dTTP ratio and generates probes that can be used to detect single copy genes on a Southern blot. Regarding the primer design for the DIG-labelling PCR it is advised that probes should have a minimum length of 400 bp.

Table 12: Primer for Southern Blot Probes

primer	sequence	Product length (bp)
actin_sonde_FW	GTTCCAGCCTTCTACGTTTCTATTCA	631
actin_sonde_BW	ACGGAGTACTTTCTTTCTGGTGGAG	
HSA_sonde_FW	CTCGAGTGCCAAACAGAGACTCAA	516
HSA_sonde_BW	GCGGCACAGCACTTCTCTAGAG	

In order to have enough material a PCR with non-labelled dNTPs is carried out in advance with a suitable plasmid as template. With this derived gel-purified PCR-product the labelling

PCR is done applying once again a Touchdown PCR with the same conditions as described in Table 9.

Table 13: 50µL mix for PCR

µL	component
1	template
5	10X Buffer (Biotools)
5	Mg^{2+}
1	dNTPs (10 mM)
1	Primer fw (10 pM)
1	Primer rv (10 pM)
1	Taq polymerase (Biotools)
35	HQ-water

Table 14: 50µL mix for labelling PCR

µL	component
1	PCR-product
5	10X buffer (Biotools)
5	Mg2+
5	DIG dNTPs (10 mM)
1	Primer fw (10 pM)
1	Primer rv (10 pM)
1	Taq polymerase (Biotools)
31	HQ-water

To check if the probes are now labelled, an electrophoresis was carried out. The labelled probe appears to be slightly heavier than the non-labelled probe.

After verification of a positive result by agarose gel analysis the DIG-labelled probe was applied to a preparative agarose gel electrophoresis and purified with the gel extraction kit (metabion) (see 2.4.4.2). The purified product was taken up in 50 µL of Tris/Cl 10 mM, pH 8.

The yield was determined using the NanoDrop ND-1000 Spectrophotometer. The labelled probe can be stored at -20°C until use.

2.14.2. Separation of DNA with agarose gel electrophoresis

The isolated genomic DNA was used after quantitation with the Nanodrop as well as with the Hoechst-dye 33258. The chosen samples were diluted to both 30 ng / µL and 45 ng / µL because 300 ng or 450 ng gDNA respectively (10 µL of the diluted solution) appeared to be suitable for the Southern Blot analysis (300 ng were used for the Actin, and 450 ng for the HSA southern blot analysis).

The desired fragment of genomic DNA has been cut with suitable restriction enzymes to create a characteristic band pattern. For the HSA fragment the enzymes AgeI and SbfI from NEB were used – creating a desired band of 2174 bp, and for the Actin fragment AdeI and Alw44I from Fermentas with a band of 1021 bp.

In order to create a standard for the southern blot analysis aliquots of cut and purified fragments corresponding to 0.5 to 10 copies of the two genes were used.

As templates served for the HSA fragment a corresponding pPuzzle plasmid, and for the Actin fragment genomic DNA from a X-33-wild type strain.

The required amount for a solution containing one copy of the target gene was calculated as follows: The size of the *P. pastoris* genome is approximately 10^7 bp, which was equated to 100 %. Accordingly were the standard-fragments HSA: $2.174*10^{-2}$ and Actin: $1.021*10^{-2}$ % of the genome size.

Referring to the applied amount of genomic DNA in the Southern Blot of 300 ng or 450 ng this corresponds to 65.22 pg of HSA-fragment and 45.95 pg of Actin-fragment for one gene copy.

For the restriction digest the standard solutions with the theoretical gene copies 0.5, 1, 2, 3, 4, 6 and 10 were made by using the respective volume of template.

Table 15: Restriction digest for the Actin Southern Blot

µL	component
10	diluted gDNA (45 ng / µL)
1	AdeI
1	Alw44I
1	Buffer G

Table 16: Restriction digest for the HSA Southern Blot

µL	component
10	diluted gDNA (30 ng / µL)
1	AgeI
1	SbfI
1	Buffer 4

The components were mixed with the gDNA and incubated at 37°C for at least 2 h to ensure a complete digest of the DNA.

The digested samples and standards were mixed with 2 µL loading dye (6x, MBI Fermentas). The DNA fragments were separated by agarose gel electrophoresis (0.7 %, 100 mL) overnight at 30 V (at least 2 h at 90 V). Additionally, two different marker were applied: 5 µL of the MassRuler DNA Ladder Mix (Fermentas) to check if the electrophoresis worked properly, and 7 µL of a DIG-labeled DNA molecular weight marker (0.01 µg / mL, Roche) to be able to verify that the chemiluminescent detection worked as well.

After taking a quick photo of the gel, exposure times had to be as short as possible to avoid modifications of the DNA, it was incubated in alkaline transfer buffer under shaking conditions at room temperature for 30 min. This step is necessary for denaturation of the DNA fragments to open the double strand, so that the complementary probe can attach.

2.14.3. Materials and set-Up of the Southern blot

- Alkaline transfer buffer: 0.4 M NaOH
  1 M NaCl
- Neutralizing buffer: 0.5 M Tris*HCl with pH 7.2 adjusted with 4 N HCl
  1 M NaCl

- 20x SSC stock solution: 3 M NaCl
  0.3 M NaAc*3H_2O
  pH 7.4 was adjusted with 4 N HCl, the solution had to be autoclaved
- 2x SSC, 0.1% SDS: 50 mL 20x SSC
  2.5 mL 20% SDS
  447.5 mL AD
- 0.5x SSC, 0.1% SDS: 12.5 mL 20x SSC
  2.5 mL 20 % SDS
  485 mL AD
- Maleic acid buffer: 0.1 M maleic acid (disodiumsalt – dihydrate)
  0.15 M NaCl
  pH was adjusted to 7.5 with HCl, the buffer had to be autoclaved
- Washing buffer: 0.3% Tween20 were added to Maleic acid buffer
- Blocking Reagent stock solution: 10 g Blocking Reagent were added to 100 mL Maleic acid buffer and stirred at 60°C for approximately 1 hour.
- Blocking buffer: Blocking Reagent stock solution diluted 1:10 with Maleic acid buffer
- 10 % N-lauroylsarcisine stock solution: 10 % in AD, stirred for 1 h and filtered with a Millex-GP membrane.
- 1M Sodium-Phosphate buffer, pH 7.0: 1 M NaH_2PO_4
  1 M Na_2HPO_4*$12H_20$
  NaH_2PO_4 was provided and the pH was adjusted with Na_2HPO_4
- High SDS concentration hybridization buffer: 250 mL 100% formamide
  124.5 mL 20x SSC
  100 mL blocking buffer
  25 mL 1 M sodium-phosphate, pH 7.0
  5 mL 10 % N-lauroylsarcisine
  The solution was poured onto 35 g SDS and heated while stirring to dissolve the SDS.
  The solution was then aliquoted in 40 mL portions and stored at -20°C.
  Before use it has to be heated to 65°C
- Detection buffer: 100 mM TrisHCl pH 9.5
  100 mM NaCl
- Substrate: CDP-Star (NEB)

- Set-Up:

6 Whatman paper, the nylon membrane (Nytron SPC 0.45 µm Nylon Transfer Membrane, Whatman) and about 10 cm Kleenex paper (all cut in size of the gel) were prepared. The blots were assembled in following order in an electrophoresis chamber:

Bottle of water for weight and pressure
~10 cm Kleenex paper
3 x Whatman paper
nylon membrane
gel (0.7 %) upside down
3 x Whatman paper
wick made of Whatman paper sopped in alkaline transfer buffer

The blotting was performed by upward capillary transfer and occurred for approximately 5 h, occasionally the Kleenex was changed to stabilize the blot.
After disassembly of the blot, the slots of the gel have been carefully marked on the membrane, the blot was incubated in neutralizing buffer with shaking for 10 min at room temperature. To fix the DNA bands to the membrane, the blot was baked on Whatman paper in a suitable oven for 30 min at 120°C. The baked membrane can then be stored dry at 4°C or be prehybridized immediately.

2.14.4. Prehybridization and hybridization

For prehybridization the membrane was incubated in 40 mL of pre-warmed (42°C) high SDS buffer while rotating for 2 hours at 42°C in the hybridization oven. Prehybridization decreases the background, because non-specific nucleic acid binding sites on the membrane become blocked.
Meanwhile the DIG labelled probe was heated for 5-10 minutes at 99°C and put immediately on ice to avoid rehybridization of the single stranded probe. The optimal concentration of the probe should be between 5 to 25 ng / mL buffer according to DIG-systems user guide (Boehringer), in this case 15 ng / mL have been applied.
After two hours of incubation the prehybridization solution has been discarded (storage at -20°C) and high SDS buffer containing the labelled probe was added to the blot. Hybridization was allowed to occur overnight at 42°C under constant shaking.
The hybridization buffer with probe was then saved and stored for next use at -20°C, because it could be re-used for 4 to 5 times.

2.14.5. Washing and detection of the Southern blot

As the labelled probes can also hybridize unspecifically to sequences that bear homology, these mismatches have to be removed by certain washing steps. The stringency of the washing can be adjusted to the specific blot by varying the salt concentrations and the temperature. Best results have been achieved under the following conditions:

After removal of the hybridization solution, the blot was washed in 2x SSC, 0.1 % SDS twice for 5 minutes respectively at room temperature to clean the blot from unbound probe, which would otherwise lead to high background.

The blot was then put in the hybridization oven at 68°C and washed again with 0.5x SSC, 0.1 % SDS twice for 5 minutes. The following steps have all been prepared at room temperature. The blot was equilibrated in a special washing buffer (DIG-System) for 1 minute, and then the membrane was blocked in 50 mL of blocking solution by gently agitating for 30 min. The Anti-DIG-Fab-AP antibody was diluted 1:10,000 in 50 mL of fresh blocking solution and the membrane was incubated for another 30 min with gentle shaking. After having discarded the antibody solution, the blot was washed twice in the special washing buffer (DIG-system) for 15 min. Before applying the detection substrate, the blot was equilibrated in detection buffer for 2 min. The chemiluminescent substrate CDP-Star, which has been diluted 1:100 in detection buffer, was then applied to the blot for 5 min of incubation, always covering solution with tin foil to avoid light. The detection buffer with CDP-Star was also saved and stored for re-use at 4°C. The blot was weld in a plastic bag to prevent drying and detected with the Lumi-Imager (Boehringer Mannheim) by chemiluminescence.

3. EXPERIMENTS AND RESULTS

The following figure summarizes the work performed in the present study.

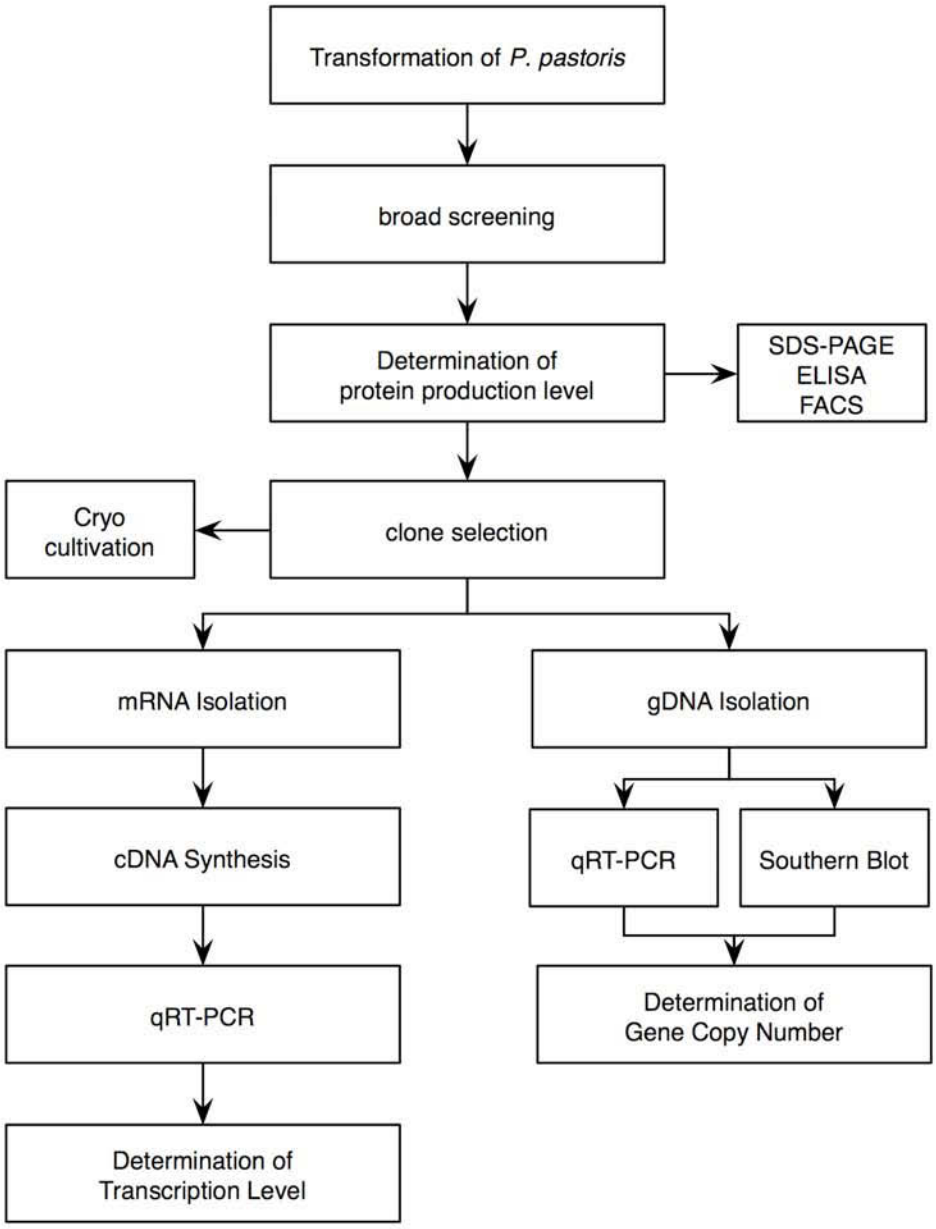


Figure 7: Experimental set-up

P. pastoris clones expressing either an intracellular reporter (enhanced green fluorescent protein, eGFP) or a secreted heterologous protein (human serum albumin, HSA) were cultivated under the control of five different promoter sequences and the amount of recombinant reporter protein was quantified. Transcription levels and gene copy number of selected clones were analyzed using quantitative real-time PCR (qRT-PCR).

3.1. Expression of the model proteins under the control of different promoters

In the present study the following promoter sequences were analyzed, summarized in Table 17. Gene name and function are given according to literature data or the respective *S. cerevisiae* homologues (www.yeastgenome.org).

Table 17: Gene name and function of the analyzed promoter sequences

gene name	protein function
GAP	glyceraldehyde-3-phosphate dehydrogenase
PET9	major ADP/ATP carrier of the mitochondrial inner membrane
TEF1	translation elongation factor EF-1α
ENO1	Enolase I
THI11	protein involved in synthesis of the thiamine precursor hydroxymethylpyrimidine

These promoter sequences were selected due to their interesting regulatory properties in previous experiments (see Fig.1), with P_{GAP} as control.

3.1.1. Transformation of *P. pastoris*

Vectors of the used pPUZZLE series are based on a modular design, allowing an easy exchange of almost all vector components by restriction enzyme digest. To gain information on the properties and activities of the different promoters and to avoid chromosome organization effects of different integration loci on reporter gene expression levels the expression vectors of each construct were linearized with AscI in the 3'-AOX1 integration region and transformed into *P. pastoris* by electroporation.

Table 18: Used strains and vector combinations

strain	plasmid	selection marker	code	colour scheme
X-33	pPUZZLE_PGAP_eGFP	Zeocin	gap	red
X-33	pPUZZLE_PPET_eGFP	Zeocin	pet	blue
X-33	pPUZZLE_PTEF_eGFP	Zeocin	tef	green
X-33	pPUZZLE_PENO_eGFP	Zeocin	eno	black
SMD1168H	pPUZZLE_PGAP_HSA	Zeocin	GAP	red
SMD1168H	pPUZZLE_PPET_HSA	Zeocin	PET	blue
SMD1168H	pPUZZLE_PTEF_HSA	Zeocin	TEF	green
SMD1168H	pPUZZLE_PENO_HSA	Zeocin	ENO	black

In case of the intracellular reporter eGFP the *P. pastoris* X-33 wild type was used as host strain, for the secreted expression of HSA *P. pastoris* SMD1168H was utilized, as the protease A deficient strain has been proven to be beneficial for HSA stability previously (unpublished data).

For better comparability all *GAP* promoter constructs appear labelled in red, blue for P_{PET9}, green for P_{TEF1} and black for the *ENO1* promoter constructs.

3.1.2. Analysis of promoter activity in *P. pastoris* - Screening

To test the ability of the selected promoter sequences for recombinant protein production a first round of screening was carried out.

3.1.2.1. Expression of the intracellular model protein eGFP

For each promoter construct 10 individual single clones were cultivated and used to inoculate a pre-culture on YPD and grown at 28°C, 180rpm for 24 h. Aliquots (OD_{600} = 0.1) were transferred to shake flask cultures to test for promoter activity. Glucose was added every 12h to a final concentration of 0.5 % (w/v). Fluorescence of eGFP was quantified using flow cytometry as described in 2.9.1.1.

The results of cells grown on YPD medium for 48 h are represented in figure 8.

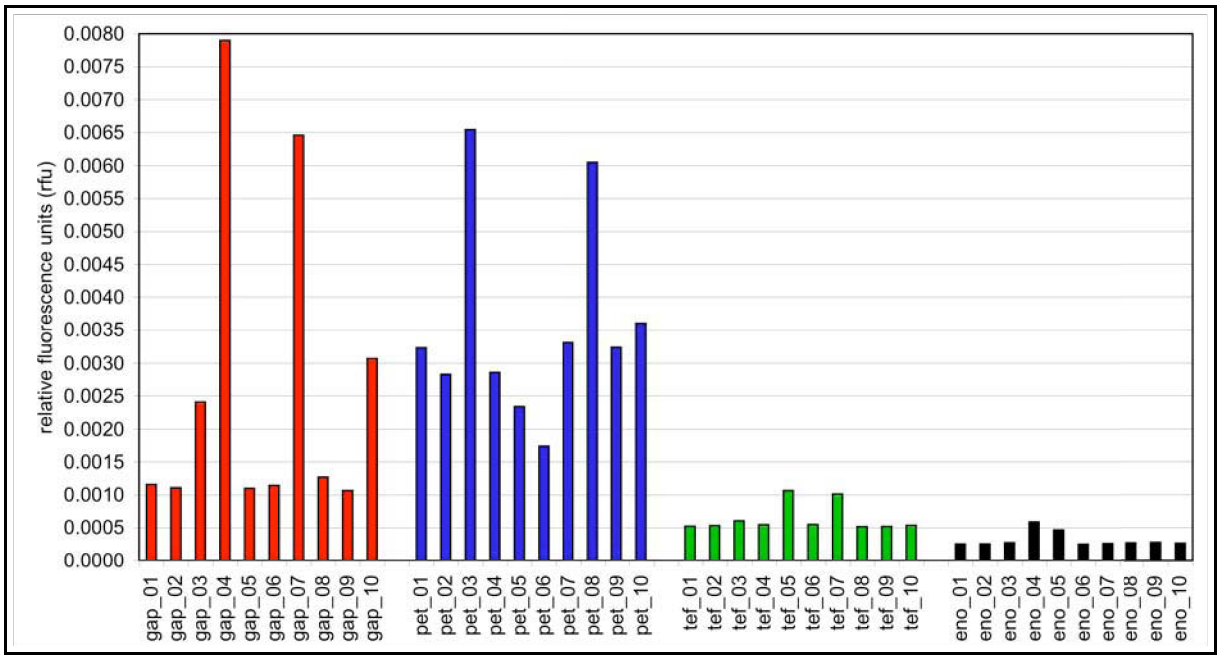


Figure 8: Broad screening – expression levels of eGFP – Promoter driven expression on complex glucose medium (YPD) after 48 h cultivation. eGFP was quantified using flow cytometry and correlated to the cell size according to equation1. Depicted data represent the results of one out of two independent experiments.

All tested clones exhibited eGFP expression levels, ranging from a very low expression level in case of the *ENO1* promoter to a remarkably high level in case of the *PET9* promoter. As can be seen in figure 8, the expression of eGFP is very heterogenous within a clone family leading to the question of gene copy number effects. A second screening was done under the same conditions and gave reproducible results. On this account for each construct four clones were chosen for further analysis (Fig. 9).

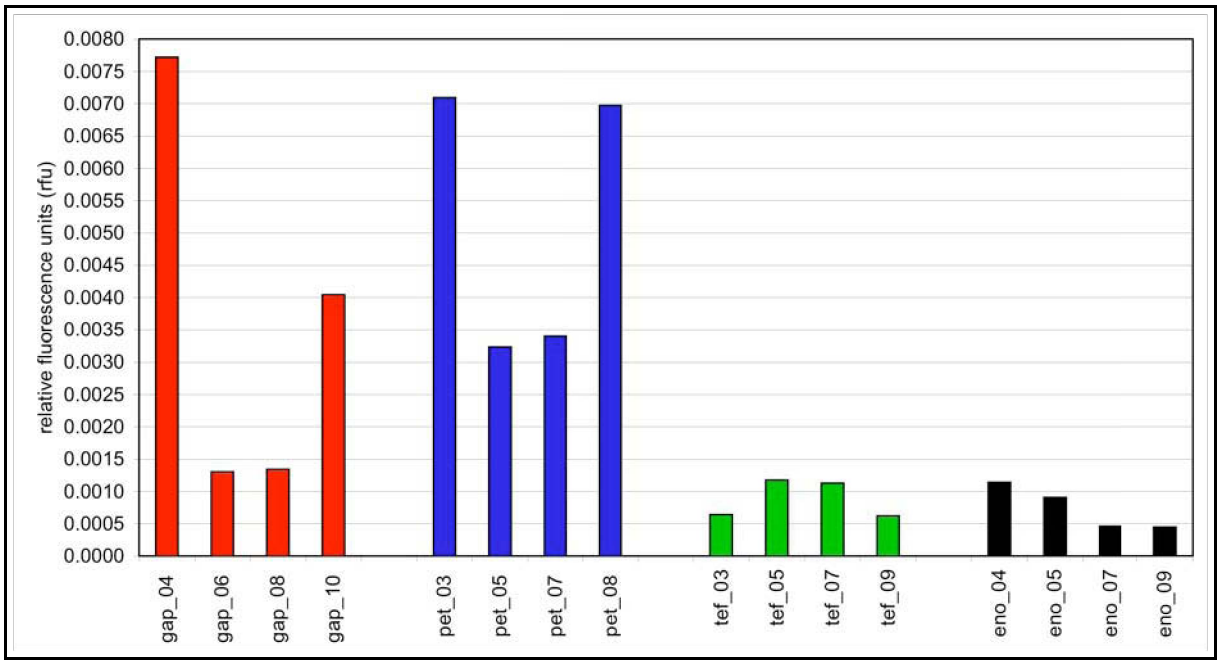


Figure 9: Selected clones – expression levels of eGFP – Promoter driven expression on YPD medium after 48 h cultivation. eGFP was quantified using flow cytometry, depicted data represent the results of one out of two independent experiments.

Purposefully, two clones with rather low expression levels and two clones with high expression levels were chosen for further analysis, as different gene copy numbers present in the selected clones may give more insight on the strength of a promoter construct.

3.1.2.2. Expression of the secreted model protein HSA

The potential of the selected promoters to drive expression of secreted heterologous proteins was analyzed using HSA as reporter. The amount of secreted HSA in the supernatant was taken as value for promoter activity. As secretion leader the native leader of HSA was used.

At first – for checking if the correct product was secreted into the supernatant – a SDS-PAGE analysis was performed. After 48 h of cultivation at 28°C in BM medium with glucose feeding, the yeast cultures were centrifuged, and the undiluted supernatant was used as sample.

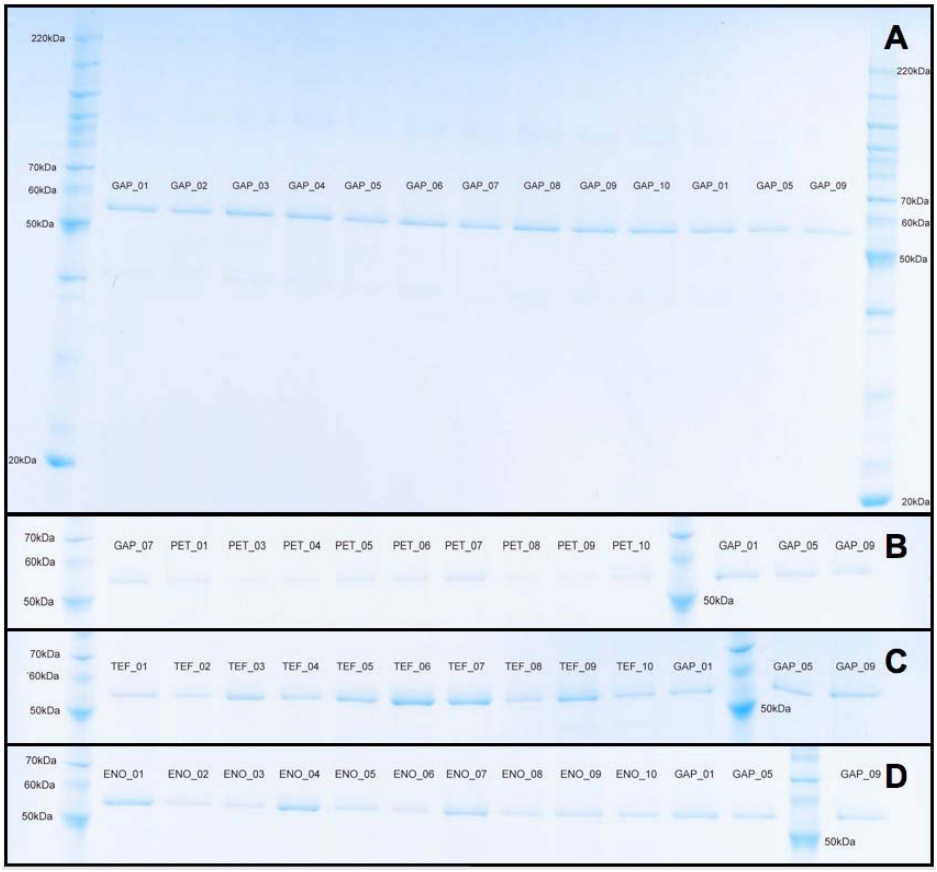


Figure 10: Characterization of expression levels of HSA – SDS-PAGE analysis was performed using undiluted supernatant after 48 h shake flask cultivation on glucose medium (BM) **A:** *GAP* **B:** *PET9* **C:** *TEF1* and **D:** *ENO1* promoter driven expression

All transformed clones show a prominent protein band with the correct size (66.5 kDa). Other protein bands were only scarcely visible (see Figure 10A), showing that there was no proteolytic degradation of the protein.

Interestingly, clones under control of the *PET9* promoter (displayed in Figure 10B) gave only a very faint protein band, whereas the *TEF1* clones (Figure 10C) present very intense protein bands. Clones under control of the P_{ENO1} showed higher intensities as expected from the eGFP expression but they also appeared very heterogenous.

To gain information on the amount of expressed and secreted HSA, a specific HSA-ELISA (enzyme-linked immunosorbent assay) was carried out. The supernatant was diluted accordingly to the intensity of the SDS-PAGE gel bands 1:50 or 1:100 in order to fit to the calibration curve.

Table 19: Screening of HSA expressing clones

clone #	HSA conc. (µg/mL)	yeast wet mass (mg/mL)	clone #	HSA conc. (µg/mL)	yeast wet mass (mg/mL)
GAP_01	8.75	48.50	**TEF_01**	8.96	48.00
GAP_02	4.98	46.30	**TEF_02**	6.22	48.40
GAP_03	1.09	47.70	**TEF_03**	14.35	47.40
GAP_04	9.30	47.40	**TEF_04**	5.73	49.20
GAP_05	9.27	48.50	**TEF_05**	14.07	48.20
GAP_06	7.03	45.00	**TEF_06**	25.31	47.80
GAP_07	6.96	44.70	**TEF_07**	21.37	49.00
GAP_08	6.27	46.60	**TEF_08**	7.02	48.10
GAP_09	1.69	48.10	**TEF_09**	11.63	48.80
GAP_10	1.20	47.40	**TEF_10**	6.50	46.70
PET_01	1.18	50.10	**ENO_01**	14.95	47.00
PET_02	-	-	**ENO_02**	2.99	48.10
PET_03	0.02	49.10	**ENO_03**	4.17	48.30
PET_04	1.37	48.60	**ENO_04**	13.03	48.80
PET_05	2.52	41.80	**ENO_05**	5.22	48.10
PET_06	8.02	50.60	**ENO_06**	4.38	48.60
PET_07	9.16	51.60	**ENO_07**	7.63	48.60
PET_08	1.36	48.10	**ENO_08**	3.27	49.80
PET_09	0.98	51.20	**ENO_09**	5.09	46.80
PET_10	7.20	52.00	**ENO_10**	3.93	50.00

With these measured data the product to biomass yield $Y_{P/X}$ can be calculated as follows:

$$Y_{P/X} = \frac{mg/L\ product}{g/L\ yeast\ wet\ mass} \qquad \text{(Eq. 3)}$$

The results of the promoter dependent HSA product yields after 48 h of cultivation on glucose medium are depicted in Figure 11.

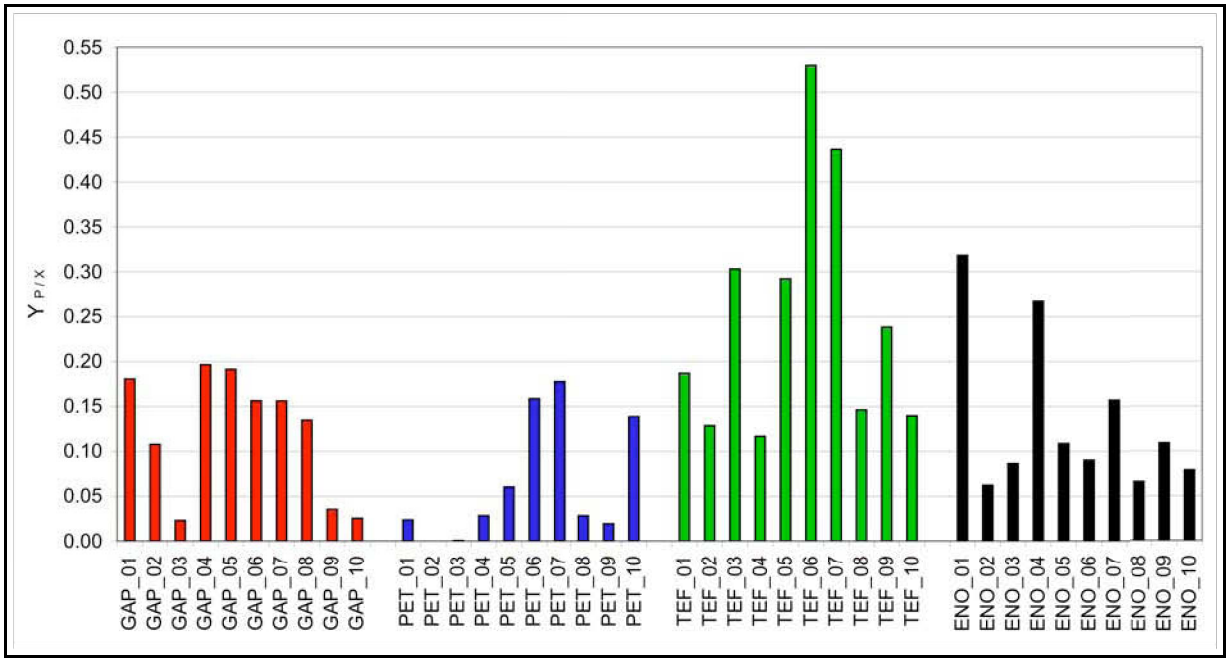


Figure 11: Product yield to biomass $Y_{P/X}$ of the HSA expressing clones – HSA yield was determined using equation 3 after 48 h of cultivation on glucose medium (BM). Depicted data represent the results of one out of two independent experiments.

Also for the HSA expressing clones reproducible results in a second independent screening were obtained. From each promoter construct again four clones were selected and shake flask cultivated.

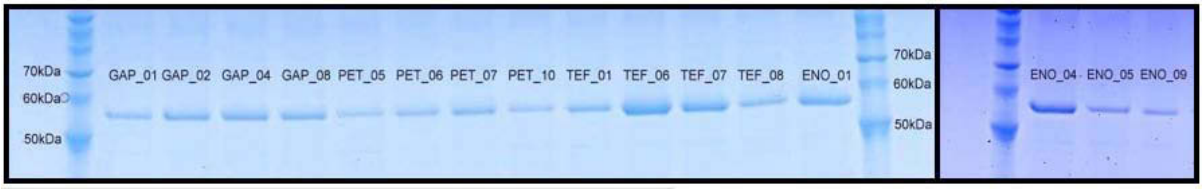


Figure 12: Expression levels of selected HSA expressing clones after 48 h cultivation in BM – SDS-PAGE analysis of undiluted culture supernatant stained with Coomassie Blue.

The samples taken after 48 h of cultivation show the same intensity pattern as in the first screening round (Fig. 12).

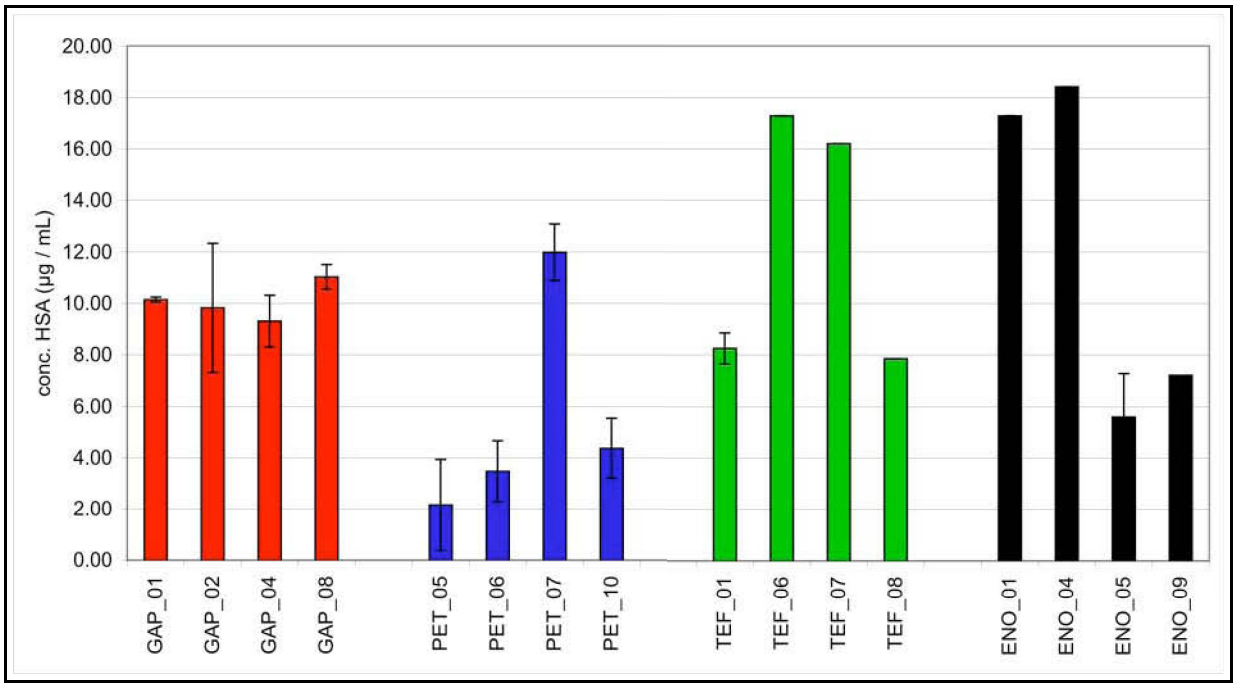


Figure 13: Expression levels of selected HSA secreting clones – amount of HSA was determined using diluted culture supernatant (1:50) with the HSA-ELISA after 48 h of cultivation on glucose medium (BM).

The data of the HSA experiment differ from the results obtained with eGFP. While eGFP expression under control of the P_{TEF1} and the P_{ENO1} gave low fluorescence levels, in case of HSA secretion similar or even higher amounts compared to the *GAP* promoter were observed. This effect can be explained by subsequent accumulation of the secreted protein in the supernatant over 48 h cultivation period instead of “actual picture” in case of eGFP expression.

Interestingly, a different behaviour was seen for expression under control of the *PET9* promoter. Contrary to high expression of eGFP, only low promoter activity was measured in case of HSA (Fig. 12 and 13).

These results were supported by analysis of other experiments (Stadlmayr et al. 2010a). In all analyzed cases expression under the P_{PET9} resulted in high expression levels of intracellular eGFP and low expression levels of the secreted HSA as is delineated in figure 14 in detail. But as figure 14 is also highlighting HSA secretion as well as eGFP expression show a very remarkable intersubject variability in expression levels, further highlighting the question of gene copy number effects.

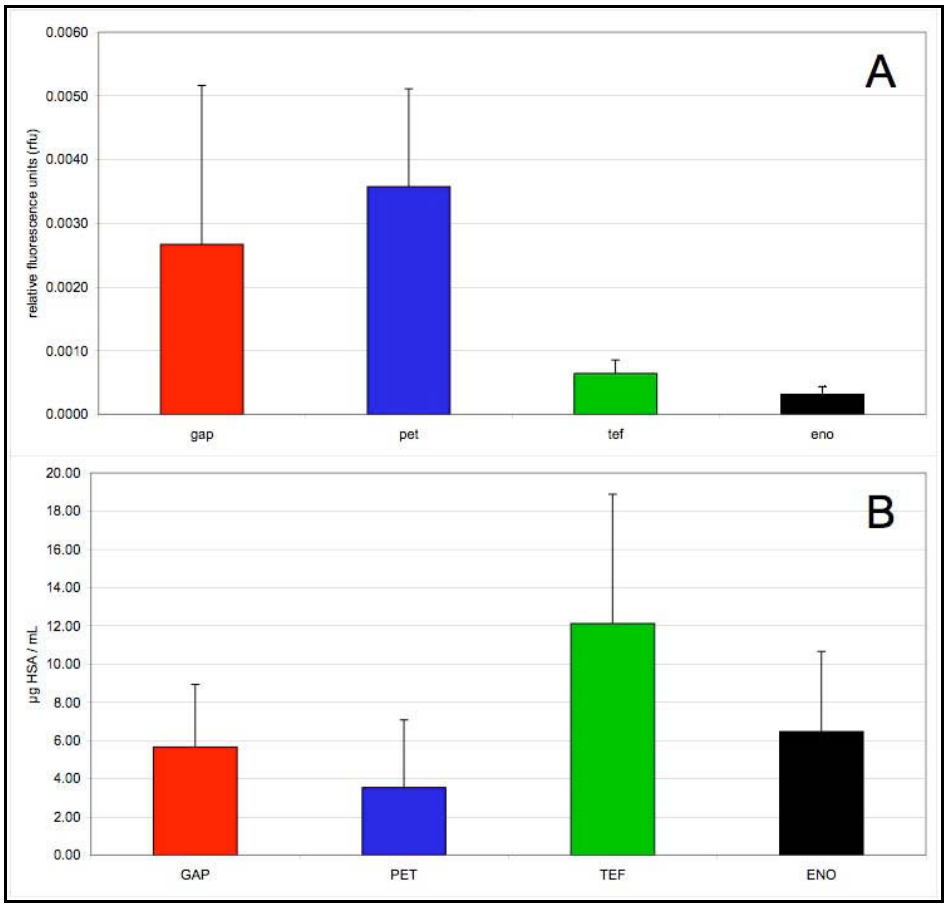


Figure 14: Mean value of expression of 10 clones each – A: FACS analysis data: eGFP was quantified using flow cytometry and correlated to the cell size according to equation 1**; B: HSA-ELISA data:** HSA was quantified using diluted supernatant.

3.2. Characterization of selected promoter constructs

For further characterization of these different promoters, the selected constructs were investigated for promoter activity not only indirectly by measurement of the amount of gene product expressed but also directly by determination of the transcript level using quantitative real time PCR.

As gene copy number has a strong impact on protein production, and even a weak promoter can give high protein levels with a high gene dosage, purposefully clones with expected different gene copy number were chosen. From each promoter-reporter construct 4 clones – 2 with high and 2 with low expression levels – were selected and shake flask cultivated (BM for HSA, YP for eGFP expressing clones).

As RNA isolation yields best results in exponentially growing cells, the next step was the determination of expressed protein after 6 h of shake flask cultivation.

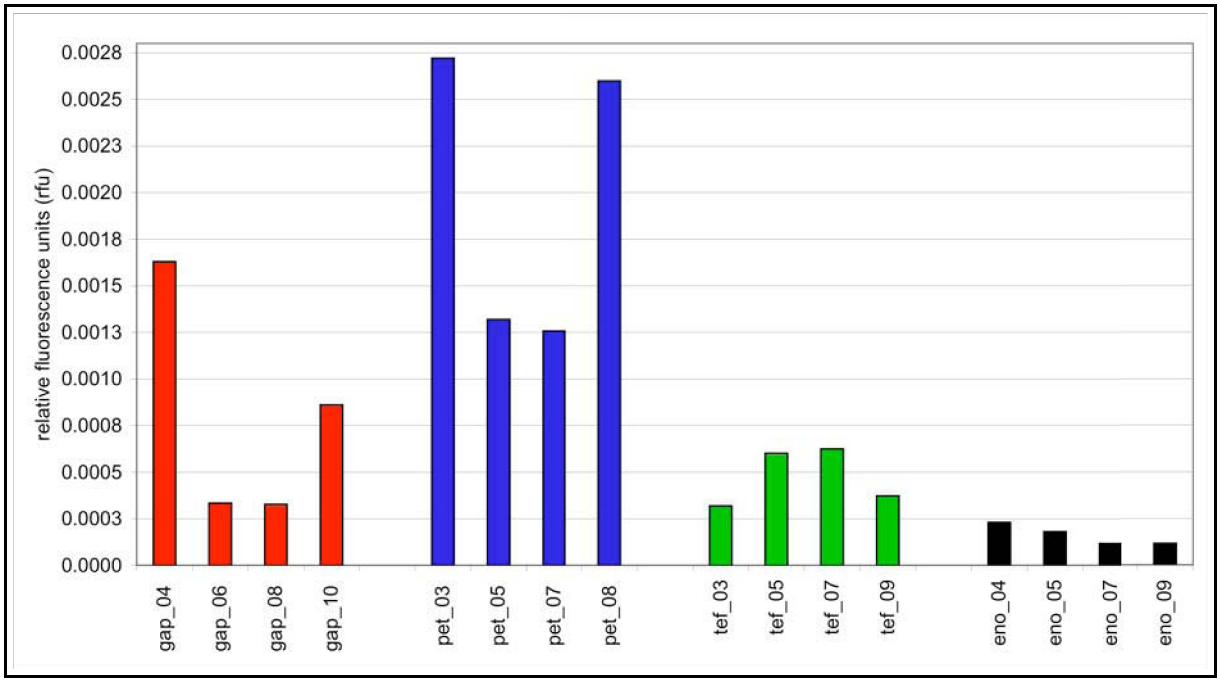


Figure 15: Expression levels of selected eGFP clones after 6 h cultivation – the selected clones were cultivated in YPD for 6 h to obtain exponentially growing cells, again eGFP was quantified using flow cytometry and correlated to the cell size according to equation 1.

Again flow cytometry was applied to measure eGFP expression (Fig. 15) and the HSA-ELISA to determine the HSA expression level in exponentially growing cells (Fig. 17). After 6 h of shake flask cultivation, the expression levels were very low as expected, only the clones under control of the P_{PET9} gave outstandingly high eGFP expression levels in exponentially growing cells.

To exclude the possibility that the different promoters might have an impact on the fluorescence properties of eGFP, an alternative method for eGFP quantitation was chosen. An eGFP-ELISA was performed after cell-disruption using shear forces when vortexing with glass beads to determine the amount of eGFP.

Table 20: comparison of FACS analysis and ELISA

clone #	rfu	GFP conc. (µg/mL)	clone #	rfu	GFP conc. (µg/mL)
gap_04	1.63E-03	0.83	**tef_03**	3.19E-04	0.09
gap_06	3.34E-04	0.14	**tef_05**	6.01E-04	0.29
gap_08	3.27E-04	0.14	**tef_07**	6.24E-04	0.24
gap_10	8.60E-04	0.40	**tef_09**	3.72E-04	0.13
pet_03	2.72E-03	1.63	**eno_04**	2.30E-04	0.07
pet_05	1.32E-03	0.76	**eno_05**	1.80E-04	0.06
pet_07	1.26E-03	0.61	**eno_07**	1.16E-04	0.03
pet_08	2.60E-03	1.10	**eno_09**	1.15E-04	0.03

The results of the eGFP-ELISA were highly consistent to the data obtained with flow cytometric analysis. (correlation coefficient R^2 = 0.9502)

The selected HSA expressing clones were also shake flask cultivated with the same conditions as described previously. The first sample was taken after 6 h to gain information of the secretion level of exponentially growing cells and a second sample after 16 h of cultivation.

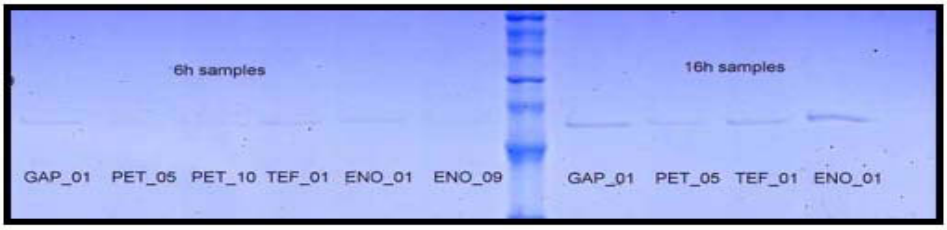


Figure 16: Expression level of HSA after 6 h and 16 h cultivation in BM – SDS-PAGE analysis with undiluted culture supernatant

Unsurprisingly after 6 h of shake flask cultivation nearly no product could be detected in the supernatant with a SDS-PAGE but after 16 h a distinct protein band was visible (Fig. 16).

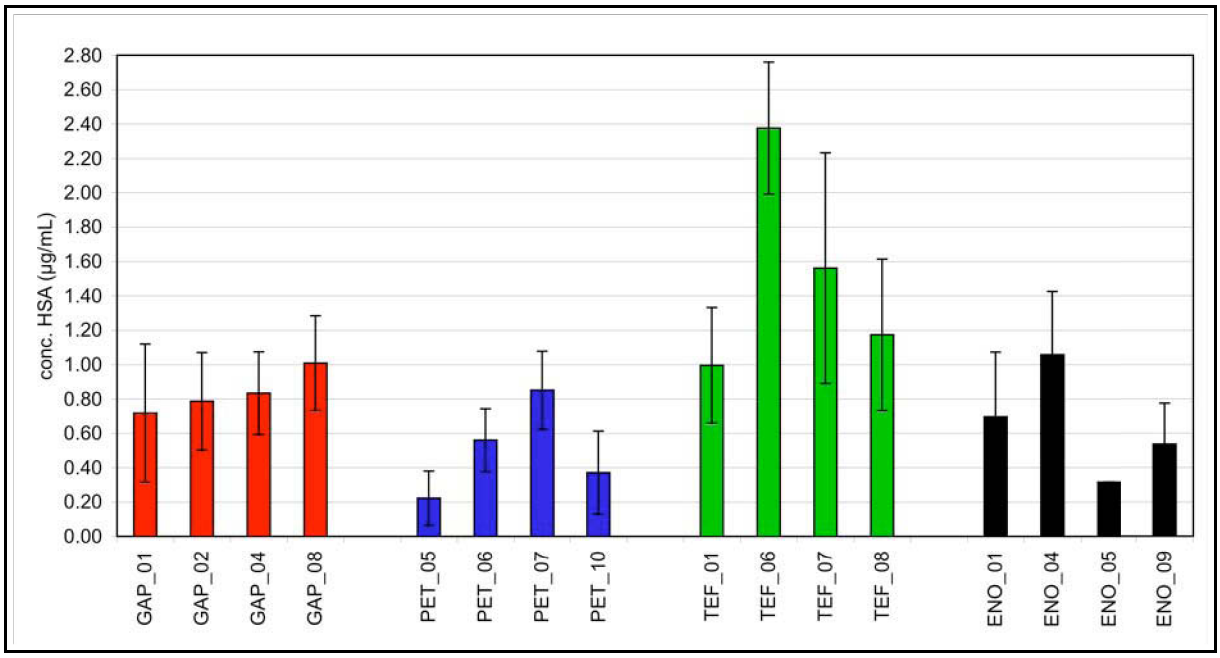


Figure 17: HSA amount after 6 h of shake flask cultivation in BM – diluted culture supernatant (1:2) was used as sample for ELISA.

Figure 17 demonstrates the low HSA expression levels after 6 h of shake flask cultivation. Only the clones under control of the P_{TEF1} gave high HSA expression levels in exponentially growing cells. Again secretion driven by the *PET9* promoter remained low. Despite the purposeful clone selection with different expression levels, the four *GAP* clones appear homogenous in the exponential growth phase.

Total RNA was isolated immediately after quenching of the cells, harvested in the exponential growth phase.

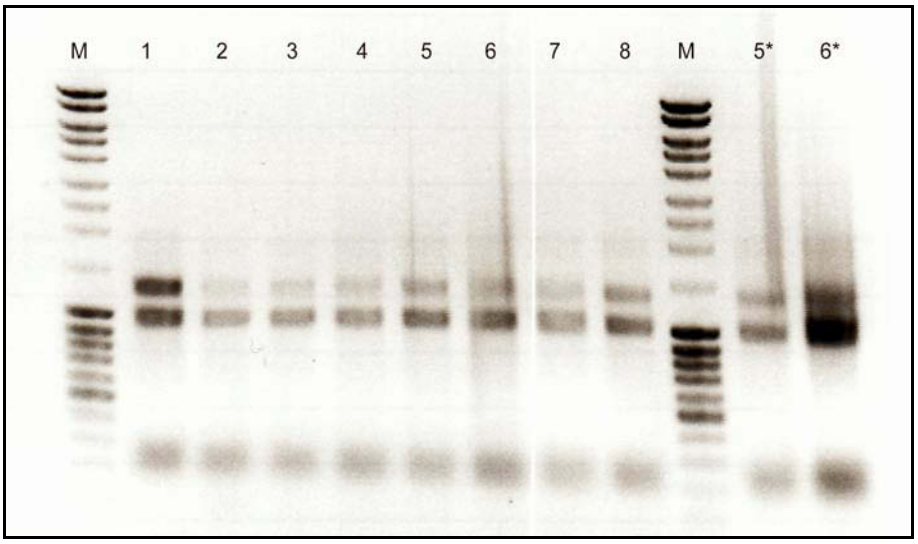


Figure 18: RNA quality check on a 1% agarose gel – M: molecular weight marker; lanes 1-8 isolated total RNA with TRI Reagent after DNA-free treatment; lanes 5* and 6* prior to DNA-free treatment.

As RNA quality has a high impact on reverse transcription and therefore also on the following PCR reactions (Fleige et al. 2006), it was quickly checked with the Nanodrop ND-1000 and also visually checked by agarose gel electrophoresis. Gel electrophoresis can show genomic DNA contamination and RNA decay. Genomic DNA will be visible as a tight DNA band of high molecular weight, ribosomal RNA (18S and 26S rRNA, small band of 5 and 5.8S rRNA) as two sharp bands half way down the gel and mRNA appears as smear in the background. Figure 18 shows undegraded and after DNA-free treatment non-contaminated RNA.
Samples were considered for further analysis only when the RNA was of high purity as indicated by the ultraviolet absorbance spectrum and appeared undegraded on the agarose gel. After quantification with the Nanodrop spectrophotometer (A_{260} for RNA: 40 mg/L RNA has a A_{260} of 1) First-Strand-cDNA Synthesis was accomplished using the Superscript system from Invitrogen.

Additionally, the promoter of the thiamine biosynthesis gene P_{THI11} , which showed interesting features in the work of Stadlmayr (2010a), was included into ongoing experiments. As the *THI11* promoter is only active in thiamine free conditions, the P_{THI11} clones were cultivated using a synthetic medium consisting only of salts, biotin and glucose.

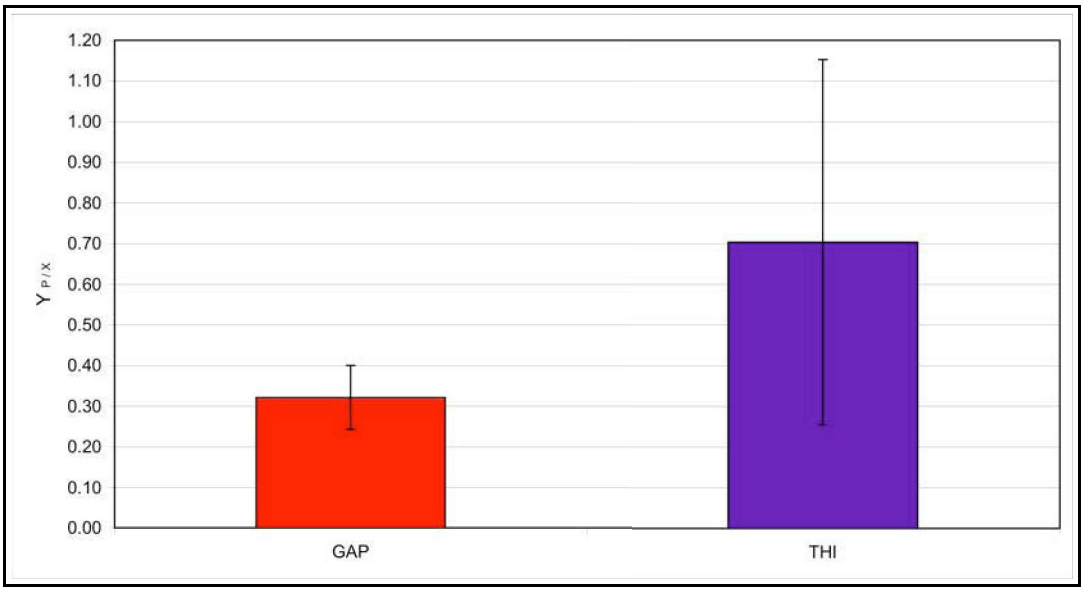


Figure 19: Product yield of HSA expressing clones under control of the P_{GAP} and the P_{THI} – HSA productivity was determined using equation 3 after 48 h of cultivation in a synthetic medium (SSC). Data represent the results of 10 clones each, with the depicted standard error of the mean.

As can be seen in figure 19, high level expression using the P_{THI11} was possible on thiamine free media using glucose as carbon source. Addition of thiamine to the culture completely inhibits expression, indicating that the *THI11* promoter is responsive to thiamine availability and active in thiamine limited conditions (data not shown).
Again in this screening it was obvious, that gene copy number has a high impact on the outcome of an expression analysis.

3.2.1. Determination of transcript level

Given that the intention of this study is the characterization of promoter activities, a Realtime-PCR system had to be developed making a relative comparison among different promoters possible. Lacking a known control with for example one-gene-copy, a 10-fold dilution series of PCR product ranging from 10^8 to 10^4 copies per reaction was used (see Fig. 6). The numbers of copies per microliter was calculated using Avogadro's number (see 2.13.2). Increased levels of fluorescence are directly related to the accumulation of PCR product and are detected during each cycle of amplification through the use of specialized instrumentation (Rotor-Gene 6000™; Corbett Life Science, Australia). The cycle thresholds (C_t) are assigned automatically to each sample according to the cycle at which the fluorescence exceeds a specific level above background (Fig. 20).
The fluorescence threshold value was calculated using the Rotorgene 6000 system software. During the SYBR Green I reaction the software detects the accumulation of PCR product by the accumulation of fluorescence. Normalized fluorescence is plotted versus cycle number.

A C_t value is obtained by drawing an arbitrary cutoff through the reactions so that the line passes through the log phase of each reaction.

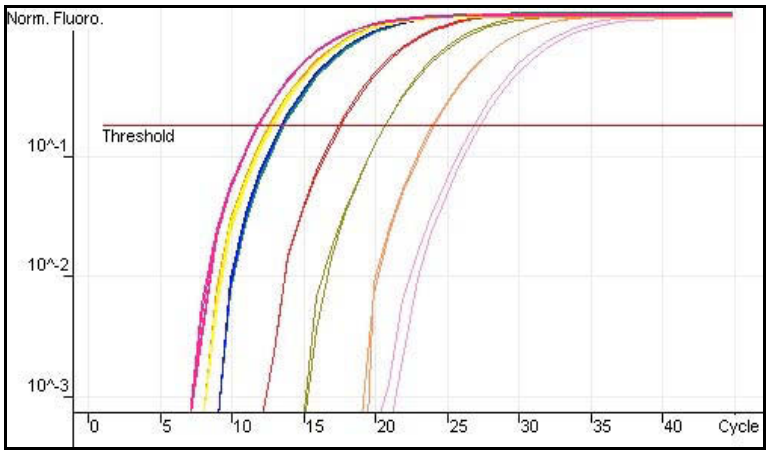


Figure 20: Example of SYBR Green I results representing amplification plots from serial dilutions of PCR product using a logarithmic scale for the y axis – the cycle threshold (red horizontal line) provides an arbitrary cutoff at which a C_t value for each sample is assigned.

Samples with higher levels of template at the beginning of the reaction will amplify to detectable levels more quickly and therefore yield a lower C_t.

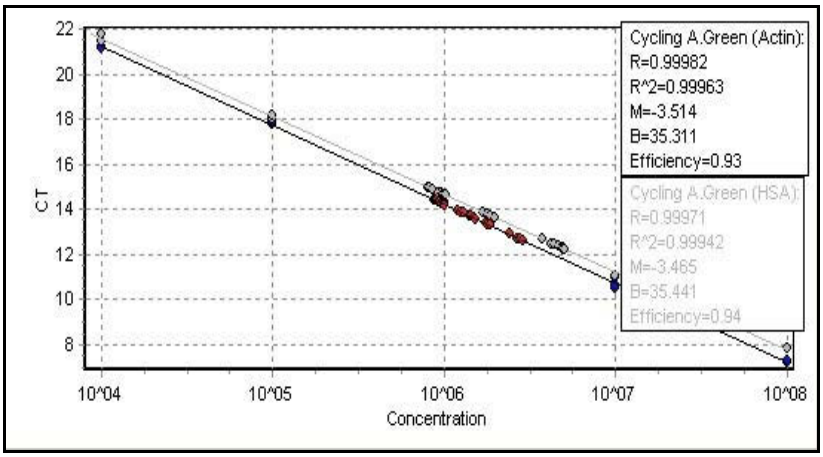


Figure 21: standard curves for actin and HSA

Normalization of the data was achieved using the house-keeping gene actin as a reference, meaning that each sample was analyzed twice. First to identify the actin level, second to determine the level of the gene of interest (GOI). The gene of interest C_t value (either *HSA* or *GFP*) is compared to the actin C_t (act1) value to normalize the C_t value of each reaction to yield ΔC_t [ΔC_t = C_t (GOI) – C_t (act1)]. This calculation removes any variation contributed by unequal template input in reactions, due to different reverse transcription efficiency of different samples.

With these data a calibration curve was generated (example Fig. 21).

All depicted qRT-PCR data represent the results of one out of at least two independent experiments.

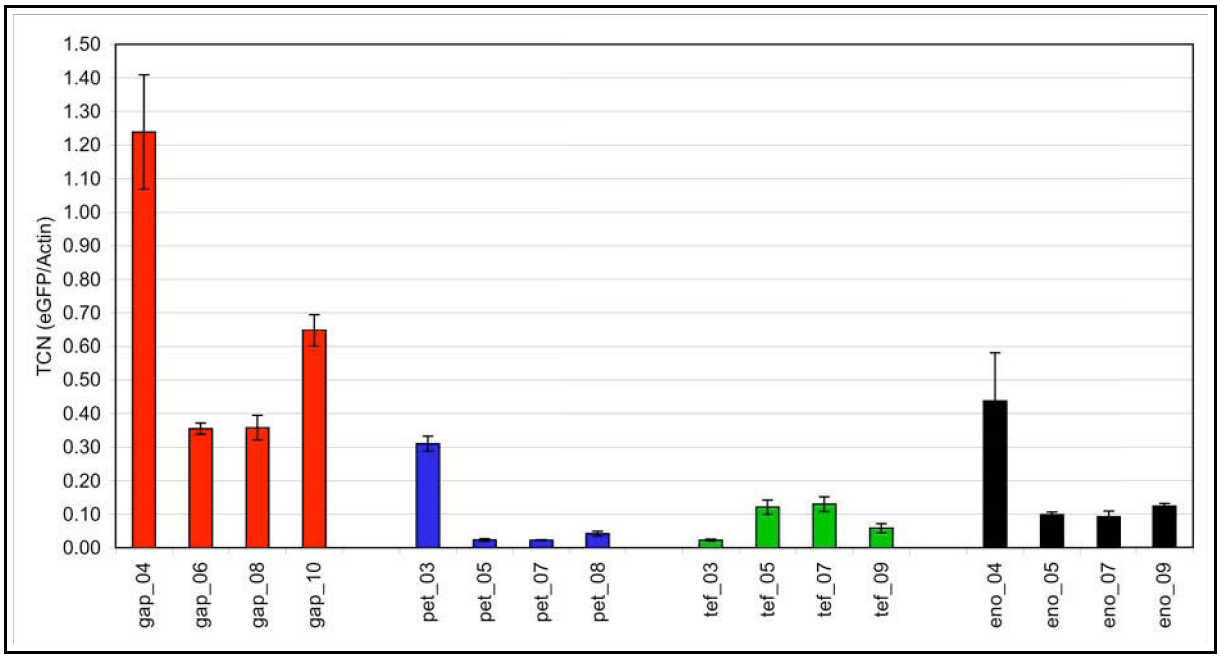


Figure 22: Transcript Copy Number (TCN) of eGFP expressing clones – data derive from qRT-PCR normalized to Actin. Each bar resembles a quadruple measurement.

Surprisingly, the transcription pattern achieved from the qRT-PCR experiment (Fig. 22) did not resemble the picture obtained from measuring eGFP protein levels. Expression of eGFP driven by the P_{PET9} nearly has no detectable transcript copy level opposed to the highest levels of protein expression. Only the pet_03 clone gave similar results as the clones under control of the *GAP* promoter. Also very interesting were the results from the *TEF1*- and *ENO1*-clones. *ENO1* promoter driven expression of eGFP always remained below the expression levels of the P_{TEF1}, however the results from the qRT-PCR experiment indicate a higher TCN (transcript copy number) for the P_{ENO1}.

In order to rule out that the behaviour of the P_{PET9} is growth rate related, transcript levels of two clones each expressing eGFP under the control of P_{GAP} and P_{PET9} were followed during batch cultivation over 51h in rich glucose medium to investigate mRNA levels and thereby promoter activity in correlation to growth phase. The results are shown in figure 23.

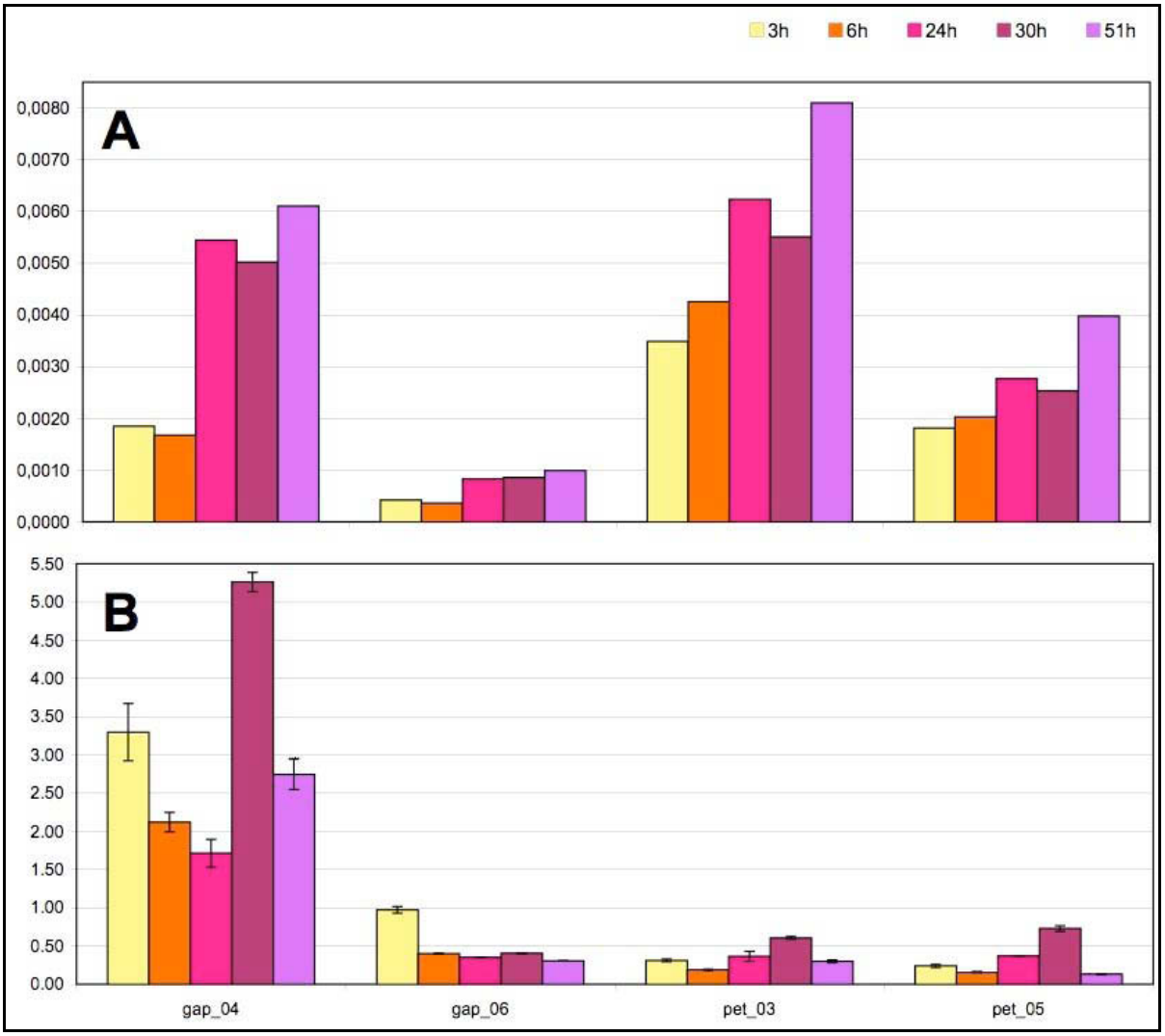


Figure 23: time course over 51 h of four eGFP expressing clones – samples were taken 5 times, after 3 h (yellow bars), 6 h (orange bars), 24 h (light pink bars), 30 h (dark pink bars) and 51 h (violet bars) **A: protein level:** eGFP was quantified using flow cytometry and correlated to the cell size according to equation 1 **B: transcript level:** data derive from quantitative Realtime-PCR normalized to Actin. Each bar depicts a quadruple measurement.

During this time course the eGFP transcript level under the control of the P_{PET9} remained disproportionately low compared to the product concentration.

The transcript level of the HSA expressing clones (Fig. 24) shows a good correlation to the measured HSA protein level (Fig. 17). Clones under control of the P_{PET9} showed a higher transcript level as *TEF1*-clones. *ENO1* promoter driven expression appeared in the range of P_{GAP} expression levels, but the results from the qRT-PCR experiment indicate a higher TCN as well.

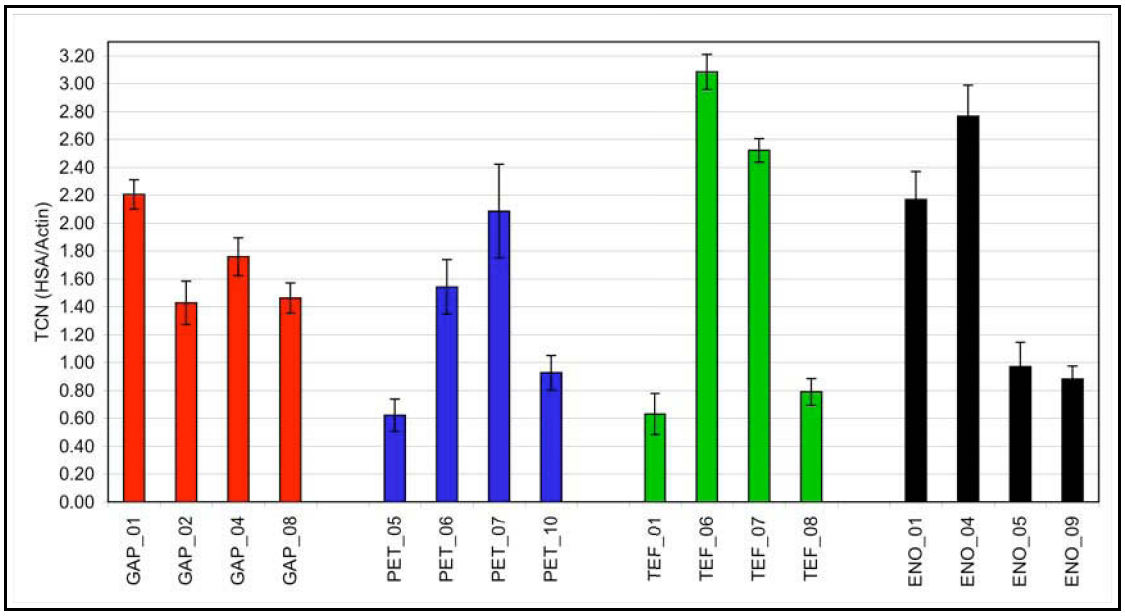


Figure 24: Transcript level of HSA secreting clones – data derive from quantitative Realtime-PCR normalized to actin. Each bar depicts a quadruple measurement.

Generally, for the HSA and eGFP expressing clones a good correlation of protein secretion to transcript level can be seen for all the individual clones, with the exception of P_{PET9}. When using the *PET9* promoter construct transcript levels of HSA and eGFP were rather low, but resulted in high protein levels in case of eGFP.

3.2.2. Gene copy number determination

The previous experiments lead to the question of gene dosage present in the selected clones. As was already observed gene copy number has a high impact on gene expression, therefore genomic DNA from the selected *P. pastoris* clones was prepared using the DNeasy Blood&Tissue Kit (Qiagen) (Fig. 25). Clones were cultivated overnight in rich medium with glucose. The quality and concentrations of all isolated DNA samples were determined with the Nanodrop ND-1000 spectrophotometer and the A260 for double-stranded DNA: 50 mg/L double-stranded DNA has an A260 of 1. Once again samples were considered for further analysis only when the DNA was of high purity and appeared undegraded.

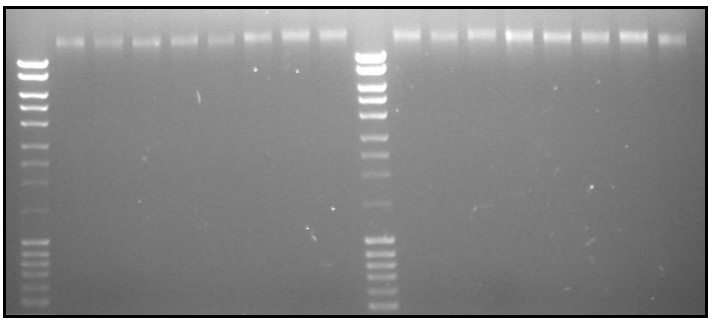


Figure 25: isolated genomic DNA of *P. pastoris*

The DNA extracted from the *P. pastoris* samples was adjusted to a concentration of 4 ng / µL. This concentration was proven to be beneficial for experiments using Realtime-PCR (Konigshoff et al. 2003). For each experiment 1 µL (4 ng) of DNA was used.

For determination of gene copy number, the same experimental set-up was used as in the foregone transcription analysis. Again a 10-fold dilution series of PCR product ranging from 10^8 to 10^4 copies per reaction was used and for normalization the level of the reference gene actin was determined.

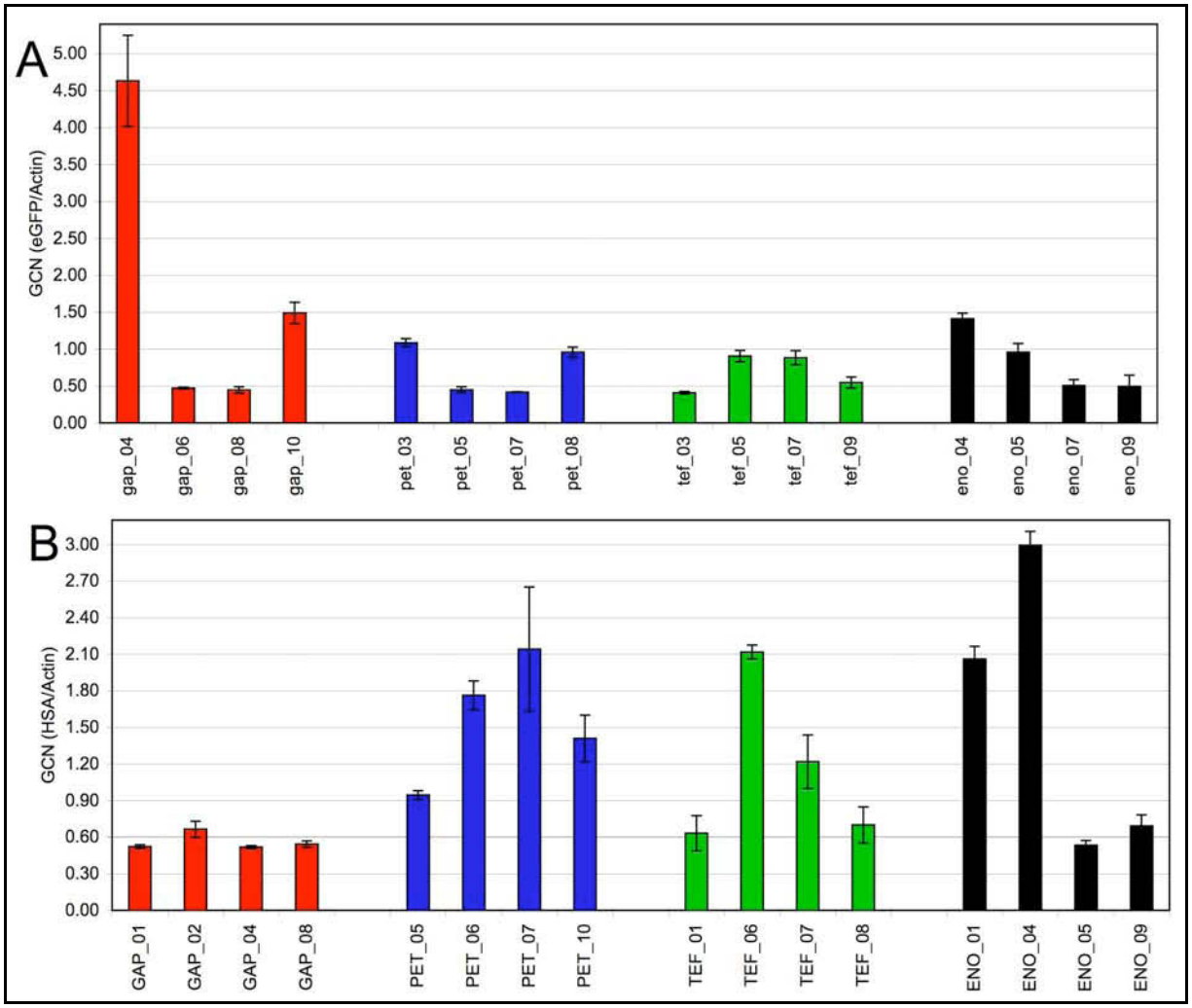


Figure 26: Gene Copy Number (GCN) determination – A: eGFP expressing clones; B: HSA expressing clones: data derive from quantitative Realtime-PCR normalized to actin. Each bar represents a quadruple measurement.

Again, the HSA expressing clones showed a good correlation of protein secretion to gene dosage, as can be seen in detail in figure 27 for the clones under control of the P_{THI11}, but now the eGFP clones showed as well a good accordance (Fig. 26).

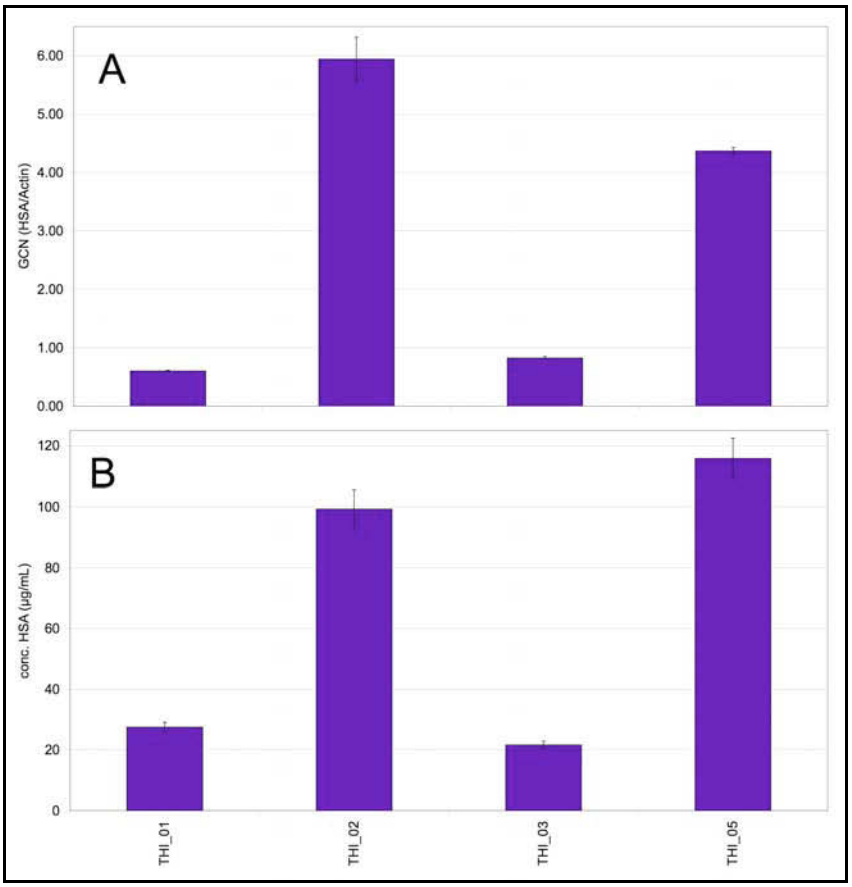


Figure 27 A: GCN determination – data derive from quantitative Realtime-PCR normalized to actin. Each bar represents a quadruple measurement. **B: HSA amount after 48 h of shake flask cultivation in SSC** – diluted supernatant was used as sample for ELISA.

3.3. Evaluation of qRT-PCR analysis

3.3.1. TCN

To address the problem of the low results in the qRT-PCR experiments regarding the eGFP-expressing clones (see for example Fig. 22), additionally a different primer-pair was tested (eGFP NEW).

With this second eGFP primer pair, a Touchdown PCR (as described in Tab. 9) was carried out and visualized with agarose gel electrophoresis. As can be clearly seen in figure 28 the product quantity was considerably higher with the newly designed eGFP primer pair.

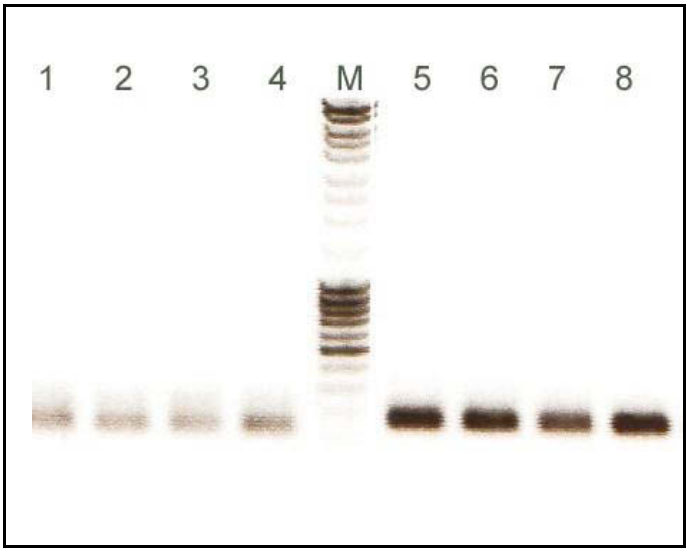


Figure 28: Agarose gel after Touchdown-PCR – M: molecular weight marker, lanes 1-4 samples gap_04, pet_03, tef_03 and eno_04 with the old eGFP primer-pair; lanes 5-8 the same samples with the new eGFP primer-pair.

However, no differences in TCN were obtained in qRT-PCR results when using the new eGFP primer pair (Tab. 21).

Table 21: comparison of eGFP TCN determination with qRT-PCR

	TCN (GFP NEU/Actin)	STDEV	**TCN (GFP/Actin)**	STDEV
gap_04	1.26	0.13	1.24	0.17
gap_06	0.40	0.06	0.35	0.02
gap_10	0.57	0.41	0.65	0.47
pet_03	0.30	0.04	0.31	0.02
pet_07	0.02	0.00	0.02	0.00
tef_03	0.13	0.24	0.02	0.00
tef_05	0.09	0.04	0.12	0.02
tef_09	0.10	0.08	0.06	0.01
eno_04	0.46	0.14	0.44	0.14
eno_05	0.09	0.01	0.10	0.01

3.3.2. GCN

To verify the GCN data achieved by qRT-PCR quantitative Southern Blot analysis was carried out (see 2.14).

After separating the digested DNA fragments of the genomic samples as well as the standard dilutions by gel electrophoresis, a southern blot was prepared.

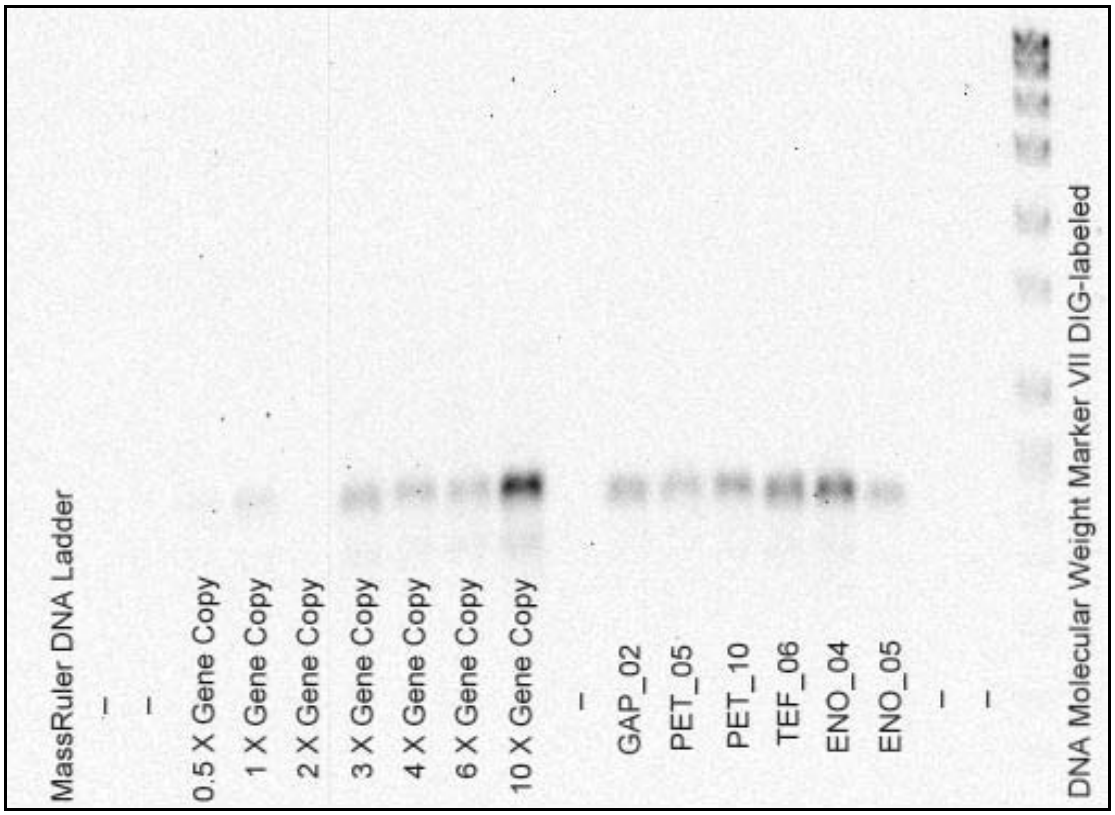


Figure 29:Southern Blot probed with a HSA-specific probe

Unfortunately all prepared Southern Blots could not be quantified due to problems with the applied standards (Fig. 29), therefore relative comparison of the promoter samples was performed.

A correlation of gene copy number determined by qRT-PCR to max BLU (maximal Boehringer Light Units) from the Southern Blot analysis is depicted in figure 30, it shows the same trend.

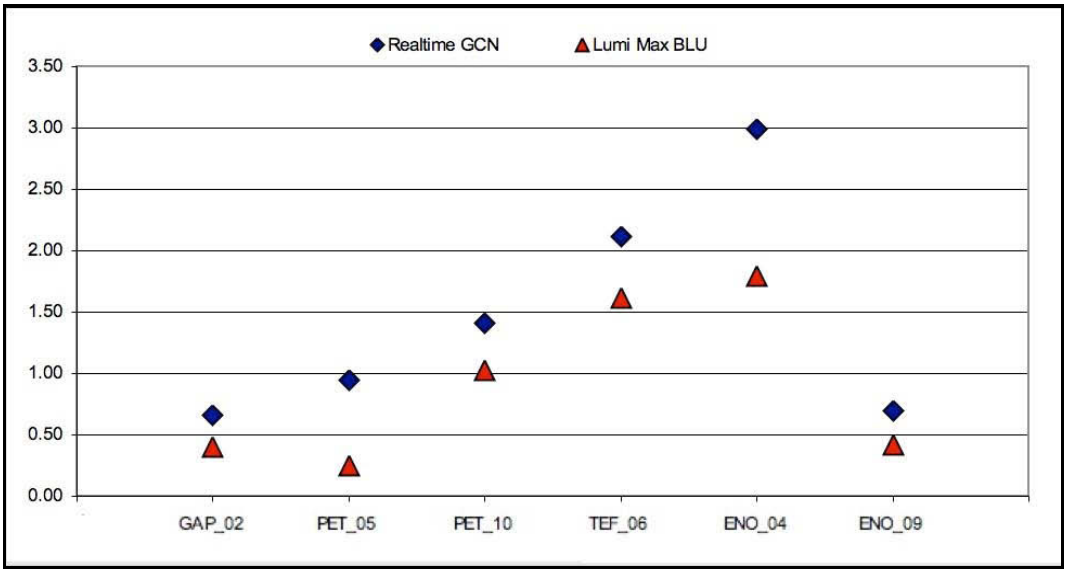


Figure 30: Plot of gene copy number determined by Realtime PCR and Lumi Max BLU (Boehringer Light Units) from Southern Blot analysis

3.4. Correlation of promoter activity to specific growth rate

In our study (Stadlmayr et al. 2010a), the potential of three promoters (P_{GAP}, P_{TEF1} and P_{THI11}) for the expression of secreted heterologous proteins was also evaluated in fed batch cultivations. For this purpose clones containing a single copy of the HSA expression cassette were selected, even though these clones would produce only low total amounts of HSA, to ensure a comparable genetic background for all promoters.

Figure 31 shows the cultivation kinetics during the fed batch phase using glucose as the carbon source. Typical for the applied feeding strategy, developed by Maurer et al. (2006), the specific growth rate profiles showed a steadily decreasing profile. As desired all cultures reached approximately the same biomass concentration.

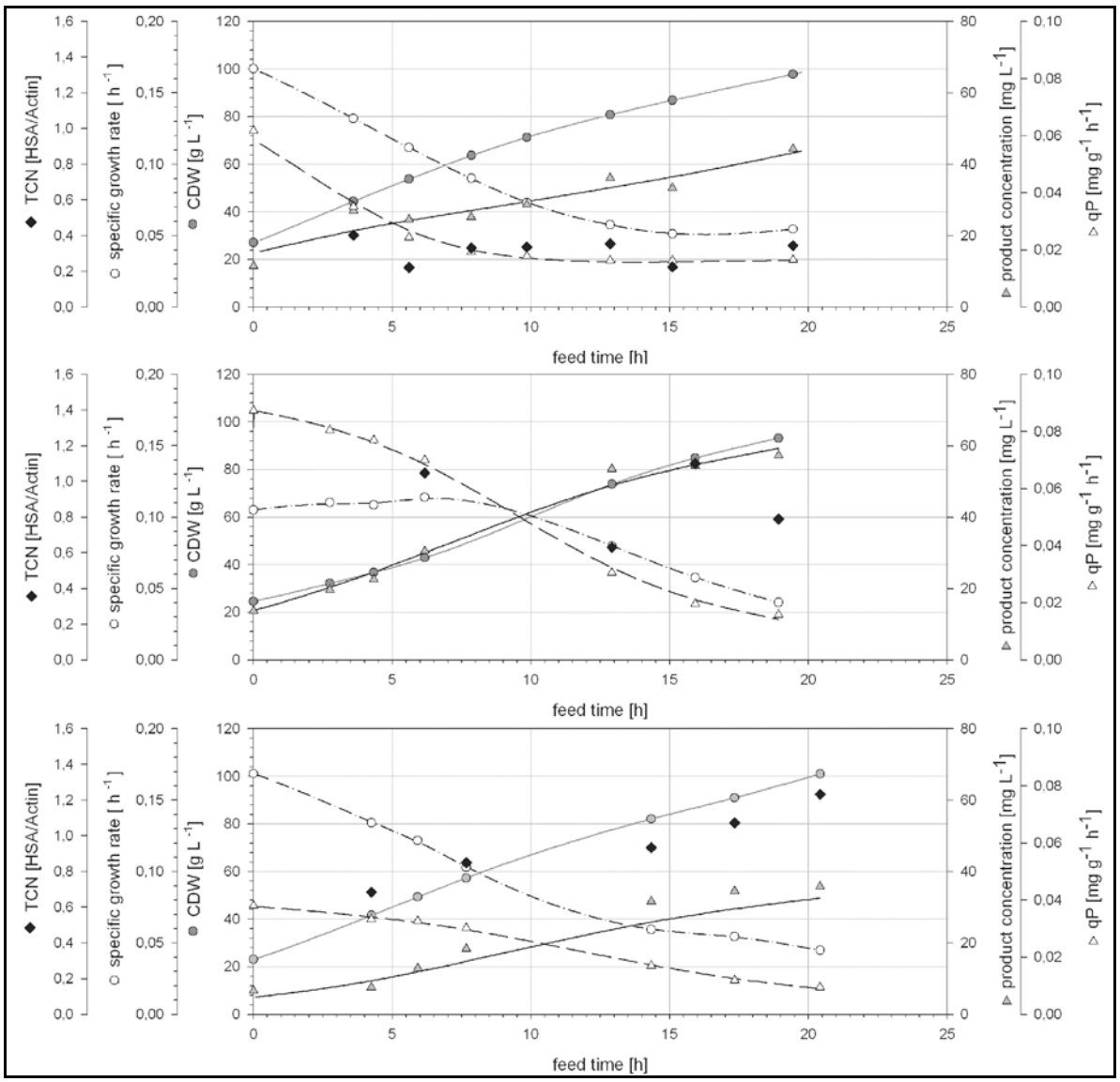


Figure 31: Fed batch cultivations of HSA clones with P_{GAP}, P_{TEF1} and P_{THI11} – Feed phases of the fed batch cultivations of *P. pastoris* SMD1168H expressing recombinant HSA (taken from Stadlmayr (2010a) **A:** under control of the *GAP* promoter, **B:** under the control of the *TEF1* promoter and **C:** under control of the *THI11* promoter. Full circles represent dry cell mass concentration [g L^{-1}], open circles specific growth rate [h^{-1}], full triangles product concentration [mg L^{-1}], open triangles specific production rate [mg L^{-1} h^{-1}] and full diamonds represent normalized transcript levels (TCN) of HSA per actin

The highest product titer was reached expressing HSA under the control of P_{TEF1} with 57.2 mg L^{-1} HSA, followed by the P_{GAP} with 44.2 mg L^{-1} and 31.3 mg L^{-1} with P_{THI11}. The expression under control of the *THI11* promoter starts after the batch at a lower level (approximately 6 mg HSA L^{-1}), this indicates a possible repression in this phase.

The different characteristics and properties of these promoters are depicted in figure 32. P_{GAP}, a constitutive promoter leads to specific growth rate independent constant expression levels. As expected by the results of Ahn et al. (2007) the amount of HSA mRNA expressed under control of P_{TEF1} is higher at high specific growth rate, whereas P_{THI11} driven expression levels are increased at lower specific growth rates.

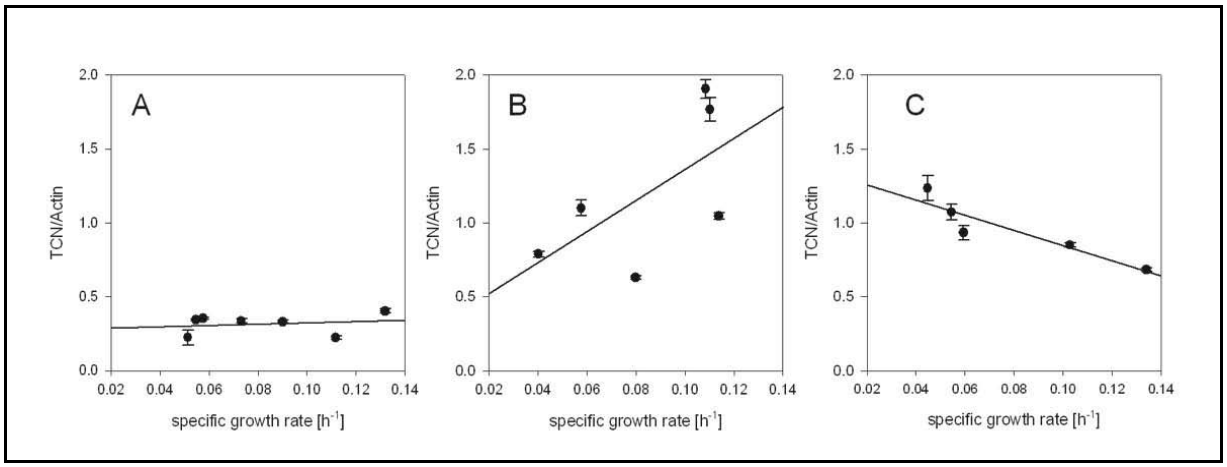


Figure 32: Correlation of relative transcript levels to specific growth rate – Correlation of HSA mRNA levels (relative transcript copy number TCN) to the specific growth rate during the fed batch cultivations, also (taken from Stadlmayr (2010a)) **A:** expression under the *GAP* promoter, **B:** *TEF1* promoter and **C:** *THI11* promoter.

4. DISCUSSION

In this work a focus was set on the characterization of novel promoter activities in *P. pastoris*. Promising constructs were selected to investigate the transcription levels and gene copy number in comparison to the commonly used constitutive promoter P_{GAP}.

4.1. Characterization of the selected promoters

As outstandingly high relative expression levels could be found in case of the combination of P_{PET9} and eGFP as reporter, this construct was selected, as well as the growth rate dependent P_{TEF1}, which gave stronger results with the secreted HSA as reporter. The P_{ENO1} was chosen because of its activity on all tested carbon sources (Stadlmayr et al. 2010a). During the course of this work a fourth promoter was selected, P_{THI11}, as it displayed interesting, yet uncharacterized regulatory properties, both in terms of the derepression mechanism as well as the growth dependency, and was compared to the previously characterized promoters.

First, promoter activities were once again determined indirectly by measurement of the amount of gene product expressed from the promoter to be able to estimate the potential of the promoter constructs. The results of these experiments can be seen in Fig. 14.

All tested clones exhibited eGFP expression levels, ranging from a very low expression level in case of the *ENO1* promoter to remarkably high levels in case of the *PET9* promoter. The outcome of the screening experiments with the secreted HSA as product, showed a different pattern. While eGFP expression under control of the P_{TEF1} and the P_{ENO1} gave low fluorescence levels, in case of HSA secretion similar or even higher amounts compared to the *GAP* promoter were observed, whereas a different behaviour was seen for expression under control of the *PET9* promoter. Contrary to high expression of eGFP, only low expression levels were measured in case of HSA. These results were in accordance to the data of the first broad screening round of Stadlmayr and colleagues. But already at this point the question arises why such high intersubject variability in expression levels occur within a clone family, pointing towards gene copy number effects. So it was of interest to not only evaluate the transcriptional force of the selected promoter elements, but according to these experiments to additionally determine the gene copy number present in the selected clones.

After ruling out the possibility that the different promoters might have an impact on the expression of eGFP and therefore could change the fluorescence properties, by applying a different method for eGFP quantitation (Tab. 20), the next step was the observation of the transcriptional level in tandem with gene copy number determination with the means of qRT-PCR.

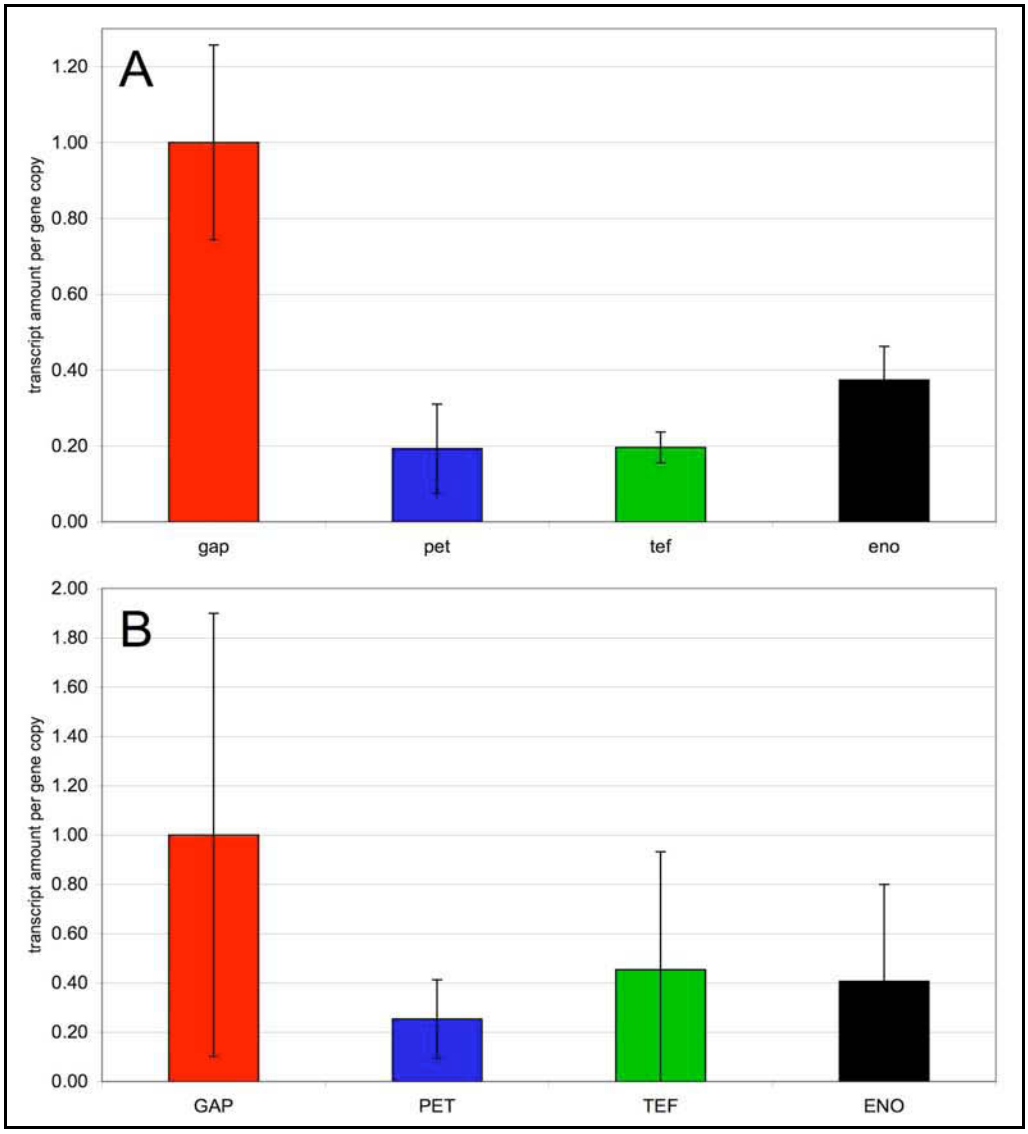


Figure 33: Mean value of transcript copy number per gene copy number of four clones each per construct – A: clones expressing **eGFP, B:** clones secreting **HSA**. Data were obtained applying qRT-PCR.

qRT-PCR analysis revealed that the *GAP* promoter constructs produce noticeable higher rates of transcript than all the other selected promoter constructs. Figure 33 illustrates the highest transcriptional strength of the P_{GAP} for both reporter proteins. Relative to the P_{GAP} clones secreting eGFP under control of the P_{PET9} and P_{TEF1} roughly produce one fifth of transcript copy in comparison to P_{GAP} (Fig. 33A). Expressing HSA the transcriptional strength of P_{TEF1} increases to two fifth compared to P_{GAP} (Fig. 33B), whereas P_{ENO1} has a consistent transcript amount per gene copy of about two fifth of the P_{GAP}.

Product levels per gene copy number of both reporter proteins were equally high for P_{TEF1} and P_{GAP} (see Fig. 34). In clones expressing eGFP under P_{PET9} the ratio product to gene copy number is much higher as in the other constructs (see Fig. 34A).

But analyzing the data of the eGFP expressing clones is very difficult due to the fact of the very low transcript copy numbers obtained through qRT-PCR. It is necessary to consider the fact that the eGFP transcript might be unstable or not completely accessible by PCR.

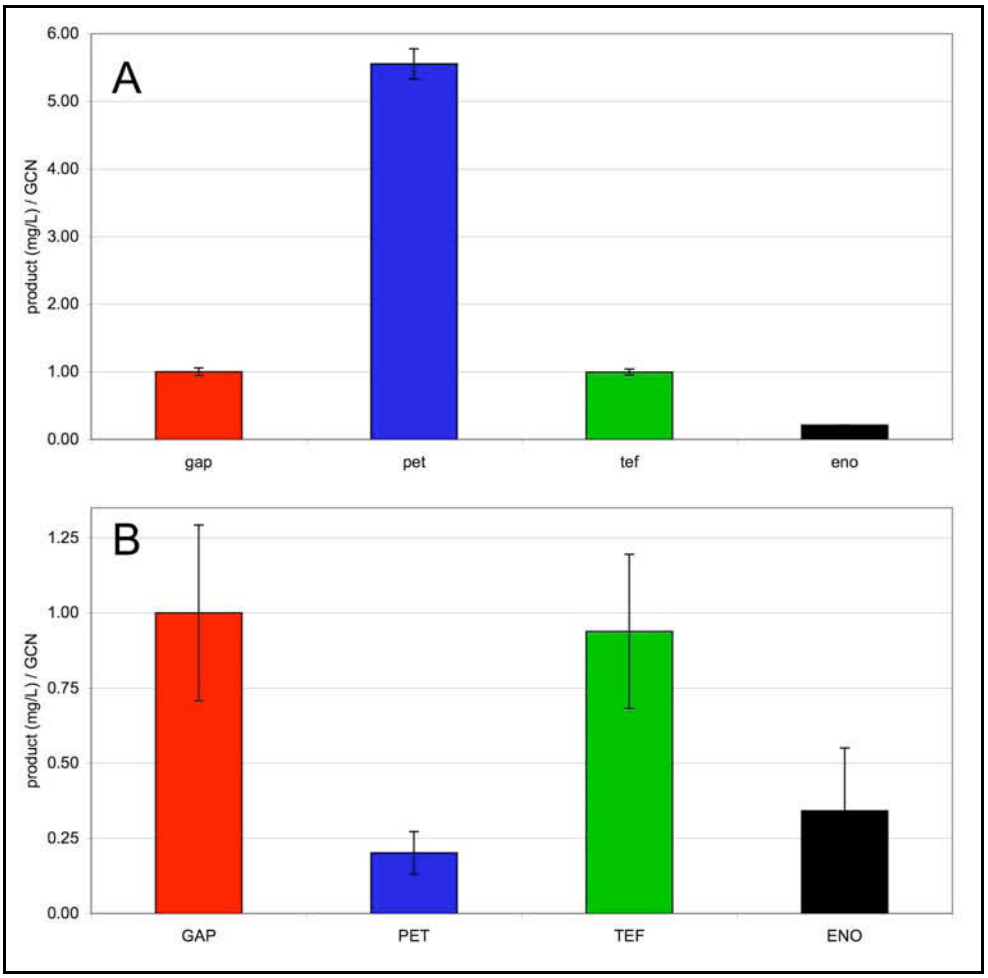


Figure 34: Mean value of product concentration per gene copy number of four clones each per construct – A: clones expressing **eGFP, B:** clones secreting **HSA.** Data were obtained applying ELISA and qRT-PCR.

While for HSA good correlation of protein secretion to both gene dosage and transcript copy number can be seen, for the intracellular reporter eGFP a good correlation of protein expression to both transcript level and gene dosage can be seen for all individual clones with the exception of the P_{PET9} constructs. However, according to the report of Xuan (2009), their results show that the changing trend of transcription levels was consistent with that of the eGFP expression.

4.2. Gene copy number effects

In *P. pastoris* multiple copies of the heterologous gene expression cassette can be integrated into the same genome. A high copy number usually results in high-level expression of the recombinant protein. This variety in gene copy number probably occurs through homologous recombination during random integration events (10% frequency, Cereghino et al. 2000). Neglecting gene copy numbers can therefore easily lead to false interpretations of experimental results from promoter studies and co-expression of helper factors as demonstrated in an example by Abad very recently (2010).

As already mentioned, a high correlation between expression level and gene dosage for both the intracellular reporter as well as the secreted reporter protein was observed.

When relating gene copy number to the level of secreted product by HSA measurement of selected clones, a linear correlation between expression level and gene dosage could be obtained to a certain point, similar correlations were obtained by Marx et al. (2009). These results suggest a strong correlation of gene copy number and productivity (dependent on the specific promoter strength). Even promoters with a rather low transcriptional strength such as P_{ENO1} can lead to protein production levels comparable to highly active promoters, when clones with a higher gene copy number are analyzed, see Fig. 26, five copies under P_{ENO1} equal one copy of P_{GAP} or P_{TEF1}. The combination of promoters with reasonable strength and an efficient method for copy number amplification, as described by Marx (2009), will be an efficient method to optimize transcription and thereby protein production levels.

But depending on the target gene, high copy number expression strains do no necessarily lead to higher expression rates. As already stated for some secreted proteins rising numbers of expression cassettes can result in a decreased expression due to bottlenecks other than transcript levels (Hohenblum et al. 2004; Gasser et al. 2006; Inan et al. 2006).

4.3. Impact of reporter gene

Comparing the promoter-dependent, relative expression levels of the two different reporter genes tested in this study, an impact of the reporter gene on the indirect promoter activity (amount of recombinant protein) was observed. As already stated, an outstandingly high impact of the reporter could be seen in case of the *PET9* promoter. Despite an identical cloning strategy and no differences in the promoter sequence upstream of the start codon, the relative productivities of eGFP and HSA varied in a very broad range. qRT-PCR analysis revealed that both the eGFP as well as the HSA transcript level were quite low when using the *PET9* promoter, and resulted in disproportionately high protein levels in case of eGFP. Recapitulating one can say, that additional investigations focusing on promoter architecture would be necessary to gain an explanation for this phenomenon.

4.4. Correlation of promoter activity to specific growth rate

Information on the kinetics of promoter activity over different growth phases may have influences for the optimal course of a fed-batch process. By applying a promoter dependent growth kinetic during bioreactor cultivation, more productive processes can be obtained. Therefore the correlation between specific growth rate and specific transcript levels has to be investigated. Especially the performance of the thiamine responsive P_{THI11} was compared to the strong promoters P_{GAP} and P_{TEF1} in bioreactor cultivations. P_{PET9} was excluded at this point, as the high production levels of eGFP could not be repeated for the secreted model protein.

By analyzing the relative transcript level over a broad range of specific growth rates (by applying a fed-batch cultivation as described by Maurer et al. 2006) a clear promoter dependent correlation between specific growth rate and relative transcript levels of the reporter gene HSA could be shown. In case of the commonly used P_{GAP} no effect of the specific growth rate on the amount of HSA mRNA could be found. For P_{TEF1} and P_{THI11} a different picture could be seen. As has been described previously for a batch cultivation (Ahn et al. 2007), the relative transcription level (promoter activity) of the P_{TEF1} is directly proportional to the specific growth rate. In glucose limited fed batch, up to a more than 4 fold higher transcript level can be reached with P_{TEF1} compared to P_{GAP}. Also for P_{THI11} an overall higher transcript level compared to P_{GAP} was detected, but in this case an indirectly proportional relation between the HSA transcript level and specific growth rate could be seen (see Fig. 32).

To evaluate the potential of promoters that have a positive or negative correlation to the specific growth rate, respective feeding regimes have to be established in future.

4.5. Interpretation of qRT-PCR results

For quantifying mRNA levels, the SYBR-Green based real-time reverse transcription PCR method is widely used (Bustin 2002). The key advantages of qRT-PCR over other PCR-based methods for quantification of nucleic acids are an extremely wide dynamic range (over more than 8 orders of magnitude as stated by Heid et al. 1996) and the high degree of reliability of the results, as not only the quantity of the product is analyzed after the PCR, but the kinetics of the entire reaction are visible. The knowledge of the reaction curves can be used to quickly and easily determine the initial quantity of the template (Higuchi et al. 1993).

But naturally such a powerful tool also has disadvantages. Critical issues defining the reliability of the obtained data are among others the choice of the house-keeping gene: an ideal housekeeping gene should always have the same level of expression, and of course the influences of sample preparation (Klein 2002) have to be mentioned. As at the time this

study started no reports for real-time PCR performed in *P. pastoris* were available, the *ACT1* gene, coding for Actin, was chosen to be used as an endogenous control, as this is a single copy gene of the host and should be constitutively expressed. Meanwhile, also other investigators reported the use of *ACT1* to calibrate transcript levels in *P. pastoris*.

According to Fleige and Pfaffl (2006) qRT-PCR performance is affected by the RNA integrity, therefore RNA integrity was quickly checked with the Nanodrop ND-1000 (Nanodrop, USA) and contaminations with DNA in mRNA analysis were overcome by applying a DNase treatment step.

Another aspect is the intra- and inter-assay variation in interpreting qRT-PCR data. Intra-assay variation was calculated measuring four replicates (quadruples) amplified in the same PCR run. Inter-assay variation was measured not only by repeating the experiment on another day, but reproducibility of qRT-PCR was also proven by repeated measurements testing a different primer-pair (see Tab. 21) and also by applying a different gDNA isolation method (data not shown). Mean error between PCR runs was 2.8% and 12% between transcript quantification of independently repeated cultures.

Consequently, data interpretation is the last step in the workflow. At the end of each reaction, the recorded fluorescence intensity is used for the following calculations by the software of the system used. Many different approaches to quantify the amount of template exist (Livak et al. 2001; Pfaffl 2001; Ginzinger 2002).

The standard curve method determines the input copy number of the template of interest, usually by relating the signal to a standard curve. The standard curve is first constructed from a template of known concentration. This curve is then used as a reference standard to gain quantitative information for targets of unknown concentrations (Giulietti et al. (2001), also reviewed by Bustin (2000)). This method seemed suitable for this work and was therefore employed.

In the relative fold change method (Pfaffl et al. 2002), one of the experimental samples is the calibrator. Each of the normalized target values is divided by the calibrator normalized target value to generate the relative expression levels. All quantities are expressed as an n-fold difference relative to the calibrator. As the calibrator should be available at large enough quantities to be included in each run, this method could not be employed in this work due to lack of a suitable calibrator at the beginning of the experiments.

The comparative threshold (C_t) method, also referred to as $\Delta\Delta C_t$ method (Livak et al. 2001) uses no known amount of standard but compares the relative amount of the target sequence to any of the reference values chosen and the result is given as relative to the reference value. Before using the $\Delta\Delta C_t$ method for quantitation a validation experiment has to be performed to demonstrate that efficiencies of target and reference are approximately equal. The advantage of using the comparative Ct method is that the need for a standard curve is

eliminated, therefore not only more wells are available for samples, but also the adverse effects of any dilution or calculation errors made in creating the standard curve samples are eliminated. But as for this method also a calibrator sample is needed, this method was not employed in this study.

During the analysis of the bioreactor samples, the idea occurred to try a comparison with some of the analyzing options offered by the Rotorgene software. For all the relative methods the sample taken during the batch phase was set as calibrator (B2). Figure 35 shows that the different methods are very similar in the outcome.

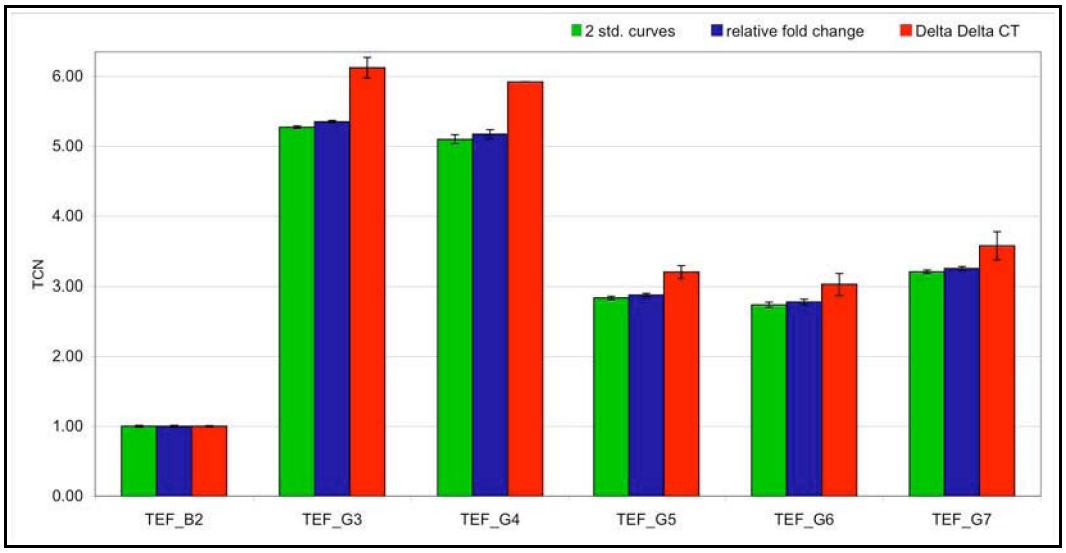


Figure 35: Comparison of three different data evaluation methods – (green) standard curve method, (blue) relative fold change method, (red) $\Delta\Delta C_t$ method; displayed are the samples taken from the P_{TEF1} bioreactor cultivation, relative fold change and $\Delta\Delta C_t$ analysis were assessed with the Rotorgene 6000 system software**,** data for the standard curves were exported to MS Excel for further analysis.

According to Pfaffl (2004) the question which might be the best RT-PCR quantification strategy to express the exact content in a sample has still not been answered to universal satisfaction.

Real-time PCR has the advantage of sensitivity, speed, high throughput and high degree of potential automation compared to the conventional quantification methods such as Northern Blot analysis or RNase protection assay. However, up to now, no single standard operating protocol exists, as Bustin already stated in his review (2002). Crucial points are among others: lysis of cells, RNA extraction, DNase treatment, cDNA synthesis, real-time PCR and last but not least data evaluation. In general qRT-PCR needs to be optimized to obtain reliable and accurate data because many aspects of the multi-step qRT-PCR procedure can influence the outcome. Furthermore, comparing clones at the same point in the growth phase is a crucial factor for getting reliable results (Werbrouck et al. 2007).

Here we could show the applicability of qRT-PCR to study gene regulation as well as gene copy number in recombinant *P. pastoris* when using *ACT1* as reference gene.

4.6. Promoter regulation

Inducible promoters have been essential tools in basic and applied biological research, for example in studies of essential genes (Zhang et al. 2003) and naturally for the development of strains engineered to produce toxic proteins or metabolites as reviewed by Porro (2004). All these applications require very specific induction properties. For example an ideal inducible promoter for the industrial fermentation of yeast must, according to Nevoigt (2007), be tightly regulated, be inexpensive to induce, express at high levels after induction and be easy to handle.

In a first round of fed batch cultivations, the thiamine responsive P_{THI11} was shown to be a promising new regulatory sequence. The bioprocess cultivations were performed on fully synthetic minimal medium without addition of thiamine. As pre-cultures were carried out on complex medium, any residual thiamine had to be removed by the performed wash step prior to reactor inoculation. The experimental design targeted fully induced conditions for the P_{THI11}, however, the applied growth regime prevented constantly high promoter activity. To evaluate this possibility further experiments should be carried out investigating the induction condition in more detail. This derepressible promoter can be used for the regulation of target gene expression by the addition or absence of thiamine from the cultivation medium, independently from the preferred carbon source, thereby enabling a very easy induction strategy (Stadlmayr et al. 2010a).

4.7. Conclusion and further prospectives

As the number of available promoter sequences is rather limited in *P. pastoris*, and also restricted to promoters having a strong activity, a broader range of promoter activities opens up the field of controlled protein expression. Improving the production of secreted recombinant proteins is one of the main goals in host cell engineering. Very promising results of increased production and secretion rates could be gained by co-overexpression of chaperones, foldases (reviewed by Gasser et al. 2008) and other cellular host proteins (Stadlmayr et al. 2010b). However all these coexpressed genes were set under the control of the strong *GAP* or *AOX* promoter, that were also used for the expression of the target protein. A correlation between gene copy number of the coexpressed secretion enhancing factor *PDI1* and the yield of secreted target protein in *P. pastoris* was described by Inan and co-workers (2006). Also, a high correlation between expression level and gene dosage was

observed in the present study. However, a strong overexpression of secretion enhancing genes can also have negative effects on cell viability as it was described for the overexpression of *PDI1* in *K. lactis* (Bao et al. 2000), making the controllable expression level of target protein an important tool in host cell engineering applications.

In conclusion, the production of model proteins did not surpass the expression levels obtained with the standard promoter P_{GAP}, but higher transcription levels with especially P_{TEF1} and P_{THI11} were detected. Additionally P_{THI11} appears to be the first regulated promoter for *P. pastoris* that can be controlled independently of the main nutritional components such as carbon and nitrogen source.

Finally, the results presented in this work demonstrate that future process optimization with these promoters is opening a new path in engineering production hosts.

5. REFERENCES

Abad, S., K. Kitz, et al. (2010). "Real-time PCR-based determination of gene copy numbers in Pichia pastoris." Biotechnol J **5**(4): 413-20.

Ahn, J., J. Hong, et al. (2007). "Translation elongation factor 1-alpha gene from Pichia pastoris: molecular cloning, sequence, and use of its promoter." Appl Microbiol Biotechnol **74**(3): 601-8.

Ahn, J., J. Hong, et al. (2009). "Phosphate-responsive promoter of a Pichia pastoris sodium phosphate symporter." Appl Environ Microbiol **75**(11): 3528-34.

Bao, W. G., K. K. Huo, et al. (2000). "Protein disulphide isomerase genes of Kluyveromyces lactis." Yeast **16**(4): 329-41.

Bellin, T., M. Pulz, et al. (2001). "Rapid detection of enterohemorrhagic Escherichia coli by real-time PCR with fluorescent hybridization probes." J Clin Microbiol **39**(1): 370-4.

Brierley, R. A. (1998). "Secretion of recombinant human insulin-like growth factor I (IGF-I)." Methods Mol Biol **103**: 149-77.

Bustin, S. A. (2000). "Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays." J Mol Endocrinol **25**(2): 169-93.

Bustin, S. A. (2002). "Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems." J Mol Endocrinol **29**(1): 23-39.

Cereghino, J. L. and J. M. Cregg (2000). "Heterologous protein expression in the methylotrophic yeast Pichia pastoris." FEMS Microbiol Rev **24**(1): 45-66.

Cesarone, C. F., C. Bolognesi, et al. (1979). "Improved microfluorometric DNA determination in biological material using 33258 Hoechst." Anal Biochem **100**(1): 188-97.

Chen, H. L., C. C. Yen, et al. (2006). "Production and characterization of human extracellular superoxide dismutase in the methylotrophic yeast Pichia pastoris." J Agric Food Chem **54**(21): 8041-7.

Cregg, J. M., K. R. Madden, et al. (1989). "Functional characterization of the two alcohol oxidase genes from the yeast Pichia pastoris." Mol Cell Biol **9**(3): 1316-23.

de Almeida, J. R., L. M. de Moraes, et al. (2005). "Molecular characterization of the 3-phosphoglycerate kinase gene (PGK1) from the methylotrophic yeast Pichia pastoris." Yeast **22**(9): 725-37.

De Schutter, K., Y. C. Lin, et al. (2009). "Genome sequence of the recombinant protein production host Pichia pastoris." Nat Biotechnol **27**(6): 561-6.

Don, R. H., P. T. Cox, et al. (1991). "'Touchdown' PCR to circumvent spurious priming during gene amplification." Nucleic Acids Res **19**(14): 4008.

Fleige, S. and M. W. Pfaffl (2006). "RNA integrity and the effect on the real-time qRT-PCR performance." Mol Aspects Med **27**(2-3): 126-39.

Gasser, B., M. Maurer, et al. (2006). "Engineering of Pichia pastoris for improved production of antibody fragments." Biotechnol Bioeng **94**(2): 353-61.

Gasser, B., M. Saloheimo, et al. (2008). "Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview." Microb Cell Fact **7**: 11.

Gasser, B., M. Sauer, et al. (2007). "Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts." Appl Environ Microbiol **73**(20): 6499-507.

Ginzinger, D. G. (2002). "Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream." Exp Hematol **30**(6): 503-12.

Giulietti, A., L. Overbergh, et al. (2001). "An overview of real-time quantitative PCR: applications to quantify cytokine gene expression." Methods **25**(4): 386-401.

Hamilton, S. R. and T. U. Gerngross (2007). "Glycosylation engineering in yeast: the advent of fully humanized yeast." Curr Opin Biotechnol **18**(5): 387-92.

Hartner, F. S., C. Ruth, et al. (2008). "Promoter library designed for fine-tuned gene expression in Pichia pastoris." Nucleic Acids Res **36**(12): e76.

Heid, C. A., J. Stevens, et al. (1996). "Real time quantitative PCR." Genome Res **6**(10): 986-94.

Higuchi, R., C. Fockler, et al. (1993). "Kinetic PCR analysis: real-time monitoring of DNA amplification reactions." Biotechnology (N Y) **11**(9): 1026-30.

Hohenblum, H., N. Borth, et al. (2003). "Assessing viability and cell-associated product of recombinant protein producing Pichia pastoris with flow cytometry." J Biotechnol **102**(3): 281-90.

Hohenblum, H., B. Gasser, et al. (2004). "Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant Pichia pastoris." Biotechnol Bioeng **85**(4): 367-75.

Inan, M., D. Aryasomayajula, et al. (2006). "Enhancement of protein secretion in Pichia pastoris by overexpression of protein disulfide isomerase." Biotechnol Bioeng **93**(4): 771-8.

Killgore, G. E., B. Holloway, et al. (2000). "A 5' nuclease PCR (TaqMan) high-throughput assay for detection of the mecA gene in staphylococci." J Clin Microbiol **38**(7): 2516-9.

Klein, D. (2002). "Quantification using real-time PCR technology: applications and limitations." Trends Mol Med **8**(6): 257-60.

Kobayashi, K., S. Kuwae, et al. (2000). "High-level expression of recombinant human serum albumin from the methylotrophic yeast Pichia pastoris with minimal protease production and activation." J Biosci Bioeng **89**(1): 55-61.

Konigshoff, M., J. Wilhelm, et al. (2003). "HER-2/neu gene copy number quantified by real-time PCR: comparison of gene amplification, heterozygosity, and immunohistochemical status in breast cancer tissue." Clin Chem **49**(2): 219-29.

Lekanne Deprez, R. H., A. C. Fijnvandraat, et al. (2002). "Sensitivity and accuracy of quantitative real-time polymerase chain reaction using SYBR green I depends on cDNA synthesis conditions." Anal Biochem **307**(1): 63-9.

Liu, H., X. Tan, et al. (1995). "PER3, a gene required for peroxisome biogenesis in Pichia pastoris, encodes a peroxisomal membrane protein involved in protein import." J Biol Chem **270**(18): 10940-51.

Livak, K. J. and T. D. Schmittgen (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." Methods **25**(4): 402-8.

Macauley-Patrick, S., M. L. Fazenda, et al. (2005). "Heterologous protein production using the Pichia pastoris expression system." Yeast **22**(4): 249-70.

Martell, M., J. Gomez, et al. (1999). "High-throughput real-time reverse transcription-PCR quantitation of hepatitis C virus RNA." J Clin Microbiol **37**(2): 327-32.

Marx, H., A. Mecklenbrauker, et al. (2009). "Directed gene copy number amplification in Pichia pastoris by vector integration into the ribosomal DNA locus." FEMS Yeast Res.

Mattanovich, D., A. Graf, et al. (2009). "Genome, secretome and glucose transport highlight unique features of the protein production host Pichia pastoris." Microb Cell Fact **8**: 29.

Maurer, M., M. Kuhleitner, et al. (2006). "Versatile modeling and optimization of fed batch processes for the production of secreted heterologous proteins with Pichia pastoris." Microb Cell Fact **5**: 37.

Menendez, J., I. Valdes, et al. (2003). "The ICL1 gene of Pichia pastoris, transcriptional regulation and use of its promoter." Yeast **20**(13): 1097-108.

Mhlanga, M. M. and L. Malmberg (2001). "Using molecular beacons to detect single-nucleotide polymorphisms with real-time PCR." Methods **25**(4): 463-71.

Minghetti, P. P., D. E. Ruffner, et al. (1986). "Molecular structure of the human albumin gene is revealed by nucleotide sequence within q11-22 of chromosome 4." J Biol Chem **261**(15): 6747-57.

Mocellin, S., C. R. Rossi, et al. (2003). "Quantitative real-time PCR: a powerful ally in cancer research." Trends Mol Med **9**(5): 189-95.

Morlino, G. B., L. Tizzani, et al. (1999). "Inducible amplification of gene copy number and heterologous protein production in the yeast Kluyveromyces lactis." Appl Environ Microbiol **65**(11): 4808-13.

Neuvians, T. P., I. Gashaw, et al. (2005). "Standardization strategy for quantitative PCR in human seminoma and normal testis." J Biotechnol **117**(2): 163-71.

Nevoigt, E., C. Fischer, et al. (2007). "Engineering promoter regulation." Biotechnol Bioeng **96**(3): 550-8.

Overbergh, L., D. Valckx, et al. (1999). "Quantification of murine cytokine mRNAs using real time quantitative reverse transcriptase PCR." Cytokine **11**(4): 305-12.

Paul, J. H. and B. Myers (1982). "Fluorometric Determination of DNA in Aquatic Microorganisms by Use of Hoechst 33258." Appl Environ Microbiol **43**(6): 1393-1399.

Peters, T., Jr. (1985). "Serum albumin." Adv Protein Chem **37**: 161-245.

Pfaffl, M. W. (2001). "A new mathematical model for relative quantification in real-time RT-PCR." Nucleic Acids Res **29**(9): e45.

Pfaffl, M. W., G. W. Horgan, et al. (2002). "Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR." Nucleic Acids Res **30**(9): e36.

Pfaffl, M. W., A. Tichopad, et al. (2004). "Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations." Biotechnol Lett **26**(6): 509-15.

Porro, D. and D. Mattanovich (2004). "Recombinant protein production in yeasts." Methods Mol Biol **267**: 241-58.

Reischer, H., I. Schotola, et al. (2004). "Evaluation of the GFP signal and its aptitude for novel on-line monitoring strategies of recombinant fermentation processes." J Biotechnol **108**(2): 115-25.

Reisinger, H., W. Steinfellner, et al. (2008). "The absence of effect of gene copy number and mRNA level on the amount of mAb secretion from mammalian cells." Appl Microbiol Biotechnol **81**(4): 701-10.

Romanos, M., C. Scorer, et al. (1998). "The generation of multicopy recombinant strains." Methods Mol Biol **103**: 55-72.

Rossini, D., D. Porro, et al. (1993). "In Saccharomyces cerevisiae, protein secretion into the growth medium depends on environmental factors." Yeast **9**(1): 77-84.

Sauer, M., P. Branduardi, et al. (2004). "Differential gene expression in recombinant Pichia pastoris analysed by heterologous DNA microarray hybridisation." Microb Cell Fact **3**(1): 17.

Saunders, C. W., B. J. Schmidt, et al. (1987). "Secretion of human serum albumin from Bacillus subtilis." J Bacteriol **169**(7): 2917-25.

Sears, I. B., J. O'Connor, et al. (1998). "A versatile set of vectors for constitutive and regulated gene expression in Pichia pastoris." Yeast **14**(8): 783-90.

Shen, S., G. Sulter, et al. (1998). "A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast Pichia pastoris." Gene **216**(1): 93-102.

Sleep, D., G. P. Belfield, et al. (1991). "Saccharomyces cerevisiae strains that overexpress heterologous proteins." Biotechnology (N Y) **9**(2): 183-7.

Sreekrishna, K., R. G. Brankamp, et al. (1997). "Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast Pichia pastoris." Gene **190**(1): 55-62.

Sreekrishna, K., R. H. Potenz, et al. (1988). "High level expression of heterologous proteins in methylotrophic yeast Pichia pastoris." J Basic Microbiol **28**(4): 265-78.

Stadlmayr, G., K. Benakovitsch, et al. (2010b). "Genome-scale analysis of library sorting GALibSo: Isolation of secretion enhancing factors for recombinant protein production in Pichia pastoris." Biotechnol Bioeng **105**(3): 543-55.

Stadlmayr, G., A. Mecklenbräuker, et al. (2010a). "Identification and characterisation of novel Pichia pastoris promoters for heterologous protein production." J Biotechnol **in press**.

Suzuki, T., P. J. Higgins, et al. (2000). "Control selection for RNA quantitation." Biotechniques **29**(2): 332-7.

Tschopp, J. F., P. F. Brust, et al. (1987). "Expression of the lacZ gene from two methanol-regulated promoters in Pichia pastoris." Nucleic Acids Res **15**(9): 3859-76.

Tsien, R. Y. (1998). "The green fluorescent protein." Annu Rev Biochem **67**: 509-44.

Tyagi, S. and F. R. Kramer (1996). "Molecular beacons: probes that fluoresce upon hybridization." Nat Biotechnol **14**(3): 303-8.

Vassileva, A., D. A. Chugh, et al. (2001). "Effect of copy number on the expression levels of hepatitis B surface antigen in the methylotrophic yeast Pichia pastoris." Protein Expr Purif **21**(1): 71-80.

Voytas, D. (2001). "Agarose gel electrophoresis." Curr Protoc Mol Biol **Chapter 2**: Unit2 5A.

Waterham, H. R., M. E. Digan, et al. (1997). "Isolation of the Pichia pastoris glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter." Gene **186**(1): 37-44.

Werbrouck, H., N. Botteldoorn, et al. (2007). "Quantification of gene expression of Listeria monocytogenes by real-time reverse transcription PCR: optimization, evaluation and pitfalls." J Microbiol Methods **69**(2): 306-14.

Whitcombe, D., J. Theaker, et al. (1999). "Detection of PCR products using self-probing amplicons and fluorescence." Nat Biotechnol **17**(8): 804-7.

Xuan, Y., X. Zhou, et al. (2009). "An upstream activation sequence controls the expression of AOX1 gene in Pichia pastoris." FEMS Yeast Res **9**(8): 1271-82.

Zhang, A. L., J. X. Luo, et al. (2009). "Recent advances on the GAP promoter derived expression system of Pichia pastoris." Mol Biol Rep **36**(6): 1611-9.

Zhang, N., M. Osborn, et al. (2003). "Using yeast to place human genes in functional categories." Gene **303**: 121-9.

Zimmer, M. (2009). "GFP: from jellyfish to the Nobel prize and beyond." Chem Soc Rev **38**(10): 2823-32.